Performance Evaluation of VoIP Services using Different CODECs over a UMTS Network

Jianguo Cao

School of Electrical and Computer Engineering
RMIT University
Melbourne, VIC 3000 Australia
Email: j.cao@student.rmit.edu.au

Mark Gregory

School of Electrical and Computer Engineering
RMIT University
Melbourne, VIC 3000 Australia
Email: mark.gregory@rmit.edu.au

Abstract - **In this paper, we evaluate the performance of Voice over Internet Protocol (VOIP) services that use different compression and decompression (CODEC) schemes, over a hybrid network that includes a Universal Mobile Telecommunications System (UMTS) network segment. We focus on the VoIP transmission end-to-end delay. We found that different CODECs provide very different results depending on the hybrid network. The research found that for VoIP services to operate over a hybrid network including a UMTS network segment, with an end-to-end delay comparable to that of circuit switched voice service, there is a requirement for careful comparison and design on choosing the CODEC scheme.**

I. INTRODUCTION

In recent years, Voice over Internet Protocol (VoIP) has attracted the attention of the network engineering research and operation communities. The phenomenal growth of VoIP is the result of rapid commercial solutions and network improvements. Other factors include the ongoing decrease in quality differences between existing Public Switched Telephone Networks (PSTN) telephony and VoIP [1] and the increase in bandwidth available to commercial and residential customers over which VoIP may be transported [2].

Most IP networks today were not designed for real-time, delay-sensitive voice or video traffic [3]. Current IP networks only provide a best-effort service and subsequently there is no guarantee that VoIP speech quality will be equivalent to what is provided by the existing PSTN telephony services.

VoIP services in wireless networks, such as UMTS [4], are being progressively implemented. The main benefit of VoIP over a wireless IP network is the mobility that is provided to users. Wireless networks have their own characteristics [5]. It is more challenging when the wireless network is combined with a real-time application's transmission requirement to meet the minimum Quality of Service (QoS). In order to gain a greater understanding of the VoIP transmission QoS, it is necessary to evaluate the performance of VoIP services over the hybrid wireless network.

End-to-end path delay has significant impact on the quality for the real-time delay sensitive VoIP services. In this research, the VoIP packet end-to-end delay will be investigated and analysed.

A number of encoder decoders (CODEC) are in common use today and each has different characteristics [6]. Transmission results may also vary when one CODEC is used and the information being transmitted is arranged into frames of different sizes. This effect is most noticeable when the segmentation of the transport stream occurs into Real-time Transport Protocol (RTP) packets. Generally, one or more voice frames can be put into one RTP packet. Figure 1 shows that RTP packets are then put into UDP packets and finally into IP packets prior to transmission across the network. The IP packets are encapsulated into MAC frames and transmitted from node to node. Delays are accumulated during the processing that occurs at each node across the path from source to destination.

The rest of the paper is structured as follows: (1) background information about VoIP, CODECS, QoS and end-to-end delay is provided in the next section, (2) in Section 3 the simulation environment for VoIP over UMTS and different CODEC schemes is discussed, (3) preliminary results are presented in Section 4, and (4) conclusions are presented in Section 5.

II. VOIP

In this section the literature review on VoIP, CODECs, QoS and end-to-end delay is provided.

A. CODEC

A CODEC is an algorithm used to encode and decode the voice stream. Since the voice stream is analog, it needs to be digitized so it can be transmitted over the Internet. Once it reaches the other end it needs to be decoded to restore the analog stream. There are a variety of different ways this encoding and decoding can be done.

Three popular CODECs are used and discussed in this paper. The CODECs are G.711, G729A and GSM. CODEC uses different algorithms to compress and decompress the voice stream and each CODEC will contribute a processing delay to the overall end-to-end delay [6].

B. VoIP Protocols

In a typical VoIP system, the voice stream is digitalized into voice frames. The voice frames are encapsulated and transmitted in RTP packets which are managed using RTCP which provides stream control and statistical information [7].

Signalling protocols have been introduced to provide overall management of VoIP calls. In the recent years, the Session Initiation Protocol (SIP) [8] has become very popular and is playing a major role in advancing VoIP. Figure 1 shows the protocol relationships in a SIP based VoIP system.

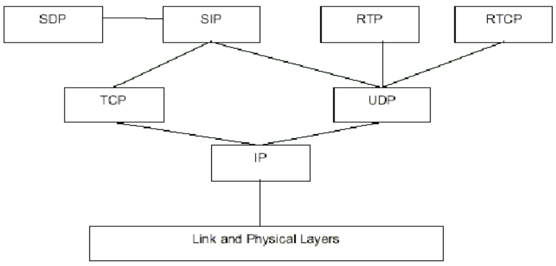

Figure 1. Protocol hierarchy for a SIP based VoIP system

IP packets used to transmit voice not only include the voice data, but also include protocol headers. The protocol headers are IP, UDP and RTP. An IPv4 header is 20 octets; a UDP header is 8 octets and an RTP header is 12 octets. The total length of this header information is 40 octets (bytes), or 320 bits, and the headers are sent with every packet containing voice data. The addition of headers as part of the encapsulation process adds to the bandwidth requirement for voice transmission.

C. *QoS*

In a PSTN network, phone calls are provided with a fixed bandwidth [9]. In a best effort packet network bandwidth is not fixed and may change over time [10]. When bandwidth drops in a best effort network for a voice call data packets may be lost and this causes voice call quality degradation. To achieve a minimum voice call QoS in a packet network it is necessary to apply a mechanism to control network parameters. The parameters used for VoIP QoS include end-to-end delay, jitter and packet loss [11]. The total end-to-end delay is the sum of a packet assembly at the source, the network delay and the receiver delay including packet disassembly. Jitter is the variance of packet arrival time due to different processing times and delays across the network [12]. Packet loss has a large effect on VoIP call QoS. Generally, a packet loss rate of 5% will annoy users and a higher packet loss rate can cause users to terminate a call. Packet loss may occur when there is congestion on the transmission path which causes the router buffers to overflow. The stability of the network heavily influences the packet loss rate [12].

D. *End-to-end delay*

Delay in VoIP networks is caused by propagation delay and processing delay [11].

Propagation delay is a characteristic of the transmission medium which may be fibre optic, copper medium or radio frequency.

Processing delay is caused by the network devices that handle VoIP traffic. Processing delay may be affected by the network device capacity and also the current network traffic at each device along a transmission path.

III. Hybrid Network Model

A hybrid UMTS model was used to simulate end-to-end delay. The UMTS functionality can be outlined into three groups [13]:

1. User Equipment (UE)
2. UMTS Terrestrial Radio Access Network (UTRAN). UTRAN consists of NodeB and Radio Network Controller (RNC), refer to Figure 2.
3. Core Network. The Core Network comprises two basic nodes: Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). GGSN provides internetworking with external packet switched networks such as IP networks.

In this research work Opnet Modeler 14.0 [14] was used to simulate the hybrid UMTS network and implement VoIP applications across the hybrid network i.e. combination of UMTS and IP network.

A. *Simulation Environment*

In Opnet Modeler, a simulation scenario was setup as shown in Figure 2.

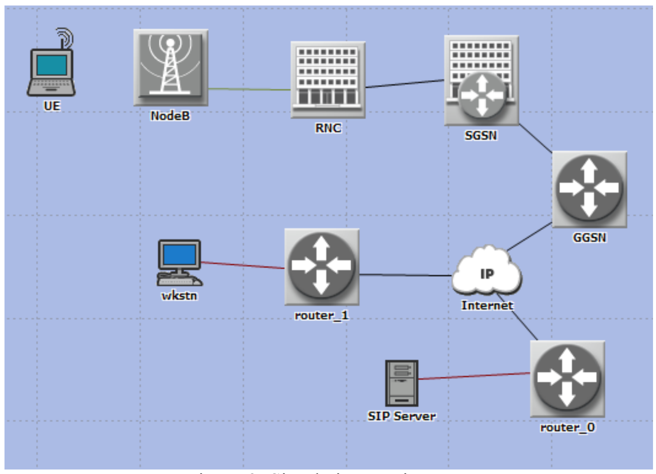

Figure 2. Simulation Environment

As shown in Figure 2 the UE, NodeB, RNC, SGSN and GGSN form a simple UMTS network which is connected to the Internet via the GGSN. A workstation and a SIP server are connected to the Internet using a router. The SIP Server is used as a signaling server to establish VoIP calls between the UE and the workstation.

B. *VoIP and CODEC Schemes*

VoIP calls between UE and the workstation were configured and the SIP server was used for call control. The simulation run-time was set to one hour, and during the simulation VoIP calls were made repeatedly with five-minute duration and three minute intervals.

The initial scenario was duplicated and different CODECs schemes were used for the VoIP calls between the UE and the workstation. Three CODECs (GSM FR, G.711 and G.729A) were used. Simulations were run for each CODEC with the frame size varied to 4 ms, 10 ms, 20 ms and 30 ms.

Additional simulations were run with the number of voice frames per packet varied. The same network was used for each of the different simulations.

In this research different frame sizes were used for the same CODEC to analyse the relationship between the frame size and the end-to-end delay. Table 1 summarizes the end-to-end delay assumption for each network component. The frame size of a CODEC can effect the NodeB scheduling and Hybrid Automatic Retransmission reQuest (HARQ) delay [15]. For a typical VoIP call, an acceptable end-to-end delay is less than 200 milliseconds.

TABLE I
SUMMARY OF END-TO-END DELAY COMPONENT

Delay component	Delay assumption
Voice encoder	Equal to CODEC frame size
NodeB scheduling + HARQ	Max. 110ms
NodeB ROHC, RLC+MAC processing Downlink propagation	30ms
UE scheduling + HARQ	40ms
UE processing, buffering, etc Uplink propagation	30ms
Backhaul delay (GSN-NodeB)	30ms
IP network delay	About 42ms

Normally the data rate for VoIP is less than 20kbps (except G.711), much smaller than the bandwidth UMTS can provide. If a very small VoIP packet is transferred in each MAC frame (2ms), the total available bandwidth would be under utilized.

As the number of VoIP packet is increased the amount of overhead grows by adding the IP/UDP/RTP headers. Research has been done in an effort to address this situation by aggregating several successively generated voice frames into one VoIP packet in order to improve bandwidth utilization [16].

IV. RESULTS ANALYSIS

In this section the simulation results will be presented and discussed. The simulations carried out were to identify if changes in the number of voice packets and CODECs affected the performance overall. The network configuration as shown in Figure 2 represents a typical network setup for voice calls.

A. Voice Frame Size and End-to-End Delay

In Figure 3, six simulation results are displayed. The results in Figure 3 show that for GSM-FR, G.711 and G.729A with frame size 4ms and 10 ms, the end-to-end delay results are similar, about 1580 ms delay, and not acceptable for VoIP calls. It is concluded that 4ms and 10ms frame sizes are not suitable for VoIP implementation over UMTS networks.

Figure 4 shows the results of GSM-FR with 20 ms and 30 ms frame sizes and Figure 5 shows the results of G.729A with the same frame sizes.

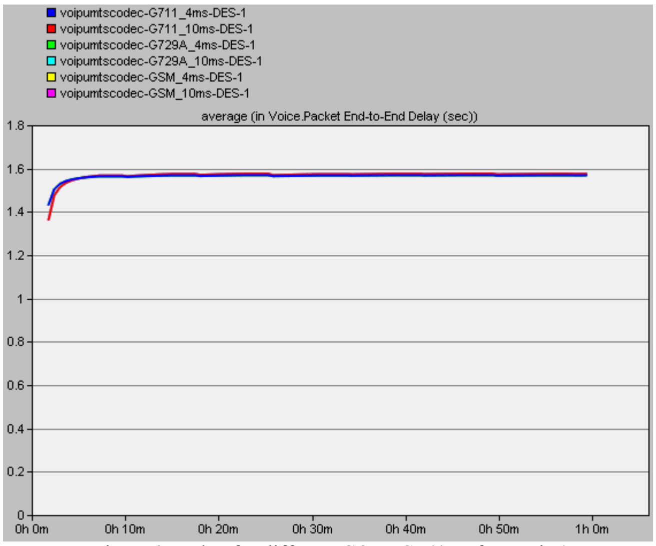

Figure 3. Delay for different CODECs (4ms frame size)

Figure 4 and Figure 5 show similar characteristics. From Figure 4 and Figure 5, it can be said that both GMS-FR and G.729A get less end-to-end delay when they are using 20ms frame sizes than 30 ms frame sizes. The actual end-to-end delay time is very close. Both CODECs are getting about 150ms end-to-end delay for a 20 ms frame size and 155ms end-to-end delay for a 30 ms frame size. The results shown in Figure 4 and Figure 5 are acceptable for VoIP calls.

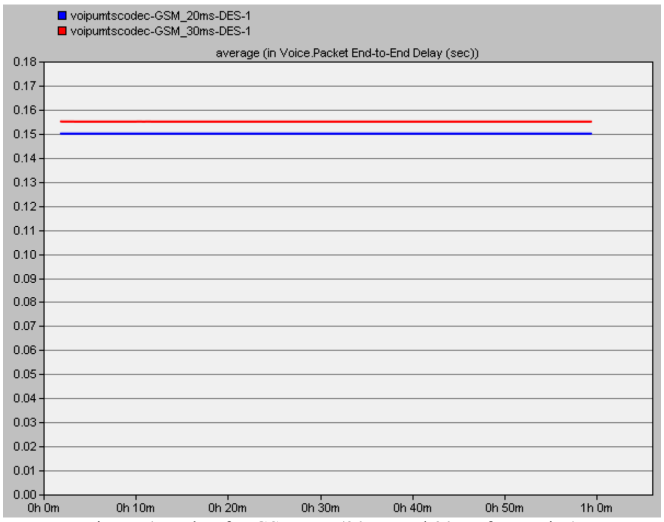

Figure 4. Delay for GSM-FR (20ms and 30ms frame size)

For the G.711 CODEC, the results found were different than that found for the GSM-FR and G.729A CODECs. Very high packet-loss rates were found in the G.711 CODEC simulations which had not been expected and may be considered in future research.

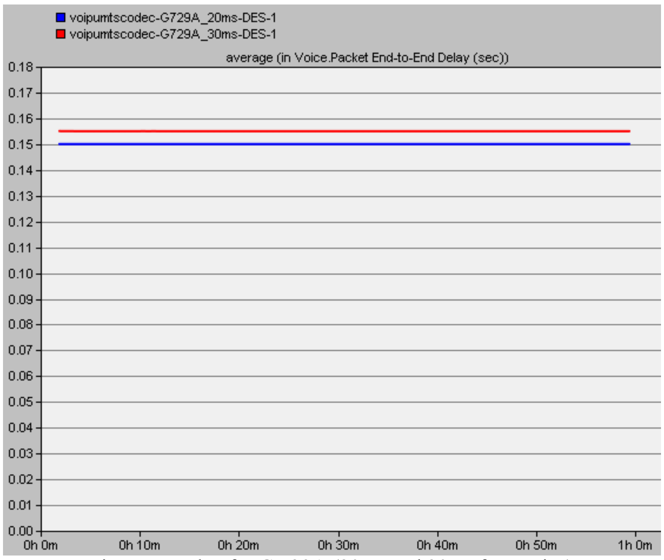

Figure 5. Delay for G.729A (20ms and 30ms frame size)

B. *Voice Frames per Packets and End-to-End Delay*

From the section, simulation results showed that voice calls with 20 ms frame sizes achieved better end-to-end delay results than voice calls with 30 ms frame sizes and much better results than for the other frame sizes. In this section voice calls with a 20 ms frame size will be used to investigate the relationship between end-to-end delay and voice call frames per packet.

Figure 6 shows the results for GSM-FR. Six simulations have been done using the hybrid network. The GSM-FR CODEC with a 20 ms frame size has been configured and used with a varied number of voice call frames in each VoIP packet. As shown in Figure 6, increasing the number of voice call frames per VoIP packet will increase the end-to-end delay. The 20 ms frame size still performs better than other length frame sizes with about 150 ms end-to-end delay.

Similar results were found for the G.729A CODEC. Figure 7 shows the results collected from the group of simulations using G.729A with a 20 ms frame size.

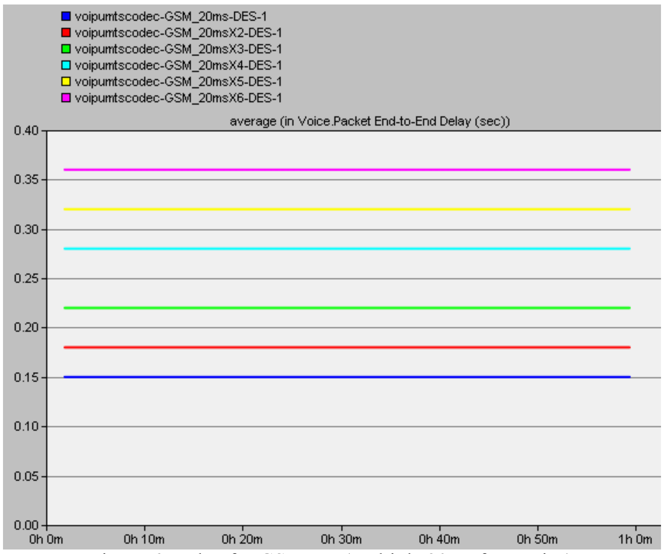

Figure 6. Delay for GSM-FR (multiple 20ms frame size)

From Figure 6 and Figure 7, the results found for the GSM-FR and G.729A CODECs identified that one voice call frame per packet will result in a lower end-to-end delay. Overall the results showed that as more voice call frames are added to each VoIP packet the end-to-end delay will increase.

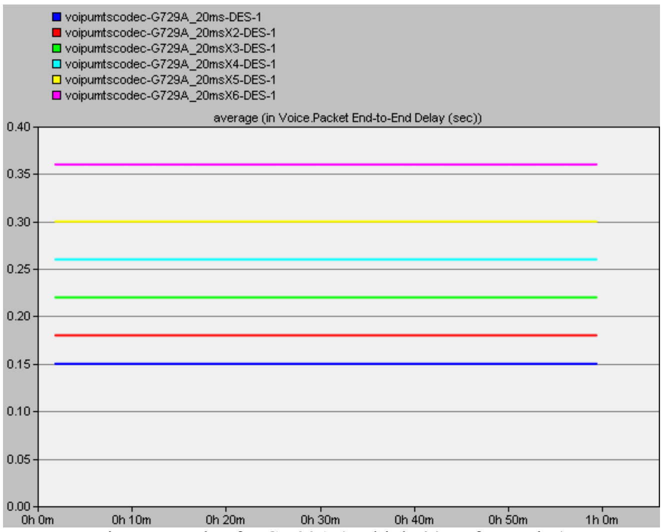

Figure 7. Delay for G.729A (multiple 20ms frame size)

V. CONCLUSION

In this research, voice call simulations were conducted using the one hybrid UMTS network to evaluate the performance of VoIP services with different CODECs. The VoIP packet end-to-end delays were examined for each configuration and compared to evaluate the VoIP call performance. The simulation results showed that there is a correlation between the end-to-end delay, the CODEC used and the number of voice frames in each VoIP packet. The G.711 CODEC was found to be unsuitable for VoIP over a hybrid UMTS network. For the GSM-FR and G.729A CODECs the best results were achieved for a 20 ms voice frame size and one voice frame for each VoIP packet. Overall, the simulations showed that for a hybrid UMTS network the selection of CODEC, the voice frame size and how the voice traffic is packaged into VoIP packets will affect the overall end-to-end delay.

VI. FUTURE WORK

In this research, the G.711 CODEC was found to perform very differently to GSM-FR and G.729A CODECs. To highlight this result, Figure 8 shows the voice traffic when a 30 ms frame size was used and Figure 9 represents the voice traffic when a 20 ms frame size was used.

Simulations using the G.711 CODEC with 4ms and 10ms frame sizes result in similar voice traffic with packet-loss rates about 50% as shown in Figure 8.

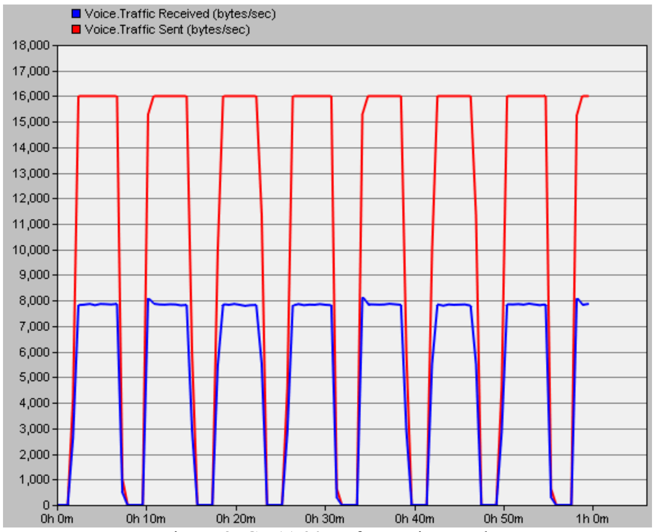

Figure 8. G.711 30ms frame size results

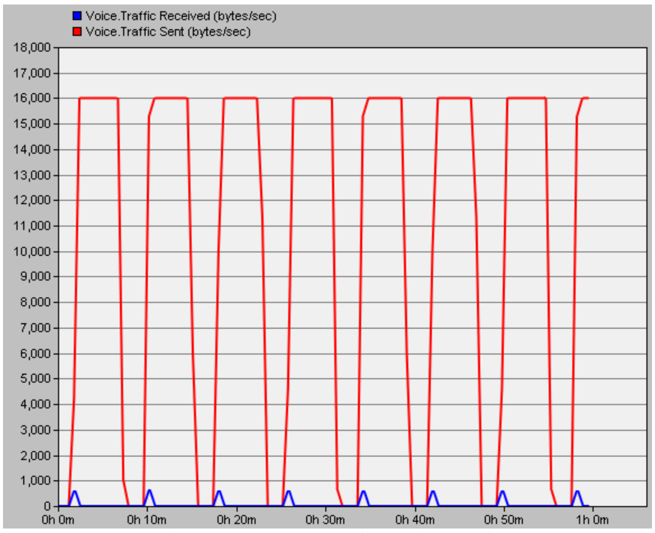

Figure 9. G.711 20 ms frame size results

For the G.711 CODEC with a 20 ms frame size varying the number of voice frames per VoIP packet did not affect the increased packet loss rate as shown in Figure 9.

REFERENCES

[1] T. Wallingford, "Switching to VoIP", Publisher: O'Reilly, ISBN: 0-596-00868-6, Pub Date: June 2005
[2] T. Kwok, "Residential broadband Internet services and applications requirements" Communications Magazine, IEEE Volume 35, Issue 6, June 1997 Page(s):76 - 83
[3] R. Dominach "Quality of service for real-time IP applications", TMCnet. March 2004
[4] The International Engineering Consortium, "UMTS Protocols and Protocol Testing", 2006
[5] A. Samukic, "UMTS Universal Mobile Telecommunications System Development of Standards for the Third Generation", 1998
[6] ITU-T P.830, "Subjective performance assessment of telephone-band and wideband digital codecs," 1996.
[7] H. Schulzrinne "RTP: A Transport Protocol for Real-Time Applications" (RFC3550), 2003
[8] IETF Network Working Group "SIP: Session Initiation Protocol" (rfc 3261), June 2002
[9] T. Wallingford, "Switching to VoIP", Publisher: O'Reilly, Pub Date: June 2005, ISBN: 0-596-00868-6
[10] K. Milutin, D. Damir, "IP telephony network saving capacity due to substitution of PSTN by IP network" Software in Telecommunications and Computer Networks, 2006. SoftCOM 2006. International Conference on Sept. 29 2006-Oct. 1 2006 Page(s):336 - 341
[11] J. Janssen et al., "Assessing Voice Quality in Packet-Based Telephony," IEEE Internet Computing, vol. 6, no. 3, 2002, pp. 48–57.
[12] F. Li "Quality and Reliability of Telephone Calls on the Internet Pre-study Report", 2001
[13] 3GPP, Network Architecture, TS 23.002, March 2000
[14] OPNET Modeler Radio, http://www.opnet.com
[15] 20053GPP TR 25.853, Delay Budget within the Access Stratum
[16] G. Rittenhouse and H. Zheng, "Providing VoIP Service in UMTS-HSDPA with Frame Aggregation," Proc. ICASSP,