

Performance Analysis of VoIP Data over IP Networks

Dariusz Strzëciwilk

Abstract—The paper presents the results of research and analysis of voice data transmission quality in IP packet networks. It analyses mechanisms allowing for the assessment of packet telephony data transmission quality. Possible transmission quality levels and adequate quality metrics, applicable in the recommendations of standardisation organisations, as well as suggested limit values conditioning acceptable voice data transmission quality were indicated and discussed. A packet network model was designed and tested, taking into account VoIP architecture supporting various audio codecs used for voice compression. Transmission mechanisms based on audio codecs G.711, G.723, G.726, G.728 and G.729 were investigated. It was shown that for delay-sensitive traffic which fluctuates beyond its nominal rate, selected codecs have an advantage over others and allow for better transmission quality of VoIP traffic with guaranteed bandwidth and delay.

Keywords—MOS, VoIP, RTCP, QoS, audio codecs, transmission quality

I. INTRODUCTION

MODERN Internet based on packet data transmission provides us with many complex multimedia communication services. Examples include VoIP (*Voice over IP*) telephony, IPTV (*Internet Protocol Television*), video conferencing, P2P exchange of files containing multimedia content, VoD (*Video on Demand*). VoIP packet telephony has become a common method of making voice calls. Voice over IP communication gains still greater importance in telecommunications industry. The advantages of VoIP technology are numerous, ranging from the convergence of two network infrastructures into one to lower or no voice call prices. Packet telephony is willingly used by operators in the backbone of the network, which is due to the high efficiency of resource use, i.e. bandwidth. The growing importance of VoIP Internet telephony technology and the increase in multimedia data transmitted over IP packet networks means that voice data transmission may deteriorate. Hence, in order to implement VoIP services with a full QoS (*Quality of Service*) guarantee, the data transmission technologies used must be carefully examined. QoS transmission and security mechanisms are among the most important challenges arising during the design and maintenance of modern computer networks and NG (*Next-Generation*) networks [1], [2]. Guaranteeing adequate quality of service is of particular importance for real-time applications, such as data transmission in VoIP networks [3], [4]. These

services are particularly delay-sensitive and require guaranteed bandwidth [5]. VoIP applications and systems can choose from a wide variety of speech codecs, which vary in bandwidth, quality of speech listened to, and resistance to quality degradation in case of packet loss. VoIP codecs also exhibit differences in aspects such as computational complexity of compression algorithms or traffic generated, which in turn affect energy consumption and CPU usage [6]. When implementing VoIP networks and services, customer expectations in terms of voice quality must be considered and met. One of the most important challenges that arise during the design and maintenance of both modern computer networks and next generation networks is the quality of service QoS [7]. Current IP networks are based on BE (*Best Effort*) services. Queuing of traffic in IP networks and lack of stringent QoS control can cause packet loss, delay and jitter, which directly affect the quality of VoIP services. The current IP network architecture needs to be enhanced with mechanisms that guarantee QoS especially for VoIP applications. Hence, the aim of the research was to investigate how VoIP transmission affects voice quality. Moreover, the aim of the research was to analyse and investigate the transmission quality of audio codecs used for voice compression in computer networks supporting VoIP technology. An attempt was also made to determine to what extent selected audio codecs affect the quality of transmitted voice. Popular audio codecs G.711, G.723.1, G.726, G.728 and G.729 were used in the study. The computer network supporting VoIP technology was built using Cisco 2900 routers and CME (*Call Manager Express*) software, which allowed the control and management of calls between VoIP phones.

II. VOICE TRANSMISSION AND QOS

The introduction of the service of speech transmission via packet transmission, popularly known as Voice over IP (VoIP) to the Internet is currently a very popular method of making voice calls. It consists in sending voice in IP packets over the Internet. A specification describing Internet telephony is presented in document RFC 741 [8]. Internet telephony is a very broad and extensive issue. Details on the implementation of VoIP technology can be found in many works (see [9]–[11]). The quality of voice calls is essentially affected by two groups of factors that depend on the mechanisms used during packet voice transmission. The first group includes parameters of voice data transmission such as packet loss, delay and jitter (delay variation). The second group includes factors that

D. Strzëciwilk is with Institute of Information Technology, University of Life Sciences, Warsaw, Poland (e-mail: dariusz_strzëciwilk@sggw.pl).



© The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>), which permits use, distribution, and reproduction in any medium, provided that the Article is properly cited.

depend on the codec used, mechanisms counteracting packet loss (PLC), mechanisms counteracting transmission distortion (FEC) and dynamic or static buffer jitter. Excessive jitter causes audio to be played at a variable rate. The use of a packet buffer on the receiving end counteracts its effect. Guaranteeing quality of service is particularly important for real-time applications such as VoIP voice transmission. From the perspective of the computer network, VoIP systems perform two functions - controlling calls and transferring voice. Voice transmission in the VoIP technology requires guaranteeing appropriate conditions for proper data transmission. To match the quality of the traditional (analogue) voice signal, the digital (packet) data stream must be transmitted very quickly and reliably. In addition, the packets must arrive in the same order in which they were emitted. Increased speed is guaranteed by the use of the UDP protocol in packet networks. However, the UDP mechanism does not take error checking into account. Furthermore, it does not ensure that packets are delivered in the correct order. It is therefore necessary to use QoS mechanisms to allow voice data to be transmitted in the first order. The implementation of QoS mechanisms in VoIP networks can be based on the marking of the 8-bit TOS (*Type of Service*) field for data transmitted using the IPv4 protocol (See Fig. 1) and the TC (*Traffic Clas*) field when using the IPv6 protocol [12]. The mechanisms and uses of the TOS field [13] (IPv4) and the Traffic Class octet (IPv6), which take the name DS (*Differentiated Services*) are described in RFC 2474 [14]. The marking of voice packets is closely related to the DS model of the QoS system. The DS model assumes that the marking of packets for QoS purposes takes place when the packet enters the network and can be implemented by setting the value of the IP Precedence field, the DSCP field [15] or the Traffic Class field in the case of IPv6 packets.

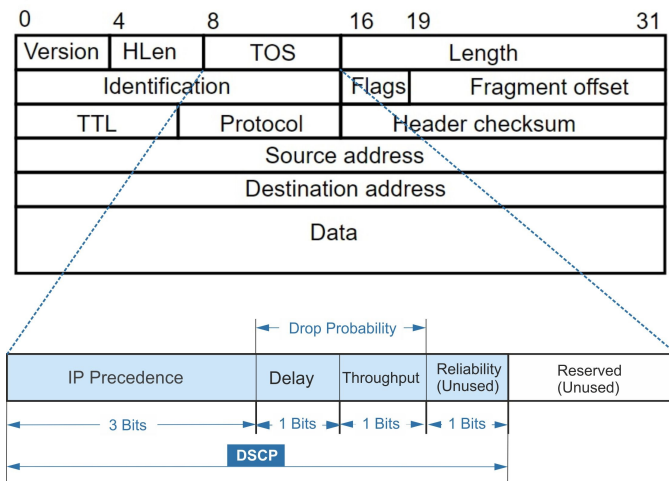


Fig. 1. ToS field in the IP packet header [16]

Recommendation RFC 2475 [17] defines the division of the DSCP field into two parts: the Class Selector field, which provides compatibility with earlier solutions (*Type of Service*) by analogy with the three-bit IP Precedence field in the ToS field, and the Drop Precedence field, which specifies the

level of packet loss probability. The ECN field identifies the congestion in the network. The Class Selector field identifies network traffic by appropriately dividing the traffic into classes of network services and prioritising the packets (datagrams) belonging to that traffic. Packets with the same DSCP class value should be subject to similar handling rules in the network nodes.

Determining the right QoS parameters and priorities for voice traffic is not a straightforward opinion and is highly dependent on the design and business needs of the organisation. However, there are rules of thumb that indicate good practice for QoS parameters for voice traffic:

- Do not set IP Precedence classes of service with a value equal to 6 or 7, as these are reserved for Internetwork Control and Network Control,
- RTP voice packets should be marked with DSCP class 5:EF,
- Data traffic should have a lower priority.

Using the highest priority for RTP (*Real Time Protocol*) data can result in RTP traffic without additional granularity causing problems for voice calls. The reason could be that this type of data is blocked by, for example, RTP traffic carrying video conferences. Telephone call data is carried using the RTP protocol, but it should be remembered that RTP is also designed to transport different types of information, including video streams. The use of the RTP protocol during voice data transmission allows for real-time data transmission, taking into account time stamps and sequence numbers, which enable the detection of lost voice packets. However, RTP does not include error correction mechanisms. In addition, voice calls require duplex transmission and control of this transmission, that is, negotiation and establishment of voice or audio-visual calls. In VoIP telephony, telephone numbers are converted into IP addresses of the end units forming the real-time data stream. Communication related to control is realised in the DC (*Data Channel*), while the RTP stream is determined by the BC (*Bearer Channel*) transport channel.

III. DELAYS AND LOSS OF VOICE PACKETS

Providing the desired QoS service for voice packets along the entire path from sender to receiver has been the subject of research work for many years [18]–[20]. The analysis and results of work in this area show that this is not a simple task. The quality degradation in a wired network can occur due eventual packet losses, which may be associated with overloaded routers and other network infrastructure problems [21]. Digital voice processing is also associated with latency problems. Delay is a quantity that reflects the amount of time it takes for a packet to travel through the network. The packet E2E (*end-to-end*) delay is measured by calculating the delay from the speaker to the receiver (including the compression and decompression delays) [22].

In IP telephony, ITU-T Recommendation G.114 [23] states that the delay in one direction must not exceed 150 ms (See results on Table I). The delay value is usually affected by propagation delay and computational delay. In the case of propagation delay, we rather have no influence on it. It

TABLE I
ITU-T PRECEPT FOR VOICE QUALITY.

Network Parameters	Good	Acceptable	Poor
Delay (ms)	0 - 150	150 - 300	>300
Jitter (ms)	0 - 20	20 - 50	>50

depends on the propagation time of the signal on the path between phones and the speed of propagation of the signal in the medium. Computational delay results from the need to process packets by the devices that transmit them and by the end devices. Examples include routers, DSPs (*Digital Signal Processors*), PBXs or telephones, which contribute some additional computational delay. Ways to reduce computational delay can be to reduce the number of devices along the packet path, to replace devices with newer ones of higher efficiency, or to reduce the complexity of audio compression algorithms. The result of the delay effect is a prolongation of audio pauses or the creation of an echo effect. The laboratory data on which the ITU-T G.114 [23] recommendations were based were collected by various research centres and laboratories. The picture of the impact of delay on the quality of voice transmission is highly simplified and Recommendation G.114 is only the result of a compromise between very different research results obtained in different research centres [24]. Therefore, when designing computer networks in which voice transmission technology (VoIP) is to be implemented, it is worth conducting research and voice transmission quality tests. Designing and implementing a well-managed VoIP network that will ensure very good quality parameters is not an easy process and requires taking into account various factors affecting the quality of transmission. For example, the G.723 codec adds a constant delay of 30 ms. If additionally the delay resulting from data interception by the gateway is added, the total delay contributed by one codec will be 32 to 35 ms. When selecting different codecs, it is therefore important to bear in mind that the delay may adversely affect voice quality or increase the demand for bandwidth. Apart from codecs, which have a clear impact on the final quality of transmitted voice in VoIP networks, there are also other undesirable factors. These include, among others:

- delays,
- delay variation (jitter),
- packet loss,
- echo.

Hence, the aim of the study was to assess the significant features and parameters of packet voice transmission, VoIP using selected audio codecs used in voice compression. Guaranteeing adequate quality of service is of particular importance in the case of networks supporting VoIP transmission. The study aims to show to what extent the selected audio codec affects the quality of transmitted voice.

IV. THE RESULTS OF RESEARCH AND DISCUSSION

A. Network topology

In order to carry out the tests, a sample network supporting VoIP transmission (See Fig. 3.) was designed. The network

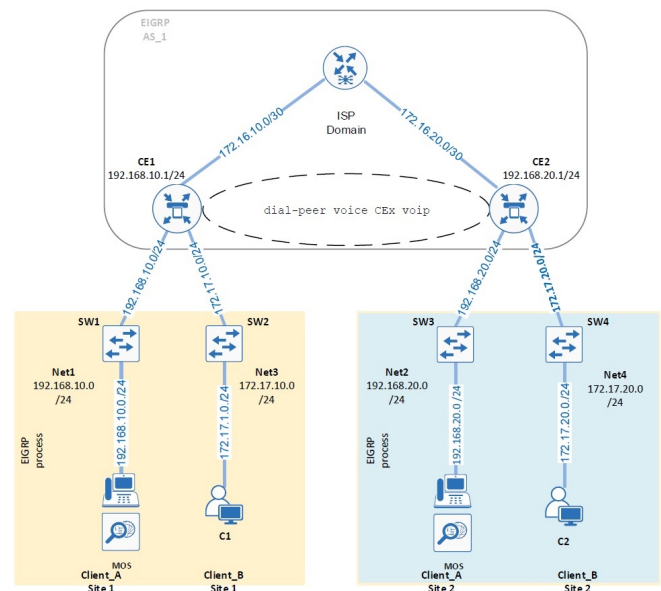


Fig. 2. Topology of the research network.

topology consisted of two locations connected to an ISP (*Internet Service Provider*). Cisco ISR routers and L2 switches were used in the topology. Cisco 7960 phones were used for VoIP communication. A hybrid routing protocol EIGRP (*Enhanced Interior Gateway Routing Protocol*) was used on all routers in the network, which uses DUAL (*Diffusing Update Algorithm*) for fast convergence and reduction of potential routing loops. Customer networks with the internal addresses

192.168.10.0/24, 192.168.20.0/24

were connected to the ISP network by sharing the EIGRP process on the CE (*Customer Edge*) routers. In addition, VoIP telephony was configured for each location on edge routers CE1, CE2 connecting customer networks to the provider network. The configured Dial Peers allowed the identification of source and destination endpoints for VoIP calls and define the characteristics that are applied to each call. Cisco phones allow for ongoing monitoring of the MOS value of the call in progress. In addition, we can monitor the current value of delay and jitter parameters. Additionally, taking into account information about the type of audio coding algorithm used, it is possible to obtain quite detailed information about the transmission quality of the conducted conversation. Additionally, DHCP servers have been configured on the customer network edge routers (CE1, CE2) for VoIP addressing. On the L2 switches located in the customer networks, voice VLANs were created, which enable separation of data and voice traffic and implementation of QoS tags in the headers of IP packets carrying voice data. The network prepared in this way was used to carry out voice transmission quality tests. The aim of the tests was to evaluate the MOS (*Mean Opinion Score*) parameters, required bandwidth, compression delay, One Way Delay and End to End Delay.

B. Methods for evaluating sound quality

There are various methods for assessing sound quality described in ITU (*International Telecommunication Union*) recommendations [25], [26]. Among the most commonly used methods we can include the following:

- MOS (*Mean Opinion Score*) [27],
- PSQM/PSQM+ (*Perceptual Speech Quality Measure*) [28],
- MNB (*Measuring Normalized Blocks*) [29],
- PESQ (*Perceptual Evaluation of Speech Quality*) [28],
- PAMS (*Perceptual Analysis Measurement System*) [30],
- E-Model [31].

Sound quality assessment methods can be divided into two groups. The first group includes subjective methods, which rely on the subjective evaluation of sound by the user. The second group consists of objective methods, which are based on the evaluation of the quality of voice samples of the transmitted speech with the received speech sample. Subjective tests are carried out in a laboratory environment. Their result, known as MOS (*Mean Opinion Score*) are the most reliable, but the test conduction is expensive and time-consuming. On the other hand, the objective methods use an algorithm to predict a MOS index value [21]. For VoIP services, the subjective evaluation of quality can be expressed on the MOS scale, which is a 5 degree scale where a value of 5 indicates the best quality and a value of 1 indicates the worst quality (See results on Table II).

TABLE II
MOS EVALUATION.

Score	Call Quality	Listening effort	Volume
5	Excellent	Total relaxation	Much louder than necessary
4	Good	Slight Attention	Louder than necessary
3	Fair	Moderate attention	Volume as needed
2	Poor	Considerable listening effort	Quieter than necessary
1	Bad	Lack of understanding, high attention	Much quieter than necessary

The primary purpose of object-oriented methods for assessing the quality of voice transmission is to measure the service provided as a human would. Subjective methods are the most authoritative, but due to the rather expensive and time-consuming testing, these methods can be replaced by testing using a computer program. There are also methods to measure the quality of VoIP services in a non-intrusive way, carried out only on the basis of traffic transmitted over the network. The ITU-T has developed a method to approximate the subjective assessment of VoIP call quality expressed on the MOS scale by objective parameters such as delay and packet loss. The developed method called *E-model* is described in ITU-T Recommendation G.107 [32]. Obtaining a score expressed in the MOS scale requires the calculation of a transmission rating

factor R value, which is then mapped to a value expressed in the MOS scale, according to [31] the following formula:

$$MOS = \begin{cases} 1 & R < 0 \\ 1 + 0.0035R + R(R - 60) & 0 < R < 100 \\ 4.5 & R > 100 \end{cases} \quad (1)$$

The value of R is described by the relation

$$R = R_0 - I_s - I_d - I_e + A \quad (2)$$

where:

R_0 - basic value calculated from the values of analogue parameters,

I_s - component representing analogue distortions occurring simultaneously with the useful signal,

I_d - component modeling the influence of packet delay on voice quality,

I_e - component modeling the impact of packet loss taking into account the specific codec used,

A - Advantage factor, a component representing the expectations of the user participating in a conversation.

The value of I_d is described by the relation:

$$I_d = 25 \left\{ (1 + X^6)^{1/6} - 3 \left(1 + \left| \frac{X}{3} \right|^6 \right)^{\frac{1}{6}} + 2 \right\} \quad (3)$$

where:

$$X = \frac{\log\left(\frac{d}{100}\right)}{\log 2} \quad (4)$$

The value of I_e , on the other hand, is described by the relation:

$$I_e = a + b \ln \left(1 + c \frac{IPLR}{100} \right) \quad (5)$$

where:

a, b, c - parameters characterising the codec,
 $IPLR$ - level of loss expressed in percentage.

The R -index can be used in the planning of the telephone network to be established. The R -value determines the quality level of the system, taking into account the expected speech quality. It is assumed that the higher the R -value, the better the signal quality. The definition of the *E-model* implies that the quality of the connection is strongly dependent not only on the level of packet loss and delay, but also on the type of codec used. To emphasise that certain quality measures are closely related to human perception, the QoE (*Quality of Experience*) model is used, which can be translated as "*perceived quality*". Thus, the term QoE is used to assess quality in a way that is more representative of human perception [33]. The quality of the system in the *E-model* is expressed by the value of the R -factor, which ranges from 0 to 100. A value of 0 indicates a system with very poor quality, while a value of 100 indicates a

system with high connection quality. The quality parameters of NG (*Next Generation*) networks should be determined during the network requirements phase, and with regard to the quality of the transmitted speech, the ITU-T Recommendation G.107 [32], which defines the E-model, should be followed.

C. Audio codecs

Although the assessment of the quality of transmitted voice in IP networks can be made at various levels, the quality mainly depends on the speech codecs used and the state of the network infrastructure. Codecs used for voice transmission in VoIP networks differ in coding rate, coding scheme, compression, used network bandwidth and algorithmic delay. The task of audio data compression is to reduce the number of bits needed to faithfully represent a speech signal in order to transmit it over a distance and then play it back. This process is carried out using an encoder and a decoder (CODEC - *COder-DECoder*). The main purpose of speech signal compression in VoIP telephony is therefore the reduction of the information stream, thanks to which the required bandwidth can be many times smaller. In a VoIP service, we can use at least several audio codecs, but it should be remembered that audio codecs are the main element affecting the quality of the transmitted voice. Hence, this paper attempts to evaluate the quality of audio codecs used in VoIP technology. VoIP service allows the use of various audio codecs that use different compression algorithms. There are several codecs that are more popular than others. The most commonly used codecs are G.711, G.722, G.723.1 and G.729, but they have different performance and parameters [34]:

- G.711: Minimum bandwidth needed is 128 kbps and its speech transmission is precise,
- G.722: Different compression is possible,
- G.723.1: Voice quality is high but consumes high processor power,
- G.726: Version of G.723 and G.721,
- G.729: Has efficient utilization of bandwidth (license required).

Table 3 lists the most popular codecs used in VoIP technology [35] along with their MOS characteristics. The highest sound quality according to the MOS scale has the G.711 codec for which the MOS value is 4.4. It is a codec based on PCM (*Pulse-Code Modulation*). In the case of the remaining codecs, despite different throughputs of the codecs, the MOS parameters have similar values (See results on Table III). It is worth noting the high value of the MOS parameter and low throughput (8 kbps) for the G.729 codec. This codec uses the CS-ACELP algorithm (*Conjugate Structure-ACELP*) algorithm, which is a modification of the CELP algorithm [36]. The biggest advantage of this codec is its lower computational complexity, which affects its quality parameters.

It is extremely important for the planning of network infrastructure to be used for the transmission of voice data to determine how much of the available bandwidth will be used for VoIP calls. In order to determine the bandwidth requirements for VoIP calls, it is necessary to determine the characteristics (Relevant Bandwidth) for individual codecs. Analyzing

TABLE III
VOIP CODECS AND MOS SCALE.

Codec	Bit Rate (kbps)	MOS
G.711	64	4.4
G.723.1	6.3	3.9
G.726	32	3.85
G.728	16	3.61
G.729	8	3.92

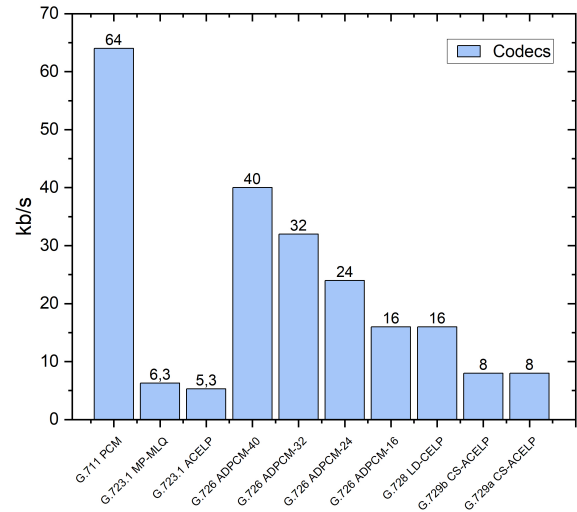


Fig. 3. VoIP codecs and Relevant Bandwidth.

VoIP codecs in terms of their bandwidth requirements, one can see significant differences in bandwidth requirements for selected codecs (See Fig. 3). Codec G.711, due to the lack of data compression, requires a bandwidth of 64 kbps. The G.723 codec has very high data compression and thus the bandwidth requirement is relatively low at 5.3 kb/s and 6.3 kb/s respectively depending on the type of coding. In the case of the G.726 codec, there are four variants which differ in the required bandwidth, i.e. 40 kb/s, 32 kb/s, 24 kb/s and 16 kb/s respectively. The G.728 codec has a bandwidth requirement of 16 kb/s. In the case of the G.729 codec, the bandwidth requirement is identical in the two coding versions and is 8 kb/s.

Telephone communication is very sensitive to delays occurring in the VoIP connection chain. For the caller, a total RTD (*Round Trip Delay*) of more than 250 ms becomes noticeable. The delay generated by the codecs is the packetisation delay, which is the time required to compress the analogue signal into a digital signal and is one of the components of delay found in networks that affect the quality of voice transmission. Fig. 4 shows the value of one-way packetisation delay for selected voice codecs. As expected, due to the lack of compression, the G.711 codec has the lowest packetisation delay of 0.75 ms. The G.723 codec is characterised by high data compression, which translates into a high packetisation delay value of 30 ms. For the G.726 codec, the packetisation delay value is low at 1 ms. For the G.728 codec, the packetisation delay value is

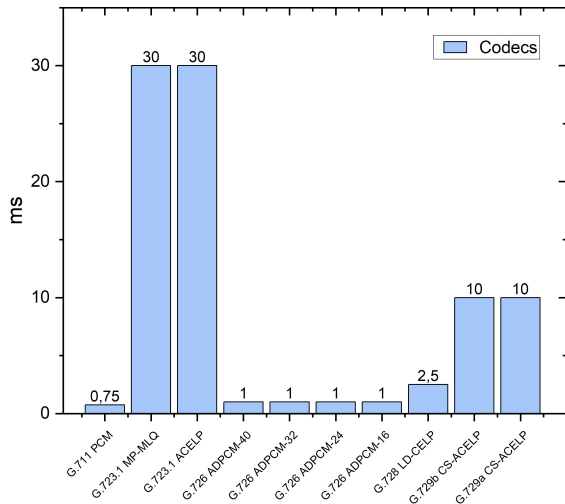


Fig. 4. Packetization delay of selected codecs.

slightly higher to the earlier codec and is 2.5 ms. For the G.729 codec, the one-way packetisation delay is already significantly higher as it is 10 ms. (See Fig. 4). It means that the lower the value of packetization delay, the lower the total delay of voice packets. Apart from the packetization delay, the total packet transmission time from the sender to the receiver should also include:

- queuing,
- network delay,
- delay variation (buffer jitter).

On the other hand, too much delay variation may cause packets to be discarded because data will be sent to the buffer jitter input in the wrong order. In order to check the relationship between the variable delay in the network and the parameters of the mechanism compensating for its impact, the value of the jitter buffer was also examined.

The jitter buffer can be either dynamic or static. The topology studied uses Cisco devices that use a dynamic buffer. This means that the buffer can grow and shrink as it is needed. The static buffer may be too large or too small, which may result in packet loss that has a very negative impact on the quality of voice transmission. Hence, in VoIP transmission, packets after reaching the recipient are queued in the jitter buffer (playout buffer). This is to smooth out the delay fluctuations that occur during transmission. However, the size of the buffer should be set in such a way as to maintain the appropriate proportion between delay and quality. Analysing VoIP codecs in terms of the jitter (playout buffer) parameter, it can be concluded that the times of the analysed codecs are very similar to each other and fall within the range of 3.1 to 3.8 ms (See Fig. 5). The highest value was observed for the G.723 codec in both its coding versions. On the other hand, values in the range of 3.3 to 3.5 ms were found for the codec G.726 depending on the coding version, and the lowest value of 3.1 ms was shown by the codecs G.729a and G.729b.

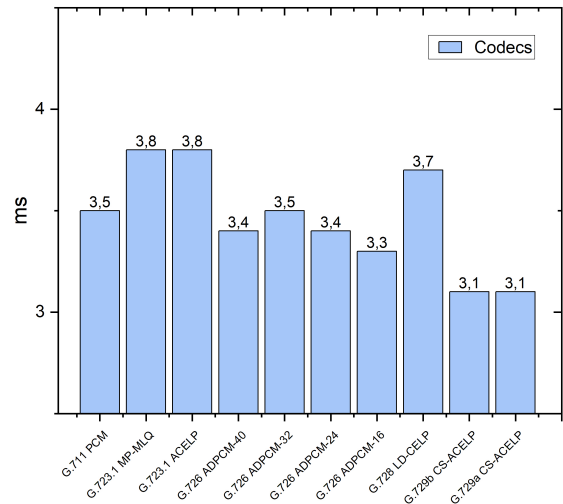


Fig. 5. Average jitter buffer value of selected codecs.

Another factor that determines the quality of voice calls is the average one-way data packet transmission rate. This is the total time it takes the packet to travel from the sender to the receiver. This time includes:

- packetization time,
- queuing,
- network delay,
- delay variation (buffer jitter).

The average packet transfer time one way was measured during a 3 minute phone call. For the measurement, 10 packets with different transmission times were selected and the arithmetic mean was calculated for the tested packets. The comparison of obtained results is shown in Fig. 6. It was found that codecs G.711, G.726 and G.728 show very similar values, which fall within the range of 21 to 24 ms. However, in the case of codec G.723, this value is twice as high and amounts to 53 ms and 51 ms respectively, depending on the coding algorithm used. The G.729 codec, on the other hand, is characterised by transmission times of 30 ms and 31 ms depending on the type of coding algorithm selected (See Fig. 6).

The next test examined the parameter, which was the average round-trip time of a voice packet. This is the time needed for the packet to reach the recipient and return with feedback to the sender. The test was carried out under the same conditions as the One Way Delay test, the measurement was made during a 3 minute phone call, and the arithmetic mean was calculated from 10 random packets. The summary of obtained results is shown in Fig. 7. It was found that the maximum time occurs for the G.723 codec with MP-MLQ coding algorithm and amounts to 112 ms, while for the ACELP algorithm this time slightly decreases to 103 ms. For the other codecs, the times are similar to each other and range from 56 to 72 ms. The G.726 codec in 40 kbps, 32 kbps, 24 kbps and 16 kbps versions is characterised by transmission times of 68 ms, 71 ms, 72 ms and 67 ms respectively. For the G.728

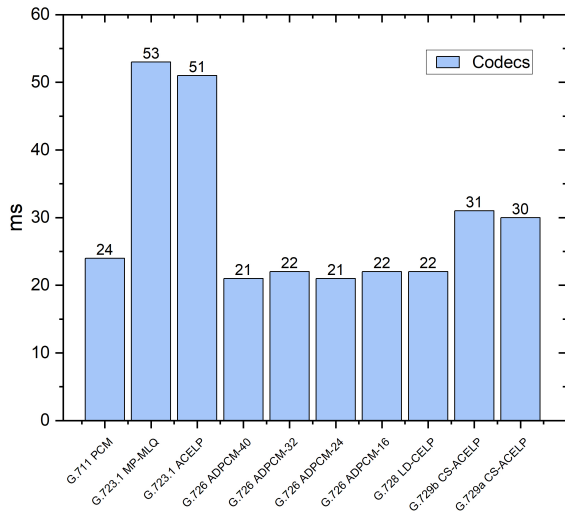


Fig. 6. One Way Delay.

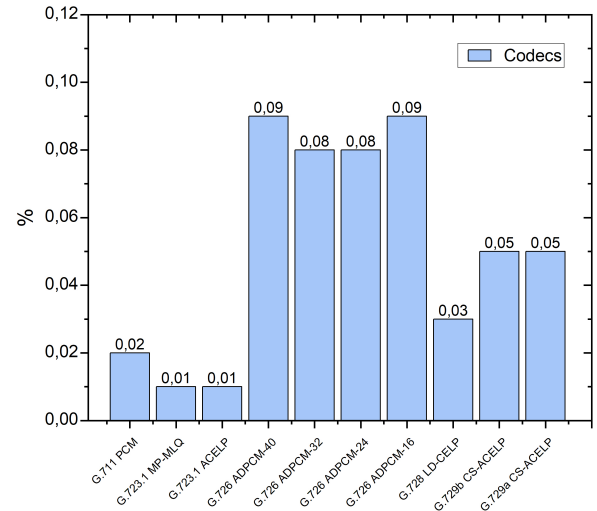


Fig. 8. Average packet loss for selected codecs.

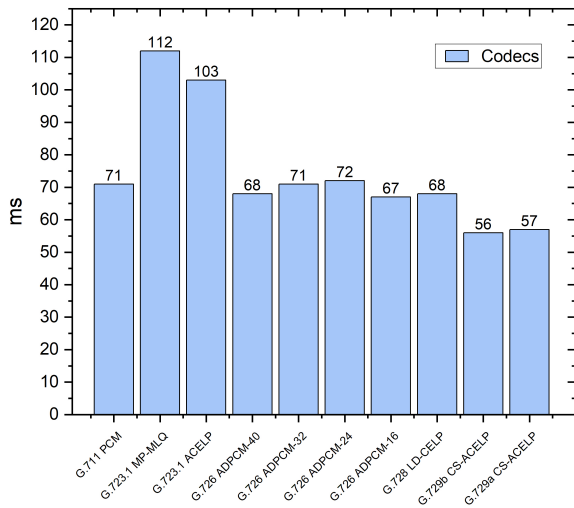


Fig. 7. End to End Delay.

codec, the transmission time is similar at 68 ms. The lowest transmission times occur for codecs G.729b and G.729a at 56 ms and 57 ms respectively (See Fig. 7).

In the last stage, the parameter of packet loss during a telephone call was studied. The lost packets were measured during a 5 minute phone call. The result of the measurement is the percentage of packets sent to the number of packets received. It was found that for all tested codecs the percentage of lost packets is at a very low, acceptable level and the values of lost packets will not significantly affect the quality of the conversation. For the G.711 codec, the lost packets did not exceed 0.02%. For G.723 codec, packet loss was 0.01% in both coding algorithms. The highest number of lost packets was observed for codec G.726: 0.09%, 0.08%, 0.08% and 0.09% in its coding versions 40 kbps, 32 kbps, 24 kbps and

16 kbps respectively. Similar level of lost packets have codecs G.728 where packet loss is 0.03% and codec G.729 in both its versions, the level of lost packets is slightly higher and is 0.05%.

V. CONCLUSION

QoS requirements placed on IP-based networks used for voice transmission (VoIP) are very restrictive, as a drop in call quality is immediately perceived by the end user. This issue is important because the share of multimedia, voice transmission, sensitive applications in IP networks is growing and the demand for guaranteed services will increase. The main goal of a voice session in IP-based networks is to meet QoS recommendations and at the same time to achieve the highest possible MOS value even in the case of network congestion. For this purpose, a network model based on Cisco devices, based on IP protocol with VoIP architecture supporting various audio codecs used for voice compression, has been designed and studied. The conducted tests of selected audio codecs for different coding algorithms showed that each codec has its own advantages and disadvantages. Under network congestion conditions, low bit rate codecs allow better performance than high bit rate codecs. Conversely, when there is no network congestion, high bit rate codecs show better performance. The conducted research has shown that the best quality of a telephone connection is achieved by the G.711 codec, but it requires the fastest link among the codecs tested. We can relatively easily determine the quality parameters of audio codecs, but the choice of codec mainly depends on the needs of the organisation, as well as the type and performance of the equipment and technology used in the company. It seems that the best performance can be achieved by a compromise between the bandwidth requirements for the codec and the desired transmission quality.

REFERENCES

- [1] S. K. Puspita FM and S. Z. Taib BM, "Improved models of internet charging scheme of single bottleneck link in multi qos networks," 2013. [Online]. Available: <http://ddms.usim.edu.my:80/jspui/handle/123456789/15429>
- [2] A. R. Modarressi and S. Mohan, "Control and management in next-generation networks: challenges and opportunities," *IEEE Communications Magazine*, vol. 38, no. 10, pp. 94–102, 2000. [Online]. Available: <https://doi.org/10.1109/35.874976>
- [3] D. Strzëciwilk, K. Ptaszek, P. Hoser, and I. Antoniku, "A research on the impact of encryption algorithms on the quality of vpn tunnels' transmission," in *ITM Web of Conferences*, vol. 21. EDP Sciences, 2018, p. 00011. [Online]. Available: <https://doi.org/10.1051/itmconf/20182100011>
- [4] H. J. Kim and S. G. Choi, "A study on a qos/qoe correlation model for qoe evaluation on iptv service," in *2010 The 12th International Conference on Advanced Communication Technology (ICACT)*, vol. 2. IEEE, 2010, pp. 1377–1382.
- [5] D. Strzëciwilk, "Examination of transmission quality in the ip multi-protocol label switching corporate networks," *International Journal of Electronics and Telecommunications*, vol. 58, pp. 267–272, 2012. [Online]. Available: <http://doi.org/10.2478/v10177-012-0037-z>
- [6] A. J. Estepa, R. Estepa, J. M. Vozmediano, and P. Carrillo, "Dynamic voip codec selection on smartphones," *Netw. Protoc. Algorithms*, vol. 6, no. 2, pp. 22–37, 2014. [Online]. Available: <https://doi.org/10.5296/npa.v6i2.5370>
- [7] W. M. Zuberek and D. Strzëciwilk, "Modeling traffic shaping and traffic policing in packet-switched networks," *Journal of Computer Sciences and Applications*, vol. 6, no. 2, pp. 75–81, 2018. [Online]. Available: <http://pubs.sciepub.com/jcsa/6/2/4>
- [8] D. Cohen, "Specifications for the network voice protocol," UNIVERSITY OF SOUTHERN CALIFORNIA MARINA DEL REY INFORMATION SCIENCES INST, Tech. Rep., 1976. [Online]. Available: <https://www.rfc-editor.org/info/rfc741>
- [9] J. Davidson, J. Peters, J. Peters, and B. Gracely, *Voice over IP fundamentals*. Cisco press, 2000.
- [10] S. Ganguly and S. Bhatnagar, *VoIP: wireless, P2P and new enterprise voice over IP*. John Wiley & Sons, 2008.
- [11] B. Hartpence, *Packet Guide to Voice over IP: A system administrator's guide to VoIP technologies*. " O'Reilly Media, Inc.", 2013.
- [12] S. Deering and R. Hinden, "Rfc2460: Internet protocol, version 6 (ipv6) specification," 1998.
- [13] K. Ramakrishnan, S. Floyd, and D. Black, "Rfc3168: The addition of explicit congestion notification (ecn) to ip," 2001.
- [14] K. Nicholas, "Definition of the differentiated services field in the ipv4 and ipv6 headers," *RFC 2474*, 1998.
- [15] F. Baker, J. Polk, and M. Dolly, "A differentiated services code point (dscp) for capacity-admitted traffic," *Internet Engineering Task Force (IETF)*, 2010.
- [16] D. Strzëciwilk, R. Nafkha, and R. Zawiślak, "Performance analysis of a qos system with wfq queuing using temporal petri nets," in *International Conference on Computer Information Systems and Industrial Management*. Springer, 2021, pp. 462–476. [Online]. Available: https://doi.org/10.1007/978-3-030-84340-3_38
- [17] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," 1998.
- [18] D. C. Dowden, R. D. Gitlin, and R. L. Martin, "Next-generation networks," *Bell Labs technical journal*, vol. 3, no. 4, pp. 3–14, 1998. [Online]. Available: <https://doi.org/10.1002/bltj.2125>
- [19] G. R. Ash, *Traffic engineering and QoS optimization of integrated voice and data networks*. Elsevier, 2006.
- [20] M. H. Miraz, S. A. Molvi, M. A. Ganie, M. Ali, and A. H. Hussein, "Simulation and analysis of quality of service (qos) parameters of voice over ip (voip) traffic through heterogeneous networks," *arXiv preprint arXiv:1708.01572*, 2017. [Online]. Available: <https://arxiv.org/abs/1708.01572>
- [21] E. T. Affonso, R. D. Nunes, R. L. Rosa, G. F. Pivaro, and D. Z. Rodriguez, "Speech quality assessment in wireless voip communication using deep belief network," *IEEE Access*, vol. 6, pp. 77 022–77 032, 2018. [Online]. Available: <https://doi.org/10.1109/ACCESS.2018.2871072>
- [22] J. Yu and I. Al-Ajarmeh, "Call admission control and traffic engineering of voip," in *2007 Second International Conference on Digital Telecommunications (ICDT'07)*. IEEE, 2007, pp. 11–11.
- [23] T. ITU, "Recommendation g. 114, one-way transmission time," *Series G: Transmission Systems and Media, Digital Systems and Networks, Telecommunication Standardization Sector of ITU*, 2000.
- [24] J. H. James, B. Chen, and L. Garrison, "Implementing voip: a voice transmission performance progress report," *IEEE Communications Magazine*, vol. 42, no. 7, pp. 36–41, 2004. [Online]. Available: <https://doi.org/10.1109/MCOM.2004.1316528>
- [25] J. G. Beerends, C. Schmidmer, J. Berger, M. Obermann, R. Ullmann, J. Pomy, and M. Keyhl, "Perceptual objective listening quality assessment (polqa), the third generation itu-t standard for end-to-end speech quality measurement part i—temporal alignment," *Journal of the Audio Engineering Society*, vol. 61, no. 6, pp. 366–384, 2013. [Online]. Available: <http://resolver.tudelft.nl/uuid:91d98cbc-d802-40d3-a1bb-a58d67668728>
- [26] R. D. Nunes, R. L. Rosa, and D. Z. Rodríguez, "Performance improvement of a non-intrusive voice quality metric in lossy networks," *IET Communications*, vol. 13, no. 20, pp. 3401–3408, 2019. [Online]. Available: <https://doi.org/10.1049/iet-com.2018.5165>
- [27] B. Naderi and R. Cutler, "An open source implementation of itu-t recommendation p. 808 with validation," *arXiv preprint arXiv:2005.08138*, 2020. [Online]. Available: <https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.21437%2FInterspeech.2020-2665&v=69f1738e>
- [28] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (pesq)-a new method for speech quality assessment of telephone networks and codecs," in *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*, vol. 2. IEEE, 2001, pp. 749–752.
- [29] S. Voran, "Objective estimation of perceived speech quality. i. development of the measuring normalizing block technique," *IEEE Transactions on speech and audio processing*, vol. 7, no. 4, pp. 371–382, 1999. [Online]. Available: <https://doi.org/10.1109/89.771259>
- [30] M. Coto-Jimenez, J. Goddard-Close, L. Di Persia, and H. L. Rufiner, "Hybrid speech enhancement with wiener filters and deep lstm denoising autoencoders," in *2018 IEEE International Work Conference on Bioinspired Intelligence (IWOB)*. IEEE, 2018, pp. 1–8. [Online]. Available: <https://doi.org/10.1109/IWOB.2018.8464132>
- [31] L. Ding and R. A. Goubran, "Speech quality prediction in voip using the extended e-model," in *GLOBECOM'03. IEEE Global Telecommunications Conference (IEEE Cat. No. 03CH37489)*, vol. 7. IEEE, 2003, pp. 3974–3978. [Online]. Available: <https://doi.org/10.1109/GLOCOM.2003.1258975>
- [32] J. A. Bergstra and C. Middelburg, "Itu-t recommendation g. 107: The e-model, a computational model for use in transmission planning," 2003.
- [33] R. Jain, "Quality of experience," *IEEE multimedia*, vol. 11, no. 1, pp. 96–95, 2004. [Online]. Available: <https://doi.org/10.1109/MMUL.2004.10000>
- [34] A. Eskandar, M. Syed *et al.*, "Performance analysis of voip over gre tunnel," *International Journal of Computer Network & Information Security*, vol. 7, no. 12, 2015. [Online]. Available: <http://doi.org/10.5815/ijcnis.2015.12.01>
- [35] R. S. Ramakrishnan and P. V. Kumar, "Performance analysis of different codecs in voip using sip," in *The Conference on Mobile and Pervasive Computing*, 2008, pp. 142–145.
- [36] S. Ragot, B. Kovesi, R. Trilling, D. Virette, N. Duc, D. Massaloux, S. Proust, B. Geiser, M. Gartner, S. Schandl *et al.*, "Itu-t g. 729.1: An 8-32 kbit/s scalable coder interoperable with g. 729 for wideband telephony and voice over ip," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 4. IEEE, 2007, pp. IV–529. [Online]. Available: <https://doi.org/10.1109/ICASSP.2007.366966>