



PERFORMANCE EVALUATION OF VOICE OVER IP USING MULTIPLE AUDIO CODEC SCHEMES

L. Audah¹, A. A. M. Kamal¹, J. Abdullah¹, S. A. Hamzah¹ and M. A. A. Razak²

¹Optical Communication and Network Research Group, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia

²Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

E-Mail: shipun@uthm.edu.my

ABSTRACT

The evolution of Voice over IP (VoIP) has made it one of the most popular applications over the wired/wireless Internet system due to its flexibility in technology integration and low cost of services. Telco and service operators have used the communication resources to optimize the VoIP architecture in order to provide better quality of service (QoS) to end consumers. The VoIP is a delay-sensitive traffic which requires minimum delay for general applications and minimum loss ratio for specific applications as the key QoS performance parameters. This paper compares the end-to-end (e2e) QoS performance parameters of VoIP codec schemes against multiple traffic connections transmitted over the Internet system. Background traffics are included in the simulations to closely match the real-world Internet scenario. Simulations analysis of bidirectional VoIP communications are done from the network layer perspective to compare the QoS performances of G.711, G.729A, G.723.1 and GSM.AMR codec schemes against the incremental of active connections in the network system. The results show that the G.729A produces at least 2.81% better in term of average accumulative e2e delay. The G.711 produces at least 21.89% better in term of average accumulative e2e jitter but produces the worst e2e packet loss ratio. In addition, GSM.AMR shows the best e2e effective transmission rate ratio ranges between 42.67% and 89.82%. This study has investigated the QoS performance variations of VoIP codecs so that the results could be used as guidelines to estimate the optimal network resources for various traffic requirements as early as in the design stage. As for future works, this study suggests the adaptive priority queue and packet scheduling at Internet gateway to regulate the traffic based on per flow QoS requirements.

Keywords: VoIP, audio codecs, QoS, internet, simulation, NS-2.

INTRODUCTION

Voice over IP (VoIP) applications have been very popular among the millions of Internet users for audio communications. The packet based framework of VoIP and low cost of packet data usage have made it the most preferable choice of future Internet communications to replace the conventional circuit switch telephone network. The flexibility of VoIP application to be implemented as software/hardware based and the ability to connect multiple devices simultaneously through wired/wireless channels have become one of the hot topics in the realm of Internet of Things (IoT). Big players from the telco operators and service providers have participated in the billion dollars market opportunities of VoIP applications for future smart-home and smart-cities.

In order to optimize the VoIP applications in future communications, various research works have been done from the physical to application layers. Previous research paper by (El-brak *et al.*, 2011) has compared the VoIP performance over small mobile ad-hoc network (MANET) using a pair of source and destination nodes. Besides that, (Ashouri *et al.*, 2014) has analyzed the performance of VoIP using different encryption methods over a local wireless network. Moreover, (Kim *et al.*, 2014) has studied the performance of VoIP QoS over long

term evolution (LTE) system from the user perspective by varying the speed, distance and number of mobile nodes. In addition, research by (Cocker *et al.*, 2014) has analyzed from the transport layer perspective for the buffer requirements over time and path taken by VoIP traffic over the Internet system. None of the related research has compared the QoS performance of VoIP codec schemes from the network layer perspective and its robustness against various competing traffics.

This paper analyzes and compares the e2e QoS performance parameters of bidirectional VoIP (i.e. delay, jitter, loss ratio and throughput) using different types of audio codecs mainly used for VoIP applications (i.e. G.711, G.729A, G.723.1 and GSM.AMR). Simulations analyses have been done from the network layer perspective using multiple connections of VoIP and background traffics to closely match the real-world Internet system scenario. The simulation results provide insight on the performance comparison of each VoIP codec used for audio communication over the Internet system. The remainder of this paper is organized as follows: Section II explains the NS-2 simulation configuration. The results and analysis are discussed on Section III. Finally, Section IV concludes the findings and suggests future research works.



SIMULATION SETUP

The simulation setup characterizes a simple ubiquitous e2e Internet system as depicted in Figure-1. The network simulations are done using NS-2.34 and AWK programming tools. The next subsections explain the details of network elements parameters involved in the simulations.

Network configuration

The NS-2 network configurations consist of 2 VoIP remote LANs and 2 background traffics remote LANs

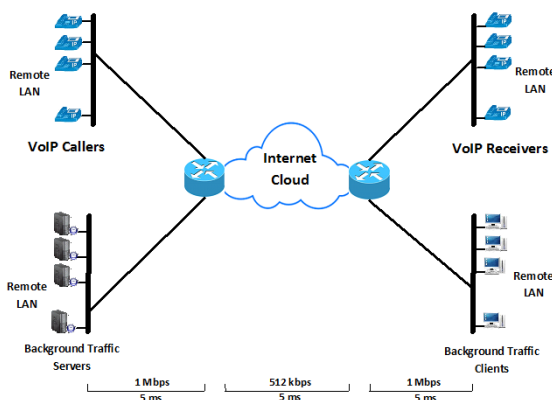


Figure-1. NS-2 network simulation scenario.

connected via the Internet. Each remote LAN is connected to the Internet gateway router via 1 Mbit/s full-duplex link with 5 ms of link delay. The Internet network cloud is represented with a 512 kbit/s full-duplex link and also 5 ms of link delay. Each link interface contains a simple Droptail queue with a 50 packet buffer size.

The link bandwidths are set to such small scale values to represent a much larger broadband network system because of the computer hardware/software and NS-2 simulations time limitations. The network system design does not include the bit-error-rate (BER) for link error model because it is assumed that the BER for wired broadband network is very low and could be neglected.

APPLICATION MODELLING

There are 2 types of application services used in the NS-2 simulations which are the bidirectional VoIP with multiple codec schemes and also the background traffic. Multiple connections are generated from each traffic type during 500 s of a simulation time. The simulations are done separately for 4 different VoIP codec which are the GSM.AMR, G.711, G.723.1 and G.729A. Simulations for each VoIP codec are done for 10 times with incremental of VoIP and background traffic connections. The VoIP connections are increased by 50

between 50 and 500 connections while the background traffic increased by 5 between 5 and 50 connections for the simulations of each VoIP codec type. Details of VoIP and background traffic configurations will be explained in the next subsections.

Bidirectional VoIP

The bidirectional VoIP application system model as in (Bacioccola *et al.*, 2007) and (Andreozzi *et al.*, 2010) is used for the simulations in NS-2. Standard codec type like GSM.AMR (Sjoberg *et al.*, 2007), G.711 (ITU-T G.711, 1988), G.723.1 (ITU-T G.723.1, 2006) and G.729A (ITU-T G.729, 2012) are used at the source side to encode the Voice Activity Detection (VAD) (e.g. talkspurt and silence events) into a few chunk of bytes before sending the data over the Internet network using UDP transport protocol (Schulzrinne and Casner, 2003). The talkspurt and silence events between the communicating VoIP clients are not independent and modelled using modified Brady's model.

The bidirectional audio streams between caller and listener can be in any of four states which are single talk, double talk, short silence and mutual silence (Bacioccola *et al.*, 2007) and (Andreozzi *et al.*, 2010).

Table-1 shows the codec's sample size and bit rate used in the simulations as suggested in (Schulzrinne and Casner, 2003). The VoIP sample size data are excluding the IP/UDP/RTP headers size which approximately 40 bytes. At the receiver side, the transmitted audio packets are demultiplexed and inserted into a H.323 playout buffer to ensure that the packets are decoded at the same interval at which they are generated by the encoder. The H.323 protocol provides real-time audio communications on point-to-point, multipoint or broadcast network configurations (ITU-T H.323, 2009).

Background traffic

Multiple connections of the background traffic are used in the NS-2 simulations to represent the other types of Internet applications that shared the network resources with VoIP application. The background traffic is modelled as in (Harfoush, 2000) by using a Pareto On/Off UDP source. The Pareto shape parameter (α) is set to 1.2. The traffic is generated using constant packet size of 200 bytes with 2 s of burst time and 1 s of idle time. The packets are transmitted at 3.6 kbit/s bit rate.

RESULTS AND ANALYSIS

Each NS-2 simulation is carried out for the duration of 500 s. The simulations are done 10 times (i.e. 10 average values of new VoIP connection inter-arrival time) for each VoIP codec. The simulations results and analysis are divided into 4 QoS categories which are the average e2e delay, jitter, packet loss ratio and throughput. The QoS parameters are calculated based on each simulation output trace file using AWK programming



script and then presented in the form of graphs as shown in the next sub-sections.

Average end-to-end delay

Average e2e delay is the common QoS parameter used to describe the level of service interactivity and smooth playback. The estimated e2e delay (one-way) for a VoIP connection could be estimated using equation (1).

$$D_t \approx D_p + D_q + D_n \quad (1)$$

where D_t is the one way packet delay from VoIP caller to receiver, D_p is the one way propagation time, D_q is the queuing delay and D_n is the additional delay due to traffic and other network factors.

Table-1. VoIP codec sample size and bit rate.

Codec Type	Sample Size (Bytes)	Bit Rate (kbit/s)
GSM.AMR	32	4.75 – 12.2; Toll quality speech = 7.4
G.723.1	24	6.3
G.711	80	64
G.729A	20	8

The one way propagation delay (D_p) could be estimated by summing up all links delays from caller to receiver. Based on the simulation scenario in Figure 1, the one way propagation delay from caller to receiver is 15 ms (i.e. delay from caller to Internet gateway + propagation delay within the Internet network cloud + delay from the Internet gateway to receiver). The queuing delay in this case mainly occurs at the Internet gateway which is calculated using equation (2).

$$D_q = \frac{N \times P \times 8}{B_l} \quad (2)$$

where N is the current total number of packets in queue buffer, P parameter refers to the packet size in bytes while B_l is the link bandwidth measured in bit per second (bit/s). D_q parameter for each of the Internet gateway is $1.56NP \times 10^{-5}$ second depending on codec types for N and P parameters if we includes the Internet link bandwidth of 512 kbit/s. In addition, the D_n parameter is the delay due to the packet transmission time and also the network dynamic factors like congestion and link failure.

The average e2e delays for all VoIP connections in a simulation, D is then calculated by summing up all of the one way connection delay and then divides with the total number of established VoIP connections (i.e. n

parameter) during the simulation time as shown in equation (3).

$$D = \frac{\sum_{i=1}^{i=n} (D_t)_i}{n} \quad (3)$$

Figure-2 shows the average end-to end delay against average number of established VoIP connections in a simulation. The lower the delay the better would be the transmitted voice quality across the network system. The delays are proportional to the increment of average VoIP connections. The more new connections established between caller and receiver, the higher would be the delay for each type of VoIP codec. The average delays for all VoIP codecs except G.711 remain below 25 ms if the average generated VoIP connections within the network system are between 50 and 150. After that, the delays increased and reach the maximum for approximately 121 ms. In contrast, the maximum delay for G.711 VoIP codec is 244 ms.

The average accumulative delays for GSM.AMR, G.723.1, G.711 and G.729A codecs are 81.92 ms, 69.16 ms, 196.7 ms and 67.22 ms respectively as shown in Figure-3. All VoIP codecs satisfy the International Telecommunication Union (ITU) delay recommendation to be less than 150 ms for best audio reception quality except G.711 codec (Demichelis and Chimento, 2002). One of the main reasons is that the sample size for G.711 is 80 bytes excluding the 40 bytes of IP/UDP/RTP headers.

This codec sample size is approximately 2.5 times more than the others. The other reason is that the G.711 codec utilizes A-law/ μ -law companding algorithm with low compression rate which eventually accommodate larger link bandwidth fraction compared to the other codecs.

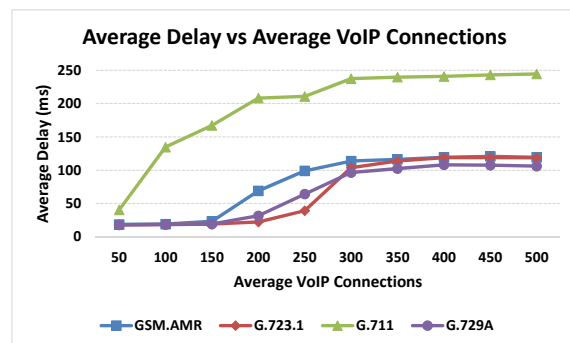


Figure-2. Average e2e delay.



Average end-to-end jitter

Jitter is the e2e one way delay variation between packets transmitted from caller to receiver by ignoring any lost packets (Szigeti and Hattingh, 2014). Jitter causes the packets to arrive at different timing and possibly in different order. At certain level, jitter may cause audio anomalies like stuttering, uneven audio and abnormal speech rhythm. Equation (4) is the general equation used to calculate jitter per connection.

$$J(i+1) = J(i) + \left| (R(i+1) - S(i+1)) - (R(i) - S(i)) \right| \quad (4)$$

where

$S(i)$ – Time at which packet ‘i’ is transmitted from the caller.

$R(i)$ – Time at which packet ‘i’ is received at the receiver.

The average e2e jitter for all VoIP connections in a simulation, J is then calculated by summing up all of the one way connection jitter and then divide with the total number of established VoIP connections (i.e. n parameter) during the simulation time as shown in equation (5).

$$J = \frac{\sum_{i=1}^{i=n} (J(x))_i}{n} \quad (5)$$

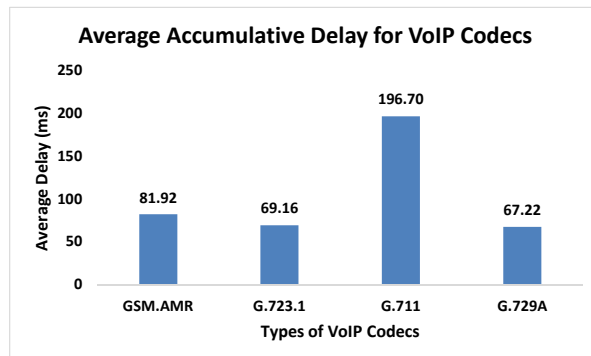


Figure-3. Average accumulative e2e delay of VoIP codecs.

The average jitters shown in Figure-4 are steadily increased between 50 and 250 VoIP connections for all codec types. After that the values maintain between 4.0 ms and 5.3 ms except for G.711 which decrease and maintain between 2.2 ms and 2.3 ms after 300 VoIP connections. Besides that, the average accumulative jitter for GSM.AMR, G.723.1, G.711 and G.729A codecs are 3.86 ms, 3.93 ms, 2.89 ms and 3.70 ms respectively as shown in Figure-5.

There are no specific VoIP jitter limitation specified by the ITU but it is strongly advisable to be less than 30 ms for one way e2e average jitter when designing the VoIP communications over LAN, WAN and VPN systems (ITU-T G.114, 2003). In the simulations results as in Figure-4, the networks becomes congested and saturated with many competing VoIP and background traffics after 300 VoIP connections. The decrement of jitter values for G.711 codec after 300 connections is because too many packets have been dropped in congested links as the network moves towards saturation point and the jitter counted in the simulations ignored the loss packets. In addition, larger packet size as in G.711 codecs has severely affected by network dynamics compare to smaller packets size like in the other codec types. Although playout buffer is used in this VoIP system to dampen the jitter and rearrange the out order packets but the severity of congestion at the bottleneck link has degraded the jitter performance.

Average end-to-end packet loss ratio

The average e2e packet loss ratio over the Internet system is another important QoS performance parameters for VoIP. Packet loss ratio is the ratio of total packet loss over total packet sent from caller to receiver. The packet loss ratio for a VoIP connection in a simulation is measured as in equation (6).

$$L(i) = \frac{\sum_{j=1}^{j=y} (P_l)_j}{\sum_{k=1}^{k=z} (P_s)_k} \quad (6)$$

where $\sum P_l$ is the total packets loss and $\sum P_s$ is the total sending packets from the caller during a VoIP connection session. By considering all generated VoIP connections (n) in a simulation run time, the average e2e packet loss ratio is calculated as in equation (7).

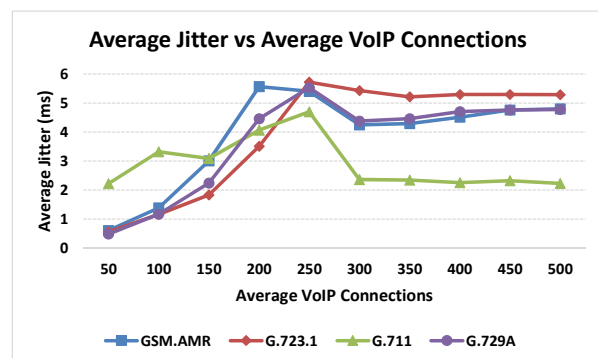


Figure-4. Average e2e jitter.

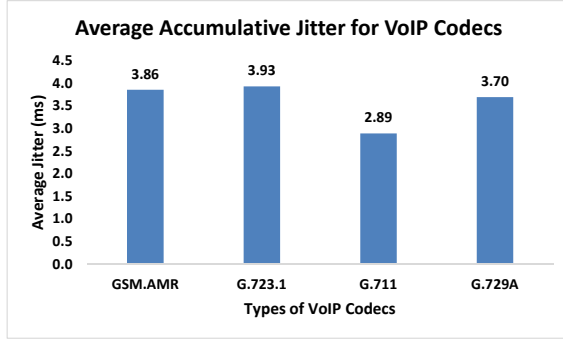


Figure-5. Average accumulative e2e jitter of VoIP codecs.

$$L = \frac{\sum_{i=1}^{i=n} L(i)}{n} \quad (7)$$

Figure-6 shows the average e2e packet loss ratio against average VoIP connections generation in a simulation run time for all codec types. The average e2e packet loss ratio is proportional to the increment of average generated VoIP connections in the network system. The lower the loss ratio the better would be the transmitted audio quality. The average e2e packet loss ratio for all codec types except G.711 and G.723.1 remain almost constant below 1.2 % for average connection between 50 and 150 depending on the codec types.

Only the G.723.1 codec successfully maintains packet loss ratio below 1% until 250 connections while the G.711 suffers packet losses between 1% and 64.7%. After that, all codecs shows rapid increment of packet loss between 250 and 450 connections and then remains almost constant between 48.4% and 79.0%. In addition, the average accumulative packet loss ratio as shown in Figure 7 for GSM.AMR, G.723.1, G.711 and G.729A are 27.48%, 20.76%, 56.21% and 23.15% respectively.

The packet loss occurs when the queue buffer on network link becomes overflow as the results of network congestion. The packet loss event is not a critical factor for the delay-sensitive traffic like VoIP compared to the throughput-sensitive traffic. However, in order to transmit high quality audio over the network system it is necessary to keep the losses at minimum. The ITU does not specifically mention the standard of maximum packet loss for VoIP. However, the ITU has suggested that the loss ratio should be less than 1% for a VoIP connection in order to achieve the best audio quality (ITU-T G.114, 2003). If the loss ratio is much higher, the VoIP connections may experience audio anomaly like echo, stuttering and etc. The G.723.1 could achieve that goal for average connections below 250 while the others except G.711 only below 150. The G.711 codec transmission must maintain below 50 connections in order to achieve best audio quality.

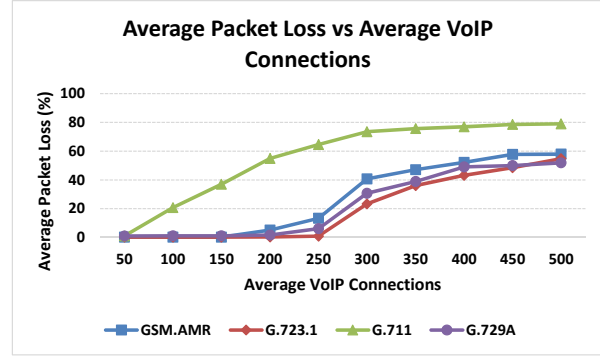


Figure-6. Average e2e packet loss ratio.

Apart from the network dynamic conditions, packet loss occurs because of the VoIP codec unique characteristics. The good performance of G.723.1 in this network simulation scenario is also due to high packet compression ratio, low transmission bit rate and lossy commanding algorithm utilization. In contrast, the low performance of G.711 codec might be due to its built in framework design which utilizes the A-law/ μ -law commanding algorithm in order to deliver precise speech transmission. The codec also produce higher bit rate compare to the others which is about 64 kbit/s. However, the codec is more sensitive towards packet losses due to poor packet loss interpolation.

Average end-to-end throughput

The e2e throughput concludes the previous QoS performance parameters as it measures the rate of successfully received packets at the receiver. The throughput might not be as critical as the delay and jitter QoS parameters for the delay sensitive traffic like VoIP but it may provide references in designing the e2e network system with QoS to accommodate high quality voice traffics.

The average e2e throughput in bit/s for all VoIP connections (n) in a simulation is calculated using equation (8).

$$T = \frac{\sum_{i=1}^{i=n} \left(\frac{P_b \times 8}{t_r - t_c} \right)_i}{n} \quad (8)$$

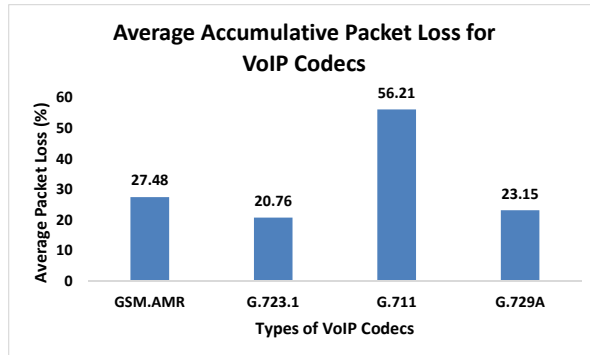


Figure-7. Average accumulative e2e packet loss ratio of VoIP codecs.

where P_b is the total received packets at the receiver in bytes, t_r is the packet received time and t_c is the packet sending time from the caller for a VoIP connection.

Figure-8 shows the average e2e throughput for average VoIP connections between 50 and 500. The throughput values for all VoIP codecs except G.711 remain almost constant between 3.6 kbit/s and 6.6 kbit/s for average connections between 50 and 250 depending on codec types. After that, the throughputs slightly decrease and then remain almost constant again between 2.0 kbit/s and 3.68 kbit/s for average VoIP connections between 400 and 500. In contrast, the G.711 codec shows rapid decrement from 28.0 kbit/s to 7.5 kbit/s between 50 and 300 average connections and then remain almost constant between 6.0 kbit/s and 6.97 kbit/s. In addition, the average accumulative throughput as in Figure-9 for GSM.AMR, G.723.1, G.711 and G.729A codec types are 5.13 kbit/s, 2.97 kbit/s, 12.52 kbit/s and 4.34 kbit/s respectively. The G.711 codec achieved the highest average throughputs for all VoIP connections while the G.723.1 codec achieved the lowest compare to the others.

The average e2e throughput patterns are inversely proportional to the increment of average VoIP connections. There are 2 main factors that affecting the throughput which are the delay and packet losses. The delay and packet loss results for all codec types except G.711 have shown significant increment after 200 average VoIP connections generated within the network during a simulation time. Consequently, the throughputs show significant decrement and then remain almost constant until 500 connections depending on the codec types. Since the new VoIP connections are continuously generated without waiting for the previous connections to be completed, more VoIP connections will compete with the background traffic to access the network in best-effort manner. At certain point of time, the network becomes congested and then the packets need to queue for network access which eventually increases the delay. When the buffer is full, the packets start to be dropped and reduce the number of successful received packets. As for the G.711 codec, the rapid decrements of throughput between

50 and 300 connections not only due to the delay and packet loss effects but also due to the codec's unique specification with larger codec sample size, short interval time between packets and also lossless data compression method utilization. Simulations results have shown that those characteristics severely degrade the G.711 codec performances in certain high density network traffic system.

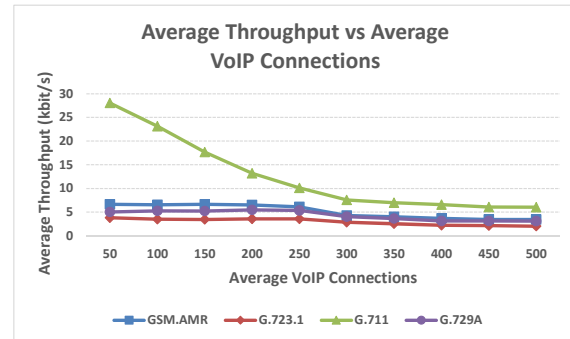


Figure-8. Average e2e throughput.

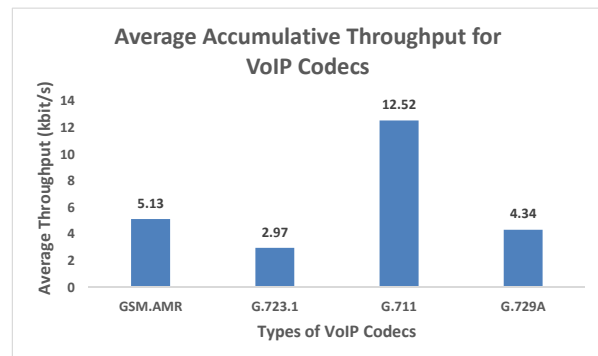


Figure-9. Average accumulative e2e throughput of VoIP codecs.

Table-2. VoIP codec effective transmission rate ratio.

Codec type	Effective transmission rate ratio	
	Min (%)	Max (%)
GSM.AMR	46.72	89.82
G.723.1	32.11	60.54
G.711	9.43	43.79
G.729A	38.72	68.10

This paper also estimates the effective transmission rate as the ratio of measured throughput over source bit rate in term of percentage. Table-2 shows the range of minimum and maximum effective transmission rate ratios for all codec types. The minimum and



maximum throughput data are collected from the NS-2 simulations and then compare against the codec source bit rate. The highest range of values is achieved by GSM.AMR codec which are between 46.72% and 89.82% while the lowest is G.711 codec which are between 9.43% and 43.79%. The effective transmission rate does not dictate that certain VoIP codec are the best or the worst but rather provides information to decide on which application does the codec can be used optimally (e.g. wired or wireless network, high or low audio quality, broadband or narrowband applications, etc.).

CONCLUSION AND FUTURE WORK

The simulations study shows the e2e QoS performances for 4 different VoIP codec types transmitted over the best effort network system. The network scenario is designed with the standard parameters without specific QoS improvement modifications in order to closely match the real-world VoIP applications (i.e. multiple VoIP and background traffic connections are transmitted over the network system to produce more realistic services behaviors).

Firstly, the study shows that the G.729A produces the best performance in term of lower average accumulative e2e delays while the G.711 codec shows the opposite. All codec types except G.711 produce delays below 150 ms for all generated VoIP connections as recommended by the ITU for high quality audio communications. The G.711 mouth-to-ear delays are within the acceptable range and below the upper bound of 400 ms for general network planning. Secondly, the study shows that the G.723.1 produce the least jitter at lower average e2e VoIP connections and the highest jitter after 250 generated VoIP connections while the G.711 shows the opposite. The jitters for all codec types are below the recommended 30 ms of acceptable audio quality. Thirdly, the e2e loss ratios for all codec types except G.711 are less than 1% for average VoIP connections below 150. The G.723.1 codec shows the lowest of overall losses while the G.711 shows the highest after 50 average connections. For high quality audio quality transmission, it is recommended to keep the packet loss ratio to be below than 1%.

Fourthly, the G.711 codec shows the highest e2e throughput while the G.723.1 shows the opposite. The G.711 is mostly affected by the increment of VoIP connections while others show slow variation. This might be due to the differentiation of codec's source bit rate characteristic. In term of average throughput over bit rate ratio, the GSM.AMR shows the highest performance which is between 46.72% and 89.82% of minimum and maximum values respectively. Fifthly, the overall QoS performances of G.729A in term of average e2e delay, jitter and loss ratio are better than the GSM.AMR in most cases of network dynamics.

This paper does not dictate the best VoIP codec for all types of applications but rather to provide a

guideline for future researchers to design better QoS performances of VoIP system application. The future works aims at designing a QoS aware routing, scheduling and queuing schemes for Next Generation Internet (NGI) network system that can adaptively maintains each traffic type QoS requirements at optimum level. In addition, the VoIP protocols could also be enhanced with adaptive coding and modulation to suite with the network dynamic characteristics.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Education Malaysia for the generous financial support under Research Acculturation Grant Scheme (R054).

REFERENCES

- Andreozzi, M. M., Migliorini, D., Stea, G. and Vallati, C. 2010. Ns2Voip++, an advanced module for VoIP simulations. Proceedings of 3rd International ICST Conference on Simulation Tools and Technique, SIMUTools '10, Malaga, Spain, article no. 50.
- Ashouri, S., Khosraviroshkhari, H., Seno, S. A. H and Moghaddam, M. H. Y. 2014. VoIP performance comparison in wireless networks with different encryption methods. Proceedings of the 4th International Conference on Computer and Knowledge Engineering (ICCKE 2014), Mashhad, Iran, pp. 753-757.
- Bacioccola, A., Cicconetti, C. and Stea, G. 2007. User-level performance evaluation of VoIP using ns-2. Proceedings of 2nd Performance Evaluation Methodologies and Tools, ValueTools '07, Nantes, France, article no. 20.
- Cocker, E., Ghazzi, F., Speidel, U., Dong, M. C., Wong, V., Vink, A. J. H., Yamamoto, H., Yokoo, H., Morita, H., Ferreira, H., Emle, A., McFadzein, R., Palelei, S. and Eimann, R. (2014). Measurement of buffer requirement trends for real time traffic over TCP. Proceedings of the 15th IEEE International Conference on High Performance Switching and Routing (HPSR 2014), Vancouver, Canada, pp. 120-124.
- Demichelis, C. and Chimento, P. 2002. IP packet delay variation metric for IP performance metrics (IPPM). RFC 3393.
- El-brak, S., Bouhorma, M., and Boudhir, A.A. 2011. VoIP over MANET (VoMAN): QoS & performance analysis of routing protocols for different audio codecs. International Journal of Computer Applications, 36(12), pp. 22-26.



Harfoush, K., Bestavros, A. and Byers, J. 2000. Robust identification of shared losses using end-to-end unicast probe. Proceedings of the 8th IEEE International Conference on Network Protocols, ICNP 2000, Osaka, Japan, pp. 22-33.

ITU-T Recommendation G.114 2003. One-way transmission time.

ITU-T Recommendation G.711 1988. Pulse code modulation (PCM) of voice frequencies.

ITU-T Recommendation G.723.1 2006. Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s.

ITU-T Recommendation G.729 2012. Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP).

ITU-T Recommendation H.323 2009. Packet-based multimedia communications systems.

Kim, J., Niyaz, Q. and Javaid, A.Y. 2014. Performance evaluation of VoIP broadcasting over LTE for varying speeds and distances of mobile nodes. Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB 2014), Beijing, China, pp. 1-5.

Schulzrinne, H. and Casner, S. 2003. RTP profile for audio and video conferences with minimal control. RFC 3551.

Sjoberg, J., Westerlund, M., Lakanien, A. and Xie, Q. 2007. RTP payload format for the adaptive multi-rate (AMR) and adaptive multi-rate wideband (AMR-WB) audio codecs. RFC 4867.

Szigeti, T. and Hattingh, C. 2014. QoS design overview. End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs. 1st ed. Indianapolis, IN: Cisco Press, ch. 2, pp. 33-38.