

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN
UNIVERSIDAD POLITÉCNICA DE CARTAGENA



Proyecto Fin de Carrera

Estudio experimental de calidad de servicio de voz sobre IP: comparativa subjetiva versus objetiva



AUTOR: Salvador Moreno Urrea
DIRECTOR: María Dolores Cano Baños

Septiembre / 2012



Autor	Salvador Moreno Urrea
E-mail del Autor	salvadormorenourea@gmail.com
Director(es)	María Dolores Cano Baños
E-mail del Director	mdolores.cano@upct.es
Codirector(es)	
Título del PFC	Estudio experimental de calidad de servicio de voz sobre IP: comparativa subjetiva versus objetiva
Descriptorios	
Resumen	
<p>El objetivo del proyecto es realizar un estudio comparativo (análisis objetivo vs análisis subjetivo) de la calidad entre dos códecs de audio de dos aplicaciones de VoIP. Para ello se tendrán en cuenta tanto las opiniones del usuario como las medidas objetivas que se realizarán. En un escenario cableado Ethernet controlado, se realizarán una serie de pruebas (en total 60). Para cada prueba realizaremos dos llamadas, utilizando en cada una de ellas una aplicación de Voz sobre IP (Jitsi y Skype).</p> <p>Para Jitsi se utilizará el códec de audio G.722 y para Skype el códec Silk, con el fin de contrastar los resultados obtenidos. Tras cada llamada el usuario deberá rellenar un cuestionario que se habrá elaborado anteriormente y también se capturará el tráfico en el destino y en el origen de la llamada. Con los datos obtenidos se procederá a realizar el estudio propiamente dicho de la calidad del servicio.</p> <p>Este proyecto ha reflejado la existencia de una relación entre los resultados obtenidos en el estudio de QoS y los obtenidos en el estudio de QoE.</p>	
Titulación	Ingeniería Técnica de Telecomunicaciones
Intensificación	Telemática
Departamento	Tecnologías de la información y las comunicaciones
Fecha de Presentación	Septiembre - 2012

ÍNDICE DE CONTENIDOS

CAPITULO 1

Introducción **8**

1.1 Motivación.....8

1.2 Objetivos.....9

CAPITULO 2

Voz Sobre IP **10**

2.1 Definición de VOIP.....10

2.2 Especificaciones del Servicio de VOIP.....11

2.3 Arquitectura.....12

2.4 Códecs de Audio.....13

2.5 Protocolos de Señalización.....14

 2.5.1 H.232.....14

 2.5.2 SIP.....15

 2.5.3 MGCP-MEGACO.....15

 2.5.4 IAX.....16

CAPITULO 3

Medida de la Calidad de Voz en Redes VOIP **17**

3.1 Calidad de Servicio (QoS).....17

 3.1.1 Ancho de Banda.....19

 3.1.2 Pérdida de Paquetes.....19

 3.1.3 Retardo.....20

 3.1.4 Factor de Compresión.....22

3.2 Calidad Experimentada por el Usuario (QoE).....22

 3.2.1 Metodologías de Estimación de Calidad de Experiencia.....24

 3.2.1.1 Métodos Subjetivos.....24

 3.2.1.2 Métodos Objetivos.....25

CAPITULO 4

Desarrollo de las Pruebas y Obtención de Resultados **27**

4.1 Entorno y Características de los Componentes.....	27
4.1.1 Grabaciones de Señales en la Fuente.....	28
4.1.1.1 Entorno de Grabación.....	28
4.1.1.2 Sistema de Grabación.....	28
4.1.1.3 Material de Conversación.....	28
4.1.1.4 Procedimiento de Grabación.....	28
4.1.1.5 Participantes en las Pruebas.....	29
4.1.1.6 Niveles de Señal Vocal.....	29
4.1.1.7 Sistema Emisor.....	29
4.1.2 Concepción del Experimento.....	30
4.1.3 Procedimiento de la Prueba de Escucha.....	30
4.1.3.1 Sistema de Escucha.....	30
4.1.3.2 Oyentes.....	31
4.2 Análisis QoE de los Códecs SILK y G.722.....	31
4.2.1 Cuestionario.....	31
4.2.2 Identificador.....	32
4.2.3 Descripción Teórica de los Cálculos.....	33
4.2.4 Resultados del Análisis de QoE.....	34
4.2.4.1 Perfil de los encuestados.....	34
4.2.4.2 Dificultades para oír.....	34
4.2.4.3 Calidad y esfuerzo de comprensión.....	36
4.2.4.4 Skewness y Kurtosis.....	37
4.2.4.5 ANOVA One-Way.....	38
4.3 Análisis QoS de los Códecs SILK y G.722.....	39
4.3.1 Descripción Teórica de los Cálculos.....	41
4.3.2 Scripts AWK.....	41
4.3.2.1 Script utilizado en Extremo Transmisor.....	41
4.3.2.2 Script utilizado en Extremo Receptor.....	45
4.3.2.3 Script para el Cálculo del Ancho de Banda del Tráfico no VOIP... 48	48
4.3.3 Resultados del Análisis de QoS	49
4.3.3.1 Ancho de Banda.....	49
4.3.3.2 Tamaño Medio de los Paquetes.....	52
4.3.3.3 Desviación Estándar del Tamaño de los Paquetes.....	54

4.3.3.4 Media de la Diferencia de Tiempos.....	55
4.3.3.5 Desviación Estándar de la Diferencia de Tiempos.....	56
4.3.3.6 Pérdida de Paquetes.....	57
4.3.3.7 Ancho de Banda No VoIP.....	58
4.3.3.8 Retardo de la Red.....	59

CAPITULO 5

Conclusiones	60
---------------------	-----------

ANEXO A

Métodos Subjetivos para Audio	62
--------------------------------------	-----------

ANEXO B

E-Model	64
----------------	-----------

ANEXO C

Instalación de los Programas	66
C.1 Instalación Wireshark.....	66
C.2 Instalación Skype.....	69
C.3 Instalación Jitsi.....	73

ANEXO D

Cuestionario para el Análisis de QoE	75
---	-----------

ANEXO E

Demostración Intervalo de Confianza y Tabla t-student	76
E.1 Construcción de un Intervalo de Confianza con σ^2 Desconocida	76
E.2 Tabla t-students.....	78

Referencias	79
--------------------	-----------

CAPÍTULO 1

Introducción

1.1 Motivación

El abrumador crecimiento de Internet de los últimos años ha llevado a la proliferación de múltiples servicios que actualmente se ofrecen a través de la misma. La tendencia al uso de Internet como la portadora de todos estos servicios es una realidad tangible. Sin embargo, este crecimiento desmedido no ha sido acompañado de un cambio estructural real que permita asegurar garantías de calidad a los usuarios finales: Internet es aún una red de mejor esfuerzo. Pero es claro que sin garantías de calidad no es posible la tarificación de servicios y la idea de red convergente quedaría en simplemente eso: una idea.

Este problema ha generado un creciente interés por parte de los proveedores de servicio en estimar la calidad ofrecida. El problema con la calidad de servicio (*Quality of Service - QoS*) es qué medir y cómo hacerlo. Tradicionalmente se ha estimado la QoS en base al estado de la red de transporte, midiendo valores promedio de probabilidad de pérdida de paquetes, retardos, ancho de banda disponible y demás. Algunos alegan sin embargo que las medidas tradicionales no son las adecuadas ya que no representan correctamente la QoS experimentada por el usuario final.

En los últimos años surge un nuevo enfoque para el problema de calidad de servicio en Internet. Este se basa en que la calidad de un servicio es un tema completamente subjetivo y depende directamente de lo que el usuario perciba del mismo, independientemente del estado de la red que lo transporta. Esto es muy claro en los servicios de audio y video; en base a distintas técnicas de codificación de la señal y al uso de algoritmos de información redundante es posible que el usuario final experimente niveles de calidad aceptables aún frente a problemas serios en la red. Aparece así el concepto de calidad de experiencia (*Quality of user Experience*).

Existen en la actualidad distintos algoritmos para estimar la calidad de experiencia en voz y video sobre IP. Muchos se basan en técnicas utilizadas en el diseño de redes de telefonía tradicional y en el desarrollo de nuevos algoritmos de codificación de audio y video; si bien en estas áreas han probado ser buenos estimadores, su pasaje a Internet no ha sido del todo exitoso. Otras técnicas recientes han demostrado mejores resultados, pero de momento no está clara la generalidad de los mismos.

Las directrices y los principios de este proyecto serán centran en el estudio de los parámetros de calidad en la utilización de dos programas de VoIP distintos (Skype y Jitsi) que utilizan códecs distintos (SILK y G.722).

1.2 Objetivos

Una de las características fundamentales de este proyecto es la información de carácter didáctico. El objetivo del proyecto será realizar un estudio comparativo (análisis objetivo vs análisis subjetivo) de la calidad entre dos códecs de audio de dos aplicaciones de VoIP.

Las redes de telecomunicación modernas proporcionan una extensa gama de servicios vocales que utilizan numerosos sistemas de transmisión. En particular, la rápida expansión de las tecnologías digitales ha hecho crecer la necesidad de evaluar las características de transmisión de los nuevos equipos. En muchas circunstancias es necesario determinar los efectos subjetivos causados por algún nuevo equipo de transmisión o por la modificación de las características de transmisión de una red telefónica.

En este proyecto se realizará un estudio comparativo (análisis objetivo vs análisis subjetivo) de la calidad entre dos códecs de audio de dos aplicaciones de VoIP. Para ello se tendrán en cuenta tanto las opiniones del usuario como las medidas objetivas que se realizarán.

Para una mejor comprensión de los estudios que se van a llevar a cabo se comenzará este proyecto con unos capítulos en los que explicará teóricamente aspectos tales como la voz sobre IP, los códec utilizados en VoIP, la calidad en comunicaciones de VoIP (QoS y QoE), etc.

Se elegirán dos códecs diferentes, utilizados por aplicaciones VoIP diferentes y sobre un escenario de pruebas controlado se realizarán una serie de pruebas, en las que se obtendrán datos necesarios para la realización del estudio, tanto para QoS como para QoE. Dichas pruebas van a consistir en realizar una serie de conversaciones (dos por cada oyente, una por cada códec), en las que se va a emitir una grabación y cada oyente rellenará un cuestionario por cada aplicación de VoIP. Dicho cuestionario será necesario para el estudio de QoE. Para el estudio de QoS se realizará la captura de paquetes en la conversación utilizando una herramienta específica para ello.

Una vez se tengan todos los datos recopilados se llevarán a cabo los estudios anteriormente citados y se observará si hay alguna relación entre ellos.

CAPÍTULO 2

Voz sobre IP

En este capítulo se introducen los conceptos generales de calidad de servicio (*QoS -Quality of Service*) en servicios multimedia de tiempo real sobre IP. Un resumen de los parámetros tradicionales de calidad de servicio se presenta en la sección 2.1. Al mismo tiempo y motivado en lograr un mayor entendimiento de la gran variedad de elementos que componen estos servicios multimedia, se describe en la sección 2.3 las características fundamentales de los servicios de voz sobre IP. Se brinda un análisis de la influencia de los parámetros que afectan la calidad de estos servicios.

2.1 Definición de VoIP

Voz sobre IP es una tecnología que permite realizar llamadas telefónicas sobre redes de ordenadores, como puede ser *Internet*. Voz sobre IP también es conocida como la telefonía de *Internet*. La telefonía tradicional consiste en transportar la señal analógica sobre cable de cobre, pero la tecnología de VoIP convierte la voz analógica en paquetes de datos digitales que soportan la comunicación sobre el protocolo IP y que pueden emplear protocolos para aplicaciones en tiempo real como el RTP (*Real Time Protocol*) [HSRV96].

Así, el servicio de VoIP sustituye el tradicional teléfono residencial por un teléfono de gran ancho de banda que utiliza *Internet* para hacer y recibir llamadas. Si se realiza una llamada a un número de teléfono tradicional, la señal es convertida en el otro extremo. Dependiendo del tipo de servicio VoIP, se puede hacer una llamada VoIP desde el ordenador, desde un teléfono especial VoIP o desde un teléfono tradicional con o sin adaptador.

A continuación se enumeran las distintas funciones que caracterizan VoIP y por las que resulta tan ventajoso su uso. En primer lugar, permite realizar más de una llamada telefónica simultáneamente. Además, las llamadas entrantes pueden ser automáticamente dirigidas al teléfono VoIP independientemente de dónde se esté conectado a la red. Muchos operadores de VoIP incluyen servicios gratuitos que en las PSTN (*Public Switched Telephone Network*, Red Pública de Telefonía Conmutada) la mayoría de las compañías de telecomunicación normalmente cargan como una tarifa extra. Ejemplos pueden ser una llamada a tres, llamadas al extranjero, re-llamada automática o identificador de llamada.

VoIP puede ser más segura ya que permite el uso de protocolos como el *Secure Real-time Transport Protocol* (SRTP, Protocolo de Seguridad de Transporte en Tiempo Real) [MDNE04]. Sólo es necesario cifrar y autenticar la trama de datos. Asimismo, VoIP es independiente del lugar, sólo es necesaria una conexión a *Internet* para conseguir una conexión a un proveedor VoIP. Los teléfonos VoIP pueden integrarse con otros servicios disponibles en *Internet*, incluyendo conferencia de audio o de video, mensajes o intercambio de archivos de datos en paralelo con la conversación.

Finalmente, habilita la transmisión en un mayor ancho de banda que puede significar mejorar la calidad de la conversación y la música.

Sin embargo, existe un importante inconveniente que ha hecho que la expansión de la VoIP no sea tan rápida como se esperaba: la dificultad en ofrecer QoS. En la transmisión de voz es necesario que todos los paquetes lleguen ordenados, que no haya pérdidas y garantizar una mínima tasa de transmisión lo que implica la necesidad de QoS. En otros servicios como el correo, ofrecer QoS no es crítico, ya que si un paquete no ha llegado al destino se solicita su retransmisión; pero esto no es posible en la VoIP, ya que se trata de un servicio en tiempo real. La solución radica en diferenciar los paquetes de voz de los paquetes de datos, priorizar la transmisión de los paquetes de voz y evitar que la transmisión de los paquetes no supere los 150 milisegundos, tal y como se especifica en la recomendación ITU-TG 114 [DOHE04].

La calidad de servicio se está logrando en base a los siguientes criterios:

- La supresión de silencios y VAD (*Voice Activity Detection*), otorga más eficiencia a la hora de realizar una transmisión de voz, ya que se aprovecha mejor el ancho de banda al transmitir menos información.
- Compresión de cabeceras aplicando los estándares RTP/RTCP (*Real Time Protocol*).
- Cancelador de eco
- Priorización de los paquetes que requieran menor latencia.
- La implantación de IPv6 que proporciona mayor espacio de direccionamiento y la posibilidad de tunneling.

2.2 Especificaciones del Servicio de VoIP

En esta sección se introducen los conceptos básicos para la comprensión del servicio de voz sobre IP. Al conjunto de tecnologías que hacen posible la transmisión de la voz sobre redes que utilizan el protocolo IP se le llama VoIP.

Se pueden distinguir dos categorías dentro de estas tecnologías:

- Las que se encargan del tratamiento de la señal de voz. Digitalización, compresión, empaquetado, etc.
- Las encargadas de la configuración de la llamada (*call setup*).

En la primera categoría encontramos las que hacen posible la transmisión eficiente de la señal de voz por la red. El primer paso es pasar la señal del dominio analógico al dominio digital. Luego esta señal es codificada para reducir la tasa de información a transmitir. La señal codificada de manera eficiente es paquetizada para enviarla por la red. En el receptor el proceso es el inverso.

En la segunda categoría se ubican las tecnologías de intercambio de información sobre la configuración de la llamada. Existen en la actualidad varios protocolos VoIP: H.323, SIP, IAX, MGCP/MEGACO, etc. Esta información es utilizada para autorización, autenticación, resolución de direcciones, etc.

2.3 Arquitectura

Para la transmisión de voz sobre una red IP, el estándar define tres elementos fundamentales en su estructura:

- **Terminales:** son los puntos finales de la comunicación y pueden ser implementados como:
 - **Hardware:** un teléfono IP es un terminal que tiene soporte VoIP nativo y puede conectarse directamente a una red IP
 - **Software:** un *softphone* es una aplicación audio ejecutable desde PC que se comunica con las PABX a través de la LAN. Para interactuar con el usuario se basa en la utilización de un micrófono y altavoz o mediante un teléfono USB.
- **Servidor:** provee el manejo y funciones administrativas para soportar el enrutamiento de llamadas a través de la red. Este servidor puede adoptar diferentes nombres dependiendo del protocolo de señalización utilizado. Así en un sistema basado en el protocolo H.323, el servidor es conocido como *Gatekeeper*; en un sistema SIP, **servidor SIP**; y en un sistema basado en MGCP o MEGACO, *Call Agent* (Agente de llamadas). El servidor es un elemento opcional, normalmente implementado en software, y en caso de existir, todas las comunicaciones pasarían por él.
- **Gateways:** enlace de la red VoIP con la red telefónica analógica o RDSI. Se encarga de adaptar las señales de estas redes a VoIP y viceversa, actuando de forma totalmente transparente para el usuario. El Gateway posee, además de puertos LAN, interfaces de conexión a estas redes: FXO, FXS, E&M, BRI, PRI, G703/G.704
- **Red IP:** provee conectividad entre todos los terminales. La red IP puede ser una red IP privada, una Intranet o Internet.

Los distintos elementos pueden residir en plataformas físicas separadas o bien pueden convivir varios elementos en la misma plataforma. De este modo es bastante habitual encontrar juntos servidor y gateway. La figura 2.1 muestra la estructura de una red VoIP.

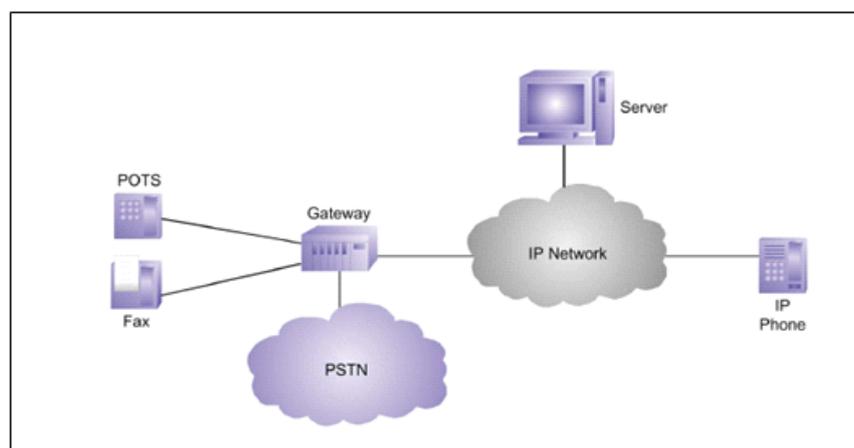


Figura 2.1 Estructura de una red VoIP [ROHE06]

2.4. Códecs de Audio

La señal de audio ha de ser digitalizada, comprimida y codificada antes de ser transmitida por la red IP. Para ello se utilizan algoritmos matemáticos implementados en software llamados códecs (acrónimo de codificador-descodificador aunque actualmente se le atribuye también las funciones de compresor-descompresor). Existen diferentes modelos de códecs de audio utilizados en VoIP, y dependiendo del algoritmo escogido en la transmisión, variará la calidad de la voz, el ancho de banda necesario y la carga computacional. El objetivo principal de esta tecnología es encontrar un equilibrio entre eficiencia y calidad de la voz. Aunque el sistema auditivo humano es capaz de captar las frecuencias comprendidas entre 20 Hz y 20 kHz, la gran mayoría de códecs procesan aquella información dentro de la banda de 400 Hz – 3,5 kHz [WIKI11] ya que a la hora de reconstruir la señal, esta sigue siendo inteligible.

Podemos dividir los códecs en tres categorías según su principio de funcionamiento:

- **Códecs de forma de onda:** Estos códecs se basan en almacenar información sobre la forma en el tiempo de la señal. Por lo general son los que más ancho de banda consumen, dado que no utilizan ninguna característica especial de la señal. Pueden ser utilizados para transmitir cualquier tipo de señal, no solamente voz (por ejemplo música, fax).
- **Códecs vocales:** Este tipo de códecs se basan en un modelo de cómo el sonido fue creado. Reconstruyen la señal solamente con la información del modelo, la forma de la señal es descartada por completo. El modelo es construido teniendo en cuenta el sistema vocal humano, básicamente un resonador y un tubo. Las desventajas que introducen estos códecs son: alta complejidad (cálculos necesarios) y dado que procesan la señal en tramas introducen un retardo significativo. Este tipo de códecs utilizan muy poco ancho de banda (2.4 Kb/s) pero se pierde completamente la naturalidad de la voz debido a la simplificación del modelo.
- **Códecs híbridos:** Dentro de esta categoría se encuentran los códecs que están a medio camino entre los anteriores. Incorporan elementos de los dos, resultando en tasas de transmisión más altas que los vocales pero menor que los de forma de onda. La complejidad es menor que los vocales y mayor que los de forma de onda. En términos de calidad se aproximan a los codificadores de forma de onda.

A continuación se enumeran y describen, entre otros, los códecs más utilizados en VoIP:

- **G.726:** este estándar de la ITU, también conocido como ADPCM (*Adaptive Differential Pulse Code Modulation*), sustituyó al obsoleto estándar G.721 en 1990. Permite trabajar a velocidades de 16 kbps, 24 kbps y 32 kbps. La gran ventaja de este códec es la disminución de ancho de banda requerido sin aumentar en gran medida la carga computacional [ITUT90].
- **G.723.1:** este algoritmo, estandarizado en 1995 por la ITU, puede operar a 6.3 kbps o 5.3 kbps. Si es utilizado en una aplicación comercial es necesario pagar una licencia [ITUT06].

- **G.722:** es una UIT-T estándar de 7 kHz de banda ancha. Este códec que opera en 48, 56 y 64 kbit/s fue aprobado por la UIT-T en noviembre de 1988. Dicho códec está basado en sub-banda ADPCM (SB-ADPCM) [ITUT88].
- **GSM (*Global System Mobile*):** estándar que opera a 13 kbps con una carga de CPU aceptable. No requiere el pago de una licencia [WIKI10].
- **Speex:** software libre creado por *Xiph.Org Foundation* que implementa un algoritmo capaz de variar la velocidad de transmisión dependiendo de las condiciones actuales de la red (VBR: *Variable Bit Rate*). El ancho de banda puede variar desde 2.15 a 22.4 kbps [SPEE11].
- **SILK:** Este códec puede utilizar una frecuencia de muestreo de 8 kHz, 12, 16 o 24 kHz y una tasa de bits 6 a 40 kbit/s. También puede utilizar un algoritmo de bajo retardo de 25 ms [SILK09].

En la figura 2.2 se muestra una tabla con características de varios códecs de audio.

Codificadores estandarizados ITU	Descripción	Ancho de banda (Kbps)
G.711	PCM	64
G.721	ADPCM	32, 16, 24,40
G.722	PCM-ADPCM	48, 56, 64,
G.728	LD-CELP	16
G.729	CS-ACELP	8
G.723.1	CELP	6.3, 5.3

Figura 2.2: Tabla con los codificadores estandarizados establecidos por la ITU.

2.5 Protocolos de Señalización [TELE10]

La señalización en VoIP tiene un papel muy importante en la red, ya que es la encargada de establecer, mantener, administrar y finalizar una conversación entre dos puntos. Además de ofrecer funciones de supervisión, marcado, llamada y retorno de tonos de progreso; también se encarga de proveer QoS en cada canal de transmisión. En los siguientes apartados se describe alguno de los protocolos más importantes utilizados en VoIP.

2.5.1 H.323

H.323 es una familia de estándares desarrollado por la ITU en 1996 con el objetivo de ofrecer un mecanismo de transporte para servicios multimedia sobre redes que no garantizan QoS, aunque su uso se ha extendido sobre todo al uso sobre redes IP. Pese a que inicialmente fue definido como un protocolo de videoconferencia, rápidamente ha ido evolucionando para cubrir todas las necesidades de la VoIP. De hecho el protocolo VoIP generaliza los conceptos introducidos por H.323. Además especifica aspectos basados en el sistema de señalización número 7 (SS7) para la interconexión con la PSTN [ITUT09].

Se trata de una recomendación bastante cerrada donde se define los códecs a utilizar, tanto en audio como en video, y los protocolos de transporte de la información. De hecho fue el primer estándar en adoptar como medio de transporte el protocolo RTP, siendo capaz de aplicar algoritmos de cifrado de la información, evitando de esta manera añadir elementos de seguridad adicionales a los requeridos para la conexión a Internet.

Pese a que técnicamente es un protocolo potente y maduro, el interés por parte de los usuarios y empresas actualmente ha disminuido debido principalmente a su complejidad y a ciertas ineficiencias detectadas en conferencias entre un número elevado de terminales.

2.5.2 SIP

SIP (*Session Initial Protocol*) es un protocolo desarrollado por el IETF en 1999 para el control de llamadas multimedia y la implementación de servicios telefónicos avanzados [WIKI09].

SIP está basado en HTTP (*HyperText Transfer Protocol*) adoptando las características más importantes de este estándar como son la sencillez de su sintaxis y una estructura cliente/servidor basada en un modelo petición/respuesta. Otra de las ventajas de SIP es su sistema de direccionamiento. Las direcciones SIP tienen una estructura parecida a la de un correo electrónico dotando a sus clientes de una alta movilidad facilitando una posible integración en comunicaciones móviles. Cabe destacar que aunque originalmente SIP tenía como objetivo la simplicidad, en su estado actual se ha vuelto tan complejo como H.323.

Pero el gran potencial de SIP reside en su flexibilidad ya que ofrece la posibilidad de programar nuevos servicios no definidos por la propia recomendación. Entornos de programación como CGI (*Common Gateway Interface*) o sencillos lenguajes de programación como CPL (*Call Processing Language*) son algunas de las herramientas para la implementación de servicios sin que conlleve a un peligro para la integridad del sistema. Esta es la característica principal por la que SIP actualmente goza de un mayor éxito que H.323.

Los clientes SIP llamados *peers* o *user agents* usan el puerto 5060 en TCP (*Transmission Control Protocol*) y UDP (*User Datagram Protocol*) para conectar con los servidores SIP. SIP es usado simplemente para iniciar y terminar llamadas de voz y video. Todas las comunicaciones de voz/video van sobre RTP.

2.5.3 MGCP-MEGACO

Media Gateway Control Protocol (MGCP) es otro estándar de señalización para VoIP desarrollado por la IETF. MGCP está basado en un modelo maestro/esclavo donde el *Call Agent* (servidor) es el encargado de controlar al gateway. De esta forma se consigue separar la señalización de la transmisión de la información, simplificando la integración con el protocolo SS7 [CISC11].

Esta importante ventaja propició la colaboración conjunta entre el IETF y la ITU para el desarrollo de una nueva especificación basada en MGCP que fuera complementaria a SIP y H.323. El resultado fue MEGACO [ITUT01] aunque la ITU se refiere a este protocolo

como H.248 [ITUT01]. En definitiva, SIP y H.323 se utilizan para la señalización en los extremos, mientras que MEGACO es óptimo para los grandes operadores de telefonía.

2.5.4 IAX

Inter-Asterisk eXchange protocol (IAX) fue desarrollado por *Digium* para la comunicación entre centralitas basadas en *Asterisk* aunque actualmente se ha implementado clientes que también soportan este protocolo [WIKI08].

El principal objetivo de IAX es minimizar el ancho de banda utilizado en la transmisión de voz y vídeo a través de la red IP y proveer un soporte nativo para ser transparente a los NATs (*Network Address Translation*). La estructura básica de IAX se fundamenta en la multiplexación de la señalización y del flujo de datos sobre un simple puerto UDP, generalmente el 4569.

El protocolo original ha quedado obsoleto en favor de su segunda versión conocida como IAX2. Se caracteriza por ser robusto y simple en comparación con otros protocolos. Permite manejar una gran cantidad de *códecs* y transportar cualquier tipo de datos.

CAPÍTULO 3

Medida de la Calidad de Voz en Redes VoIP

La VoIP enfrenta problemáticas propias de las redes de datos, que se manifiestan como degradaciones en la calidad del servicio percibida por los usuarios.

Estas degradaciones pueden deberse por ejemplo a retardos, jitter (diferencia de retardos) y pérdida de paquetes, entre otros factores. Para que la tecnología de VoIP pueda ser utilizada en las empresas, es esencial garantizar una calidad de voz aceptable. Para ello se han desarrollado métodos para medirla. Estos métodos se dividen en *subjetivos* y *objetivos*. Los métodos *subjetivos* de medida de la calidad de servicio, se basan en conocer directamente la opinión de los usuarios. Típicamente resultan en un promedio de opiniones por ejemplo, el valor MOS (Mean Opinión Score). Los métodos *objetivos*, a su vez se subdividen en *intrusivos* (se inyecta una señal de voz conocida en el canal y se estudia su degradación a la salida) y *no intrusivos* (monitorean ciertos parámetros en un punto de la red y en base a estos permite establecer en tiempo real la calidad que percibiría un usuario). En la figura 3.1 se puede ver un esquema gráfico sobre los métodos de asignación de calidad.

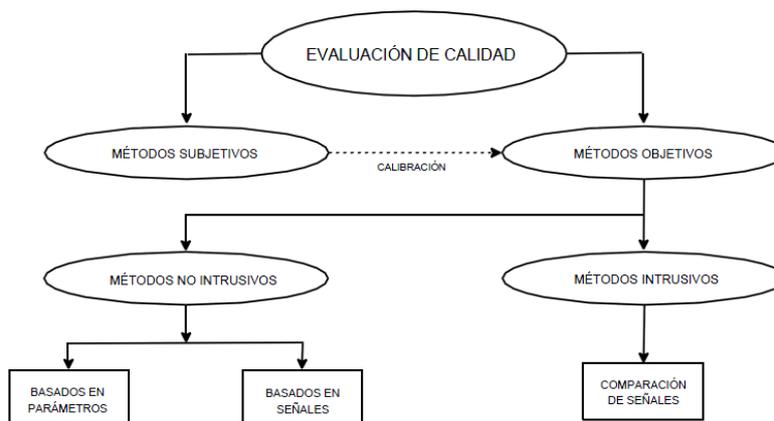


Figura 3.1: Clasificación de los métodos de asignación de calidad

3.1 Calidad de Servicio (QoS)

El concepto de QoS (*Quality of Service*) representa hoy en día un elemento fundamental a la hora de hablar de venta y consumo de servicios. Si bien no es un concepto nuevo, es muy fácil escuchar cada vez más la palabra QoS. La razón es simple: los consumidores de servicios son cada día más exigentes y la necesidad de poder cuantificar la calidad ofrecida se ha tornado un tema fundamental.

QoS hace referencia a la capacidad que tiene un sistema de asegurar que se cumplen los requisitos de tráfico para un flujo de información determinado, y puede definirse como el

valor de un conjunto de parámetros de prestaciones que aseguran al usuario de un servicio niveles aceptables de calidad. Como distintos tipos de servicio mantienen características particulares, cada uno tendrá su propia QoS. Por ejemplo, en el caso de telefonía tradicional se puede definir QoS como el tener un canal de 64Kbps durante el tiempo que dure la conversación y una disponibilidad de servicio de 99,999 %. En el caso de Internet, las características heterogéneas de los distintos servicios que transporta hacen del tema calidad de servicio un problema mayor. Es difícil identificar en cada caso cuales son los parámetros de prestaciones que aseguran niveles aceptables de calidad.

En la RFC 2475 (*An Architecture for Differentiated Services*) se establecen algunas características significativas, cualitativas o estadísticas, a tener en cuenta en la transmisión de un paquete en una dirección a través de una o más rutas en una red: caudal (*throughput*), demora (*delay*), variación de demora (*jitter*), pérdidas, etc.

Es conocido que Internet ofrece un único nivel de servicio, *Best Effort*, donde no existe una preasignación de recursos, ni plazos conocidos, ni garantía de recepción correcta de la información. Sencillamente, se ofrece el mejor nivel de servicio posible en ese momento.

La QoS tiene como objeto solventar los problemas que han ido surgiendo en Internet y en las redes IP en general ante las nuevas aplicaciones y acceso masivo de usuarios. La demanda de servicios garantizados, y el despliegue de aplicaciones multimedia y servicios síncronos como voz y video en tiempo real, han generado la necesidad de definir y establecer calidad de servicio en las redes para el buen funcionamiento de estas aplicaciones.

Para el usuario final, la QoS implica tener un nivel de conexión y servicio en términos de rapidez, fiabilidad, rendimiento y disponibilidad, además de un servicio de atención al cliente con respuesta ágil y eficaz.

La VoIP ha migrado el tráfico de voz que de forma tradicional utilizaba como soporte una red de conmutación de circuitos (RTPC), a una red de conmutación de paquetes. Esto implica que la información de voz es fragmentada creando un flujo de paquetes independientes que viajan por diferentes caminos de la red, llegando al destino de forma desordenada y con diferentes retardos acumulados. Debido a esto, en la integración de la voz y los datos sobre una estructura única de conmutación de paquetes, existen algunas limitaciones que deben ser consideradas en el diseño e implantación de una solución de este tipo:

- El ancho de banda necesario para la transmisión de las comunicaciones de voz.
- El retardo con el que llegan los paquetes.
- La variación del retardo en la transmisión (*jitter*).
- Las pérdidas de paquetes.

Con el objetivo de paliar los problemas que generan estas limitaciones, QoS actúa a diferentes niveles:

- Nivel de dispositivo de la red a través del uso de colas, la planificación y la adaptación del tráfico.
- Nivel de señalización para coordinar la QoS de extremo a extremo.
- Nivel de gestión para controlar y administrar el tráfico extremo a extremo.

Por tanto, en términos cuantitativos la calidad de servicio se refleja en una serie de parámetros o factores que pueden ser medidos y ajustados para satisfacer el grado de servicio demandado.

Los factores que se enumeran a continuación determinan la calidad de servicio:

- El ancho de banda que tiene la red para cursar llamadas puede provocar retardos.
- La limitación del ancho de banda y la congestión de los *routers*, provocan pérdida de paquetes.
- El procesado al que es sometida la voz y el recorrido que realizan los paquetes a través de la red provocan retardos.
- Como consecuencia de que cada paquete se transmite de forma independiente al resto se produce *jitter*.
- El acoplo que sufre la señal de voz entre los distintos sentidos de la comunicación genera eco.

3.1.1 Ancho de Banda

El ancho de banda (*BandWidth*, BW) va a establecer la velocidad máxima de transferencia de datos entre dos extremos de la red.

Relativo a la VoIP, en función del algoritmo para la codificación de la voz y de la compresión del códec utilizado, existen unos requerimientos mínimos para la transmisión de la señal de voz en cuestiones de BW.

El ancho de banda disponible en la red se reparte entre las aplicaciones que hay sobre ésta. Debido a ello y con objeto de arbitrar este reparto, han surgido algunos mecanismos de QoS. La técnica empleada para la gestión del ancho de banda en la red es determinante en el retardo que sufrirán los paquetes, y en consecuencia, en la calidad de la voz.

3.1.2 Pérdida de Paquetes

La pérdida de paquetes es un fenómeno común en las redes de conmutación de paquetes. La información al ser fragmentada y viajar de forma independiente por la red puede atravesar en su camino hacia el destino diferentes equipos. En este contexto, cuando se produce congestión en el *router* (se han llenado las colas y el equipo no puede aceptar más paquetes), se producen pérdidas. Para el tráfico en tiempo real como la voz, la retransmisión de tramas perdidas en la capa de transporte no es práctica por ocasionar retardos adicionales, derivando de esto que sea habitual el transporte UDP. Trabajando con este protocolo, los paquetes que no son recibidos en el destino no son reenviados.

El problema de la pérdida de paquetes puede quedar resuelto cambiando el *router* y/o sustituyendo el enlace por uno de mayor capacidad, pero puede suponer una solución cara por lo que se buscan otras alternativas.

Cada paquete contiene de forma aproximada, pues depende del esquema de codificación utilizado, unos 20 ms de señal vocal que se corresponden con la duración media de un fonema. Debido a esto, la pérdida de un paquete en la transmisión representa la pérdida de un fonema en el proceso de reproducción de la señal de voz. En consecuencia, la pérdida de paquetes de voz es una pérdida de calidad de la señal. Cuando esta tasa de pérdida es pequeña, el cerebro puede reconstruir los fonemas perdidos, lo que permite que la calidad en la conversación pueda ser aceptable dejando de ser así para pérdidas superiores. Por esto existen mecanismos que mitigan estos efectos como pueden ser: técnicas de corrección, distribución, contención y recuperación.

El impacto de la pérdida de paquetes es medido en términos de I_e (que puede ser observado en la figura 3.2), el factor de deterioro, y cuanto más grande sea éste, el empeoramiento de la señal es más severo. En la siguiente tabla (figura 3.2) se muestra el impacto de las tramas perdidas en el factor de deterioro, según la ITU-T G.113. Destacar que incluso con un 0% de pérdida de paquetes, los códecs G.729A y G.723.1 tienen un factor de deterioro de 11 y 15, respectivamente.

Codec	I_e (0% pérdida)	I_e (2% pérdida)	I_e (5% pérdida)
G.711 sin PLC	0	35	55
G.711 con PLC	0	7	15
G.729A	11	19	26 (valor para 4% pérdidas)
G.723.1 (6,3 Kbps)	15	24	32 (valor para 4% pérdidas)

Figura 3.2: Impacto de las tramas perdidas sobre I_e

3.1.3 Retardo

Los retardos que se producen en la transmisión de los paquetes de la señal de voz causan dos problemas:

- **Eco.** Es consecuencia de las reflexiones de la señal. Cuando el retardo es superior a cierto umbral, que la ITU sitúa en 5ms, el hablante empieza a escuchar una versión molesta y retardada de sus propias palabras, y si este retardo alcanza niveles muy elevados, se llega al punto en que mantener la conversación resulta imposible.
- **Talker overlap** (solapamiento de la voz de los interlocutores). Durante la conversación se establecen pausas que invitan al otro interlocutor a dar una

respuesta, si durante un tiempo razonable no ha llegado la respuesta, el hablante original puede seguir hablando. En este caso, si llega la respuesta del segundo interlocutor debido al retraso, se da una colisión que no permite la comunicación. A partir del umbral de retardo de 150 ms aproximadamente empieza a aparecer este fenómeno.

Por estos inconvenientes, el retardo de extremo a extremo es un factor crítico en el diseño de una red de transporte de VoIP. Normalmente, la calidad de la voz empieza a degradarse a partir de retardos superiores a 150 ms, pero si condiciones extraordinarias lo requieren, los usuarios habitualmente pueden aceptar hasta 400 ms.

Las principales fuentes de retardo en la VoIP son dos:

- **Retardo debido al gateway o pasarela.** El procesado de la voz puede provocar diferentes tipos de retardos:
 - ✓ Retardo de procesamiento del códec. La digitalización de la voz y la compresión de la señal repercute en un retardo inevitable.
 - ✓ Retardo de serialización. Relacionado directamente con la tasa del reloj de la transmisión, es el tiempo requerido para transmitir un paquete IP.
 - ✓ Retardo de empaquetamiento. Referente al tiempo necesario para llenar un paquete de información (carga útil) con la conversación ya codificada y comprimida, depende del tamaño de bloque requerido por el codificador de voz y el número de bloques de una sola trama.
 - ✓ Retardo de supresión de *jitter*, como se puede observar en la figura 3.3. Con objeto de eliminar la variabilidad del retardo, se realiza un almacenamiento temporal del flujo de paquetes antes de recuperar la forma de onda de la señal de voz en un *buffer*.

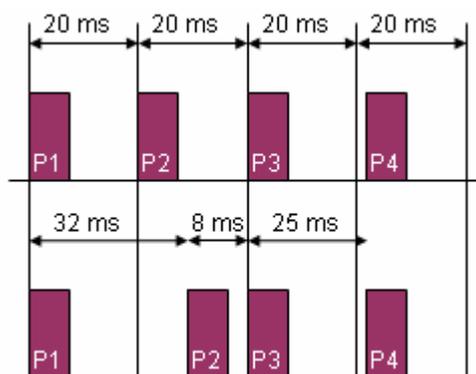


Figura 3.3: Ejemplo de *jitter*

- **Retardo debido a la red.** Los paquetes con la información viajan por la red y sufren de fuentes de retardo adicionales:
 - ✓ Retardo de encolado. Ocasionado por la espera de los paquetes para ser transmitidos por una línea de salida. Existen diferentes maneras de reducirlo a través del empleo de líneas con mayor ancho de banda o mediante técnicas de gestión de BW.

- ✓ Retardo de conmutación de la red. Es referente al tiempo que necesita la red para transportar la información de origen a destino y hay de diferentes tipos: retardo de acceso al medio, de propagación o asociado a servicios de red.

3.1.4 Factor de Compresión

Para poder transmitir la voz a través de una red de datos, es necesario realizar previamente un proceso de digitalización. En telefonía clásica, éste proceso se realiza utilizando CÓDECs, obteniendo una señal digital de 64 kb/s. Este proceso, se realiza de acuerdo a la recomendación G.711 de la ITU-T. Sin embargo, cuando se dispone de velocidades de red reducidas, es conveniente tratar de minimizar el ancho de banda requerido por las señales de voz. Para ello, se han desarrollado varias recomendaciones, que reducen la velocidad de transmisión requerida, a expensas de degradar la calidad de la voz.

La Figura 3.4 resume las recomendaciones de la ITU-T respecto a los algoritmos estandarizados de compresión de voz:

Algoritmo	Descripción
G.711	Audio encoding at 64 k bit/s (μ -law and A-law)
G.722	7 kHz speed at 48, 56 and 64K bit/s (hi-fi voice)
G.723.1	Dual Rate Speed at 6.4 and 5.3 k bit/s
G.728	16 k bit/s speech
G.729 Annex A	8 k bit/s speech (Conjugate structure- algebraic code excited linear prediction or CS-ACELP). Reduce Complexity
G.729 Annex B	8 k bit/s speech (Conjugate structure- algebraic code excited linear prediction or CS-ACELP). Silence Compression
G.729 Annex AB	8 k bit/s speech (Conjugate structure- algebraic code excited linear prediction or CS-ACELP). Reduce Complexity & Silence Compression

Figura 3.4: Recomendaciones de la ITU-T

3.2 Calidad Experimentada por el Usuario (QoE)

Una definición de QoE sería cómo se perciben los resultados totales de un sistema desde el punto de vista del usuario. QoE mide la actuación final de todos los servicios desde la perspectiva del usuario. Es una medida que refleja cómo de bien la red satisface las necesidades del usuario final. No obstante, la QoE está relacionada con la QoS, la cual trata de realizar objetivamente mediciones del servicio entregado. Mientras que la QoE es una medida subjetiva de la experiencia que tiene el cliente de un servicio ofrecido.

El término de la calidad en la experiencia del usuario, QoE, comienza a ser importante a partir de los años 90 cuando el usuario, la experiencia que éste posee de la aplicación y el diseño centrado en sus necesidades comienzan a ser importantes para las empresas. Las empresas buscan la QoE debido a las ventajas que este factor les aporta. Si una aplicación proporciona su servicio con una alta QoE, los clientes estarán satisfechos, de lo contrario

se le estará aportando al cliente una gran diferencia entre lo que él necesita y lo que la aplicación le ofrece. Si no se alcanza la QoE esperada por el usuario, se estará propiciando una pérdida de clientes ya que buscarán otras aplicaciones que le aporten la QoE que ellos esperan. Podemos determinar que hoy en día la calidad es una necesidad debido a la competición que existe entre aplicaciones.

Las medidas realizadas para determinar la QoE deben corresponderse directamente con la calidad que percibe el usuario de lo contrario la QoE hallada no se correspondería con la real. Existen numerosos factores que afectan a la calidad de la voz en la comunicación y como consecuencia a la QoE y a la QoS. A continuación los enumeraremos:

- **Latencia:** los paquetes de VoIP son muy sensibles a los retardos que se producen en la red.
- **Jitter:** Los paquetes de VoIP se encaminan al igual que el resto de paquetes por distintos nodos, sin embargo estos paquetes necesitan llegar en orden y que la diferencia de tiempo entre paquetes sea similar a la diferencia con la que se transmitieron.
- **Pérdida de paquetes:** la voz se transmite en paquetes, si estos paquetes se pierden, también se pierde el trozo de conversación que se transportaba en ellos.
- **Inestabilidad de las llamadas:** se produciría cuando una vez establecida una llamada, se produce inesperadamente una finalización de la conexión. Se debe al fallo de algún equipo o a la congestión de la red.
- **Bloqueo de las llamadas:** también se debe a la congestión de la red y se produce cuando se quiere iniciar una llamada pero esta no se consigue conectar.

Algunos de los factores que producen impacto en la QoE son:

- Factores Humanos: la expectación que tienen de la aplicación, la edad del usuario, la familiaridad que tiene con las aplicaciones,...
- Compresión y codificación de la voz: el algoritmo usado para comprimir la voz, el algoritmo para codificar la voz, la supresión del silencio, la cancelación del eco...
- Temas de la red: retraso, variación de retraso, paquetes perdidos, disponibilidad de la red
- Características de servicio: ancho de banda disponible para transmitir la llamada, características de las llamadas, número de llamadas, *Wireless* o conexión de cable.

En la figura 3.5 se muestra un esquema que complementa la información anterior.

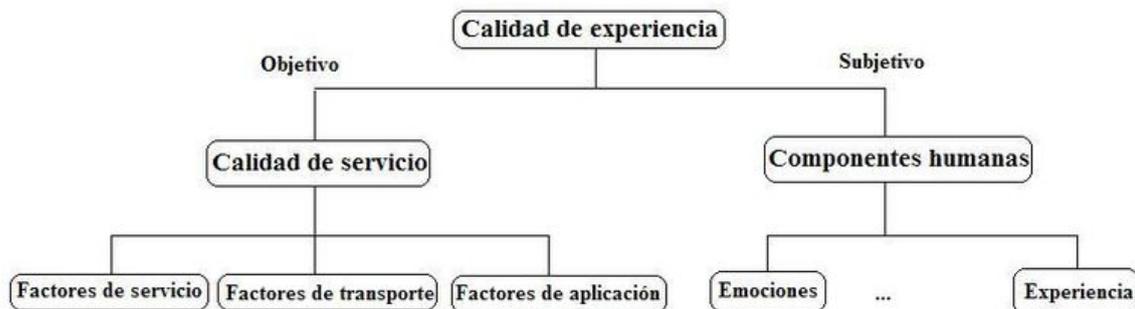


Figura 3.5: Factores de la calidad de experiencia.

3.2.1 Metodologías de Estimación de Calidad de Experiencia

3.2.1.1 Métodos Subjetivos

Los distintos métodos subjetivos están normalizados por la ITU en las recomendaciones ITU BT.500 [ITUR02] para video e ITU P.800 [ITUP96] para audio.

Básicamente se pueden clasificar en *Absolute Category Rating* (ACR) que tiene como resultado el *Mean Opinion Score* (MOS), y *Degradation Category Rating* (DCR) que tiene como resultado el *Degradation Mean Opinion Score* (DMOS).

Los distintos tests son realizados normalmente en condiciones controladas en un laboratorio (por ejemplo cuartos aislados acústicamente). También se requiere mucho cuidado en el momento de definir las condiciones e interpretar los resultados.

En el Anexo A se presentan de forma más detallada los distintos métodos subjetivos explicados en las recomendaciones de la ITU para audio, ya que el audio es en lo que se centra este proyecto.

Absolute Category Rating (ACR)

En este tipo de test, los participantes deben asignar un valor global de calidad a la señal (imagen, video o audio) que se les presenta (por lo general ya transmitida), sin tener acceso a la señal original. De ahí el nombre de absoluto. Los valores de calidad se asignan de acuerdo a la siguiente tabla (Figura 3.6).

Valor de calidad	Calidad de la señal
5	Excelente
4	Buena
3	Regular
2	Pobre
1	Mala

Figura 3.6 : Valores de calidad según ACR

El valor medio asignado por los participantes es el MOS. En el caso que se simulen conversaciones el resultado es el MOSc.

Degradation Category Rating (DCR)

Cuando las señales son de buena calidad, los métodos ACR son insensibles a los pequeños cambios de calidad. En este tipo de caso se utilizan los métodos tipo DCR, en los cuales a los participantes se les presentan dos señales y deben asignar un valor a la degradación de la calidad de una respecto de la otra, de acuerdo a la siguiente escala (figura 3.7):

Valor de calidad	Nivel de Degradación
5	Imperceptible
4	Perceptible pero no molesta
3	Un poco molesta
2	Molesta
1	Muy molesta

Figura 3.7: Valores de degradación de calidad según DCR

El valor medio asignado por los participantes es el DMOS.

Existen diversas variantes en este tipo de método dependiendo de las señales que se presenten juntas. Lo más común es presentar la señal original y luego la distorsionada (ya transmitida). Esto permite medir la fidelidad del sistema de comunicación utilizado.

Otros métodos Subjetivos

Se han propuesto recientemente nuevos métodos para una mejor asignación de calidad en servicios multimedia con gran variabilidad de calidad en el tiempo. Para ello se realiza una asignación continua de calidad por parte de los participantes mediante el uso de cursores electrónicos.

3.2.1.2 Métodos Objetivos

➤ **Métodos Intrusivos de estimación de QoE**

Los métodos intrusivos normalmente utilizan dos señales de entrada, una de referencia (original) y una distorsionada (ya transmitida). Son considerados intrusivos debido a la introducción de señales auxiliares y la utilización de la red.

Existe una gran variedad de métodos tanto para audio como para video. Se pueden clasificar en dos grandes grupos. Están los que realizan comparaciones en el dominio del tiempo, como el *Signal to Noise Ratio* (SNR) o su versión para imagen y video PSNR. Estos métodos son muy simples de implementar, pero la correlación con las medidas subjetivas no es muy buena. El segundo grupo realiza medidas relevantes a la percepción, transformando las señales al dominio de la percepción utilizando modelos de la percepción humana, ya sea de la audición o de la visión. Estos métodos son más complejos que los primeros pero presentan mejor correlación con los métodos subjetivos.

Métodos típicos de medida en el dominio de la percepción son el *Perceptual Speech Quality Measure* (PSQM), *Measuring Normalizing Blocks* (MNB), *Enhanced Modified Bark Spectral Distorsion* (EMBSD) y *Perceptual Evaluation of Speech Quality* (PESQ) para audio. Para video existen unos cuantos algoritmos como el basado en el índice de similitud estructural (SSIM), basado en distorsión espacio-temporal del *Institute for Telecommunication Science* (ITS), pero difieren en qué es lo importante a la percepción. Cabe señalar que a diferencia del audio, donde el modelo de percepción está bastante aceptado, en imágenes y video el conocimiento del mismo es limitado lo que hace que los distintos algoritmos tengan supuestos diferentes sobre la visión.

El resultado que se obtiene con estos algoritmos es una medida de la distorsión relevante a la percepción. Es necesario llevar estos valores a una escala común, que permita compararlos entre sí y que además sea de fácil comprensión para las personas. La escala que surge naturalmente es la utilizada en los métodos subjetivos. Para unificar las escalas se ajusta de forma paramétrica la relación entre valores subjetivos y objetivos de cada método. Por lo tanto es necesaria la realización de test subjetivos para todas las secuencias utilizadas.

En el caso del PESQ, este incluye el ajuste, por lo que la salida se da directamente en escala de DMOS.

➤ **Métodos No Intrusivos de estimación de QoE**

A diferencia de los métodos intrusivos antes presentados donde el servicio debe ser interrumpido para inyectar las señales, los métodos no intrusivos pueden ser utilizados durante el servicio. Aquí cabe aclarar que no siempre es posible utilizar estos métodos en servicio, debido a que si bien no utilizan señales extra, sí pueden inyectar algún tipo de tráfico para estimar el estado de la red.

Estos métodos se pueden clasificar en basados en parámetros o basados en señales. Los últimos predicen la calidad utilizando la señal distorsionada sin necesidad de referencia. A este tipo de método se los denomina *Null Reference*. Los otros predicen la calidad a partir del valor de parámetros de la red IP (por ejemplo probabilidad de pérdida, jitter, retardo) y de parámetros no específicos de la red (códec utilizado, eco, tasa de bits del video, etc). Ejemplo de estos métodos son el E-Model y el uso de redes neuronales.

El E-Model es un modelo empírico matemático estandarizado por la ITU en la recomendación G.107 [ITUT05]. Es un conjunto de fórmulas que tienen como entrada parámetros de la red tradicional de circuitos conmutados y de la red de paquetes conmutados, y tiene como salida el factor de calidad el cual se puede mapear en MOSc. Si bien es una herramienta para la planificación de redes, actualmente es muy utilizada para predecir calidad percibida en VoIP.

Las redes neuronales se utilizan para aproximar la relación no lineal que existe entre calidad percibida (mejor dicho MOS) y el conjunto de parámetros considerado. Un conjunto de parámetros de entrada posible sería el formado por: la probabilidad de pérdida, retardo, jitter, códec utilizado, tasa de bits del video, lenguaje en audio, etc. Para lograr el mapeo deseado se debe generar una base de entrenamiento que consiste en un conjunto de valores de los parámetros y el correspondiente valor de calidad obtenido mediante tests subjetivos.

El obtener una buena base de entrenamiento, es decir un rango considerable de variación de los parámetros, es el principal limitante debido al costo de los tests subjetivos. El E-Model se analiza con mayor profundidad en el anexo B.

CAPÍTULO 4

Desarrollo de las Pruebas y Obtención de Resultados

4.1 Entorno y Características de los Componentes

A continuación se describen las características técnicas de todos los elementos que intervienen en la realización de las pruebas y por tanto son relevantes en la obtención de los resultados.

Este proyecto se ha realizado en un entorno cableado Ethernet como se muestra en la figura 4.1.

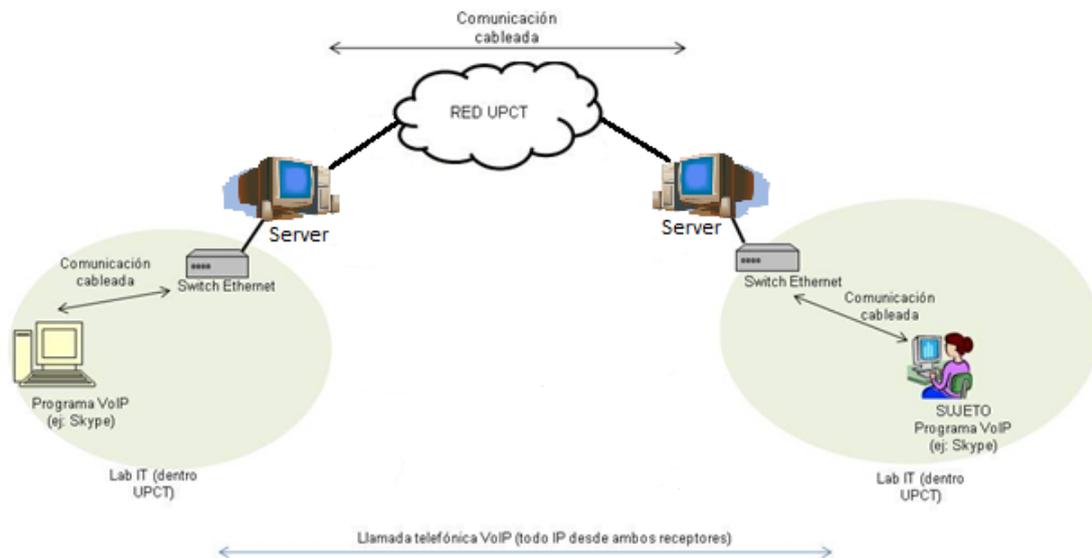


Figura 4.1: Esquema del escenario de pruebas

Concretamente se ha realizado entre los laboratorios IT-3 e IT-4 de la UPCT. Se ha comprobado mediante la utilización del comando *tracert* en la consola de Windows que entre el emisor y el oyente hay tres saltos, tal y como se muestra en la figura 4.1. Los servidores de los laboratorios son los nodos que separan al emisor del receptor.

Para la correcta obtención de los cálculos observando los paquetes intercambiados en la conversación de VOIP (se verá más adelante) cabe destacar que los servidores de estos laboratorios tiene las siguientes direcciones ip's públicas.

- **IT-3:** 212.128.44.132
- **IT-4:** 212.128.44.50

Es importante resaltar que en el caso de la utilización de Skype la comunicación entre ambos extremos de la conversación es directa, mientras que en la utilización de Jitsi se utiliza un nodo “retransmisor” que actúa de intermediario entre ambos extremos de la

conversación. Dicho “retransmisor” se encuentra a 18 saltos del extremo emisor de la conversación.

Parte imprescindible de este proyecto es la utilización de los dos software mencionados anteriormente (Skype [SKYP11] y Jitsi [JITS11]) y otro programa, llamado Wireshark [WIRE11], utilizado para capturar los paquetes intercambiados en la conversación de VoIP para realizar el análisis de QoS. La instalación y configuración de dichos programas se muestran en el anexo C.

4.1.1 Grabaciones de Señales en la Fuente

En primer lugar se han realizado las grabaciones. Para eliminar variaciones no deseadas en la fuente de las señales vocales, en primer lugar se han grabado muestras de señales vocales con las propiedades normalizadas deseadas, de la forma siguiente (Ubicación: Estudio Polimedia, Universidad Politécnica de Cartagena), ver anexos F, G y Z con la descripción técnica de los equipos. Se han grabado un total de cuatro grabaciones (dos de hombre y dos de mujer).

4.1.1.1 Entorno de Grabación

El hablante se sentó en una sala silenciosa, de un volumen entre 30 y 120 m³ y con un tiempo de reverberación inferior a 500 ms. El nivel de ruido de sala fue inferior a 30 dBA, sin crestas dominantes en el espectro.

4.1.1.2 Sistema de Grabación

El sistema de grabación debe ser de alta calidad (de estudio) y puede adoptar una de las siguientes formas:

- a) Magnetófono convencional de dos pistas. Debe indicarse el tipo de ecualización, pero se recomienda utilizar el de la CEI. En todo momento deben utilizarse cintas de alta calidad (de bajo nivel de ruido y baja transferencia magnética).
- b) Procesador de audio digital de dos canales equipado con un magnetoscopio de alta calidad o un equipo de cinta magnética de audio digital (DAT, *digital audio tape*).
- c) Sistema de almacenamiento digital controlado por computador.

El tercer sistema es el más adecuado y resulta muy flexible, pero por razones de orden práctico a menudo se elige uno de los otros dos. En estos sistemas una de las dos pistas debe emplearse para grabar la conversación y la otra para insertar señales de control con nivel y frecuencia determinados, a fin de evitar problemas de diafonía.

4.1.1.3 Material de Conversación

El material de conversación consistió en frases sencillas, breves y con significado, elegidas al azar y fáciles de entender. Con estas frases se confeccionó una lista de forma aleatoria, de tal manera que no hubiese ninguna conexión evidente entre el significado de una frase y la siguiente. Se evitaron las frases muy cortas y muy largas pues el objetivo es que cada frase pudiera incluirse en un intervalo de tiempo de 2 a 3 segundos.

4.1.1.4 Procedimiento de Grabación

Las señales vocales se grabaron mediante un micrófono lineal y un amplificador de bajo nivel de ruido con respuesta en frecuencia uniforme, como indica la Publicación 581-5 de la CEI. El micrófono se colocó a una distancia entre 140 mm y 200 mm de los labios del participante.

La relación entre el nivel vocal activo y el nivel de ruido ponderado sofométricamente (para la definición véase 8.2.3/P.830), SNR(p), en el medio de grabación tiene un valor > 40 dB, tomándose como objetivo el valor de 50 dB.

4.1.1.5 Participantes en las Pruebas

Participaron en las grabaciones cuatro personas (dos hombres y dos mujeres).

Los hablantes pronunciaron las frases de manera fluida pero sin énfasis y no presentaron deficiencias de dicción como tartamudeo; adoptaron un volumen de voz que les resulte confortable y que pudo ser mantenido de forma constante.

Como los procesos complejos a menudo afectan de forma distinta a las voces femenina y masculina, el plan del experimento contempla dos tipos de voces como factor de equilibrio; las notas para las señales vocales femenina y masculina se evalúan diferenciadamente.

Además, para disminuir el riesgo de que los resultados dependan en gran medida de las peculiaridades de las voces elegidas, es fundamental el haber utilizado más de una voz masculina y más de una voz femenina, de acuerdo con un plan equilibrado. En concreto se han utilizado dos voces masculinas y dos femeninas.

4.1.1.6 Niveles de la Señal Vocal

Una vez completadas, las grabaciones se reprodujeron y se midió el nivel vocal activo de cada frase con un medidor conforme a la Recomendación P.56. A continuación, se volvieron a grabar las listas en un segundo sistema introduciendo los ajustes de ganancia necesarios para que cada grupo de frases tuviera el nivel vocal activo normalizado, manteniendo las relaciones de tiempo adecuadas entre las frases y las señales de tono en el otro canal.

Para las señales vocales en banda estrecha, el nivel normalizado se obtiene midiendo y ajustando directamente la señal de banda estrecha grabada; el valor objetivo recomendado es -26 dB (+ 0,5 dB) con respecto al nivel de saturación de cresta del sistema de grabación.

4.1.1.7 Sistema Emisor

Las grabaciones fueron reproducidas en un ordenador portátil utilizando, en nuestro caso los dos programas de VoIP (Skype y Jitsi), usando los códecs SILK y G.722 respectivamente, para establecer la conexión telefónica con el extremo receptor, cuyas características se explicarán más adelante.

Como lo que se pretendía transmitir eran las grabaciones y nada por micrófono, se desactivó el micrófono del equipo, y se configuró el equipo y cada programa para que se transmitiera en la conversación el sonido que se reprodujera en el equipo, lo cual se

consiguió utilizando como dispositivo de entrada de sonido la Mezcla Estéreo del equipo. Con esto se evitaba la transmisión de sonidos indeseados.

Las características más relevantes del equipo transmisor se muestran en la figura 4.2:



Figura 4.2: Características del equipo transmisor en la conversación de VoIP.

Destacar también que la tarjeta de sonido utilizada es una *NVIDIA High Definition Audio*.

4.1.2 Concepción del Experimento

Para el diseño del experimento, además de cumplir con lo expuesto anteriormente, también hubo que tener en cuenta el efecto del orden de reproducción de las grabaciones. Como las pruebas se hicieron para un total de 60 personas, se reprodujeron las grabaciones en orden diferente para cada grupo de 20 personas, es decir, para las 20 primeras personas se reprodujo las grabaciones siendo las dos grabaciones de los hombres primero y las dos de las mujeres después; para las 20 siguientes personas se reprodujo primero las dos de las mujeres y las dos de los hombres después; y para las últimas 20 personas se reprodujeron de forma intercalada.

Para un grupo determinado de participantes, la prueba viene limitada en tiempo por la máxima duración posible de la sesión sin que aparezca el cansancio. Idealmente ninguna sesión debía durar más de 20 minutos y en ningún caso debe rebasar los 45 minutos. En este proyecto cada oyente (encuestado) escuchó dos veces el grupo de grabaciones (cada grabación duró 5 min. y 17 seg.) que le correspondió (una primera vez para evaluar el códec SILK utilizado por el software Skype y otra vez para evaluar el códec G.722 utilizado por el software Jitsi), por lo que cada sesión, en ninguno de los casos, duró más de 15 minutos.

4.1.3 Procedimiento de la Prueba de Escucha

4.1.3.1 Sistema de escucha

En el extremo receptor, que es donde se encuentra el oyente, se utilizó unos auriculares con las siguientes características:

- Respuesta en frecuencia (Hz):** 18-20000
- Potencia entrada máx. (mW):** 200
- Cable (m):** 2,5 (99,99% OFC, unilateral)
- Eficiencia (dB SPL/V):** 115
- Impedancia (ohms):** 32

Estos auriculares estaban conectados a un ordenador con las características que se muestran en la figura 4.3:

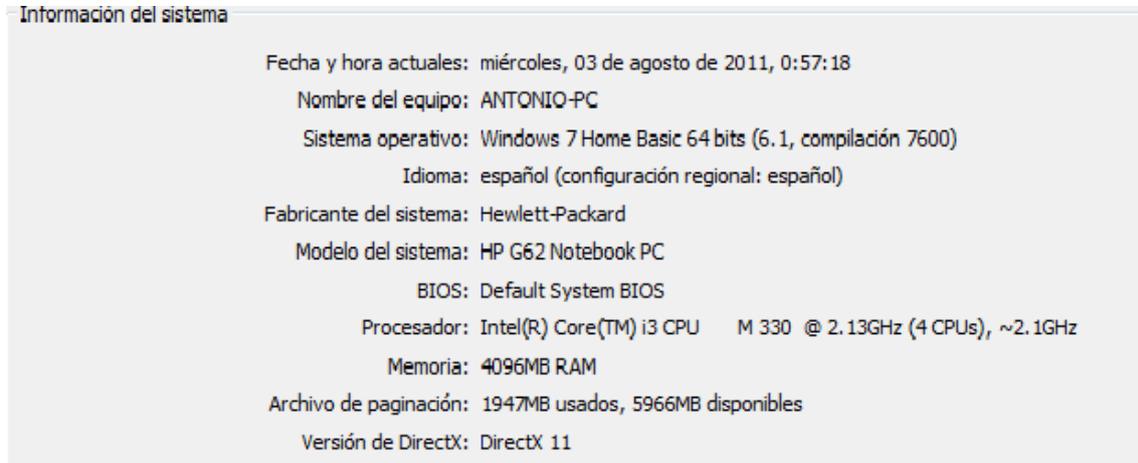


Figura 4.3: Características del equipo receptor en la conversación de VOIP.

Destacar también que la tarjeta de sonido utilizada es una *Realtek High Definition Audio*.

4.1.3.2 Oyentes

Los participantes en las pruebas de escucha se escogieron al azar entre la población que utiliza normalmente el servicio telefónico, fijando las siguientes condiciones:

- a) que no hayan participado directamente en trabajos relacionados con la evaluación de la calidad de transmisión de los circuitos telefónicos o tareas afines, como codificación de la palabra.
- b) que no hayan participado en pruebas subjetivas, de cualquier naturaleza, al menos durante los seis meses precedentes, ni en pruebas de opinión sobre la escucha al menos durante un año.
- c) que no hayan oído antes la misma lista de frases.

4.2 Análisis QoE de los Códecs SILK y G.722

4.2.1 Cuestionario

Para la realización del análisis de QoE sobre los códecs SILK (Skype) y G.722 (Jitsi) ha sido imprescindible la elaboración de un cuestionario que tuvo que rellenar cada oyente para cada uno de los software. Es decir, cada oyente tuvo que escuchar dos grabaciones (una para cada software) y rellenar dos cuestionarios iguales pero cada uno para cada software. Dicho cuestionario se elaboró según las especificaciones de la recomendación

UIT-T P.82. Este cuestionario recoge la opinión de cada usuario según la experiencia que tuvo en la realización de las pruebas, y engloba valoraciones como:

- La valoración ACR de la calidad de la señal vocal;
- Dificultad para oír por el enlace;
- Esfuerzo necesario para comprender el significado de las frases;
- Grabaciones escuchadas con mayor claridad según sea masculina o femenina;
- Frecuencia con la que utiliza el sujeto internet y aplicaciones de VoIP.

Para poder diferenciar cada uno de los cuestionarios se utilizó un identificador único para cada uno de los cuestionarios, que será explicado en el siguiente apartado. En el anexo D se muestra el cuestionario completo.

4.2.2 Identificador

Este identificador ha sido utilizado tanto para identificar cada uno de los cuestionarios como para identificar cada uno de los archivos resultantes de capturar el tráfico de la conversación que son utilizados para el análisis de QoS, explicado más adelante. Con esto se consiguió tener en todo momento identificados todos los archivos de capturas de tráfico con sus correspondientes cuestionarios y así saber en todo momento qué cuestionario se correspondía con cada archivo de captura de tráfico.

El identificador sigue el siguiente formato:

AÑO_MES_DIA_TIPORED_NºENCUESTA_GRABACION_SOFTWARE_PUNTO

- ✓ **TIPO_RED** = {ETH todo cableado; WIFI extremos wifi; MIX un extremo eth y el otro wifi}
- ✓ **Nº ENCUESTA** = Numeración de las personas encuestadas
- ✓ **GRABACION** = 1H (primer hombre), 2H (segundo hombre), 1M (primera mujer), 2M (segunda mujer). Para el cuestionario se pusieron los cuatro seguidos según el orden en que se escuchó.
- ✓ **SOFTWARE** = SK (Skype) o JIT (Jitsi).
- ✓ **PUNTO** = TX origen de la llamada, RX destino de la llamada

Un ejemplo de utilización es

2011_05_18_ETH_5_1H2M2H1M_SK_RX

En este caso el cuestionario se realizó el 18/05/2011; sobre Ethernet: el oyente nº5; el orden de grabación fue 1H2M2H1M; se utilizó Skype y el extremo donde se captura este tráfico es el de destino de la llamada.

4.2.3 Descripción Teórica de los Cálculos

En este apartado se muestran fórmulas relevantes a la hora de realizar los cálculos.

Skweness o Sesgo

Esta función caracteriza el grado de asimetría de una distribución con respecto a su media. Si el sesgo es igual a cero, la distribución es simétrica; si el sesgo es positivo la distribución una tendrá una cola asimétrica extendida hacia los valores positivos. Un sesgo negativo indica una distribución con una cola asimétrica extendida hacia los valores negativos.

$$\alpha_3 = \frac{\sum(x_i - \bar{x})^3 n_i}{N S^3}, \text{ siendo } S \text{ la desviación estándar de la muestra}$$

Curtosis

La curtosis caracteriza la elevación o el achatamiento relativo de una distribución, comparada con la distribución normal. Una curtosis positiva indica una distribución relativamente elevada, mientras que una curtosis negativa indica una distribución relativamente plana.

$$\alpha_4 = \frac{\sum(x_i - \bar{x})^4 n_i}{N S^4}, \text{ siendo } S \text{ la desviación estándar de la muestra}$$

ANOVA (Análisis de varianza)

Es una técnica estadística diseñada para medir si existen diferencias entre los valores medios de una variable dependiente calculados para los distintos grupos que se pueden obtener con otra variable independiente y nominal.

En el caso de que la variable independiente tuviera solo dos alternativas, sería suficiente aplicar un test T de diferencia de medias.

La variable o variables independientes, reciben el nombre de Factor y deben ser variables de tipo nominal, y sus distintos valores el de tratamientos, mientras que la variable dependiente debe ser métrica, puesto que sobre ella se debe calcular los valores medios objeto del análisis de la varianza.

La hipótesis nula a contrastar es que se consideran iguales las medias en todos los grupos, o lo que es lo mismo, no existen diferencias entre las medias obtenidas para cada uno de los grupos formados por la variable independiente o factor.

Se rechaza la hipótesis nula con que al menos una de las medias sea significativamente diferente de las demás.

4.2.4 Resultados del Análisis de QoE

4.2.4.1 Perfil de los encuestados

En la figura 4.4 y 4.5 se muestran el porcentaje de encuestados de cada sexo y la frecuencia de uso de aplicaciones VoIP e internet de los mismos.

Del total de los encuestados el 76,67% son hombres y el 23,33% son mujeres. En una escala MOS (1-5), dichos encuestados determinan su frecuencia de uso en 2,60 para aplicaciones de VoIP y 4,31 para Internet como muestra la figura 4.5 (CI, Intervalo de Confianza).

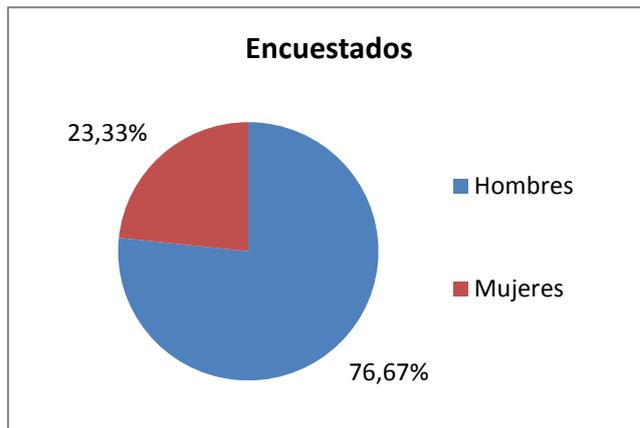


Figura 4.4: Porcentaje de sexos encuestados.

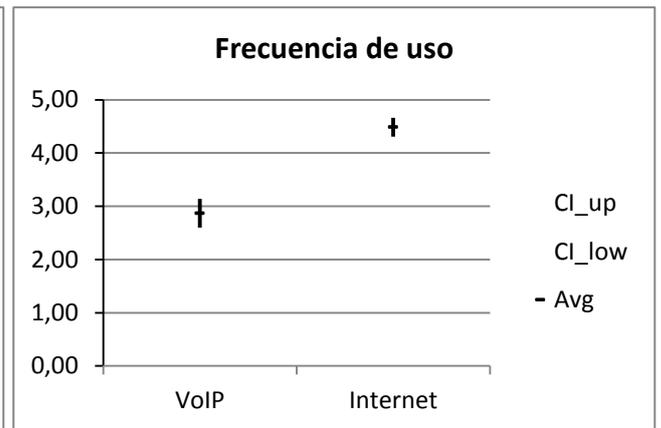


Figura 4.5: Frecuencia de uso de aplicaciones VoIP e Internet.

4.2.4.2 Dificultades para oír

En la siguiente figura (figura 4.6) se muestran los porcentajes de oyentes que han tenido alguna dificultad para oír las conversaciones al utilizar ambas aplicaciones.

En ambas aplicaciones el porcentaje de oyentes que han tenido alguna dificultad para oír es bastante inferior que el que no la han tenido, pero es en el caso de Jitsi en la que la diferencia es mayor, ya que en Skype han tenido dificultad para oír el 32% frente al 68% que no la han tenido, y con Jitsi la han tenido el 22% frente al 78% que no la han tenido.

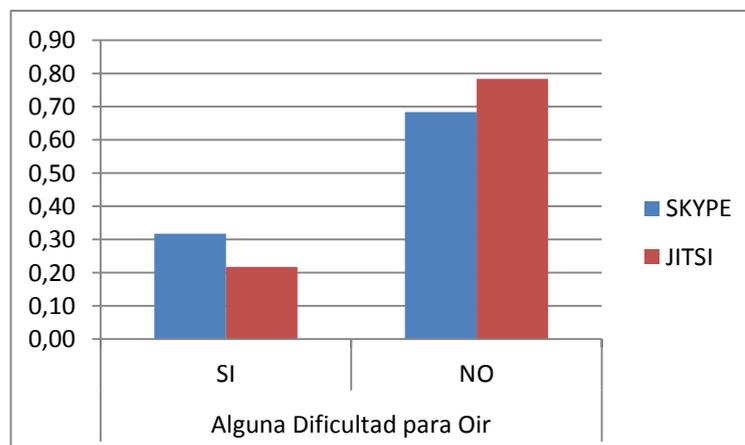


Figura 4.6: Porcentajes de dificultades para oír.

A continuación, como se puede observar en la figura 4.7, se muestra una comparativa de los distintos tipos de dificultad que han tenido los oyentes al realizar las pruebas con cada una de las aplicaciones.

- **Skype:** El ruido o zumbido es la dificultad más dada con un 52,63% de los encuestados, seguida de la distorsión y la diafonía (con valores muy parecidos) y siendo las dificultades de poco volumen y variaciones de nivel e interrupciones de las menos numerosas con un 5,26%. En este caso no se han dado casos de eco ni diafonía.
- **Jitsi:** En este caso no hay una dificultad tan diferenciada de las demás sino que son dos (el ruido o zumbido, y variaciones de nivel e interrupciones) las más dadas y con un porcentaje menor que la dificultad más dada en Skype, en este caso con un 38,46% cada una. No se han dado casos de diafonía ni poco volumen.

La única dificultad que se ha dado en un porcentaje similar en ambas aplicaciones la distorsión, con un 21,05% en Skype y un 15,38% en Jitsi.

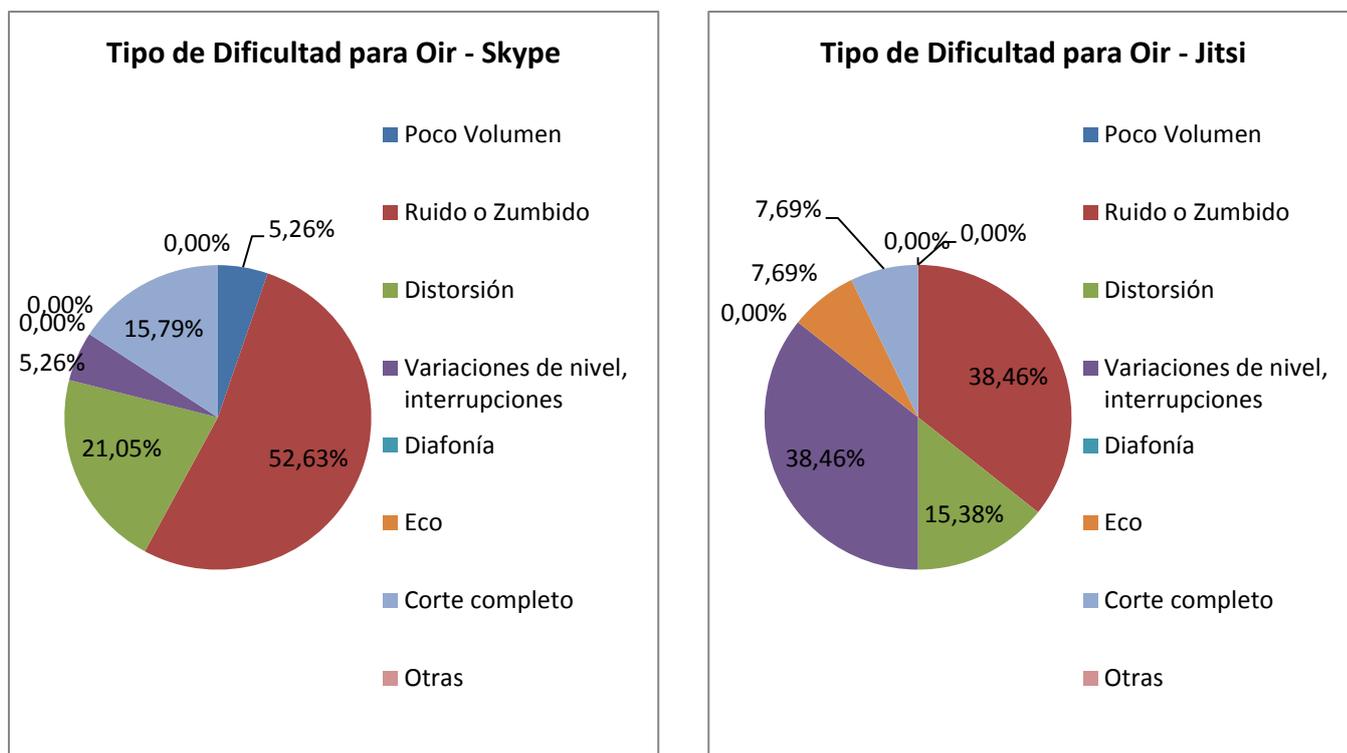


Figura 4.7: Porcentajes de los tipo de dificultad

4.2.4.3 Calidad y esfuerzo de comprensión

En la figura 4.8 se muestran los porcentajes de oyentes que escucharon las conversaciones con mayor claridad según el sexo de la voz que escuchaban.

En ambas aplicaciones no se produce una distinción del sexo a la hora de calificar la calidad de la conversación, ya que para la gran mayoría es similar la calidad de la conversación en la que habla un hombre que en la que habla una mujer.

- **Skype:** El 10% ha calificado las conversaciones con voz de hombre de mayor calidad, frente al 20% que ha calificado las conversaciones con voz de mujeres de mayor calidad. El resto (70%) no ha encontrado diferencia de calidad entre las conversaciones con voces de distinto sexo.
- **Jitsi:** El 7% ha calificado las conversaciones con voz de hombre de mayor calidad, frente al 18% que ha calificado las conversaciones con voz de mujeres de mayor calidad. El resto (75%) no ha encontrado diferencia de calidad entre las conversaciones con voces de distinto sexo.

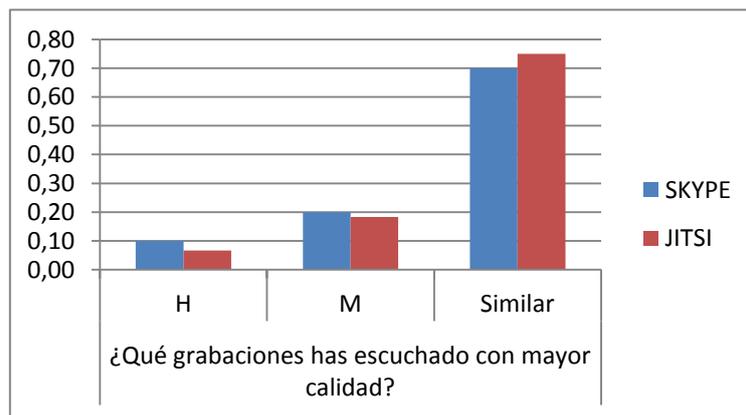


Figura 4.8: Grabaciones escuchadas con mayor calidad

En la figura 4.9 y 4.10 se muestran la calidad de la señal vocal y el esfuerzo de comprensión respectivamente.

Se puede observar que el esfuerzo de comprensión de ambas aplicaciones es muy similar y que la calidad de señal vocal de Jitsi es ligeramente superior que la de Skype, con una media (utilizando valoración MOS) de 4,48 y 3,92 respectivamente.

	Calidad de la señal vocal			Esfuerzo necesario para comprender el significado de las frases (1-5)		
	Cl_up	Cl_low	Avg	Cl_up	Cl_low	Avg
SKYPE	4,13	3,70	3,92	4,55	4,21	4,38
JITSI	4,69	4,28	4,48	4,77	4,46	4,62

Figura 4.9: Calidad de la señal vocal y esfuerzo de comprensión

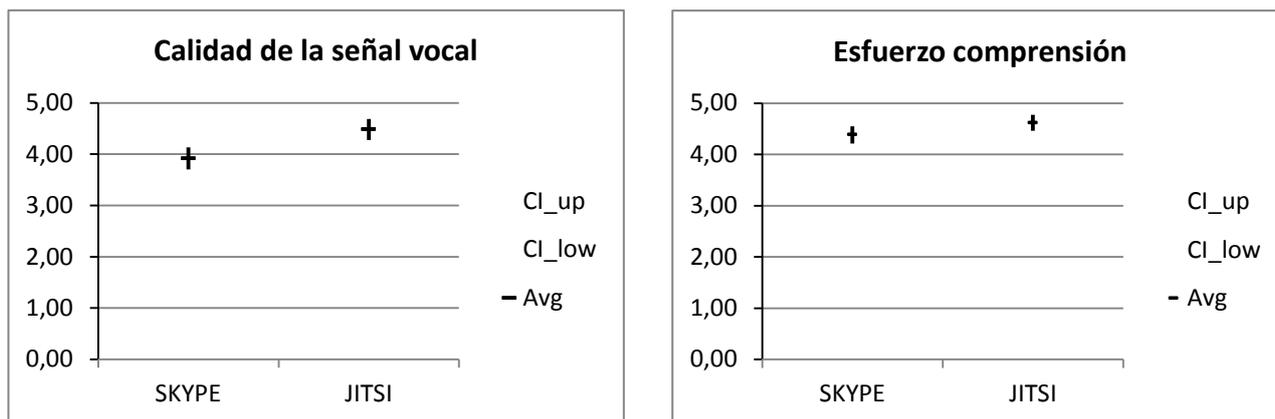


Figura 4.10: Representación de la calidad de la señal vocal y esfuerzo de comprensión.

4.2.4.4 Skewness y Kurtosis

La figura 4.11 muestra la tendencia que siguen (respecto a la media) la valoración, el uso de aplicaciones de VoIP y el uso de internet, haciendo uso del cálculo estadístico skewness.

En lo único en lo que se aprecia una clara diferencia entre ambas aplicaciones es en la distribución de la valoración, ya que aunque en ambos casos el skewness es negativo y por tanto se produce una tendencia hacia valores mayores que la media correspondiente, en el caso de Jitsi el skewness es más negativo que el de Skype.

Para los parámetros restantes se produce una distribución igual en ambas aplicaciones. En el uso de internet se produce un sesgo positivo pero muy cercano al cero, por lo que se produce una ligera tendencia hacia valores menores que la media correspondiente. En el uso de internet se produce un sesgo negativo, por lo que se produce una tendencia hacia valores mayores que la media correspondiente.

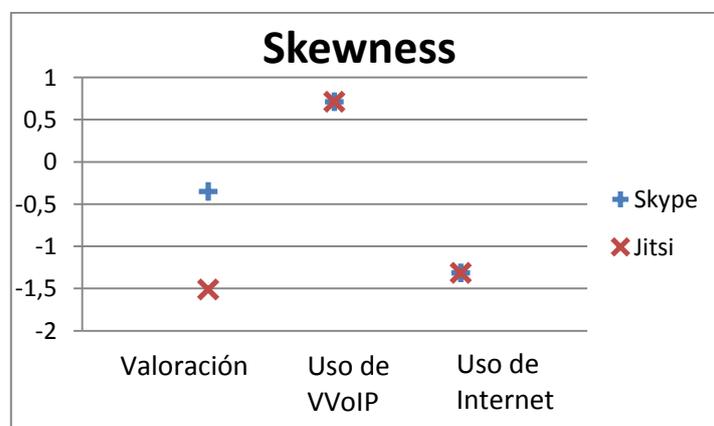


Figura 4.11: Skewness

La figura 4.12 muestra la elevación o el achatamiento relativo de la distribución, respecto con la distribución normal, de la valoración, el uso de aplicaciones de VoIP y el uso de internet, haciendo uso del cálculo estadístico kurtosis.

En lo único en lo que se aprecia una clara diferencia entre ambas aplicaciones es en la distribución de la valoración, ya que en el caso de Skype la kurtosis tiene un valor negativo y cercano al cero, por lo que la distribución de los datos es levemente mas chata que la distribución normal. En el caso de Jitsi la kurtosis tiene un valor positivo, por lo que la distribución es más elevada (pronunciada) respecto a la distribución normal.

Para los parámetros restantes de produce una distribución igual en ambas aplicaciones. En el uso de internet se tiene una kurtosis negativa y cercana al cero, por lo que la distribución de los datos es levemente más chata que la distribución normal. En el uso de Internet se tiene una kurtosis positiva, por lo que la distribución es más elevada (pronunciada) respecto a la distribución normal.

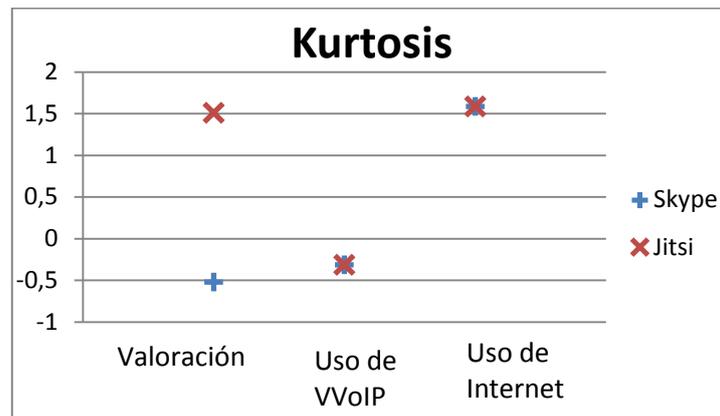


Figura 4.12: Kurtosis.

4.2.4.5 ANOVA One-Way

En la figura 4.13 se muestran los resultados del análisis de varianza de un solo factor (ANOVA One-Way) para cada uno de los factores de este estudio, que nos da la herramienta para distinguir si un factor afecta la respuesta en promedio.

Hay que tener en cuenta que:

- Si el valor p es menor o igual a un nivel predeterminado de significancia (nivel α , en este caso $\alpha=0,05$), se rechaza la hipótesis nula y da crédito a la alternativa.
- Si el valor p es mayor que el nivel α , no se rechaza la hipótesis nula y no se puede dar crédito a la hipótesis alterna. No hay una diferencia estadísticamente significativa entre las medias de la población.

Aplicando lo anterior a este proyecto se obtiene que en todos los casos que se muestran en la figura 4.13 el valor de P es mayor de 0,05 (α) y por tanto se acepta la hipótesis nula (H_0) y se puede concluir entonces que:

- El género, el uso de Internet y el uso de aplicaciones VoIP no afectan a la valoración global que se ha hecho de cada aplicación de VoIP.
- El género y el uso de Internet no afectan al uso de aplicaciones de VoIP.
- El género no afecta al uso de Internet.

En el último caso ocurre lo contrario, es decir, que el valor de P es menor de 0,05 (α) y por tanto se rechaza la hipótesis nula (H_0) y se puede concluir entonces que el uso de aplicaciones de VoIP sí afecta al uso de Internet.

	Skype		Jitsi	
	F	P	F	P
VALORACION vs. GENERO	1,93	0,170	0,21	0,647
VALORACION vs. USO INTERNET	0,75	0,528	2,34	0,083
VALORACION vs. USO VoIP	1,04	0,396	0,41	0,799
USO VoIP vs. GENERO	1,41	0,239	1,41	0,239
USO VoIP vs. USO INTERNET	2,06	0,116	2,06	0,116
USO INTERNET vs. GENERO	2,02	0,161	2,02	0,16
USO INTERNET vs. USO VoIP	3,23	0,019	3,23	0,019

Figura 4.13: ANOVA One-Way

4.3 Análisis QoS de los Códecs SILK y G.722

Para esta parte del proyecto ha sido imprescindible la utilización de la herramienta de captura y análisis de paquetes de red llamada Wireshark. Con dicha aplicación se ha realizado la captura de los paquetes que se intercambiaron en cada una de las llamadas de VoIP para cada uno de los participantes en las pruebas. Como cada oyente tuvo que escuchar una grabación para cada uno de los software (Skype y Jitsi) y la captura de paquetes se realizó tanto en el origen como en el destino de la llamada, por cada oyente se obtuvo un total de cuatro ficheros de captura de paquetes, los cuales quedaron totalmente identificados y relacionados con sus respectivos cuestionarios gracias al identificador, el cual ha sido explicado en el apartado 4.2.2 de esta memoria.

Para este análisis se tendrán en cuenta los siguientes cálculos:

- Ancho Banda Total (Bytes/segundo).
- Media de los tamaños de los paquetes (Bytes).
- Desviación estándar de los tamaños de los paquetes (Bytes).
- Intervalo de confianza (95%) de los tamaños de los paquetes
- Media de las diferencias de tiempos entre la recepción de dos paquetes consecutivos (segundos).
- Desviación estándar de las diferencias de tiempos entre la recepción de dos paquetes consecutivos (segundos).
- Intervalo de confianza (95%) de las diferencias de tiempos entre la recepción de dos paquetes consecutivos
- Paquetes perdidos
- Ancho banda NO VoIP (Bytes/segundo).

Todos estos cálculos excepto el último se calcularon filtrando los archivos que contenían los paquetes capturados dejando solamente los paquetes referentes a la conversación de VoIP. El último parámetro se calculó filtrando dejando solamente el tráfico que no era referente a la conversación de VoIP, es decir el tráfico causado por otros servicios o aplicaciones ajenos a la llamada de VoIP.

En esta parte del proyecto es imprescindible destacar que cada fichero resultante de la captura de tráfico (cuyas extensiones eran .pcap) tuvo que ser pasado a formato texto (.txt) para poder ser entendido por el script que se programó en awk. Dicho script será explicado detalladamente en el siguiente apartado (4.3.1). A continuación se muestra la figura 4.14 en la que se observa una capturas de pantalla de cómo queda un fichero con wireshark y una vez pasado a formato texto.

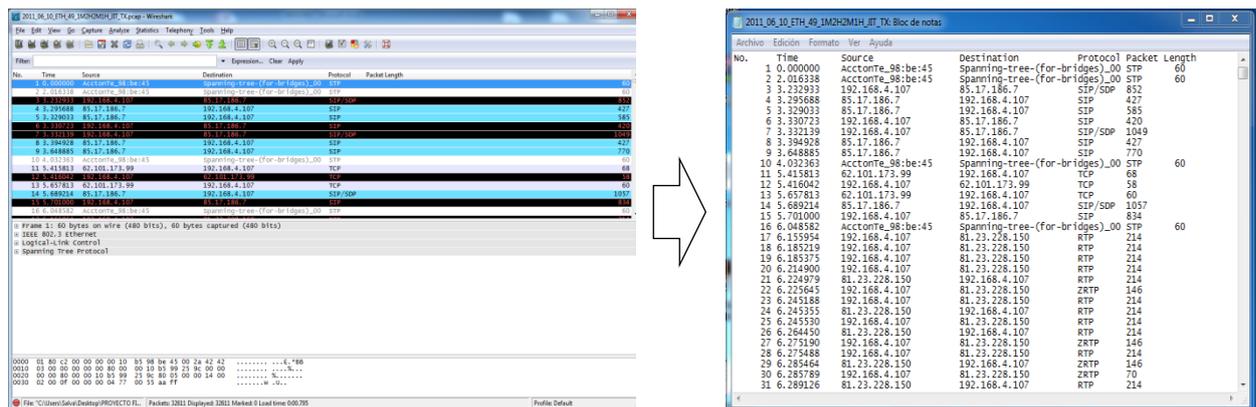


Figura 4.14: Capturas de pantalla de un fichero .pcap y un fichero .txt

4.3.1 Descripción Teórica de los Cálculos

En este apartado se muestran fórmulas relevantes a la hora de realizar los cálculos.

Media muestral

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

Varianza muestral

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{N}{N-1} [\overline{X^2} - \bar{X}^2] \quad (2)$$

Siendo:

$$\overline{X^2} = \frac{1}{N} \sum_{i=1}^N x_i^2 \quad (3)$$

Desviación Típica (Estándar)

$$S = \sqrt{S^2} \quad (4)$$

Intervalo de Confianza (con σ^2 desconocida)(población desconocida)

$$(\bar{x} \pm t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}) \quad (5)$$

La demostración de cómo se obtiene este intervalo de confianza, así como la tabla t-student (necesaria para el cálculo de este intervalo de confianza) se encuentra en el anexo E.

4.3.2 Scripts AWK

En concreto se han utilizado tres scripts diferentes, uno para las capturas obtenidas en el extremo transmisor, otro para las capturas obtenidas en el extremo receptor, cuya única diferencia con respecto al utilizado en el extremo transmisor es que en vez de calcular el número de paquetes enviados se calcula el número de paquetes recibidos (con el fin de obtener los posteriormente los paquetes perdidos), y un tercero utilizado para calcular solamente el ancho de banda total (el cual se utilizó para calcular el ancho de banda total utilizado por servicios o aplicaciones que no tenían nada que ver con la conversación de VoIP).

4.3.2.1 Script Utilizado en Extremo Transmisor

A continuación se muestra al código completo del script utilizado para realizar los cálculos utilizando los ficheros de captura de tráfico obtenidos en el extremo transmisor. Cabe recordar que para utilizar este script, el fichero (ya pasado a .txt) que fue utilizado fue filtrado para que solo apareciera el tráfico referente a la conversación VoIP.

```

BEGIN{
    comp_tiempo=10;

    pasada_filal=0;

    t=1.96; #como el número de muestras siempre va a ser mayor de 120 y como
el intervalo de confianza siempre es del 95%, el valor siempre será el mismo.

    ind=-1;
}

{

    if(pasada_filal==1){ #Para no tener en cuenta la primera fila, que
contiene los títulos de los campos.

        #####
        ###MEDIA DIFERENCIA TIEMPO ENTRE DOS PAQUETES CONSECUTIVOS###
        #####
        dif_tiempo=$2-tiempo_anterior;

        tiempo_anterior=$2;

        tiempo_total=tiempo_total+dif_tiempo;

        #####
        ###TAMAÑO MEDIO PAQUETES###
        #####
        suma_tam_paquetes=suma_tam_paquetes+$6;

        #####
        ###ANCHO DE BANDA CADA 10 SEGUNDOS###
        #####
        if($2>comp_tiempo){ #comp_tiempo irá cambiando en intervalos de 10
segundos para obtener el valor del ancho de banda cada 10 segundos

            ind=ind+1;
            comp_tiempo=comp_tiempo+10;
            ancho_banda_parcial=(suma_tam_paquetes-suma_tam_paquetes_ant)/($2-
tiempo_ant);

            valores_ancho_banda_parcial[ind]=ancho_banda_parcial;

            suma_tam_paquetes_ant=suma_tam_paquetes;
            tiempo_ant=$2;

        }

        #####
        ###ESTADISTICOS###
        #####
        suma_cuadrados_dif_tiempo=suma_cuadrados_dif_tiempo+(dif_tiempo^2);
        suma_cuadrados_tam_paquetes=suma_cuadrados_tam_paquetes+($6^2);

        #####
        ###PAQUETES ENVIADOS###
        #####

        if($4=="85.17.186.6"||$4=="81.23.228.129"||$4=="81.23.228.150"||$4=="212.1
28.44.132"){ #Todo lo enviado a estas direcciones SIEMPRE va al extremo RX, ya

```

que las tres primeras direcciones son los "retransmisores" entre tx y rx cuando utilizamos jitsi, y la última es la ip pública del laboratorio de rx, utilizada cuando usamos skype.

```
        paq_enviados=paq_enviados+1;
    }

    numero_paquetes=numero_paquetes+1;
}
pasada_filal=1;
}

END{
    media_dif_tiempo=tiempo_total/(numero_paquetes-1); #-1 porque el número de
diferencias de tiempos entre dos paquetes consecutivos será n-1, siendo n el
número de paquetes.

    tam_medio_paquetes=suma_tam_paquetes/(numero_paquetes);

    ancho_banda_total=suma_tam_paquetes/$2;

    media_suma_cuadrados_tam_paquetes=suma_cuadrados_tam_paquetes/numero_paque
tes;
    desv_estandar_tam_paquetes=sqrt((numero_paquetes)*(media_suma_cuadrados_ta
m_paquetes-tam_medio_paquetes^2)/(numero_paquetes-1));

    desvio_tam_paquetes=t*desv_estandar_tam_paquetes/sqrt(numero_paquetes);
    ext_inferior_inter_confianza_tam_paquetes=tam_medio_paquetes-
desvio_tam_paquetes;

    ext_superior_inter_confianza_tam_paquetes=tam_medio_paquetes+desvio_tam_pa
quetes;

    media_suma_cuadrados_dif_tiempo=suma_cuadrados_dif_tiempo/(numero_paquetes
-1);
    desv_estandar_dif_tiempos=sqrt((numero_paquetes-
1)*(media_suma_cuadrados_dif_tiempo-media_dif_tiempo^2)/(numero_paquetes-2)); #-1
porque el número de muestras es una menos que el número de paquetes, y -1 porque
la fórmula es el número de muestras menos uno.

    desvio_dif_tiempos=t*desv_estandar_dif_tiempos/sqrt(numero_paquetes-1);

    ext_inferior_inter_confianza_dif_tiempos=media_dif_tiempo-
desvio_dif_tiempos;

    ext_superior_inter_confianza_dif_tiempos=media_dif_tiempo+desvio_dif_tiempos;

    printf("\n%s;%s;%s;%s;[%s,%s];%s;%s;[%s,%s]; ;%s; ;
";,ARGV[1],ancho_banda_total,tam_medio_paquetes,desv_estandar_tam_paquetes,ext_in
ferior_inter_confianza_tam_paquetes,ext_superior_inter_confianza_tam_paquetes,med
ia_dif_tiempo,desv_estandar_dif_tiempos,ext_inferior_inter_confianza_dif_tiempos,
ext_superior_inter_confianza_dif_tiempos,paq_enviados);
    for(j=0;j<=ind;j++){
        printf("%s;",valores_ancho_banda_parcial[j]);
    }
}
```

Con este script se consiguió obtener los cálculos necesarios para el análisis de QoS utilizando las fórmulas descritas en el apartado anterior (4.3.1).

Se ha decidido utilizar, para una mejor comprensión del código, variables con nombres significativos con el dato que contienen, aunque esto conllevara a una mayor longitud del nombre de cada variable.

Para entender cómo funciona hay que tener en cuenta que en AWK se pasa este script secuencialmente a cada una de las filas del fichero ascii, y para referirse a los datos de cada campo del fichero se utiliza \$+número de columna, en nuestro caso:

- **\$2:** Para referirse al segundo campo (columna 2). Tiempo en segundos en el que se recibe el paquete.
- **\$4:** Para referirse al cuarto campo (columna 4). Dirección IP de Destino del paquete.
- **\$6:** Para referirse al sexto campo (columna 6). Tamaño en bytes del paquetes

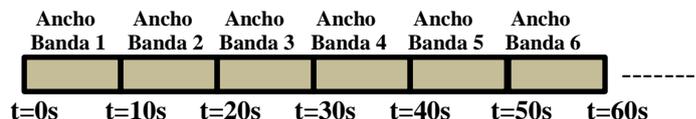
El patrón BEGIN solamente se cumple cuando se arranca el programa de awk, es decir justo antes de procesar la primera fila. Por ello se utilizó para inicializar variables. En este bloque cabe destacar que la variable t es la utilizada para calcular el intervalo de confianza y se corresponde con el dato $t_{n-1, 1-\alpha/2}$ (ver apartado 4.3.1). Dicho valor se debe averiguar mirando la tabla t-student (Anexo E), como a partir de 121 muestras siempre es el mismo, el valor de dicho dato siempre será 1,96. A continuación se muestra cómo se obtuvo el valor de este dato para el intervalo de confianza del 95% (para un mejor entendimiento completar la siguiente información con la que aparece en el primer apartado del Anexo E).

$$0.95 = 1 - \alpha \quad \Rightarrow \quad \alpha = 0.05$$

$$1 - \alpha/2 = 0.975$$

Por tanto el valor que hubo que mirar en la tabla *t-student* fue el correspondientes a $t_{\infty, 0.975}$ (∞ porque el número de muestras es superior a 120).

En la parte central (que es ejecutada para cada una de las líneas) está todo bien comentado en el código. Comentar que el bloque llamado “ANCHO DE BANDA CADA 10 SEGUNDOS” el ancho de banda parcial de cada fichero está calculado de forma NO acumulativa, obteniendo así los valores del ancho de banda utilizado en diferentes momentos (cada 10 segundos) de la transmisión.



Los resultados al calcular el ancho de banda, que se fueron obteniendo cada 10 segundos fueron almacenados en un array del cual se obtuvieron todos sus valores más adelante.

También aclarar que en el bloque “PAQUETES ENVIADOS” cuando se habla de “retransmisores” se refiere a los nodos intermedios que actúa de intermediario entre ambos extremos de la conversación (en el caso de la utilización de Jitsi). Las direcciones de los

retransmisores utilizados en todas las pruebas son siempre alguno de los tres que se especifican en el código.

El patrón END sólo se cumple cuando se han procesado todas y cada una de las filas del fichero de texto, y en este proyecto es utilizado para realizar los cálculos pertinentes (siguiendo las fórmulas del apartado 4.3.1) e imprimir los resultados. Los resultados de los cálculos se imprimieron en una sola línea y separados con “;”, de tal manera que al copiarlos en la plantilla que se hizo en Excel, dicho programa consiguiera entenderlo de forma que cada dato separado por “;” quedase en una columna.

La ejecución de los scripts en awk se realizó en Linux (Ubuntu) y todos los resultados se enviaron a un fichero de texto. Para conseguir esto se ejecutó por línea de comandos (para cada uno de los ficheros de texto que contenían los paquetes de tráfico capturado) el siguiente comando:

```
awk -f programa captura_trafico.txt >> resultado.txt
```

Se obtuvieron unos ficheros de texto con los resultados de los cálculos y fueron utilizados para rellenar la plantilla de Excel que se diseñó para tal fin. Un pequeño fragmento de uno de estos ficheros de texto resultante se puede observar en la figura 4.15.

```
2011_05_25_ETH_2_1H2H1M2M_JIT_RX.txt;21907.2;217.433;21.6841;[21  
7.199,217.667];0.00992549;0.00807379;[0.0098383,0.0100127];
```

Figura 41.5: Fragmento ejemplo del fichero en formato texto.

4.3.2.2 Script Utilizado en Extremo Receptor

A continuación se muestra al código completo del script utilizado para realizar los cálculos utilizando los ficheros de captura de tráfico obtenidos en el extremo receptor. Cabe recordar que para utilizar este script, el fichero (ya pasado a .txt) que fue utilizado fue filtrado para que solo apareciera el tráfico referente a la conversación VoIP.

```
BEGIN{  
    comp_tiempo=10;  
  
    pasada_filal=0;  
  
    t=1.96;#como el número de muestras siempre va a ser mayor de 120 y como el  
intervalo de confianza siempre es del 95%, el valor siempre será el mismo. Mirar  
tabla t-student  
  
    ind=-1;  
  
}  
  
{  
  
    if(pasada_filal==1){ #Para no tener en cuenta la primera fila (contiene  
los títulos de los campos)
```

```

#####
###MEDIA DIFERENCIA TIEMPO ENTRE DOS PAQUETES CONSECUTIVOS###
#####
dif_tiempo=$2-tiempo_anterior;

tiempo_anterior=$2;

tiempo_total=tiempo_total+dif_tiempo;

#####
###TAMAÑO MEDIO PAQUETES###
#####
suma_tam_paquetes=suma_tam_paquetes+$6;

#####
###ANCHO DE BANDA CADA 10 SEGUNDOS###
#####
if($2>comp_tiempo){ #comp_tiempo irÃ; cambiando en intervalos de 10
segundos para obtener el valor del ancho de banda cada 10 segundos

        ind=ind+1;
        comp_tiempo=comp_tiempo+10;
        ancho_banda_parcial=(suma_tam_paquetes-suma_tam_paquetes_ant)/($2-
tiempo_ant);

        valores_ancho_banda_parcial[ind]=ancho_banda_parcial;

        suma_tam_paquetes_ant=suma_tam_paquetes;
        tiempo_ant=$2;

    }

#####
###ESTADISTICOS###
#####
suma_cuadrados_dif_tiempo=suma_cuadrados_dif_tiempo+(dif_tiempo^2);
suma_cuadrados_tam_paquetes=suma_cuadrados_tam_paquetes+($6^2);

#####
###PAQUETES RECIBIDOS###
#####

    if($3=="85.17.186.6"||$3=="81.23.228.129"||$3=="81.23.228.150"||$3=="212.1
28.44.50"){ #Todo lo recibido de estas direcciones SIEMPRE procede del extremo
TX, ya que las tres primeras son los "retransmisores" entre tx y rx cuando
utilizamos jitsi, y la última es la ip publica del laboratorio de tx, utilizada
cuando usamos skype.
        paq_recibidos=paq_recibidos+1;
    }

    numero_paquetes=numero_paquetes+1;

}
pasada_filal=1;
}

END{
    media_dif_tiempo=tiempo_total/(numero_paquetes-1); #-1 porque el número de
fierencias de tiempos entre dos paquetes consecutivos será n-1, siendo n el
número de paquetes.

```

```
tam_medio_paquetes=suma_tam_paquetes/(numero_paquetes);
ancho_banda_total=suma_tam_paquetes/$2;
media_suma_cuadrados_tam_paquetes=suma_cuadrados_tam_paquetes/numero_paquetes;

desv_estandar_tam_paquetes=sqrt((numero_paquetes)*(media_suma_cuadrados_tam_paquetes-tam_medio_paquetes^2)/(numero_paquetes-1));

desvio_tam_paquetes=t*desv_estandar_tam_paquetes/sqrt(numero_paquetes);

ext_inferior_inter_confianza_tam_paquetes=tam_medio_paquetes-desvio_tam_paquetes;

ext_superior_inter_confianza_tam_paquetes=tam_medio_paquetes+desvio_tam_paquetes;

media_suma_cuadrados_dif_tiempo=suma_cuadrados_dif_tiempo/(numero_paquetes-1);

desv_estandar_dif_tiempos=sqrt((numero_paquetes-1)*(media_suma_cuadrados_dif_tiempo-media_dif_tiempo^2)/(numero_paquetes-2)); #-1 porque el número de muestras es una menos que el número de paquetes, y -1 porque la fórmula es el número de muestras menos uno.

desvio_dif_tiempos=t*desv_estandar_dif_tiempos/sqrt(numero_paquetes-1);

ext_inferior_inter_confianza_dif_tiempos=media_dif_tiempo-desvio_dif_tiempos;

ext_superior_inter_confianza_dif_tiempos=media_dif_tiempo+desvio_dif_tiempos;

printf("\n%s;%s;%s;%s;[%s,%s];%s;%s;[%s,%s]; ;%s; ;",ARGV[1],ancho_banda_total,tam_medio_paquetes,desv_estandar_tam_paquetes,ext_inferior_inter_confianza_tam_paquetes,ext_superior_inter_confianza_tam_paquetes,media_dif_tiempo,desv_estandar_dif_tiempos,ext_inferior_inter_confianza_dif_tiempos,ext_superior_inter_confianza_dif_tiempos,paq_recibidos);
for(j=0;j<=ind;j++){
    printf("%s;",valores_ancho_banda_parcial[j]);
}
}
```

Este script es exactamente igual que el anterior pero con la única diferencia de que en este caso no se calcula el número de paquetes enviados, sino el número de paquetes recibidos. De este modo al introducir los datos en la tabla Excel se pudo calcular el número de paquetes perdidos restando a los paquetes que se enviaron en el extremo tx los recibidos en el extremo rx. Para identificar cada fichero de captura de tráfico y saber así que ficheros pertenecían a una misma conversación se utilizó el identificador explicado en el apartado 4.2.2.

4.3.2.3 Script para el Cálculo del Ancho de Banda del Tráfico no VOIP

A continuación se muestra al código completo del script utilizado para realizar el cálculo del ancho de banda total utilizado por otros servicios y aplicaciones que no tenían nada que ver con la conversación de VoIP, utilizando los ficheros de captura de tráfico obtenidos tanto en el extremo receptor como en el emisor. Cabe recordar que para utilizar este script, el fichero (ya pasado a .txt) que fue utilizado fue filtrado para que sólo apareciera el tráfico que no tenía nada que ver con la conversación VoIP.

```
{
    if(pasada_filal==1){ #Para no tener en cuenta la primera fila
    (contiene los títulos de los campos)

        #####
        ###TAMAÑO MEDIO PAQUETES###
        #####
        suma_tam_paquetes=suma_tam_paquetes+$6;

    }

    numero_paquetes=numero_paquetes+1;
}
pasada_filal=1;
}

END{

    ancho_banda_total=suma_tam_paquetes/$2;

    printf("\n%s;%s;",ARGV[1],ancho_banda_total);

}
```

4.3.3 Resultados del Análisis de QoS

4.3.3.1 Ancho de Banda

Podemos establecer una relación directamente proporcional entre el ancho de banda y la calidad de los datos transmitidos. No obstante, necesitamos aplicaciones que no consuman un ancho de banda excesivo de lo contrario estarían requiriendo un uso exclusivo del canal para su transferencia, pero necesitamos que sí tengan el necesario para su transmisión con éxito.

En las siguientes gráficas (figuras 4.16 y 4.17) podemos ver el ancho de banda en origen y en destino, para cada una de las dos aplicaciones analizadas.

Tanto en transmisión como en recepción la aplicación que más ancho de banda utiliza en todas y cada una de las capturas es Jitsi. Cada aplicación utiliza aproximadamente el mismo ancho de banda medio en recepción y en transmisión, variando en Skype entorno a los 12500 bytes por segundo (Bps) y en Jitsi entorno a los 21000 bytes por segundo (Bps).

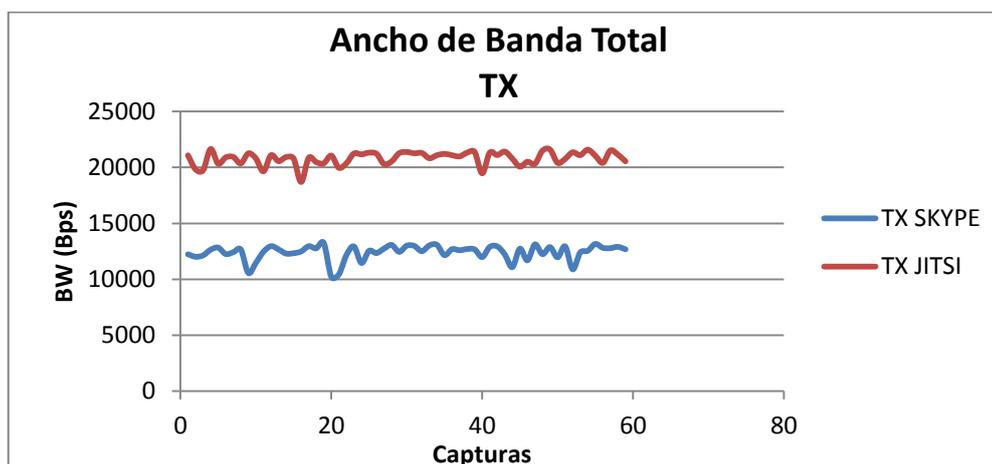


Figura 4.16: Ancho de banda total en TX.

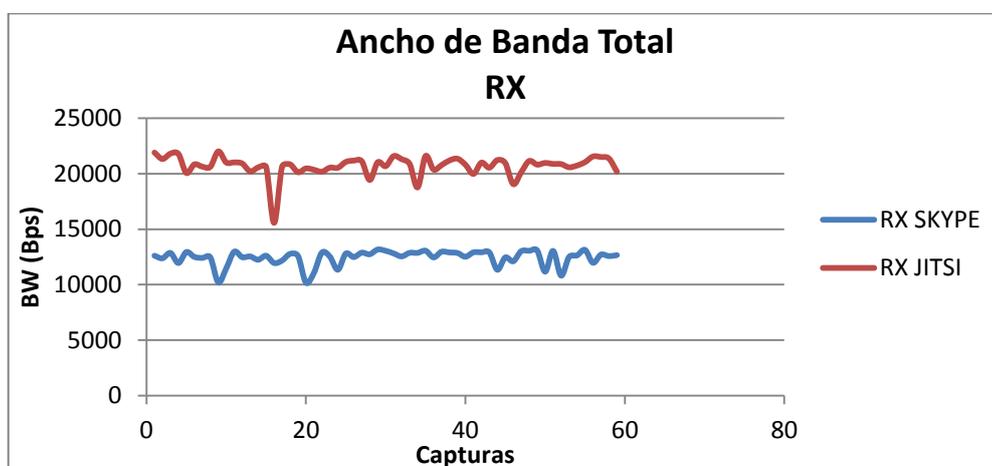


Figura 4.17: Ancho de banda total en RX.

Para una mejor comprensión del ancho de banda utilizado por cada aplicación se aportan las siguientes gráficas (figuras 4.18, 4.19, 4.20 y 4.21) en las que se muestran el ancho de banda que se ha ido utilizando a lo largo de la realización de cada una de las conversaciones de VoIP, en cada uno de los extremos de dicha conversación y para cada uno de los programas. Dichos valores del ancho de banda han sido calculados en periodos de tiempo de 10 segundos desde que comenzó la conversación hasta el fin de la misma. Para evitar obtener demasiadas gráficas que dificulten la comprensión de los datos se han seleccionado 5 conversaciones al azar de entre las 60.

Como ocurría con el ancho de banda total, en este caso cada aplicación utiliza aproximadamente el mismo ancho de banda tanto en recepción como en transmisión.

- **Skype:** El menor ancho de banda utilizado fue al inicio de las conversaciones y éste fue aumentando en los primeros 40 segundos, a partir de los cuales se estabilizó entorno a los 13000 Bps.
 - Transmisión: Los valores del ancho de banda utilizado al inicio de las conversaciones es diferente para cada una de ellas, tomando valores desde los 35 Bps a los 4889 Bps.
 - Recepción: Los valores del ancho de banda utilizado al inicio de las conversaciones es diferente para cada una de ellas, tomando valores desde los 100 Bps a los 12360 Bps.
- **Jitsi:** El menor ancho de banda utilizado fue al inicio de las conversaciones y éste fue aumentando en los primeros 25 segundos, a partir de los cuales se estabilizó entorno a los 21800 Bps.
 - Transmisión: Los valores del ancho de banda utilizado al inicio de las conversaciones es diferente para cada una de ellas, tomando valores desde los 61 Bps a los 12086 Bps.
 - Recepción: Los valores del ancho de banda utilizado al inicio de las conversaciones es diferente para cada una de ellas, tomando valores desde los 140 Bps a los 21536 Bps.

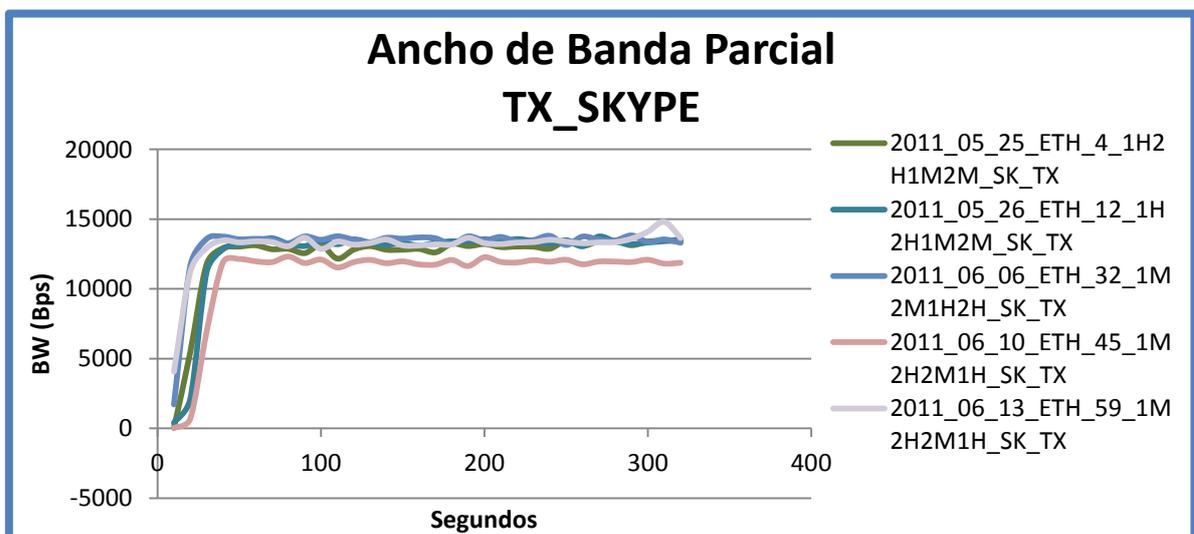


Figura 4.18: Ancho de banda parcial en TX_SKYPE.

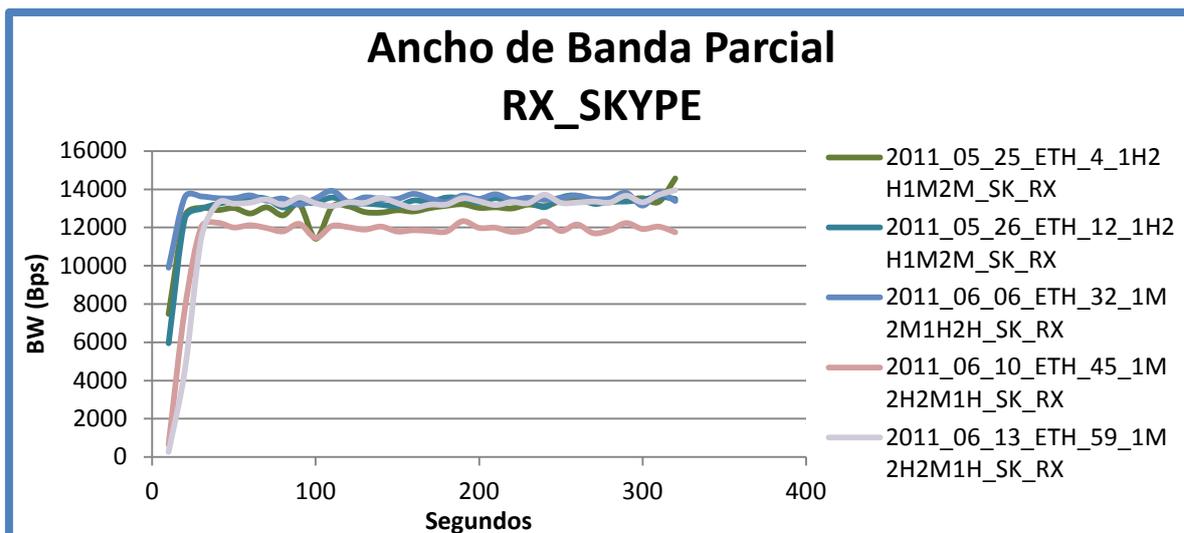


Figura 4.19: Ancho de banda parcial en RX_SKYPE.

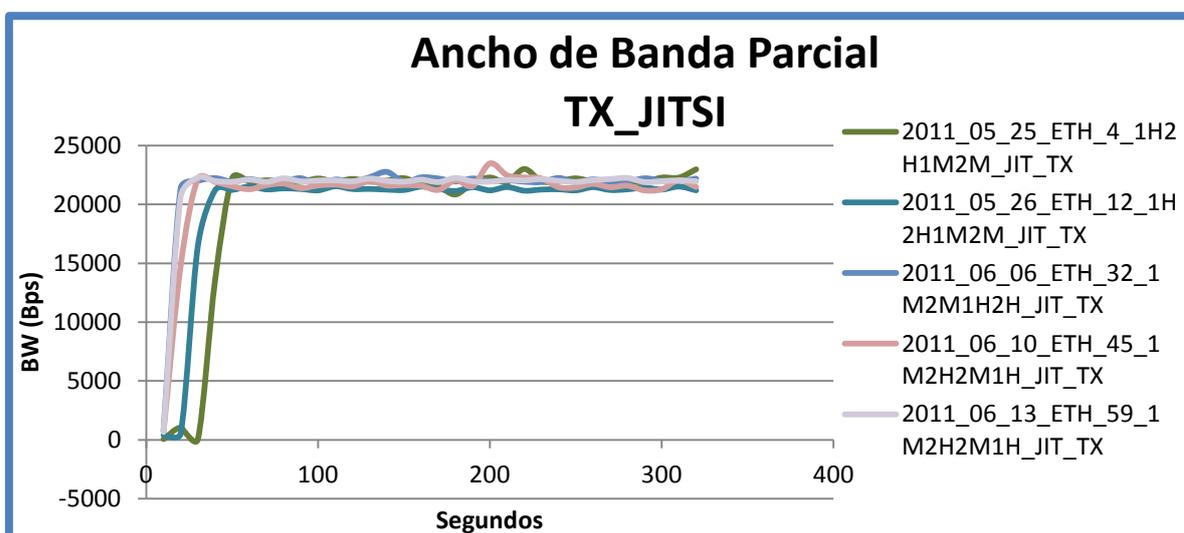


Figura 4.20: Ancho de banda parcial en TX_JITSI.

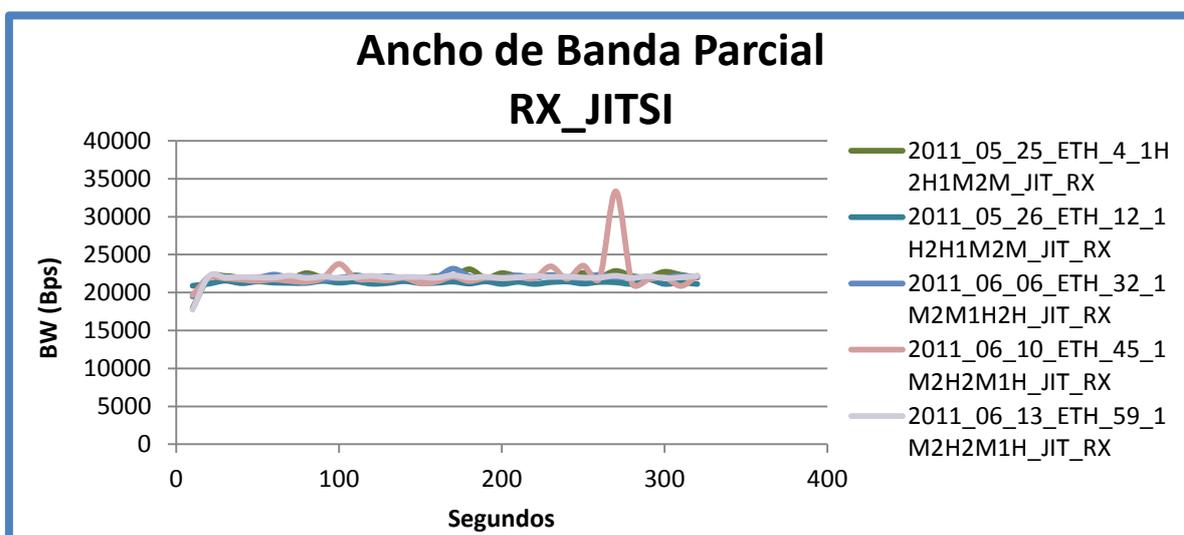


Figura 4.21: Ancho de banda parcial en RX_JITSI.

4.3.3.2 Tamaño Medio de los Paquetes

En las siguientes gráficas (figuras 4.22 y 4.23) podemos ver la comparación del tamaño medio de los paquetes en origen y en destino, de cada una de las aplicaciones analizadas.

La aplicación que utiliza paquetes con mayor tamaño es Jitsi, tanto en el extremo transmisor como en el extremo receptor.

- En el caso de Skype, el tamaño medio de los paquetes varía desde los 124 Bytes a los 152 Bytes.
- En el caso de Jitsi, el tamaño medio de los paquetes varía desde los 213 Bytes a los 217 Bytes, lo que demuestra que en este caso el tamaño medio de los paquetes se mantiene mucho más estable comparado con Skype.

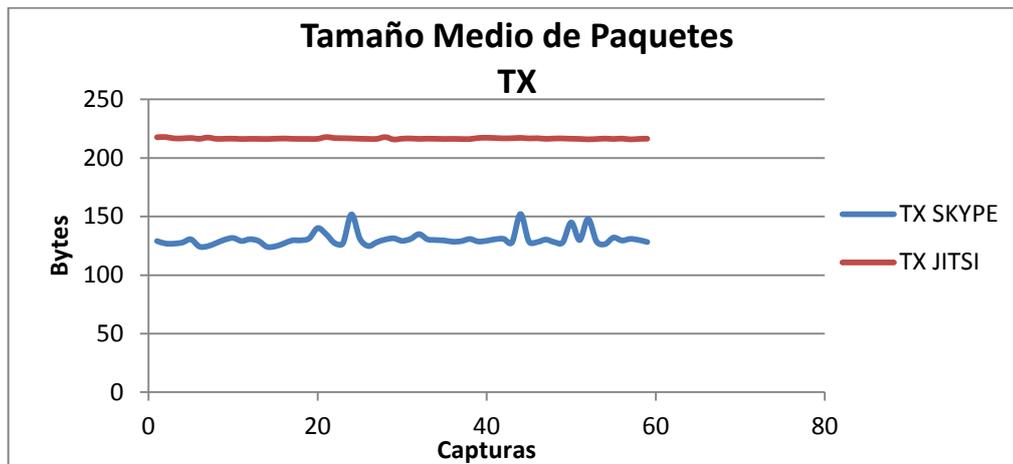


Figura 4.22: Tamaño medio de los paquetes en TX.

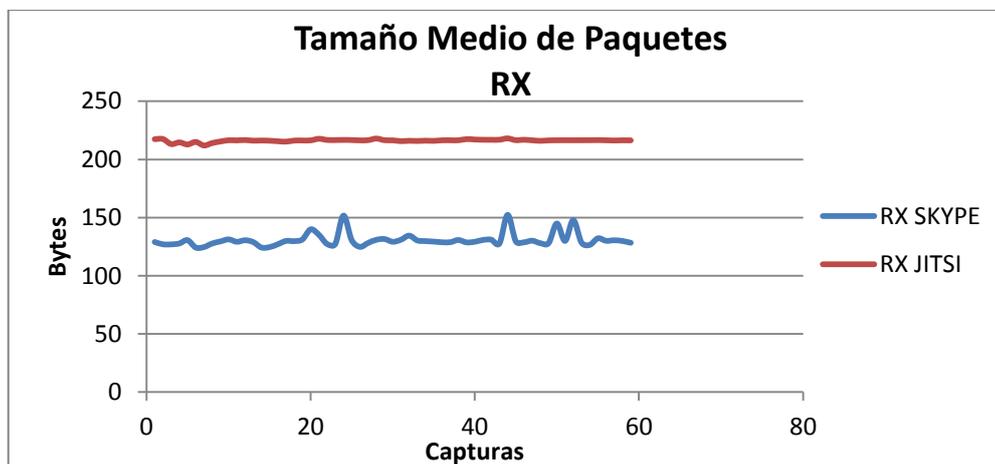


Figura 4.23: Tamaño medio de los paquetes en RX.

Si observamos los datos obtenidos en función de la aplicación (figuras 4.24 y 4.25), es decir, teniendo en una misma gráfica, para cada aplicación, el tamaño medio de los paquetes enviados y el de los recibidos tenemos:

- **Skype:** Aunque se produce una mayor variación (con respecto a Jitsi) entre el tamaño medio de los paquetes para las distintas conversaciones realizadas, ocurre que para una misma conversación el tamaño medio de los paquetes es muy similar en transmisión y recepción.

- **Jitsi:** Aunque se produce una menor variación (con respecto a Skype) entre el tamaño medio de los paquetes para las distintas conversaciones realizadas, ocurre que para una misma conversación el tamaño medio de los paquetes no es tan similar (como ocurría en Skype), sino que para algunas conversaciones el tamaño medio de los paquetes en transmisión y recepción varía en el caso máximo 5 Bytes.

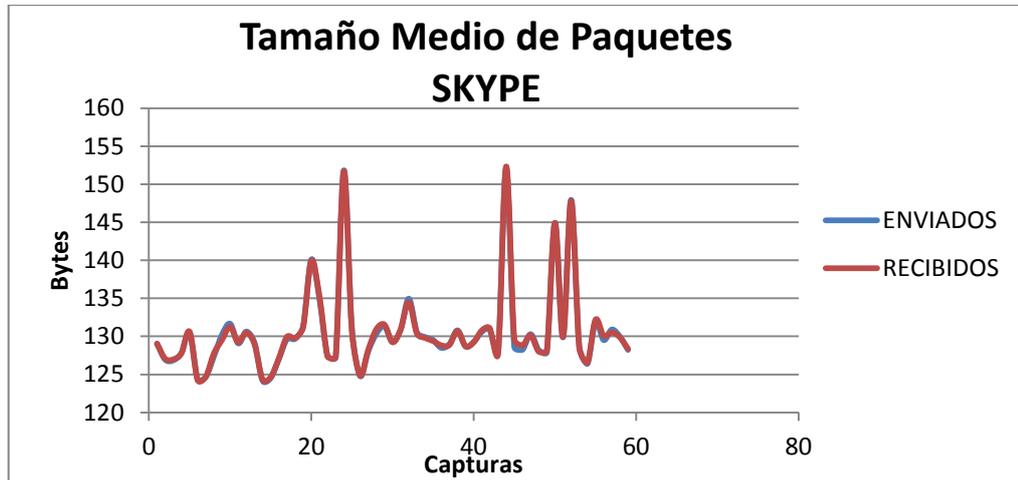


Figura 4.24: Tamaño medio de los paquetes SKYPE.

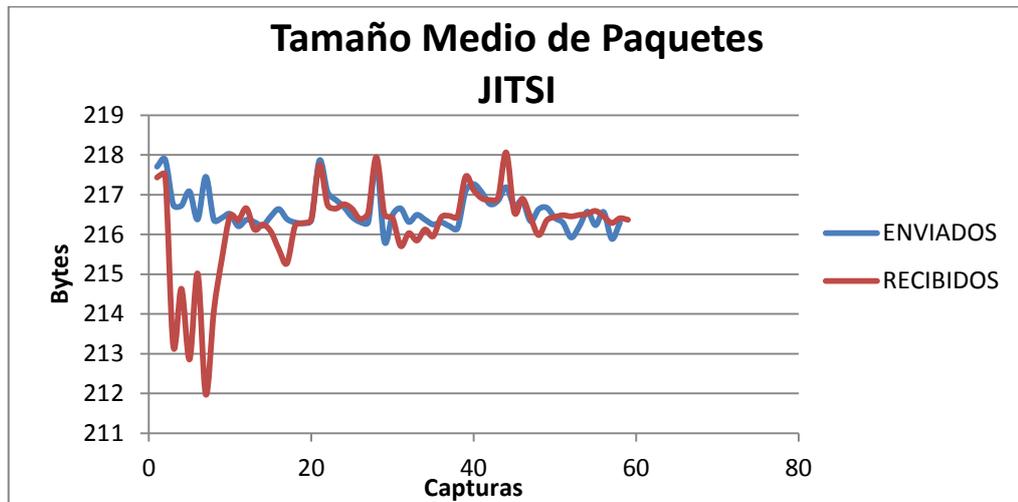


Figura 4.25: Tamaño medio de los paquetes JITSI.

4.3.3.3 Desviación Estándar del Tamaño de los Paquetes

En las siguientes gráficas (figuras 4.26 y 4.27) podemos ver la desviación estándar del tamaño de los paquetes en origen y en destino, para cada una de las dos aplicaciones analizadas.

En Transmisión los valores de la desviación estándar del tamaño de los paquetes son parecidos en ambos programas, aunque en la mayoría de las capturas este valor es levemente mayor en Skype que en Jitsi. La variación de los valores es también similar entre los dos programas.

- Jitsi: Se tienen valores entre los 26 Bytes y los 37 Bytes.
- Skype: Se tienen valores entre los 26 Bytes y los 40 Bytes.

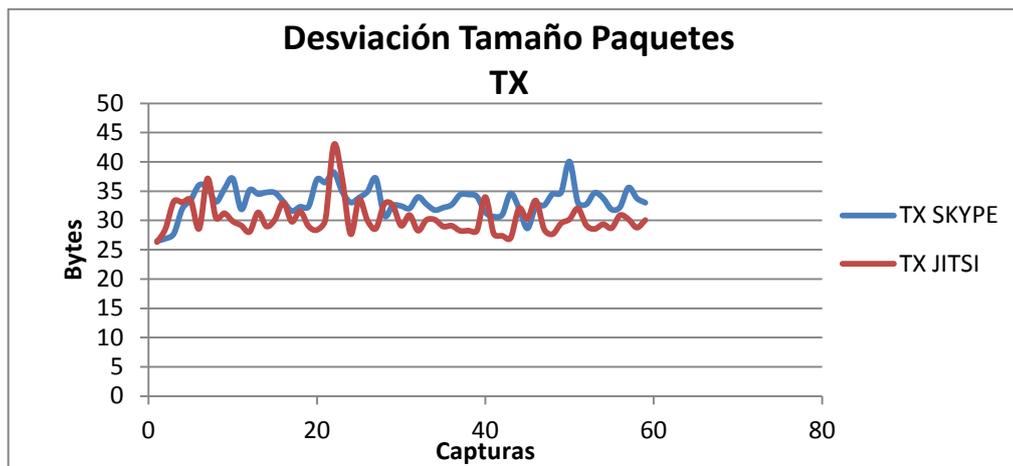


Figura 4.26: Desviación estándar del tamaño de los paquetes en TX.

En recepción los valores de la desviación estándar del tamaño de los paquetes son parecidos en ambos programas, aunque en la mayoría de las capturas, este valor es levemente mayor en Skype que en Jitsi (en este caso la diferencia entre ambos programas es menor que en transmisión). La variación de los valores es también similar entre los dos programas.

- Jitsi: Se tienen valores entre los 21 Bytes y los 75 Bytes aproximadamente, aunque en este caso el valor máximo es producido por una de las variaciones muy grandes que se producen en algunas capturas, concretamente es el caso de la captura 45, que tiene un valor de 75,6617 Bytes.
- Skype: Se tienen valores entre los 24 Bytes y los 37 Bytes.

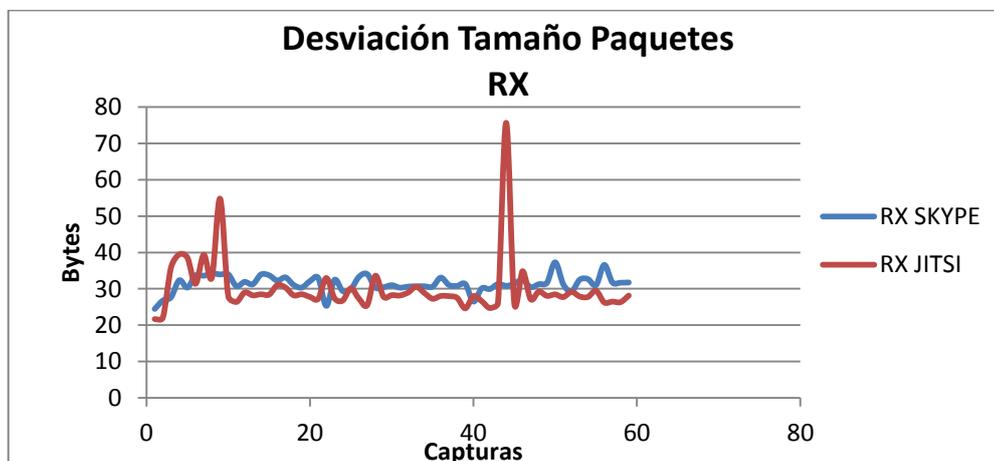


Figura 4.27: Desviación estándar del tamaño de los paquetes en RX.

4.3.3.4 Media de la Diferencia de Tiempos

El tiempo que debe existir en la recepción entre los paquetes de voz tiene que ser un valor bajo ya que si fuera alto el oyente percibiría que la conversación no es fluida y se entrecortaría debido al retraso que se provocaría de un paquete a otro.

Para garantizar que la voz percibida por el receptor no sufra modificaciones debido a la diferencia de llegada de un paquete a otro, los paquetes de voz IP deben ser enviados cada 20 ms. Esta diferencia tiene que ser constante y mantener una diferencia que el oído humano no note que se están transmitiendo en paquetes distintos. Como se verá a continuación estos requisitos se cumplen en nuestro proyecto, ya que el valor más alto obtenido en las diferentes capturas es de 13 ms.

En las siguientes gráficas (figuras 4.28 y 4.29) podemos ver la media de diferencia de tiempos en origen y en destino, para cada una de las dos aplicaciones analizadas.

Para cada una de las conversaciones y para cada uno de los programas la media de la diferencia de tiempos es muy similar en transmisión y recepción. En Skype se producen más variaciones (con respecto a Jitsi) de la media de diferencia de tiempos entre las distintas capturas.

En transmisión:

- Skype: Toma valores entre 0,0098 segundos y 0.013 segundos.
- Jitsi: Toma valores entre 0,0097 segundos 0,011 segundos.

En recepción:

- Skype: Toma valores entre 0,0098 segundos y 0.013 segundos.
- Jitsi: Toma valores entre 0,0097 segundos 0,013 segundos.

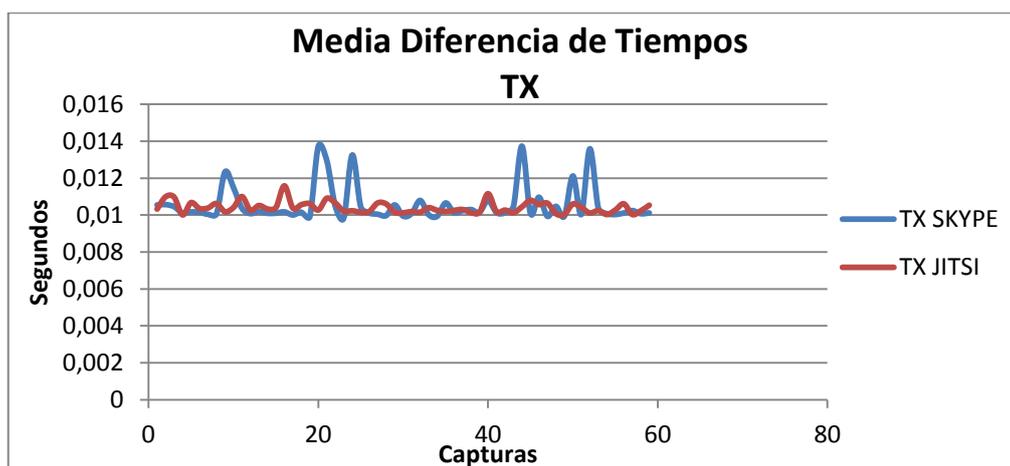


Figura 4.28: Media de la diferencia de tiempos en TX.

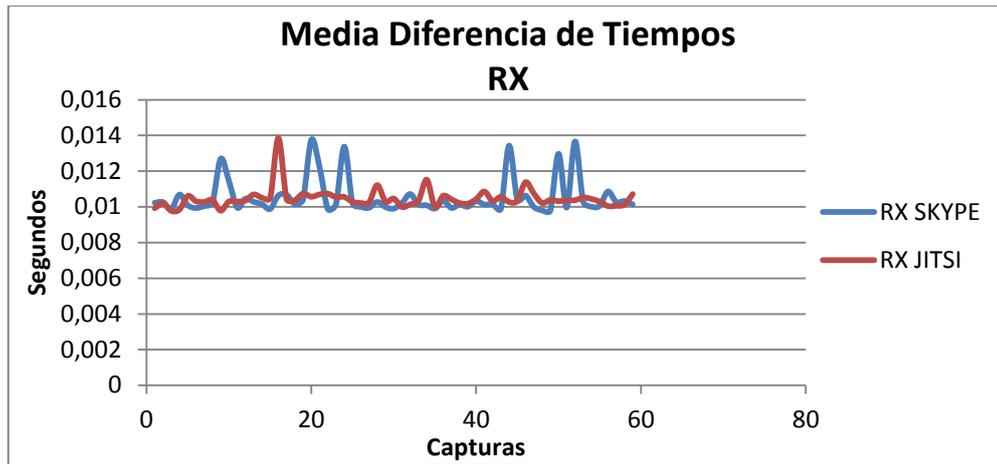


Figura 4.29: Media de la diferencia de tiempos en RX.

4.3.3.5 Desviación Estándar de la Diferencia de Tiempos

En las siguientes gráficas (figuras 4.30 y 4.31) podemos ver la desviación estándar de diferencia de tiempos entre paquetes consecutivos en origen y en destino, para cada una de las dos aplicaciones analizadas.

Para cada uno de los programas la desviación estándar de la diferencia de tiempos es muy similar en transmisión y recepción. Se puede observar que se producen grandes variaciones en las gráficas, pero al tratarse de unidades tan pequeñas (centésimas de segundo), la diferencia en la desviación estándar de la desviación de la diferencia de tiempos en cada una de las capturas es muy poco notable.

En transmisión:

- Skype: Toma valores entre 0,012 segundos y 0,042 segundos.
- Jitsi: Toma valores entre 0,0159 segundos y 0,0484 segundos.

En recepción:

- Skype: Toma valores entre 0,0068 segundos y 0,0399 segundos.
- Jitsi: Toma valores entre 0,014 segundos y 0,0737 segundos, siendo este último valor un dato puntual muy atípico con respecto a los demás.

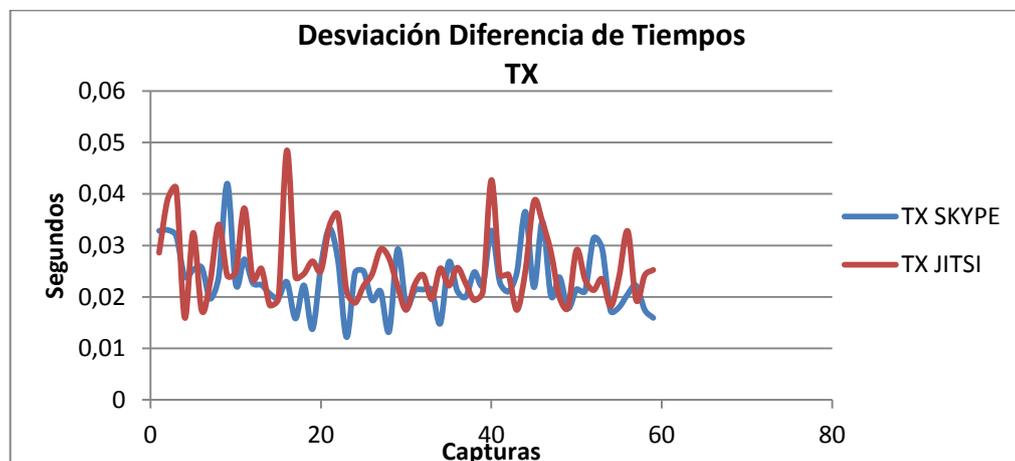


Figura 4.30: Desviación estándar de la diferencia de tiempos en TX.

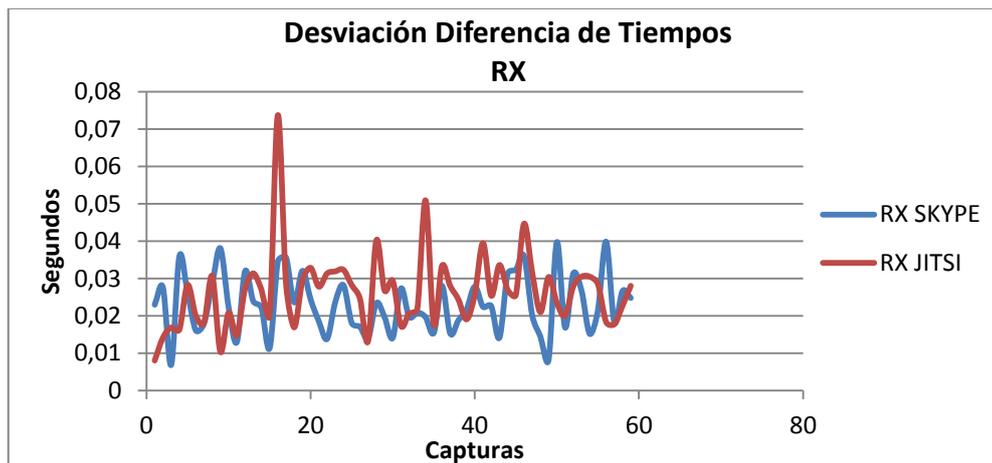


Figura 4.31: Desviación estándar de la diferencia de tiempos en RX.

4.3.3.6 Pérdida de Paquetes

El principal problema de la voz IP es que las aplicaciones están basadas en tiempo real y usan el protocolo UDP. Este protocolo de transporte es poco tolerante a la pérdida de paquetes ya que no tiene posibilidad de retransmisión. Por ello la pérdida de paquetes supone una disminución de la calidad en la conversación.

La calidad de servicio ofrecida por una de estas aplicaciones no se ve disminuida con un porcentaje de pérdidas inferior al 5% [CINI01], más de esto supone una gran degradación en la calidad de la conversación.

En la figura 4.32 podemos ver la pérdida de paquetes producida en cada una de las capturas para cada uno de los programas.

Los valores que se obtienen no superan el 3.01% de paquetes perdidos por lo no se supera en ningún caso el requisito de 5% para que sea aceptable la pérdida de paquetes.

En este caso y refiriéndose a ambos programas no se observa una monotonía en los valores de las distintas capturas, produciéndose las siguientes variaciones:

- Skype: Se tienen valores entre 0,26% y 3.01%.
- Jitsi: Se tienen valores entre 0,052% y 1,97%.

Como se puede observar el rango es mayor en skype que en jitsi.

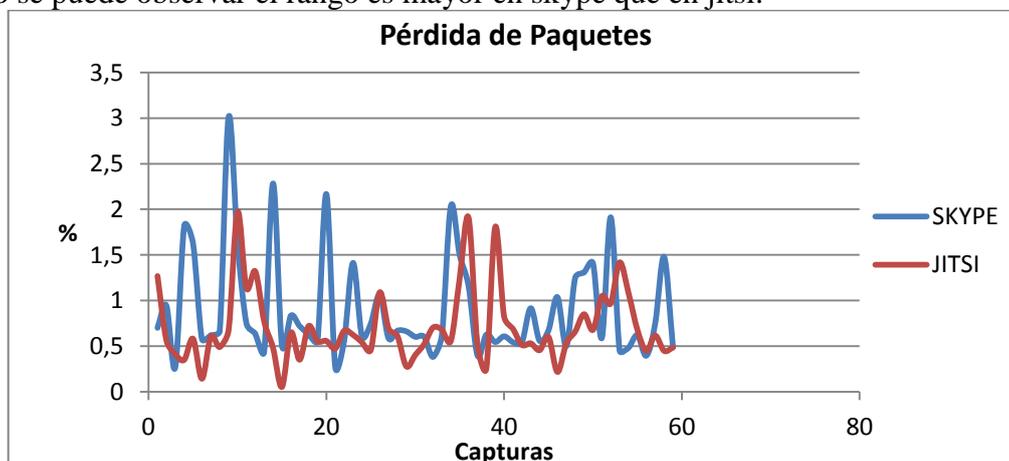


Figura 4.32: Pérdida de paquetes.

4.3.3.7 Ancho de Banda No VoIP

En este apartado analizamos el ancho de banda no perteneciente a las aplicaciones VoIP bajo estudio que podría afectar al desempeño final (figura 4.33 y 4.34).

Tanto en transmisión como en recepción el ancho de banda utilizado por otras aplicaciones distintas a las de VoIP es muy bajo, dándose algunos picos en algunas capturas en concreto cuando se utilizaba Skype.

En transmisión:

- Skype: Toma valores entre 47,6 Bytes y 10168,2 Bytes.
- Jitsi: Toma valores entre 2,68 Bytes y 168 Bytes.

En recepción:

- Skype: Toma valores entre 63,3 Bytes y 8767,98 Bytes.
- Jitsi: Toma valores entre 42,63 Bytes y 504,92 Bytes.

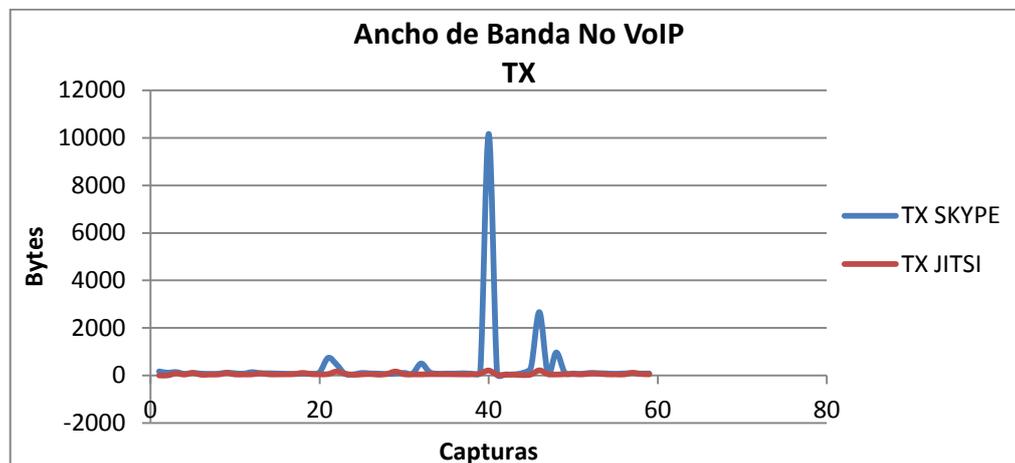


Figura 4.33: Ancho de bando no VoIP en TX.

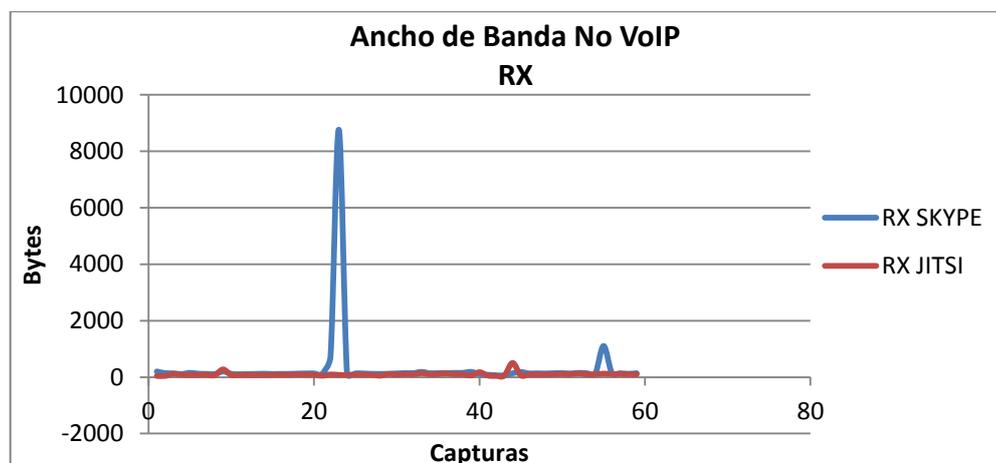


Figura 4.34: Ancho de bando no VoIP en RX.

4.3.3.8 Retardo de la Red

El retardo es también otro factor fundamental que deriva de la calidad de la transmisión y de presentarse, el usuario percibiría que en su conversación, lo que dice se superpone con lo que escucha. En caso de que el retraso sea mayor a 300ms, las conversaciones pueden llegar a ser muy molestas, razón por la cual, la recomendación G.114 de la UIT-T [ITU03] establece que el retraso en ida-vuelta no puede ser mayor a 300ms.

Antes de la realización de cada una de las capturas se realizó un ping al otro extremo para obtener el retardo de red en cada caso. Los resultados obtenidos se pueden observar en la siguiente gráfica (figura 4.35).

Se obtienen valores comprendidos entre los 15 ms y los 17 ms, siendo bastante menores de 300 ms, por lo cual se satisface la recomendación G.114 de la UIT-T y por tanto las necesidades de los usuarios.

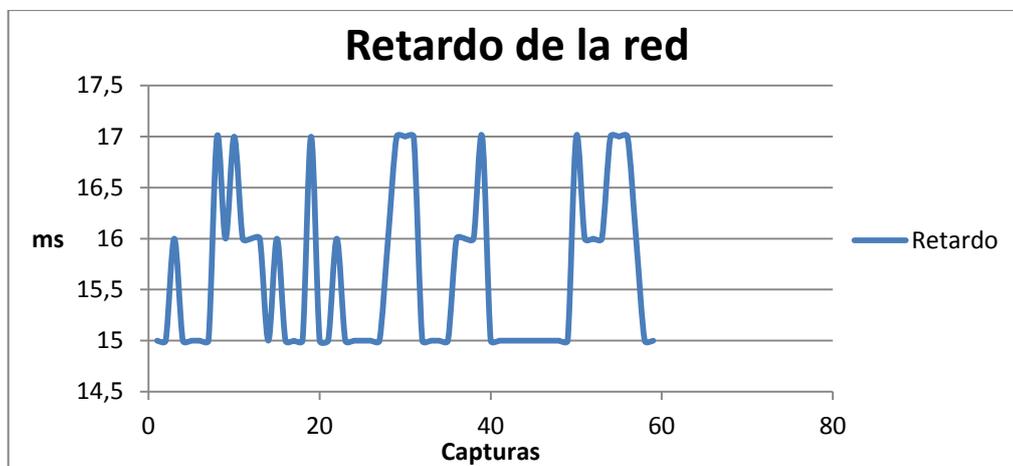


Figura 4.35: Retardo de la red.

CAPÍTULO 5

Conclusiones

Las conclusiones obtenidas tras la realización de las mediciones de QoS y de las valoraciones de QoE son las que detallamos a continuación.

En las mediciones de los parámetros de QoS se obtuvieron unos resultados que reflejan que en general podemos definir a Jitsi (código G.722) como la mejor aplicación frente a Skype (código SILK). Cabe destacar que los resultados que han sido considerablemente diferentes entre ambas aplicaciones, y que por tanto se puede decir que son los parámetros que realmente influyen en la calidad de la conversación son el tamaño medio de los paquetes y el ancho de banda utilizado. En ambos casos Jitsi obtiene un resultado mayor que Skype. Los demás parámetros son similares entre ambas aplicaciones, e incluso en algunos casos demasiado variantes.

En cuanto a los resultados obtenidos para las mediciones de QoE, también se refleja una superioridad de Jitsi en la valoración de los encuestados. La diferencia es mínima, pero Jitsi ha sido valorada mejor en la mayoría de los resultados, como es el caso de la calidad de la señal vocal, que es superior para Jitsi, aunque el esfuerzo de comprensión sea superior en este caso que en caso de Skype. Como se ha visto anteriormente, la curtosis y el sesgo muestran también en la valoración que Jitsi ha sido valorada más positivamente que Skype. Cabe destacar que a través de ANOVA se obtiene que el género, el uso de internet y el uso de aplicaciones VoIP no afectan a la valoración.

Como conclusión se obtiene que hay una clara relación entre los resultados obtenidos en las pruebas de QoS y los obtenidos en las pruebas de QoE. Aunque lo más normal es que se pueda pensar que una mejor o peor valoración de los encuestados depende fundamentalmente del retardo de red en este caso no es así. Esto puede que se deba a que en realidad la diferencia del retardo entre cada una de las conversaciones es mínima (2 ms). La valoración por parte de los encuestados, en este proyecto, está íntimamente ligada con el tamaño medio de los paquetes y por tanto, el ancho de banda utilizado por cada una de las aplicaciones. La aplicación que claramente utiliza un mayor tamaño de paquetes y mayor ancho de banda (Jitsi) ha sido la mejor valorada. Por esto se puede tomar el tamaño de los paquetes y el ancho de banda como un factor decisivo a la hora de valorar una aplicación de VoIP.

Como líneas futuras a este proyecto se propone realizar un estudio de ambas aplicaciones pero orientado a la transmisión de video sobre IP. Ya que se tendría el complemento perfecto y quedaría un estudio más completo

ANEXO A

Métodos Subjetivos para Audio

En la recomendación P.800 [ITUP96] se describen los distintos métodos separados en dos grandes categorías, test de conversación y test en los que solo se escucha (*Listening test*).

Test de conversación

En este tipo de test la idea es reproducir en un laboratorio una conversación telefónica con las condiciones reales de servicio. Para eso se colocan dos personas en cabinas acústicamente aisladas. La recomendación especifica la dimensión de las cabinas, la decoración, la atenuación al sonido externo, la densidad espectral de ruido ambiente, etc. Al final de cada conversación los participantes le asignan un valor a la calidad de acuerdo a la figura A.1.

Excelente	5
Buena	4
Razonable	3
Pobre	2
Mala	1

Figura A.1 : Valores de calidad según ACR

A la media aritmética de la colección de resultados se le llama valor medio de opinión de la conversación (*mean conversation-opinion score*) y se representa por el símbolo MOSc. Si bien este test es lo más adecuado para asignar valores subjetivos de calidad a conversaciones telefónicas los requisitos técnicos son inalcanzables para nosotros, así como la implementación de las conversaciones complicada. Por esto optamos en utilizar el segundo tipo de test, los *Listening test*.

Listening test-Absolute Category Rating (ACR)

Estos test consisten en que distintas personas escuchen muestras de audio (frases cortas) y les asignen un valor a la calidad utilizando la escala presentada anteriormente. El nombre de *absolute category rating* viene de que las personas juzgan la calidad solo escuchando la señal "distorsionada" (ya transmitida) y no tienen acceso a la señal original. En este método no se espera el mismo realismo que en los test de conversación (no se está simulando una conversación) y por lo tanto las restricciones son menos severas. Al igual que en el otro método se especifica el lugar en que se deben realizar los test y la calidad del sistema de reproducción y grabación de las muestras de audio.

Las muestras de audio consisten en cierto número de frases simples, con sentido y de corta duración. Las frases se agrupan de manera aleatoria de forma que no tengan sentido entre sí. La duración debe ser de entre 2 y 3 segundos.

La persona escucha las distintas muestras de audio, separadas entre si por unos diez segundos. En este tiempo el sujeto debe dar un valor de calidad a la muestra que acaba de escuchar, siguiendo la escala antes mencionada. A la media aritmética de los resultados de las distintas personas se denomina resultado de opinión media (*mean opinion score*) MOS.

Listening test-Degradation Category Rating (DCR)

Este método es una modificación del anterior, las muestra se presentan en parejas, separadas entre sí por 1 segundo, donde la primera es la señal de referencia (sin deterioro) y la segunda es la señal ya transmitida. Una vez que escucharon ambas los individuos tienen que asignarle un valor a la degradación sufrida de acuerdo a la siguiente escala (figura A.2).

La degradación es inaudible	5
La degradación es audible pero no molesta	4
La degradación es un poco molesta	3
La degradación es molesta	2
La degradación es muy molesta	1

Figura A.2: Valores de degradación de calidad según DCR

El valor medio de los resultados obtenidos es el DMOS (*degradation mean opinion score*).

Listening test-Comparison Category Rating (CCR)

Este método es muy similar al anterior. Las muestras se presentan en parejas, señal de referencia y señal distorsionada. Pero a diferencia del DCR donde la primera señal es siempre la de referencia, en el CCR el orden es aleatorio. Las personas asignan un valor a la calidad de la segunda señal comparada con la primera de acuerdo a la siguiente escala (figura A.3).

Mucho mejor	3
Mejor	2
Apenas mejor	1
Igual	0
Apenas peor	-1
Peor	-2
Mucho peor	-3

Figura A.3: Valores de calidad según CCR

El valor medio de los resultados es el CMOS (*comparison mean opinion score*).

La ventaja de los dos últimos métodos (DCR y CCR) es que permiten evaluar la influencia que tiene el sistema de transmisión en la degradación de la calidad.

ANEXO B

E-Model

El E-Model, abreviado de *European Telecommunications Standards Institute* (ETSI) *Computational Model*, definido en el ETSI *Technical Report* ETR 250 [ETSI96] y luego estandarizado en la *ITU-T Recommendation G.107* [ITUT05], surge como una herramienta para la planificación de redes telefónicas híbridas (mezcla de redes de circuitos conmutados (SCN) y de paquetes conmutados (PSN)). Esta herramienta permite a los proyectistas ver como los distintos parámetros de transmisión afectan la calidad que percibirán los usuarios finales. Si bien no existe un acuerdo para esto [ITUT05], el E-Model está siendo ampliamente utilizado como un método no intrusivo para estimar la calidad de servicio (QoS) en aplicaciones de voz sobre IP (VoIP).

El principio fundamental del E-Model se basa en un concepto introducido por J. Allnatt: “Factores psicológicos en la escala psicológica son aditivos”; esto simplifica bastante lo que al principio podría ser una relación muy compleja entre los distintos factores que determinan la calidad. La salida del E-Model es el factor de clasificación de la transmisión (*transmission rating factor*) R , el cual se calcula como:

$$R = R_0 - I_s - I_d - I_e - \text{eff} + A \quad (6)$$

Donde:

- **R_0** (relación señal a ruido básica, en el punto de 0 dBr): representa los efectos del ruido. Depende de los distintos ruidos presentes, desde el ruido ambiente hasta el ruido introducido por los circuitos que forman la red.
- **I_s** es la suma del deterioro que se produce de forma simultánea con la transmisión de la voz (ruido de cuantización, sidetone, nivel de recepción de la voz). Tanto este factor como el anterior están determinados por parámetros de la telefonía convencional.
- **I_d** es la suma del deterioro que se produce de forma retardada. Está determinado por el eco del lado receptor y transmisor y el retardo absoluto boca-oído. Este factor además del eco da cuenta de la pérdida de interactividad que sufren los usuarios por retardos muy grandes. Los parámetros que influyen son el retardo medio de ida, ida y vuelta, y el retardo desde el lado receptor hasta el punto en la conexión donde se produce el acoplamiento de señales como fuente de eco. Por interesarnos en una conexión sobre IP exclusivamente, tomamos este último igual que el retardo medio de ida.

- **Ie-eff** tiene en cuenta la distorsión de la voz producida por el códec que se utilice y por la pérdida de paquetes. Se calcula como:

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}} \quad (7)$$

donde

- ✓ **Ie** representa la distorsión introducida por el códec, y está cuantificada de forma empírica para distintos códecs, en [ITUT02] se encuentran valores para los diferentes códecs.
 - ✓ **Ppl** es la probabilidad de pérdida de paquetes, que incluye tanto los paquetes perdidos en la red (por ejemplo descartados en las colas de los routers), como los paquetes perdidos en el buffer del códec debido al jitter.
 - ✓ **Bpl** (*Packet-loss Robustness*) es una medida de la robustez del códec frente a la pérdida de paquetes, depende de si el códec implementa algún algoritmo de reconstrucción. En [ITUT02] se encuentran valores para distintos codecs.
 - ✓ **BurstR** (*Burst Ratio*) permite incluir el efecto de pérdidas en ráfagas. Se define como el cociente entre el largo medio de las ráfagas observadas en la secuencia de llegada sobre el largo medio de ráfagas esperado suponiendo pérdidas independientes. Un valor de uno implica pérdidas independientes mientras que un valor mayor implica pérdidas en ráfagas.
- **A** (factor de expectativa o ventaja) se debe a los diferentes niveles de exigencia de acuerdo al sistema de comunicación utilizado (por ejemplo teléfono cableado A=0, celular A=10)

El E-Model será utilizado en una red de conmutación de paquetes, simplificándose así el modelo implementado. De esta forma el resultado sólo dependerá de los parámetros de transporte de la red IP.

Para lograr esta simplificación se tomaron los valores por defecto dados en [ITUT05] para los parámetros clásicos de una red SCN, quedando como variables del modelo la probabilidad de pérdida de paquetes, el *Burst Ratio*, el retardo absoluto de ida y vuelta, y los parámetros dependientes del códec (**Ie** y **Bpl**).

Una vez obtenido el valor R, es posible mapearlo a parámetros utilizados en las medidas subjetivas de calidad como el MOSc (*Conversational MOS*).

Si bien el E-Model es atractivo para la predicción de calidad presenta varias limitaciones. Está basado en fórmulas empíricas y por lo tanto es aplicable solamente a un cierto conjunto de códecs y condiciones de la red. Para cada configuración es necesaria la validación del modelo de forma experimental, comparando los resultados con medidas subjetivas de calidad (costosas en tiempo y recursos). Actualmente ha sido verificado para cierto rango de valores de los parámetros de entrada [ITUT05].

ANEXO C

Instalación de los Programas

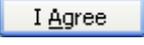
C.1 Instalación Wireshark

1. Una vez que se obtiene el instalador de la aplicación desde el siguiente link, <http://www.wireshark.org/download.html> se ejecuta el archivo wireshark-setup-1.0.0.exe (en este caso la versión es 1.0.0) para iniciar la instalación. Es importante mencionar que las librerías necesarias como WinPcap están incluidas en el instalador.

Se muestra la siguiente pantalla del asistente (figura C.1):



Figura C.1: Instalación Wireshark 1.

2. Presionando el botón  se despliega la especificación de la licencia y al presionar el botón  se despliega la siguiente ventana (figura C.2) para seleccionar los componentes que se desean instalar.

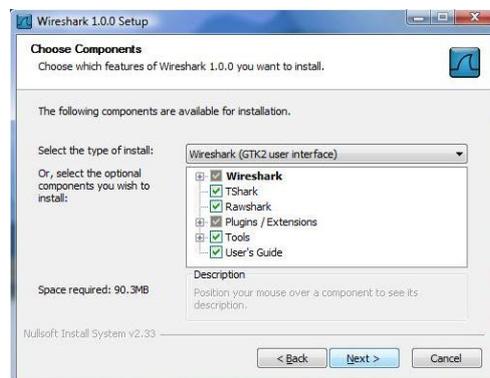


Figura C.2: Instalación Wireshark 2.

Para esta instalación se seleccionarán los siguientes:

- Wireshark, GUI del analizador de protocolos.
 - TShark, línea de comando del analizador de protocolos. *Plugins/Extensions*, especificar *plugins* y extensiones para TShark y Wireshark en este punto deberá seleccionar todos los ítems listados.
 - Tool, ofrece herramientas adicionales para aplicar a los archivos que contienen los paquetes para su análisis seleccionar todas las ofrecidas.
 - *Editcap*, para manipular los archivos.
 - *Text2Pcap*, convierte un archivo ASCII en formato libpcap.
 - *Mergecap*, permite obtener un archivo desde la combinación de 2 o más archivos de paquetes capturados.
 - *Capinfos*, es un programa que proporciona información de los paquetes capturados.
3. La siguiente pantalla (figura C.3) permite seleccionar si se desea crear un acceso directo a la aplicación en el escritorio, crear un menú de inicio y visualizar el icono en la barra de tareas. Adicionalmente se tiene la posibilidad de permitir, que los archivos generados por otros analizadores de tráfico puedan ser visualizados con Wireshark (opción que debemos seleccionar).

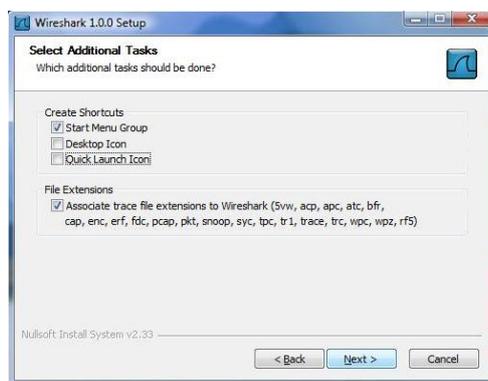


Figura C.3: Instalación Wireshark 3.

4. A continuación se deberá seleccionar el directorio donde se instalará la aplicación, en este punto se acepta el indicado por defecto en el instalador.

El instalador de WireShark contiene una versión de WinPcap se verifica si se debe actualizar versión en el PC donde se está realizado la instalación y ofrece la opción de agregar un servicio para que usuarios que no tiene privilegios de administrador pueda capturar paquetes, como se observa en la figura C.4. En este punto se seleccionan ambos ítems.



Figura C.4: Instalación Wireshark 4.

Se presiona el botón  para iniciar el proceso de instalación. Aparecerá una ventana como la figura C.5.

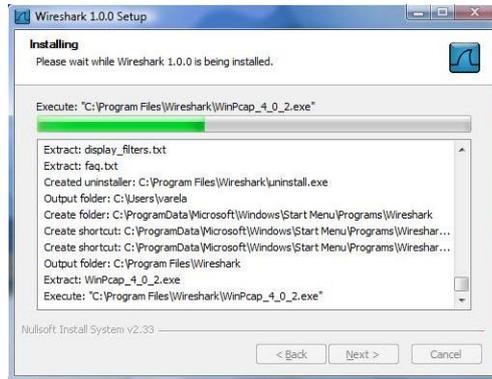
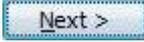


Figura C.5: Instalación Wireshark 5.

- Como se mencionó anteriormente el instalador de WireShark para Windows permite hacer la instalación de las librerías, *plugins*, servicios, etc. Particularmente para el caso de WinPcap se interrumpe la instalación en el punto que muestra la pantalla arriba e inicia el asistente para la instalación de WinPcap. Se debe seleccionar  hasta finalizar la instalación, tal y como se muestra en la figura C.6:

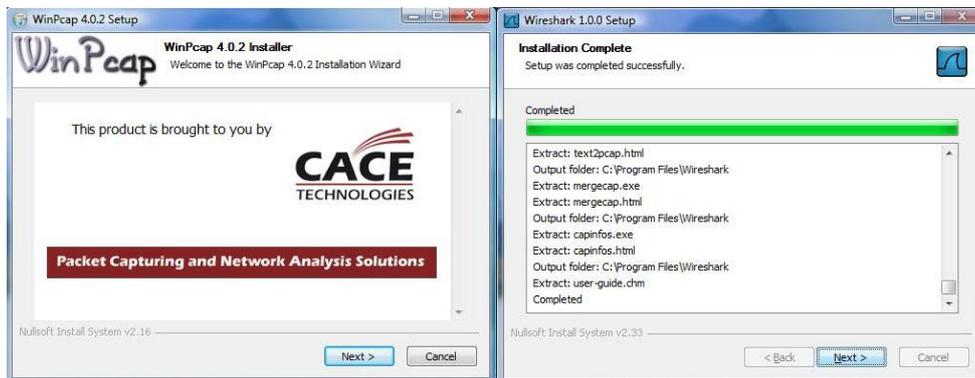


Figura C.6: Instalación Wireshark 6.

La siguiente pantalla (figura C.7) indica que la instalación ha finalizado exitosamente.



Figura C.7: Instalación Wireshark 7.

C.2 Instalación Skype

Como Skype es una aplicación gratuita se puede descargar directamente desde la página oficial de dicha aplicación, por lo que solamente hay que seguir el siguiente link:

<http://www.skype.com/intl/es/download/skype/windows/downloading/>

Es posible que tu navegador no comience la descarga inmediatamente. Va a estar esperando que confirmes la descarga desde la barra amarilla que se mostrará en la parte superior. (Ver figura C.8)



Figura C.8: Instalación Skype 1.

Haz click sobre esta barra y elige la opción Descargar Archivo. Inmediatamente se mostrará el diálogo de descarga, tal como se muestra en la figura C.9.

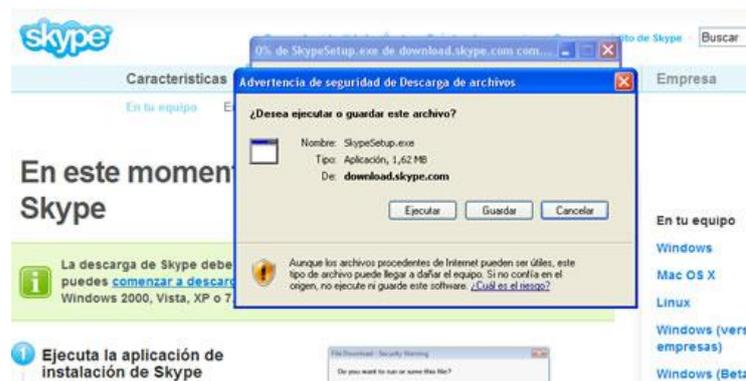


Figura C.9: Instalación Skype 2.

Elije Ejecutar y acto seguido comenzará la descarga. (Figura C.10)



Figura C.10: Instalación Skype 3.

Al finalizar la descarga, se lanzará la aplicación de instalación de Skype. Deberás elegir el idioma para que el instalador de Skype comience a descargar e instalar la aplicación en tu PC. (Figura C.11).



Figura C.11: Instalación Skype 4.

Una vez finalizado el proceso, la instalación te propone crear una nueva cuenta, tal y como se muestra en la figura C.12. Ingresas los datos solicitados y haz click sobre Acepto Crear una Cuenta. Si ya tienes una cuenta de Skype, debe presionar en la opción Iniciar Sesión.



Figura C.12: Instalación Skype 5.

Una vez creada la cuenta aparecerá una ventana como la que se observa en la figura C.13, ingresa los datos solicitados para completar tu perfil y confirma haciendo click en Aceptar.



Figura C.13: Instalación Skype 6.

En la Bienvenida a Skype (figura C.14) vas a poder acceder a una Guía para aprender a utilizarlo, y a una herramienta para chequear que los parlantes y micrófonos de tu PC funcionen correctamente.



Figura C.14: Instalación Skype 7.

Para la realización de este proyecto además de la instalación de Skype se ha tenido que realizar la siguiente configuración.

Pinchar en la pestaña *tools > option*, y aparecerá una ventana como la que se observa en la figura C.15.

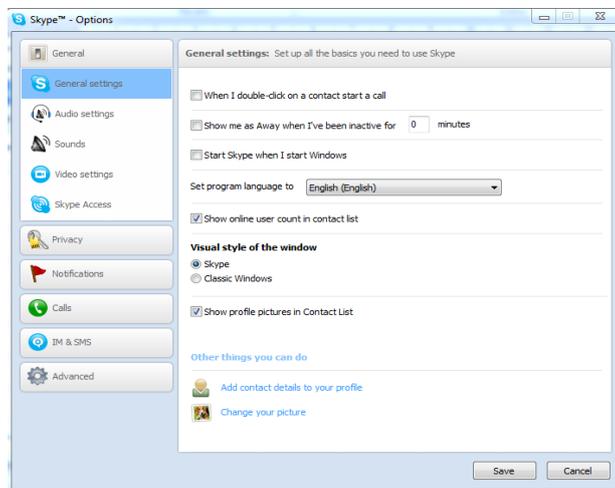


Figura C.15: Instalación Skype 8.

Pinchamos en *Audio settings* y tendremos una ventana igual a la figura C.16. Elegimos en la opción del micrófono “Mezcla Estéreo”.

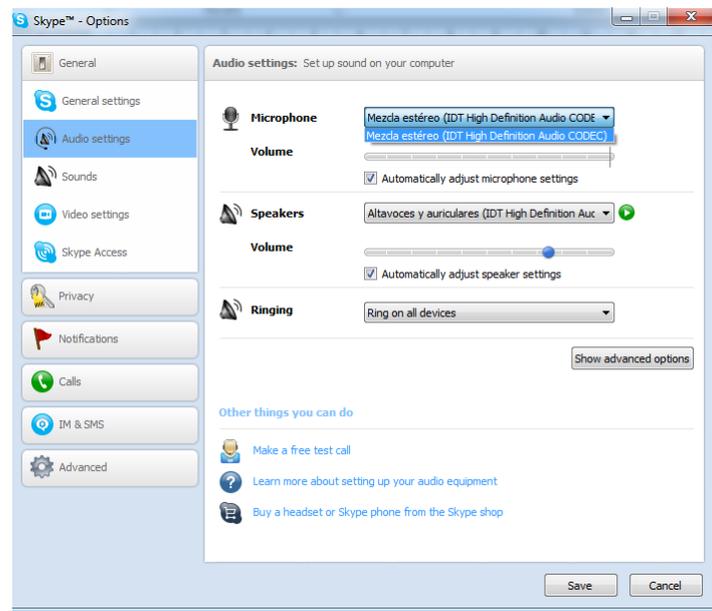


Figura C.16: Instalación Skype 9.

C.3 Instalación Jitsi

Como Jitsi es una aplicación de código abierto se puede descargar directamente desde la página oficial de dicha aplicación, por lo que solamente hay que seguir el siguiente link:

<http://download.jitsi.org/jitsi/windows/>

Al ejecutar el instalador nos aparecerá la siguiente ventana (figura C.17).



Figura C.17: Instalación Jitsi 1.

Ahora aparecerá la siguiente ventana (figura C.18), en la que se debe señalar la opción de aceptación de los términos de licencia y pinchar en Next.



Figura C.18: Instalación Jitsi 2.

En la siguiente ventana se debe pulsar Next y aparecerá la siguiente ventana (figura C.19), en la que se deben marcar las opciones que se observan en la imagen y volver a pinchar en Next.

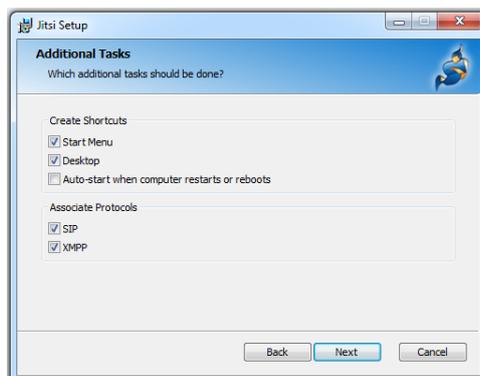


Figura C.19: Instalación Jitsi 3.

En la ventana que aparece a continuación (figura C.20) pulsar en *Install*.

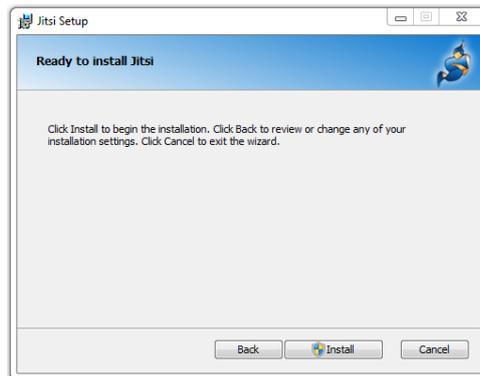


Figura C.20: Instalación Jitsi 4.

Con estos pasos ya se habrá instalado correctamente Jitsi.

Para la realización de este proyecto además de la instalación de Jitsi se ha tenido que realizar la siguiente configuración.

Pinchamos en *Herramientas > opciones*, y en la opción de Audio se eligen los parámetros tal y como aparecen en la figura C.21.

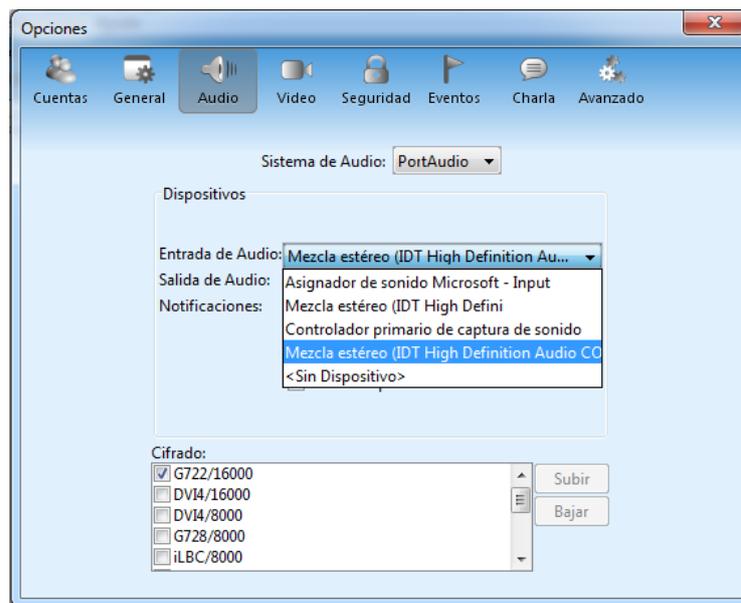


Figura C.21: Instalación Jitsi 5.

ANEXO D

Cuestionario para el Análisis de QoE

ID Cuestionario:

SOFTWARE:

Códec:

ORDEN GRABACIÓN:

Edad:

Sexo: Hombre Mujer

Valore según la escala la calidad de la señal vocal.

Excelente	5	<input type="checkbox"/>
Buena	4	<input type="checkbox"/>
Regular	3	<input type="checkbox"/>
Mediocre	2	<input type="checkbox"/>
Mala	1	<input type="checkbox"/>

¿Tuviste dificultades para hablar u oír por este enlace? Sí No

En caso afirmativo, indique cuál:

– poco volumen.....	<input type="checkbox"/>
– ruido o zumbido.....	<input type="checkbox"/>
– distorsión.....	<input type="checkbox"/>
– variaciones de nivel, interrupciones.....	<input type="checkbox"/>
– diafonía.....	<input type="checkbox"/>
– eco.....	<input type="checkbox"/>
– corte completo.....	<input type="checkbox"/>
– otras (especifíquense):	<input type="checkbox"/>

Esfuerzo necesario para comprender el significado de las frases.

Audición perfecta; ningún esfuerzo	5	<input type="checkbox"/>
Cierta atención es necesaria; ningún esfuerzo apreciable ...	4	<input type="checkbox"/>
Esfuerzo moderado	3	<input type="checkbox"/>
Esfuerzo considerable	2	<input type="checkbox"/>
Significado incomprensible, aun con el mayor esfuerzo ...	1	<input type="checkbox"/>

¿Qué grabaciones has escuchado con mayor calidad?

Grabaciones Hombre	<input type="checkbox"/>
Grabaciones Mujer.....	<input type="checkbox"/>
Ambas calidad similar.....	<input type="checkbox"/>

Indique con qué frecuencia utiliza usted aplicaciones de VoIP(similares a las utilizadas en estas pruebas).

Muy a menudo	5	<input type="checkbox"/>
A menudo	4	<input type="checkbox"/>
De vez en cuando.....	3	<input type="checkbox"/>
Alguna vez	2	<input type="checkbox"/>
Nunca	1	<input type="checkbox"/>

Indique con qué frecuencia utiliza usted Internet.

Muy a menudo	5	<input type="checkbox"/>
A menudo	4	<input type="checkbox"/>
De vez en cuando.....	3	<input type="checkbox"/>
Alguna vez	2	<input type="checkbox"/>
Nunca	1	<input type="checkbox"/>

ANEXO E

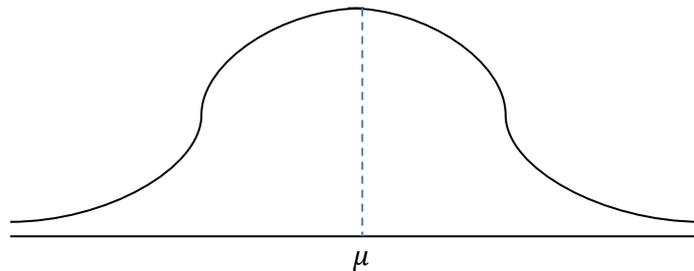
Demostración Intervalo de Confianza y Tabla t-student

E.1 Construcción de un Intervalo de Confianza con σ^2 Desconocida.

- Primero buscamos intervalos simétricos en probabilidad

$$\Pr(\mu \in (a_{x_1, \dots, x_n}, b_{x_1, \dots, x_n})) = (1 - \alpha) \quad (8)$$

$$\Pr(\mu > b_{x_1, \dots, x_n}) = \Pr(\mu < a_{x_1, \dots, x_n}) = \alpha/2 \quad (9)$$

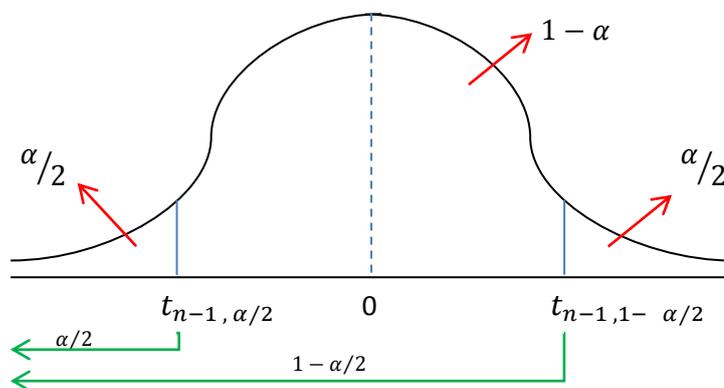


- Buscamos un estimador de μ con distribución conocida para poder calcular probabilidades. En este caso el estimador será:

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1} \quad (10)$$

- Fijado $1 - \alpha$. Sabemos:

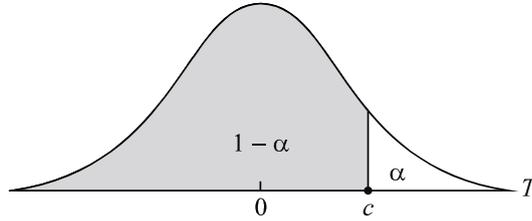
$$t_{n-1, \alpha/2} = -t_{n-1, 1 - \alpha/2} \quad (11)$$



$$\begin{aligned}
 & \Pr\left(t_{n-1, \alpha/2} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{n-1, 1-\alpha/2}\right) = 1 - \alpha \quad \Rightarrow \\
 & \Rightarrow \Pr\left(t_{n-1, \alpha/2} S/\sqrt{n} \leq \bar{X} - \mu \leq t_{n-1, 1-\alpha/2} S/\sqrt{n}\right) = 1 - \alpha \quad \Rightarrow \\
 & \Rightarrow \Pr\left(-\bar{X} + t_{n-1, \alpha/2} S/\sqrt{n} \leq -\mu \leq -\bar{X} + t_{n-1, 1-\alpha/2} S/\sqrt{n}\right) = 1 - \alpha \quad \Rightarrow \\
 & \Rightarrow \Pr\left(\bar{X} - t_{n-1, \alpha/2} S/\sqrt{n} \geq -\mu \geq \bar{X} - t_{n-1, 1-\alpha/2} S/\sqrt{n}\right) = 1 - \alpha \quad \Rightarrow \\
 & \Rightarrow \Pr\left(\mu \in \left(\bar{X} - t_{n-1, 1-\alpha/2} S/\sqrt{n}, \bar{X} + t_{n-1, 1-\alpha/2} S/\sqrt{n}\right)\right) = 1 - \alpha \\
 & \quad \boxed{\left(\bar{X} \pm t_{n-1, 1-\alpha/2} S/\sqrt{n}\right)} \quad (12)
 \end{aligned}$$

E.2 Tabla t-student

La tabla da áreas $1 - \alpha$ y valores $c = t_{1-\alpha, r}$, donde, $P[T \leq c] = 1 - \alpha$, y donde T tiene distribución t-Student con r grados de libertad..



k	$t_{0.75}$	$t_{0.80}$	$t_{0.85}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

Referencias

- [CINI01] Luis Gabriel Sienna. "VoIP: Una Nueva Alternativa en Telefonía". Julio 2001. Acceso web en la URL <http://www.cinit.org.mx/articulo.php?idArticulo=3>
- [CISC11] *Media Gateway Control Protocol (MGCP)*. CISCO. 2011.
- [DOHE04] Doherty S.; *The Survivor's Guide to 2004: Converged Voice, Video and Data* (2003). *Network Computing*: <http://www.networkcomputing.com/showitem.jhtml?articleID=17000077&pgno=7>
- [ETSI96] ETSI, "Speech Communication Quality from Mouth to Ear of 3.1 kHz Handset Telephony across Networks," Technical Report. ETR 250, 1996.
- [HSRV96] H. Schulzrinne, S. Casner, R. Frederick y V. Jacobson "RTP: A Transport Protocol for Real-Time Applications", January 1996.
- [ITUP96] *ITU-P, Recommendation P.800, Methods for subjective determination of transmission quality*, 1996.
- [ITUR02] ITU-R, Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, 2002.
- [ITUT01] *ITU-T, Recommendation H.248.1, "Gateway control protocol: Version 3"*. Septiembre 2005.
- [ITUT02] ITU-T, Recommendation G.113, "Transmission impairments due to speech processing," Mayo 2002.
- [ITUT03] ITU-T, Recomendación G.114, "Sistemas y medios de transmisión, sistemas y redes digitales. Tiempo de transmisión en un sentido", Mayo 2003.
- [ITUT05] ITU-T, Recommendation G.107, "The E-model, A Computational Model for Use in Transmission Planning," Marzo 2005.
- [ITUT06] *ITU-T, Recommendation G.723.1, "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s"*, Mayo 2006.
- [ITUT09] *ITU-T, Recommendation H.323, "Packet-based multimedia communications systems"*. Diciembre 2009.
- [ITUT88] *ITU-T, Recommendation G.722, "7 kHz audio-coding within 64 kbit/s"*. Noviembre 1988.
- [ITUT90] *ITU-T, Recommendation G.726, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)"*, Diciembre 1990.
- [JITS11] Jitsi. <http://www.jitsi.org>. 2011
- [LPGB05] L. Aspirot, P. Belzarena, G. Perera, B. Bazzano, End to end quality of service prediction based on functional regression, Conferencia HET-NET 2005. <http://iie.fing.edu.uy/investigacion/grupos/artes/publicaciones/hetnet05.pdf>
- [MDNE04] M. Baugher, D. McGrew, M. Naslund, E. Carrara y K. Norrman "The Secure Real-time Transport Protocol (SRTP)", March 2004.
- [ROHE06] Rosario Villarreal M.A., Herrera Vega F.; *El estándar VoIP. Redes y servicios de banda ancha*. (2006). *Monografias.com*: <http://www.monografias.com/trabajos33/estandar-VoIP/estandar-VoIP2.shtml>
- [SILK09] Skype [SILK Data Sheet](#), Retrieved 2009-09-01

- [SKYP11] Skype. <http://www.skype.com/intl/es/home>. 2011
- [SPEE11] Speex: A Free Codec For Free Speech.. <http://www.speex.org/>, 2011.
- [TELE10] Telefónica., “Las telecomunicaciones de nueva generación” *Cap. 7 La capa de control*. Páginas 153-162. Acceso web en la URL http://www.telefonica.es/sociedaddelainformacion/pdf/publicaciones/telecomunicacion_esng/capitulos/07_la_capa_de_control.pdf, 2010
- [WIKI08] IAX. *Wikipedia The Free Encyclopedia*: http://en.wikipedia.org/wiki/Inter-Asterisk_eXchange, 2011.
- [WIKI09] *Session Initiation Protocol. Wikipedia The Free Encyclopedia*: http://en.wikipedia.org/wiki/Session_Initiation_Protocol. Septiembre 2011.
- [WIKI10] GSM. *Wikipedia The Free Encyclopedia*: <http://en.wikipedia.org/wiki/GSM>, 2011.
- [WIKI11] *Speech Encoding. Wikipedia The Free Encyclopedia*: http://en.wikipedia.org/wiki/Speech_encoding
- [WIRE11] Wireshark. www.wireshark.org. 2011