

Comparative Study and Analysis of various VoIP coding Algorithms

Hardeep Singh
Department of Physics
Guru Nanak Dev University
Amritsar, 143005 (India)

M. Mian
Department of Physics
Guru Nanak Dev University
Amritsar, 143005 (India)

ABSTRACT

Voice over internet protocol (VoIP) is a method of providing phone services over dedicated public IP networks. It allows significant cost savings over traditional Public Switched Telephone Networks (PSTN). Speech quality, as perceived by the users of VoIP telephony, is critically important. Signal quality of the VoIP system is degraded by various network layer problems, which include delay, packet loss and jitter. The implementation of signal through various coding algorithms and digital signal processor can improve the quality of degraded VoIP signal. The present work deals with comparative study and analysis of VoIP Codecs (G.711, G.729, AMR, AMR-WB etc.). The VoIP simulations are conducted for G.711, G.729 and AMR-WB speech coders for different network conditions. The coding algorithms are implemented on VoIP speech signal. The results are validated through the measurement of enhancement signal using perceptual evaluation of speech quality (PESQ)/MOS measurement.

General Terms

Signal Processing, Simulation, Coding Algorithms.

Keywords

VoIP, QoS, G.711, G.729, AMR, AMR-WB.

1. INTRODUCTION

Nowadays, communication via Internet (VoIP) is the most widely used technology used for transmission of speech signal over packet-switched IP Networks. A phone call that uses VoIP converts an analog input signal to digital signal, place those digital signals into packets with independent source and destination network addresses and finally send the packetized information over the Internet networks instead of the traditional PSTN lines. However, it is important to note that both the transmitting (source) and receiver (destination) terminals must support the particular codec for proper encoding and decoding [1, 2]. The bandwidth requirements of speech transmission over IP networks are relatively low than PSTN lines. It takes only 6-8 Kbps or less of bandwidth, whereas PSTN takes 64Kbps of bandwidth to make a call. But to provide QoS is one of the problems in implementation of voice over IP. The quality of voice is the characteristic of the IP network and it is affected by various parameters such as delay, jitter, packet loss and network congestion. The effect of these parameters results in harmful effects on the quality of VoIP. [3, 4]

2. SPEECH CODECS

Speech coding describes the conversion of analog voice to the digital form and then compress these digitized samples to reduce the consumption of network bandwidth required to transmit the speech signals. It uses speech specific parameter

estimation using audio signal processing techniques to model the speech signal so as to represent modeled parameters of signal in a compact bit stream of packets. A speech decoder receives coded frames and synthesizes reconstructed speech signal at the receiver. Standards typically dictate the input output relationships of both coder and decoder. Speech coding has two most important applications, mobile telephony and Voice over IP. The VoIP network can support several different codecs for audio. The codecs differs in type of voice coding, coding bit rate and frame size, complexity (operations per second), delay (ms), and perceptual quality of the synthesized speech. Narrowband (NB) coding refers to coding of speech signals whose bandwidth is less than 4 kHz (8 kHz sampling rate), while wideband (WB) coding refers to coding of 7-kHz-bandwidth signals (14–16 kHz sampling rate). NB coding is more commonly used than WB coding because of its narrowband nature. Recently, there has been an increased effort in wideband speech coding because of its use in several applications such as videoconferencing [3-5]. Following codecs are frequently used in VoIP communication.

2.1 ITU-T G.711

G.711 is a codec that was introduced by ITU in 1972 for use in digital telephony. The G.711 describes a simple way to digitize analog data using pulse code modulation (PCM) technique. It has two variants based on type of coding algorithm, A-law and Mu-law. It uses a logarithmic compression. The encoded stream is 64 kbps, consisting of 8 kHz sampling of 8 bit signals. The frame length is eight 125 μ s samples, or 1 ms. Using G.711 for VoIP gives the best voice quality. It is the same codec used in the PSTN network and Integrated Services Digital Network (ISDN) lines. Benefits of G.711 are simple implementation and very good perceived audio quality, having MOS value 4.1. The only downside is that it utilizes more bandwidth than other codecs, up to 84 Kbps. [6, 7]

2.2 ITU-T G.729

G.729 is based on conjugate-structure algebraic-code-excited linear prediction (CS-ACELP) coding model. It is an audio data compression algorithm for voice that compresses digital voice in packets of 10 milliseconds duration. [1] G.729 operates at a bit rate of 8 kbit/s, but there are extensions, which provide rates of 6.4 kbit/s and 11.8 kbit/s for marginally worse and better speech quality, respectively. G.729 has been extended with various features, commonly designated as G.729a and G.729b. Because of its low bandwidth requirements, G.729 is mostly used in Voice over Internet Protocol (VoIP) applications (such as Skype) [5, 8].

2.3 Adaptive Multi-rate Wideband (AMR-WB)

Adaptive Multi-Rate Wideband Codec (AMR-WB) is a speech coder standard introduced by the 3rd Generation Partnership Project (3GPP), which is a partnership project of various standards organizations, for compressing the toll quality speech (16,000 samples/second). The AMR-WB Codec has been approved by the ITU-T standards body and is referred to as G.722.2. This codec has nine basic bit rates,

23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.6 kbit/s. This codec works on the principle of Algebraic Code Excited Linear Prediction (ACELP) for all bit rates. AMR-WB offers even lower bit rate compressions as well as the ability to quickly adapt to varying compression as per the network conditions. The bandwidth is automatically conserved when network congestion is high. When the congestion returns to a normal mode, a lower-quality bit rate is restored. [5, 9, 10] Here table 1 shows the comparative analysis of various codecs on the basis of coding algorithm, bit rates, sampling rate and Mean opinion score

Table 1. Specifications of Different Codecs

Number	Standard By	Description	Bit rate (kb/s)	Sampling rate (kHz)	Frame Size (ms)	MOS
G.711	ITU-T	Pulse code modulation (PCM)	64	8	Sampling	4.1
G.726		40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)	16/24/32/40	8	Sampling	3.85
G.729	ITU-T	Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)	8	8	10	3.91
AMR	3GPP		23.85, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.6	8		4.14

3. MODELING AND SIMULATION OF VoIP SYSTEM

In Voice over Internet Protocol (VoIP) the voice signal is processed through the IP network and the quality of the VoIP signal is degraded by delay, delay variation (jitter) and packet loss. To analyze the effect of these degradation parameters, a VoIP system has been designed as shown in the figure 1 below.

Basic Steps carried out in the design of VoIP system are listed as below:

- Step 1: The original speech signal as an input to the system.
- Step 2: The speech signal is encoded using G.711, G.729, AMR coders at different bit rates and it converts the signal to digital bit stream then compress this stream.

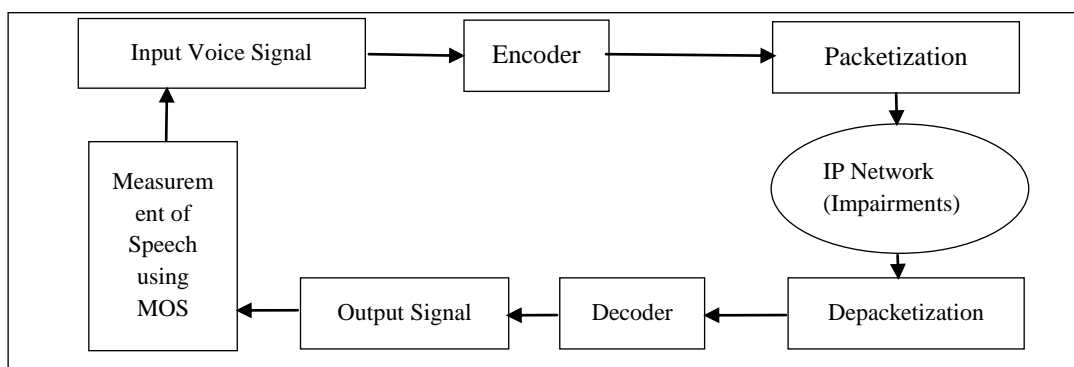


Figure 1: Design Layout of VoIP system

Step 3: The compressed version of the signal is then packetized and send over the IP network.

Step 4: The digitized signal has been degraded at this level because of network impairments. These includes packet loss, delay, jitter and noise.

Step 5: Degraded signal is then depacketized and decoded at the receiver end using decoders.

Step 6: Performance of the perceived VoIP signal is analyzed using the PESQ measurement given by ITU-T recommendation P.862. [12]

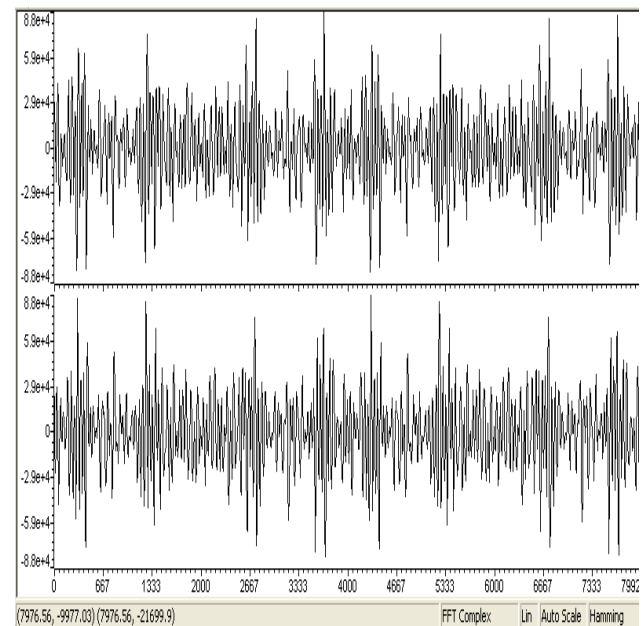
The signal was encoded using G.711, G.729 (8 kbps) and AMR-WB (6.6 kbps) coders. To study the effect of various impairment parameters like delay, packet loss on the quality of VoIP signal after passing through channel, the resultant VoIP signal was analyzed with various DSP algorithms. Assessment of quality of the received VoIP signal is done using mean opinion score (MOS). The Mean Opinion Score (MOS) is a subjective number indicating how people feel about the quality of the voice signal. MOS is measured on a scale from 1-5 where 1 is the lowest and 5 the highest. Delay is an important factor which adversely affects the quality of the VoIP signal. A delay between 0 to 150 ms is acceptable for most user applications; and between 150 to 400 ms for international connections. If the one way delay through a voice network exceeds the threshold value, the natural conversation is strained, due to this reason most network deployments keep the delay well below that threshold value. So in any case a delay of more than 400 ms, is unacceptable for a general network. [11, 12] The spectrum of the degraded signal is analyzed in time and frequency both. The various signal processing algorithms. [13]

FFT Spectrum: A common use of FFT's is to find the frequency components of a signal buried in a noisy time domain signal. The spectrogram of the speech signal is an intensity plot of the STFT magnitude. STFT is a sequence of

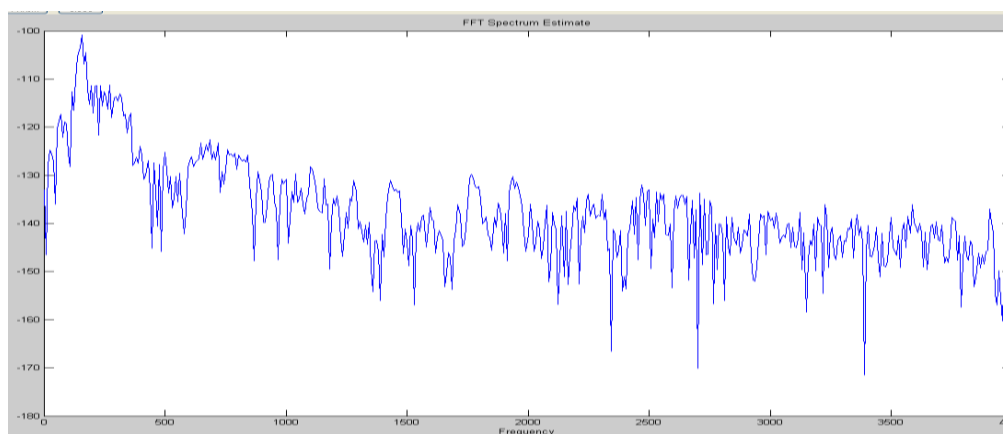
FFTs of windowed data segments where the windows are usually allowed to overlap in time [5]. Here in the present work a time axis for data is, running from $t=0$ until $t=.25$ in steps of 1 millisecond. This generates spectrum of speech signal at 50 Hz and 250 Hz.

4. RESULT AND DISCUSSION

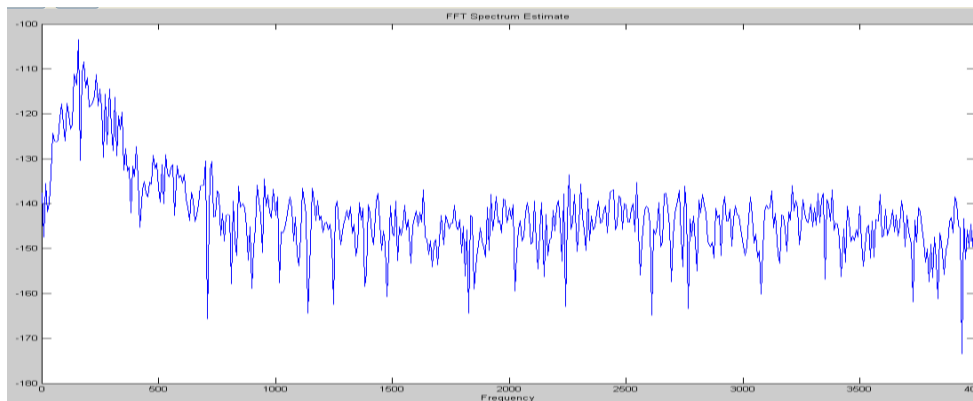
This section presents the results of effect of coding algorithms G.711, G.729 and AMR – wideband, on VoIP signal processed under the influence of various network degradation parameters like delay, packet loss. Figure 3. (a) Shows the spectrum analysis of VoIP signal for G.729 and AMR – WB codec under the influence of noise.



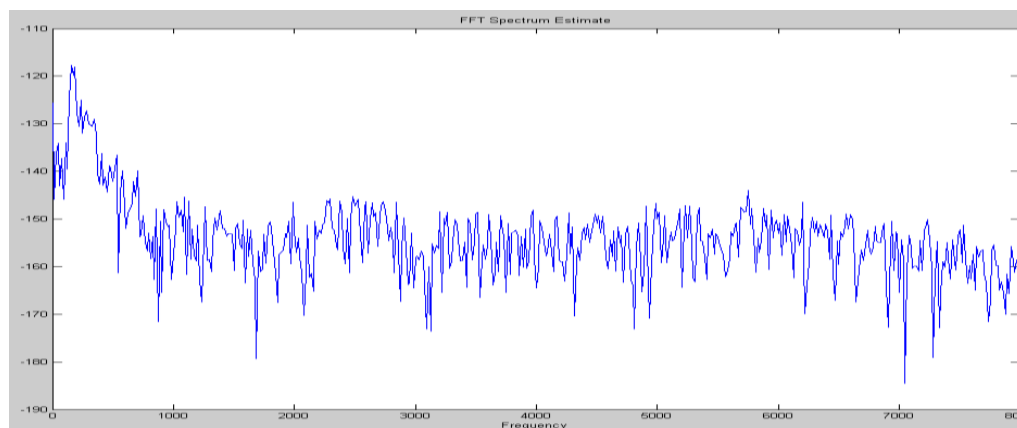
(a)



(b)



(c)



(d)

Figure 3 (a) Analysis for G.729 and AMR Coder [5], (b) FFT analysis for G.711 Coder, (c) Analysis for G.729 coder, (d) FFT analysis for AMR coder.

Figure 3. (b), (c) and (d) shows the analysis of the VoIP signal for G.711, G.729 and AMR-WB codecs for different delay levels above the threshold value. The analysis of degraded VoIP signal indicates that the signal quality can be improved by using efficient coding algorithm for VoIP signal. The MOS score for the output of VoIP system using AMR-wideband coder is approximately above 4.0, which is very well accepted for signal quality.

5. CONCLUSION

The paper presents the overview of mostly used codecs for VoIP communication. The study is useful to choose the coding algorithms for adaptive voice processing over internet protocol. The network impairments affects the speech signal. In case of additional delay, noise is observed in the waveform with the VoIP speech signal. The enhancement in the speech quality has been also observed through the simulations conducted in this work, which further suggests that the AMR codecs gives better results than G.711 and G.729 coders, as the bandwidth for telephony is extended from 4 kHz to 7 kHz. The analysis of the simulated results for VoIP signal with different speech processing algorithms clearly indicate the significant reduction in degradation when VoIP signal is passed through AMR coders. In future, the study can be useful for improving the speech quality using various signal processing algorithms performed at higher frequency digital signal processors.

6. ACKNOWLEDGEMENT

This work is dedicated to Late Prof. Jasvir Singh, Department of Electronics Technology, G.N.D. University, Amritsar, for

guidance & motivation without whom it was impossible to carry out the work. Authors are especially thankful to Dr. Suyeab A. Khan, Shiv Shankar Institute of Engineering & Technology, Patti and Dr. Harjit Pal Singh, CT Institution of Engineering, Management and Technology, Jalandhar, for providing the technical help and guidance to perceive this work.

7. REFERENCES

- [1]. Sulovic, M., Raca, D., Hadzialic M., and Hadziahmetovic, N. 2011 Dynamic Codec Selection Algorithm for VoIP. In The Sixth International Conference on Digital Telecommunications, IARIA Press (Budapest, Hungary). 74–79.
- [2]. Ramakrishnan, R.S., and Kumar, P.V. 2008 Performance Analysis of Different Codecs in VoIP Using SIP. In Mobile and Pervasive Computing, Allied Publishers, (Chennai, India).142-145.
- [3]. Johnson, M. H., and Alwan, A. 2003 Speech coding: fundamentals and applications.
- [4]. Rodman, J. 2008 VoIP to 20 kHz: Codec Choices for High Definition Voice Telephony. White paper.
- [5]. Singh, H.P., Singh, S., and Singh, J. 2011 Comparison of Narrowband and Wideband VoIP using TMS320C6713 DSP Processor. In International Symposium on Devices MEMS, Intelligent Systems & Communication. 25-29.

- [6]. A high quality low complexity algorithm for packet loss concealment with G.711. 1999 ITU, Geneva, Switzerland. Rec. ITU-T G.711, App 1.
- [7]. Mehta, P., and Udani, S. 2001 Voice over IP Sounding good on internet. IEEE Potentials, 36-40.
- [8]. ITU, ITU-T G.729: CS-ACELP speech coding at 8 kbit/s, ITU 1998.
- [9]. 3GPP TS 26.171, "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description", 2008-12.
- [10]. Bessette, B., Salami, R., Lefebvre, R., and Jelinek, M., 2002 The adaptive multirate wideband speech codec (AMR-WB) IEEE transactions on speech and audio processing. **10** (8), 620-636.
- [11]. Markopoulou, A., Tobagi, F., and Karam, M. 2002 Assessment of VoIP quality over Internet backbones. In Proc. of IEEE INFOCOM, (New York, USA).
- [12]. ITU-T Recommendation, 2001, P.862: Perceptual evaluation of speech quality (PESQ); an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codec.
- [13]. Proakis, J. D., and Manolakis, D. G. 2000 Digital Signal Processing: Principles, algorithm and applications, Third edition.