Performance Analysis of Interleaving Scheme in Wideband VoIP System under Different Strategic Conditions

Harjit Pal Singh[1], Sarabjeet Singh[1] and Jasvir Singh[2]

[1]Dept. of Physics, Dr. B.R. Ambedkar National Institute of Technology, Jalandhar, India

[2]Dept. of Electronics Technology, Guru Nanak Dev University, Amritsar, India

harjit_nit@yahoo.co.in

Abstract

In Voice over Internet Protocol (VoIP) system, the speech signal is degraded when passed through the network layers. The speech signal is processed through the best effort policy based IP network, which leads to the network degradations including delay, packet loss and jitter. The packet loss is the major issue of the degradation in the VoIP signal quality; even a single lost packet may generate audible distortion in the decoded speech signal. In addition to these network degradations, the quality of the speech signal is also affected by the environmental noises and coder distortions. The signal quality of the VoIP system is improved through the interleaving technique. The performance of the system is evaluated for various types of noises at different network conditions. The performance of the enhanced VoIP signal is evaluated using wideband extension-perceptual evaluation of speech quality (WB-PESQ) measurement for wideband signal.

Keywords: *Voice over Internet Protocol, Interleaving, Packet Loss, Packet Size*

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. This results in harmful effects on the quality of VoIP, since speech packets can be discarded when routers or gateways are congested. 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. One of the most difficult problems in such networks is the packet loss issue. Even a single lost packet may generate audible distortion in the decoded speech signal. To reduce the effect of packet loss on perceived speech quality, the lost packets have to be regenerated at the receiver using packet loss concealment algorithms.

The brief description of the related work is presented in Section 2. The Section 3 presents the brief description of the interleaving packet concealment technique. The modeling of the IP network and VoIP simulations are presented in Section 4. The various performance evaluation results are presented in the Section 5. The conclusion and future scope of the work is presented in the last section.

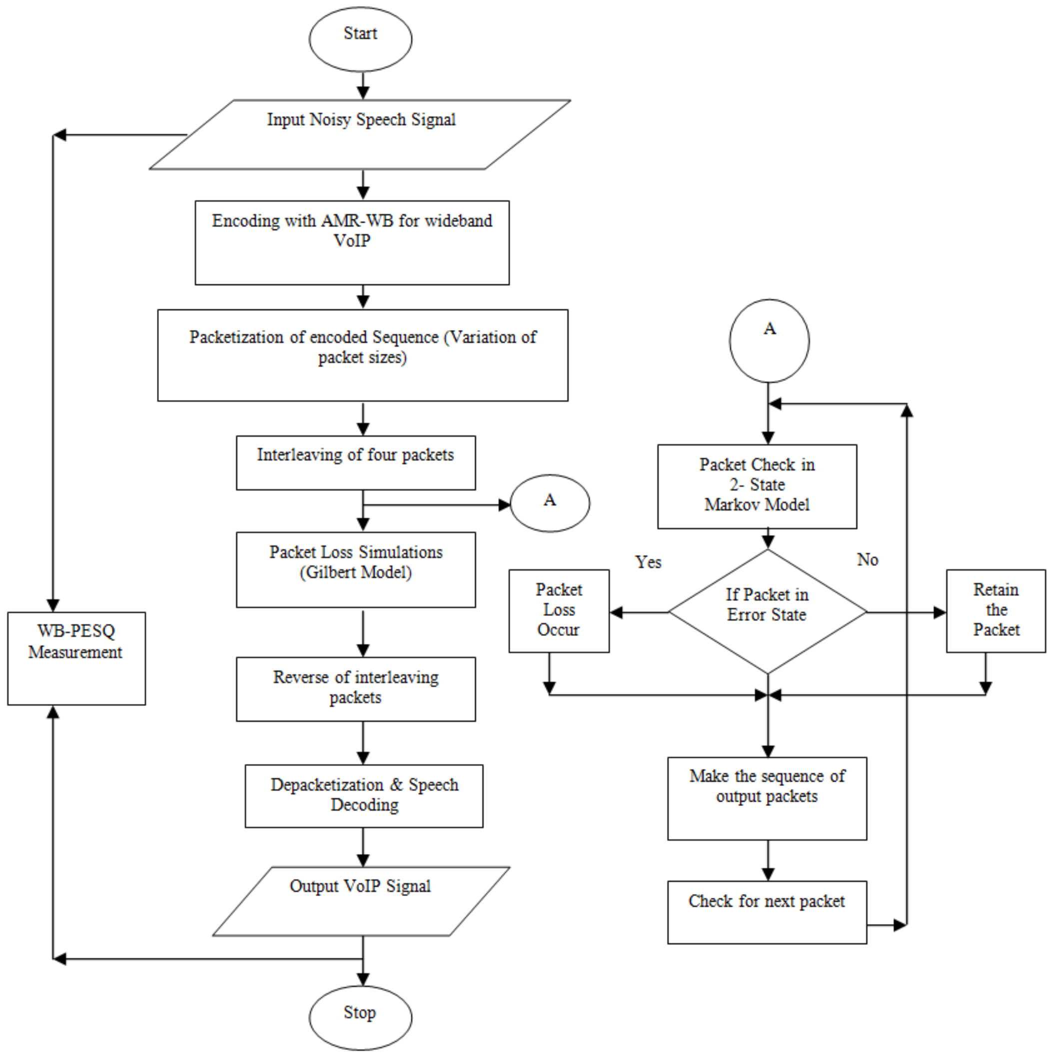

Figure 1. Flow Chart for Proposed System

2. Related Work

Goodman et.al had used Waveform substitution algorithms successfully for pulse code modulation (PCM) speech coder. The lost packets were also regenerated with the use of time scale modification algorithms [3, 4]. The lost packets were also regenerated with the use of time scale modification algorithms [5, 6]. Voicing classification technique had also been used to conceal the lost packets by estimating the excitation signal of the missing packets [7]. Feng et.al implemented ITU-T G.729 and G.723.1 speech codecs on TMS320C6201 DSP processor. The optimization methods used in the work had reduced the speech processing time. G.729 codec was able to process concurrently 20 voice channels and G.723.1 codec able to process 18 voice channels with single TMS320C6201 chip in IP telephony gateway [8, 9]. Langi [10, 11] had implemented ITU-T G.723.1 voice codec algorithm for VoIP gateways on TMS320C5402 DSP processor and use optimization techniques to improve processing time. Han et.al [12] had raised the issue of noise reduction for VoIP speech codecs. The authors proposed a modified Wiener filter based noise reduction scheme optimized to the estimated SNR at each frequency bin as a logistic function.

The work in this paper presents the improvement in the signal quality of the VoIP system through interleaving technique. The noisy speech signal is encoded with various wideband speech coders and the efficiency of the system is tested at varying packet loss rate and varying packet sizes network conditions.

3. Interleaving

To reduce the effect of packet loss on perceived speech quality, the lost packets have to be regenerated at the receiver using packet loss concealment algorithms. Interleaving technique is widely used packet loss concealment technique, when the multiple frames are used in single VoIP packet for concealment of the lost packets during VoIP communications. The interleaving scheme is incorporated into the VoIP system as presented in the Figure 1. In interleaving method, the data in N consecutive frames can be mixed together before transmission. The effect of packet loss during the transmission is assumed to spread over the N consecutive mixed frames in the process of interleaving. In this way, the loss of single packet destroys only a few bits from each frame. The process of interleaving reduces the effect of loss but at the expense of the substantial delays [13].

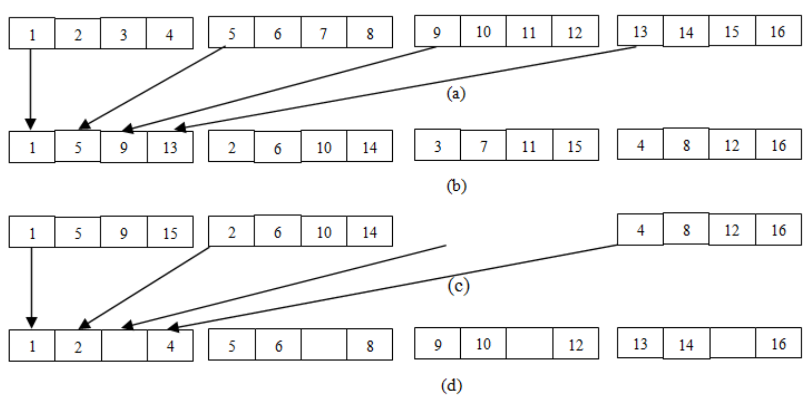

Figure 2. Interleaving Packet Loss Concealment Scheme: (a) Original Four Frames; (b) Four Frames Interleaved; (c) Frame Loss; (d) Reconstructed Frames

In this work, the four consecutive frames are interleaved before transmission and each frame is divided into 4 sub-frames for AMR-WB coder. The first sub-frames of each are grouped to form the first frame. The second sub-frames of each frame are concatenated to form the second frame. The third sub-frames of each frame are concatenated to form third frame. Then, the fourth sub-frames of each frame are concatenated to form the fourth frame. This interleaved data packets are then transmitted into IP network, where the packet loss occurs. After transmission, the loss of a single packet from an interleaved data stream results in multiple small gaps in the reconstructed stream. The whole interleaving process over four consecutive frames is depicted in Figure 2.

4. IP Network Modeling

The simulation of VoIP system was performed where each packet contains one frame. Packet losses are not independent on a frame-by-frame basis, but appear in bursts. The packet loss can be approximated by Markovian loss model such as Gilbert model, as discussed by Bolot in [14]. Most research in VoIP networks uses a Gilbert Model to represent packet loss characteristics, since the complexity is increased through the higher order Markov models [15]-[18]. Thus simulation of IP network was performed by using a 2-state Gilbert Model. The model has two states reflecting whether the previous packet is received or lost. The state “0” represents that a packet being correctly received and state “1” represents that a packet being lost. The Gilbert model is shown in Figure 3.

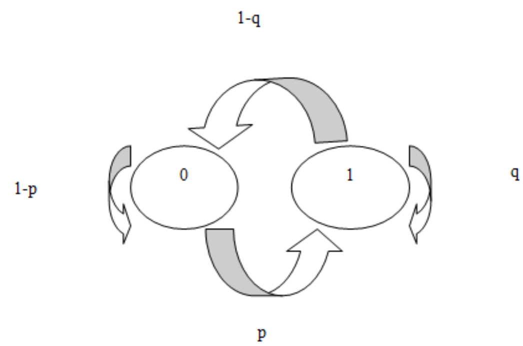

Figure 3. Two State Gilbert Model

Let p be the transition probability for the network model to drop a packet given that the previous packet is delivered i.e. the probability for network model to go from state “0” to state “1”. Let q is the probability for the network model to drop a packet given that the previous packet is dropped, i.e. the probability for the network model to stay in state “1”. This probability is also known as the conditional loss probability.

Let $p0$ and $p1$ denote the probability of the network model to be in state 0 and 1.The probability for a packet to be dropped regardless whether the previous packet is delivered or dropped i.e. the unconditional loss probability is exactly the probability for the network model to be in state 1 ($p1$).

$$p0 = \frac{q}{p+q}$$
$$p1 = \frac{p}{p+q} \qquad (1)$$

The transition matrix is given as

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \qquad (2)$$

Table 1. Simulated Loss Rates

PLR (%)	p	q
0	0	0
2	0.0032	0.15
4	0.0120	0.25
8	0.0250	0.25
10	0.1000	0.85

4.1. VoIP Simulations

The speech signal is encoded into VoIP frames using AMR-WB [19] codec at different data rates according to the network conditions in wideband VoIP system. The network impairments were introduced into VoIP frames with the modeling of the IP network through above discussed Gilbert model. The speech signal of VoIP system is degraded at different packet loss rates (PLR), which are simulated through different combinations of the p and q, as discussed in Table.I. The performance is evaluated with wideband extension-perceptual evaluation of speech quality (WB-PESQ) measurement defined by ITU-T recommendation P.862.2 [20] for wideband VoIP system. After comparing the degraded signal with the original one, the WB-PESQ measurement gives the subjective measurement as Mean Opinion Scores (MOS) value.

5. Results & Discussion

5.1. Wideband VoIP System without Background Noise

The simulations for the VoIP system are performed through 2-state Gilbert model and for the concealment of the lost packets, interleaving is incorporated into the system. The speech signal is encoded with various AMR-WB codec at bit rate of 6.6 kbps and the performance results of the interleaving based system are compared for various packet loss rates and for varying VoIP packet sizes. The speech samples for both male and female are taken from [21]. The overall perceptual speech quality is measured between the reference speech signal and the degraded speech signal with loss by WB-PESQ (P.862.2) for wideband VoIP signal.

5.1.1. Variation of Packet Size and Packet Loss Rate: The gain in WB-PESQ MOS scores is taken at different packet loss rates for interleaved output. The results for wideband VoIP system are presented in Table II. The significant improvement can be noticed in the results obtained with interleaving scheme in VoIP system at various network conditions. The proposed system with interleaving packet loss concealment gives much better results for various packet sizes.

Table 2. Comparison Results for Interleaving for AMR-WB 6.6 kbps based Wideband VoIP Signal

PLR (%)	Packet Size=20 ms		Packet Size=40 ms		Packet Size=80 ms		Packet Size=120 ms	
	None	Inter-leaving	None	Inter-leaving	None	Inter-leaving	None	Inter-leaving
2	3.49	3.72	3.41	3.63	3.39	3.57	3.34	3.52
4	3.37	3.64	3.30	3.58	3.24	3.48	3.15	3.37
8	3.24	3.53	3.16	3.45	3.08	3.33	2.89	3.14
10	3.11	3.43	3.00	3.32	2.89	3.21	2.60	2.86

Table 3. Comparison Results Interleaving with AMR-WB at 15.85 kbps Coded VoIP Signal in Noisy Environment

SNR	PLR (%)	Packet Size=20 ms		Packet Size=40 ms		Packet Size=80 ms		Packet Size=120 ms	
		None	Inter-leaving	None	Inter-leaving	None	Inter-leaving	None	Inter-leaving
	2	2.51	2.78	2.49	2.74	2.42	2.67	2.38	2.60
0 dB	4	2.37	2.68	2.36	2.63	2.29	2.56	2.26	2.50
	8	2.16	2.48	2.10	2.38	2.04	2.33	1.98	2.24
	10	2.07	2.41	1.99	2.29	1.95	2.27	1.88	2.18
	2	2.79	3.07	2.74	3.00	2.67	2.94	2.60	2.84
5 dB	4	2.63	2.95	2.59	2.88	2.57	2.86	2.52	2.78
	8	2.42	2.76	2.31	2.61	2.23	2.55	2.13	2.41
	10	2.27	2.63	2.19	2.53	2.11	2.45	2.01	2.32
	2	3.18	3.48	3.08	3.37	3.03	3.29	2.97	3.24
10 dB	4	3.04	3.36	2.94	3.25	2.88	3.19	2.82	3.11
	8	2.95	3.30	2.84	3.19	2.75	3.09	2.69	3.01
	10	2.85	3.22	2.76	3.13	2.66	3.03	2.56	2.93
	2	3.59	3.89	3.49	3.79	3.41	3.68	3.34	3.64
15 dB	4	3.47	3.81	3.36	3.69	3.27	3.60	3.20	3.53
	8	3.32	3.65	3.18	3.57	3.07	3.47	3.01	3.37
	10	3.11	3.51	3.01	3.42	2.92	3.33	2.85	3.25

5.2. Wideband VoIP System with Background Noise

The simulations for VoIP system are performed in noisy environment with different types of background noises including babble, car and street at SNR of 0, 5, 10, 15 dB. The speech signal is encoded with various wideband AMR-WB coder and the overall perceptual speech quality is measured between the reference speech signal and the degraded speech signal with loss by WB-PESQ for wideband VoIP signal. The noisy and clean speech samples for both male and female are taken from [22]. The performance results of the proposed system are evaluated for varying packet loss rates and for VoIP packet sizes in various noisy conditions.

5.2.1. Variation of Packet Size and Packet Loss Rate: The average gain in PESQ MOS scores is taken for various noise types at different packet loss rates for interleaved output. The comparison results for wideband VoIP system are presented in Table III- Table IV. The significant improvement can be noticed in the results obtained with interleaving scheme in VoIP system at various network conditions. The proposed system with interleaving packet loss concealment gives much better results for various packet sizes.

Table 4. Comparison Results for Interleaving with AMR-WB 18.25 kbps Coded VoIP Signal in Noisy Environment

SNR	PLR (%)	Packet Size=20 ms		Packet Size=40 ms		Packet Size=80 ms		Packet Size=120 ms	
		None	Inter-leaving	None	Inter-leaving	None	Inter-leaving	None	Inter-leaving
	2	2.55	2.88	2.52	2.81	2.47	2.71	2.40	2.66
0 dB	4	2.48	2.83	2.44	2.77	2.39	2.68	2.32	2.63
	8	2.39	2.76	2.35	2.69	2.30	2.59	2.22	2.54
	10	2.31	2.69	2.30	2.66	2.24	2.56	2.16	2.50
	2	2.80	3.12	2.77	3.05	2.71	2.98	2.65	2.91
5 dB	4	2.61	2.95	2.66	2.99	2.61	2.92	2.51	2.81
	8	2.59	2.95	2.55	2.90	2.48	2.83	2.38	2.66
	10	2.48	2.87	2.44	2.83	2.39	2.76	2.29	2.65
	2	3.20	3.55	3.17	3.47	3.11	3.38	3.03	3.31
10 dB	4	3.12	3.50	3.09	3.41	3.04	3.33	2.94	3.25
	8	3.05	3.46	3.01	3.37	2.94	3.29	2.85	3.20
	10	2.99	3.42	2.96	3.35	2.90	3.29	2.72	3.10
	2	3.61	3.97	3.57	3.88	3.50	3.82	3.43	3.72
15 dB	4	3.54	3.92	3.51	3.85	3.44	3.77	3.35	3.69
	8	3.44	3.85	3.40	3.77	3.34	3.70	3.25	3.59
	10	3.32	3.77	3.29	3.71	3.24	3.65	3.15	3.58

5.2.2. Comparison for Various Types of Noises: The comparison results of AMR-WB (15.85 kbps and 18.25 kbps) based VoIP system at various packet sizes with varying packet loss rates for different noise types is presented in Fig.4-Fig.5 respectively. The average gain in MOS scores is taken for various SNR at different packet loss rates for both filtered and interleaved output. For wideband VoIP systems, the interleaving scheme results significant improvement for each type of noise used in the present work.

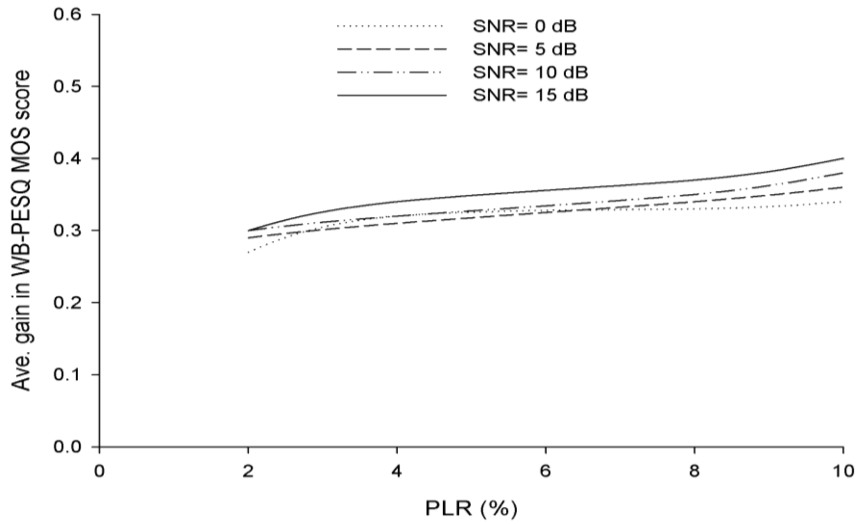

Figure 4. (a) Packet Size= 20 ms

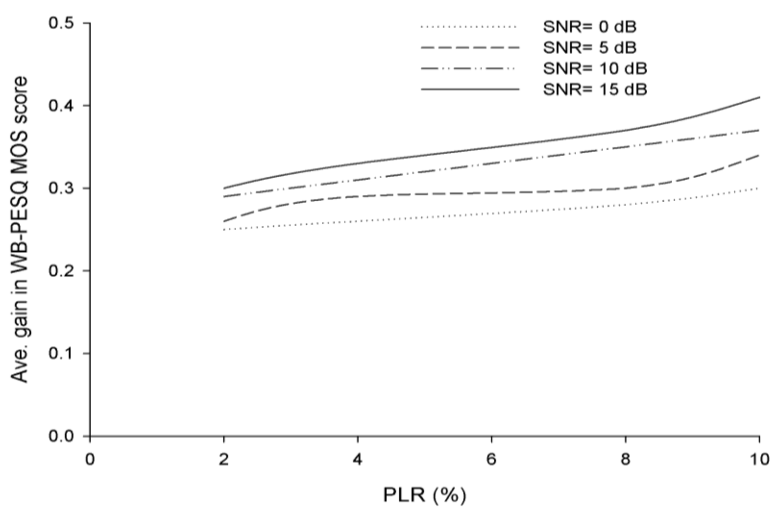

Figure 4. (b) Packet Size= 40 ms

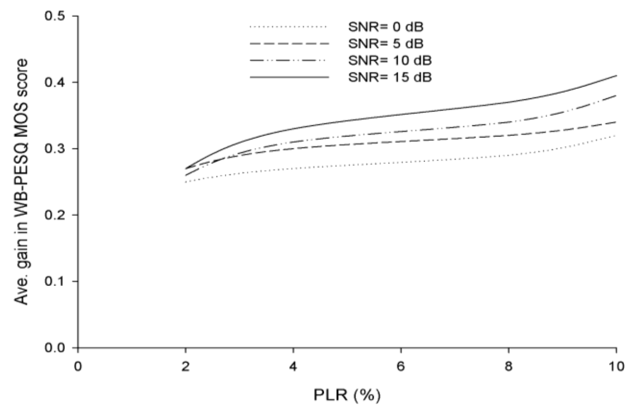

Figure 4. (c) Packet Size= 80 ms

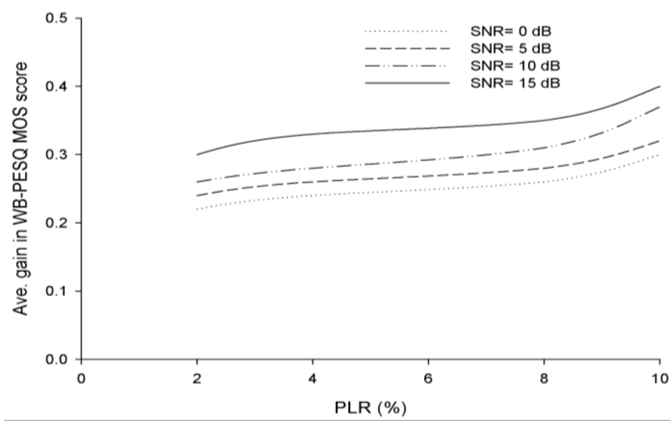

Figure 4. (d) Packet Size= 120 ms

Figure 4 Results for noises at different SNR for interleaved VoIP speech signal coded with AMR-WB at 15.85 kbps codec.

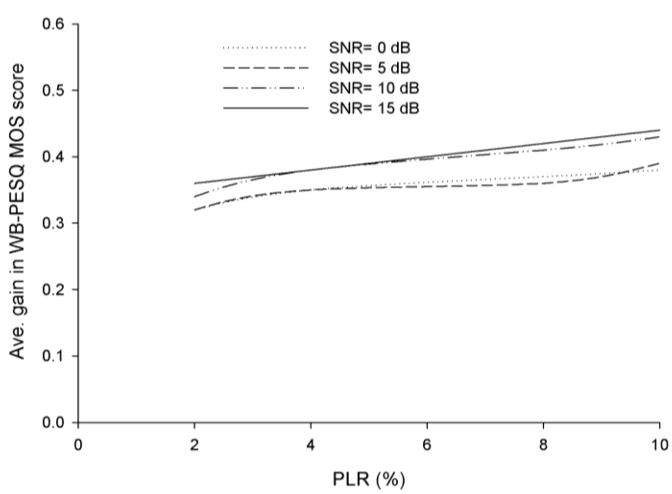

Figure 5. (a) Packet Size= 20 ms

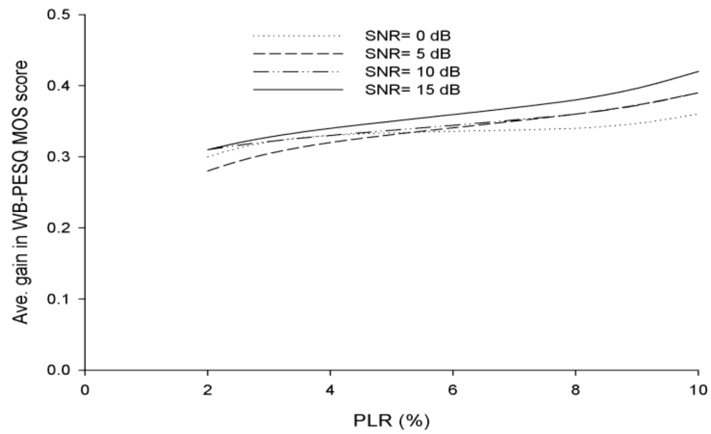

Fig. 5 (b) Packet Size= 40 ms

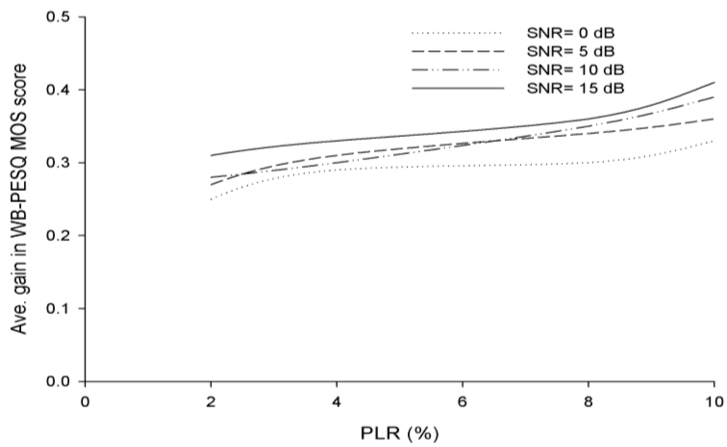

Figure 5. (c) Packet Size= 80 ms

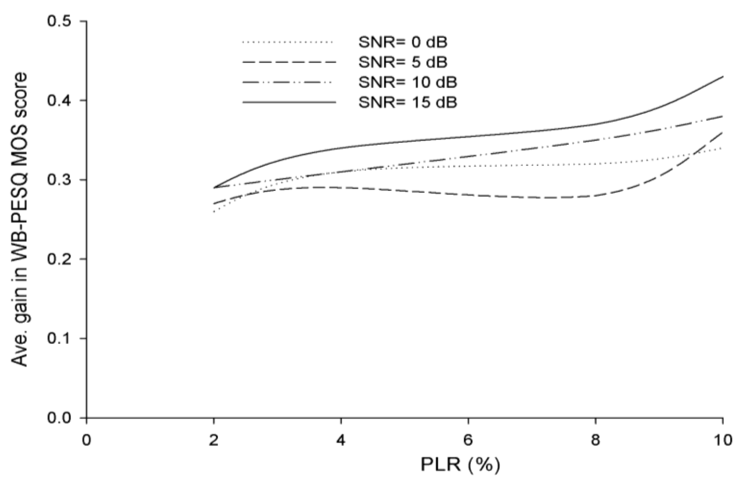

Figure 5. (d) Packet Size= 120 ms

Figure 5 Results for noise at different SNR for interleaved VoIP speech signal coded with AMR-WB at 18.25 kbps codec.

5.3. Discussion

The significant improvement in the quality of the VoIP signal is achieved with the interleaving scheme for different network conditions. The average increase of 0.33 and 0.26 in MOS scores is achieved with the interleaving scheme for single frame in each packet and six frames in each packet in AMR-WB based wideband VoIP system without the background noise. The interleaving scheme is also very much effective in the noisy environment. The average gain of 0.33 and 0.38 is achieved for single frame in each packet for AMR-WB at 15.85 kbps & 18.25 kbps based wideband VoIP system. For multiple voice frames in each packet in noisy VoIP system, the average gain in MOS scores for AMR-WB 15.85 kbps & 18.25 is 0.0.30 and 0.32 for six voice frames in each packet respectively. The significant improvement in signal quality with the interleaving scheme is observed for various type environmental noises as show in Fig.4 and Fig.5 for AMR-WB at 15.85 kbps and 18.25 kbps. The interleaving scheme not only reduces the background noise but also conceals the lost packets due to network impairments to improve the speech quality of the VoIP signal. Thus interleaving scheme can efficiently be used for VoIP applications.

6. Conclusion

The work in this paper presented an efficient method for reconstruction of lost packets and quality improvement of the VoIP system. The four frames were interleaved to spread out the distortion in the signal due to the lost of the packets. The performance of the interleaving based VoIP system was evaluated for varying packet loss rate and varying packet size in noisy environment. The significant increment in the MOS scores was obtained with interleaving of the packets in wideband VoIP system in presence of the various types of background noises, too.

References

[1] J. W. Goralski and C. M. Kolon, “IP Telephony”, 1st ed., McGraw-Hill, New York, USA, **(2000)**.
[2] U. Varshney, A. Snow, M. Mcgivern and C. Howard, “Voice over IP, Communications of the ACM”, vol. 45, no. 1, **(2002)**, pp. 89-96.
[3] D. Goodman, G. Lockhart, O. Wasem and W. Choong, “Waveform substitution techniques for recovering missing speech segments in packet voice communications”, IEEE Transactions on Acoustics, Speech and Signal processing, vol. 34, no. 6, **(1986)**, pp. 1440-1448.
[4] O. J. Wasem, D. Goodman, C. A. Dvorak and H. G. Page, “The effect of waveform substitution on the quality of PCM packet communications”, IEEE Transactions on Acoustics, speech and signal processing, vol. 36, no. 3, **(1988)**, pp. 342-348.
[5] W. Verhelst and M. Roelands, “An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech”, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, **(1993)** April 27-30, pp. 554-557, Minneapolis, USA.
[6] A. Stenger, K. B. Younes, B. Girod and H. Sanneck, “A new technique for audio packet loss concealment”, Proceedings of IEEE Global Telecommunication Conference, Communications: The key to Global Prosperity, **(1996)** November 18-22, pp. 48-52, London, UK.
[7] S. Jelassi, H. Youssef, C. Hoene and G. Pujolle, “Voicing aware parametric speech quality models over VoIP networks”, Proceedings of the Second international conference on Global Information Infrastructure Symposium, **(2009)** June 22-26, pp. 120-127, Hammamet, Tunisia.
[8] H. Y. Feng and Z. J. Ling, “Implementation of ITU-T G.729 speech codec in IP telephony gateway”, Wuhan University Journal of Natural Sciences, vol. 5, no. 2, **(2000)**, pp. 159-163.

[9] Y. F. Huang, J. Yang, S. Zhang and J. L Zhang, "Implementation of ITU-T G.723.1 dual rate speech codec based on TMS320C6201 DSP", Proceedings of 5th International Conference on Signal Processing, vol. 2, **(2000)** August 21-25, pp. 679-682, Beijing, China.
[10] A. Z. R. Langi, "Rapid development of a real-time speech coder on a TMS320C54x DSP", Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering, vol. 2, **(2002)** May 12-15, pp. 1045-1048, Manitoba, Canada.
[11] A. Z. R. Langi, "A DSP implementation of a voice transcoder for VoIP gateways", Proceedings of Asia Pacific conference on circuits & systems, vol. 1, **(2002)** December 16-18, pp. 181-186, Singapore.
[12] S. Han, S. Jeong, H. Yang and J. Kim, "Noise reduction for VoIP speech codecs using modified Weiner filter", Advances and Innovations in Systems, Computing Sciences and Software Engineering, **(2007)**, pp. 393-397.
[13] J. L. Ramsey, "Realization of optimum interleavers", IEEE Transactions on Information Theory, vol. 16, no. 4, **(1970)**, pp. 338-345.
[14] J. C. Bolot, "End-to-end packet delay and loss behavior in the internet", Proceedings of ACM Symposium on Communications Architectures, Protocols and Applications, **(1993)** September 13–17, pp. 289-298, San Francisco, CA, USA.
[15] H. Sannech and N. T. L. Le, "Speech property-based FEC for Internet Telephony applications", Proceedings of the SPIE/ACM/SIGMM Multimedia Computing and Networking Conference, (2000), pp. 38-51, San Jose, CA.
[16] O. Hohlfeld, G. Rudiger and G. Halblinger, "Packet loss in real-time services: Markovian Models generating QoE impairments", Proceedings of 16th International workshop on quality of service, **(2008)** June 2-4, pp. 239-248, Enschede, The Netherlands.
[17] C. C. Wu, K. T. Chen, C. Y. Huang and C. L. Lei, "An Empirical evaluation of VoIP playout buffer dimensioning in Skype", Google Talk and MSN messenger, Proceeding of 19th international workshop on network and operating system support for digital audio and video, **(2009)** June 3-5, pp. 97-102, Williamsburg, Virginia.
[18] H. P. Singh, S. Singh and J. Singh, "Computer modeling and performance analysis of VoIP under different strategic conditions", in Proceedings of 2nd IEEE International Conference on Computer Engineering and Applications, vol. 1, **(2010)** March 19-21, pp. 611-615, Bali, Indonesia.
[19] 3GPP TS 26.190, "Speech codec processing functions; Adaptive Multi-Rate Wideband (AMR-WB) speech codec; Transcoding functions, **(2009)**.
[20] ITU-T Recommendation P.862.2: Wideband Extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs, **(2005)**.
[21] Open Speech Repository, http:// www.voiptroubleshooter.com/open_speech/.
[22] http://www.utdallas.edu/~loizou/speech/noizeus/.