IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Issue-7 E-ISSN: 2347-2693

Media Performance Analysis to Increase the Audio Quality in Multimedia Services

Sheela D[1*] and Anil Kumar K M[2]

[1*,2] *Dept. of Computer Science & Engineering, SJCE, India*

www.ijcseonline.org

Received: Jun/04/2015 Revised: Jun/28/2015 Accepted: July/18/2015 Published: July/30/ 2015

***Abstract*—** Mobile communication has drastically increased in the software industry. Multimedia services in mobile communication has also increased its importance. Circuit Switched is now replaced by packet switched by most of the industries. These packets transmission is affected by the varied network parameters. The Audio is also transmitted in packets. The loss of packet or delay of the packet reduces the audio quality on the receiver side. To manage the packets and to maintain smooth audio quality various jitter buffer algorithms are implemented. The Performance of the jitter buffer algorithm affects the quality of audio. Our system analysis the jitter buffer management algorithm used on the receiver side to manage the packets arrived out of order, packet loss or duplicate packets received. The Media Performance Analysis System/tool checks the performance for varied network condition, different audio codecs and for different network profiles. It helps in easy analysis for the performance of the JBM .

***Keywords*—** Jitter buffer management, Amr-NB, Amr-WB, SIP, RTP, MTSI, Forward Error Correction VOIP

I. INTRODUCTION

Multimedia services, such as video conferencing, Internet telephony and streaming audio, have recently been introduced for the millions of users of the Internet. Factors like the type of audio codec used, latency, jitter and jitter buffer, packet loss, packet size, silence suppression, echo and other network parameters that affect the call quality for VOIP applications

Removing jitter involves collecting packets and holding them in the jitter buffer. This allows slower packets to arrive in time to be played out at the appropriate times. Generally the larger the jitter buffer is, the bigger the added delay and the more packets that are successfully played out. Unfortunately this additional delay lowers the perceived QoS. On the other hand, if the playout delay is set too low, the network-induced delay will cause some packets to arrive too late for playout and thus be lost, which also lowers the perceived QoS. The main objective of jitter buffering is to keep the packet loss rate under 5% and to keep the end-to-end delay as small as possible.

There are variety of jitter buffer management algorithms implemented to increase the audio performance on the receiver side. The buffering time of the audio packets received depends on the varied network conditions. The audio quality depends on the network condition, buffering time and JBM algorithm, So there are various JBM algorithms implemented to enhance the audio quality.

The selection of JBM algorithm effects the audio quality. The JBM used in VOIP or volte in multimedia should meet the 3GPP standards and work accurately for varied network conditions. In our work the JBM performance for varied network condition is analyzed. The selection of JBM for the audio quality for different audio codecs like AMRNB, Amr-wb with varied network should meet the standard for smooth audio playback.

The buffering time of the RTP packet, drop of the silence packets on the JBM must not reduce the audio quality. .

II. BACKGROUND

The following sections present a review of existing literature relating to audio quality and jitter buffer management within the scope of this study

A. *Need for Jitter buffer management*

Transmitting real time digitized speech over a best-effort packet networks (such as the Internet), is appealing for a number of reasons, including the wide availability and low cost of such networks. For that reason, Voice over IP (VoIP) has become very popular. Nevertheless, to achieve the high quality standards users expect from commercial solutions, a number of issues inherent to the network need to be addressed. Three issues may particularly affect call quality: packet loss, delay jitter, and clock drift. Packet loss is inherent to the "best effort" characteristics of these networks, and it becomes more pronounced in the Wifi [1] and other lower quality connections. Traditional forward error correction (FEC) techniques would introduce significant extra delay, and thus are not appropriate for real-time communication. So, most VoIP systems use none or simpler FEC techniques (e.g., packet repetition), and rely aggressively on loss concealment techniques. Delay, and delay jitter is another problem. If a packet has not been

received by the time it has to be played out, the decoder will have to interrupt the speech stream. This is generally referred to as "late loss". The old solution was to buffer enough packets to make sure the probability of a late loss is small, but this requires a large buffer, and therefore a long delay. More recent solutions use an adaptive playout strategy, to keep a small buffer size, while reducing late losses. But current adaptive playout technology still incur in late losses. The effect of packet losses can be mitigated by forward error correction or error-resilient audio coding techniques. The occurrence of late loss is traditionally minimized by introducing a *jitter buffer*, which stores packets and provides them to a decoder at more regular time intervals. If the size of the jitter buffer were at least as long
The difference between the smallest and largest possible delays, the late loss would be eliminated entirely. On the other hand, the jitter buffer also introduces an extra delay, which is undesirable for real-time conversations. So, in practice, the buffer size is set to some length which is a compromise between late loss and delay. In general, the size of this buffer can only be adjusted during silence periods. Managing the jitter buffer became crucial for the smooth audio playback.

B. Protocols used in transmitting audio packets

The Transmission Control Protocol (TCP) is the most widely used transport-level protocol in the Internet. However, there are several facts that make TCP quite unsuitable for the real-time traffic. Firstly, TCP includes an in-built retransmission mechanism, which may be useless with strict real-time constraints. Secondly, TCP is a point-to-point protocol without direct support for multicast transmission. Thirdly, there is not any timing information carried, which is needed by most real-time applications. The other widely-used transmission protocol, User Datagram Protocol (UDP), does not either include any timing information. So, a new transport level protocol, called Real Time Transport Protocol (RTP), was specified within the Internet Engineering Task Force (IETF) to cope with the before mentioned problems with the real-time traffic. The IETF's Audio/Video Transport (AVT) working group [1] has since then been the main forum for RTP related discussion and specification work. The International Telecommunications Union (ITU) has also adopted the RTP as the transport protocol for the multimedia. The ITU-T recommendation H.323 [2], and furtherly the recommendation H.225.0[3] include RTP as the transport protocol of multimedia sessions.

C. RTP –Real time Transport Protocol

RTP [4] is a real-time end-to-end transport protocol. However, considering RTP as a transport protocol may be misleading because it is mostly used upon UDP, which is also considered as a transport protocol. On the other hand, RTP is very closely coupled to the application it carries. So, RTP is best viewed as a framework that applications can use to implement a new single protocol. RTP doesn't guarentee timely delivery of packets, nor does it keep the packets in sequence. RTP gives the responsibility for recovering lost segments and resequencing of the packets for the application layer. There are a couple of benefits in doing so. The application may accept less than perfect delivery and with video or speech there usually is no time for retransmissions. What RTP then provides, is:

- Payload type identification
- Source identification
- Sequence numbering
- Timestamping

Which are required by most multimedia applications. The accompanying RTP Control Protocol (RTCP) provides feedback of the quality of the data delivery and information about session participants. A RTP session usually is composed of a RTP port number (UDP port), a RTCP port number (consecutive UDP port) and the participant's IP address. The RTP packet format (Figure 1) is in detail reviewed in the following.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |V=2|P|X|  CC   |M|     PT      |       sequence number         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                           timestamp                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |           synchronization source (SSRC) identifier            |
 +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
 |            contributing source (CSRC) identifiers             |
 |                             ....                              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Format of the RTP packet

The first 32 bits of the header consists of several control bits. The version number (V) is currently 2. The padding bit (P) indicates if there is padding octets inserted at the end of this packet. Padding may be required by some applications with fixed length packet sizes. The extension (X) bit indicates if there is an experimental extension after the fixed header. The count field (CC) tells the number of contributing source identifiers (CSRC) following the fixed header. The marker bit (M) may be used as general marker, f.g. indicating the beginning of a speech burst. The sequence number is an incrementing counter which is started by a source from a random number. The timestamp corresponds to the generation
instant of the first octet in the payload. The synchronization source identifier (SSRC) is a randomly generated value that uniquelly identifies the source within a session. Even if it is very unlikely that two sources generate the same SSRC number, every RTP implementaton should have a mechanism to cope with this chance. Following the fixed

header there are one or more contributing source identifiers which are supplied
by the mixer and the payload.

D. *RTP –Payload*

Before RTP may be used for a particular application the payload codes and the actual payload formats should be defined in a profile specification, which may also describe some application specific extensions or modifications to RTP.
These payload types include for example G.721, GSM Full Rate, G.722 and G.728 speech codecs and JPEG and H.261 video codecs. A new revision of this RFC is about to come, which adds some new types including G.723, G.729 and H.263 codecs. Additionally, there are several separate RFCs or drafts for different codecs (f.g. for MPEG1/2/4, JPEG, H.261 and H.263) which define the payload formats and transport policies in more detail.

E. *Mixers and Translators*

As RTP is designed to support multicast transmission the RTP packet includes a source identifier (SSRC) which identifies the particular sender from the group. There are, however two special kinds of sources: a mixer and a translator. A mixer combines packets from multiple senders and forwards them to one or more destinations. The mixer assigns itself as the sender of the packet and it also resynchronizes the sending (SSRC). The identifiers of all contributing sources (CSRC) are attached to the combined RTP packet. A translator may change the format of the data in the packet, for example if there is a difference in the allowable transfer rate of the end-points.

F. *Jitter Buffer Management in terminals*

Jitter Buffer Management (JBM) denotes the actual buffer as well as any control, adaptation and media processing algorithm (excluding speech decoder) used in the management of the jitter induced in the transport channel. An illustration of an exemplary structure of an MTSI speech receiver with adaptive jitter buffer is shown in figure 1. The blocks "network analyzer" and "adaptation control logic" together with the information on buffer status form theactual buffer control functionality, whereas "speech decoder" and "adaptation unit" provide the media processing functionality. Note that the external playback device control driving the media processing is not shown in figure 2.

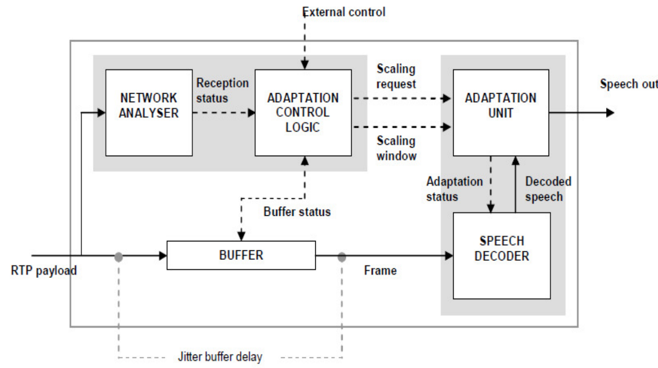

Figure2. Example structure of an MTSI speech receiver

The functional requirements for the speech JBM guarantee appropriate management of jitter which shall be the same for all speech JBM implementations used in MTSI clients. A JBM implementation used in MTSI shall support the following requirements, but is not limited in functionality to these requirements. They are to be seen as a minimum set
of functional requirements supported by every speech JBM used in MTSI. Speech JBM used in MTSI shall:

- Support all the audio codec.
- Support source-controlled rate operation as well as non-source-controlled rate operation;
- Be able to receive the de-packetized frames out of order and present them in order for decoder consumption;
- Be able to receive duplicate speech frames and only present unique speech frames for decoder consumption;
- Be able to handle clock drift between the encoding and decoding end-points.

The jitter buffering time is the time spent by a speech frame in the JBM. It is measured as the difference between the decoding start time and the arrival time of the speech frame to the JBM. The frames that are discarded by the JBM are not counted in the measure. The minimum performance requirements consist of objective criteria for delay and jitter-induced concealment.

III. METHODOLOGY

The following sections will discuss the methodologies adopted for the purpose of this study body.

A. *Aim*

The aim of this study is analyze the performance of the various jitter buffer management algorithms. This helps in efficient usage of JBM for increasing the audio quality.

B. Research Methods

The jitter buffer management algorithms implemented can be tested for varied network conditions. The buffering time of the RTP packets on receiving in the JBM may vary for different network parameters like packet loss, delay and jitter.

The buffering time and the packet drop in the JBM should not affect the audio quality. The performance can be analysed with different network parameters. The methodologies used in our approach is inducing different network parameters using the delay error profiles and the audio packets. The audio packets can be of any audio codec with DTX on and off cases.

The JBM is also tested by extracting the RTP packets from the pcap file generated by the wireshark. The loopback method along with network parameters induced will helps in analysis of packet drop and playout time of RTP packets on the receiver side.

C Computation

The network parameters are given by user. The audio files and the delay error profile files is used in analysis. The system analysis the performance of the JBM by extracting the buffering time of the RTP packet on the receiver side. The JBM tested must meet the 3GPP standards and the audio quality should not be degraded by the extra delay in the jitter buffer.

IV. ALGORITHM

Input: Audio files
Output: JBM delay. JBM performance result
[1] The network parameters is altered.
[2]The JBM performance is affected by the network parameters.
[3] Tested for different audio codec like amrnb and amrwb
[4] Tested using wireshark.

V. DATA COLLECTION

There are different delay and error profiles used to check the tested JBM for compliance with the minimum performance requirements. The profiles span a large range of operating conditions in which the JBM shall provide sufficient performance for the MTSI service. All profiles are 7500 IP packets long. The Delay and error profiles are listed with their characteristics in Table 1.

Table 1: Delay and error profile overview

Profile	**Characteristics**	**Packet loss rate (%)**	**Filename**
1	Low-amplitude, static jitter characteristics, 1 frame/packet	0	dly_error_profile_1.dat
2	Hi-amplitude, semi-static jitter characteristics, 1 frame/packet	0.24	dly_error_profile_2.dat
3	Low/high/low amplitude, changing jitter, 1 frame/packet	0.51	dly_error_profile_3.dat
4	Low/high/low/high, changing jitter, 1 frame/packet	2.4	dly_error_profile_4.dat
5	Moderate jitter with occasional delay spikes, 2 frames/packet (7 500 IP packets, 15 000 speech frames)	5.9	dly_error_profile_5.dat
6	Moderate jitter with severe delay spikes, 1 frame/packet	0.1	dly_error_profile_6.dat

The data is stored as RTP packets, formatted according to "RTP dump" format . The input to these files is AMR or AMR-WB encoded frames, encapsulated into RTP packets using the octet-aligned mode of the AMR RTP payload format.

The table 2 list the different audio files with their codec. The number of frames per packet is also mentioned.

Table 2. Input files for JBM performance evaluation

Codec	Frames per RTP	packet Filename
AMR (12.2 kbps)	1	test_amr122_fpp1.rtp
AMR (12.2 kbps)	2	test_amr122_fpp1.rtp
AMR-WB (12.65 kbps)	1	test_amrwb1265_fpp1.rtp
AMR-WB (12.65 kbps)	2	test_amrwb1265_fpp1.rtp

VI. RESULTS

The JBM is tested in loopback, using pcap and the audio files with different audio codec with varied network parameters. The varied JBM performance for different audio files and delay error profiles are recorded.

Table 3.Testing of audio packet for Profile1

Audio Packet	codec	Packet loss %	JBM Performance	Result
test_amrnb122_3gpp_dtx_fpp1.rtp	amrnb	0	Meets 3GPP standards	Pass
test_amrnb122_3gpp_nodtx_fpp1.rtp	amrnb	0	Meets 3GPP standards	Pass
test_amrnb122_clean_dtx_fpp1.rtp	amrnb	0	Meets 3GPP standards	Pass
test_amrnb122_clean_nodtx_fpp1.rtp	amrnb	0	Meets 3GPP standards	Pass
test_amrwb1265_3gpp_dtx_fpp1.rtp	amrwb	0	Meets 3GPP standards	Pass
test_amrwb1265_3gpp_nodtx_fpp1.rtp	amrwb	0	Meets 3GPP standards	Pass
test_amrwb1265_clean_dtx_fpp1.rtp	amrwb	0	Meets 3GPP standards	Pass
test_amrwb1265_clean_nodtx_fpp1**.rtp**	amrwb	0	Meets 3GPP standards	Pass

The table3 list test cases for profile1 with all the audio files of different codecs. The performance of the JBM for audio with DTX and no DTX passes and meets the 3GPP standard. The media performance system test accurately the JBM implemented.

Table 4. Testing of audio packets for Profile 5

Audio Packet	codec	Packet loss %	JBM Performance	Result
test_amrnb122_3gpp_dtx_fpp1.rtp	amrnb	5.9	Fails to meet 3GPP standards	Pass
test_amrnb122_3gpp_nodtx_fpp1.rtp	amrnb	5.9	Meets 3GPP standards	Pass
test_amrnb122_clean_dtx_fpp1.rtp	amrnb	5.9	Fails to meet 3GPP standards	Pass
test_amrnb122_clean_nodtx_fpp1.rtp	amrnb	5.9	Meets 3GPP standards	Pass
test_amrwb1265_3gpp_dtx_fpp1.rtp	Amrwb	5.9	Meets 3GPP standards	Pass
test_amrwb1265_3gpp_nodtx_fpp1.rtp	Amrwb	5.9	Meets 3GPP standards	Pass
test_amrwb1265_clean_dtx_fpp1.rtp	Amrwb	5.9	Meets 3GPP standards	Pass
test_amrwb1265_clean_nodtx_fpp1**.rtp**	Amrwb	5.9	Meets 3GPP standards	Pass

The table 4 gives the experimentation of audio packets for the delay error profile 5 which has packet loss. The JBM performance meets the 3GPP standards for all the profiles. The MPAT is used in testing both on the system and on the target. Based on the JBM algorithm developer test case results the MPAT results are passed.

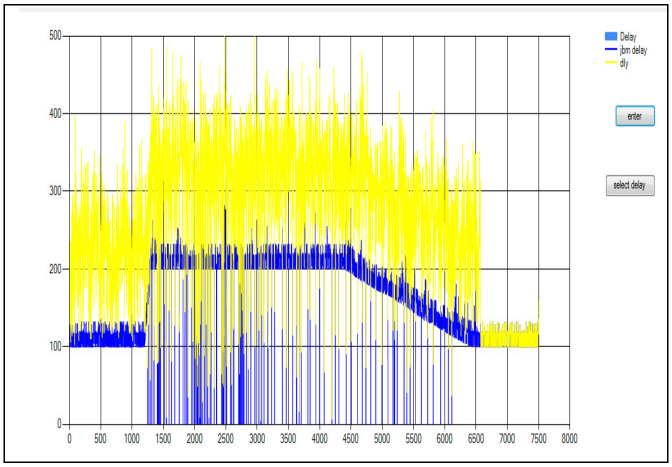

Figure3. Behavior of JBM for Profile5 and amrnb122_3gpp_dtx_fpp1.rtp

Figure 3 shows the behavior of the JBM for profile 14 and audio packet with DTX. The delay along with the network delay should not exceed the jbm delay. The MPAT plots a graph using the JBM delay. The network delay is added to the JBM delay and is plotted in yellow. This graph helps in analyzing for which packet or at what time is the behavior of the JBM changes.

The network parameters effecting the performance of the JBM will behave differently for the varied conditions. The variation of the JBM delay indicates the loss of packet or increase in the buffering time of the RTP packet.

The buffering algorithm in this test case shows changes of JBM delay drastically which indicates the buffering time increases as the network parameters become worse. The profile 14 has a bad network condition. The drastic change in delay indicates the performance of the JBM.

The performance of the media performance analysis system is accurate to 97%. It is calculated based on the developers review.

CONCLUSION

The approach presented in this work has been tried and tested with design diagrams obtained from various concept engineers working in various domain. So far, the results are positive and encouraging. The approach increases the efficiency of the user as the time taken to analyze the performance of Jitter buffer management algorithm implemented will be reduced. This approach also helps in increasing the audio quality in most of multimedia services. There are various jitter buffer algorithms implemented to increase the audio quality. To analyze the performance of these algorithm and use in multimedia services will be effective using our application.

This approach helps in effective use of jitter buffer management in various multimedia services like VOIP call, volte calls. The JBM used should meet the 3GPP standards for the audio quality.

REFERENCES

[1] Enhanced Adaptive Playout Scheduling and Loss Concealment Techniques for Voice over IP Networks. Dinei Florencio and Li-Wei He Microsoft Research One Microsoft Way, Redmond, WA 98052 I. F. Akyildiz, and X. Wang, "A Survey on Wireless Mesh Networks", IEEE Radio Communications, Vol.43, Issue-3, **2005,** pp.**23-30.**

[2] ITU-T Recommendation H.323 : Packet-Based Multimedia Communications Systems. February 1998.

[3] ITU-T Recommendation H.225.0 : Media Stream Packetization and Synchronization on Non- Guaranteed QualityofServiceLANs.November.

[4] A System for Buffer Management in Voice Communication. World Applied Sciences Journal 21 (9): 1305-1308, 2013 ISSN 1818-4952 © IDOSI Publications, 2013 DOI: 10.5829/idosi.wasj.2013.21.9.2886

[5] On Packet Loss Performance under Varying Network Conditions with Path Diversity Richard Haywood and Xiao-Hong Peng Electronic Engineering Aston University Birmingham, United Kingdom haywoorj@aston.ac.uk

[6] Dynamical Management of Dejitter Buffers Based on Speech Quality. Raquel F. do Valle, Leandro S. G. de Carvalho, Regeane B. Aguiar, Edjair S. Mota, Dio´genes Freitas Computer Science Department – Federal University of Amazonas (UFAM) Manaus – Amazonas – Brazil

[7] Technical Specification Group Services and System Aspects (2006), IP Multimedia Subsystem (IMS), Stage 2, TS 23.228, 3rd Generation Partnership Project

[8] Alexander Harrowell, Staff Writer (October 2006), A Pointless Multimedia Subsystem?, Mobile Communications International, archived from the original on September 2010.

[9] Rajinder Singh, "4G By WiMAX2 and LTE-Advance", International Journal of Computer Sciences and Engineering, Volume-01, Issue-03, Page No(36-38).

