Journal of Interdisciplinary and Multidisciplinary Research (JIMR)
E-ISSN:1936-6264| Impact Factor: 8.886| UGC CARE II
Vol. 20 Issue 12, Dec- 2025
Available online at: https://livejimrjournal.in/
(An open access scholarly, peer-reviewed, interdisciplinary, monthly, and fully refereed journal.)

Real – Time Analysis of various VoIP Coding Algorithms

Hardeep Singh[1*], Dhruva Kumar[2]
[1]*Department of Physics, Guru Nanak College, Budhlada 151502, India*
[2]*Department of Chemistry, Guru Nanak College, Budhlada 151502, India*
* *hardeep.phy@gmail.com, ohm23dhruva@gmail.com*

Abstract:

Voice over Internet Protocol (VoIP) is a method of providing phone services over dedicated public IP networks. Maximum people of the world are using mobiles to communicate with each other using internet telephony and Mobile VoIP is one of the fastest-growing trends. With the extensive availability of 5G networks, mobile devices will become even more essential for business communication. The Speech quality, is critically important for the users of VoIP telephony, Signal quality is degraded by various problems, which include delay, packet loss and jitter. The implementation of degraded signal through various coding algorithms and digital signal processor can improve the quality of VoIP signal. The work deals with comparative study and analysis of VoIP Codecs (G.711, G.729, AMR, AMR-WB etc.) in real time environment. The study analyzes the performance of various codecs for VoIP, using MATLAB. The VoIP simulations are conducted for G.711, G.729A and AMR-WB speech coders for different network conditions. The results are validated through the perceptual evaluation of speech quality (PESQ) measurement.

Keywords: VoIP, MOS, G.729, AMR, PESQ.

1. Introduction:

These days, communication via the Internet (VoIP) is the most widely used technology used for the transmission of speech signals over packet-switched IP networks. For a phone call using VoIP, an analog input signal is converted into a 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 specific codec for proper encoding and decoding [1, 2]. The bandwidth requirements of speech transmission over IP networks are relatively lower than those of PSTN lines. It takes only 6–8 kbps or less of bandwidth, whereas PSTN takes 64 kbps of bandwidth to make a call. But providing QoS is one of the problems in the implementation of voice-over IP. The quality of voice is a characteristic of the IP network, and it is affected by parameters such as delay, jitter, packet loss, and network congestion. The effect of these parameters results in harmful effects on the quality of VoIP.

Speech coding describes the conversion of analog voice to digital form and then compressing these digitized samples to reduce the consumption of network bandwidth required to transmit the speech of audio signals containing speech. It uses speech-specific parameter estimation using audio signal processing techniques to model the speech signal and represent the modeled parameters of the signal in a compact bit stream of packets. A speech decoder receives coded frames and produces a reconstructed speech signal at the receiver. Standards typically dictate the input-output relationships of both the coder and the decoder. Speech coding has two most important applications: mobile telephony and voice-over-IP. [3] [11]

2. Speech Codecs:

Speech codecs differ primarily in bit rate (measured in bits per sample or bits per second), complexity (measured in operations per second), delay (measured in milliseconds between recording and playback), 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 common than WB coding mainly because of the narrowband nature of the wire line telephone channel (300–3600 Hz). Recently, there has been an increased effort in wideband speech coding because of several applications such as videoconferencing.

G.711

G.711 is a codec that was introduced by the International Telecommunication Union (ITU) in 1972 for use in digital telephony. The G.711 describes a simple way to digitize analog data by using a semi-logarithmic scale called the computed pulse code modulation (PCM). It has two variants based on the type of coding algorithm: A-law and Mu-law. It uses logarithmic compression. Its goal is to increase the resolution of small signals, while large signals are treated proportionally. The encoded stream is 64 kbps, consisting of 8 kHz sampling of 8-bit signals. The frame length is eight 125 µs samples, or 1 ms. Using G.711 for VoIP gives the best voice quality. It also has the lowest latency because there is no need for compression, which costs processing power. Its benefits are simple implementation and very good perceived audio quality, having a MOS value of 4.1. The problem with this codec is that it takes more bandwidth than other codecs, up to 84 kbps, including all TCP/IP overhead. However, with increasing broadband bandwidth, this should not be a problem. G.711 was the most widely used codec for communication. [3-4]

G.729

G.729 is an audio data compression algorithm that compresses digital voice into packets of 10 millisecond duration. It is officially described as the coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP). [1] Because of its low bandwidth requirements, G.729 is mostly used in voice over Internet Protocol (VoIP) applications (such as Skype), where bandwidth must be conserved. Standard G.729 operates at a bit rate of 8 kbit/s, but there are extensions that provide rates of 6.4 kbit/s and 11.8 kbit/s for marginally worse and better speech quality, respectively. G.729 has been drawn-out with various features, commonly designated as G.729a and G.729b. Fax transmissions, DTMF tones and high-quality audio cannot be transported reliably with this codec. DTMF requires the use of the RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals as specified in RFC 2833. [1-5]

Adaptive Multi-Rate (AMR-WB)

The 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. AMR-WB codec works on the principle of Algebraic Code Excited Linear Prediction (ACELP) for all bit rates. To reduce average bit rate, this codec supports the discontinuous transmission (DTX), using Voice Activity Detection (VAD) and Comfort Noise Generation (CNG) algorithms. The coder works on a frame of 320 speech samples (20 msec), and a look ahead of 5 msec is required. So the algorithmic delay for the coder is 25 msec. AMR-WB provides excellent speech quality due to a wider speech bandwidth of 50–7000 Hz compared to narrowband speech coders which in general are optimized for POTS wireline quality of 300–3400 Hz. AMR-WB is codified as G.722.2 by ITU-T, formally known as Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB). The corresponding 3GPP specifications are TS 26.190 for the speech codec and TS 26.194 for the Voice Activity Detector. [1-3,10] A common file extension for AMR-WB file format is. awb. This format is the 3GPP-specified 3GP container format based on ISO base media file format. [4,6-9]

Comparative analysis of various codecs:

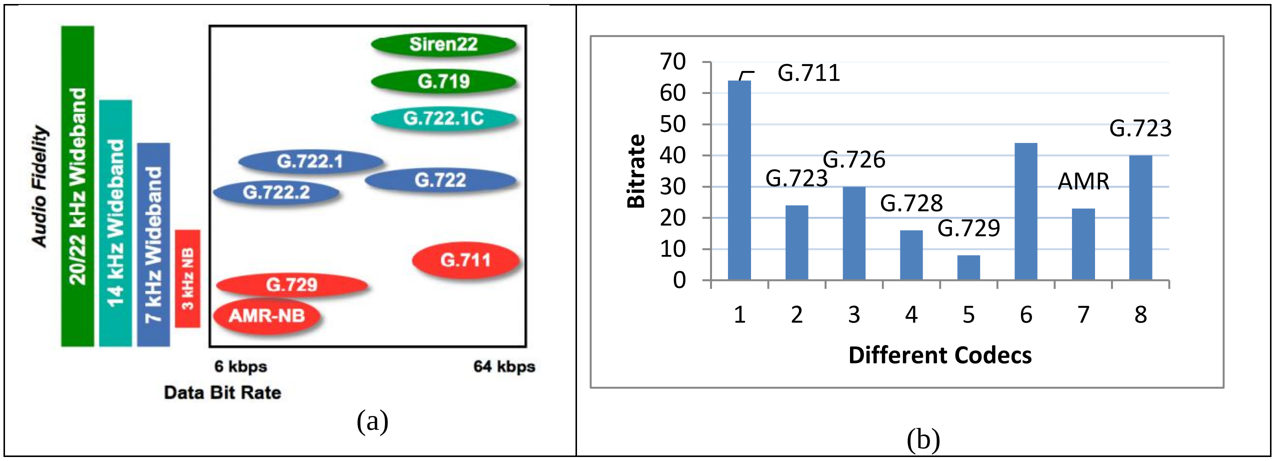


Figure 1. To comparison of different VoIP codecs. (a) Audio bandwidth versus bit rate for different codecs [4], (b) Bitrate requirements for various Speech Codecs.

The fundamental concern for VoIP QoS is the quality of the received voice. One way to measure customers' observation about the quality of these systems, is to conduct a subjective test involving panels of human subjects. However, these tests are expensive and unsuitable for real-time monitoring applications. 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. Table1 shows the correlation between MOS scores and the listening quality [10-11].

MOS	Quality	impairment
5	Excellent	Imperceptible
4	Good	Imperfections can be perceived
3	Fair	Slightly Annoying
2	Bad	Annoying
1	Poor	Very Annoying

Table. 1 Mean opinion score

Here table 2 shows the comparatative analysis of various codecs on the basis of coding algorithm, bit rates, sampling rate and Mean opinion score.

Coder	Type	Rate [kbps]	Packetization period [ms]	Frame size [ms]	Algorithmic delay [ms]	Codec delay [ms]
G.711	PCM	64	20	0.125	0	0.125
G.723.1	MPC-MLQ	5.33	30	30	7.5	37.5
G.723.1	ACELP	6.4	30	30	7.5	37.5
G.726	ADPCM	32	20	10	0	10
G.728	LD-CELP	16	30	0.625	0	0.625
G.729A	CS-ACELP	8	20	10	5	15
AMR	ACELP, DTX, VAD & CNG	4.75 to 12.2 kbit/s	20	8-20	20ms per frame	4.14

Table. 2 Specifications of Different Codecs

4. Modeling and Simulation of VoIP System

In Voice over Internet Protocol (VoIP) the voice signal is processed through the IP network, the quality of the VoIP signal is degraded by delay, delay variation (jitter) and packet loss. To analyze the effect of different codecs, a VoIP system had been designed. The real-time VoIP environment has been created using PC to PC communication approach. [7]

Basic Steps in 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.

Step 3: The compressed version of the signal is then packetized into packets to 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.

The experimental set-up used for the real time voice quality evaluation for VoIP system is shown in the figure 2. It is a software based telephone system for which the software has been installed on PC that allows the simulation of key processes in voice over IP and speech quality

Journal of Interdisciplinary and Multidisciplinary Research (JIMR)
E-ISSN:1936-6264| Impact Factor: 8.886| UGC CARE II
Vol. 20 Issue 12, Dec- 2025
Available online at: https://livejimrjournal.in/
(An open access scholarly, peer-reviewed, interdisciplinary, monthly, and fully refereed journal.)

measurement. The Linphone, x-Lite, 3CX, Skype softwares were tried for VoIP traffic analysis. The recorded calls details have been utilized to measurement of voice quality of the VoIP system. Objective voice quality measurements were made with the ITU-T recommendation P.862 for Perceptual Evaluation of Speech Quality (PESQ) (ITU-T, 2001). [8-11]

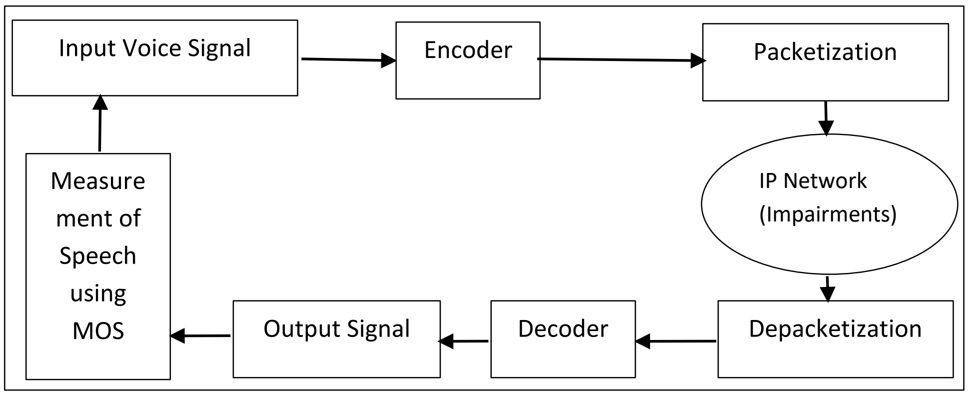
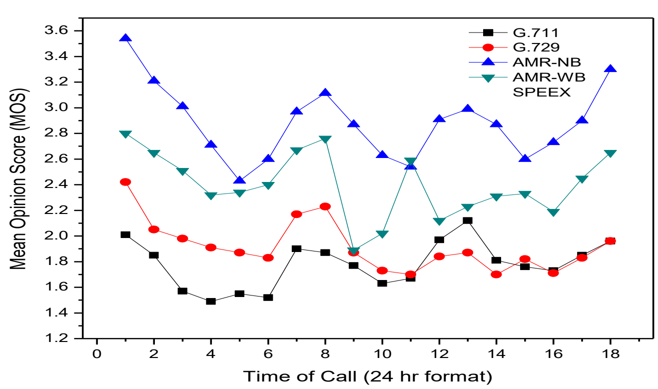


Figure 2. Design of Real - time VoIP system

It has been observed from the results obtained that during the start of the day the network bandwidth is available to large extent, this time period is termed as 'normal hours' (7:00 to 10:00hrs) and the MOS score obtained ranged between 2.4 to 3.3 for different codecs. MOS values decreased during the 'peak hours' (10:00 to 13:00hrs) when most of bandwidth have been utilized and during middle of the day, it's the break time here we are using 'Average hours' (13:00 to 15:00hrs) slight increase in the MOS score has been recorded. The quality of the received call slightly improves for all the codecs. For the time period 15:00 hrs onward people were back on work and network got congested till the 17:00 hours which is off time for most of the offices. Here the figure 3 shows the variation of MOS for the different codecs for the VoIP traffic during the different hours.



Conclusion:

Today people use cell phone for all the purposes. It overtakes traditional phone technology. The codec role is very important to maintain quality. Some of the quality codecs features, bandwidth, transmission rate, compression rates are discussed in present paper. The analysis of MOS observed through the simulations conducted in this work, it has been analyzed that the AMR codecs gives better results than G.711 and G.729 coders, since the telephony bandwidth is extended from 4 kHz to 7 kHz. At present time most of the VoIP softwares such as Skype etc. [9-10], are also updating their services through various coders, to improve the quality. In future the study can be useful for improving the speech quality using various signal processing algorithms performed at higher frequency digital signal processors. [11]

Acknowledgments

This work is dedicated to Late Prof. Jasvir Singh, Department of Electronics Technology, G.N.D. University, Amritsar, without his guidance it was impossible to carry out this work. We special thanks to Dr. Suyeb A. Khan and Dr. Harjit Pal Singh, for providing us the technical help to carry out this work.

References

1. Sulovic, M., Raca, D., Hadzialic, M. & Hadziahmetovic, N. (2011) Dynamic codec selection algorithm for VoIP. In: *The Sixth International Conference on Digital Telecommunications*. IARIA Press: Budapest, pp. 74–79s
2. Ramakrishnan, R.S. & Kumar, P.V. Performance analysis of different codecs in VoIP using SIP. In: *Mobile and Pervasive Computing (CoMPC–2008)*. Allied Publishers: Chennai, India, pp. 142–145.
3. Hasegawa-Johnson, M. & 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. Sat, B. & Wah, B.W. (2007) Evaluation of conversational voice quality of the Skype, Google-Talk, Windows Live, and Yahoo! Messenger VoIP systems. In: *IEEE Publications Int.'l Workshop on Multimedia Signal Processing.*
6. Markopoulou, A.P., Tobagi, F.A. & Karam, M.J. (2002) Assessment of VoIP quality over Internet backbones. In: *Proceedings of the of IEEE Infocom*, New York, pp. 150–159.
7. Mario Di Mauro. (2022). Multivariate Time Series Characterization and Forecasting of VoIP Traffic in Real Mobile Networks. *IEEE Dataport*.
8. W. Chiang (2006) A performance study of voip applications: Msn vs. skype. *In Proc. of MULTICOMM.* pp. 13-18.
9. Zampognaro, F., Aurigemma, R., & Munarini, W. (2021, December). VoIP CODEC assessment and performance evaluation in satellite-based scenarios. In *2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)* (pp. 01-06). IEEE.
10. Maheswari, K., Balamurugan, A., Charlyn Pushpa Latha, G., & Ramkumar, S. (2021). Performance analysis of VoIP codecs in interactive streaming data environment. *Materials Today: Proceedings.* pp. 1-4.