PJETS Volume 1, No 1, 2011, 9-13
ISSN: 2222-9930 print

Digital Speech Interpolation Advantage of Statistical Time Division Multiplexer

Asadullah Shah*, Krishan Lal Khatri**

Abstract

This paper discusses Digital Speech Interpolation (DSI) advantage of Statistical Time Division Multiplexer (STDM) for Random Packet Loss scenarios. In order to harness this advantage each source speech is compressed at 6.4 kbps, Voice Activity Detector (VAD) is used for each source, and packet loss mechanism is introduced to achieve maximum DSI advantage. It is observed that for maximum of 9 users channel capacity, extra 3 users, i.e. total 12 users can be accommodated with 3% speech frame losses, which results in an advantage of 12/9 = 1.33.

Keywords: Digital Speech Interpolation (DSI), Statistical Time Division Multiplexing (STDM), Voice Activity Detector (VAD),

1. INTRODUCTION

In dialogue speech the activity of user occupies 40%, while the silence occupies 60% of the total time [1] [2]. On the other hand, in monologue the user's speech occupies 80% of time, while 20% of the time is inactive or silent. These silence gaps can be detected using Voice Activity Detector (VAD) on each input speech source of Statistical Time Division Multiplexer [3]. Considering a 64 Kbps channel, and 6.4 Kbps speech codec employed on each input line, ten users can be accommodated on a 64 Kbps transmission channel. When more than ten users are connected at the input of channel, speech (frames) of some of the users has to be blocked or discarded. The discarded frames of speech sources need to be reconstructed at the destination.

Frame Relay and X.25 systems are also categorized as STDMs. Both of these systems utilize aggregate HDLC frame structures, and both of these systems can interoperate with both Private and Public systems. The advantage of Frame Relay over X.25 is that it can support the same traffic as X.25, while, while facilitating "bandwidth on demand" requirements for "bursty" traffic (e.g. LANs). Public Frame Relay services are available, offering customers additional methods to interconnect LANs, rather than having dedicated Wide Area Network (WAN) links.

* Dr.Asadullah Shah is Professor International Islamic University of Malaysia Kulliyyah Of Information And Communication Technology Gombak , Malaysia. Email: dr_asadullahshah@hotmail.com

** Engr. Krishan Lal Khatri is Senior Lecturer, Department of Telecommunication Engineering and Management at Institute of Business Management, Karachi, Pakistan. Email: engrkrishan@yahoo.com

Frame Relay, however, cannot adequately support voice or video traffic because of variable End-To-End delivery times (e.g. variable delay). Voice and video transmissions are of a "Constant Bit Rate" (CBR) nature, and do not fare well sitting in a queue waiting for a big LAN packet to finish transmitting [8].

For transmission of voice or speech, frame discarding method is most appropriate. Simplest method among frame discarding approach is random packet loss technique [4]. During higher activities, i.e. when activity of user is higher than maximum capacity of the channel, some of the users are randomly selected and their packets are dropped. The transmitting and receiving sides are informed about the losses to their respective users. For each such loss, Lost Speech Frame Reconstruction at both ends is activated and reconstruction of the discarded frame is carried out so that the CODECS at both ends can operate in synchronization. Using VAD and packet loss mechanism, bandwidth of the channel can be saved. This in fact increases DSI advantage. This paper presents DSI advantage for proposed STDM.

2. EXPERIMENTAL SCENARIO

Voice or speech is, in fact, an analog phenomenon. Analog input must be digitalized before it is handled in the network. For this, it is sampled, quantized, and encoded by using a *codec*. The most common coding scheme, i.e. Pulse Code Modulation (PCM) encodes speech into 64 Kbps per channel. A great deal of work has been done exploring alternative coding techniques -- particularly aimed at reducing the bandwidth required to transmit the digital signal, without sacrificing much voice quality. Digital Speech Interpolation (DSI) does require a coding technique which can tolerate long bursts with no data. In a packet voice system, using a packet-switched communications network, the source merely stops generating packets during a silence interval [5].

Bandwidth efficiency is typically increased for speech by employing techniques such as Time Assignment Speech Interpolation (TASI) for analog, and DSI for digital domain. DSI is a technique through which speech samples are mixed to replace the missing ones, i.e. the silence gaps are filled with talk spurts. The main principle of TASI and DSI is utilization of the inactivity periods of each user. In case of random frame discarding technique, during momentary overloading, consecutive frame losses can occur, and the same user can suffer adjacent frame losses. For low bit rate CELP or PRELP orders 3% random frame losses can be acceptable [6] [7] but beyond this, the speech quality degrades abruptly, and is not affordable.

This research work was carried out using Communication System Software Application Package (COSSAP) considering following parameters.

- Pulse Residual Excited Linear Prediction (PRELP) speech coder of 6.4 Kbps compressed speech
- 64 Kbps channel capacity, i.e. 64/6.4 = 10 users
- VAD to detect 80% activity and 20% silence for monologue speech
- 3% frame loss is acceptable for maintaining the speech quality
- For each source (user) a 30 ms speech segment (frame length) is considered and VAD is applied

- For each source (user), 2500 frames are processed making a total length of each source speech 30ms x 2500 = 75000 ms or 75 seconds
- Speech sources arranged to represent real-life scenario, i. e. random and independent using variable random offset

3. RESULTS

The frame loss verses number of users for this design is illustrated in Fig 1. For users, analysis indicates no loss but simulation graph indicates lower than 1% frame loss. Around 3% and 1% loss is indicated for 11 users as shown in analysis and simulation graphs respectively. Similarly, for 12 users, the frame loss for analysis is around 5%, whereas for simulation it is around 2%. As long as the number of users is below 13, the frame loss received by the user is within acceptable range that is around 3%. As the number of users on the link is raised to 14, both analysis and simulation suffer a huge loss. It is believed that if the number of users increased high enough, both analysis and simulation loss statistics come close to each other.

4. DISCUSSION

The DSI advantage of Random Packet Loss based Statistical Time Division Multiplexer (STDM) has been carried out. In order to achieve maximum DSI advantage, three interpolation techniques were used, i.e. VAD, frame loss, and reconstruction. It is observed that during overloading of traffic, some users were picked up on random basis so that their packets were dropped. The dropped packets are reconstructed to mitigate effects of looser ones. With this arrangement, 12 users were allowed on a channel of 9 users' capacity. It is learnt that with an affordable compromise on quality, three extra users can be allowed and a DSI advantage of 12/9 = 1.33 is achieved.

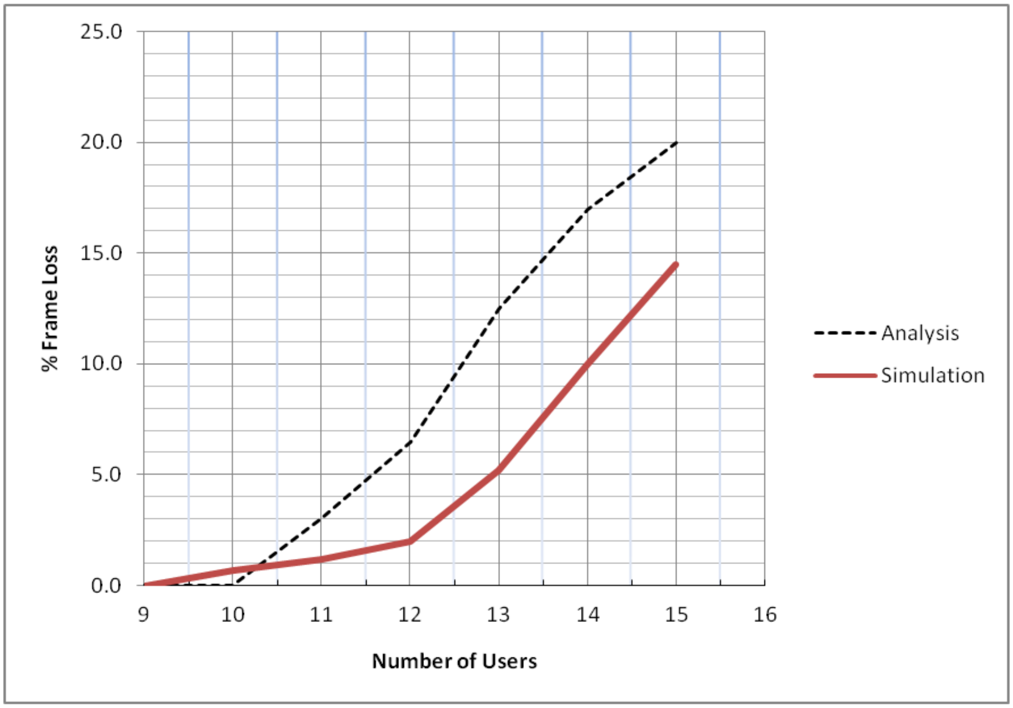

Figure 1: Losses in random multiplexer simulation

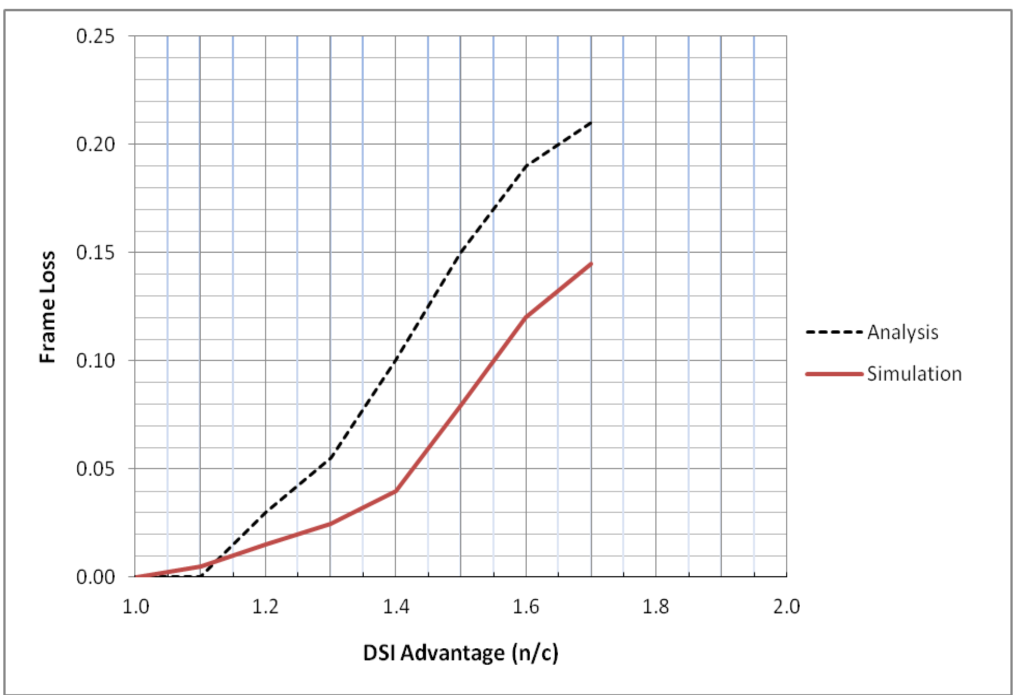

Figure 2: DSI Advantage v/s Frame loss in random multiplexer simulation

5. CONCLUSION

DSI factor is based on activity, number channels available for sharing, and capacity required per user. In order to achieve DSI above 1.33, the number of channels can be raised, activity of each user may be reduced, or further compression and frame loss mechanism can be allowed, which may result in higher DSI achievement.

REFERENCES

[1] Paul T. Bardy, "A technique for Investigating On-Off Patterns of Speech", Bell Sys. Technical Journal, Pages 1-22. November 1965

[2] Paul t. Bardy, "A Statistical Analysis of On-Off Patterns in 16 Conversations", Bell Sys. Technical Journal, Pages 73-91. January 1968

[3] Gilbert Held, "The Multiplexer Reference Manual", John Willy and Sons Ltd. 1992

[4] A. Shah, S. Laghari, "Design of Statistical Time Division Multiplexer based on Random Packet Loss Algorithm", Sindh University Research Journal. 2002

[5] John F. Shoch. "Carrying voice traffic through an Ethernet local network -- a general overview", Xerox Corporation. 1980

[6] A. Kandoz, S. Atungsiri, R. Soheili, and B. Evans. "Effective Lost Speech Frame Reconstruction for CELP Coders". In Proc. Of EUROSPEECH Conf., Volume 2, Pages 599-602, Genova, Italy, 1991

[7] A. M. Kandoz and B. G. Evans. "A High Quality Voice Coder with Integrated Echo Canceller and Voice Activity Detector for Mobile Satellite Applications". 3rd Int. Mobile Satellite Conf and Exhibition, June 1993

[8] ©Tampa Bay Interactive, Inc. USA

http://telecom.tbi.net/mux1.html accessed 25-December-2004