Pak. j. eng. technol. sci.
Volume 3, No 1, 2013, 11-22
ISSN: 2222-9930 print
ISSN: 2224-2333 online

PJETS

Maximising Bandwdith Efficency of Statistical Multiplexer Architecture using Frame Droping Methods

Asadullah Shah [1], Humaira Dar[2] and Iftikhar A Syed[3]

ABSTRACT

It is generally considered that remote and rural telephone users generate less traffic as compared to urban area users. This lowers the attraction of investment in rural areas by the telecommunications companies and service providers. The financial implications of wiring a vast area for low telephone traffic causes most telephone service providers to ignore those regions. Still, it is known that telecommunications are essential to the economic development of a region and that traffic increases rapidly as soon as the service is available. A satellite–based telephone network can provide efficient long distance telephone service to remote rural communities at a lower cost than land-based wired networks in most cases. Mobile satellite systems already provide this service, but are limited in capacity and charge high per-minute fees for the satellite link. Small earth stations and GEO satellites can provide this service more efficiently and at lower cost. On top of that, bandwidth efficient multiplexing with compressed speech, Voice Activity Detection (VAD) and Packet discarding methods can even further reduce

[1] Department of Computer Science, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia. Email: asadullah@kict.iium.edu.my

[2] Department of Computer and Information Engineering, Kulliyyah of Electrical and Electronic Engineering, International Islamic University Malaysia. Email:engg87.iium@gmail.com

[3] College of Computer Science and Information Systems, Institute of Business Management, Karachi

Pakistan. Email: syed.iftikhar@iobm.edu.pk

the cost of service for the users in rural areas. In this paper, Statistical Time Division Multiplexer (STDM) architecture was simulated. Two packet discarding methods, random packet discarding and cyclic packet discarding are used to maximize bandwidth utilization along with VAD. Results indicate that considering monologue speech source, with 80%, activity for each 6.4 kbps sources, on a channel of 64 kbps, 12 users can be allowed to be multiplexed instead of 9, therefore a Digital Speech Interpolation (DSI) advantage of 1.33 is achieved with 3% packet loss. Furthermore, it is observed that cyclic packet discarding technique perform better than random packet discarding in terms of subjective quality.

Key words: Statistical Multiplexing, Compressed Speech, Low bit Rate Channel, Rural Telephony, Bandwidth Efficiency, Digital Speech Interpolation Advantage, Loss Frame Reconstruction Techniques, Monologue speech sources.

1. INTRODUCTION

Speech has been a dominant source of communication between two parties over a telephone network, may it be fixed telephone through terrestrial network or any of the recent mobile communication networks. Users have to pay to telephone service providers for the usage and utilization of communication channel bandwidth. The channel bandwidth is the most expensive commodity in any communication system. Researchers around the globe are very much involved to develop technologies, tools and techniques to efficiently utilize the channel bandwidth so that the users can use the facilities at affordable cost[2]. Multiplexing is one of those techniques that allow users to share a communication bandwidth between many users. Traditionally, most common and well known multiplexing techniques has been in use are Frequency Division Multiplexing (FDM), and Time Division Multiplexing (TDM) [2]. The Statistical Time Division Multiplexing (STDM) is a variant of TDM. Also Digital Speech Interpolation Advantage of Statistical Time Division Multiplexer has been discussed in detail [22].The FDM is used for analogue telephone users and that has been almost replaced by TDM; a digital domain technique. The STDM initially used by the data communication purposes, not for real time speech communication, but recently, with sophisticated signal processing techniques and tools, has been considered for real time speech communications too. In addition to multiplexing, source coding, Voice Activity Detection (VAD)[3][4] and Lost Packet or Frame Reconstruction (LFR)[5] provides a significant room for maximizing bandwidth efficiency of any communication channel. Speech coding algorithms such as Pulse Residual Excited Linear Prediction (PRELP) a variant of Codebook Excited Linear Prediction (CELP) provide a good speech quality at low rates as low as 6.4 Kilo bits per second [5]. Packet reconstruction techniques [5] at 3% frame loss rate are acceptable for many applications, on top of that VAD [3][4], that detects the silence gaps within each users speech can save a bandwidth if that is efficiently utilized in multiplexer architecture design. The monologue speech, that is one way speech, like radio and TV, the activity is 80% and silences are 20%[7] and in dialogue mode of communications such as telephonic speech the activity is 40% [8][9][10] and 60% is silence at each users end. By utilizing compression, VAD and LFR an efficient multiplexer architecture has been designed and simulation results are provided in this paper, only for monologue speech sources. The paper is divided into many sections, after introduction, research methodology and simulation parameters are provide in section 2. In section 3, different architectures of STDM are explained. In section 4 results of two non criterion based frame discarding architectures are given with subjective measures. In section 5 Digital Speech Interpolation (DSI) advantage is explained. In Section 6 and 7 discussions, conclusion and future work directions are provided, respectively.

2. RESEARCH METHODOLOGY

Speech coder: A PRELP coder of 6.4 K bits per second is going to be considered for each speech source. This coding method is a variant of CELP and is known as hybrid coding scheme that uses both, waveform coding and Voice Coding (VOCODING) methods. Its performance at 6.4 K bps is rated as good as Adaptive Differential Pulse Code Modulation (ADPCM). The Mean Opinion Score (MOS) a well know subjective quality measure, attend by this scheme is around 4 out of 5 [11][12].

Transmission Channel: A 64 kbps channel such as Single Channel Per Carrier (SCPC) of Very Small aperture Terminal (VATS) is considered. VSATs are used all around the globe for various communications purposes mostly in the regions having difficult terrain conditions for establishing terrestrial networks. The VSATS are very effective in many countries of the world for rural telephony and are being used by medium scale organizations for their communication purposes [13].

Simulation Package: For simulation purpose the Communication Systems Software Package (COSSAP) is used to simulate the scenario. Due to limitations of this package some modules are developed in C-language and are integrated with existing COSSAP environment. The included modules are related with speech coding, VAD and LFR modules.

Voice Activity Detection: The VAD is used to detect the silence periods and separate talk spurts from silences. The VAD hangover used for simulation has been set for 4 frames of speech, that is, 80 mili seconds. That means even silence is detected, VAD will still declare up to 80 ms, a silence period as active speech or talk spurt. This hangover of 4 frames is so set to avoid already compromised quality of speech due to low bit rate coding from further deterioration of quality. Reducing the handover period may increase bandwidth utilization but on the cost of speech quality [3][4].

Two state Markov model: A two state Markov model is used to replace each of the speech sources except one source, in which real speech coder is used to monitor effects of the packet losses on overall quality. It is assumed that packet loss effects will be same for the remaining sources as is the case of real speech user**.** Another reason of using Markov model is to avoid extra processing requirement used by real sources. Each one of the Markov Model uses a different seed, therefore, simulating talk spurts and silences durations are randomized, it is same like the real situation as a VAD detects and generates the talk spurts and silences over real speech sources. The detailed design of Markov model is given in [14][15].

Determination of talk spurts and silence durations: In order to determine lengths and sizes of talk spurts and silences seven speech files with selected speech quality, such as thick male and female speech, thin speech, and child speech has been considered for an approximately 120 second or two minutes. The interval of observation was 20ms. After each 20 ms of speech VAD used to declare weather the frame is active speech or a silence. The number of times VAD declares each frames of speech (20ms) as active speech to determine minimum and maximum size of each talk spurt or silence respectively [16].

Loss Frame Reconstruction (LFR): The LFR is used to reconstruct dropped frames and compensate for speech quality at the receiving side. The frames or packets are forced to drop before taking on the transmission channel. At the receiving end dropped frames are reconstructed using LFR [4] scheme that used some of the parameters from previous frame to mitigate the effects of dropped frames. The limit of 3% frame loss rate has been set to avoid further speech quality loss. Since the PRELP coders are using long term and short term digital filters to model speech they are not able to bear and cope up with higher losses to recover lost frames once dropped for bandwidth efficiency purpose due to their filter memories[11][12].

Monologue speech sources: For simulation purpose a monologue speech sources are considered [7][8]. In monologue speech, activity per user is double than that of conversational speech sources. In monologue speech activity per user is 80% and 20% are silences. The activity in conversational speech is 40% and 60% is silence for each user at both the ends [9][10]. For this research only monologue speech sources are considered.

Subjective quality assessment: The most popular methods among so many others are the Mean Opinion Score (MOS). This method uses a scale from 1-5. The quality rated 1 is bad and 5 rating is given to excellent quality. In this research paper speech quality will be assessed on MOS basis, where all speech material, original, compressed without VAD, compressed with VAD having hangover of 4 frames and packet discarded and reconstructed speech will be played back to the listeners for quality rating.

3. STATISTICAL TIME DIVISION MULTIPLEXER ARCHITECTURE

The proposed architecture is based upon [16][17] Statistical Time Division Multiplexing (STDM) architecture. The TDM is known as a deterministic system that allows time slots to the users in a fixed fashion. None of the users can use slot of any other users no matter slots are empty and users slots are free. The schematic diagram of TDM is shown in Figure 1. In case of full load the bandwidth efficiency can reach to a maximum of 1. In practice and practically the bandwidth efficiency of TDM is always less than one.

Figure 1: Deterministic Time Division Multiplexing architectures

The STDM, on the other hand, uses dynamic time slot allocation on the bases of activity of the users; slots are not fixed any more. It depends upon the user's activity. During silences, time slots can be used for other users. This allows the room for extra users to be allowed to transmit their information over the same channel. How many excessive users can be allowed so that bandwidth may be utilized efficiently as well as quality of speech may be maintained to an acceptable level of perception, it depends upon few factors (a), user's activity ratio such as 80% or 40% for monologue or dialogue modes, respectively, (b) the frame loss percentage (c) VAD hangover (d) speech coder (d) buffering and delay acceptability (e) identical speech of the users (g) perceptual sensitivities (h) mixed data sources and priority settings (i) variable bit rate compression and (j) multimedia sources. A schematic diagram of STDM is shown in Figure 2.

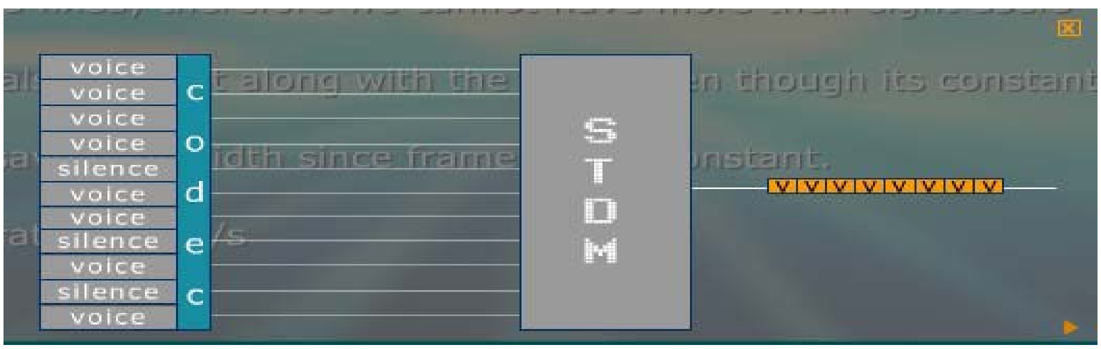

Figure 2: The schematic diagram of STDM

4. RESULTS OF STDM DESIGNS

The proposed multiplexing architecture is based upon low bit rate speech coder (PRELP) of 6.4 Kbps, VAD and packet dropping- packet reconstruction methods. There are 2 different packet dropping algorithms used in this simulation, as given below, keeping coding bit rate and VAD same for both the architectures.

1. Random Packet Dropping

2. Cyclic Packet Dropping

Both packet discarding (1 and 2) techniques are known as non criterion based methods because they do not involve perceptual sensitivities as a frame discarding criterion [18][19]. For evaluation and assessment purpose only subjective test score, very well known as MOS, is provided to measure and compare the performance of both the designs architectures. The MOS of random packet discarding method is given in Table 1. For the MOS, 35 experience researchers were involved to assess the quality. All the files were played to each of the users to score the quality on a scale of 1-5. The score rated 1 being the lowest quality and 5 the excellent quality. The number of users allowed to access the channel was higher than the channel can afford in this case (9). As soon as the active users crossed the 9 users limit the packet discarding mechanism was activated to drop packets of the users for all extra active users more than 9 users. In case of random packet discarding users were picked up on random bases. However, for cyclic packet discarding the users were forced to drop packet on a fair policy in which each user has to drop equally and to avoid consecutive packet losses. The MOS results of cyclic packet dropping are provided in Table 2. Window of observation is 20 ms, or a frame of speech. After each 20 ms, the number of users is counted to check how many are active, if that number goes higher than 9 users. The frame discarding mechanism was activated. On the receiving end lost frame was reconstructed to mitigate the effects of lost frames upon speech quality.

Table 1: MOS for random frame discarding approach

Users	Frame Loss Rate %	Speech Source 1	Speech Source 2	Speech Source 3	Speech Source 4	Speech Source 5	VAD
9	---	3.8	4.0	3.8	4.0	3.6	off
9	0	3.5	3.8	3.5	3.5	3.2	On
10	1	3.0	3.5	3.0	3.0	3.0	On
11	1	2.8	3.2	3.0	2.7	2.6	On
12	2	2.5	3.1	3.0	2.5	2.5	On

13	6	1.8	2.5	2.5	2.0	2.0	On

Table 2: MOS for cyclic frame discarding approach

Users	Frame Loss Rate%	Speech Source 1	Speech Source 2	Speech Source 3	Speech Source 4	Speech Source 5	VAD
9	--	3.8	4.0	3.8	4.0	3.6	off
9	0	3.5	3.8	3.5	3.5	3.2	On
10	1	3.5	3.8	3.4	3.2	3.0	On
11	1	3.0	3.5	3.2	3.0	2.7	On
12	2	3.0	3.5	3.0	2.7	2.5	On
13	6	2.2	2.7	2.0	2.0	2.0	On

5. DIGITAL SPEECH INTERPOLATION

The DSI advantage, previously known as Time Assign Speech Interpolation TASI [20][21] is known as the ratio of total users and users a channel can afforded without any sophisticated interpolation techniques. Theoretically, for monologue speech sources a DSI, with VAD is 1.25 and for conversational speech DSI is 2.5 because of activities of each sources as 80% and 40% respectively.

6. DISCUSSION

Two techniques involving frame discarding as a core method for bandwidth efficiency were simulated and speech quality is evaluated by using subjective measures or MOS. As seen in Table 1 and 2, first row, speech coders MOS is provided without activating VAD. The MOS rated by listeners is 3.6 lowest to 4.0 the highest. Since different speech material is used so is rated differently.

The Table 1 shows the packet dropped on random bases in case the number of active users goes beyond 9. The 9 users are the maximum that channel can accommodate. Since random packet discarding method does not consider any criterion so that users can be picked for consecutive packet discarding during higher activity periods. However, to overcome its limitations, the second method cyclic frame discarding, that uses a regulatory mechanism to avoid consecutive loss of same users for 2 frames during the higher activity time. In a result the cyclic frame discarding has performed better than random frame discarding method as is evident from MOS shown in Table 1 and 2. For example, for 12 users active, MOS for cyclic frame discarding is higher and quality is rated and perceived by the audiences more than that of random packet discarding method given in Table 1. The reason for this better quality is an arrangement of dropping in such a way that each user at least drop one frame before it is attempted for the second frame drop. This mechanism and control was not available in random dropping method. Since Lost packets are reconstructed at the receiving end. Speech coders always use previous frame memories for the next frame to be decoded. In case of consecutive loss of frames, memories are slow to pick up pace with coding side of algorithms. However, in single loss at a time provides better room for decoder memories to flash out the loss frame effects.

7. CONCLUSION AND FUTURE WORK

The statistical multiplexing architecture proposed in this paper was based upon low bit rate (compressed speech sources) The PRELP coder of 6.4 K bps was chosen for simulation along with two other signal processing techniques namely VAD and frame or packet discarding and reconstruction. Speech sources considered were monologue and activity in monologue sources is 80%, and only 20% each source is silent. For simulation purpose a channel of 64K bps such as VSAT channel was considered, where 9 users each having 6.4 kbps sources are considered. Our simulation results indicate that on top of 9 users, 3 extra users can be allowed to use same channel that can accommodate 9 users otherwise. These 3 users extra can be allowed at the cost of 3% packet loss to each user. The packet loss is mitigated by using FLR techniques. In this way a low bit rate channel such as VSAT-SCPC can be utilized effectively for rural telephony effectively. A DSI advantage of 1.33 that is (12/9) can be achieved for monologue speech sources over a channel of 64 k bps.

This work can be further extended for dialogue or conversational speech sources along with variable bit rates, perceptual sensitivities as frame dropping schemes, user's speech similarities and also using priority and multimedia services. It is expected that a DSI advantage of 3-4 could be achieved that means on top of 9 users, around 18-27 users can be allowed to access the same channel provided conversational speech or telephonic speech is used.

References

1. Johnson Adegbenga A and Yinusa A Ademola A " Performance Analysis of Statistical Time Division Multiplexing System" Leonardo Electronic Journal of Practices and Technology, issue 16, Pages 151-166 January-june 2010.
2. Held G. " The Multiplexer Reference Manual", John Willy and Sons Ltd, 1992.
3. Scouthcott C. B. Freeman et. all. " The Voice Activity Detector for the Pan-European Digital Cellular Mobile Telephone Service". In Proc. Of ICASSP, pages 369-372, Glasgow, UK, 1989.
4. Srinivsan K and Gersho A. " Voice Activity Detection for Cellular Networks". Technical Report CA 93106 USA, Center for Information Processing Research, Department of Electrical and Computer Engineering University of California.
5. Atungsiri S. et.al. " Effective Lost Speech Frame Reconstruction For CELP Coders". In Proc. Of EUROSPEECH Conf., Volume 2, Pages 599-602, Genova, Italy, 1991.
6. Makhoul J. " Linear Prediction: A Tutorial Review ", In Proceedings of IEEE, pages 561-580, April 1975.

7. Gruber J. G. " A Comparison of measured and calculated Speech Temporal Parameters Relevant to Speech Activity Detection", IEEE TRANS. ON Communications 30(4): 728-738, April 1982.
8. Brady P. T. " A Technique for Investigating ON-Off Patterns of Speech". Bell Sys. Technical Journal pages 1-22, November 1965.
9. Brady P. T. " A Statistical Analysis of On-Off Patterns in 16 Conversations". Bell Sys. Technical Journal pages 73-91, January 1968.
10. Brady P. T. " A model for Generating On-Off Speech patterns in Two –way Conversation". Bell Sys. Technical Journal, Pages 2445-2472, September 1969.
11. Kondoz A. M, et al. " Pulsed Residual Excited Linear Prediction". IEE Proc. On vision and Image Processing, 142(2): 105-110, April 1995.
12. Atal B. S and Schroeder M. R " Stochastic Coding of Speech Signal at Very Low Bit Rates", In Proc. Of IEEE Int. Conf. on Communications, Pages 1610-13, Amstredam, Netherlands, May 1984.
13. Naderi F. M. and Wu W. W. " Advanced Satellite Concepts For Future Generations VSAT Networks". IEEE Communications Magazine, Pages 13-21, September 1988.
14. Lutz Erich. " A Markov Model For Correlated Land Mobile Satellite Channels", International Journal of Satellite Communications, 14:333-339, 1996.
15. A. Shah and Siraj P. " Analysis of On-Off Patterns of Speech signals by a Two State Markov Model", Sindh University Research Journal (Science Series), Vol. 34-a(1), pages 57-62, September 2002.
16. A. Shah et. al., " Lossy Multiplexing of Low Bit Rate Speech in Thin Route Telephony. IEE Electronics Letters, 32(2): 905-97, 1996.
17. A. Shah" An Intelligent Multiplexer Architecture for Thin Route Communications", Ph.D Thesis, University of Surrey, England UK. 1998.
18. A. Shah et. al. "Statistical Time Division Multiplexing with Cyclic Packet Dropping Algorithms for low bandwidth Channel" Sindh University Research Journal (Science Series), Vol. 34-b(2), pages 11-14, September 2002.
19. A. Shah and S. Laghari. " Design of Statistical Time Division Multiplexer based on Random Packet Loss Algorithms" Sindh University Research Journal (Science Series), Vol. 34-b(2), pages 11-14, September 2002.

20. Clifford J. W. “The Tradeoff Between Delay and TASI advantage in a Packetized Speech Multiplexer”. IEE Trans. On Communication 27(11):1716-1720, November 1979.
21. Clifford J. W. “ Fractional Speech loss and Talkspurt Activity Model for TASI and for Packet-Switched Speech” The Tradeoff Between Delay and TASI advantage in a Packetized Speech Multiplexer”. IEEE Trans. On Communications 26(8):1253-1257, August 1978.
22. A Shah and K.L. Khatri “Digital Speech Interpolation Advantage of Statistical Time Division Multiplexer” PJETS volume 1, number 1, 9 -11, 2011