# Differences between real and theoretical technical values of the voice over IP codecs


Martín Ortega Ortega
Departamento de Eléctrica, Electrónica y Telecomunicaciones
Universidad de Cuenca
Cuenca, Ecuador
martin.ortega@ucuenca.edu.ec

Gustavo Chafla Altamirano
Carrera de Sistemas
Facultad de Ingeniería
Pontificia Universidad Católica del Ecuador
Quito, Ecuador
gxchafla@puce.edu.ec

Mara Falconí Abad
IEEE Member
Cuenca, Ecuador
mara.falconi@ucuenca.edu.ec



***Abstract*—The values of the voice codecs' technical features play a very important role when the network administrator is dimensioning the bandwidth that a VoIP communications network with SIP is going to use. In the following research, we obtained the real values of the voice codecs as the payload, bit rate, PPS and frame size through Wireshark. For this, we performed VoIP real calls through softphone applications installed in smartphones with eight voice codecs. The results obtained for this case study demonstrate that there are differences between the theoretical and the real values, and that the real values are constant in the proposed case scenarios. We conclude that directly use the theoretical values would oversize the number of users that could communicate at the same time in a VoIP network.**



***Keywords—VoIP, voice codec, SIP, payload, bit rate, PPS, frame size***


## I. Introduction

The VoIP calls service has had a significant development in the last years, allowing a direct communication between users without area codes. The VoIP calls are constituted by different frames that have the necessary information to carry the data to be transmitted from an origin to a destination. Among these frames we can find the RTP frame, which contains the voice codec. These codecs have technical features like sampling rate, payload, bit rate, packets per second (PPS), frame size, among others, that belong to codecs' design.

The values of these technical features play a very important role when the network administrator is dimensioning the bandwidth that a VoIP communications network is going to use. For this reason, it is essential to rely on the real values of the technical features.

Thus, there are some works in this area like Ali et al. in [1]. They emphasize that the knowledge of the exact data rate is very important for network dimensioning and planning, and that the theoretical values are high when a large number of users need to maintain VoIP calls. In their paper, they calculated the data rate required by VoIP flows in WiMAX through simulation. This was done to perform call admission control in order to decide if it is possible to accommodate more calls in the system.

Cai et al. in [2] claim that voice capacity analysis reveals how the codec rate and voice packetization interval affect voice traffic performance in WLANs, which provides an important guideline for network planning and management. They obtained the maximum number of voice connections that can be supported with satisfactory user-perceived quality given the parameters of the MAC protocol and voice codecs. They also performed simulations to validate the analytical results.

Nevertheless, these works focus only in simulated results obtained in different ways, but none of them considers the real values of the voice codecs or the difference that these values have with the theoretical ones. In the following research, we obtained the real values of the most common voice codecs' technical features like the payload, bit rate, PPS and frame size through VoIP monitoring software.

For this, we performed real VoIP calls using different case scenarios with softphone applications installed in smartphones. In the tests, we used eight different bandwidth voice codecs with the purpose of finding if there are variations between the theoretical and real values of these features. These variations are relevant when the network administrator is dimensioning the number of users that can use a VoIP communications network at the same time.

## II. Background

### A. *Headers in a VoIP frame*

A VoIP frame is constituted by headers that allow the transmission of information through the communications networks. These headers follow the OSI model. Among the main headers are IP, UDP and RTP. The RTP is in charge of transport the voice packets in real time, and it has a 12 bytes header. These packets are previously coded and then decoded by means of the voice codecs [3].

## B. *Voice codecs*

The codecs are complex algorithms of coding and decoding of the information, where the analogic voice signal is previously digitalized for being transmitted and later on reconstructed in the destination. These algorithms search for balance between voice quality, bandwidth and processing resources. The voice codecs have design technical features like sampling rate, payload, bit rate, PPS, and frame size, being these the most common ones [4].

### 1) *Voice codecs standardized by the ITU*

Right after it is detailed only the codecs used in this research.

*a) Codec G.711:* Wideband codec that is the guide for the others voice codecs because it allows a good voice quality. There are versions like the G.711 u-law and the G.711 a-law. This codec uses algorithms based on Pulse Codec Modulation (PCM). The G.711 has technical features like sampling rate of 8 kHz, payload of 160 bytes, bit rate of 64 kbps, PPS of 50 and frame size of 20 ms. This codec is used specially in the PSTN and in environments where the bandwithd consumption is not an issue [5] [6] [7].

*b) Codec G.722:* It is a high quality codec similar to the codec G.711 but improved. It works with Sub Band - Adaptive Differential Pulse Code Modulation (SB-ADPCM) algorithms. The G.722 has technical features like sampling rate of 7 KHz, payloads of 120, 140 and 160 bytes, bit rates of 48, 56 and 64 kbps, PPS of 50, and frame size of 20 ms. This codec is used in digital networks [7] [8].

*c) Codec G.726:* It is based on Adaptive Differential Pulse Code Modulation (ADPCM) algorithms. It has technical features like sampling rate of 8 kHz, payloads of 40, 60, 80 and 100 bytes, bit rates of 16, 24, 32 and 40 kbps, PPS of 50, and frame size of 20 ms. This codec is used in digital transmission systems [7] [9].

*d) Codec G.729:* It is a narrowband codec. It uses algorithms based on Conjugate Structure-Algebraic Code-Excited Linear Prediction (CS-ACELP). It has technical features like sampling rate of 8 kHz, payload of 20 bytes, bit rate of 8 kbps, PPS of 50, and frame size of 20 ms. It manages a high compression decreasing the bandwidth. It is used in digital networks, digital transmission systems and multimedia systems [7] [10] [11].

### 2) *Voice codecs do not standardized by the ITU*

According to the network infrastructure compatibility available in this work, the following are described:

*a) Codec GSM:* Standardized by the ETSI. It uses Regular Pulse Excitation-Long Term Prediction-Linear Predictive Coder (RPE-LTP) algorithms. GSM has technical features like sampling rate of 8 kHz, payload of 33 bytes, bit rate of 13 kbps, PPS of 50, and frame size of 20 ms. This codec is applicable in cellular mobile telephony networks [7] [12].

*b) Codec Speex:* Defined by the IETF. It uses algorithms based on Code Excited Linear Prediction (CELP). Speex has different technical features like sampling rates of 8, 16 and 32 kHz, payloads from 6 bytes to 110 bytes, variable bit rate from 2.15 kbps to 44 kbps, PPS of 50, and frame size of 20 ms. Speex is used for VoIP communications [13].

*c) Codec iLBC:* Standardized by the IETF. It is based on Linear-Predictive Coding (LPC) algorithms. This codec has technical features like sampling rate of 8 kHz, payload of 50 bytes corresponding to a bit rate of 13.33 kbps with a PPS of 33.33 and a frame size of 30 ms. In addition, it has a payload of 38 bytes corresponding to a bit rate of 15.2 kbps with a PPS of 50 and a frame size of 20 ms. This is an experimental codec for VoIP communications [14].

# III. METHODOLOGY

With the purpose of obtaining the real values of the codecs' design technical features in VoIP calls, we captured the RTP frames that contain them. For this, a VoIP network already implemented in an Institute of higher education in Ecuador was used. This network has a centralized architecture, where the WLAN is constituted of a Unifi-AP LR mesh technology and the LAN contains Ethernet structured cabling category 6A and OM3 multimode flexible fiber optic cable of 10 GB.

Inside this communications network relies the voice network that has a principal node constituted by the VoIP server with Elastix 2.3.0 software and SIP accounts. Two more smartphones were added, where softphone applications like Zoiper, CSipSimple and Bria were installed and configured with SIP accounts.

These softphones were chosen because the set of these three allow to conduct tests with all the eight voice codecs compatibles with the VoIP network infrastructure available in the Institute. These codecs are the G.711u, G.711a, G.722 (wideband); G.726-32 (between wideband and narrowband); and GSM 06.10, G.729, Speex-8 and iLBC-13.33 (narrowband). Moreover, an IP landline telephone and a computer with monitoring voice traffic software were added, as shown in the Fig. 1 and Fig. 2.

## A. *Case scenario 1 (Intra-WLAN of the Institute)*

In this case scenario, the softphones installed in the smartphones are configured in their SIP accounts with the private IP address of the VoIP Elastix 2.3.0 server of the Institute. The voice traffic is monitored through a remote desktop with VPN connection inside of the same network, as it is shown in Fig. 1

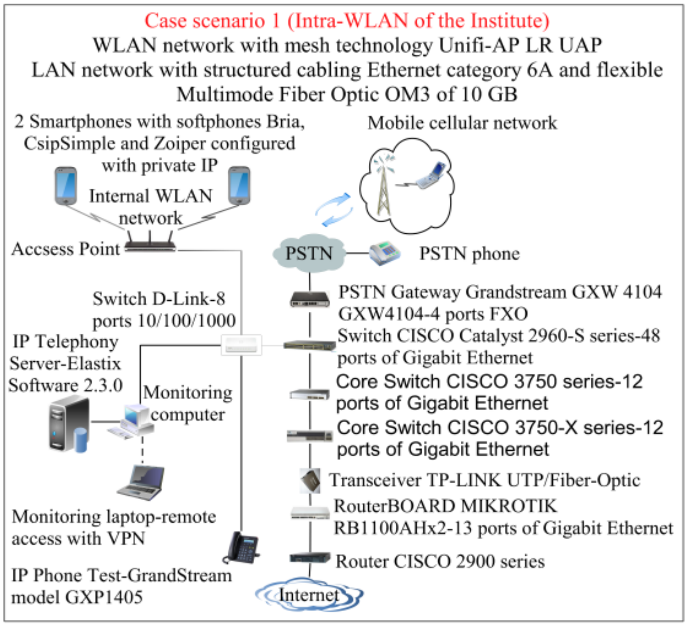


Fig. 1: Case scenario 1 (Intra-WLAN of the Institute)

### *B. Case scenario 2 (WLAN or Internet plan and mobile data)*

In this case scenario, the softphones installed in the smartphones are configured in their SIP accounts with the public IP address of the Institute, which is redirected to the gateway of the Elastix 2.3.0 VoIP server of the same Institute in order to establish the communication. The voice traffic is monitored from an external network through a remote desktop with VPN connection, as it is shown in Fig. 2.

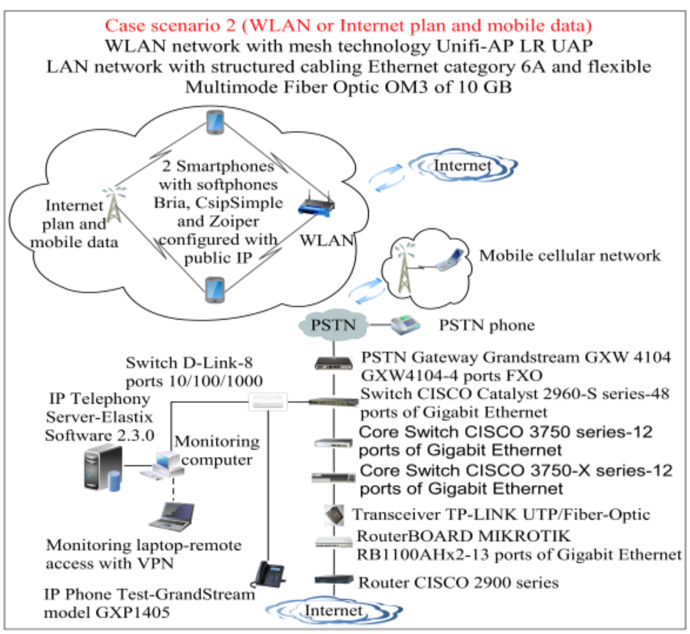


Fig. 2: Case scenario 2 (WLAN or Internet plan and mobile data)

For each case scenario, four test sub-scenarios are proposed. These are the following:

*1) Call between two softphones:* The calls are made from softphone 1 to softphone 2 or viceversa. The same voice codec is set up in the two softphones. This is done until the set of eight codecs is complete.

*2) Call between a softphone and an IP landline telephone of the Institute:* The calls are made from softphone 1 or softphone 2 to the IP landline telephone of the Institute. One codec at a time is set up until the set of eight codecs is complete, taking in consideration that the codec that the IP landline telephone uses, goes according to its own configuration. The softphone codec can be inherited by the IP landline telephone as long as this codec is available in it, otherwise the telephone will be recodified by the VoIP server.

*3) Call between a softphone and a PSTN telephone:* The calls are made from softphone 1 or softphone 2 to a PSTN telephone. Calls go through the GrandStream GXW4104 PSTN gateway, and it connects to the PSTN. In the softphone, one codec at a time is set up until the set of eight codecs is complete, considering that the codec that the PSTN gateway uses, depends on its own configuration. The PSTN gateway can inherit the softphone codec as long as it is available in the gateway, otherwise the gateway will be recodified by the VoIP server.

*4) Call between a softphone and a cellular mobile telephone:* The calls are made from softphone 1 or softphone 2 to a cellular mobile, taking in consideration that the call goes through the GrandStream GXW4104 PSTN gateway. It connects in turn with the PSTN and after that, the call is redirected to the cellular mobile telephony network. In the softphone, a given codec is set up, taking in consideration that the codec that the PSTN gateway uses, goes according to its own configuration. The softphone codec can be inherited by the PSTN gateway as long as this codec is available in it, otherwise the gateway will be recodified by the VoIP server.

For each test sub-scenario, 8 real VoIP calls were made, one for each voice codec. It means 32 calls per scenario, a total of 64 calls. In each call we captured the RTP frames with the Wireshark software. Through it, we filtered the codec contained in the RTP frame to obtain the values of the most common design technical features of the codec that are payload, bit rate, PPS and frame size. From this filtered RTP frame, we directly got the payload value, while the bit rate value was approximately obtained through a Wireshark graphic, shown in Fig. 3.

Likewise, the PPS value was approximately obtained through another Wireshark graphic, shown in Fig. 4.

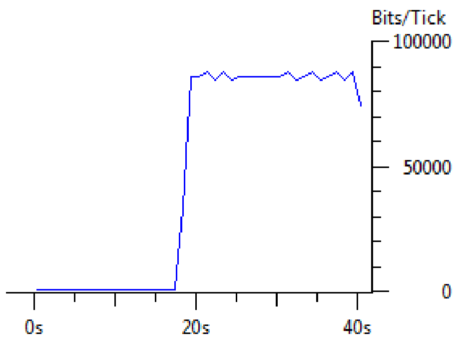


Fig. 3. Graphic measurement of the codec G.711u' bit rate in a RTP frame

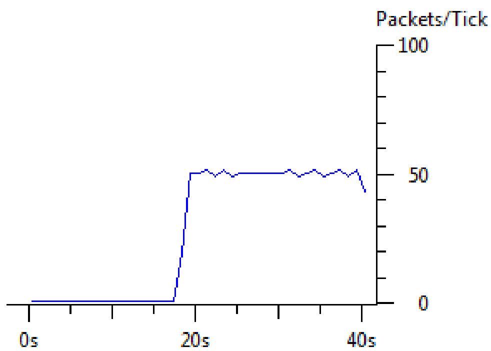


Fig. 4. Graphic measurement of the codec G.711u' PPS in a RTP frame

With the obtained payload, bit rate and PPS real values of the codec, we calculated the real value of the frame size. This was done in two ways. The first one with (1), where the PPS value is used, and the second one with (2), where the values of payload and bit rate are used.

$$frame\ size = \frac{1}{\mathrm{PPS}} \tag{1}$$

$$frame\ size = \frac{payload*8}{\text{bit rate}} \tag{2}$$

Finally, we compared the obtained real values with the corresponding technical features that the voice codecs have in the recommendations of ITU and the standards of ETSI and IETF. This procedure was applied for the eight voice codecs proposed using the different case scenarios.

## IV. Results and Analysis

In Table I and II we show the payload, bit rate, PPS and frame size obtained in the case scenario 1 and 2 and their sub-scenarios during the VoIP reals calls using the different proposed voice codecs.

**Table I.** Real Technical Features of the Voice Codecs in the Case Scenario 1

| **Sub-scenarios** | **Codec** | **Measured data** | | | **Calculated data** | |
|---|---|---|---|---|---|---|
| | | ***payload (bytes)*** | ***bit rate (kbps)*** | ***PPS*** | ***(1) frame size (ms)*** | ***(2) frame size (ms)*** |
| Call between two softphones | G.711u | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.711a | 214 | 85 | 50 | 20.00 | 20.14 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 88 | 50 | 20.00 | 19.45 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a landline IP telephone | G.711u | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.711a | 214 | 86 | 50 | 20.00 | 19.91 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 88 | 50 | 20.00 | 19.45 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a PSTN telephone | G.711u | 214 | 86 | 50 | 20.00 | 19.91 |
| | G.711a | 214 | 87 | 50 | 20.00 | 19.68 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 86 | 50 | 20.00 | 19.91 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a cellular mobile telephone | G.711u | 214 | 86 | 50 | 20.00 | 19.91 |
| | G.711a | 214 | 88 | 50 | 20.00 | 19.45 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 86 | 50 | 20.00 | 19.91 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |

In Table I, the cells in color gray indicate the wideband codecs, the cell in color yellow represents the codec between wideband and narrowband, and the orange cells indicate the narrowband codecs. Analyzing the real technical features obtained, we can see that the payload value in the same codec is identical in all the sub-scenarios, while the bit rate values have small differences between them due to their graphic measurements. These differences are until 1 kbps for the G.711u, until 3 kbps for the codec G.711a, and until 2 kbps for the G.722. For the other codecs there are no changes. Analyzing the PPS, we can see that there is no variation in its values. For the frame size, the column obtained with (1) do not show changes, but the column obtained with (2) show values with a difference between them until 0.23 ms for the G.711u, until 0.69 ms for the G.711a, until 0.05 ms for the GSM 06.10, until 0.46 ms for the G.722 and for the other codecs there is no variation.

Finally, the real values of frame size obtained with (1) and (2) show very small variations between them in spite of (1) is based on PPS and (2) is based on payload and bit rate. The reason is because in (1), PPS do not have significant variations with respect to its theoretical values, while in (2), the payload and bit rate are considerably different from their theoretical values but its variation remains constant.

**TABLE II.** REAL TECHNICAL FEATURES OF THE VOICE CODECS IN THE CASE SCENARIO 2

| **Sub-scenarios** | **Codec** | **Measured data** | | | **Calculated data** | |
|---|---|---|---|---|---|---|
| | | ***payload (byttes)*** | ***bit rate (kbps)*** | ***PPS*** | ***(1) frame size (ms)*** | ***(2) frame size (ms)*** |
| Call between two softphones | G.711u | 214 | 85 | 50 | 20.00 | 20.14 |
| | G.711a | 214 | 84 | 50 | 20.00 | 20.38 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a landline IP telephone | G.711u | 214 | 85 | 51 | 19.61 | 20.14 |
| | G.711a | 214 | 87 | 50 | 20.00 | 19.68 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a PSTN telephone | G.711u | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.711a | 214 | 87 | 50 | 20.00 | 19.68 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |
| Call between a softphone and a cellular mobile telephone | G.711u | 214 | 86 | 50 | 20.00 | 19.91 |
| | G.711a | 214 | 87 | 50 | 20.00 | 19.68 |
| | GSM 06.10 | 87 | 35 | 50 | 20.00 | 19.89 |
| | G.722 | 214 | 87 | 50 | 20.00 | 19.68 |
| | G.729 | 74 | 30 | 50 | 20.00 | 19.73 |
| | G.726-32 | 134 | 54 | 50 | 20.00 | 19.85 |
| | Speex-8 | 74 | 30 | 50 | 20.00 | 19.73 |
| | iLBC-13.33 | 104 | 28 | 34 | 29.41 | 29.71 |

The results of Table II show that the payload value in the same codec remains constant in all the sub-scenarios, while the bit rate values have small differences between them in their graphic measurements. These differences are 2 kbps for the G.711u and 3 kbps for the G.711a. For the other codecs there are no changes. The values of PPS and frame size obtained with (1) do not show changes for any codec. However, the frame size values obtained with (2) have small differences between them, corresponding to 0.46 ms for the G.711u, 0.70 ms for the G.711a, and for the other codecs there is no variation. The frame size values obtained with (1) and (2) also have smalls differences because of the reasons already explained in the case scenario 1.

After obtaining all the values of the real technical features of the voice codecs, we proceeded to compare each one of them with their respective theoretical values that appear in the recommendations and standards. This comparison is presented in Table III.

The results of Table III demonstrate that there are differences between the theoretical values of the voice codecs' technical features and their real values. Thus, with respect to the theoretical values, we can see a significant increment in the payload and bit rate, while in the PPS and frame size calculated with (1) and (2), their increment and decrement are very small.

These differences are almost the same for the case scenario 1, close to an ideal one due to the tests performed in the intra-WLAN of the Institute, as for the case scenario 2, in which the data were transmitted through a WLAN or Internet plan and mobile data.

The differences previously pointed out have a considerable influence when the administrator needs to size the bandwidth of a VoIP communications network.

**TABLE III.** VARIATIONS OF THE VOICE CODECS' REAL TECHNICAL FEATURES WTH RESPECT TO THEIR THEORETICAL FEATURES

| **Variations of the real technical features with respect to the theoretical ones** | | | | |
|---|---|---|---|---|
| **Technical features of the voice codecs** | ***Case scenarios*** | | | |
| | ***Case scenario 1*** | | ***Case scenario 2*** | |
| | ***Absolute error*** | ***Relative error (%)*** | ***Absolute error*** | ***Relative error (%)*** |
| payload (bytes) | 54 | From 33.75 to 270 | 54 | From 33.75 to 270 |
| bit rate (kbps) | From 14.67 to 24 | From 32.81 to 275 | 14.67 to 23 | From 31.25 to 275 |
| PPS | 0.67 (only for the iLBC-13.33) | From 0 to 2.01 | 0.67 to 1 | From 0 to 2.01 |
| (1) frame size (ms) | -0.59 (only for the iLBC-13.33) | From - 1.97 to 0 | From - 0.59 to 0 | From -1.96 to 0 |
| (2) frame size (ms) | From -0.55 to 0.14 | From - 0.75 to 0.70 | From - 0.32 to 0.38 | From -1.6 to 1.9 |

## V. CONCLUSIONS

It is very important to use the real values of the voice codecs' technical features when the bandwidth that is going to use a VoIP communications network is dimensioned. This is because if the theoretical values were used, the number of users that could communicate at the same time in the network would be oversized. Also, we demonstrated that the real values of payload, bit rate, PPS and frame size of the different codecs used in real IP calls, remain constant despite of the case scenarios proposed in this research are very different between them. Due to this, we can deduce that they also will remain constant for other scenarios.

With respect to their theoretical equivalents, the real values of frame size calculated with (1) have small differences because the PPS do not have significant variations. This is the same for the values of frame size calculated with (2), based on payload and bit rate, because even when they have a significant variation, it remains constant.

As new voice codecs come up to the global market, they must be subject to different evaluation studies of their technical features for their use in specific VoIP network infrastructures, being this a guide study for future research.

## ACKNOWLEDGMENT

We acknowledge Universidad Católica de Cuenca Campus Azogues for providing the VoIP communications network infrastructure to conduct this work.

## REFERENCES

[1] A. A. Ali, S. Vassilaras and K. Ntagkounakis, "A comparative study of bandwidth requirements of VoIP codecs over WiMAX access networks," *Next Generation Mobile Applications, Services and Technologies, 2009. NGMAST '09. Third International Conference on. . IEEE*, pp. 197-203, 15-18 September 2009.

[2] L. X. Cai, X. Shen and J. W. Mark, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 3, pp. 752-761, May 2006.

[3] I. E. T. Force, "IETF documents RFC 2508: Compressing IP/UDP/RTP Headers for Low-Speed Serial Links," February 1999.

[4] I. T. Union, "Recommendation ITU-T H.323: Packet-based multimedia communications systems," 14 December 2009.

[5] I. T. Union, "Recommendation G.711: Pulse code modulation (PCM) of voice frequencies," 25 November 1998. [Online]. Available: https://www.itu.int/rec/T-REC-G.711-198811-I/en.

[6] I. T. Union, "Recommendation ITU-T G.711 Appendix II: A comfort noise payload definition for ITU-T G.711 use in packet-based multimedia communication systems," 17 Febrary 2000.

[7] I. E. T. Force, "IETF document RFC 3551: RTP Profile for Audio and Video Conferences with Minimal Control," July 2003.

[8] I. T. Union, "Recommendation ITU-T G.722: 7 kHz audio-coding within 64 kbit/s," 13 September 2012.

[9] I. T. Union, "Recommendation CCITT G.726: 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," 14 December 1990.

[10] I. T. Union, "Recommendation ITU-T G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)," 29 June 2012.

[11] I. T. Union, "Recommendation ITU-T G.729.1: G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729," 29 May 2006.

[12] E. T. S. Institute, "ETSI document ETS 300961: Digital cellular telecommunications system (Phase 2+); Full rate speech; Transcoding (GSM 06.10 version 5.2.1 Release 1996)," December 2000.

[13] I. E. T. Force, "IETF document RFC 5574: RTP Payload Format for the Speex Codec," June 2009.

[14] I. E. T. Force, "IETF document RFC 3951: Internet Low Bit Rate Codec (iLBC)," December 2004.