Real Time Analysis of VoIP System under Pervasive Environment through Spectral Parameters

Harjit Pal Singh
Department of Physics
Dr.B.R.Ambedkar National
Institute of Technology
Jalandhar, India

Sarabjeet Singh
Department of Physics
Dr. B.R.Ambedkar National
Institute of Technology
Jalandhar, India

Jasvir Singh
Department of Electronics
Technology
Guru Nanak Dev University
Amritsar, India

ABSTRACT

In Voice over Internet Protocol (VoIP) system, the speech signal is degraded when passed through the network layers, since the best effort policy based IP network leads to the network degradations including delay, packet loss jitter. The transmission of voice over internet protocols requires signal processing for reliable & efficient performance over the communication platform. The work in this paper presents the effect network degradation factors on VoIP system and the spectral analysis of the VoIP signal. The lab experiment is performed to realize the VoIP system and to obtain the degraded database. The spectral analysis of VoIP signal is performed through the various signal processing algorithms. The results are validated through the quality evaluation of the VoIP signal using perceptual evaluation of speech quality (PESQ) measurement for narrowband signal.

Keywords

VoIP, WANem, Packet loss, Jitter, Delay, Digital signal processing, Spectral analysis

1. INTRODUCTION

Voice over IP is an advancing technology that is used to transmit voice over the internet or a local area network using internet protocol (IP) [1].This technology provides enhanced features such as low cost compared to the traditional Public Switched Telephone Network (PSTN). VoIP system costs as much as half the traditional PSTN system in the field of voice transmission. This is because of the efficient use of bandwidth requiring fewer long-distance trunks between switches [2].Packet switched networks like Internet, are based on the Best-effort policy which does not guarantee a minimum packet loss rate and a minimum delay of packet transmission required for VoIP system. Ferrari et.al [3, 4] had published work related to the effect of jitter and delay on real-time communication applications. Bolot et.al [5-7] also analyzed the performance of audio packets on internet for the various network characteristics. Hassan et.al [8] had analyzed the Internet telephony services for technical challenges and compared commercial products supporting the IP telephony. Karam et.al analyzed the bandwidth utilization of voice traffic over internet [9]. A performance progress of VoIP was issued in by James et.al [10]. The work in this paper presents the performance analysis of the narrowband VoIP system under varying network conditions and the signal quality is analyzed through various speech processing algorithms.

The Section II presents the brief description of the VoIP system for this work. The major degradation factors are discussed in Section III. The realization of the VoIP system in lab is presented in Section IV. The brief description of the spectral analysis algorithms is presented in Section V. The performance evaluation results and discussion are presented in Section VI. The last section concludes the work and presents the scope of the future work.

2. VoIP SYSTEM

The voice over internet protocol system is becoming the successful alternative to the traditional PSTN communication system due to its advanced features. The voice signal is processed through the internet based network during the communication. In the present work, the VoIP system is realized through the lab experiment and the experiment is performed at various network conditions [11]. The conceptual diagram of VoIP system is presented in Fig.1. The basic steps in derivation of the designed VoIP system are:

Step 1: The original speech signal is fed in to the system and the speech samples are taken from [12].

Step2: The speech signal is then encoded with G.711a and Speex speech encoders, which is the compressed version of the input signal. G.711a is the standard used for the communication purpose and is a high bit rate Pulse Code Modulation codec. It works at sampling rate of 8 kHz and uses and compresses the 16 bit audio samples into 8bits [13]. The Code Excited Linear Prediction (CELP) Speex codec is an open source codec developed for the packet network and VoIP applications [14]. The Speex supports three different sampling rates narrowband (8 kHz), wideband (16 kHz) and ultra-wideband (32 kHz).

Step3: The compressed signal is then packetized into VoIP packets to transfer it to the IP network.

Step4: The speech signal is degraded due to the various network impairments including delay, jitter and packet loss during VoIP communications. The network impairments are introduced through the lab WANem emulator [15]. The whole lab experiment is described in Section 4.

Step5: The degraded VoIP signal is depacketized and then decoded with G.711a and Speex decoders.

Step6: The performance is evaluated with Perceptual Evaluation of Speech Quality (PESQ) measurement defined by ITU-T recommendation P.862 [16]. After comparing the degraded signal with the original one, the PESQ measurement gives the

subjective measurement as Mean Opinion Scores (MOS) value from -0.5 to 4.5.

Step7: The VoIP signal is processed through various signal processing algorithms to evaluate the performance of the system.
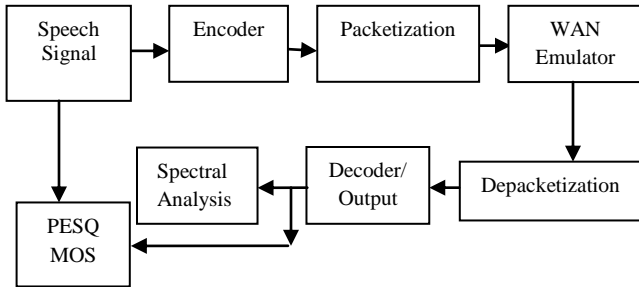
Fig 1: Conceptual diagram of VoIP System

3. FACTOR AFFECTING VoIP SIGNAL QUALITY

Voice quality in communication systems is influenced by many factors such as packet delay, jitter, packet loss and type & amount of voice compression. Due to these distortion factors, the speech signal is not of very good quality over the VoIP network. Delay is the time taken by the voice to reach from talker's mouth to the listener's ear. Round trip delay is the sum of two one-way delays that occur in the user's call. In VoIP system, the propagation delay is also affected by two additional delays such as packeting delay and the time required for propagating the packet through the network. This varies the propagation delay during the transmission. According to the ITU-T the recommended one way delay by ITU-T is 0-150 ms and this amount of delay is acceptable for most of the user applications [17]. But latency greater than 400 ms is unacceptable.
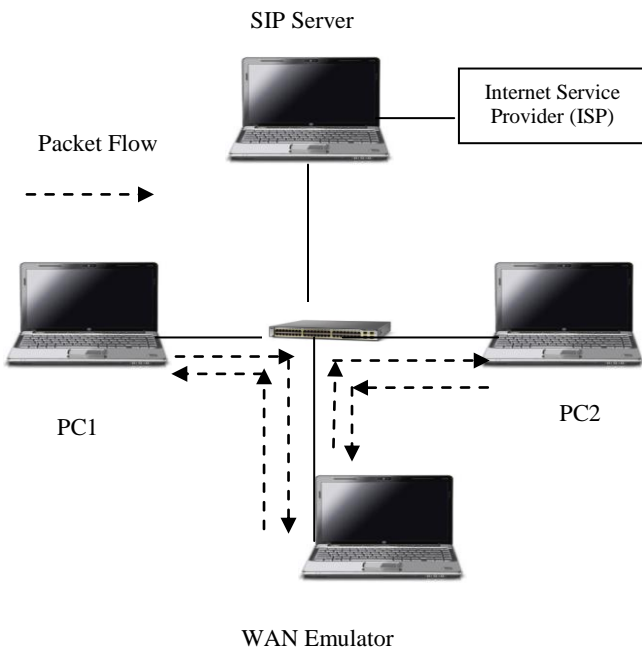

Fig 2: Lab Setup

The variation in the arrival time of the packets at the receiver end leads to jitter, which affects the perceived quality of conversation very badly. The sender is expected to transmit each voice packet at a regular interval. But jitter affects the speech in such a way that all voice packets do not arrive at the right time at the decoder and thus reconstructed speech would not be continuous, at the receiver end. The transmission time of a packet through IP network varies due to queuing effect in the interconnected network [18]. The packet loss is the percentage of the lost packets during the transportation due to various network conditions such as buffer overflow, network congestion etc. The delay and jitter also contribute to the packet losses and these results in harmful effects on the quality of VoIP signal. Due to the real time requirement for interactive speech transmission, it is usually impossible for the receivers to request the sender to retransmit the lost packets. When voice packets do not arrive before their playout time, they are considered as lost and cannot be played when they are received. Even a single lost packet may generate audible distortion in the decoded speech signal [19, 20].

4. VoIP LAB EXPERIEMNT

The lab experiment is performed to realize the VoIP system. The signal is degraded with various network impairments including delay, jitter and packet loss during VoIP communications. The different values 50, 100, 150 and 200 ms for delay and 5, 7, 10 and 15 ms for jitter are used in the performed lab experiment. The different values of the packet loss rates 0% to 10 % for have been used to realize the network impairments in this experiment. The lab setup is shown in Fig.2. The WANem emulation software, which is the product of Tata Consultancy, has been used to emulate the WAN traffic [15]. The schematic diagram of the WANem is presented in Fig.3. The X-Lite IP softphone [21] is used to establish VoIP calls between two computers. To provide the VoIP services, 3CX phone system [22] has been used as the SIP server.
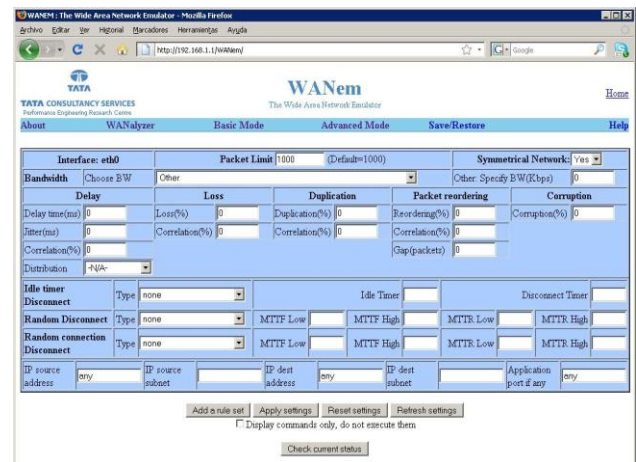
Fig.3 WANem Emulator

5. SPECTRAL ANALYSIS OF VoIP SIGNAL

To analyze the effect of packet loss on the quality of the degraded VoIP output, the spectral analysis was performed in time and frequency both. The various signal processing algorithms used for spectral analysis are discussed here:

5.1 Power Density Spectrum

The estimation of power spectral density (PSD) is carried out by dividing the time signal into successive blocks and then, averaging the squared-magnitude discrete Fourier transforms (DFTs) of the signal blocks. The Welch power spectral density estimate is given as [23],

$$P_x(f) = \frac{1}{N} |DFT(x_n)|^2$$

$$P_x(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi fn} \right|^2 \tag{1}$$

Where $x(n)$ is the input time domain signal and $P_x(f)$ is the power density spectrum function.

5.2 Spectrogram

The spectrogram of the speech signal is an intensity plot of short time Fourier transforms (STFT) magnitude. STFT is a sequence of FFTs of windowed data segments where the windows are usually allowed to overlap in time. The data to be broken up into frames, which usually overlap to reduce artifacts at the boundary and then each frame, is Fourier transformed using *K*-point FFT [24]. The STFT procedure is shown in Fig 4.

The short time Fourier transform (STFT) is expressed as:

$$S(\omega_k, l) = \sum_{n=0}^{N-1} x[n+l] w[n] e^{-j\omega_k n} \tag{2}$$

Where $\omega_k = \frac{2\pi k}{NT}, k = 0 \cdots K-1$ at given sampling frequency f_s, $x(n)$ is the time domain signal and $w(n)$ is the window function.

5.3 FFT Analysis

A fast Fourier transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) and it's inverse. A DFT decomposes a sequence of values into components of different frequencies and computing it directly from the definition is often too slow to be practical [25]. The DFT of the signal is defined as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k \frac{n}{N}}, k = 0 \ldots N-1 \tag{4}$$

An FFT is a way to compute the same result more quickly. The computational complexity reduces to $N \cdot \log(N)$ as compared to $O(N^2)$.

6. RESULTS

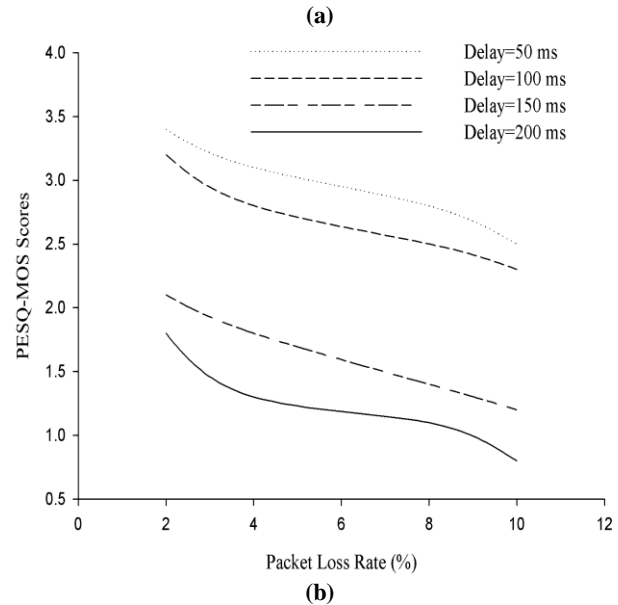The various performance analysis and spectral analysis results of the real time VoIP system are presented here:
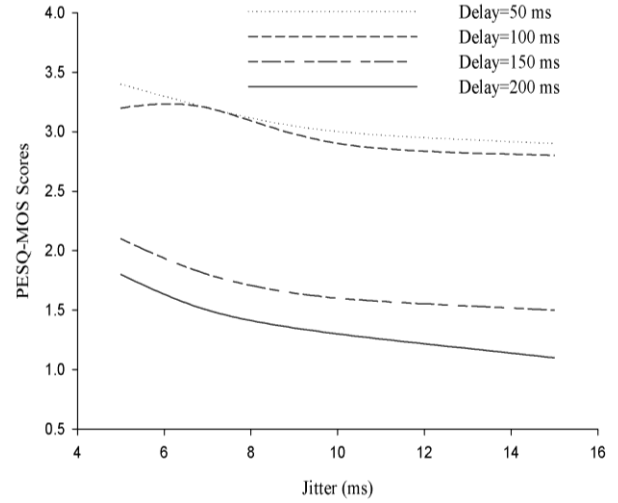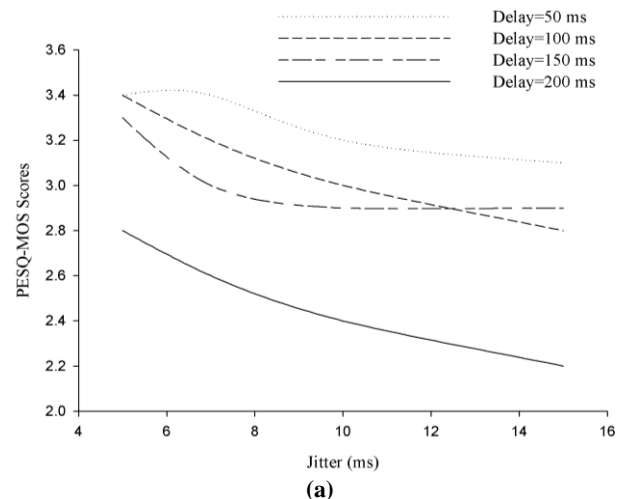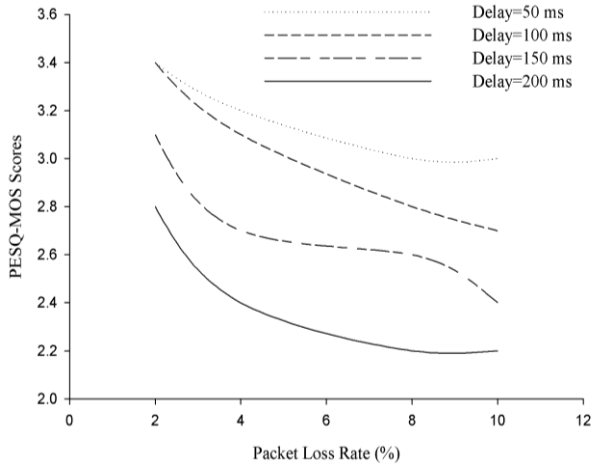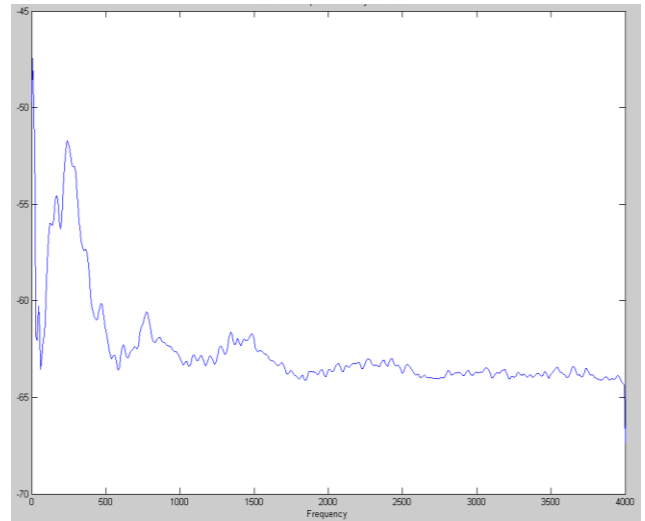

(a)

(b)

Fig.4 Performance analysis of G.711 at varying delay values for (a) Jitter (b) Packet Loss
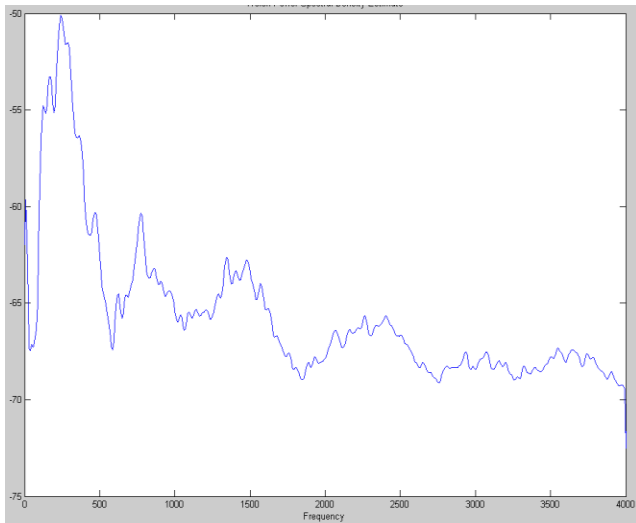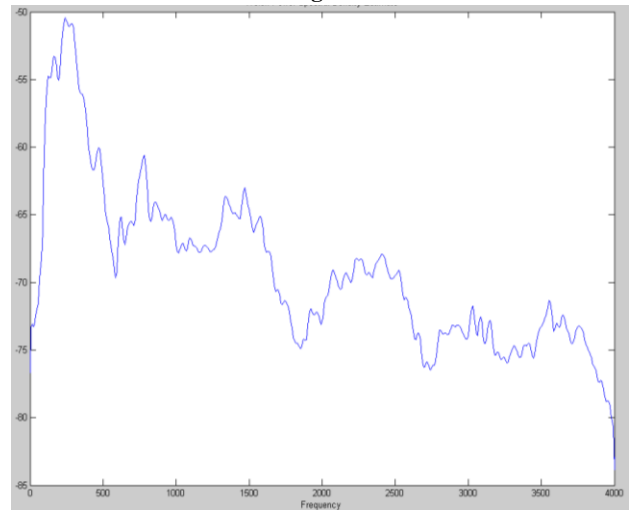

(a)

(b)

Fig.5 Performance analysis of Speex at varying delay values for (a) Jitter (b) Packet Loss
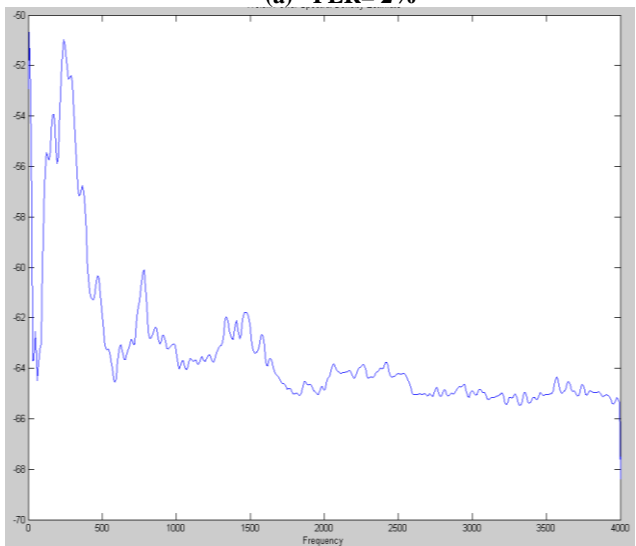

(c)   PLR= 10 %

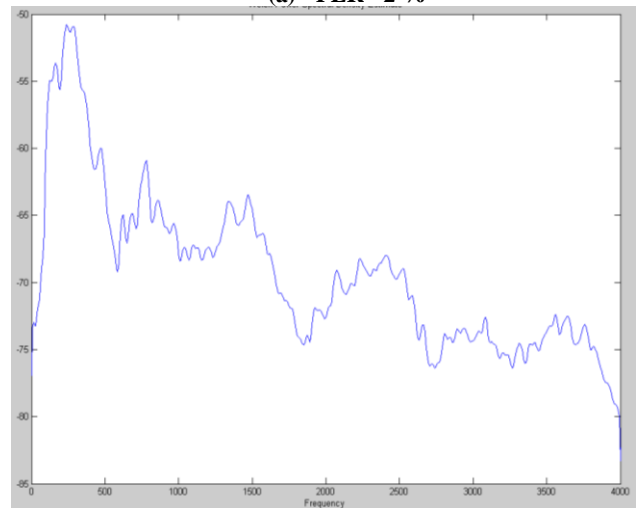**Fig.6 Power Spectral Density analysis for G.711a based VoIP Signal**
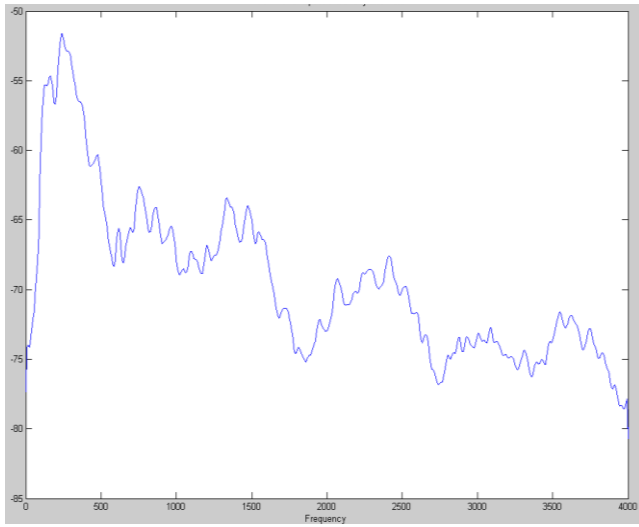

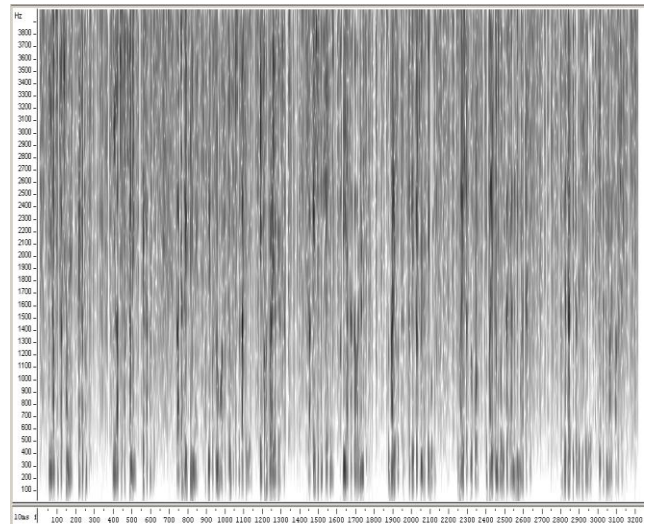**(a)   PLR= 2%**

(a)   PLR= 2 %

(b)   PLR= 8%
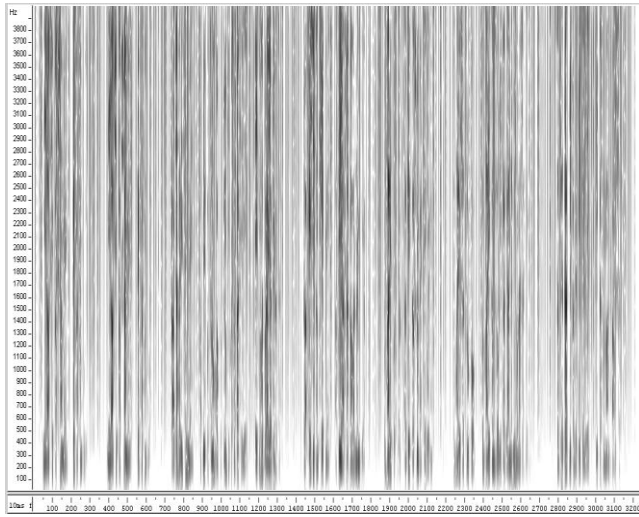

(b)   PLR= 8%

(c)   PLR= 10 %
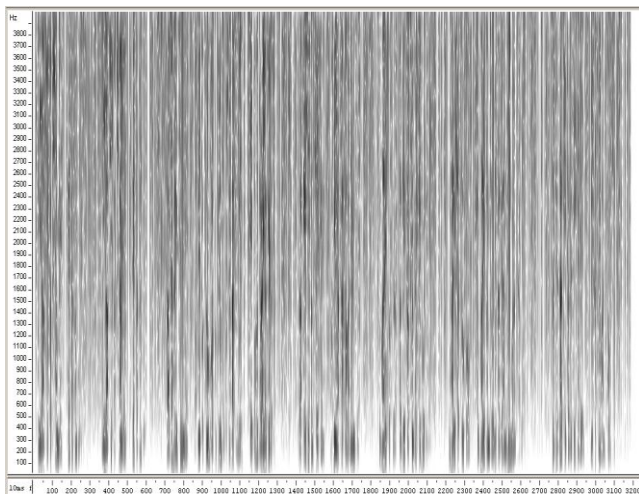**Fig.7 Power Spectral Density analysis for Speex based VoIP Signal**

(c)   PLR= 10 %
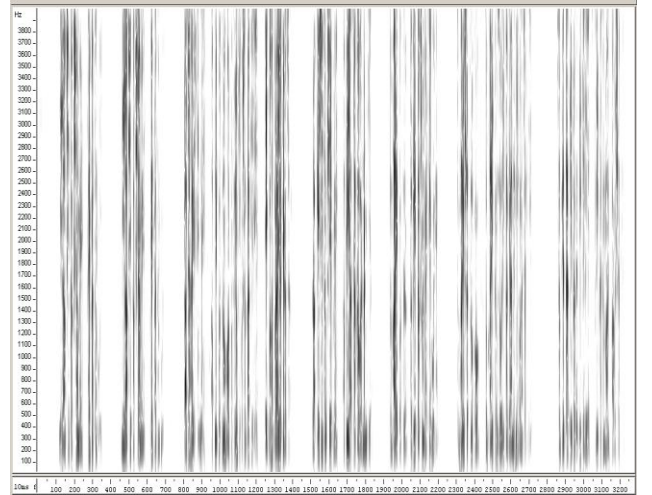**Fig.8 Spectrogram analysis for G.711a based VoIP Signal**

(a)   PLR= 2 %

(a)   PLR= 2%

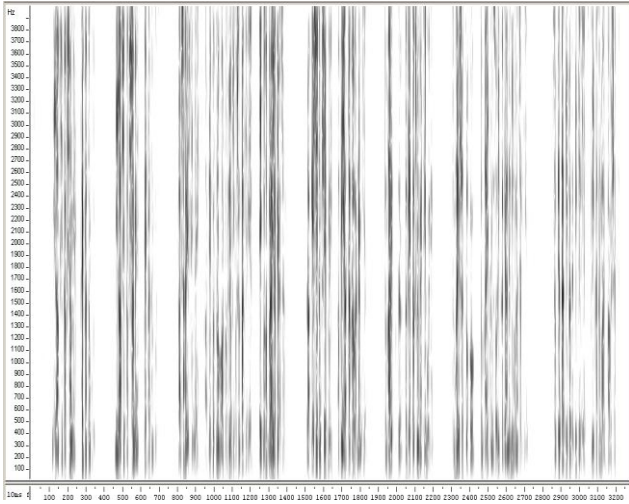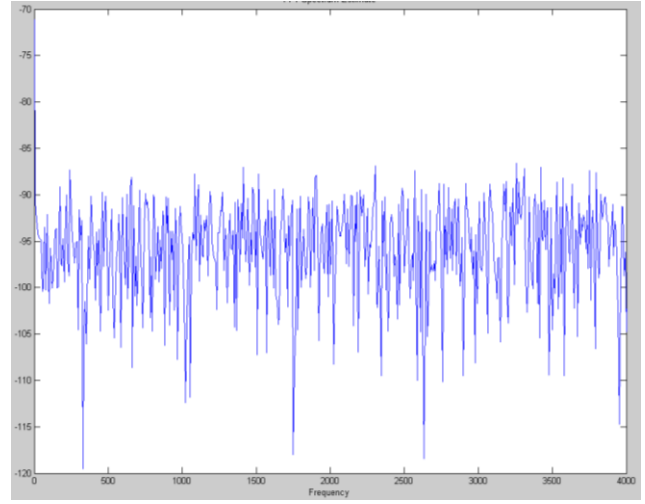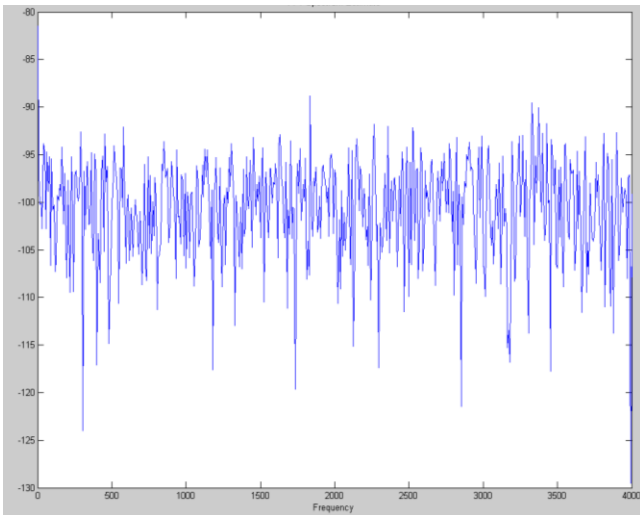**(b)   PLR= 8 %**

(b)   PLR= 8%

(c)   PLR= 10 %
Fig.9 Spectrogram analysis for Speex based VoIP Signal
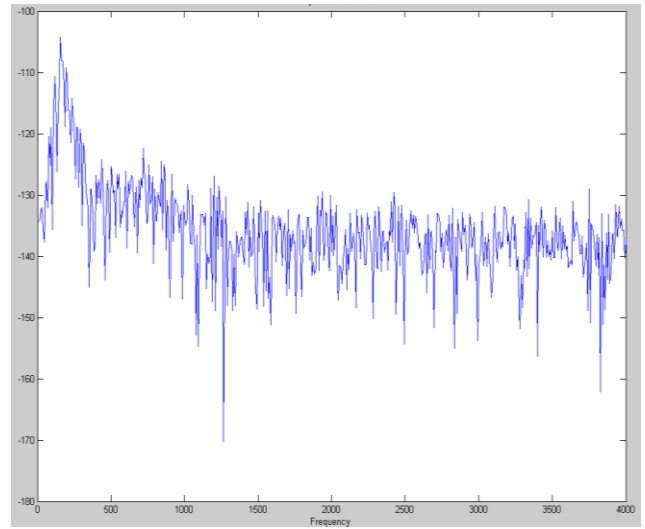

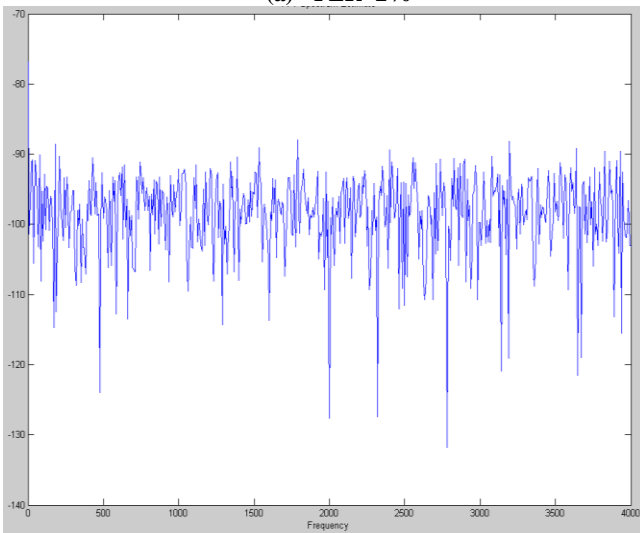
(c)   PLR= 10 %
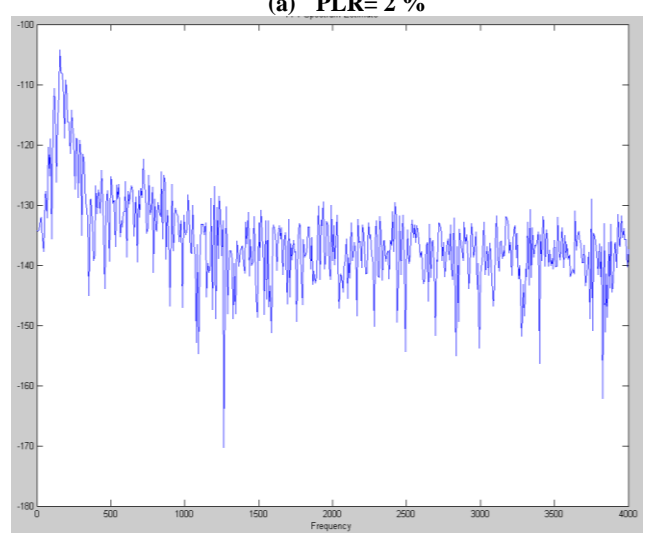**Fig.10 FFT Spectrum for G.711a based VoIP Signal**
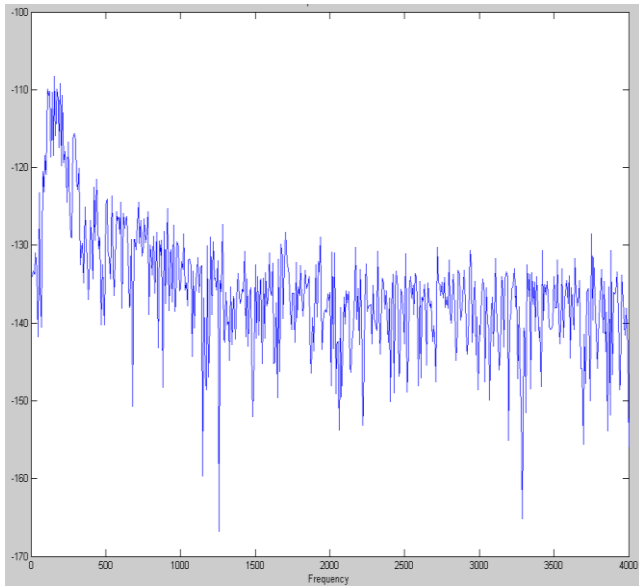

(a)   PLR=2%

(a)   PLR= 2 %

(b)   PLR= 8%

(b)   PLR= 8%

(c) PLR= 10 %

Fig.11 FFT Spectrum for Speex based VoIP Signal

6.1 Discussion

The results show that the network impairments affect adversely the speech signal. The additional noise is observed in the waveform of the VoIP speech signal when passed through IP network. The effect of delay, jitter and packet loss has been studied in this work through the lab experiment. As the delay in the system increases, jitter and packet loss increases in the system. The effect of jitter and packet loss on the perceived quality of VoIP signal at various values of the delay for G.711a codec is presented in Fig.4. At higher value of the delay, the jitter and packet loss also increases & burst losses occur; the performance of the VoIP system with G.711a codec is decreased. The PESQ-MOS decreases rapidly as the delay increases from the recommended value as presented in Fig.4. The performance analysis results of VoIP system with Speex coder at varying jitter and packet loss rates are presented in Fig.5. The Speex coder is more robust to varying network conditions in VoIP system and the performance in much better than G.711a even at the worst network conditions. The MOS scores of the VoIP signal with Speex coder are in the recommended range, even at 10% of the packet loss, 15 ms of jitter and 150 ms of delay.

The VoIP signal is then further processed through the various speech processing algorithms to evaluate the effect of network degradations on the perceived VoIP speech signal. The spectral analysis of the VoIP signal, at fixed value of delay (150 ms) & jitter (15 ms) and varying packet loss rate, is performed through power spectral density (PSD), spectrogram, FFT algorithms. Fig. 6 and Fig.7 presents the power spectral density analysis of the VoIP signal coded with G.711a and Speex respectively. The PSD graph in Fig.6 indicates that the performance of the VoIP system and quality of signal with G.711a is decreased as the packet loss increases. The PSD graph in Fig.7 indicates that the Speex based narrowband VoIP system has much capability to capture the weak signal and provides the

better signal quality. The spectrogram result for G.711a in Fig.8 shows the effect of network distortions on the VoIP signal. As the packet loss increases, the signal becomes noisy & weaker and the signal quality decreases rapidly. The VoIP signal coded with Speex has the better signal quality in comparison to the G.711a as presented in Fig.9. The VoIP signal with Speex coder has the better signal quality even at the tested peak value of loss rate. The FFT analysis in Fig.10 shows much distortion in VoIP speech signal when the signal is processed through the G.711a as compared to the Speex coder, which is depicted in Fig.11 at all used packet loss rates.

7. CONCLUSION & FUTURE WORK

The real time implementation of the VoIP system had been presented through lab the experiment and the spectral analysis of the VoIP speech signal was carried out in this work to evaluate the performance of the VoIP system with different VoIP speech coders. The VoIP signal quality was adversely affected by packet losses, since with increase of jitter and delay, consecutive packet losses occurs, which decreased the MOS scores of the signal. The spectral analysis results are similar to that of the objective measurement results for the designed VoIP system. The Speex coder presents much better signal quality in VoIP system, since it has the better capability to conceal the lost packets during voice transmission. In comparison to Speex, G.711 coder could not provide the toll quality of speech signal in VoIP system. To improve the performance of the system, these degradation parameters should be properly addressed. In future, the study could be used for improving the VoIP speech signal quality using various signal processing and filtering algorithms performed at higher frequency digital signal processors.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] Goralski, J.W., Kolon, C.M.,2000, IP Telephony, McGraw-Hill.

[2] Varshney, U., Snow, A., Mcgivern, M., Howard, C., 2002, "Voice over IP", Communications of the ACM, Vol.45,No.1, pp:.89-96

[3] Ferrari.D.,Verma, D.C., 1990, "A scheme for real time channel establishment in wide-area networks", IEEE Journal on Selected Areas in Communications, Vol. 8 No.3, pp: 368-379

[4] Verma. D.C., Zhang.H.,Ferrari.D.,1991, "Delay jitter control for real-time communication in a packet switching network", In Proceedings of IEEE conference on Communications for Distributed Applications and System, pp:35 – 43

[5] Bolot.J.,1993, "Characterizing end-to-end packet delay and loss in the Internet", Journal of High Speed Networks, Vol. 2 No. 3, pp: 305-323

[6] Bolot.J., 1993, "End-to-end packet delay and loss behavior in the internet", ACM Symposium on Communications architectures, protocols and applications, pp:289-298

[7] Bolot.J., Crepin.H., Gracia. A.V., 1995, "Analysis of audio packet loss in the internet", In Proceedings of International workshop on network and operating system support for digital audio and video, Springer Lecturer Notes in Computer Science, pp:163-174

[8] Hassan.M., Nayandoro. A., Atiquzzaman. M., 2000, "Internet telephony: services, technical challenges and products", IEEE communication magazine, Volume 38, No.4, pp:96 – 103

[9] Karam. M.J., Tobagi. F.A.,2001, "Analysis of the delay and jitter of voice traffic over the internet", In Proceeding of twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 2, pp: 824-833

[10] James. J.H., Chen. B., Garrison. L., 2004, "Implementing VoIP: A voice transmission performance progress report", IEEE communication magazine, Vol. 42, No.7, pp: 36 – 41

[11] Radhakrishanan, K., Larijani, H.,2011, "Evaluating perceived voice quality on packet networks using different random neural network architectures", Performance Evaluation, Vol. 68, No.4, pp: 347-360.

[12] Open Speech Repository, http:// www.voiptroubleshooter.com/open_speech/

[13] ITU-T Recommendations G.711, 1993, Pulse code modulation (PCM) of voice frequencies

[14] Valin, J., 2007, The Speex codec manual version 1.2 Beta 3, Speex.org,

[15] TCS, Wanem v 2.0, 2008, http://wanem.sourceforge.net/.

[16] ITU-T Recommendation P.862, 2001, Perceptual evaluation of speech quality (PESQ); an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codec

[17] ITU-T Recommendations, G.114, 2003, One way transmission time

[18] Jelassi, S., Youssef, H., Hoene, C., Pujolle, G., 2009, "Parametric speech quality models for measuring the perceptual effect of the network delay jitter", In Proceedings of the IEEE 34th conference on Local Computer Networks, pp: 193-200.

[19] Wu, C.C., Chen, K.T., Huang, C.Y., Lei, C.L., 2009 , "An Empirical evaluation of VoIP playout buffer dimensioning in Skype, Google Talk and MSN messenger", In Proceeding of 19th International Workshop on Network and Operating System Support for Digital Audio and Video, pp-97-102.

[20] Singh, H.P., Singh, S., Singh, J., 2011, "Digital signal processing approach for performance improvement in voice over internet protocol (VoIP) network", International Journal of Information and Communication Technologies, Vol.4, No.1-2, pp: 85-90.

[21] COUNTERPATH, X-lite, 2010, http://www.counterpath.com/x-lite.html.

[22] 3CX, Software based PBX, 2010. http://www.3cx.com/

[23] Proakis, J.D., Manolakis, D.G., 2000, Digital Signal Processing: Principles, algorithm and applications, Third edition

[24] Quatieri, T F, 2002, Discrete-Time Speech Signal Processing: Principle and Practice", MIT Lincoln Laboratory, Liexington, Massachusetts, Prentice Hall

[25] Mitra, S.K., 2001, Digital Signal Processing: A Computer-Based Approach: With DSP Laboratory using MATLAB, Mc.Graw Hill