

An Analysis of VoIP Communication and Overview of Bangladesh Practical Field

N. A. Shafi, *Member, IACSIT*, O. Farrok, and M. M. Ali

Abstract—In this paper the latest development G.729 which is an audio data compression algorithm has been analyzed and implemented to save a noticeable bandwidth. Different effects of AnnexB=yes and AnnexB=no has been observed in the G.729 codec which has the standard bandwidth of 8 kbps. As most of the internet subscribers of Asian countries use very low internet bandwidth the goal of this analysis is to propose to change some parameters of this system so that the quality of voice may be kept in a tolerable limit using only 5kbps to 6kbps where this codec uses at least 6.4kbps. In a real life environment it is tested practically that it is possible to transmit voice satisfactorily using less than 6.4 kbps by changing some parameter described in this paper.

Index Terms—G.729 codec, AnnexB=yes, AnnexB=no, RTP (real-time transport protocol), VAD (voice activity detector).

I. INTRODUCTION

The G.729 codec perform voice compression at bit rates that vary between 6.4 and 12.4kbps [1]. It is an audio data compression algorithm for voice that compresses digital voice in packets of 10 milliseconds duration. It is officially described as Coding of speech at 8 kbit/s using conjugate-structure algebraic code-excited linear prediction (CS-ACELP). Because of its low bandwidth requirements, G.729 is mostly used in Voice over Internet Protocol (VoIP) applications where bandwidth must be conserved. Standard G.729 operates at a bit rate of 8 kbit/s, but there are extensions, which provide rates of 6.4 kbit/s (Annex D, F, H, I, C+) and 11.8 kbit/s (Annex E, G, H, I, C+) for worse and better speech quality, respectively.

To achieve good quality low-bit-rate silence compression, a robust frame-based voice activity detector module is essential to detect inactive voice frames, also called silence or background noise frames. The achievement of bit-rate savings for coded speech at average rates as low as 4 kb/s during normal speech conversation while maintaining reproduction quality [2]. G.729 Codec is closely relation to RTP [3]. The packet-level VoIP system performance has been focused. The analogue voice signals are digitized,

compressed and then encoded into digital voice streams by the codec [4]. The VoIP protocol stack is illustrated in Fig. 1. Voice packets are transmitted over the IP network, and the reverse processes of decoding and depacketizing are accomplished at the receiver. A playout buffer is used by the receiver to smoothen the speech by getting rid of delay jitter. Packets arriving later than the playout time will simply be discarded. Some other components such as voice/silence detector, loss/error concealment and echo canceller, are also included in the system to enhance the functionality and performance of VoIP systems. The major metric to evaluate the user-perceived voice quality is the Mean Opinion Score (MOS).

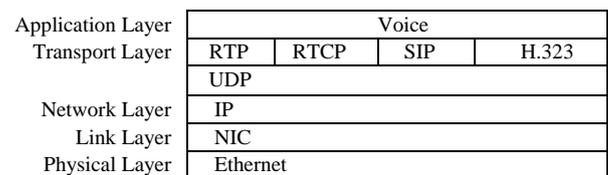


Fig. 1. VoIP protocol stack.

Voice codecs are standardized by the International Telecommunication Union-Telecommunication (ITU-T), such as G.729 with 8 kbps, G.723.1a with 5.3/6.3 kbps, etc. The output voice stream then enters the packetizer to generate constant bit rate (CBR) audio packets with RTP (RTCP)/UDP/IP header where RTP and RTCP [5] are Real-time Transport Protocol and Real-Time Control Protocol, respectively, which are designed to support real-time multimedia applications with stringent delay constraint over unreliable User Datagram Protocol (UDP). Besides these, call set-up signaling protocols, such as Session Initiation Protocol (SIP), are used for establishing VoIP connections. SIP, defined in RFC 2543 of Internet engineering task force (IETF), is a signaling protocol for Internet applications, e.g. conferencing, telephony, events notification, and instant messaging.

II. DESCRIPTION OF THE SYSTEM

G.729 has been extended in Annex B (G.729b) which provides a silence compression method that enables a voice activity detection (VAD) module. It is used to detect voice activity in the signal. It also includes a discontinuous transmission (DTX) module which decides on updating the background noise parameters for non speech (noisy frames). It uses 2-byte Silence Insertion Descriptor (SID) frames transmitted to initiate comfort noise generation (CNG). If transmission is stopped, and the link goes quiet because of no speech, the receiving side might assume that the link has been

Manuscript received November 8, 2012; revised December 18, 2012. This work was arranged and supported by Zamir Telecom Limited, Authors would like to express their deepest sense for the technical and financial support.

N. A. Shafi is with the Zamir Telecom Limited, 4-6 Davenant Street, Unit # A, London E1 5NB, England (e-mail: nahid_apee@yahoo.com).

O. Farrok is with the Department of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology (AUST), Dhaka-1208, Bangladesh (e-mail: omarruet@gmail.com).

M. M. Ali is with the Department of Electrical and Electronic Engineering, Rajshahi University of Engineering and Technology (RUET), Rajshahi-6204, Bangladesh (e-mail: mmali.ruet@gmail.com).

cut. By inserting comfort noise, analog hiss is simulated digitally during silence to assure the receiver that the link is active and operational [6].

A Voice Activity Detector (VAD) with a comfort noise generator (CNG), achieves silence compression, which is very important in modern telecommunication systems [7]. In multimedia communications a VAD guarantees simultaneous voice and data applications; in Universal Mobile Telecommunication Systems (UMTS), it reduces the average bit rate; finally, in a cellular radio system using the Discontinuous Transmission (DTX) mode, it reduces co-channel interference and power consumption in portable equipment. The paper presents a performance evaluation and comparison of recent ITU-T and ETSI voice activity detection algorithms. The last ITU-T VAD standard is Rec. G.729 Annex B [8], developed for fixed telephony and multimedia communications [9]. The G.729 Codec receives and transmits audio packet two ways in the VOS switch. AnnexB=yes is explained in Fig. 2.

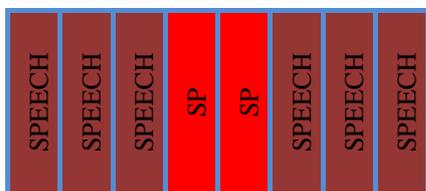


Fig. 2. Block diagram of AnnexB=yes.

There is 8kbps packet. Two packets is silence compression (SP) that means the gap of two speeches. The silence compression generates the room noise. AnnexB=no is explained in Fig. 3.

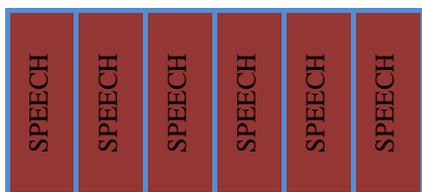


Fig. 3. Block diagram of AnnexB=no.

There is 6kbps packet. Two packets do not generate silence compression. Recently, Payload characteristic research (G.723 Annex A and G.729 Annex B) is running in the world [10].

III. DISCUSSION AND RESULT

G.729 has an optional Annex B [11], which specifies the use of silence suppression and comfort noise generation. In typical speech, only one person talks at a time. Therefore, speech consists of periods of talking (called talkspurts), followed by periods of silence. Additional compression can be achieved by discovering the silence periods. Older approaches would send either nothing for the silence periods, or would send a simple energy value, which the decoder would use to insert white noise. The practical example of the call traces are shown in Fig. 4. to Fig. 9. However, in environments with loud and non-stationary background noise, both approaches are inadequate. The algorithm operates by first making a Voice Activity Detection (VAD) decision in each frame.

Ser...	Caller Signaling	Callee Signaling	Memo
1	192.168.5.9:5060 INVITE(SDP)	192.168.3.4:5060	
2	192.168.5.9:5060 100 Trying	192.168.3.4:5060	
3			find gateway:2
4			available gateway:2
5			send to gateway OU-ROUTE
6		192.168.3.4:5060 INVITE(SDP)	
7		192.168.3.4:5060 100 Trying	
8	192.168.5.9:5060 100 Trying	192.168.3.4:5060	
9		192.168.3.4:5060 183 Session Progress(SDP)	
10	192.168.5.9:5060 183 Session Progress(SDP)	192.168.3.4:5060	
11		192.168.3.4:5060 200 OK(SDP)	
12		192.168.3.4:5060 ACK	
13	192.168.5.9:5060 200 OK(SDP)	192.168.3.4:5060	
14	192.168.5.9:5060 ACK	192.168.3.4:5060	

Fig. 4. 1st portion of call trace report for receiver end of VOS switch as AnnexB=yes.

```

Detailed Signaling
CALLFLOW|15:03:32.741544|EP1->SERVER|192.168.5.9:5060
INVITE sip:12127773456@192.168.3.4 SIP/2.0
From: "7768845830"<sip:7768845830@192.168.10.6 >,tag=72
To: <sip:12127773456@192.168.3.4 >
Call-ID: eb93e0-83809bd8d9f04b5-001b217fed16@192.168.5.9
CSeq: 1 INVITE
Via: SIP/2.0/UDP 192.168.5.9:5060 ;branch=z9hG4bK-6a152
Max-Forwards: 67
Supported: timer
Allow: INVITE,CANCEL,ACK,BYE,OPTIONS,INFO,UPDATE
Contact: <sip:7768845830@192.168.5.9:5060 >
Min-SE: 1800
Session-Expires: 1800
Content-Type: application/sdp
Content-Length: 246
v=0
o=digitalk 071547800 071547800 IN IP4 192.168.10.6
s=digitalk
i=session information
c=IN IP4 192.168.10.6
t=0 0
m=audio 17338 RTP/AVP 18 101
a=rtpmap:18 g729/8000
a=frtp:18 annexb=yes
a=rtpmap:101 telephone-event/8000
a=ptime:20
    
```

Fig. 5. 2nd portion call trace report for receiver end of VOS switch as AnnexB=yes.

Ser...	Caller Signaling	Callee Signaling	Memo	Time
1	192.168.8.12:5060 INVITE(SDP)	192.168.3.4:5060		11:57:...
2	192.168.8.12:5060 100 Trying	192.168.3.4:5060		11:57:...
3			find gat...	11:57:...
4			availabl...	11:57:...
5			send to...	11:57:...
6		192.168.3.4:5060 INVITE(SDP)		11:57:...
7		192.168.3.4:5060 100 Tryin INVITE(SDP)		11:57:...
8	192.168.8.12:5060 100 Trying	192.168.3.4:5060		11:57:...
9		192.168.3.4:5060 404 Not Found		11:57:...
10		192.168.3.4:5060 ACK		11:57:...
11			send to...	11:57:...
12		192.168.3.4:5060 INVITE(SDP)		11:57:...

Fig. 6. 1st portion of call trace report for receiver end of VOS switch as AnnexB=no.

```
Detailed Signaling
CALLFLOW|11:57:09.879385|EP1->SERVER|192.11
INVITE sip:269868801748973786@192.168.3.4
CSeq: 1 INVITE
Via: SIP/2.0/UDP 192.168.8.12:5060 ;branch=z9h
From: "624197" <sip:7903904900@192.168.8.12:5060>
Call-ID: GW79039049002819063557@192.168.8.12
To: <sip:269868801748973786@192.168.3.4 >
Contact: <sip:192.168.8.12:5060 ;transport=udp>
Content-Type: application/sdp
Content-Length: 248
v=0
o=VoipSwitch 9066 9066 IN IP4 192.168.8.12
s=VoipSIP
i=Audio Session
c=IN IP4 192.168.8.12
t=0 0
m=audio 8066 RTP/AVP 18 101
a=rtpmap:18 G729/8000/1
a=fmtp:18 annexb=no
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
a=sendrecv
```

Fig. 7. 2nd portion of call trace report for receiver end of VOS switch as AnnexB= no.

$(6.4-5/6.4) \times 100\% = 21.875\%$. So average bandwidth saving is $(6.25 + 21.875)/2 = 14.0625\%$ or approximately 14%.

```
CALLFLOW|11:57:18.123941|EP2->SERVER|192.168.3.4:5060
SIP/2.0 183 Session Progress
Content-Type:application/sdp
Allow:INVITE,ACK,OPTIONS,BYE,CANCEL,REGISTER,INFO,I
Date:Fri, 20 Apr 2012 08:32:18 GMT
Organization:Zamir
From:"7903904900"<sip:7903904900@192.168.3.4 >;tag=
To:<sip:8801748973786@192.168.2.8 >;tag=4585303036:
Call-ID:3404c24063616c6c11150cd4@192.168.3.4
CSeq:2126 INVITE
Server:TB006293/2.1
Via:SIP/2.0/UDP 192.168.3.4:5060 ;branch=z9hG4bK16b5a
Content-Length:161
v=0
o=tb640 5 1 IN IP4 192.168.2.8
s=-
c=IN IP4 192.168.2.8
t=0 0
m=audio 21856 RTP/AVP 18 101
a=fmtp:18 annexb=no
a=rtpmap:101 telephone-event/8000
```

Fig. 9. 2nd portion of call trace report for transmitter end of VOS switch as AnnexB= no.

12	192.168.3.4:3060	192.168.4.3:3060	INVITE(SDP)	11:57:
13	192.168.3.4:3060	192.168.4.3:3060	1UU Trying	11:57:
14	192.168.8.12:3060	192.168.3.4:3060	1UU Trying	11:57:
15	192.168.3.4:3060	192.168.4.3:3060	403 Forbidden	11:57:
16	192.168.3.4:3060	192.168.4.3:3060	ACK	11:57:
17			send to...	11:57:
18	192.168.3.4:3060	192.168.4.3:3060	INVITE(SDP)	11:57:
19	192.168.3.4:3060	192.168.4.3:3060	1UU Trying	11:57:
20	192.168.8.12:3060	192.168.3.4:3060	1UU Trying	11:57:
21	192.168.3.4:3060	192.168.4.3:3060	503 Service Unavailable	11:57:
22	192.168.3.4:3060	192.168.4.3:3060	ACK	11:57:
23			send to...	11:57:
24	192.168.3.4:3060	192.168.2.8:3060	INVITE(SDP)	11:57:
25	192.168.3.4:3060	192.168.2.8:3060	1UU Trying	11:57:
26	192.168.8.12:3060	192.168.3.4:3060	1UU Trying	11:57:
27	192.168.3.4:3060	192.168.2.8:3060	183 Session Progress	11:57:
28	192.168.8.12:3060	192.168.3.4:3060	183 Session Progress(SDP)	11:57:

Fig. 8. 1st portion of call trace report for transmitter end of VOS switch as AnnexB= no.

So, noise will be generated which is undesired [12]. When we are decreasing RTP packet size for low network area, it is not possible to achieve good quality for audio codec G.729. If it is possible to decrease frame size audio codec G.729, then we may get it good voice quality. We are still not getting good voice quality at low network as AnnexB=no. But AnnexB=no is better than AnnexB=yes. If the audio packet size is less than 6.4kbps as AnnexB=no, that is possible for good voice quality at low band width area. It is also possible to decrease RTP packet size. So total bandwidth saving in this technique ranges from $(6.4-6/6.4) \times 100\% = 6.25\%$ to

IV. CONCLUSION

G.729 has been extended with various features, commonly designated as G.729a and G.729b. Dual-tone multi-frequency signaling (DTMF), fax transmissions, 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. Research for this article was done by means of a literature study and practical work knowledge in Bangladesh. The literature review aims to provide business management with a review of the development of VoIP in Asian countries. The literature review included a study of published and internet articles, books and literature on the management of information systems [13]. The research department of VOIP update is G.729 codec AnnexB=no if the audio packet size is less than 6.4kbps then a good voice quality is achieved in a low band width of 5kbps to 6kbps only that saves average bandwidth of approximately 14%.

APPENDIX

VOS switch, VPS switch, various dialer, SQL database, firewall system, STM device and CISCO router & switch, OS-Windows server and Linux server etc.

ACKNOWLEDGMENT

This research work is inspired and supported by Zamir Telecom Limited, Situated in 4-6 Davenant Street, Unit # A, London E1 5NB.

REFERENCES

[1] G.729 VoIP Compression Algorithm. [Online]. Available: <http://www.adaptivedigital.com/product/vocoders/g729.htm>

- [2] ITU-T Recommendation, "G.729 Annex B: a silence compression scheme for use with G.729 optimized for V.70," *Digital Simultaneous Voice and Data Applications*, vol. 35, no. 9, pp. 64-73, Sep. 1997.
- [3] Cisco Systems. [Online]. Available: http://www.cisco.com/en/US/tech/tk652/tk698/technologies_tech_note09186a0080094ae2.shtml
- [4] L. Cai1, Y. Xiao, X. S. Shen, L. Cai, J. W. Mark, and I. J. Commun. VoIP over WLAN: Voice capacity, admission control, QoS and MAC System. *Published online in Wiley Inter Science*. 19. pp. 491-508. [Online]. Available: <http://www.interscience.wiley.com>.
- [5] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: A Transport Protocol for Real-Time Applications Available. [Online]. Available: <http://ftp.ietf.org/rfc/rfc1889.txt>.
- [6] G.729. [Online]. Available: <http://en.wikipedia.org/wiki/G.729>
- [7] R. V. Cox and P. Kroon, "Low bit-rate speech Coders for multimedia communications," *IEEE Communication Magazine*, vol. 34, no. 12, pp. 34-41, December 1996.
- [8] G. Rec and B. Annex, "A silence compression scheme for G.729 optimized for terminal conforming," Tech. Rep., ITUT, Oct 1996.
- [9] F. Beritelli, S. Casale, and G. Ruggeri, "Performance evaluation and comparison of Itu-T/Etsi voice activity detectors," Viale Andrea Doria 695125 Catania.
- [10] Offer/Answer Considerations for G.723 Annex A and G.729 Annex B. [Online]. Available: <http://tools.ietf.org/html/draft-muthu-payload-offer-answer-g723-g729-00>
- [11] ITU-T Recommendation G.729 Annex B - "A silence compression scheme for G.729 optimized for terminals conforming to Recommendation V.70," November 1996.
- [12] Recognizing and Categorizing Symptoms of Voice Quality Problems. [Online]. Available: http://www.cisco.com/en/US/tech/tk652/tk698/technologies_white_paper09186a00801545e4.shtml
- [13] Voice over internet protocol (VoIP) *African Journal of Business Management*. 3. pp. 89-94. [Online]. Available: <http://www.academicjournals.org/AJBM>



N. A. Shafi was born in Rajshahi, Bangladesh on July 23th, 1985. He received the B.Sc. and M.Sc. degrees from the Department of Applied Physics and Electronic Engineering, Rajshahi University, Rajshahi, Bangladesh, in 2006 and 2007 respectively. He is currently working toward the System Engineer in the frame of the Network Operation Center, ZAMIR TELECOM LIMITED at Bangladesh support end on September 1st, 2009, Dhaka, Bangladesh. He is a member of the International Association of Computer Science and Information Technology (IACSIT). His current research interests include VOIP, Antenna, Microcontroller and Spintronics.



O. Farrok was born in Bangladesh, on 16 July, 1982. He received Bachelor and Master degree from the University of "Rajshahi University of Engineering & Technology" in the department of Electrical and Electronic Engineering (EEE), Rajshahi-6204, Bangladesh. He has been joined since 2007 and serving as an assistant professor in the Department of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology (AUST). His current research interests include Power electronics, Digital Electronics and Communication Engineering. He is a member of The Institution of Engineers, Bangladesh.

M. M. Ali was born in Bangladesh. He received the Bachelor degree from Engineering College, Rajshahi, Bangladesh. He received Ph.D. degree from Japan. He is currently working as a professor in Rajshahi University of Engineering & Technology in the department of Electrical and Electronic Engineering (EEE), Rajshahi-6204, Bangladesh. At present he is serving as the pro vice chancellor of this university. He is also a fellow of the Institution of Engineers Bangladesh and working as Vice-Chairman in the department of academic and human resource development.