TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

PULCHOWK CAMPUS

THESIS NO.: 070MSCS656

Packet Loss Recovery and Control for VoIP

by

Dipak K.C.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRONICS AND

COMPUTER ENGINEERING IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN COMPUTER

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

LALITPUR, NEPAL

MARCH 2017

COPYRIGHT

DEPARTMENTAL ACCEPTANCE

The thesis entitled "Packet Loss Recovery and Control for VoIP", submitted by Dipak K.C. in partial fulfillment of the requirement for the award of the degree of "Master of Science in Computer System and Knowledge Engineering" has been accepted as a bona fide record of work independently carried out by him in the department.

Dr. Dibakar Raj Pant
Head of the Department

CERTIFICATE OF APPROVAL

The undersigned certify that they have read, and recommended to the Institute of Engineering for acceptance, a thesis report entitled "Packet Loss Recovery and Control for VoIP" submitted by Mr. Dipak K.C. in partial fulfillment of the requirement for the degree of Master of Science in Computer System and Knowledge Engineering.

Supervisor: Babu Ram Dawadi
Lecturer
Department of Electronics and Computer Engineering

External Examiner: Krishna Prasad Bhandari
Deputy Manager
Nepal Telecom

Committee Chairperson: Prof. Dr. Subarna Shakya
Professor
Department of Electronics and Computer Engineering

DATE OF APPROVAL: MARCH 5, 2017

ACKNOWLEDGEMENT

I am highly indebted to Department of Electronics and Computer Engineering (DOECE) for providing me the opportunity to explore my interest and ideas in the field of engineering through this Thesis.

I owe a debt of gratitude towards my thesis supervisor, **Lecturer. Babu Ram Dawadi**, for his proper guidance, encouragement and support.

Finally, I would like to thank all our teachers and friends who have helped me directly or indirectly with this thesis.

Dipak K.C.

ABSTRACT

Technology is growing rapidly which results in the rise of the VoIP technology. VoIP is a technology to transfer voice signal communication over the internet (Network). The transmission of audio data over packet switched networks faces lots of difficulties by a variety of network impairments such as packet loss, delay, echo, network security and throughput. To minimize the effect of packet loss a **Packet Loss Concealment** (**PLC**) technique is introduced in VoIP communications. Voice is a real time application and the biggest problem it faces is the loss of packets due to network congestion, delay, jitter and other network factors. The internet services implements protocols to detect and retransmit the lost packets to overcome the packet loss. However, for a real time application it is too late before a lost intermediate packet is retransmitted. This causes a need for reconstruction of the lost packets. Therefore, good reconstruction techniques must be in action to minimize the effect of packet loss during voice communication. In this thesis a concealment algorithm to reconstruct lost voice packets is reported. The algorithm is receiver based and its functionality is based on Time Scale Modifications of speech and autocorrelation of a speech signal. The technique used is Average Magnitude Difference Function (AMDF) to estimate pitch (fundamental frequency of data sample) along with Overlap and Add (OLA) technique. Method of PLC can be implemented as artificial regeneration of the packet received prior to the lost one, followed by insertion of the duplicated packet into the gap. Also lost packet can be estimated by interpolation on the basis of previously received voice packet. PLC algorithm is an end-to-end bidirectional transmission control scheme. PLC also monitor VoIP communication at both end the sender and the receiver side. Implementation of PLC minimize the effect of data loss in voice communication and overall sound quality is improved.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AMDF	Average Magnitude Difference Function
CCS7	Common Channel Signaling system
CDR	Call Detailed Records
CLI	Command line Interface
CMR	Call Management Records
CODEC	Encoder/Decoder
CPU	Central Processing Unit
DAHDI	Digium Asterisk Hardware Device Interface
DLSR	Delay Since Last Sender Report Received
FEC	Forward Error Correction
GUI	Graphical User Interface
GUI	Graphical User Interface
IAX	Inter Asterisk Exchange
IDE	Integrated Development Environment
ITU	International Telecommunication Union
LSR	Last Sender Report
MOS	Mean Opinion Score
OLA	Overlap and Add
P2P	Point to Point
PBX	Private Branch Exchange
PLC	Packet Loss Concealment
PSTN	Public Switched Telephone Network
RSVP	Resource Reservation Protocol
RTCP	Real-time Transport Control Protocol
RTP	Real Time Transmission Protocol
SCTP	Stream Control Transmission Protocol
SIP	Session Initiation Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UNIX	UNiplexed Information Computing System
VoIP	Voice Over Internet Protocol

1. INTRODUCTION

1.1. Background

Today communication services are offered through the network. It allows people to exchange and retrieve the information. The most popular and widely used service is Voice over Internet Protocol (VoIP). The reason for the popularity of VoIP is that it allows people to communicate with each other freely at low rates [1]. VoIP technology has allowed phone calls to be routed over Internet Protocol infrastructure rather than the traditional Public Switched Telephone Network (PSTN) infrastructure. The technology, called Voice over Internet Protocol (VoIP), uses the Internet Protocol (IP) to route packets containing small portions of voice conversations between the callers. In traditional technology of analog communication the poor voice quality in PSTN might result from the poor connections or old cables. In digital communication, the voice quality in VoIP is mostly dominated by the characteristics of packet networks such as delay, jitter, and packet loss. Therefore, it is important for us to take characteristics of IP network into consideration when we design a VoIP application. The transmission technology of VOIP is digital. Hence the caller's voice is digitized. The digitized voice is compressed [2]. Speech quality of VoIP may potentially be degraded by transmission errors such as packet loss, throughput delay, jitter and echo [3]. Major challenge of the VoIP network is maintaining quality i.e. packet loss which is a serious and critical issue for voice over Internet Protocol applications.

Issues on VOIP are:

- Delay in packet transmission from sender to receiver.
- Packets arriving too late at the receiver side.
- Heavy load on the network.
- Congestion of routers and gateways.
- The variations in packet inter arrival time create difference between when the packet is expected and when it is actually received is jitter.
- The loss of voice packets from sender to receiver.

1.2. Problem Statement

IP-based telephony is known to be more robust than analog-based phone network under a situation of large-scale disaster. However, under some situation such as a large scale disaster, congestion of VoIP traffic is inevitable. Not only has the traffic issue, the disaster such as an earthquake, flood or tsunami may destroy network facilities, which make the situation severe. Under a congested network, a real-time audio communication such as VoIP needs a packet loss concealment (PLC) technique [4]. For Voice Communication packet loss causes loss of information and interrupts voice communication. Also user experience and Quality of Service is affected when disturbance in voice exists.

Solution to overcome the effects of packet loss in voice communication PLC technique is used. In PLC technique lost data is interpolated using previously received good data stored in memory buffer. Voice data did not deviate more from previous data so lost voice frame can be copied from previous frame. Also to smooth the intermediate transition period between real data and copy of data from buffer memory synthesis of voice data is used. Many techniques are available for interpolation. In this thesis have used Average Magnitude Difference Function (AMDF) to estimate pitch (fundamental frequency of data sample).During the process of pitch estimate we have some length of data available in buffer. Calculation window is defined with range of frequency available. Frequency of maximum 200Hz and minimum 66Hz period is used with sample rate of 8 KHz. When pitch value is available Overlap and Add (OLA) technique is used to interpolate the lost data. Thus PLC technique is used to overcome effects of packet loss in voice communication.

1.3. Objective(s)

The main objectives of this thesis are:

- To perform analysis on Packet Loss Concealment techniques over voice signals for reliable communication.
- To develop a Packet Loss Concealment approach to optimize the control and recovery of packet loss over VoIP communication.

1.4. Scope

Many problems are faced with the issue in Quality of Service in voice communication. In network communication packet loss is a major problem. The problem cannot be avoided but it can be minimized. This thesis is about minimizing the effect due to loss of voice data during transmission. There are numerous applications of this thesis where online or offline (stored files) audio communication is involved.

2. THEORETICAL EXPLANATION & LITERATURE REVIEW

2.1. Background

M.Chetty, E.Blake and E.McPhie define VoIP as referring “to a range of protocols designed to send voice over packet switched networks, traditionally the domain of internet traffic.” In VoIP voice is sampled at a certain frequency which can be set to any desired value on the devices in use [5]. The sampled voice is then digitized and then finally packaged into packets before being sent over the IP network. VoIP uses different protocols for the call setup and the actual conversation between two communicating telephones. Signaling protocols like SIP, H.323 and IAX are used for the call setup and then the Real Time Transmission Protocol (RTP) is used to carry the voice between the telephones. RTP has been specifically optimized to for the transmission of real time data.

There are advantages of VoIP over the Traditional Public Switched Telephone Network. These advantages include the fact that VoIP makes more efficient use of bandwidth by only transmitting when something useful is being sent. This is done by silence compression which cuts out the bits when there is silence in a conversation thereby minimizing the packets sent across a network. Also VoIP unlike the traditional PSTN does not require a dedicated link between two end points since it uses packet switching instead of circuit switching for communication, and this allows the network to be used for multiple conversations concurrently [5].

W.Yu et al give a detailed suggestion of how to implement a Point to Point (P2P) network layout for a VoIP network but however do not give any comparisons between the P2P network layout and the client/server layout. W.Yuetal however mention that experimental results demonstrate that the P2P hierarchy model for conferencing applications can achieve better performance than all other VoIP models in terms of minimizing the network bandwidth overhead. This certainly suggests an angle to the deployment of the system being developed in this project should the deployment become large scale in the future [6].Asterisk VoIP system are for easy service because it allows various telephony protocols to communicate in a transparent manner and also Asterisk acts as a middleware between services and telephony technologies which makes Asterisk the ideal platform for developing services for a converged

VoIP network. Asterisk at the moment has the potential to be expanded to do much more than the current subset of telephony applications that is being used [7].

The use of VoIP services is quickly growing up. Although, there are many active corporations that offer and implement VoIP services. For example, Microsoft Net meeting and Skype application. VoIP services demand is getting increase everyday as VoIP costs is far less for the customer than the traditional telephone system in particular for international calls. One of the main advantages of VoIP services is, VoIP services provides superior management for provider involved and also cost is very less as organization could have single set-up for both data and voice as well, so it is beneficial for both customers and as well as service providers [8].

Packet Loss Control technique is beneficial for all the services that use IP communication. The concern for voice data is that it must be in real time, so it is of maximum priority that communication are robust and lossless. There always exists loss, delayed or dropped voice packet during real time communication. So PLC technique is required to monitor and interpolate what voice pattern that are received from the network. In this scenario, there is no any extra bandwidth consumption or additional request to sender. Receiver can compute and interpolate voice packet by itself and maintain the Quality of service.

The working model for VoIP system is summarized below:

- VoIP client (smartphone mobile with VoIP app) will request voice call communication for next client in a network.
- VoIP server will initiate and control all processes and allows audio communication between two clients.
- VoIP server checks performance of network between two clients for uninterruptable voice call communication in network with acceptable packet loss range.

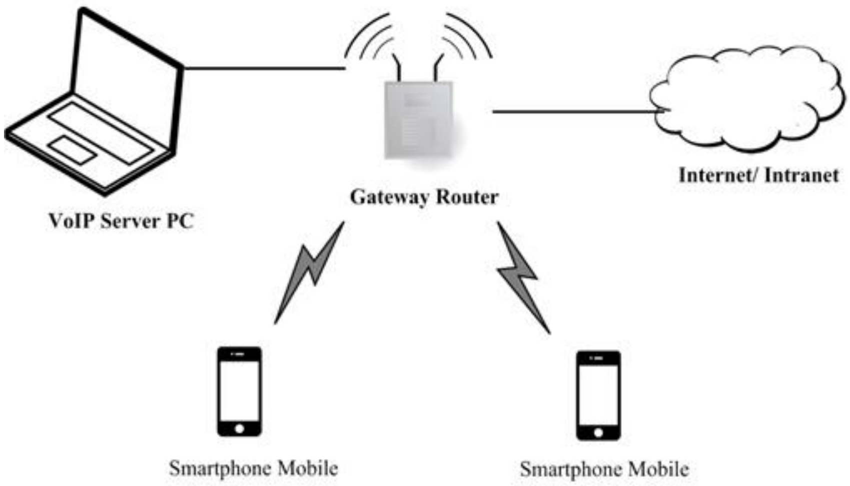

Figure 2.1-1: VoIP Model Experimental Setup

1.2. Packet loss Concealment Techniques

Packet loss concealment (PLC) is a technique to mask the effects of packet loss in VoIP communications. Voice signal is sent as packets on a VoIP network, they may travel different routes to get to destination. At the receiver a packet might arrive very late, corrupted or simply might not arrive. In a VoIP connection, error-control techniques are not feasible and the receiver should be able to cope with packet loss using PLC techniques. Packet loss concealment becomes difficult to implement effectively as the ratio of dropped packets to total packets increases. When multiple consecutive packets are dropped, PLC is less effective than in cases where a single packet is delayed or missing.

Zero Insertion: Lost speech frames are replaced with zero so there is no voice disconnect at receiver side. It is applicable for very stable network scenario (where network loss is minimum). The technique can be applied during digital-to-analog conversion at receiver end.

Waveform Substitution: Reconstruction of a lost frame by repeating a portion of already received speech. The simplest form of this would be to repeat the last received frame.

if (samples = 0)

- request stored value of previous voice frame
- replace current frame with last known voice frame

Model-Based Methods: Speech models to interpolate speech gaps in speech due to lost frame.

if (samples = 0)

- Request stored value of previous voice frame in buffer
- Use defined algorithms to interpolate
- replace current frame with interpolated frame
- receive previous voice data from stored memory
- estimate and set the playout time for packet i + 1
- Interpolate packet i with defined configuration algorithm within playout time;
- Reply interpolated packet.

PLC algorithms are divided into two categories which are:

(a) Receiver-Based Methods

Only data available at the receiver is used for the concealment. This method doesn't increase bandwidth requirements nor delays. These technique is based under the assumption that there is a small difference between two neighboring packets, which is true in case of voice.

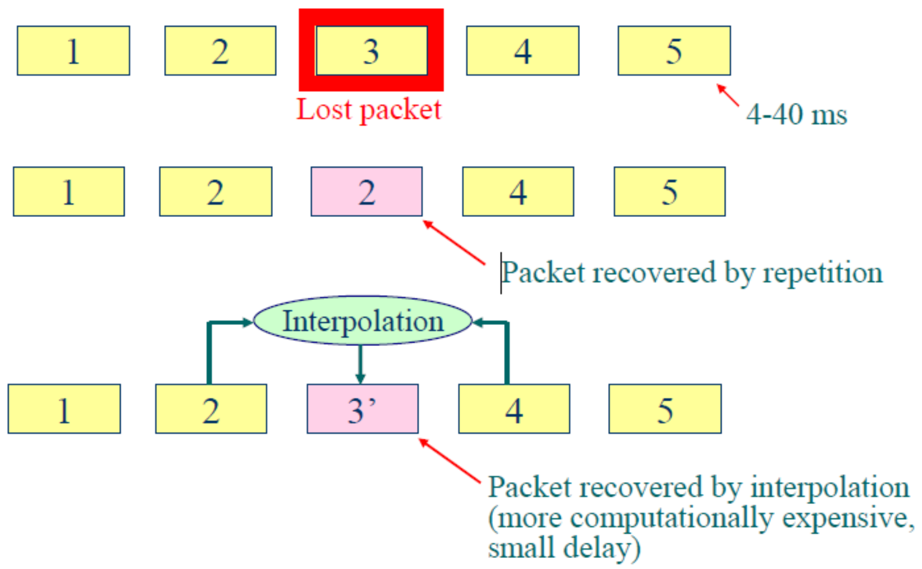

Figure 1.2-1: Receiver Based Interpolation of Voice Data

(b) Sender-Based Methods

Sender adds some redundancy or additional side information to voice frames that the receiver can later use for the concealment process.

1.3. Packet Loss Concealment Algorithms

Audio voice is continuous stream of data. Filtering of long sequence voice data is not possible due to limitation of memory size digital computer/processor. The input sequence is divided into number of blocks and output blocks are computed for respective input blocks. The input blocks are filtered together to produce the complete output sequence.

Average Magnitude Difference Function (AMDF)

AMDF is the absolute value of the difference between the original signal and the delay signal to calculate the fundamental frequency i.e. Pitch of that signal. AMDF decrease the computation complexity which make it more suitable for the real-time applications to calculate their pitch.

AMDF pitch detection algorithm:

$D_{AMDF}(m) = 1/N \left(\sum_{N=0}^{N-1} |x(n) - x(m+n)|\right)$

Where x (n) are the samples of speech. For a periodic signal with period T_0, this function is expected to have a strong minimum when the index m equals T_0. The pitch period is, in general, estimated as follows:

$T_0 = MIN(D_{AMDF}(m))$, for $m = m_{min}$ to m_{max}

Where the values of m_{min} and m_{max} are chosen to cover the expected pitch-range [9].

OLA (Overlap and Add)

Input sequence is divided into number of block size

N=L+M-1 (M-length of impulse response)

Each block consists of first L points from input sequence tailed by M-1 zero.

Consider,

X (n) its length is Ls and impulse response h (n) of length M.

Now divide the input sequence into blocks, hence it will become

x1(n)={x(0),x(1),….x(L-1),0,0...0,0,0}

First L data points M-1 zeros

From Input sequence

x2(n)={x(L),x(L+1),….x(2L-1),0,0...0,0,0}

Next L points form input   M-1 zero

x3(n)={x(2L),x(2L+1),….x(3L-1),0,0...0,0,0}

Next L points form input   M-1 zero

So by using these segmented blocks of input the respective output blocks are computed using Fourier transform. The linear convolution of sequences is always longer than the original sequences. Then output blocks are fitted together to form the complete output. In this thesis OLA with linear convolution is used.

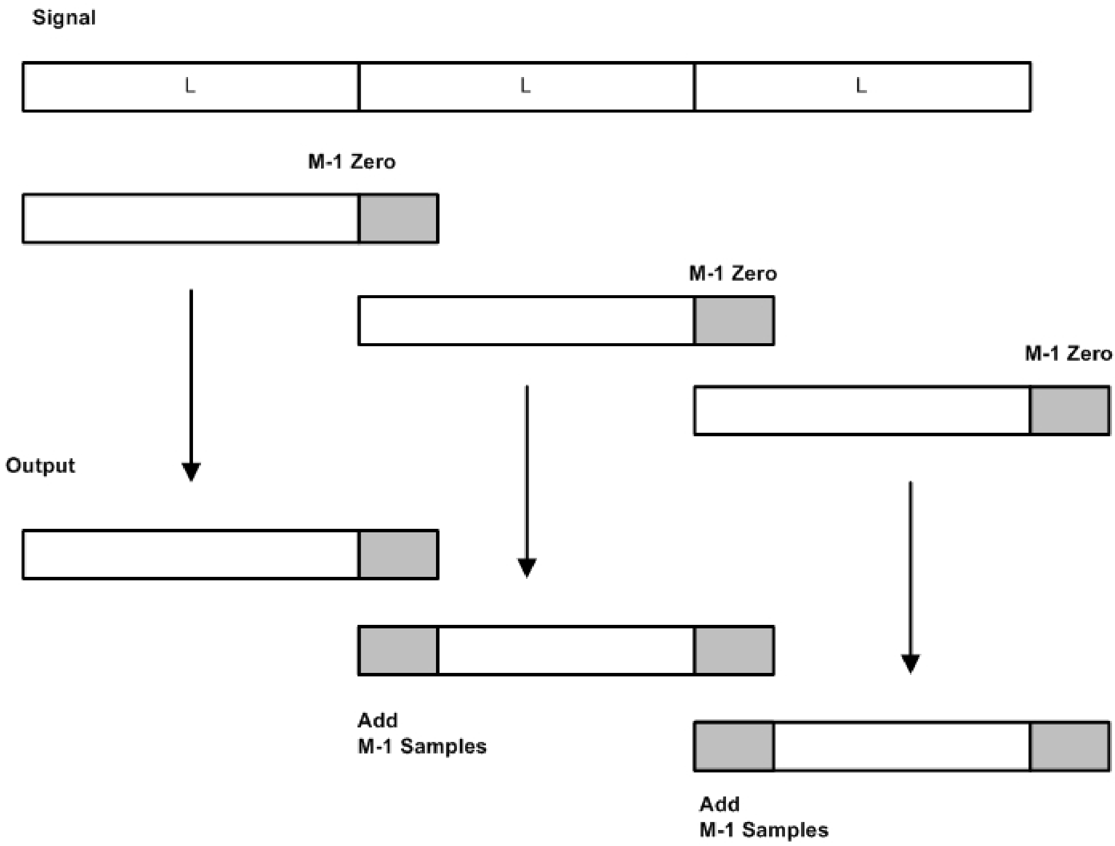

Figure 1.3-1: Overlap and Add

1.4. Theoretical Explanation of Related Terms

Voice over Internet Protocol Server

VoIP servers are internet-based telephone system. VoIP server are hardware, software or both. The VoIP server works in a similar way like a proxy server: Making a call from a VoIP or a softphone to a VoIP phone can be done via the Internet. Calling an ordinary or mobile phone is done via PSTN (ordinary telephone system). After Implementing VoIP server extensions can be configured to set services provided from server. Low-cost VoIP services like telephoning, video-calling, voice mails are some services provided from VoIP server.

Coding/ Compression

In order to reduce bandwidth consumption in the transmission of speech signals, speech coding is employed to compress the speech signals, i.e. on one hand to use as few bits as possible to represent them and on the other hand to maintain a certain desired level of speech quality.

CODECs

A knowledge of CODECs is necessary for a complete understanding of network capacity planning analysis. Voice is an analog signal that must be converted to a digital signal for transmission over a digital network. Converting analog to digital is called encoding, and converting digital to analog is called decoding. A CODEC (encoder/decoder) is used to convert in both directions [10].

Speex CODECs

Speex is an Open Source/Free Software patent-free audio compression format designed for speech. Speex is well-adapted to Internet applications. Speex is designed to compress voice at bitrates ranging from 2 to 44 kbps. Speex provide noise suppression feature and also provide narrowband (8 kHz), wideband (16 kHz), and ultra-wideband (32 kHz) compression in the same bit stream.

RTP

RTP is the Internet-standard protocol for the transport of real- time data.IT can be used for media-on-demand as well as interactive services such as Internet telephony. The RTP standard actually defines a pair of protocols: RTP and RTCP.RTP is used for exchange of multimedia data, while RTCP is the control part and is used to periodically obtain feedback control information regarding the quality of transmission associated with data flows. RTP usually runs over UDP/IP.RTP and the associated RTCP use consecutive transport-layer ports can be used over UDP. The blocks formed by such bit streams are encapsulated in RTP packets and then in UDP and IP packets.

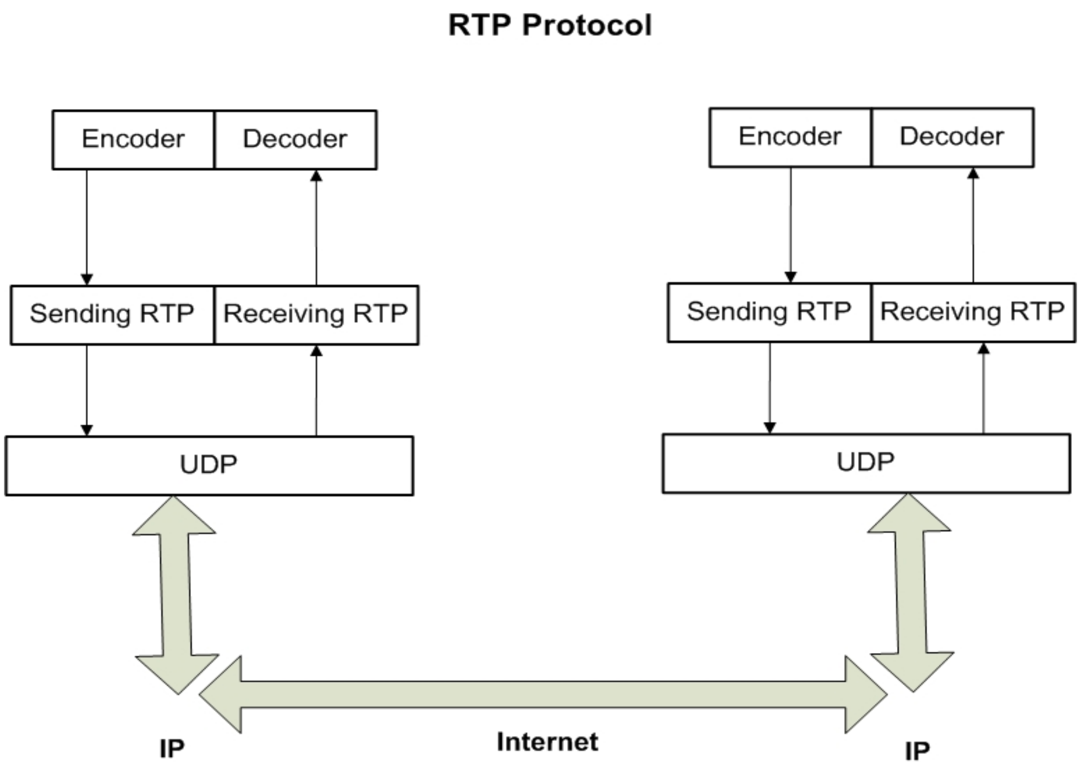

Figure 1.4-1:Real-time Transport Protocol

RTCP

RTCP is the control protocol designed to work in conjunction with RTP.RTCP packets contain sender/receiver reports that announce statistics such as number of packets sent, number of packets lost, and jitter. This may be useful for adaptive application that can used to send high or low quality data depending on the network congestion

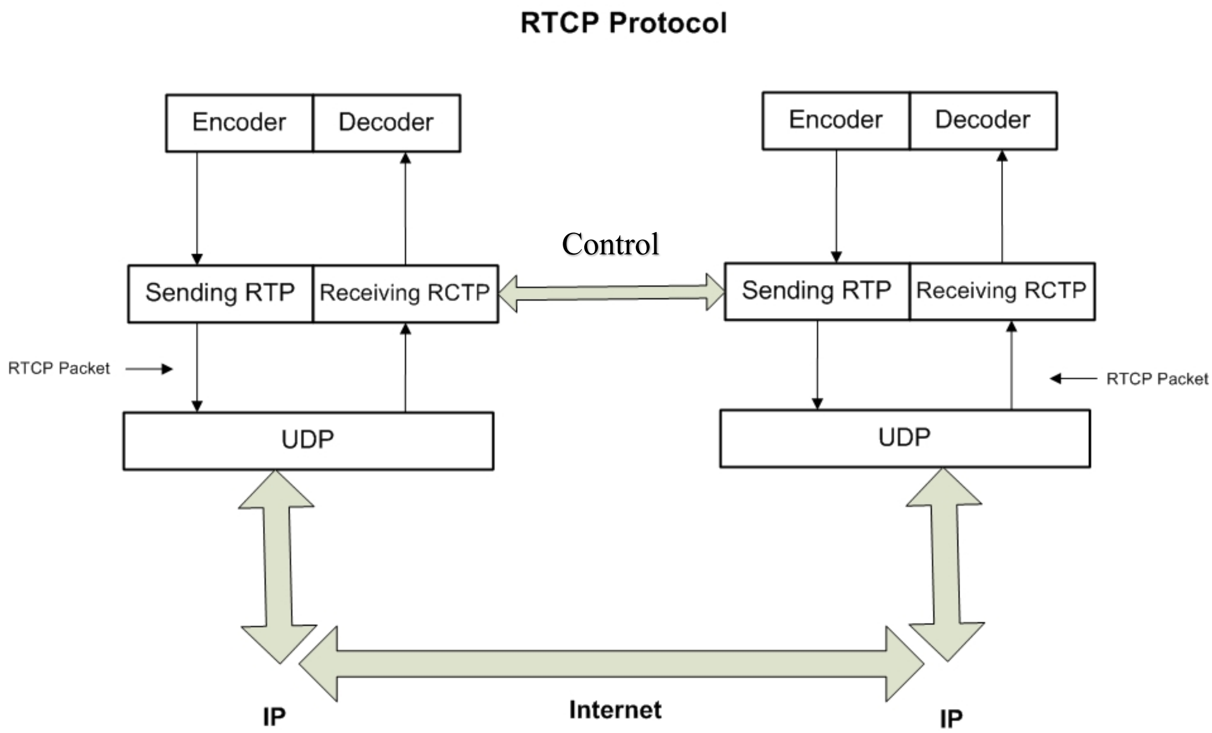

Figure 1.4-2:Real-time Transport Control Protocol

Every RTCP packet starts with a header similar to that of the RTP data packets. The payload type field identifies the type of the packet.

Round-Trip Delay

Receiver reports is used to estimate the round-trip delay between sender and receiver. The receiver report includes the LSR (timestamp from the last sender report received) and DLSR (delay since last sender report received) fields.

Sender can calculate the round-trip delay according the formula

$$D = A\text{-}LSR\text{-}DLSR \quad (1)$$

Where A is the time instant when the receiver report was received by the sender.

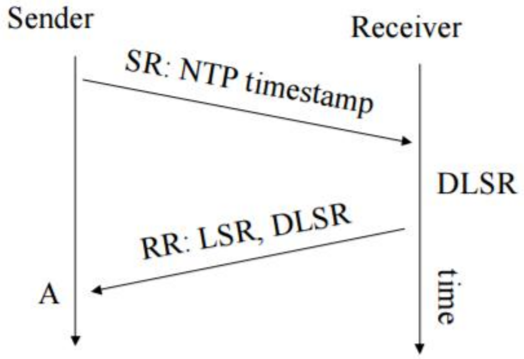

Figure 1.4-3: Calculation of Round-Trip Delay

Inter-Arrival Jitter

Receivers continuously observe the variance of the inter-arrival time of incoming RTP packets. An estimate for inter-arrival jitter is calculated as follows.

Firstly, the difference D in packet spacing at the receiver compared to the packet spacing at the sender is calculated:

$$D = (Rj-Ri)-(Sj-Si) \qquad (2)$$

Where R is the time of arrival and S is the RTP timestamp for a certain packet.

This delay variation value is calculated after every RTP packet.

To avoid temporary fluctuations the final value for inter-arrival jitter estimate is smoothed according to equation

$$Ji = (15/16)\, Ji\text{-}1 + (1/16)\, D \qquad (3)$$

Equation (3) only gives a small weight to the most recent observation. The change in this jitter estimate could indicate congestion before it leads to packet loss [10].

Fraction Loss

In RTCP the receiver reports also contain information about the lost packets. The fraction of lost packets is defined to be the number of packets lost divided by the number of packets expected, which are calculated based on actually received packets and the highest sequence number received in RTP packets. A cumulative number of packets lost is also maintained. The fraction lost field indicates the number of packets lost divided by the number of packets expected (according to the highest sequence number received) since last receiver report [10].

Table 1: Format of the Receiver Report

SSRC of the sender	
SSRC of the first source	
Fraction Lost	Cum. no of packets lost
Ext. highest sequence number received	
Inter arrival jitter estimate	
Last sender report timestamp (LSR)	
Delay since last sender report (DLSR)	
…	
Last reception report block	

The steps involved in VoIP transmissions are:

Step 1. At the source, the voice is sampled.

Step 2. The voice data is encoded by encoder to save bandwidth.

Step 3. The encoded bit stream is converted into packets.

Step 4. The packets sent along the network.

Step 5. At the destination, these packets are reassembled.

Step 6. The decoder converts the voice data back into voice samples.

Step 7. These samples are put in a buffer. [4]

VoIP Model can be divided into the following planes:

Signaling Plane: Signaling Plane is a set of application level protocol (e.g. Session Initiation Protocol (SIP)) used for call setup management. These protocols is used together to provide a broader range of VoIP services.

Support Plane: - Protocols used to support service provisioning in the signaling plane like RSVP (Resource Reservation Protocol) and RTP (Real Time Transport Protocol).

User Plane: -This plane that carries the real user traffic, in form of packetized voice data in VoIP.

UDP (User Datagram Protocol)

After the adoption of new Internet protocol there is streaming revolution in transmission network. (UDP) and new encoding techniques compressed audio files into extremely small packets of data. UDP made streaming media feasible by transmitting data more efficiently from the host server over the internet to the client. Real Time Streaming Protocol (RTSP) are making the transmission of data even more efficient.

UDP is a connectionless transport layer protocol. Meaning that the communication occurs between hosts with no previous setup, packets that are sent between two hosts may take different routes. This provides a simple and unreliable message service for transaction oriented services. Unlike the TCP, it adds no reliability, flow-control, or error-recovery functions to IP. Because of the UDP's simplicity, UDP headers contain fewer bytes and consume less network overhead than TCP and are therefore more suitable for real-time transmissions. UDP protocol is also referred to as packet switching and is used for optimization of the use of the bandwidth available in a network and to minimize the latency. . Unlike TCP transmission, when a UDP audio packet drops out, the server keeps sending information, causing only a brief glitch instead of a huge gap of silence. TCP, on the other hand, keeps trying to resend the lost packet before sending anything further, causing greater delays and breakups in the audio broadcast [11].

Packet loss

Packet loss often occurs in the Network when a router becomes congested, i.e. it receives more packets to forward than it can process. Another reason for loss can be transmission errors (bit errors) of the underlying medium. Typically the bit error ratio is extremely low however for fixed networks (but it can be significant for wireless networks).Packet loss in the Network is a frequent and also the most serious problem that speech transmissions over the Network have to face. Applications can use message sequence number of transport protocols such as RTP to detect a packet loss. In order to provide an acceptable quality, loss recovery/control must be performed. While coding schemes can exploit the redundancy within the speech efficiently for compression, together with packet loss compression can lead to even more significant degradations of the output speech quality as for speech. The time interval in which the decoder does not receive data from the network is a crucial parameter with regard to user perception (if either a loss is perceived not at all, as a glitch or as a dropout).

1.5.Recovery techniques used in VoIP

Plain Delivery: This technique simply bundles each block of encoded audio data packets into an IP packet and transmits it. It does not provide any sender-based effort to improve audio quality when packet loss occurs.

Interleaving: This technique attempts to reduce the degradation of perceptual audio quality by distributing lost data into several small gaps instead of having one large gap of lost data. This technique requires the same bandwidth Utilization that plain delivery uses since it does not transmit additional information.

Forward Error Correction: FEC is a sender-based technique for mitigating the undesired effects of packet loss. This works by transmitting redundant packets for error correction.

Retransmission: The receiving endpoint implements a loss-detection algorithm to detect lost packets; if any packets are lost this technique will resends the lost data packets upon request by the recipient. This technique is used only at the explicit request of the recipient, so it requires a round trip time that inherently induces a large end-to-end delay.

1.6.Factors Affecting VoIP Quality

In VoIP, quality means the ability to talk and listen clearly without any unwanted noise. The three major factors that affect the speech quality in VoIP are: Packet Loss, Jitter and Latency.

Packet Loss

Packet loss happens when two users talk on a VoIP phone and the call begins to “break up” during packet transfer over the IP network. This mainly happens when there is a lot of congestion on the network and it results in some part of the conversation being missed at the receiving end. Thus to minimize the effect of packet loss, all VoIP applications should have a module for packet loss concealment.

Jitter

In VoIP, when a call is established, the sender sends the VoIP packet at a constant frequency, for example 10ms or 20ms. As the packet travels over the internet it might get delayed or lost, resulting in receiver not receiving the packets with the same frequency. The difference in the expected time of arrival and the actual time of arrival of the packet is called jitter. It is usually experienced in heavily congested networks.

Delay or Latency

The total amount of time taken for speech to travel from speaker's mouth to listener's ear is measured as delay or latency. In term of VoIP packets, it is the total time taken by the packet to travel from the source to the destination.

Nepal Telecom Authority Quality of Service Specifications for VoIP:

Table 2: Nepal Telecom Authority Quality of Service Specifications for VoIP *[12]*.

Parameters	Toll Quality	Below Toll Quality
1. Packet Loss	< 0.1 %	< 2%
2. Packet Jitter	<5 ms	< 10 ms
3. Delay (One-way End-to-End)	<=150 ms (without satellite link) <=400 ms (with satellite link)	>150 ms (without satellite link) >400 ms (with Satellite Link)
4. Grade of Service (GoS)	<2%	<2%
5. CCS7 signaling delay	To conform with Q.709	To conform with Q.709
6. MOS(Mean Opinion Score) or R-Value	>=4(without satellite Link) And >=3.5 (with satellite link) or, >=80 (without satellite link) and >=70 (with satellite link)	<4(without satellite link) And <3.5(with satellite link) or, <80(without satellite link) and <70(with satellite link)

Jitter Buffer:

In Voice over IP (VoIP), a jitter buffer is a shared data area where voice packets can be collected, stored, and sent to the voice processor in evenly spaced intervals. The jitter buffer, which is located at the receiving end of the voice connection, intentionally delays the arriving so that the end user experiences a clear connection with very little sound distortion.

Jbmaxsize = 200 ms is used for analysis in this thesis.

3. RESEARCH METHODOLOGY

3.1. Concept

Modern IP network services provide for the simultaneous digital transmission of voice, video, and data. These services require loss control protocols and algorithms which can solve the packet loss parameter for voice data for Quality of control transmission for IP network. Among different method available Packet Loss Concealment (PLC) is a method to overcome data loss issue for voice data. In this thesis receiver based PLC technique is applied to overcome loss of voice frame during transmission.

3.2. Technical Framework

A VoIP server with SIP clients is set up to perform voice communication in a network. Also VoIP monitoring software is integrated in same network to view real time network and call status. Monitoring software records audio for real time voice communication, call detailed records (CDR) and call management records (CMR). This files are later used for analysis purpose.

System Model and working flowchart is attached in Annex III.

List of VoIP servers

Free SWITCH

Free SWITCH is a scalable open source cross-platform telephony platform designed to route and interconnect popular communication protocols using audio, video, text or any other form of media. It was created to fill the void left by proprietary commercial solutions. Free SWITCH also provides a stable telephony platform on which many telephony applications can be developed using a wide range of free tools.

Free PBX

Free PBX is the most widely deployed Open Source PBX platform in use across the world. The openness of the project allows Users, Resellers, Enthusiasts and Partners to utilize the Free PBX to build robust communications solutions that are powerful but at the same time easy to implement and support. In Background Free PBX uses Asterisk platform.

XiVO

XiVO is a scalable platform based on Asterisk that offers a free and complete PBX solution. It is interoperable with most market telephony systems and allows its users to benefit from a set of flexible and advanced services without requiring specialized hardware.

VoIP Switch

VoIP Switch is a Soft Switch System with routing and billing, with SIP protocol support. Basic package includes Soft switch app, IVR app, configuration manager, web dialer and portal for end clients.

Asterisk

Asterisk is an open source framework for building communications applications. Asterisk turns an ordinary computer into a communications server. Asterisk is a software implementation of a telephone Private Branch Exchange (PBX).It allows telephone to make calls to one another, and to connect to other telephone services, such as the Public Switched Telephone Network (PSTN) and Voice over Internet Protocol (VoIP) services.

Most of the open source VoIP server use Asterisk platform as their background. Different VoIP servers and its application are developed on Asterisk platform to meet their requirements. In this thesis Asterisk (VoIP Service platform) is used as a framework for development of VoIP server.

Asterisk GUI is a framework for the creation of graphical interfaces for configuring Asterisk. This system is optional using asterisk VoIP platform, However web based server management support easy configuration and data collection environment.
Asterisk GUI Graphical Interface snapshot is attached in Annex II.

System Installation

- Download Asterisk Source Code
- Download Required Project and tools as required for Asterisk (libpri, DAHDI, pjproject, required codec translation support)
- Make changes to source code as required according to our requirement.

- Compile source code.
- Configure Extension and VoIP System Requirements.

Configure Extensions

```
[dipak]
callerid="dipak" <6000>
username=6000
host=dynamic
secret=********
regcontext=dipak-internal
regexten=6000
nat=yes
```

For this thesis a working lab for VoIP server with SIP clients in a network domain is set up. Real time transmission and monitoring of data is available so there is no need of any external data source. Call status, jitter buffer status and value of packet loss, are extracted from system CLI (Command Line Interface) during call and VoIP log file as per requirement.

3.3. Measurement of Speech Quality

In VoIP, quality simply means being able to listen and speak in a clear and continuous voice, without unwanted noise. QoS (Quality of Service) is essentially a service that prioritizes certain data traffic by slowing less important data packets down. These important packets then reach their destination as quickly as possible. Speech quality is the measurement of user experience when a VoIP call is established. The measurement of speech quality is divided into two categories, Objective and Subjective. Subjective method are based on user listening tests. Users are told to rate the speech quality. These tests are tedious to perform and the accuracy of speech quality rating relies on the user's mood. The subjective methods are not practicable during the network planning phase. These methods are limited, impracticable and too expensive. In order to avoid these problems, new methods that permit the calculation of values representing the different damaging factors combinations of the network have been developed. The quality estimation, providing results as near as possible to Mean Opinion Score (MOS) values. The objective measurement techniques development uses an approach where a voice sample represents the input signal to produce a score, representing the original signal produced by the network [13].

3.3.1. Mean Opinion Score

Mean Opinion Score (MOS) is a numerical value to indicate the quality of the call from the user's perspective of the received call after compression, transmission, and decompression. MOS is a calculation based on the performance of the IP network. The Mean Opinion Score Values Taken in whole numbers, the numbers are quite easy to grade.

Table 3: Mean Opinion Score

Value	Label	Description
5	Excellent	Perfect. Like face-to-face conversation or radio reception.
4	Good	Fair. Imperfections can be perceived, but sound still clear. This is (supposedly) the range or cell phones.
3	Fair	Annoying.
2	Poor	Very annoying. Nearly impossible to communicate.
1	Bad	Impossible to communicate.

The values do not need to be whole numbers. Certain thresholds and limits are often expressed in decimal values from MOS. For instance, a value of 4.0 to 4.5 is referred to as toll-quality and causes complete satisfaction. This is the normal value of PSTN and many VoIP services aim at it, often with success. Toll Quality is a voice call quality comparable to that of an ordinary long distance call, originally placed over the analog circuit-switched Public Switched Telephone Network (PSTN).

Grade of service is the probability of a call in a circuit group being blocked or delayed for more than a specified interval. GoS reference to the busy hour when the traffic intensity is the greatest. The Grade of Service is one aspect of the quality a customer can expect to experience when making a telephone call. The CCS7 is a Common Channel Signaling system. Instead of signaling being associated with each traffic channel, a common signaling channel is used for all circuits. This signaling channel consists of one or more signaling links.

In order to analyze calls quality, voice communication systems provide a way for extracting call related data and network information. This type of information is usually provided from

VOIP Server in well-defined structures named call detailed records (CDR) and call management records (CMR). CDR files usually contain information regarding the source and destination, the duration but also some network parameters like packets sent, latency, jitter, etc.

3.4. Development Tools and Environment

3.4.1. Asterisk 11.25.1

Asterisk is an open source framework for building communications applications. Asterisk turns an ordinary computer into a communications server. Asterisk is a software implementation of a telephone Private Branch Exchange (PBX).

3.4.2. Ubuntu 14.04

Ubuntu is an operating system based on the Debian Linux distribution and distributed as free and open source software, using its own desktop environment.

3.4.3. Libpri

libpri is an open source library that encapsulates the protocols used to communicate over ISDN Primary Rate Interfaces (T1, E1). libpri is a dependency for Asterisk and DAHDI if PRI (Primary Rate Interface) signaling is used.

3.4.4. DAHDI

Digium Asterisk Hardware Device Interface (DAHDI) is a collection of open source drivers, for Linux, that are used to interface with a variety of telephony related hardware.

3.4.5. Android Studio

Android Studio is the official integrated development environment (IDE) for Android platform development. Android Studio offers features that enhance our productivity when building Android apps.

3.4.6. C Sip Simple

CSIP Simple is a free SIP client for the Android OS. This application is installed in android phone and configured with VoIP server information to make calls between clients. Also when trunk line is connected with VoIP server with phone service provider, SIP client is used as a normal phone.

CSIP Simple login interface snapshot is attached in Annex II.

3.4.7. IPTABLES

Iptables are basics of Firewall for Linux operating system. Iptables is a rule based firewall system and it is normally pre-installed on a UNIX operating system. These firewall rules are used for controlling the incoming and outgoing packets. By-default the Iptables is running without any rules, we can create, add, and edit rules. The Iptables firewall operates by comparing network traffic against a set of rules. The rules define the characteristics that a packet must have to match the rule, and the action that should be taken for matching packets.

3.4.8. VoIPmonitor

VoIPmonitor is open source network packet sniffer with commercial frontend for SIP RTP and RTCP VoIP protocols running on Linux. VoIPmonitor is a software to analyze quality of VoIP call based on network parameters - delay variation and packet loss according to ITU-T G.107 E-model which predicts quality on MOS scale. Calls with all relevant statistics are saved to database. Optionally each call can be saved to Pcap file with either only SIP protocol or SIP/RTP/RTCP protocols. VoIPmonitor can also decode speech and play it over the commercial WEB GUI or save it to disk as WAV.

International Telecommunication Union (ITU) is a specialized agency of the United Nations (UN) that is responsible for issues that concern technologies. Telecommunication Standardization sector of ITU (ITU-T) gives the algorithm for the wideband version of the E-model as the common ITU-T transmission rating model for planning speech services. This computational model can be useful to transmission planners, to help ensure that users will be satisfied with end-to-end transmission performance. The primary output of the model is a scalar rating of transmission quality. A major feature of this model is the use of transmission impairment factors that reflect the effects of different types of degradations occurring on the entire transmission path, mouth-to-ear [14].

VoIPmonitor system snaps is attached in Annex IV.

3.4.8.1. E-Model

The E-model is a transmission planning tool that provides a prediction of the expected voice quality, as perceived by a typical telephone user, for a complete end-to-end (i.e. mouth-to-ear) telephone connection under conversational conditions. The E-model takes into account a wide range of telephony-band impairments, in particular the impairment due to low bit-rate coding devices and one-way delay, as well as the "classical" telephony impairments of loss, noise and echo. It can be applied to assess the voice quality of wireline and wireless scenarios, based on circuit-switched and packet-switched technology [14]. The primary output of the E-model calculations is a scalar quality rating value known as the "Transmission Rating Factor, R". The E-model is based on a mathematical algorithm, with which the individual transmission parameters are transformed into different individual "impairment factors" that are assumed to be additive on a psychological scale.

The relation between the different impairment factors and R is given by the equation:

$$R = Ro - Is - Id - Ie, eff + A$$

Ro expresses the basic signal-to-noise ratio (received speech level relative to circuit and acoustic noise).Is represents all impairments that occur more or less simultaneously with the voice signal, such as: loud speech level, side tone, noise etc. Id sums all impairments due to delay and echo effects. Ie, eff is an "effective equipment impairment factor", which represents impairments caused by low bit-rate codecs. A is an "advantage factor", which allows for an "advantage of access" for certain systems relative to conventional systems, trading voice quality for convenience. While all other impairment factors are subtracted from the basic signal-to-noise ratio Ro, A is added and thus compensates other impairments to a certain amount. It can be used to take into account the fact that the user will tolerate some decrease in transmission quality in exchange for the "advantage of access".

Table 4: Definition of categories of speech transmission quality

Range of E-model Rating R	Speech transmission quality category	User satisfaction
90<= R <100	Best	Very satisfied
80<= R <90	High	Satisfied
70<= R <80	Medium	Some users dissatisfied
50<= R <70	Low	Many users dissatisfied
50<= R <60	Poor	Nearly all users dissatisfied

Connections with E-model Ratings R below 50 are not recommended. Although the trend in transmission planning is to use E-model Ratings R, equations to convert E-model Ratings R into other metrics, e.g. MOS, can be found in ITU-T Rec. G.107 [14] .

3.5. Working Model

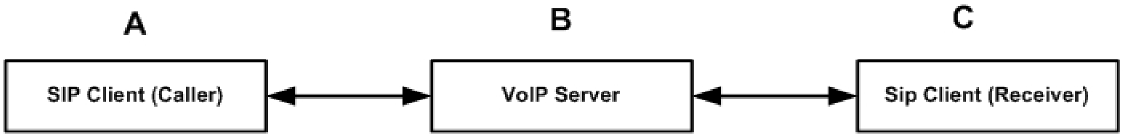

Figure 3.5-1: Deployed Model

Scenario I: When buffer memory is disable at Server B

- Client A (sender) starts a normal Call.
- Call request for client C (receiver) is acknowledged by VoIP server B.
- VoIP server B checks availability of Client C (receiver).
- Establish Call between client A and Client C via VoIP server B.

Scenario II: When buffer memory is enable at Server B

- Client A (sender) starts a normal Call.
- Call request for client C (receiver) is acknowledged by VoIP server B.
- VoIP server B checks availability of Client C (receiver).
- Server B checks status of buffer.
- Establish Call between client A (sender) and Client C (receiver) via VoIP server B using buffer memory to process voice data.
- When packet loss is detected at VoIP server B one of the two cases are used :
    - Copy loss frame from good voice frame from buffer.
    - Synthesis voice frame from previous good frame data available from buffer.

SIP (Session Initiation Protocol)

SIP provides mechanisms for establishing calls over IP. The protocol defines the messages that are sent between endpoints, which govern establishment, termination and other essential elements of a call. SIP can be used for creating, modifying and terminating sessions consisting of one or several media streams. SIP is an application layer protocol designed to be independent of the underlying transport layer. It runs on the Transmission Control Protocol (TCP), the User Datagram Protocol (UDP) or the Stream Control Transmission Protocol (SCTP).

SIP Algorithm:

- Users in a SIP network are registered and identified by unique SIP addresses.
- When a user initiates a call, a SIP request is sent to a SIP server.
- Call request includes the address of the caller and the address of the receiver (in the header field).
- Point-to-Point call is established using a SIP server.

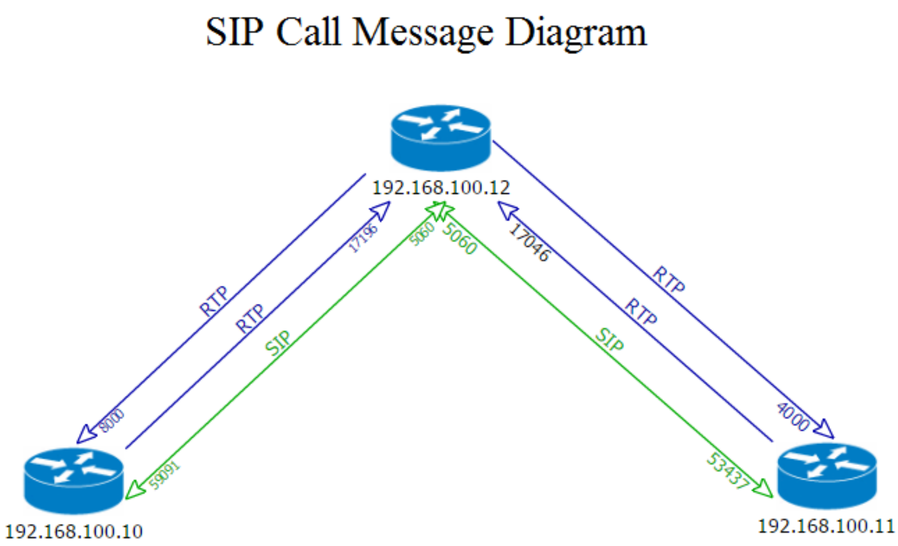

Figure 3.5-2: SIP Network Diagram

In above diagram we have a SIP client one with IP address 192.168.100.11 make a SIP phone call to SIP Client two with IP address 192.168.100.10 through Asterisk server 192.168.100.12.

4. RESULT ANALYSIS AND COMPARISION

The Analysis of loss during voice calls in Voice over IP (VoIP) is evaluated using two different technique. As voice frame signals are similar to previous good frame; when loss is observed at current lost portion, lost voice signal can be replaced by the same portion of signal from previous frame. This process does not affect Quality or degrade user perception and a normal speech communication is achieved. Next technique is interpolation, speech synthesis can be used to interpolate lost portion using the available data from previous speech frames. For lost signal some portion is interpolated using previous frame and remaining portion is just a copy of previous frame to reduce computational complexity and delay in processing.

Usually the QoS quality rating of VoIP is carried out using the subjective quality measure called Mean Opinion Score (MOS). MOS scales from 1 (lowest quality) to 5 (highest quality). This type of measurement focuses on the perceived quality provided by users. However, quality data for the study (packet loss rate, delay and jitter values) has been collected using VoIPmonitor network monitoring software. VoIPmonitor is a network packet sniffer software with commercial frontend for SIP RTP and RTCP VoIP protocols. Packet Loss values are obtained from RTP and RTCP responses from SIP client. The packet loss values are represented in the percentage form of the total RTP packets being transmitted. The delay values are obtained by calculating the difference between the RTP packet actual arrival time and the estimated arrival time. On the other hand, the jitter values have been derived from the differences in the inter-arrival time of the RTP packets. MOS is useful when considering the overall end-to-end quality of VoIP communications. For the ideal wireless LAN Environment, the default network setting with no packet loss is used.

Introducing fixed packet loss scenario in network using a UNIX operating system is attached in Annex V.

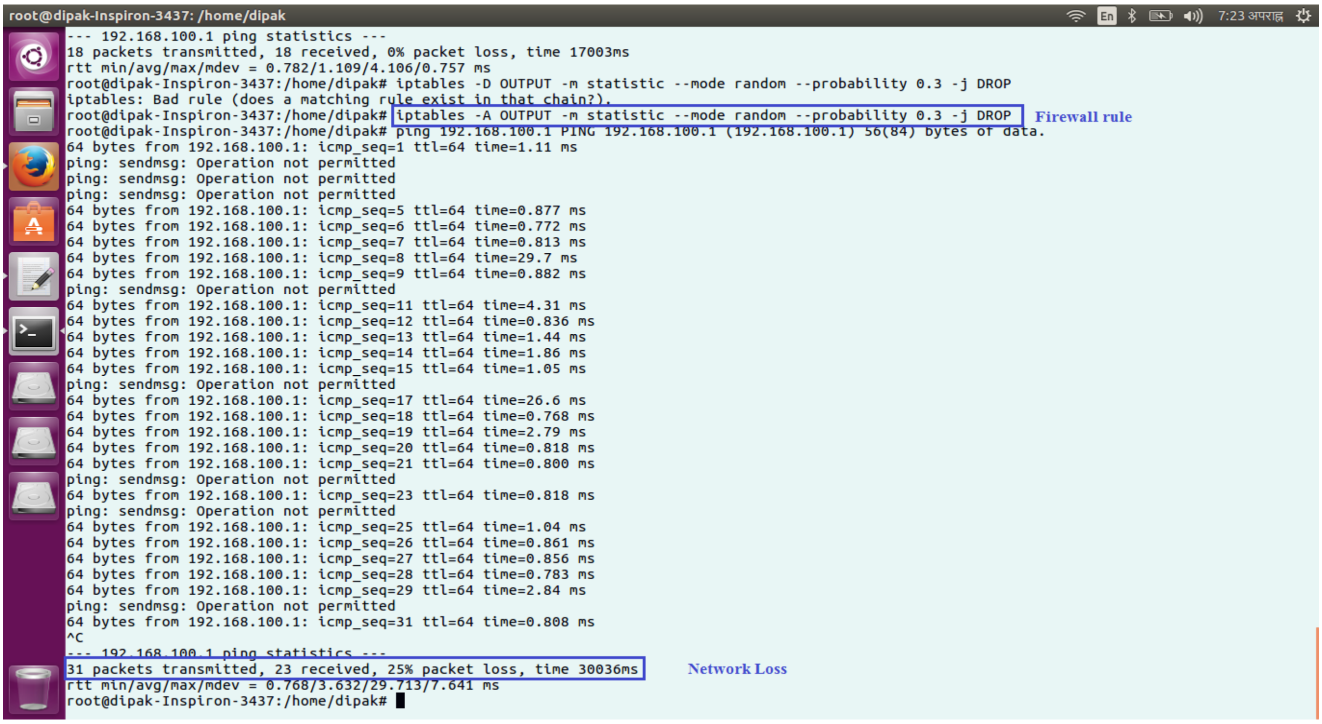

Figure 3.5-1: Linux firewall to random drop packet towards network

In Figure 3.6.1, firewall rule to drop in random probability of 30% packet is configured. During ICMP IP path ping 25% of the total packet sent is lost.

Packet Loss Status of server and client during real time voice call

Peer Call ID Duration Recv: Pack Lost (%) Jitter Send: Pack Lost (%) Jitter
192.168.100.11 29c49e8548b 00:08:23 0000024975 0000000064 (0.26%) 0.0000 0000023872 0000000211 (0.88%) 0.0136
192.168.100.8 Mo.0K5Mg-xA 00:08:23 0000024828 0000000179 (0.72%) 0.0000 0000024454 0000000223 (0.91%) 0.0135

Peer Call ID Duration Recv: Pack Lost (%) Jitter Send: Pack Lost (%) Jitter
192.168.100.11 29c49e8548b 00:08:25 0000025033 0000000064 (0.26%) 0.0000 0000023928 0000000211 (0.88%) 0.0182
192.168.100.8 Mo.0K5Mg-xA 00:08:25 0000024884 0000000179 (0.71%) 0.0000 0000024512 0000000223 (0.91%) 0.0178

Peer Call ID Duration Recv: Pack Lost (%) Jitter Send: Pack Lost (%) Jitter
192.168.100.11 29c49e8548b 00:08:25 0000025077 0000000064 (0.25%) 0.0000 0000023969 0000000211 (0.88%) 0.0173
192.168.100.8 Mo.0K5Mg-xA 00:08:25 0000024927 0000000179 (0.71%) 0.0000 0000024556 0000000223 (0.91%) 0.0174

Peer Call ID Duration Recv: Pack Lost (%) Jitter Send: Pack Lost (%) Jitter
192.168.100.11 29c49e8548b 00:08:26 0000025115 0000000064 (0.25%) 0.0000 0000024005 0000000211 (0.88%) 0.0190
192.168.100.8 Mo.0K5Mg-xA 00:08:26 0000024964 0000000179 (0.71%) 0.0000 0000024594 0000000223 (0.91%) 0.0175

Figure 3.5-2: Real Time Sip Channel status for loss and Jitter

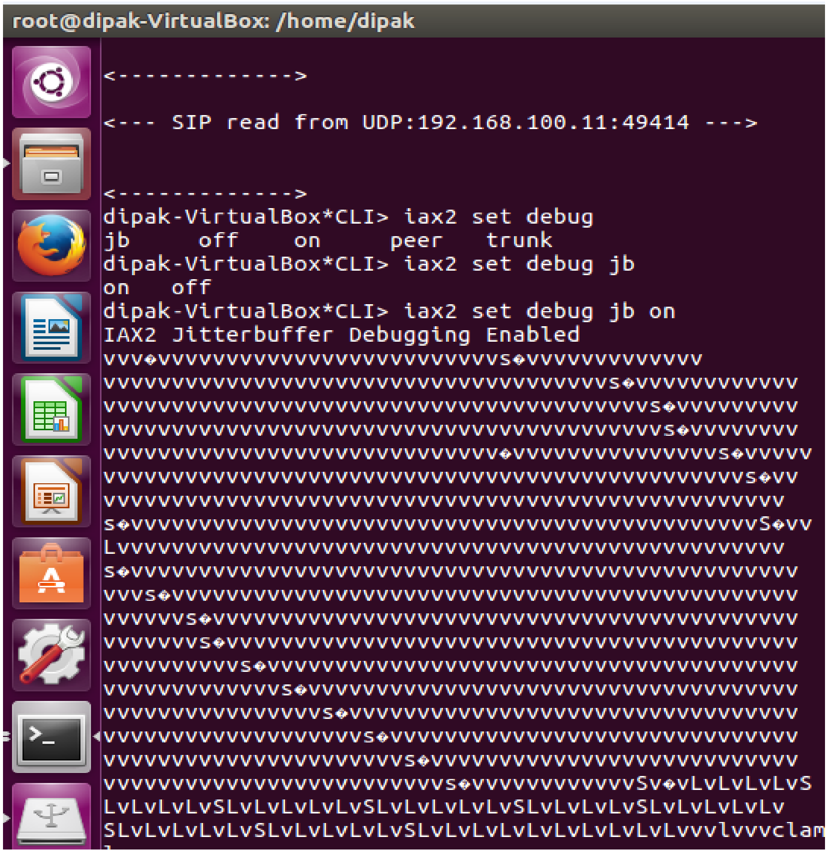

Figure 3.5-3: Real time jitter buffer status during voice call

Jitter Buffer Status output:

G: growing jitter buffer, interpolating last frame

o: not a voice frame, forwarding it to the decoder

v: early or late voice frame, forwarding it to the decoder

V: voice frame, forwarding it to the decoder

l: late voice frame, too late to play so it's dropped

L: lost frame, interpolating last frame

s: shrinking jitter buffer, frame lost

S: shrinking jitter buffer, no frame lost

The maximum jitter buffer size also plays and important role in the quality of a call if packets arrive too late, the packet will be discarded by the jitter buffer. If the burst of packets that arrives at the receiver is greater than the maximum jitter buffer size, again packet are discarded. The discarded packet results in loss of packets.

Results on voice signal when Packet Loss Concealment enabled and disabled is illustrated in Annex VI.

Receiver RTCP stream for no loss condition

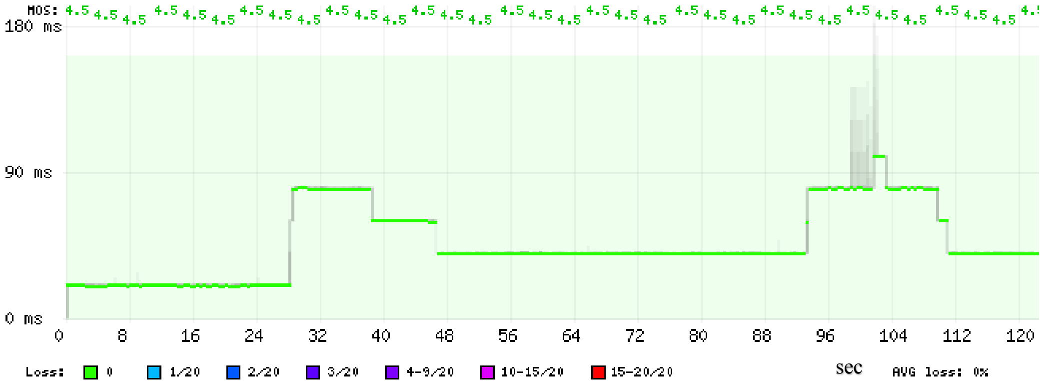

Figure 3.5-4: Receiver RTP stream for no packet loss condition

In Figure 3.6 4 RTCP stream of no packet loss condition is plotted. Delay of the voice signal is constant and Overall MOS of the voice communication is above 4.3. Thus quality of the voice communication is excellent in no loss condition.

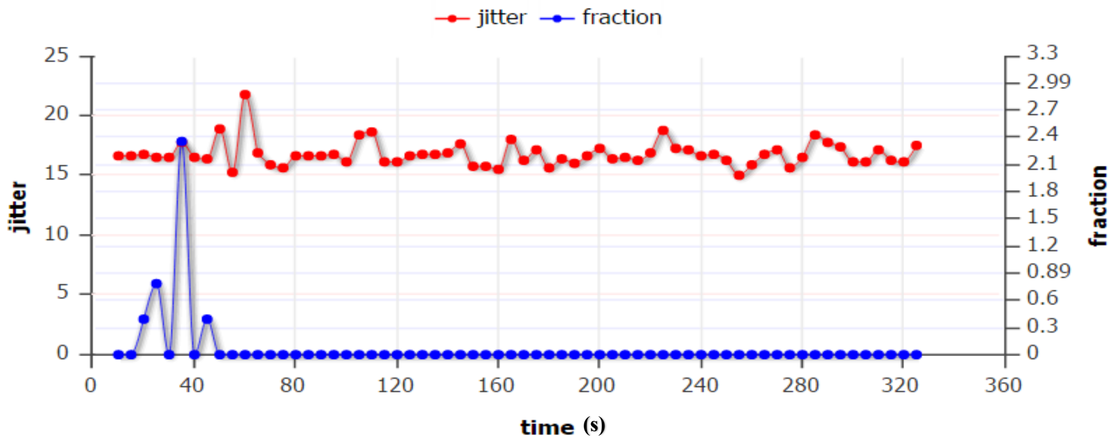

Figure 3.5-5: Receiver RTCP jitter for no packet loss condition

When PLC technique is not enabled overall jitter value is at the range of 15 ms to 22 ms which is an acceptable range in VoIP communication.

Packet Loss Concealment enabled for no loss condition

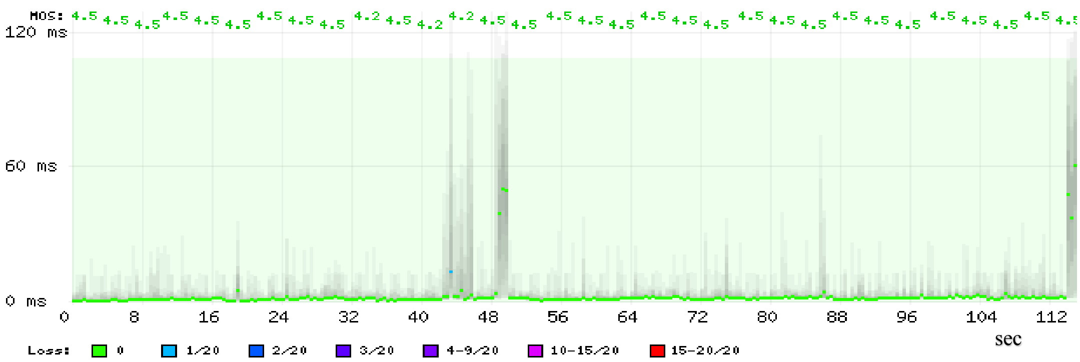

Figure 3.5-6: Caller RTP stream for no packet loss condition with jitter buffer enable

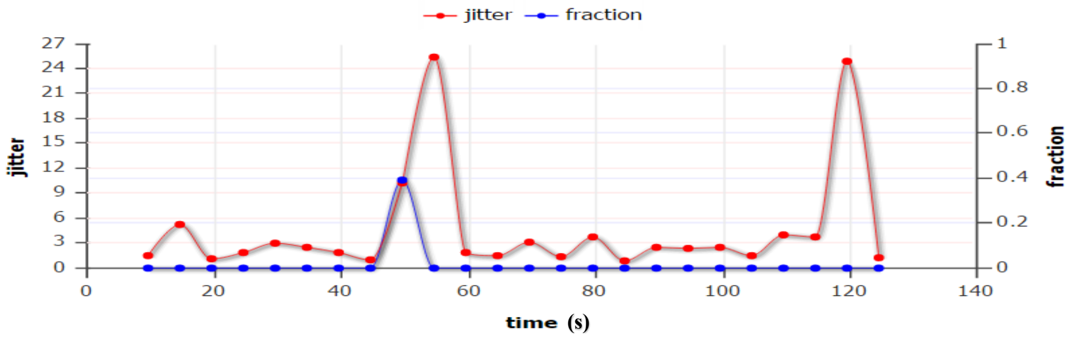

Figure 3.5-7: Caller RTCP jitter for no packet loss condition with jitter buffer enable

RTCP stream for the receiver when PLC enabled, MOS of the voice signal is above 4.2 represents excellent quality in voice communication. Also jitter value for the same voice signal is minimum when no loss is observed in the network. Thus PLC technique improves the Quality of Service in voice communication for no packet loss condition in network.

Caller RTP stream during 5% loss observed in network

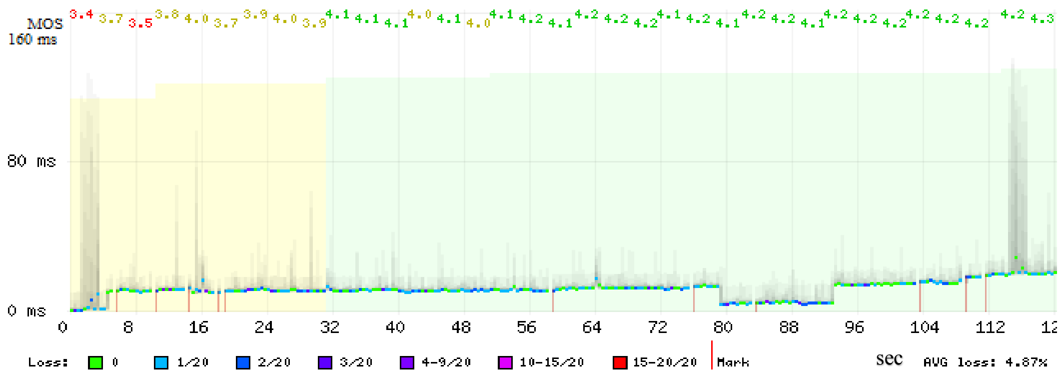

Figure 3.5-8: Caller RTP stream for 5% packet loss condition with jitter buffer disable

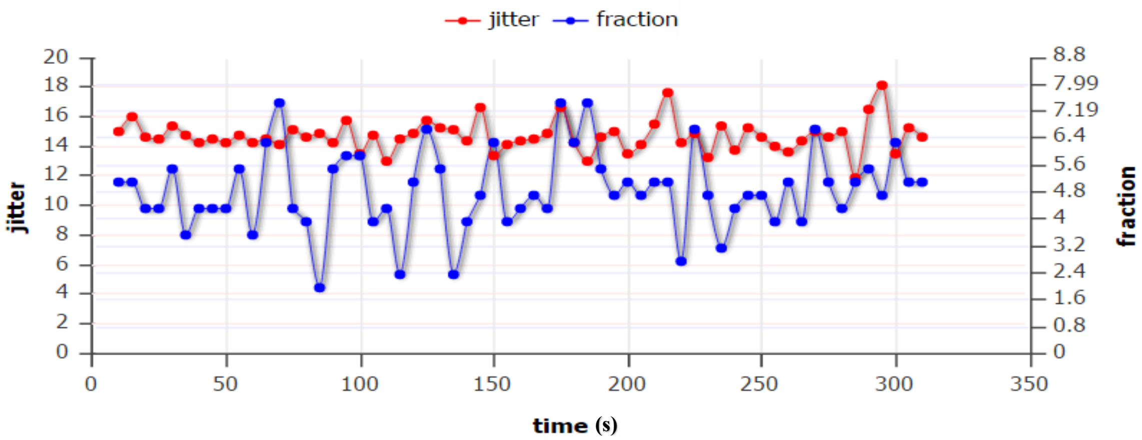

Figure 3.5-9: Caller RTCP jitter for 5% packet loss condition with jitter buffer disable

From RTCP stream of caller at 5% network packet loss, there is silence of voice call at some time interval. Silence during voice call affect user experience and communication. Also at initial MOS value decreased to 3.4. When PLC disable average jitter for the voice communication is 16 ms.

Condition when network have 5% packet loss and PLC enabled

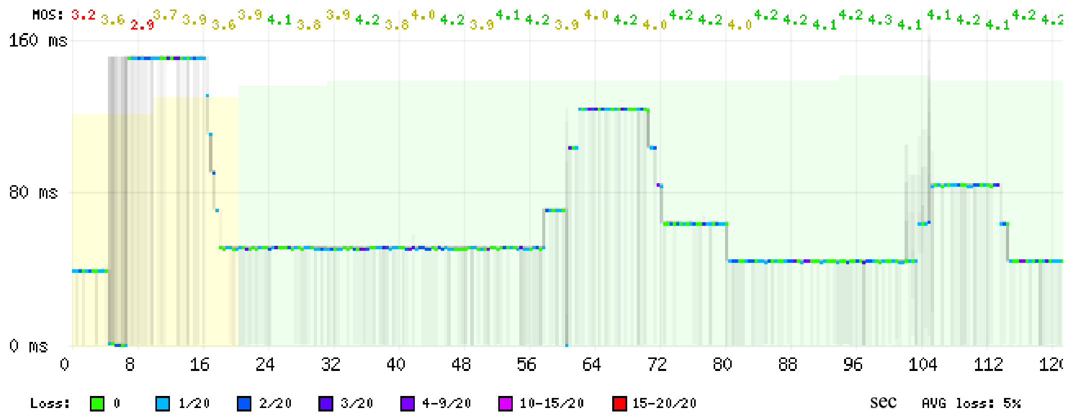

Figure 3.5-10: Receiver RTP stream for 5% packet loss condition with jitter buffer enable

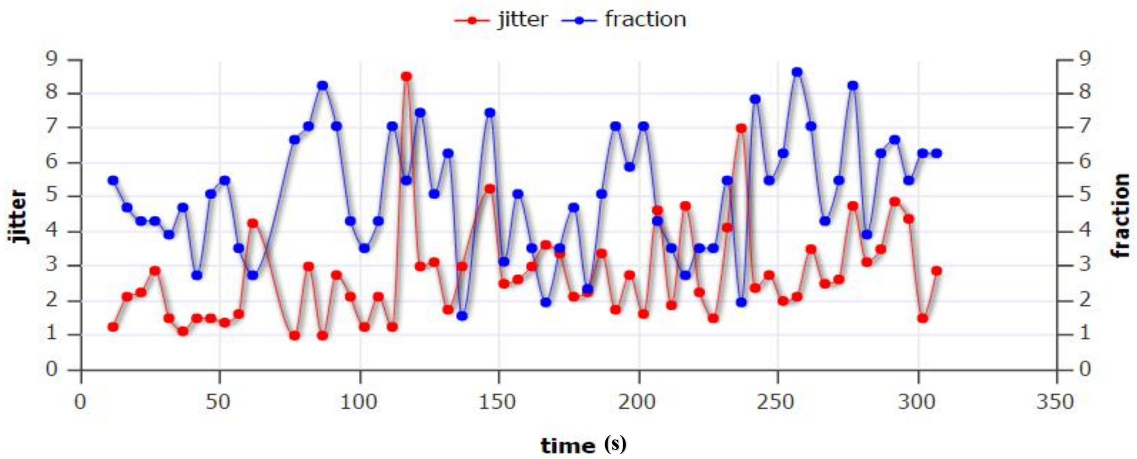

Figure 3.5-11: Receiver RTCP jitter for 5% packet loss condition with jitter buffer enable

From the analysis of RTCP stream and RTP jitter for network with packet loss 5%. Maximum jitter during PLC enabled is 8 ms which gives better voice quality then the average 12 ms jitter when PLC is disable. During VoIP communication analysis it is observed that in network for short instance of time link fluctuation occurs. Such fluctuation affects the Quality of Service and user experience. Thus implementing PLC technique Quality of voice service is improved in network with packet loss.

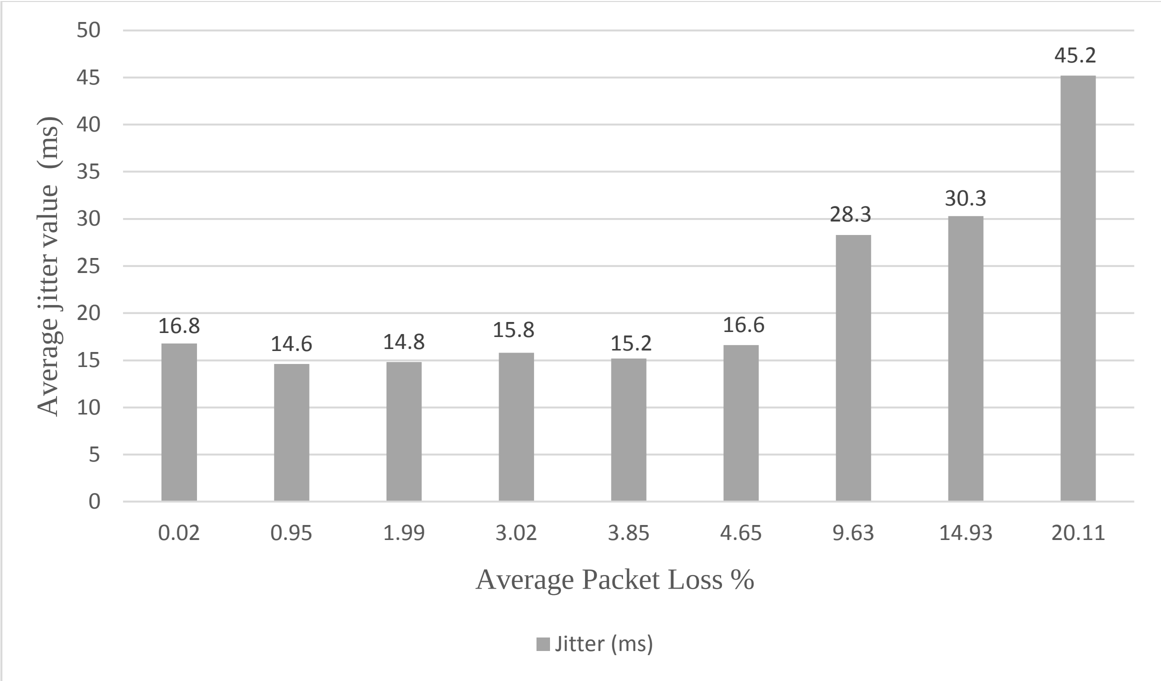

Figure 3.5-12: Average value of packet loss and jitter when Jitter Buffer Disable

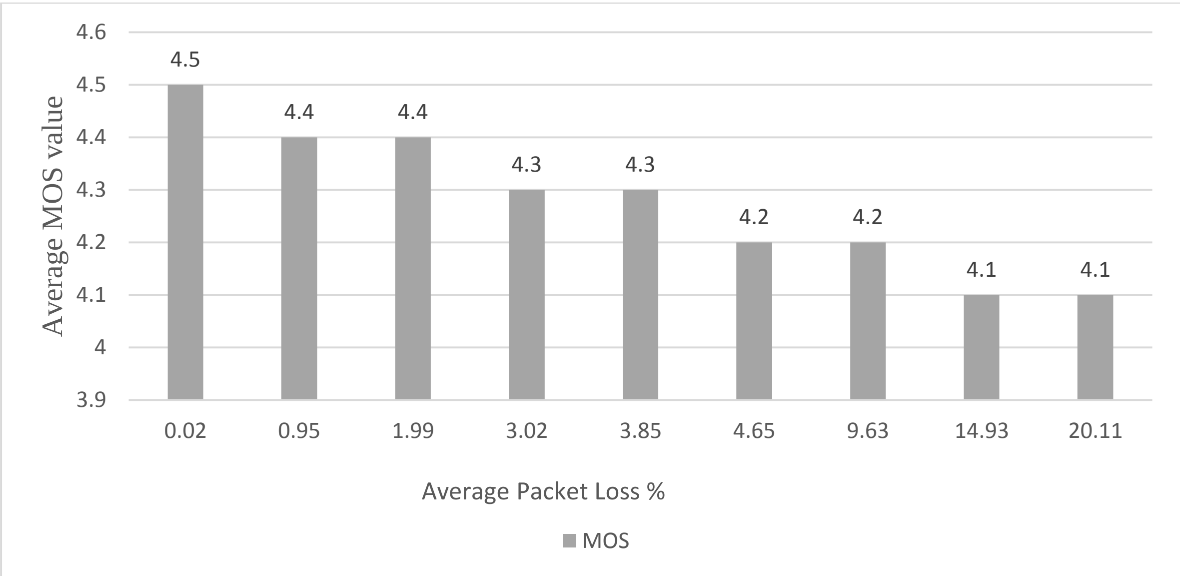

Figure 3.5-13: Average value of packet loss and MOS when Jitter Buffer Disable

In Figure 3.5-12 and Figure 3.5-13 when packet loss is greater than 15%, average jitter value is above 30ms. During RTCP observation it is found that jitter crossed 40 ms for some instance. Normal VoIP communication cannot be established for Packet loss greater than 15%.

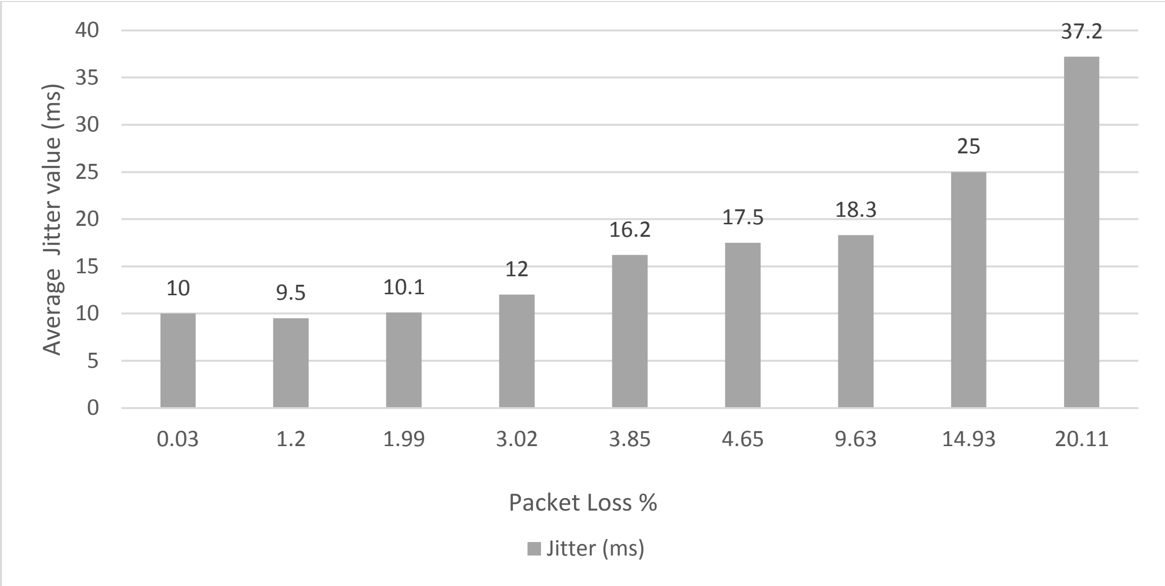

Figure 3.5-14: Average value of packet loss and jitter when Jitter Buffer Enable

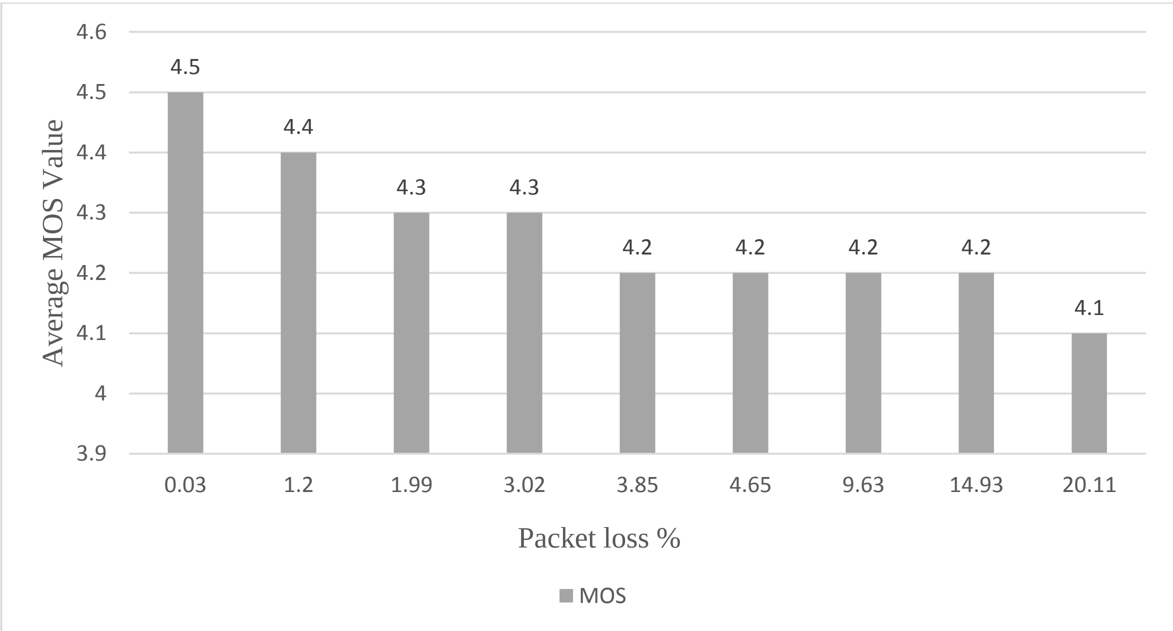

Figure 3.5-15: Average value of packet loss and MOS when Jitter Buffer Enable

In Figure 3.5-16 and Figure 3.5-17 when packet loss is greater than 15%, average jitter value is 25. Comparing Figure 20 and Figure 21, Average MOS value when Packet Loss Concealment Enable is 0.1 more than MOS value when Packet Loss Concealment Disable. Also for Jitter value when Packet Loss Concealment is enabled there is significant difference seen as shown in bar diagram.

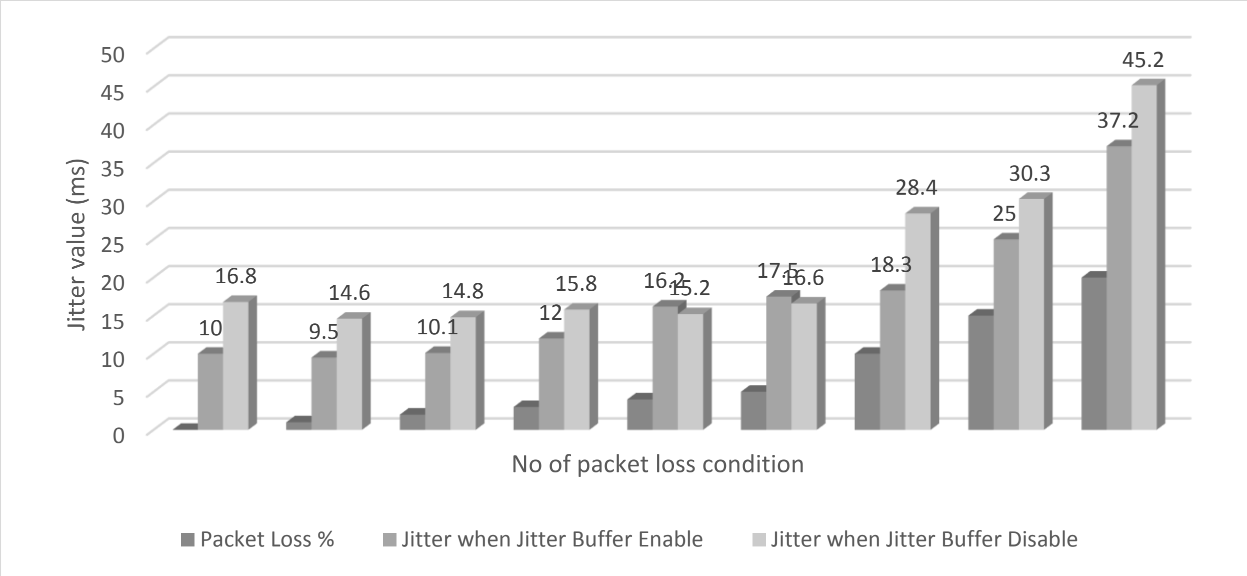

Figure 3.5-16: Relation of Jitter and Packet loss with Jitter Buffer

Results for voice communication for different network loss scenario for Packet Loss Concealment enabled and disabled is illustrated in Annex VII.

Overall analysis for objective Quality of Service (QoS) measurement in VoIP call with VoIPmonitor software is performed. It is found that average Mean Opinion Score (MOS) for different loss scenario have slight change from 4.4 to 4.1.Although average value of MOS is satisfactory when jitter value increases will affect quality of voice call communication. During Loss period MOS value is 3 or below, in which VoIP call quality is annoying or bad. Jitter value greater than 40 ms is not acceptable in VoIP communication, during loss period jitter increases to 40 ms or which will affect the VoIP call at that instance.

5. CONCLUSION

In this thesis Packet Loss Concealment technique is proposed to mitigate the effect of packet loss during voice communication in Voice over Internet Protocol. From the experiments and analysis it is observed that the quality of the voice signal is affected from various network parameters as packet loss, delay, jitter etc. Several objective tests were performed in order to test the proposed PLC technique. Initially different packet loss scenario varying from 1% to 20% were tested without implementing PLC technique. After analysis on results from different scenarios it is concluded that the loss effect on voice communication increases with the value of packet loss present in the network. Also delay, jitter value increases with packet loss and voice communication above 10% packet loss condition cannot be established. To mitigate the effect of packet loss in VoIP PLC technique is implemented for same case scenario as before.

From the results, it was observed that when the value of packet loss is low i.e. 0-1%, network quality is degraded for a short instance of time during voice call communication. After implementing PLC technique it was observed certain percent (1-2%) of packet loss in network does not affect overall quality in voice communication. In network with packet loss of 5%, jitter and MOS value is maintained under threshold which allows for normal voice communication with minor lag. For extreme packet loss condition where loss is in between 10 % to 15%, user is able to communicate with delay of 300 ms to 600 ms. Severe condition where loss is above 15 % voice call conversation can be established without PLC was not possible . Thus implementing PLC technique improves the Quality of Service of voice even in a network with packet loss. Limitation and future enhancement of this thesis is discussed in the next chapter.

6. LIMITATION AND FUTURE ENHANCEMENT

Although PLC approach minimize the effect of packet loss and recover voice quality, it does not fully mitigate the issue. Since it is a receiver based approach for Packet Loss Concealment, this method only deals with what is available in receiver jitter buffer. In PLC approach, jitter buffer size is limited to certain extent without causing delay in voice communication, hence degrading the quality of voice. Also two SIP channels exchange RTP directly, then VoIP server will never be able to process the audio frame and translation will not allow for a Packet Loss Concealment.

In this thesis a small range of 66 Hz to 200 Hz pitch window is considered due to computational complexity. Future enhancement can be done for low computational complexity approach to cover the wide range of pitch i.e. human hearing range 20Hz to 20 KHz. VoIP Quality of Service (QoS) is directly related to many network parameters as packet loss, delay, jitter which make VoIP communication and its control mechanism complex. PLC recovers voice data with high packet loss, it also introduces delay (600 ms -900 ms) is high enough to degrade Quality of Service and user experience. The interpolation of lost frame in acceptable delay range using buffer memory can be improved to get more precise synthetic voice frame. In VoIP communication further research can be done to improve network complexity of the Packet Loss Concealment technique.

References

[1] D. M. P. K. Maheswari, An Improved packet loss recovery in VoIP using combined source and Receiver based Technique, coinbatore, India: SNR Sons College Dept. of Computer Applications.

[2] N. A. G. Qutaiba Ali, "Reviewing Speaker Recognition in Strengthening VOIP Caller Authentication," *Computer Engineering Department, University of Mosul, Iraq,* vol. 1, p. 7, 2010.

[3] N. Aoki, "A Packet Loss Concealment Technique for VoIP using Steganography," Graduate School of Engineering, Hokkaido University , Sapporo, 060-8628 Japan , 2003.

[4] A. Ito, "Packet Loss Concealment of Voice-over IP Packet using Redundant Parameter Transmission Under Severe Loss Conditions," *Journal of Information Hiding and Multimedia Signal Processing,* vol. 5, p. 10, 2014.

[5] E. B. E. M. Marshini Chetty, "VoIP deregulation in South Africa: Implications for underserviced areas," College of Computing, Georgia Institute of Technology, Atlanta, Georgia, 2006.

[6] S. C. D. X. Wei Yu, "P2P/ Grid-based Overlay Architecture to Support VoIP Services in Large Scale IP Networks," Dept. of Computer and Information Science,The Ohio-State University, U.S.A, Ohio-State, U.S.A.

[7] J. Hitchcock, "Decorating Asterisk: Experiments in Service Creation for a Multi-Protocol Telephony Environment Using Open Source Tools," Computer Science Department,Rhodes University, Eastern Cape Province, South Africa, 2006.

[8] N. Younas, "Designing and optimization of VOIP PBX infrastructure," Department of computer science and technology University of Bedfordshire, United Kingdom, 2013.

[9] F. S. M. K. Abdullah-Al-Mamun, "A High Resolution Pitch Detection Algorithm Based on AMDF and ACF," King Saud University, Riyadh, Saudi Arabia , Riyadh,, 2009.

[10] Tommi Koistinen, "Protocol overview: RTP and RTCP," Nokia Telecommunications, 1999.

[11] H. Ofir, "Packet Loss Concealment for Audio Streaming," Israel Institute of Technology, Isrel, 2006.

[12] Nepal Telecom Authority, "Nepal Telecom Authority," Nepal telecom Authority, [Online]. Available: http://www.nta.gov.np/en/. [Accessed 16 1 2017].

[13] M. T. Floriano De Rango, "Overview on VoIP," *Subjective and Objective Measurement,* vol. 6, p. 14, 2006.

[14] T. S. S. O. ITU, "SERIES G: TRANSMISSION SYSTEMS AND MEDIA,DIGITAL SYSTEMS AND NETWORKS," I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n, 2011.

[15] R. S. Gokhale, "Packet loss concealment in voice over the Internet," University of South Florida, South Florida, 2003.

7. APPENDIX

A. Annex I

Linux System Install with required Software Platform

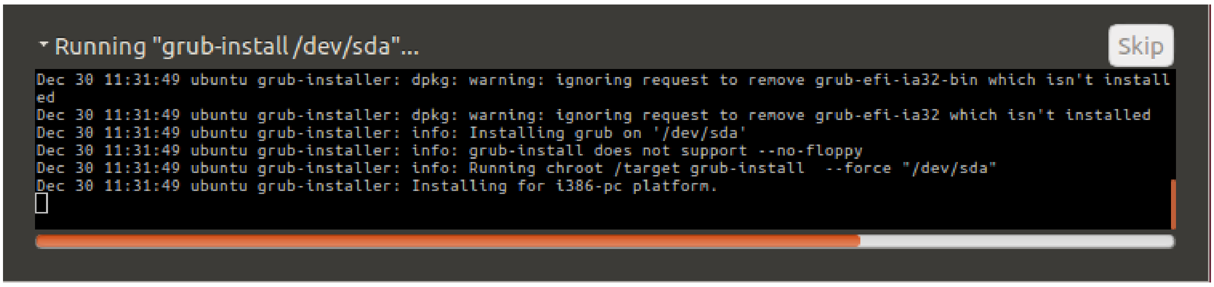

Figure A-1: Install Required Linux Platform

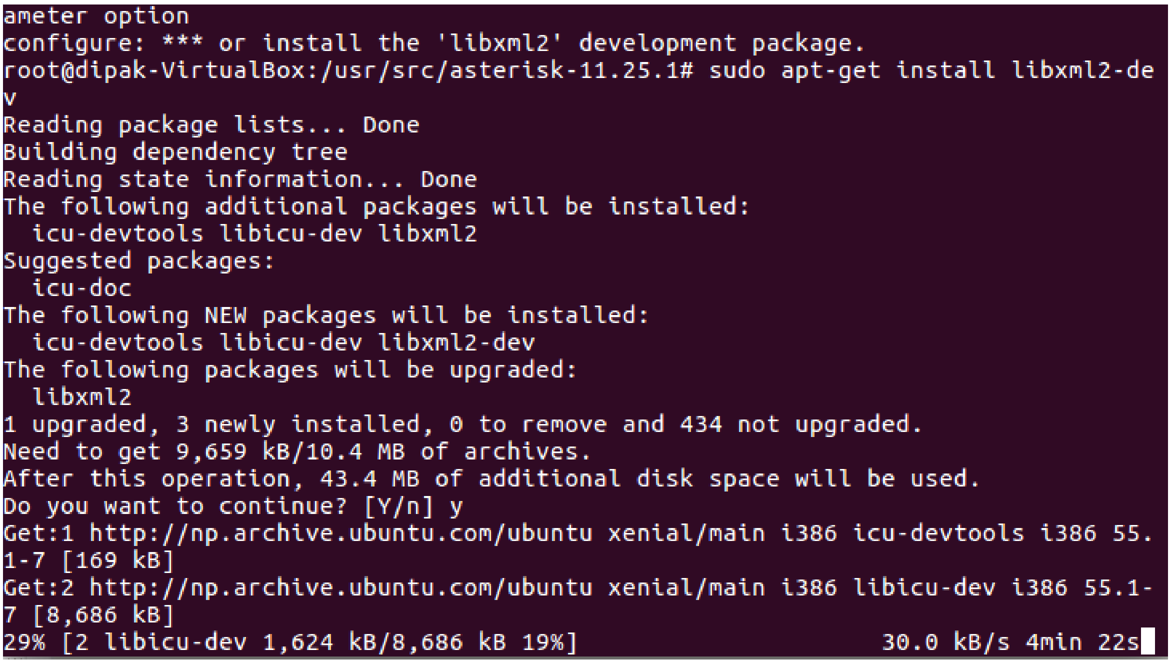

Figure A-2: Preparing System for VOIP server Configuration

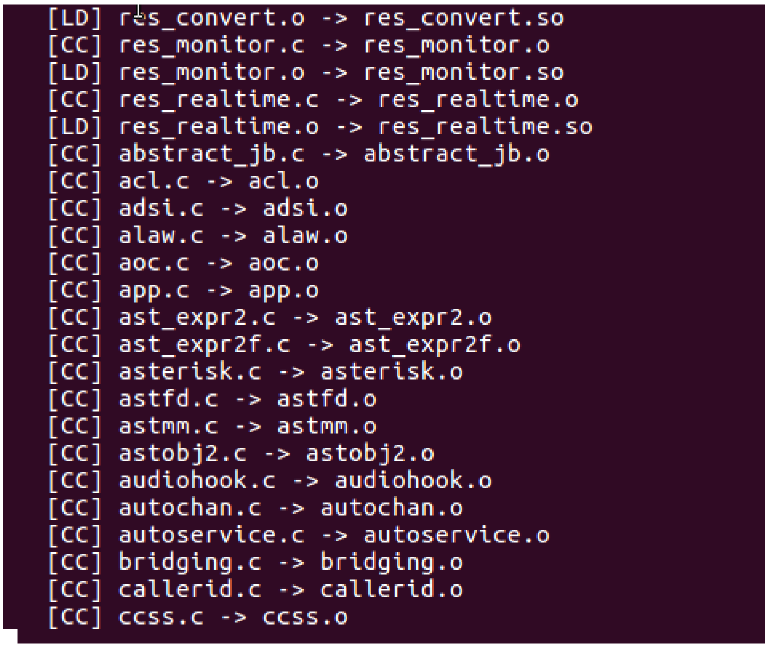

Figure A-3: Compile VOIP server Code after any changes

In Figure A-1 to Figure A-3, Linux system installation and VoIP source code compilation is shown. After any changes in source code, need to recompile to take change in to effect.

B. Annex II

Asterisk GUI Graphical Interface login from browser for easy Asterisk configuration

Figure B-1: Asterisk GUI Graphical Interface

CSipSimple VoIP SIP Client login configuration

Figure B-2: CSipSimple VoIP SIP Client

C. Annex III

System working model and flowchart as listed:

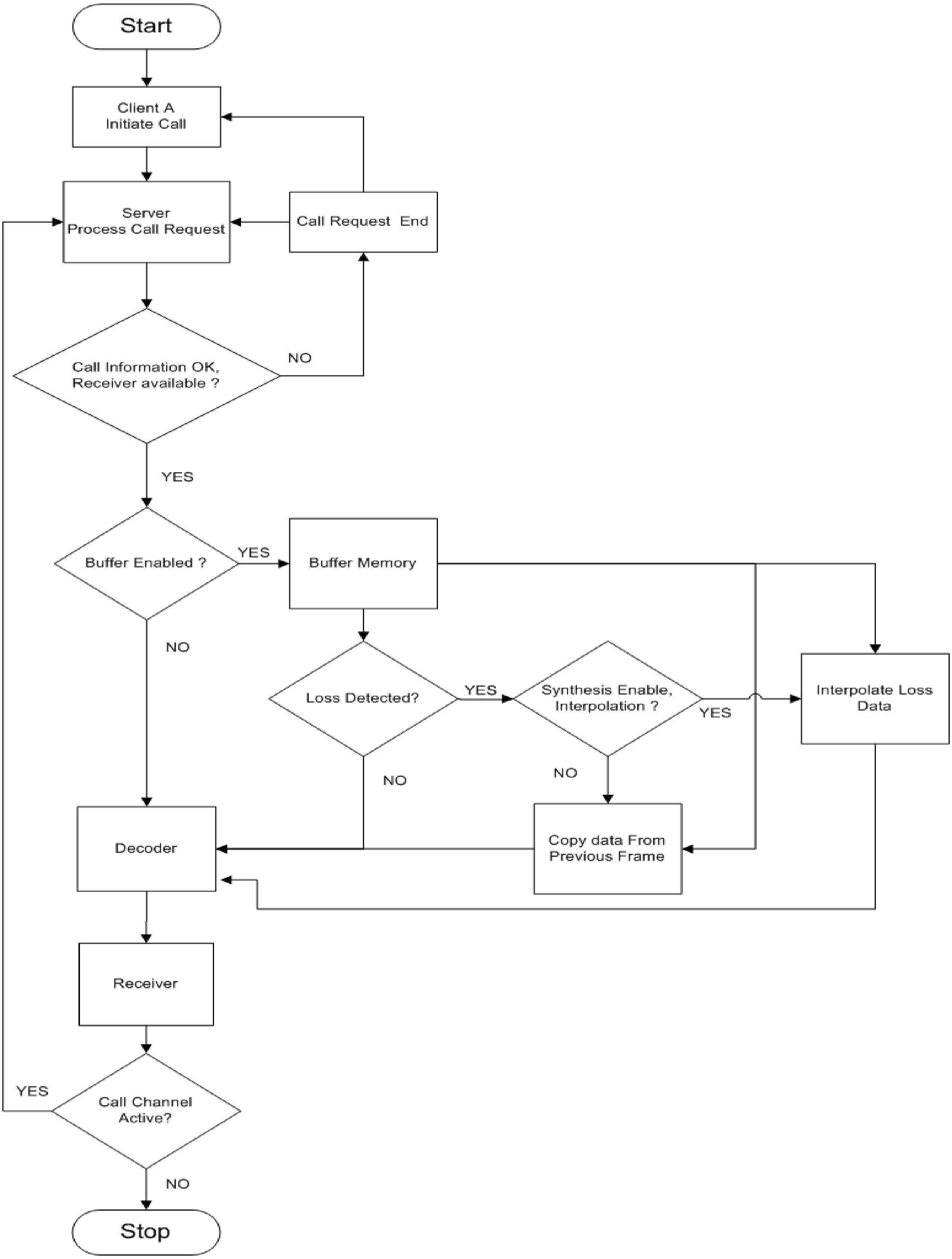

Figure C-1: System Working Flow Chart Model

D. Annex IV

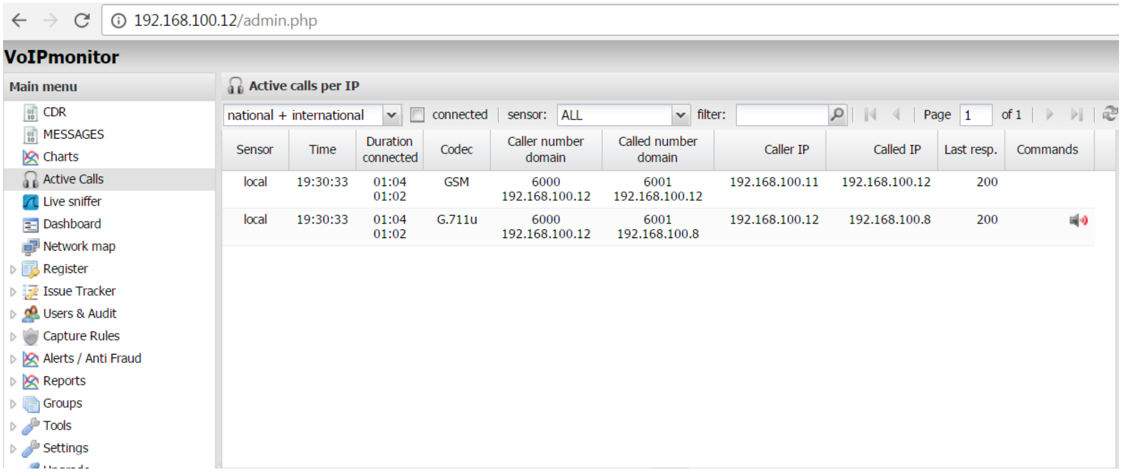

Figure D-1: VoIP Software Call statistics

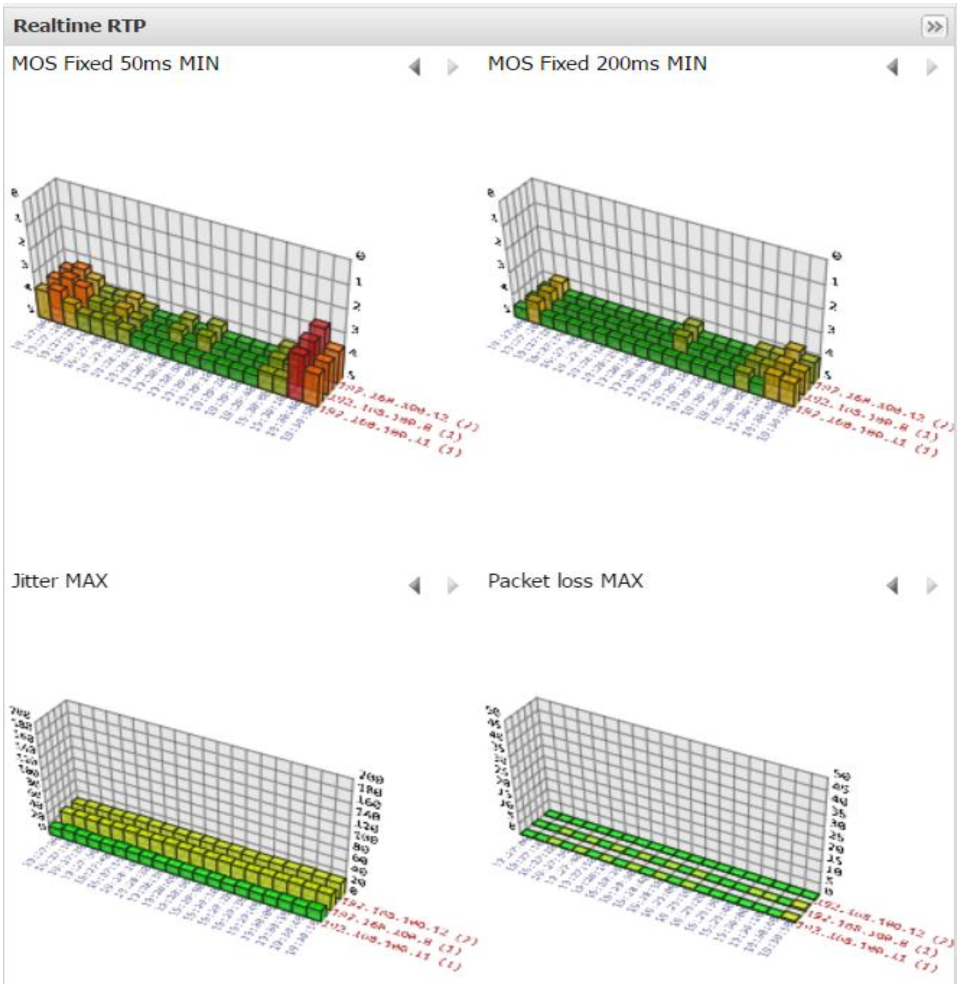

Figure D-2: Real Time RTP and MOS Monitoring

E. Annex V

Introducing fixed packet loss scenario in network

To drop a random 10% output packet from server towards client we use command as listed:

iptables –A OUTPUT –m statistic –mode random –probability 0.1 –j DROP

To drop a random 10% input packet towards server from client we use command as listed:

iptables –A INPUT –m statistic –mode random –probability 0.1 –j DROP

F. Annex VI

Audio record at VoIP monitoring software from caller and receiver when Packet Loss Concealment Technique is enable. In Figure F-1 we clearly see the lost voice signal when PLC is not enabled, when PLC is implemented lost voice signal is interpolated with synthesis voice and effect of packet loss is minimized.

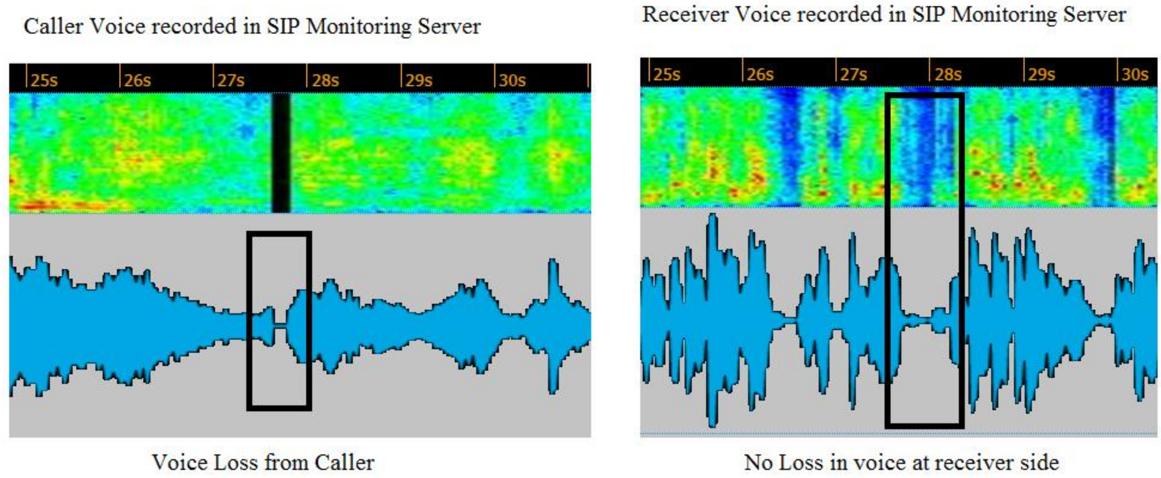

Figure F-1: Audio analysis for loss and synthesis signal

G. Annex VII

Condition I: No loss inserted in network with jitter buffer disable:

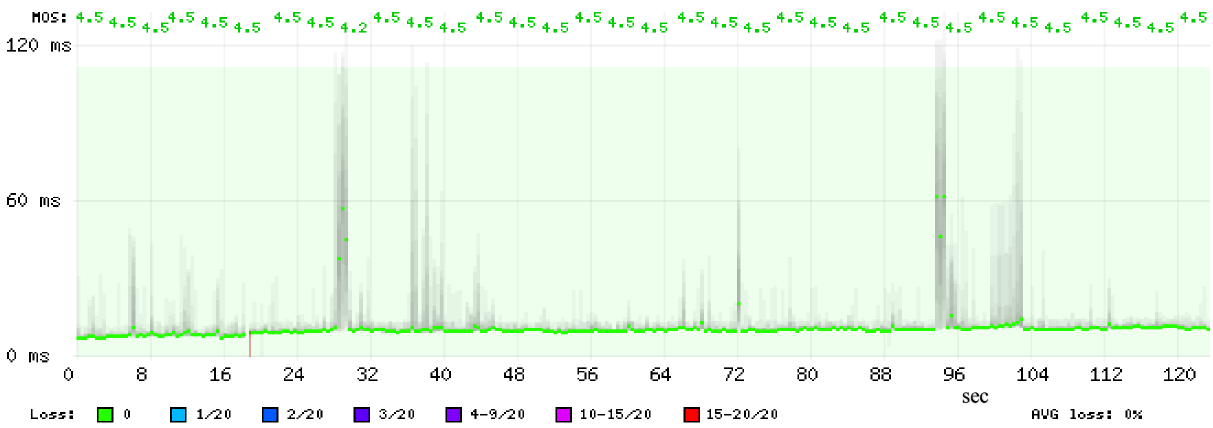

Figure G-1: Caller RTP stream for no packet loss condition

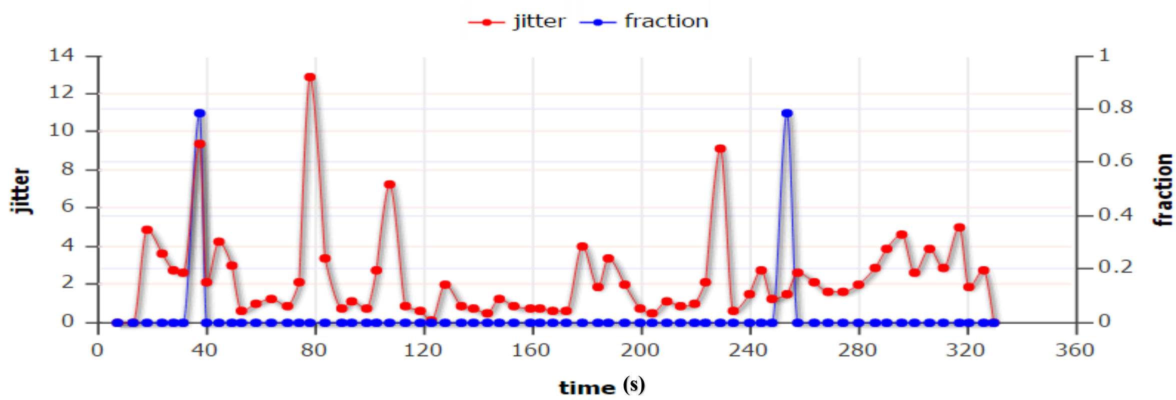

Figure G-2: Caller RTCP jitter for no packet loss condition

Condition II: No loss inserted in network with jitter buffer enable:

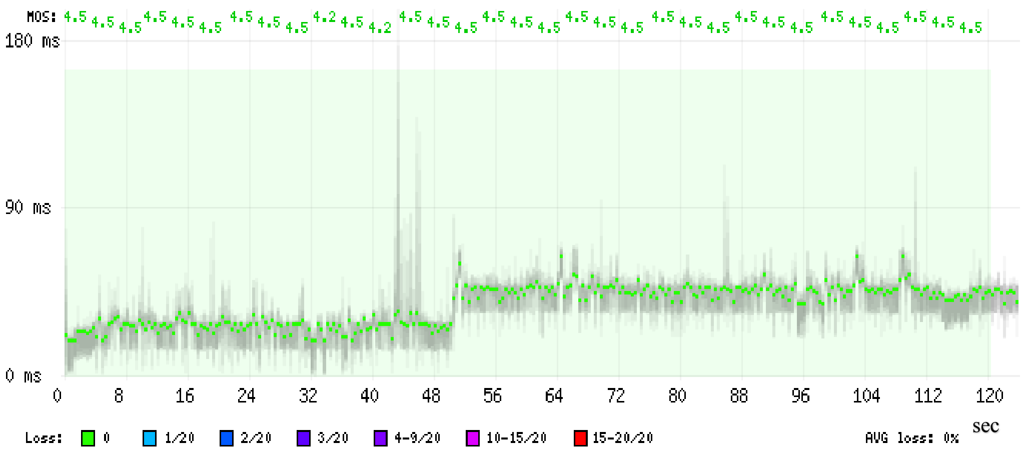

Figure G-3: Caller RTP stream for no packet loss condition with jitter buffer enable

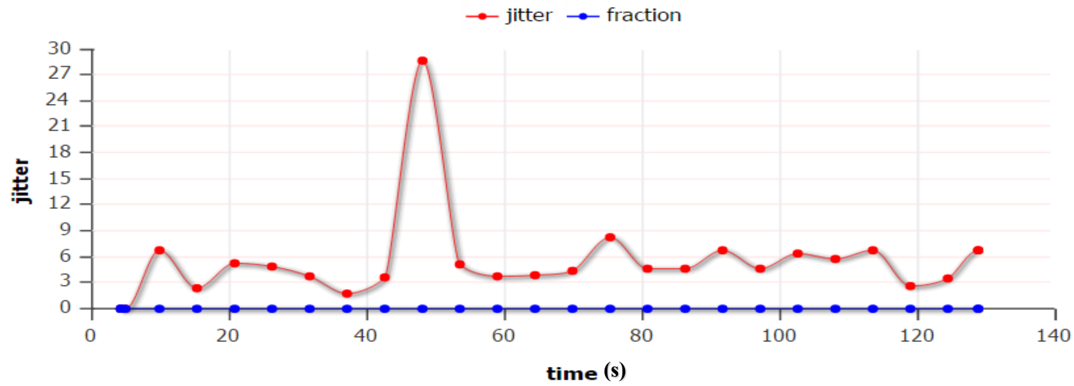

Figure G-4: Caller RTCP jitter for no packet loss condition with jitter buffer enable

Condition III: 1% Packet loss inserted in network with jitter buffer disable:

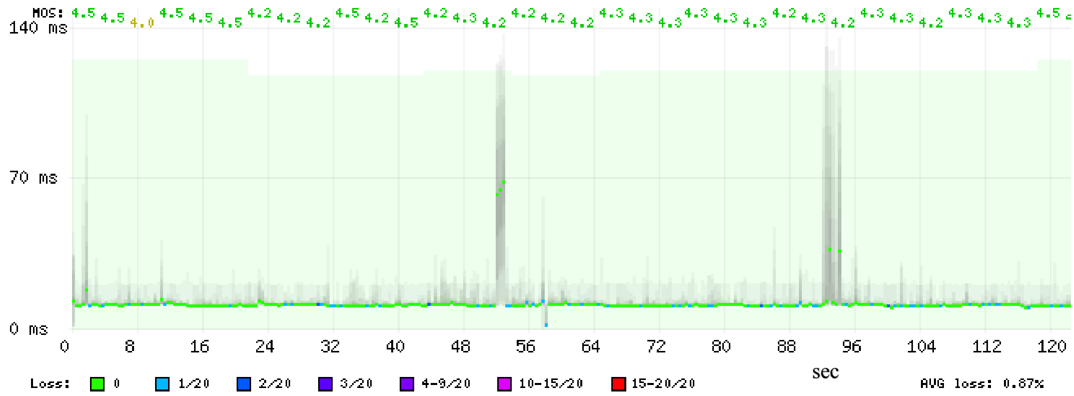

Figure G-5: Caller RTP stream for 1% packet loss condition with jitter buffer disable

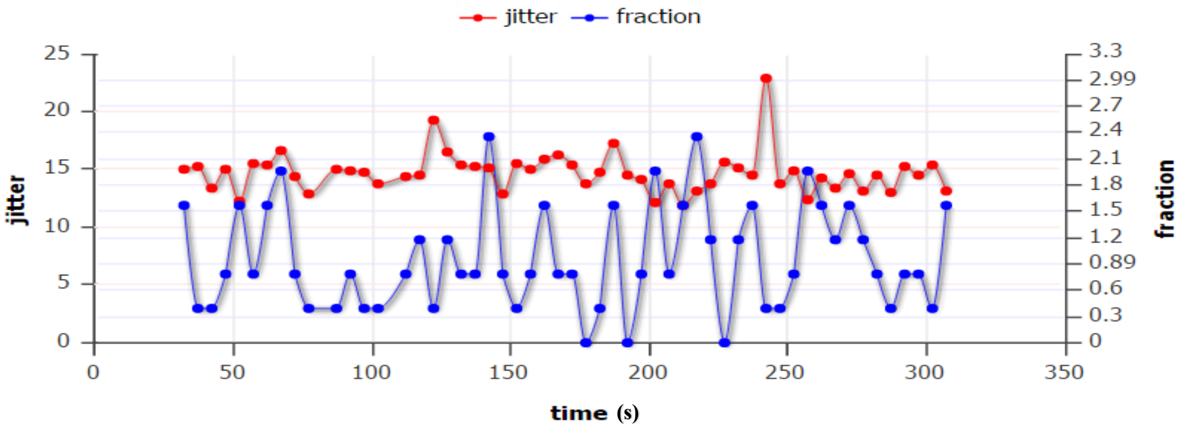

Figure G-6: Caller RTCP jitter for 1% packet loss condition with jitter buffer disable

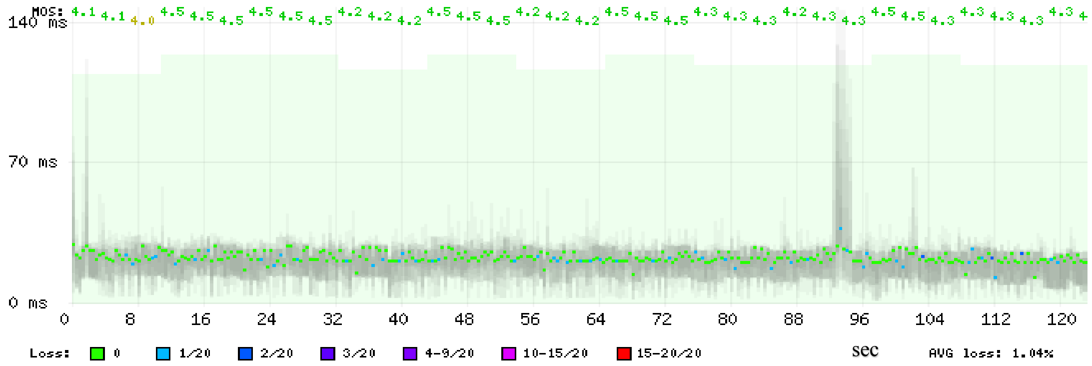

Figure G-7: Receiver RTP stream for 1% packet loss condition with jitter buffer disable

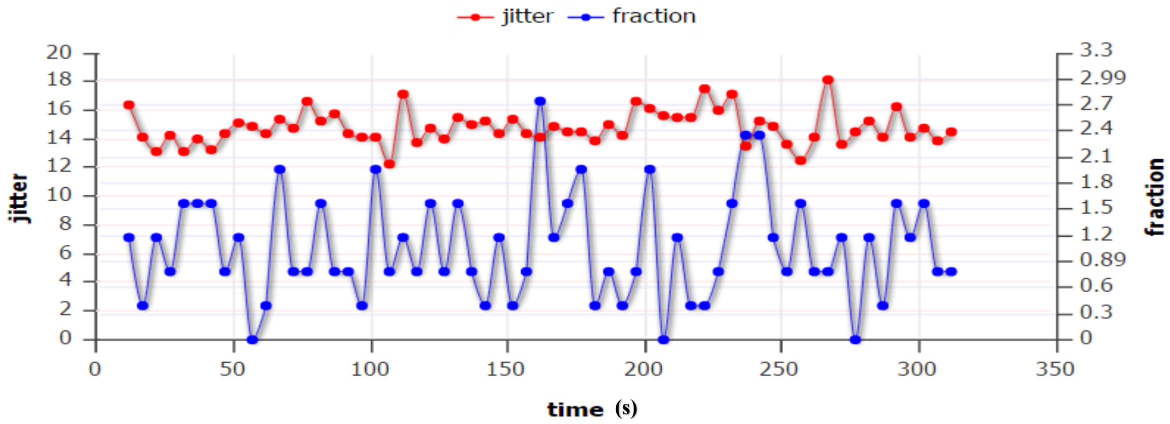

Figure G-8: Receiver RTCP jitter for 1% packet loss condition with jitter buffer disable

From Condition III, 1% loss condition MOS above 4 and jitter value is between 10 ms to 20 ms. This range of jitter is acceptable during VoIP communication.

Condition IV: 1% Packet loss inserted in network with jitter buffer enable:

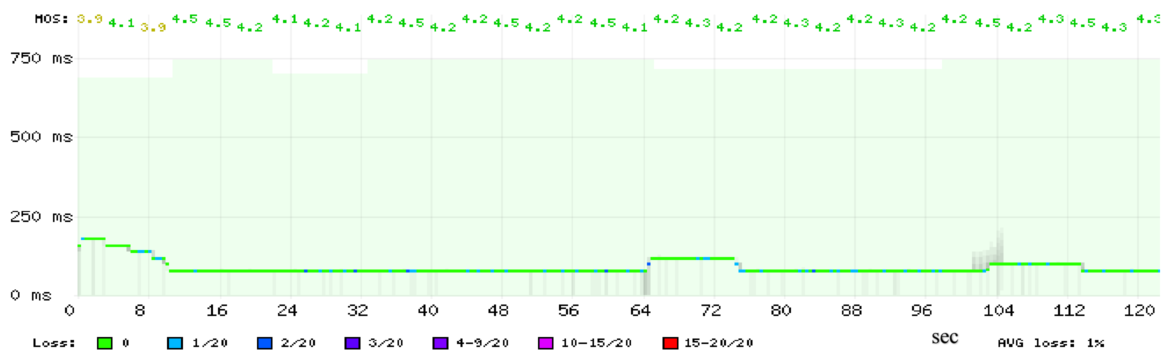

Figure G-9: Caller RTP stream for 1% packet loss condition with jitter buffer enable

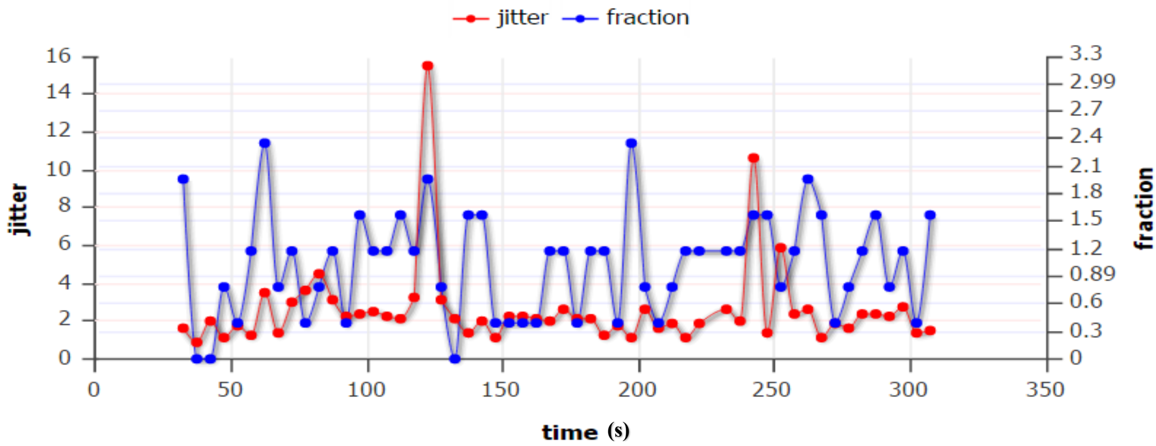

Figure G-10: Caller RTCP jitter for 1% packet loss condition with jitter buffer enable

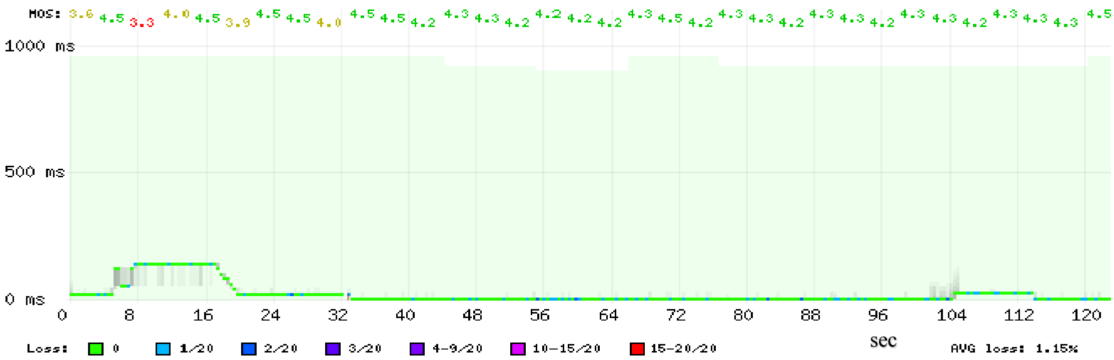

Figure G-11: Receiver RTP stream for 1% packet loss condition with jitter buffer enable

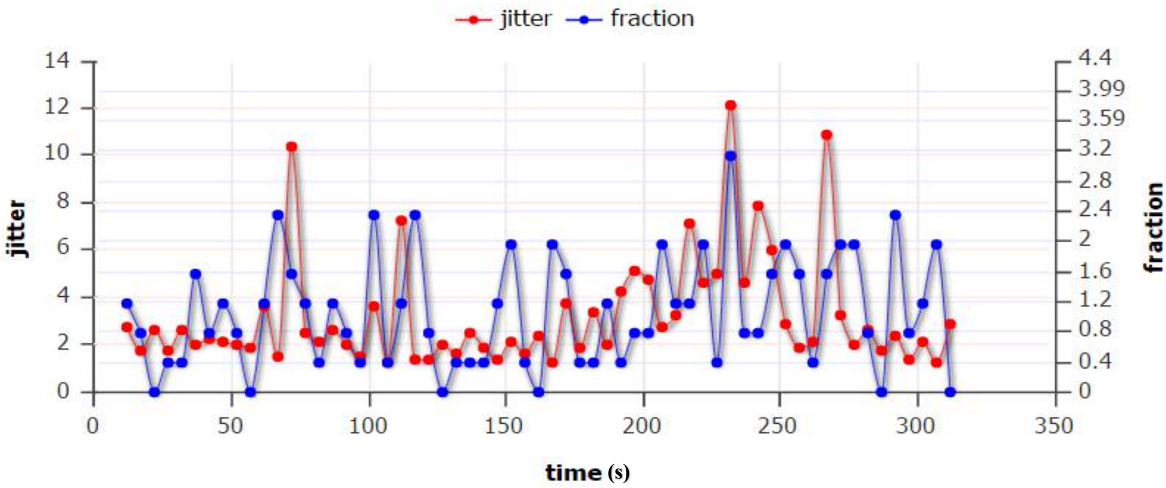

Figure G-12: Receiver RTCP jitter for 1% packet loss condition with jitter buffer enable

Condition V: 2% Packet loss inserted in network with jitter buffer disable:

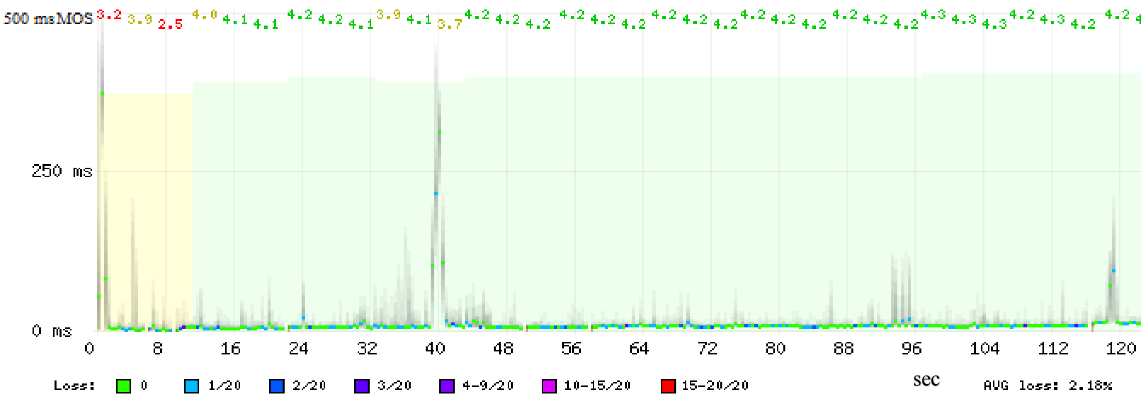

Figure G-13: Caller RTP stream for 2% packet loss condition with jitter buffer disable

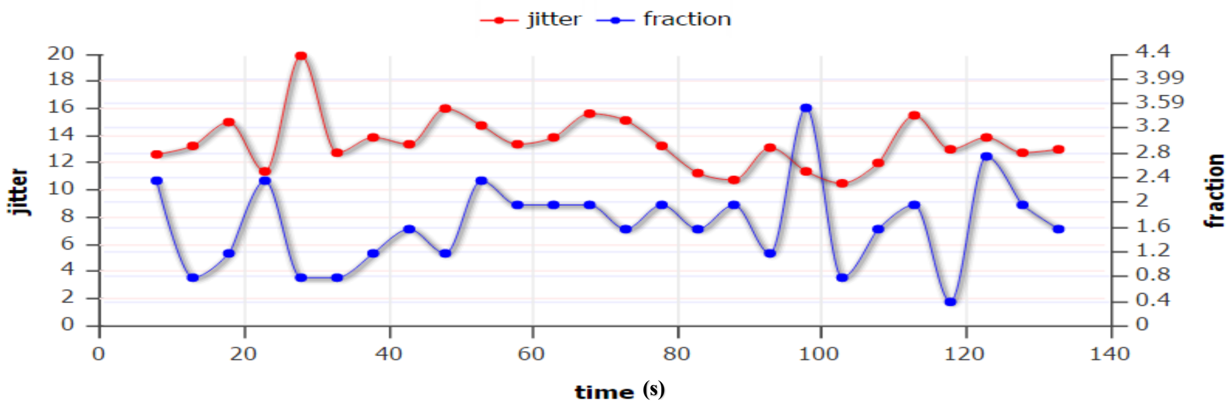

Figure G-14: Caller RTCP jitter for 2% packet loss condition with jitter buffer disable

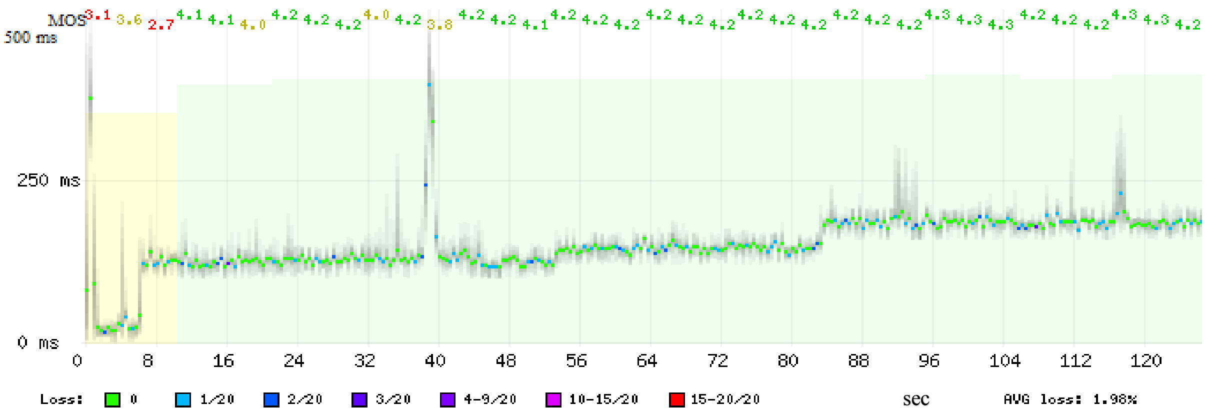

Figure G-15: Receiver RTP stream for 2% packet loss condition with jitter buffer disable

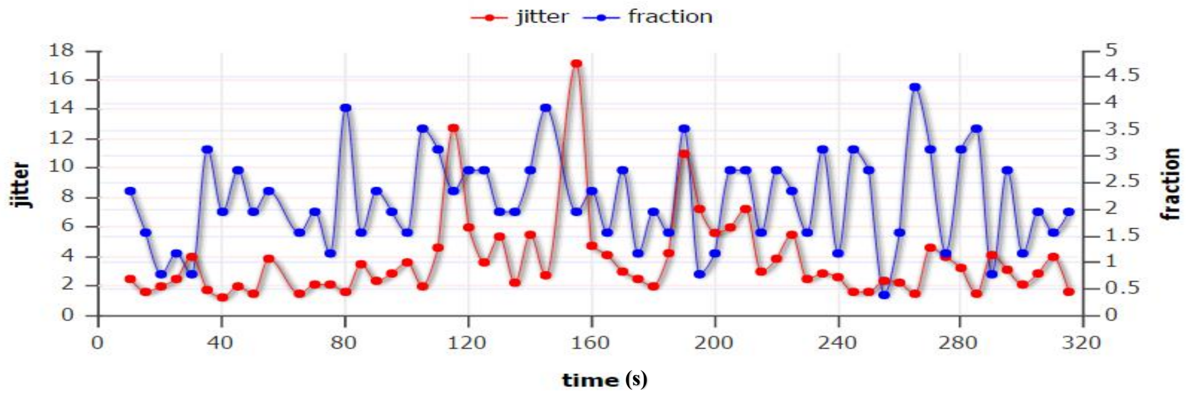

Figure G-16: Receiver RTCP jitter for 2% packet loss condition with jitter buffer disable

Condition VI: 2% Packet loss inserted in network with jitter buffer enable:

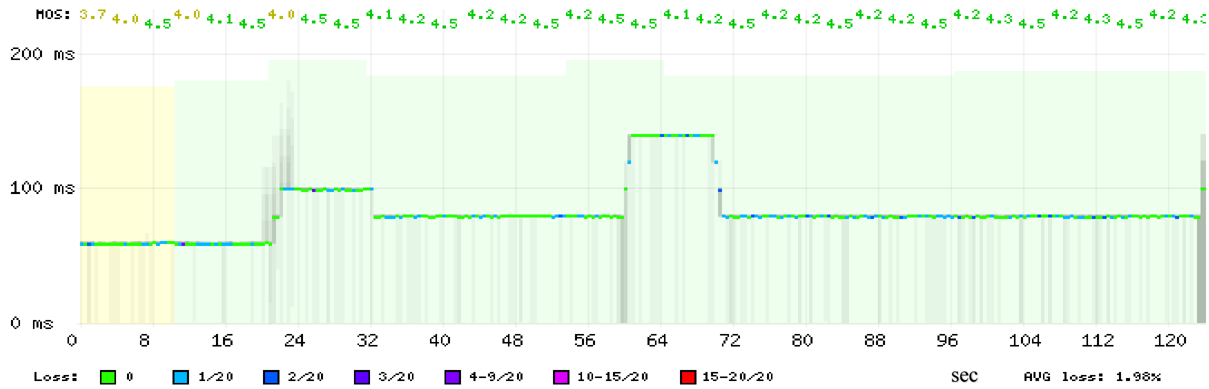

Figure G-17: Caller RTP stream for 2% packet loss condition with jitter buffer enable

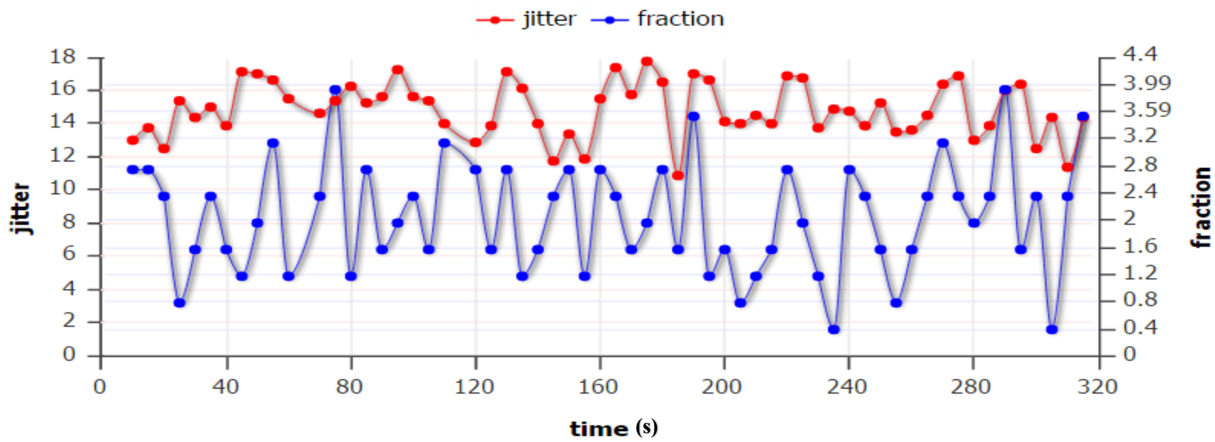

Figure G-18: Caller RTCP jitter for 2% packet loss condition with jitter buffer enable

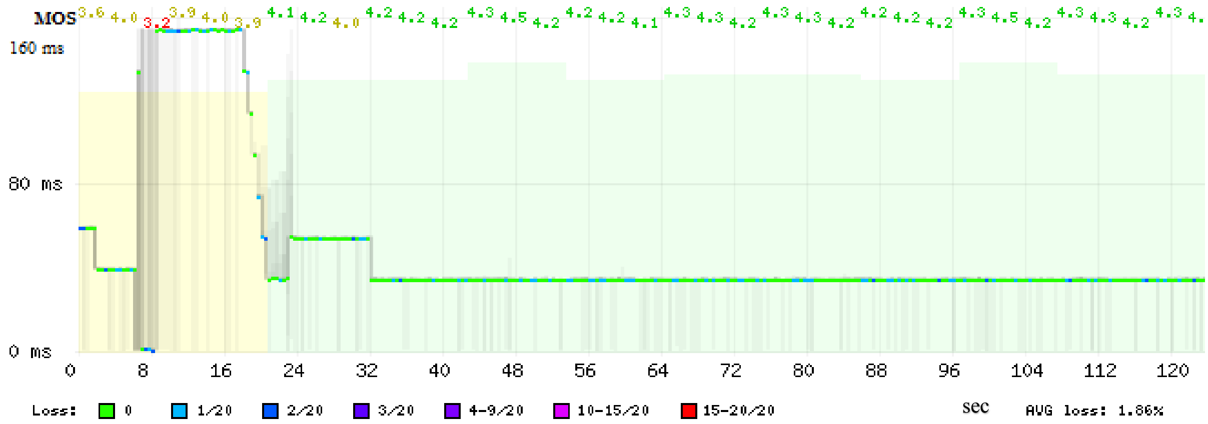

Figure G-19: Receiver RTP stream for 2% packet loss condition with jitter buffer enable

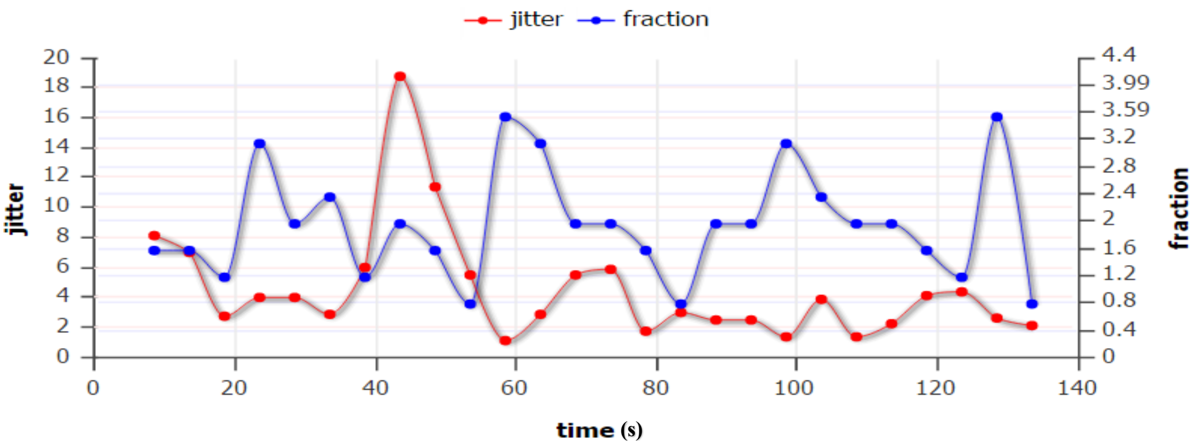

Figure G-20: Receiver RTCP jitter for 2% packet loss condition with jitter buffer enable

Condition VII: 3% Packet loss inserted in network with jitter buffer disable:

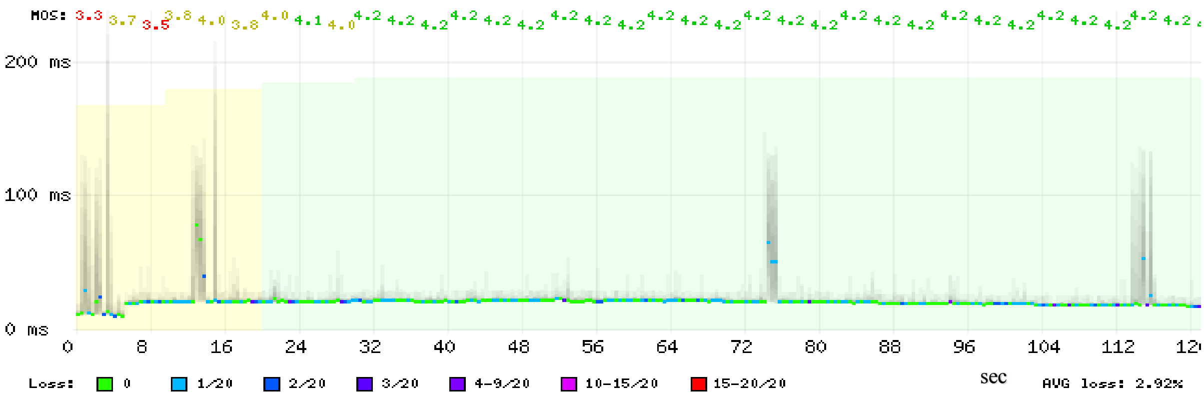

Figure G-21: Caller RTP stream for 3% packet loss condition with jitter buffer disable

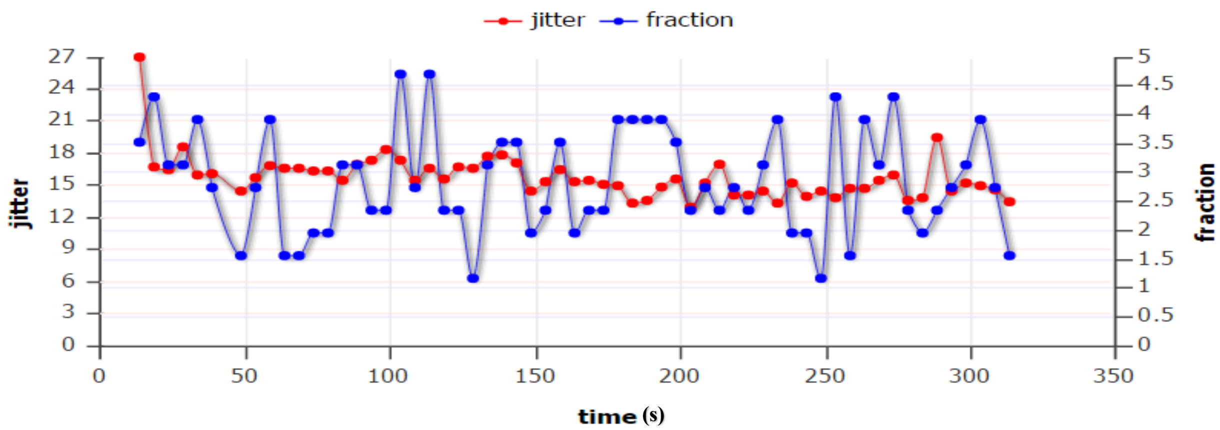

Figure G-22: Caller RTCP jitter for 3% packet loss condition with jitter buffer disable

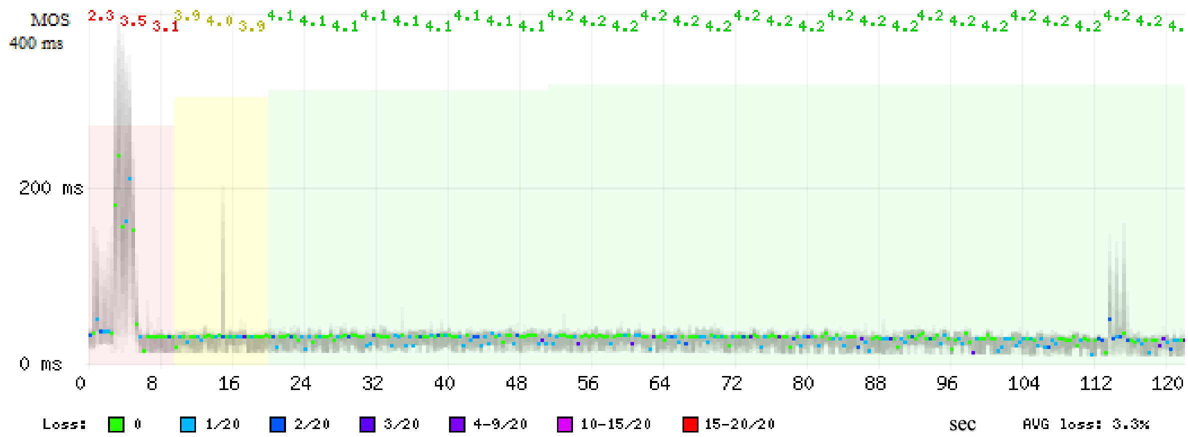

Figure G-23: Receiver RTP stream for 3% packet loss condition with jitter buffer disable

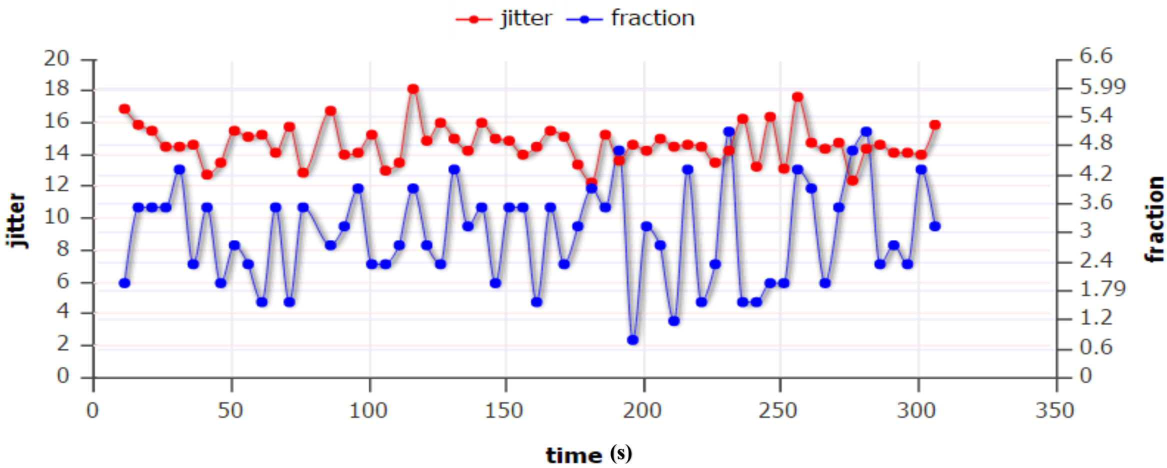

Figure G-24: Receiver RTCP jitter for 3% packet loss condition with jitter buffer disable

Condition VII: 3% Packet loss inserted in network with jitter buffer enable:

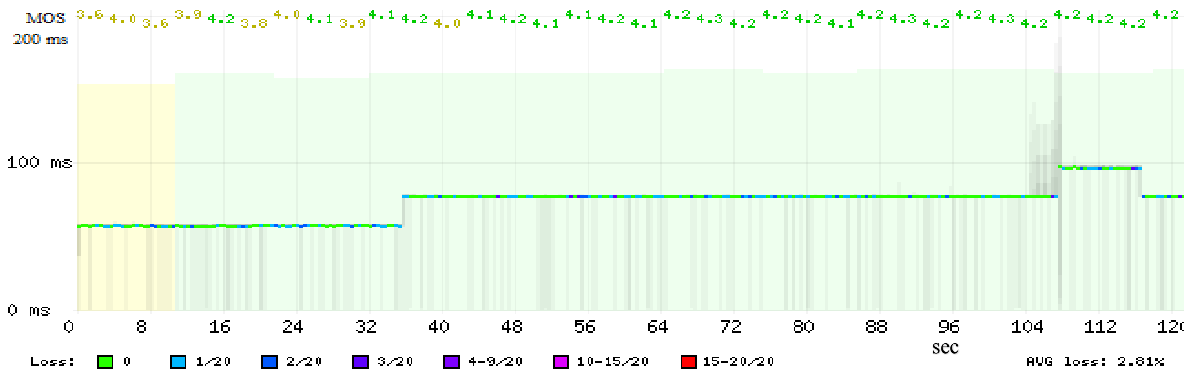

Figure G-25: Caller RTP stream for 3% packet loss condition with jitter buffer enable

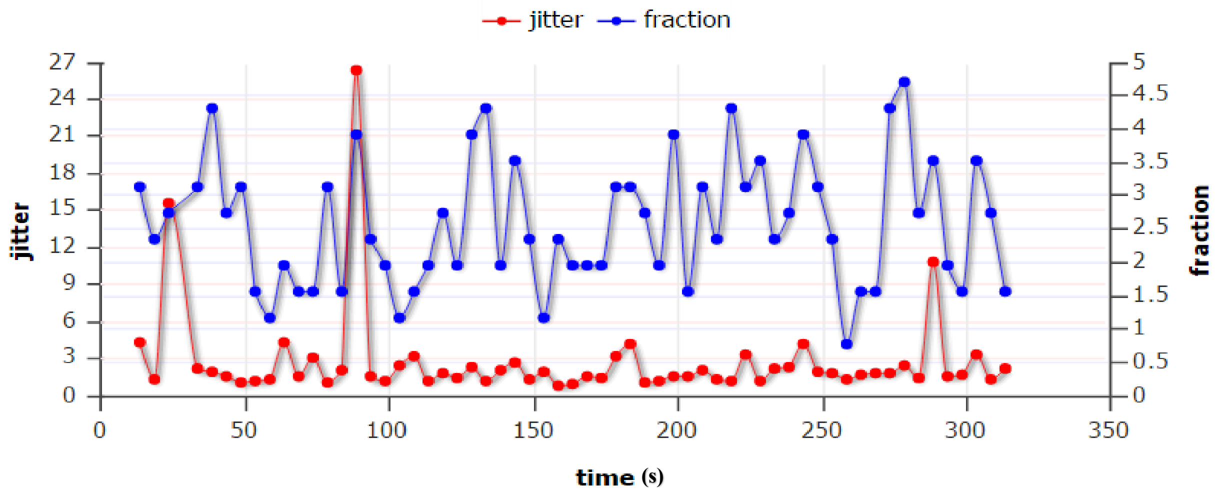

Figure G-26: Caller RTCP jitter for 3% packet loss condition with jitter buffer enable

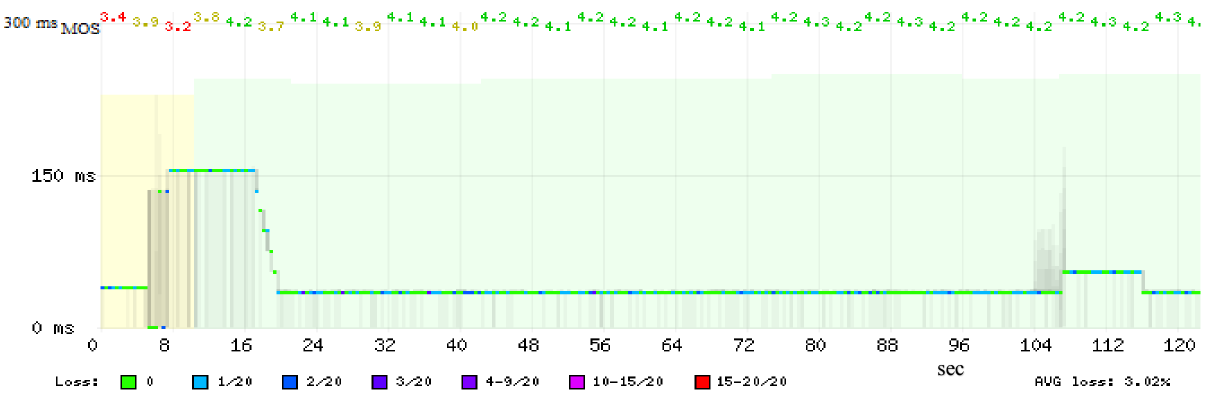

Figure G-27: Receiver RTP stream for 3% packet loss condition with jitter buffer enable

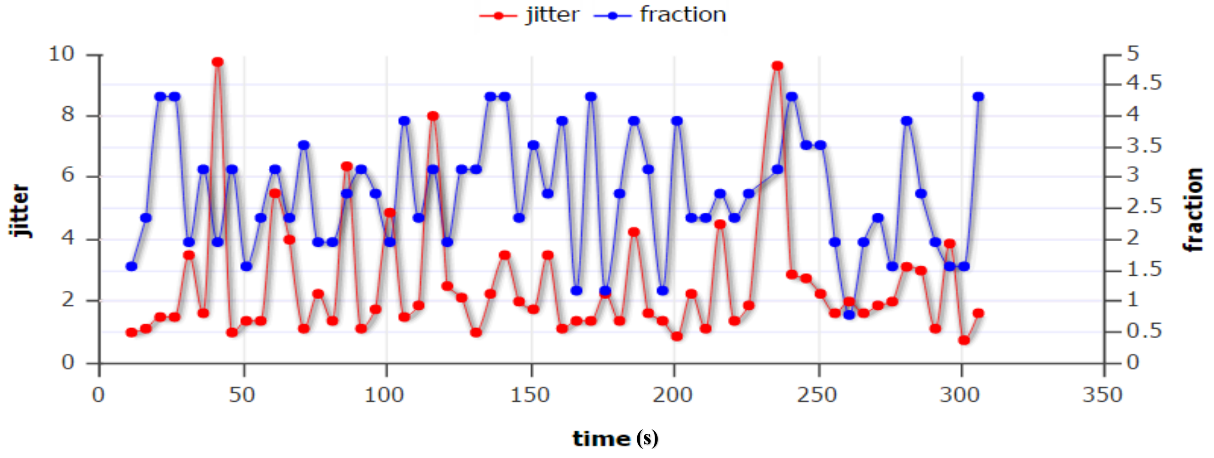

Figure G-28: Receiver RTCP jitter for 3% packet loss condition with jitter buffer enable

Condition VIII: 4% Packet loss inserted in network with jitter buffer disable:

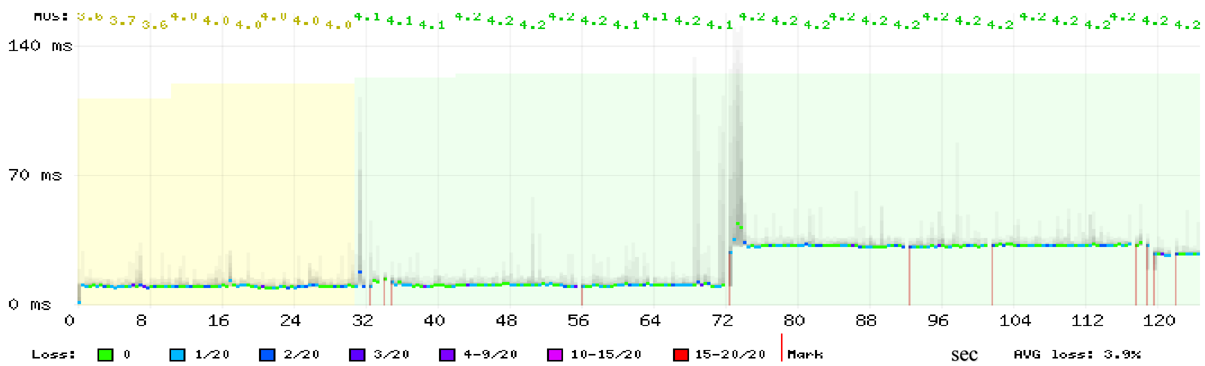

Figure G-29: Caller RTP stream for 4% packet loss condition with jitter buffer disable

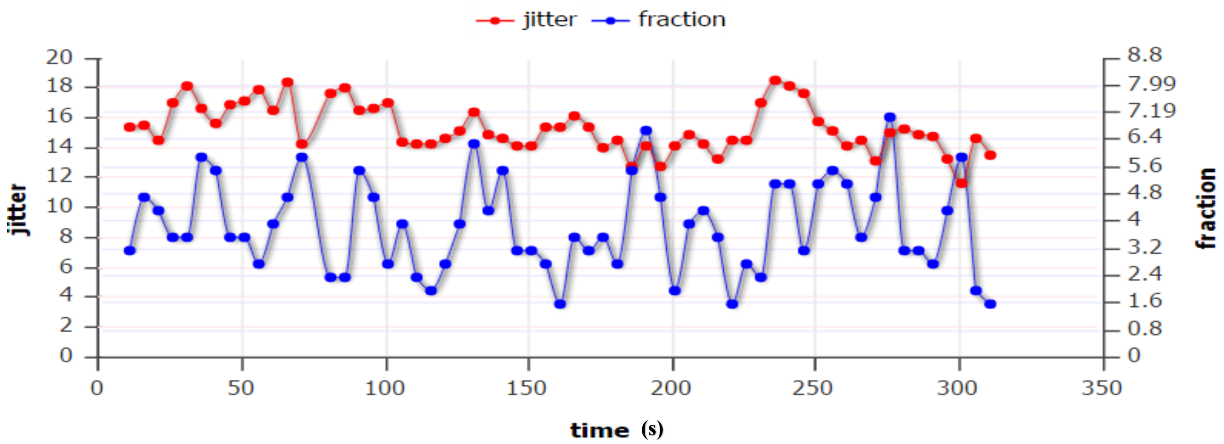

Figure G-30: Caller RTCP jitter for 4% packet loss condition with jitter buffer disable

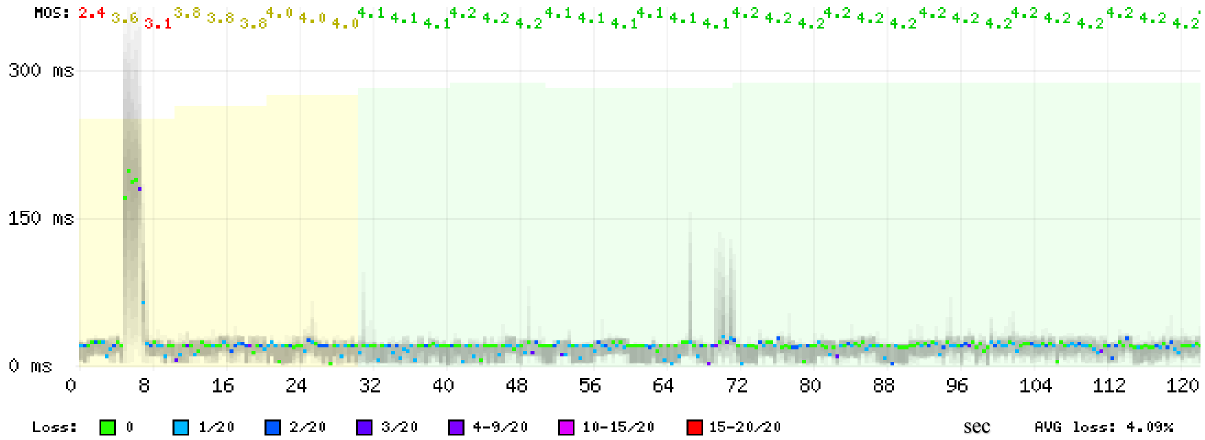

Figure G-31: Receiver RTP stream for 4% packet loss condition with jitter buffer disable

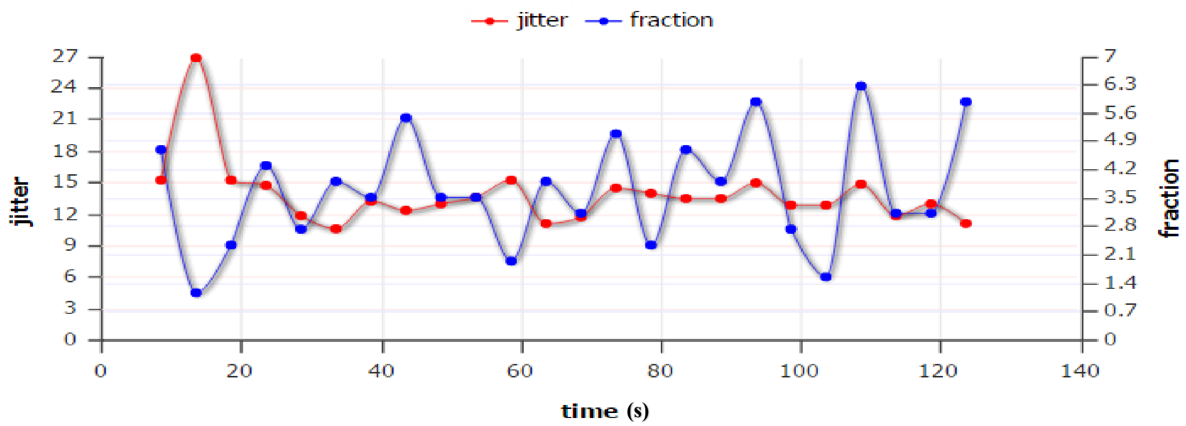

Figure G-32: Receiver RTCP jitter for 4% packet loss condition with jitter buffer disable

Condition VIII: 4% Packet loss inserted in network with jitter buffer enable:

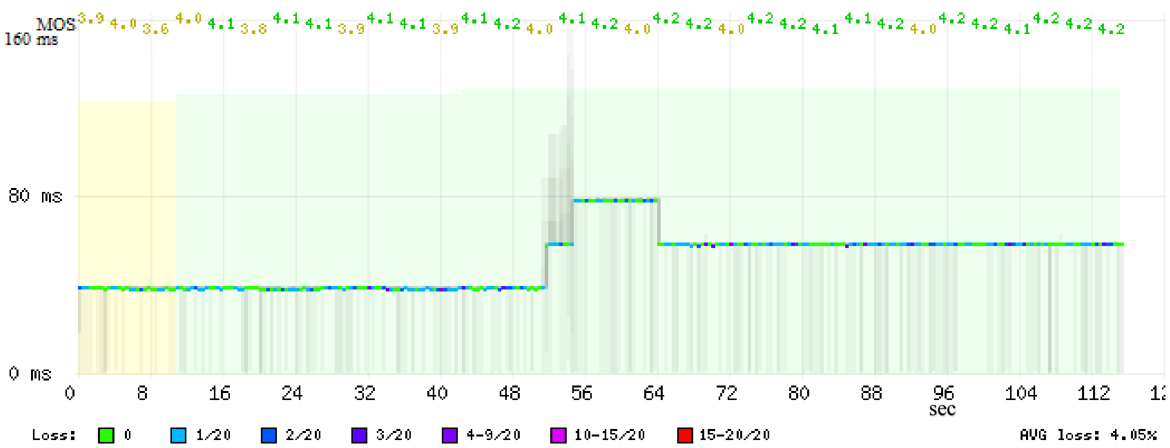

Figure G-33: Caller RTP stream for 4% packet loss condition with jitter buffer enable

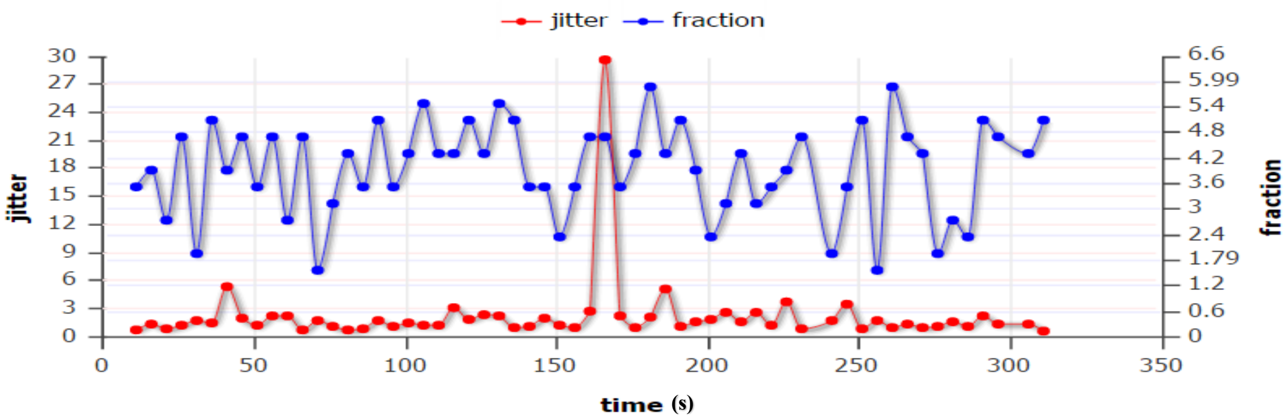

Figure G-34: Caller RTCP jitter for 4% packet loss condition with jitter buffer enable

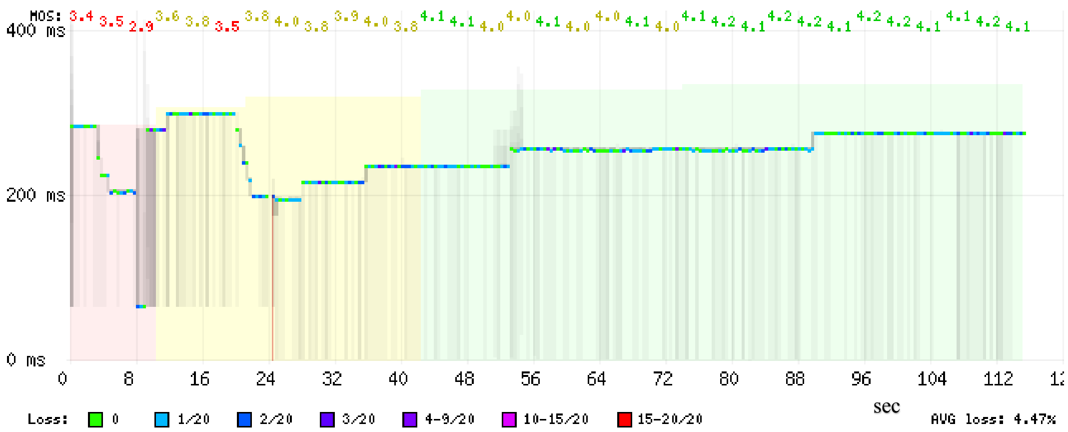

Figure G-35: Receiver RTP stream for 4% packet loss condition with jitter buffer enable

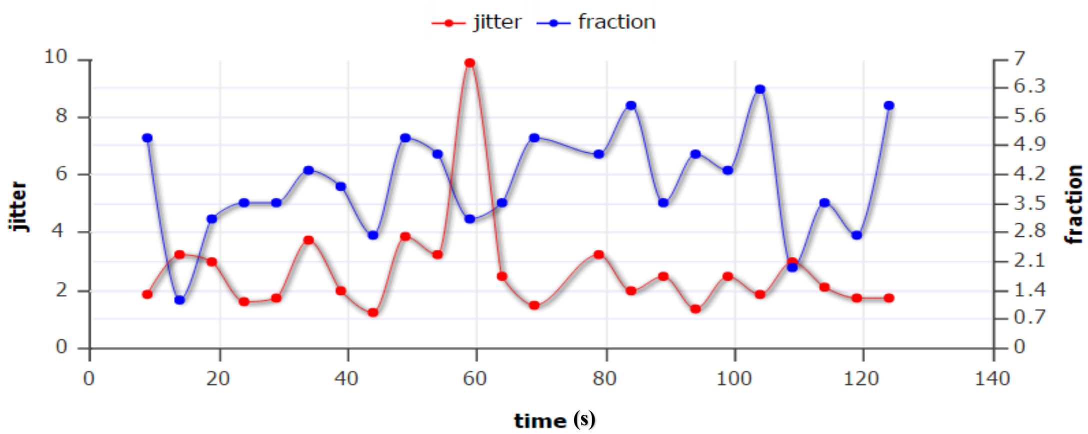

Figure G-36: Receiver RTCP jitter for 4% packet loss condition with jitter buffer enable

Condition IX: 5% Packet loss inserted in network with jitter buffer disable.

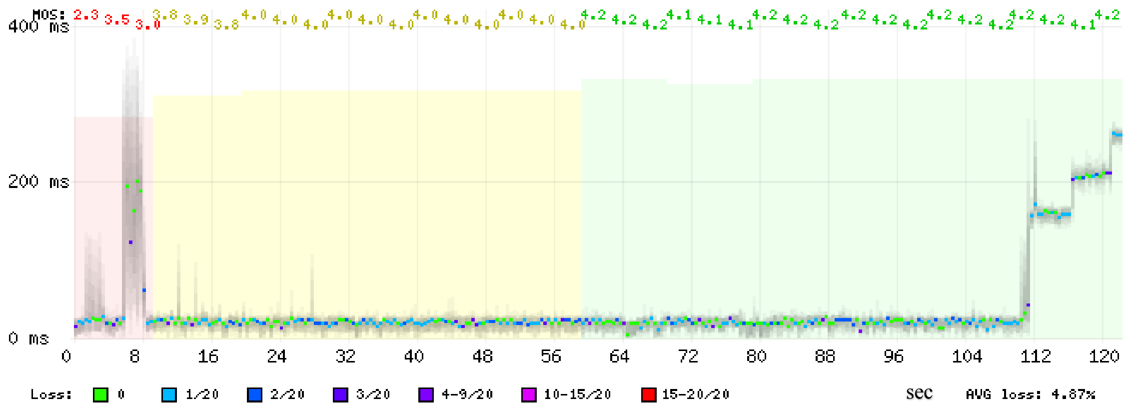

Figure G-37: Receiver RTP stream for 5% packet loss condition with jitter buffer disable

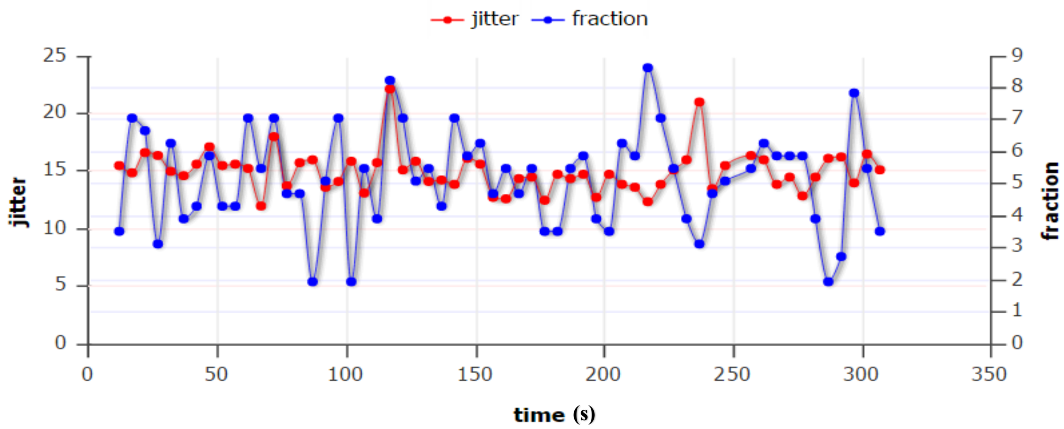

Figure G-38: Receiver RTCP jitter for 5% packet loss condition with jitter buffer disable

Condition X: 5% Packet loss inserted in network with jitter buffer enable.

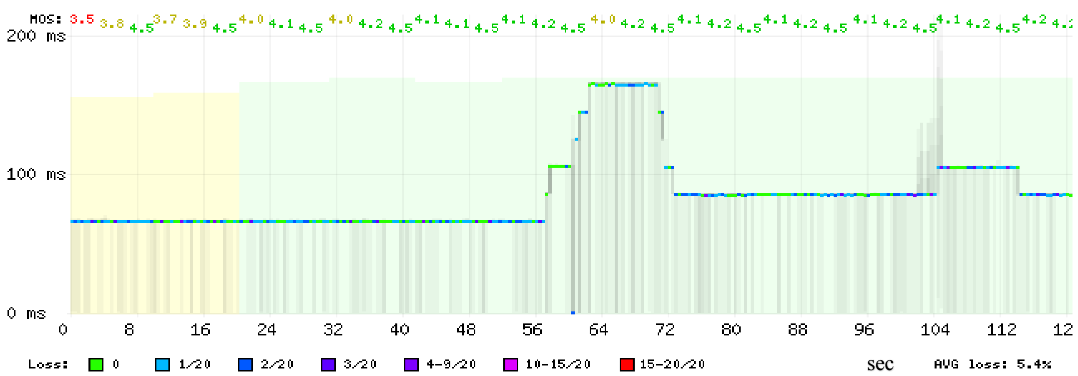

Figure G-39: Caller RTP stream for 5% packet loss condition with jitter buffer enable

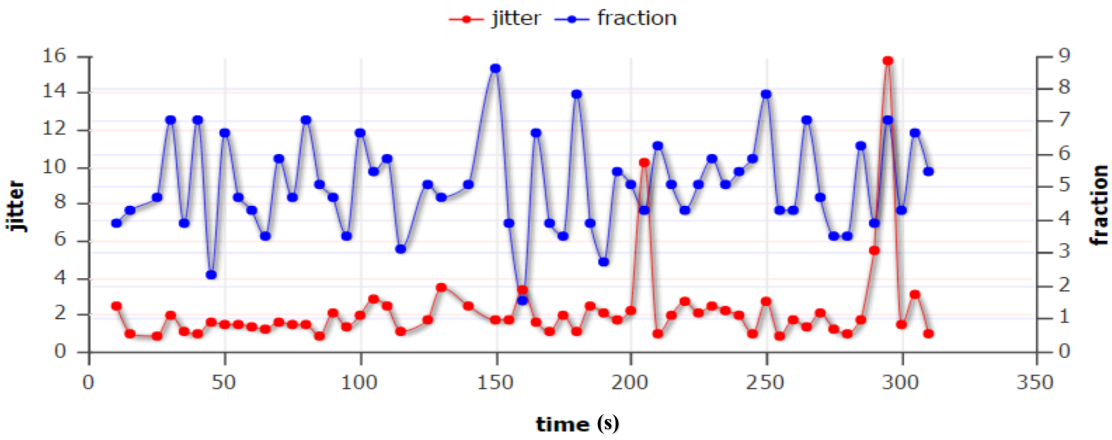

Figure G-40: Caller RTCP jitter for 5% packet loss condition with jitter buffer enable

Condition XI: 10% Packet loss inserted in network with jitter buffer disable.

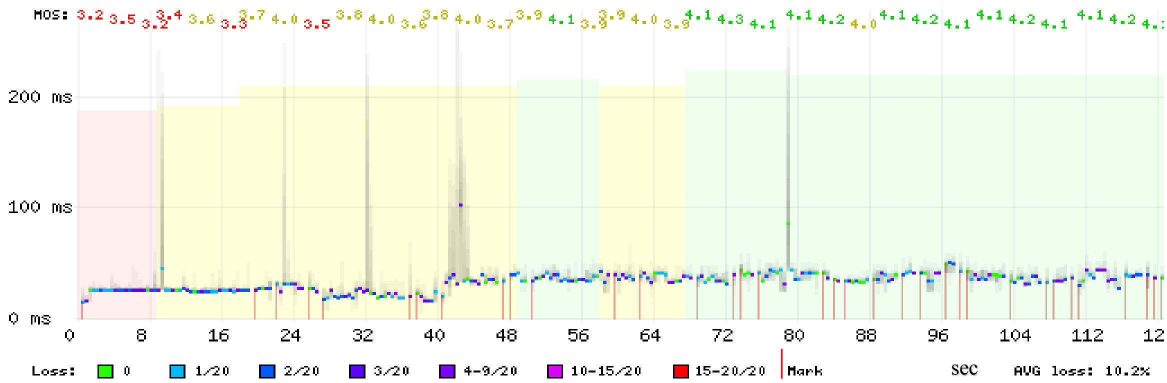

Figure G-41: Caller RTP stream for 10% packet loss condition with jitter buffer disable

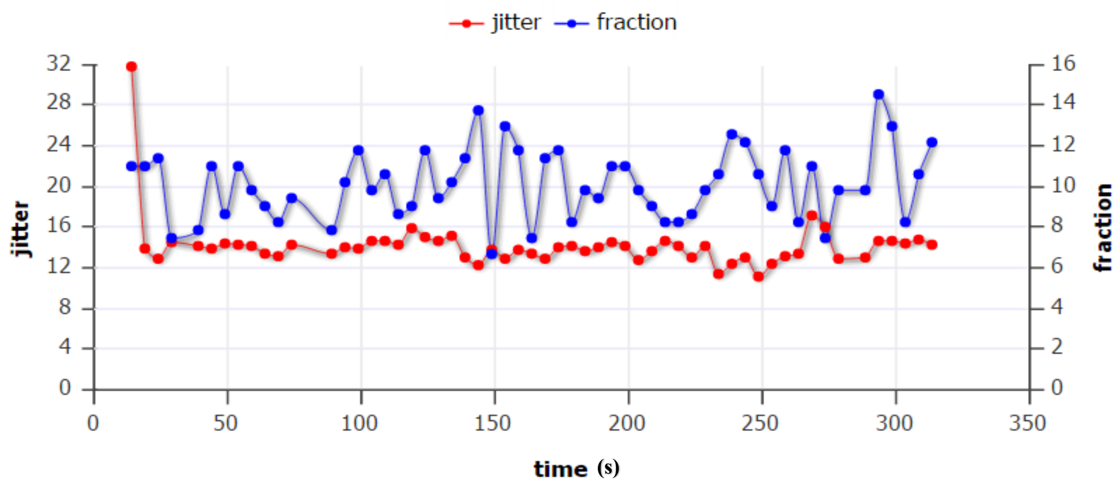

Figure G-42: Caller RTCP jitter for 10% packet loss condition with jitter buffer disable

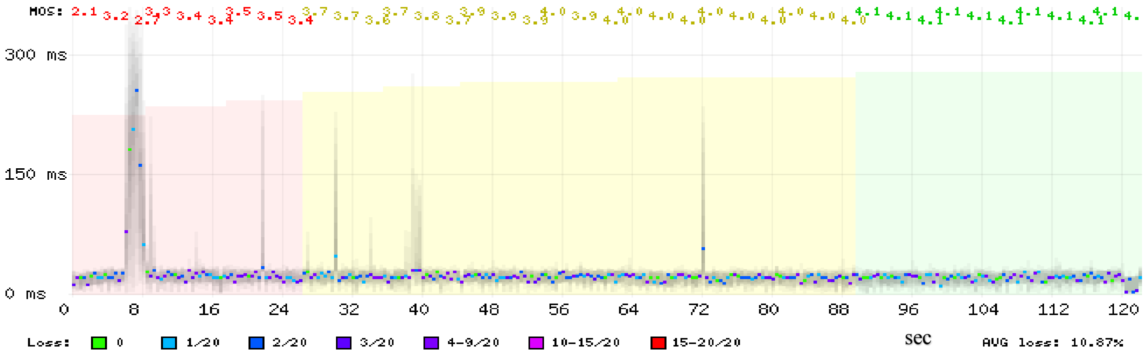

Figure G-43: Receiver RTP stream for 10% packet loss condition with jitter buffer disable

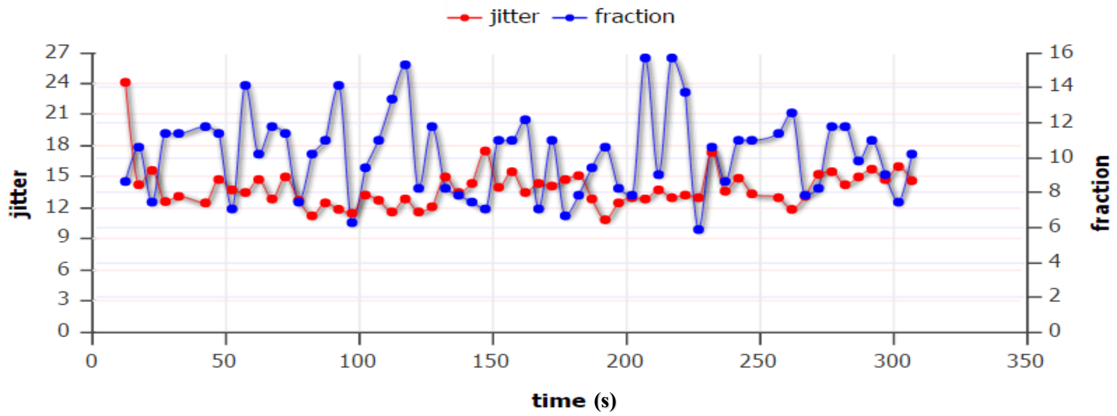

Figure G-44: Receiver RTCP jitter for 10% packet loss condition with jitter buffer disable

Condition XI: 10% Packet loss inserted in network with jitter buffer enable.

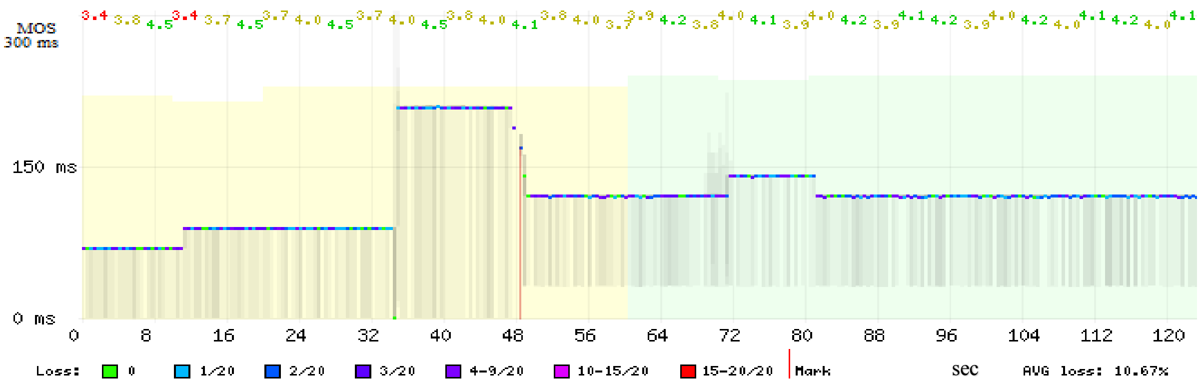

Figure G-45: Caller RTP stream for 10% packet loss condition with jitter buffer enable

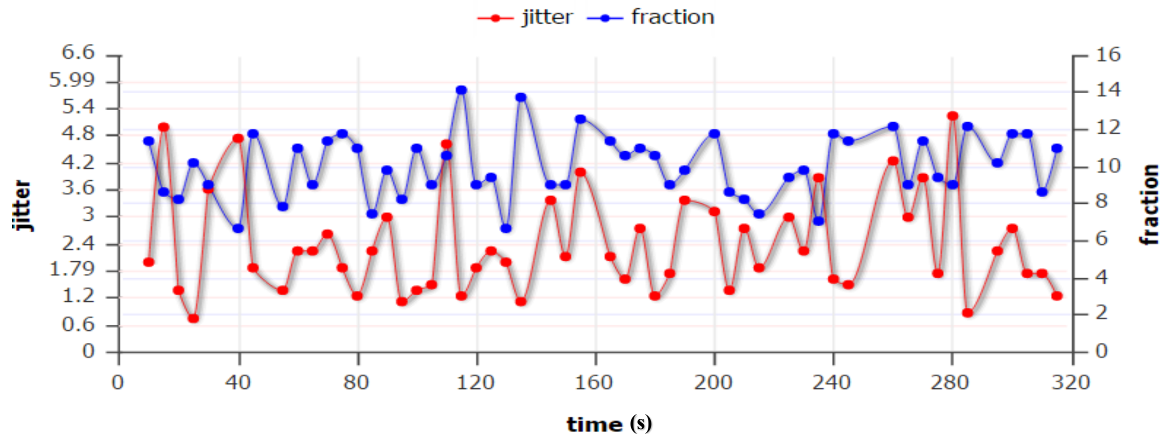

Figure G-46: Caller RTCP jitter for 10% packet loss condition with jitter buffer enable

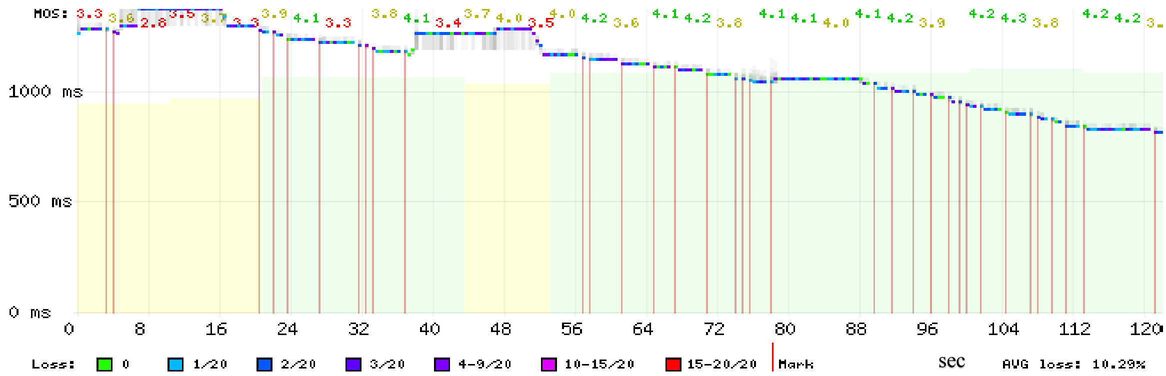

Figure G-47: Receiver RTP stream for 10% packet loss condition with jitter buffer enable

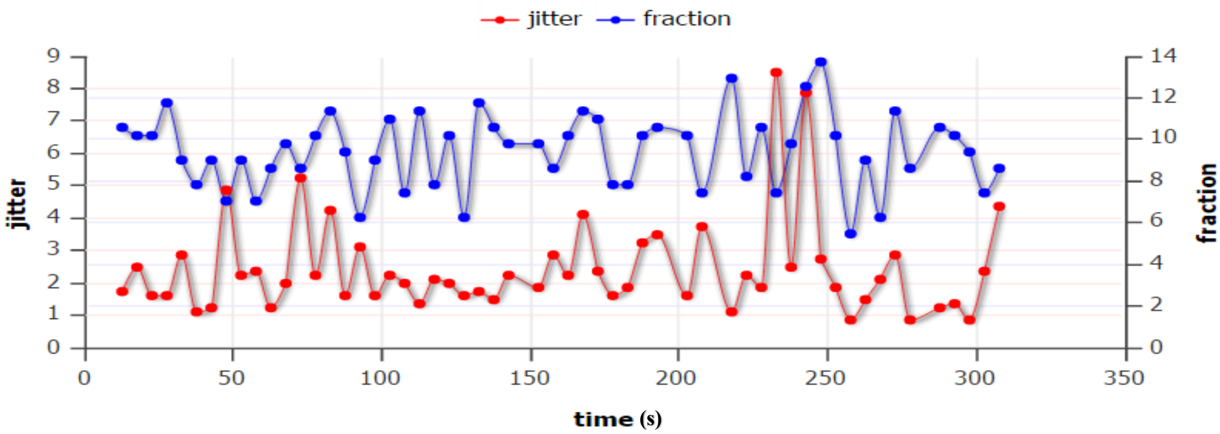

Figure G-48: Receiver RTCP jitter for 10% packet loss condition with jitter buffer enable

Condition XII: 15% Packet loss inserted in network with jitter buffer disable.

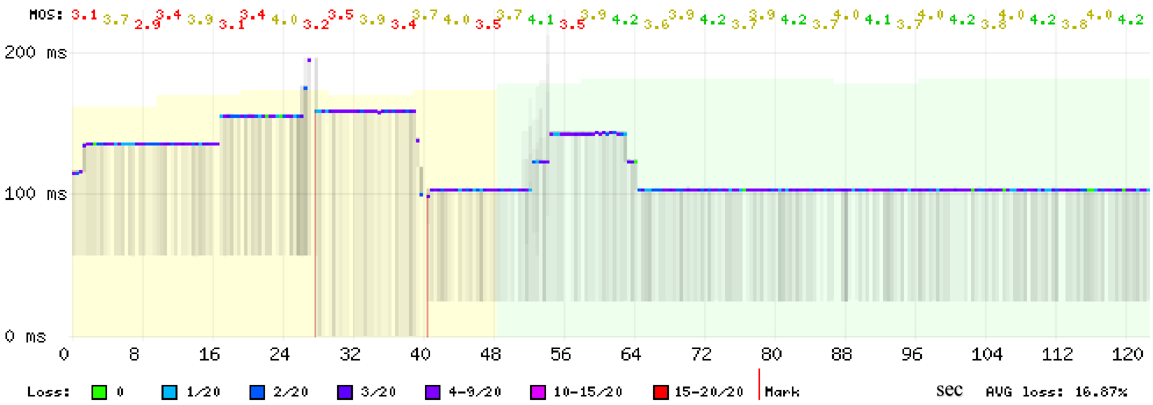

Figure G-49: Caller RTP stream for 15% packet loss condition with jitter buffer disable

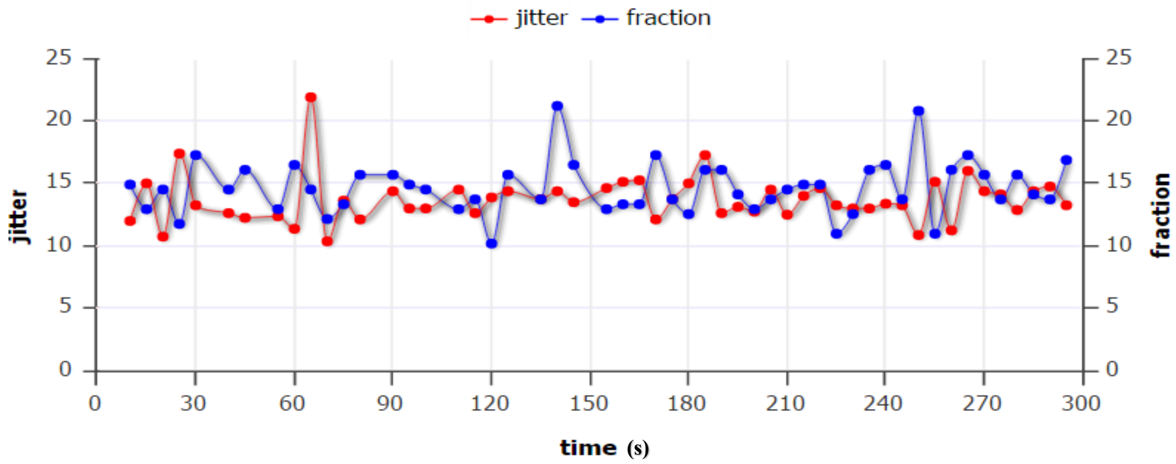

Figure G-50: Caller RTCP jitter for 15% packet loss condition with jitter buffer disable

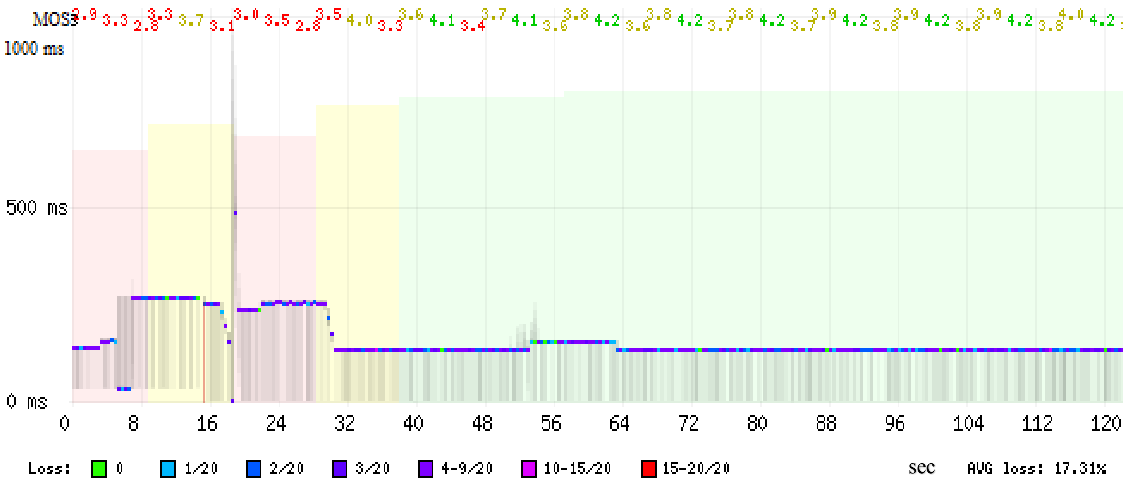

Figure G-51: Receiver RTP stream for 15% packet loss condition with jitter buffer disable

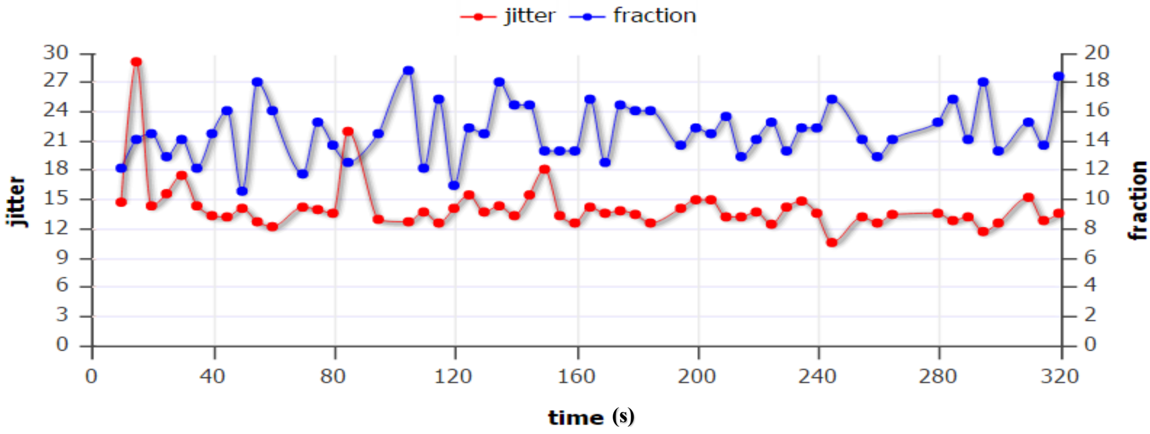

Figure G-52: Receiver RTCP jitter for 15% packet loss condition with jitter buffer disable

Condition XIII: 15% Packet loss inserted in network with jitter buffer enable.

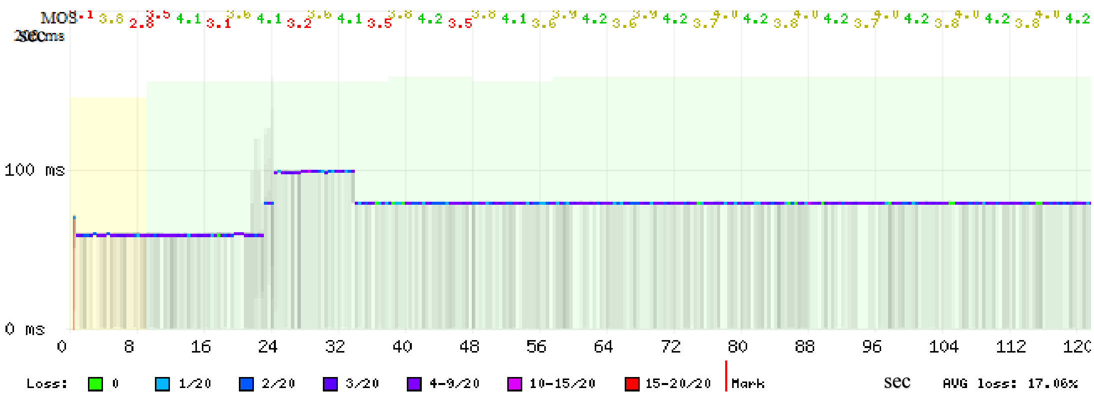

Figure G-53: Caller RTP stream for 15% packet loss condition with jitter buffer enable

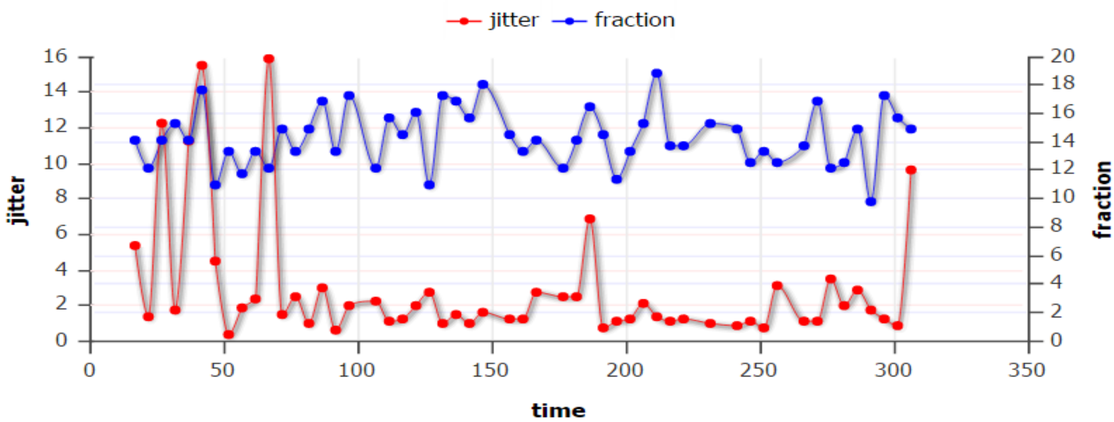

Figure G-54: Caller RTCP jitter for 15% packet loss condition with jitter buffer enable

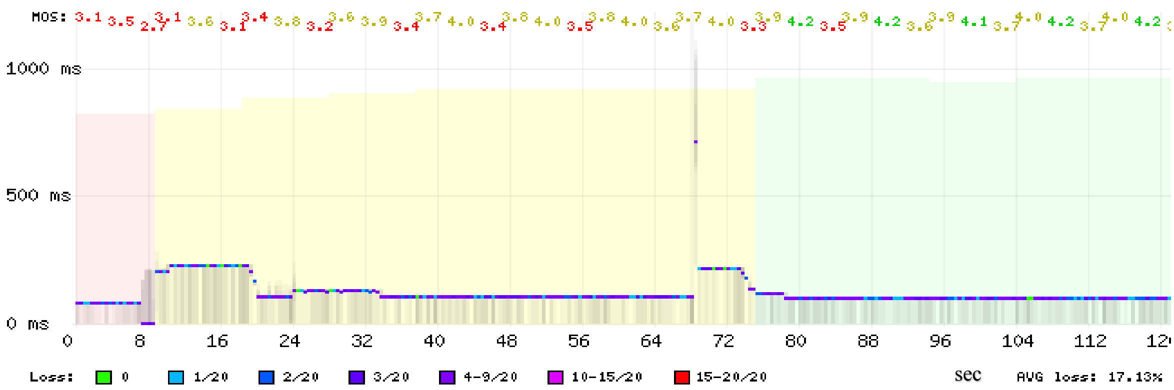

Figure G-55: Receiver RTP stream for 15% packet loss condition with jitter buffer enable

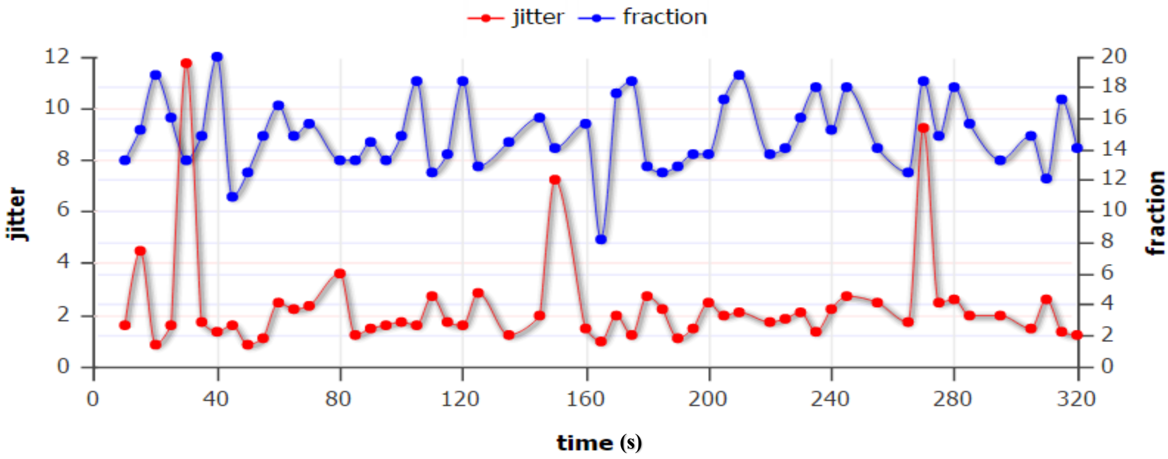

Figure G-56: Receiver RTCP jitter for 15% packet loss condition with jitter buffer enable

Condition XIII: 20% Packet loss inserted in network with jitter buffer disable.

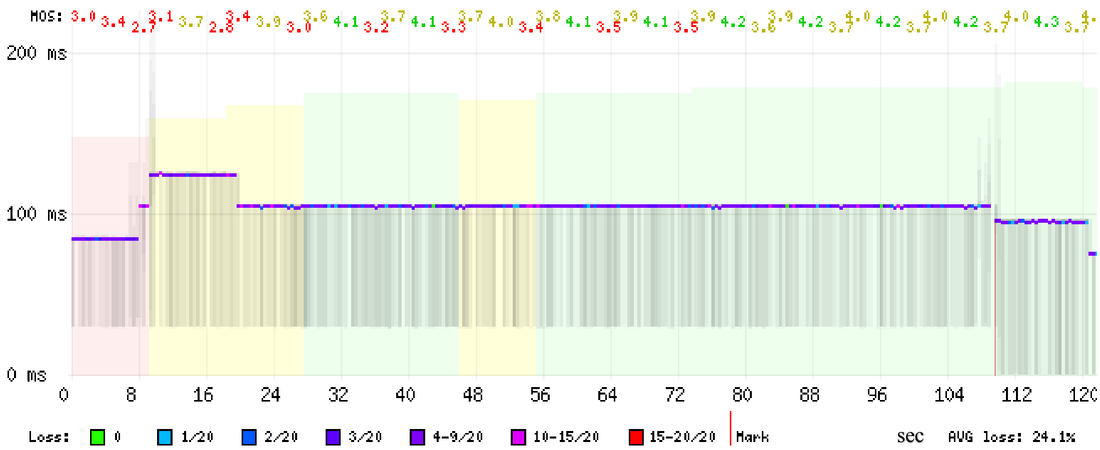

Figure G-57: Caller RTP stream for 20% packet loss condition with jitter buffer disable

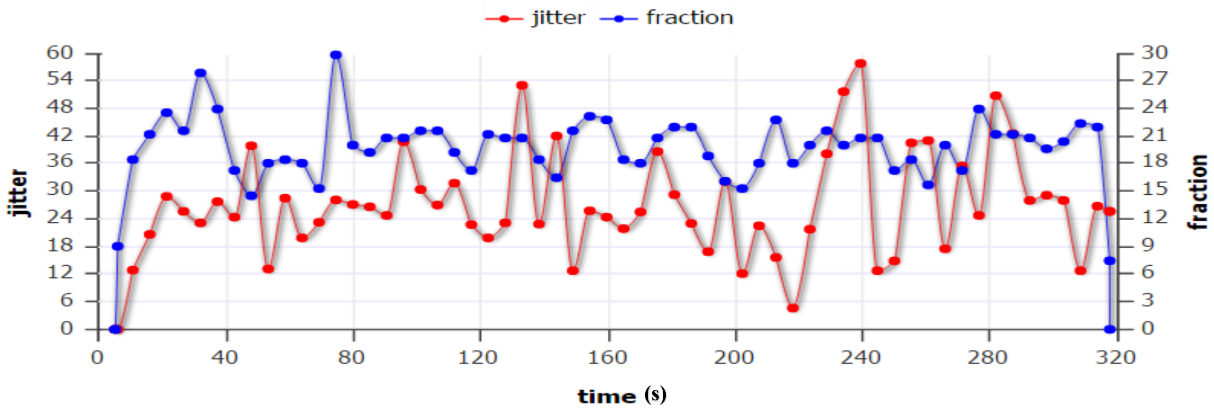

Figure G-58: Caller RTCP jitter for 20% packet loss condition with jitter buffer disable

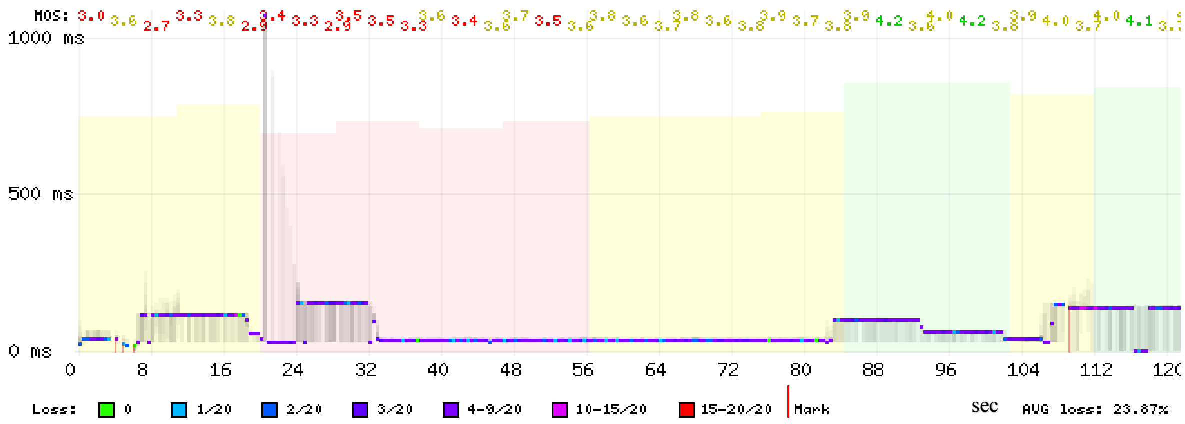

Figure G-59: Receiver RTP stream for 20% packet loss condition with jitter buffer disable

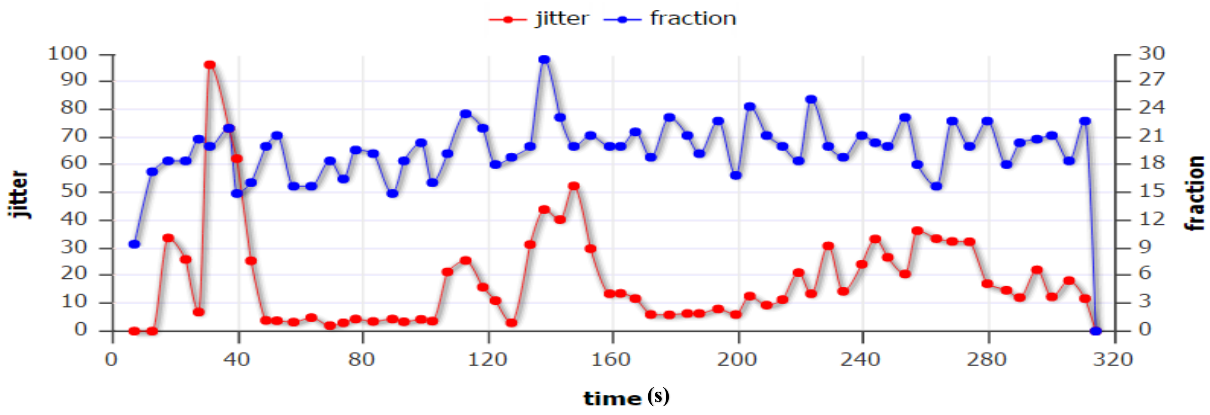

Figure G-60: Receiver RTCP jitter for 20% packet loss condition with jitter buffer disable

Condition XIV: 20% Packet loss inserted in network with jitter buffer enable.

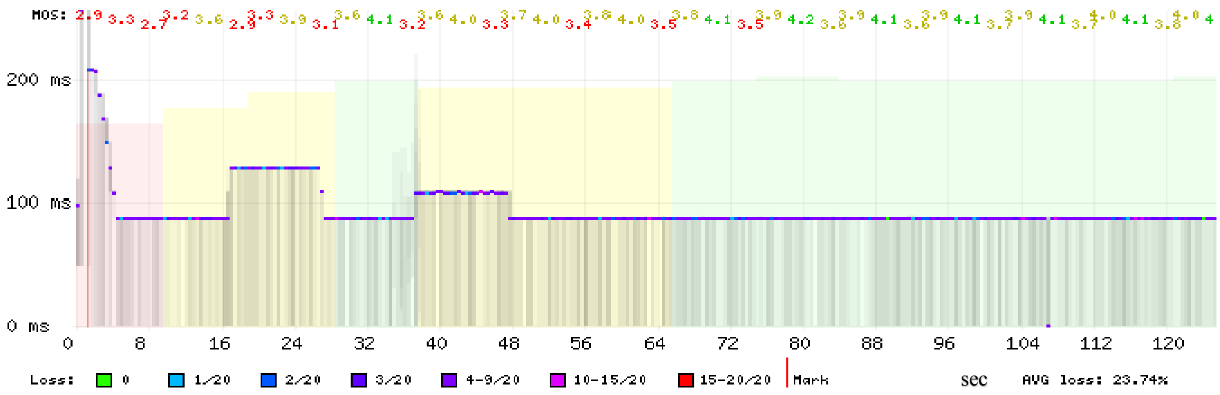

Figure G-61: Caller RTP stream for 20% packet loss condition with jitter buffer enable

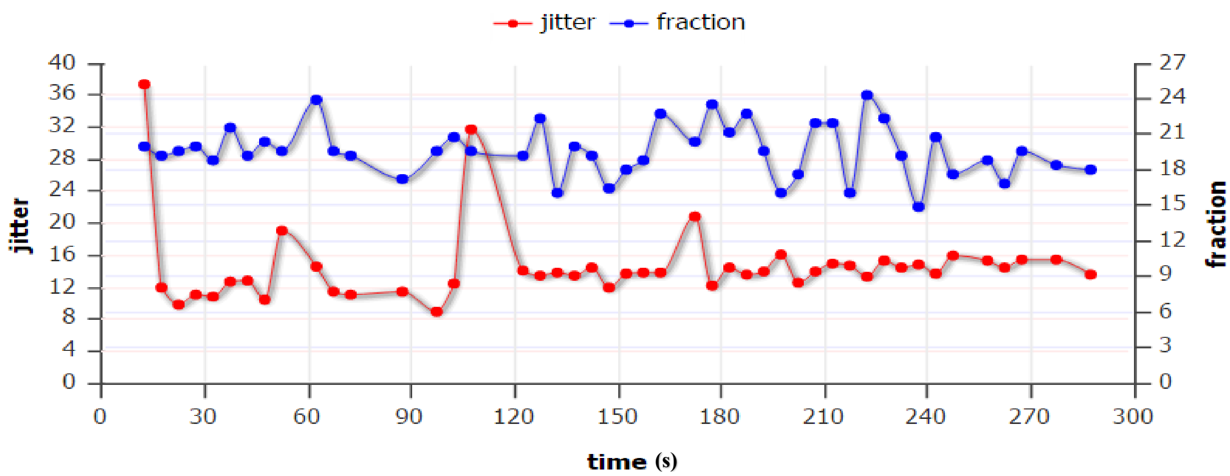

Figure G-62: Caller RTCP jitter for 20% packet loss condition with jitter buffer enable

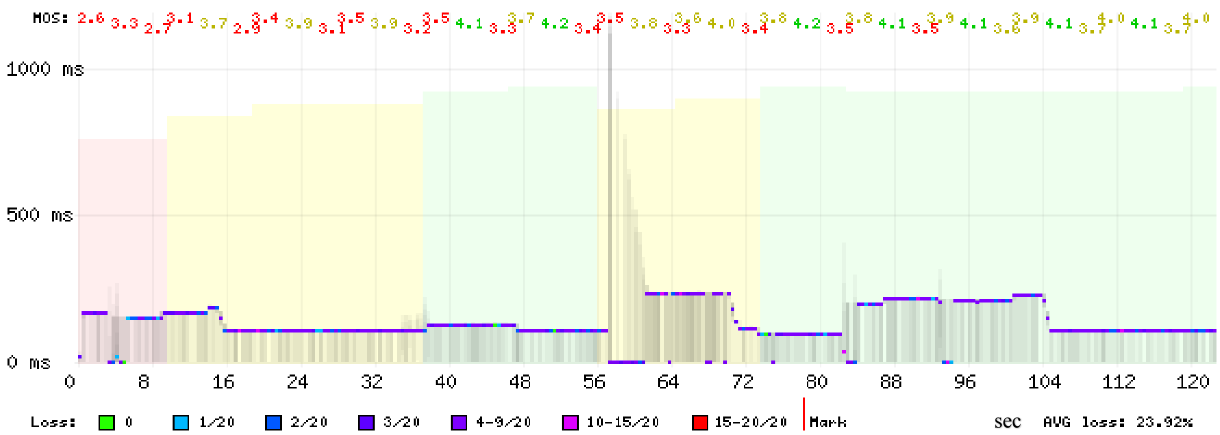

Figure G-63: Receiver RTP stream for 20% packet loss condition with jitter buffer enable

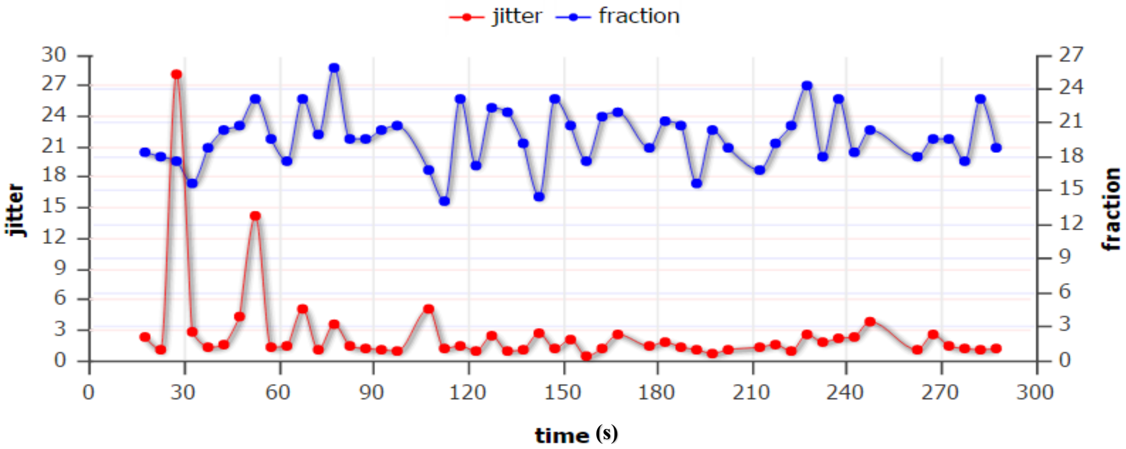

Figure G-64: Receiver RTCP jitter for 20% packet loss condition with jitter buffer enable