ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

"ESTUDIO DETALLADO DEL USO RTP/RTCP Y SERVICIOS DE QOS Y QOE EN INTERNET PARA LA VOIP"

TESIS DE GRADO

Previa a la obtención del Título de:

MAGISTER EN TELECOMUNICACIONES

Presentado por:
ING. MARIA LUISA VINUEZA BELTRAN

GUAYAQUIL – ECUADOR
2015

A G R A D E C I M I E N T O S

AGRADECIMIENTOS

Quiero agradecer a Dios por estar aquí a mi lado, siempre.

A mi amado esposo, por su inmenso amor y apoyo incondicional.

A mis padres por su amor y paciencia, por hacer este triunfo tan suyo como mío.

A Rafita y Anita, mis hermanas, gracias por creer en mí, siempre.

A Álvaro por su apoyo e inmensa paciencia.

A mi querida ESPOL por darme la oportunidad de crecer.

A la gente linda, familiares, amigos y vecinos que con su amabilidad y entusiasmo apoyaron la culminación de esta tesis a pesar de las interminables pruebas.

D E D I C A T O R I A

DEDICATORIA

A Dios, a mi esposo y a mis padres.

TRIBUNAL DE SUSTENTACION

Sara Rios, Ing.

Sub-Decana FIEC

Álvaro Suárez Sarmiento, Ph.D.

Director de Tesis

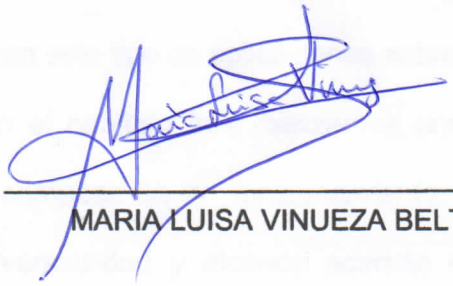
Boris Ramos, Ph.D.

Vocal Principal

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITECNICA DEL LITORAL"

Reglamento 4256 TITULO IV Capítulo II Art. 18 literal c)



MARIA LUISA VINUEZA BELTRAN

RESUMEN

Este trabajo de tesis se basa en un estudio bibliográfico investigativo sobre el protocolo RTP/RTCP y su incidencia sobre Internet tomando como referencia el desempeño de dos aplicaciones ampliamente utilizadas en la actualidad.

En el capítulo 1 realizamos un compendio sobre el presente trabajo, sus objetivos, metodología y resultados esperados. En el capítulo dos nos internamos en el mundo de la comunicación en tiempo real, nos preguntamos que necesita una aplicación para cumplir las exigencias demandadas por los usuarios que utilizan este tipo de aplicaciones sobre Internet.

Seguidamente, en el capítulo tres realizamos una visión interna al protocolo RTP/RTCP, su composición y funcionamiento. Tratamos de comprender ampliamente su versatilidad y alcance además de los conceptos sobre los cuales fue desarrollado para entender el éxito en el ámbito de VoIP y sus diferentes aplicaciones.

En el capítulo cuatro realizamos un extracto de los métodos de medición de calidad de voz, su historia y el impacto obtenido en su aplicación y los resultados obtenidos a través de los años.

Finalmente, para complementar el presente estudio, realizamos una prueba de campo realizando mediciones de calidad de voz en dos aplicaciones altamente utilizadas sobre una muestra concreta dando resultados interesantes digna de análisis para futuras investigaciones.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

CAPÍTULO 1

1. MARCO REFERENCIAL

1.1 DESCRIPCIÓN DE LA VOZ SOBRE IP

La constante evolución e innovación de la tecnología ha sido progresiva y exponencial en los últimos treinta años. Cada día las pequeñas, medianas y grandes empresas vuelcan sus esfuerzos en mecanismos que permitan proveer más y mejores servicios a los usuarios finales con calidad, fidelidad y ajustable a sus necesidades. Por esta razón, uno de los más relevantes conceptos que se maneja en la en la actualidad es la convergencia de todos los servicios de voz, vídeo y datos sobre Internet.

La *Voz sobre Internet Protocol (VoIP*) es la tecnología que engloba el transporte de la voz encapsulada en paquetes IP, transporte realizado con las exigencias propias de aplicaciones en tiempo real tales como baja latencia, poca tolerancia a pérdida de paquetes y ancho de banda reducido [1]. Dicha red de datos es una concepción diferente de los antiguos circuitos telefónicos conmutados que era útil únicamente para un propósito, envío de voz de forma dedicada.

1.2 JUSTIFICACIÓN

El auge de la convergencia de servicios en Internet crece día a día, nuevas aplicaciones, marcos de referencias, tecnologías y mayores exigencias del usuario surgen en el camino. Esta tesis ofrece un marco de apoyo técnico para investigaciones basadas en aplicaciones de tiempo real y el uso del *Real Time Protocol* (*RTP*) conociendo sus características y funcionalidad, dado el gran despliegue de este estándar por su flexibilidad que puede ser utilizado no solo para el transporte de audio y video sino también en otros escenarios como las pizarras compartidas y otros componentes de trabajo colaborativo en tiempo real.

Adicionalmente, para la comunidad juvenil con espíritu investigativo, dar herramientas y/o procesos para poder determinar técnicamente el desempeño de sus aplicaciones favoritas o personalizadas basados en parámetros de Calidad de

servicio (*QoS*, del inglés *Quality of Service*) y experiencia de usuario (*QoE*, del inglés *Quality of Experience*) tal como se describe en la parte experimental de esta tesis.

1.3 OBJETIVOS

El objetivo principal de esta tesis es estudiar detalladamente el entorno de aplicaciones de voz en tiempo real basada en RTP.

Los objetivos específicos son:

- Determinar las características y funcionalidad de RTP y su gran despliegue en aplicaciones de tiempo real.
- Conocer las métricas de calidad de servicio y experiencia disponibles para determinar el desempeño de una aplicación.
- Analizar el comportamiento y rendimiento de RTP en Internet sobre aplicaciones VoIP.
- Determinar el desempeño de aplicaciones de VoIP populares en base a parámetros de calidad.
- Poner en práctica los conocimientos adquiridos de VoIP en la Maestría de Telecomunicaciones.

1.4 METODOLOGÍA DE LA INVESTIGACIÓN

Inicialmente se realiza una investigación descriptiva sobre RTP y métricas de QoS para VoIP. Dicha investigación está basada en libros, revistas científicas, publicaciones en línea, investigaciones y trabajos de tesis de pre-grado y post-grados de ESPOL y otras universidades a nivel mundial disponibles en la Web, artículos y material público con fuente reconocida.

Luego se realiza una prueba de campo, se analizan tres aplicaciones VoIP ampliamente utilizadas como son Windows Messenger, Skype y Gtalk mediante la instalación de un analizador de paquetes en la red capturando todo el tráfico de llamadas de prueba para determinar sus parámetros de QoE y QoS tomados desde diferentes proveedores. Adicionalmente se realiza encuestas a usuarios Politécnicos de Guayaquil/Ecuador sobre dichas aplicaciones para determinar la QoE de dichas aplicaciones en base al concepto *Mean Opinion Score* (*MOS*).

1.5 RESULTADOS ESPERADOS

- Herramienta descriptiva de estudio para futuras investigaciones en base al RTP y su uso en aplicaciones de VoIP.
- La aplicación VoIP, Skype, es la más robusta y tolerante a factores negativos sobre las redes de pruebas en comparación con Windows Messenger y Gtalk en base a las métricas de QoS establecidas.

- Se espera demostrar que la aplicación Skype tiene el más alto desempeño obtenido en base al parámetro MOS basado en encuesta.

1.6 ELEMENTOS DIFERENCIADORES O INNOVADORES DE LA INVESTIGACIÓN

En la práctica existen muchos estudios sobre la QoS y la QoE de las aplicaciones típicas de VoIP. Sin embargo, es difícil encontrar el análisis de estos parámetros, concretamente en un entorno de los usuarios politécnicos de Guayaquil Ecuador. Este estudio por tanto es muy útil para aquellas empresas que deseen establecerse en Ecuador. Por otro lado, a nivel científico-técnico este es un área en constante evolución y por tanto es muy importante contar con el diseño de nuevas pruebas de rendimiento de estas aplicaciones que puedan ser la base de futuras publicaciones en las que se sintonice los trabajos teóricos con la práctica que guíe una buena innovación empresarial.

Complementar estudios técnicos antecesores sobre RTP y sus características y despliegue en nuevas aplicaciones y tecnologías, es otro elemento diferenciador.

El uso de la tecnología y software gratuitos para determinar técnicamente el desempeño de aplicaciones VoIP ampliamente utilizadas aplicando los

conocimientos adquiridos en la maestría para, en base a un criterio técnico, analizar sus resultados y generar recomendaciones a futuros usuarios.

1.7 PERFIL DE LA TESIS

Esta tesis es un estudio referencial del RTP y sus componentes e impacto sobre la comunicación de VoIP. En el capítulo 2 definimos la comunicación de voz en tiempo real, resaltando las exigencias de este tipo de tráfico sobre redes tan volubles y cambiables como el Internet, adicionalmente se revisa VoIP, sus características y protocolos ampliamente utilizados actualmente.

El capítulo 3 introduce el RTP como eje fundamental de esta tesis se realiza una revisión minuciosa de sus características, funcionalidades, limitantes y procesos de tal forma que podamos comprender el marco teórico sobre el cual se desarrollan las herramientas y aplicaciones que serán utilizadas en las pruebas de campo. Dicha revisión se realiza fundamentalmente sobre el *Request For Comment* (*RFC*) 3550 publicado por la Internet *Engineering Task Force* (*IETF*).

Posteriormente en el capítulo 4 revisamos las métricas para medir calidad de servicio para el tráfico de voz, conceptos que se aplican a las pruebas de campo que se

realizan y detallan en el capítulo 5 mediante la utilización de herramientas ampliamente utilizadas por la comunidad tecnológica.

CAPÍTULO 2

2. COMUNICACION DE VOZ EN TIEMPO REAL

2.1 DEFINICIÓN

La comunicación es un proceso de intercambio de información vital para el desarrollo del ser humano dentro de la Sociedad. Bajo este precepto el hombre ha desarrollado durante toda su existencia diferentes métodos y formas para establecer mecanismos de comunicación sin limitante de tiempo y espacio. Con el desarrollo de la tecnología y la necesidad incesante de llegar más lejos en menor tiempo, la comunicación en tiempo real como celulares, chats, VoIP, conferencias, teleconferencias, entre otros, ha tomado un impulso definitivo en nuestra era.

Un sistema en tiempo real como lo describe Donald Gillies [2] es "*aquel en el que para que las operaciones computacionales sean correctas no depende solo de la*

lógica e implementación de los programas computacionales sean correctos, sino también en el tiempo en el que dicha operación entrego su resultado. Si las restricciones de tiempo no son respetadas el sistema se dice que ha fallado". Por lo tanto para que la comunicación sea efectiva, basada en sistemas de tiempo real, debe cuidar que los interlocutores no perciban el tiempo de ejecución de transacciones, transferencia de información, camino o medio utilizado sino una respuesta menor a la sensibilidad de los sentidos, en caso de seres humanos, o a la tolerabilidad de latencia de las máquinas.

La Telefonía sobre redes dedicadas ha sido ampliamente estudiada e implementada con una calidad exquisita, sin embargo, la comunicación en tiempo real sobre Internet es aún hoy en día un problema sin resolver eficientemente. Por ello, los sistemas de VoIP son muy sensibles, dadas las características de retraso y congestión que definen a Internet, a restricciones temporales de tiempo real firme (*firm real time*) [3], por no ser un sistema de tiempo real crítico, en el que una violación de algún requisito de tiempo real puede producir un fallo generalizado del sistema. Sin embargo, aunque el sistema no falle generalizadamente, sí que es posible que cause una degradación importante en la percepción del usuario final. Esto sería una razón suficiente para que los sistemas de VoIP no fueran utilizados. Por ello es importante analizar las características de la comunicación de voz en tiempo real sobre Internet.

2.2 CARACTERÍSTICAS

En comunicaciones humanas, los sentidos que participan en la transmisión/recepción de la voz son el sentido de la audición y el gusto, por lo tanto, las variaciones de envío de paquetes de voz deben estar dentro de la sensibilidad de estos sentidos para que las fallas de la red de datos no sean perceptibles a sus interlocutores. En comunicaciones virtuales, las máquinas son más sensibles en el análisis del patrón recibido y enviado debido a su capacidad de procesamiento, es decir, el umbral de sensibilidad varía considerablemente con los sentidos humanos. A continuación se detallan dos características importantes de la comunicación ligadas a los sentidos humanos.

2.2.1 Requisitos de tiempo

"*La comunicación en tiempo real tiene como característica importante el hecho de que el valor de la comunicación depende del momento en que los mensajes llegan al destino*" [1]. Para que la voz llegue a su destino, ésta debe atravesar varios procesos para su transporte sobre la red de datos (proceso esquematizado en la Figura 2.1). Inicialmente necesitamos un primer contacto con el mundo analógico, el micrófono el cual captura la señal continua en tiempo y amplitud, la muestrea y cuantifica a una velocidad mínima determinada por la frecuencia de Nyquist. El resultado pasa al proceso de codificación y compresión, dependiendo del algoritmo a utilizar se determina el ancho de banda necesario para su transporte. Una vez lista la

información a transmitir, se fragmenta en base a la tecnología de nivel de enlace utilizada y finalmente se empaqueta para ser transportada por IP. Un proceso similar debe pasar en el destino cuando arriba el paquete y se convierta nuevamente en una señal analógica para su interlocutor.

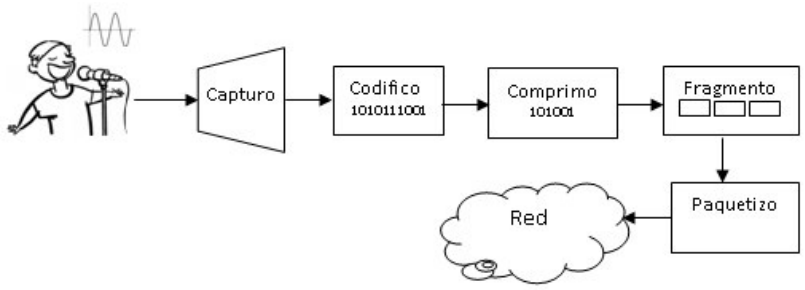


Figura 2.1. Proceso de digitalización y envío de la voz.

Una vez que la voz se empaqueta y envía a la red, existen 2 factores que influyen de forma determinante en el valor de la información, la latencia y el jitter, conceptos que revisamos más a detalle en el literal 2.4. El retraso es un factor imperante en sistemas de tiempo real pero es medianamente tolerable porque puede ser medible y/o ajustable. El tiempo de retraso o latencia de origen a destino está afectado por diferentes procesos que se realizan dentro de la red como son el retraso de propagación, transmisión, encolamiento, procesamiento de los encaminadores, entre otros. El retraso entre paquetes (*jitter*), a diferencia del retraso, es un parámetro aleatorio que afecta sobremanera el envío de información en tiempo real y debe estar en márgenes tolerables para las aplicaciones.

Todos estos procesos conllevan tiempo, el cual debe ser menor a lo percibido por el usuario final para la realidad física del ser humano a fin de que se denomine "*comunicación en tiempo real*". El retraso máximo en VoIP debe ser entre 200 y 250 ms [4] a fin de no experimentar baja calidad en los tiempos de respuesta y no sea perceptible para el oído humano.

2.2.2 Requisitos de ancho de banda

El ancho de banda necesario está sujeto a los mecanismos disponibles de codificación y compresión de los diferentes tipos de tráfico, pero mientras más compresión haya, la calidad de la información disminuiría gradualmente, por tal motivo debemos de elegir el algoritmo a utilizar para la compresión sin sacrificar la inteligibilidad del mensaje.

El ancho de banda necesario es diferente para cada tipo de tráfico, por ejemplo, el ancho de banda necesario para transmisión de video es determinada por la velocidad de *frame* (cuadro), resolución y color de la imagen y puede ocupar sin comprensión entre 90 a 270 Mbps por canal [5]; pero estos factores deben de estar sujetos a la sensibilidad del ojo humano que mínimo debe visualizar 20 cuadros por segundo para crear una ilusión de movimiento.

A diferencia del video, el audio es un factor más manejable a nivel de ancho de banda, por ejemplo en el proceso de cuantificación determinamos la calidad que deseamos mantener: la voz con calidad telefónica necesita un muestreo de 8 Khz, es decir, tomar una muestra cada 125 μs a diferencia de la voz *Hight Fidelity* (*HiFi*) debe tener un muestreo de 20 Khz. Una vez digitalizada la voz con la calidad definida, pasa al proceso de codificación y compresión, el algoritmo a utilizar para este proceso determina el ancho de banda necesario. Existen Codificadores-DECodificadores (*CODEC*) desde libres y gratuitos hasta licenciados y patentados que determinaran el ancho del canal para el envío de la voz; este lo revisamos en la sección 2.5.3.

2.3 PROTOCOLOS UTILIZADOS

Las diferentes marcas de equipos tecnológicos en el Mercado y su versatilidad en precios, el acceso a Internet con tendencias de aumento de velocidades y reducción de costos, la necesidad de comunicación a cualquier parte del Mundo en el menor tiempo posible ha motivado sobremanera el paso de tráfico multimedia sobre Internet, gracias al gran despliegue global que ésta ha tenido. Por tal motivo las aplicaciones deben de ir acorde a dichas exigencias y los desarrolladores tienen un gran desafío, generar una experiencia al usuario satisfactoria basado en las características de IP; para tal efecto, cuentan con múltiples mecanismos para su ejecución, entre ellos tenemos el estándar RTP, el más utilizado a nivel de aplicaciones.

RTP

Es un estándar que fue publicado en 1996 bajo el RCF 1889, desarrollado por el grupo de Audio y Video de IETF, posteriormente fue actualizado bajo el RFC 3550 en el 2003 [6]. Este protocolo fue creado como un marco de referencia para el transporte de tráfico multimedia en tiempo real. Usualmente trabaja, a nivel de transporte del modelo *Open System Interconnection* (*OSI*), sobre *User Datagram Protocol* (*UDP*) debido a que permite acoplarse a los niveles de exigencias de tiempo requeridos por el tráfico en tiempo real sacrificando en su media el envío confiable de paquetes. RTP no es exclusivo de IP, puede trabajar con otros protocolos diferentes a nivel de red, por ejemplo *Asynchronous Transfer Mode* (*ATM*) y puede manejar tráfico unicast o multicast.

Está compuesto de dos partes: protocolo de transferencia de datos, el cual se encarga de administrar el envío de paquetes y el protocolo de control, el cual el cual monitorea la calidad del envío, RTP y *Real Time Control Protocol* (*RTCP*) respectivamente. Ambos protocolos utilizan puertos diferentes comprendidos entre 1025 y 65535, por defecto, se usan los puertos 5004 y 5005 respectivamente.

Se considera a RTP como un protocolo flexible porque el diseñador de la aplicación multimedia particular puede armar algoritmos concretos de: reproducción, detección y corrección de errores, sincronización, se ajusta a las necesidades o características

del producto final deseado. Esta propiedad de flexibilidad es muy importante porque permite adaptarse a nuevos formatos de codificación o compresión de la voz, mecanismos avanzados de manejo de errores y seguridad… Este protocolo se revisa a profundidad en el siguiente capítulo.

Otro protocolo particular, que cabe mencionar es *Inter-Asterisk eXchange* (*IAX*), el cual se revisa en la sección 2.5.4.

2.4 PROBLEMAS DE COMUNICACIÓN DE VOZ EN INTERNET

Luego de haber revisado los requisitos de tiempo y ancho de banda en la sección anterior, queda claro que si no se cumplen estos parámetros para la transmisión de la voz sobre Internet, la comunicación en tiempo real se ve degradada adicionando además la complejidad del uso de diferentes tecnologías, administración y configuración de las redes que componen el Internet. Entre los problemas comunes o factores que influyen en las exigencias de la comunicación en tiempo real tenemos:

- *Retraso o latencia*. Es el tiempo requerido para realizar el procesamiento y transmisión de la información a través de la red. Es un factor imperante en sistemas de tiempo real pero es medianamente tolerable porque puede ser medible y/o ajustable. El retraso máximo en VoIP, como mencionamos en el apartado anterior, debe ser entre 200 y 250 ms [4].

- *Jitter*. Es la variación del retraso [5], es decir, es el máximo diferencial de tiempo en el cual un paquete es enviado y llega a su destino [1].
- *Paquetes perdidos*. Es la cantidad de paquetes enviados y no recibidos en el destino, esto puede ser causado por diferentes factores: paquetes duplicados, paquetes corruptos, paquetes encolados en buffers, pérdida de enlace en el camino a seguir, entre otros [7]. Actualmente, la prueba más utilizada por los administradores de red para verificar la existencia de paquetes perdidos es el *Packet Internet Groper* (*PING*), orden que permite diagnosticar en tiempo real, pero brevemente, el porcentaje de paquetes perdidos de extremo a extremo. Existen otras herramientas de monitoreo permanente por ejemplo el *Cacti* [8] o aplicaciones de medición más profunda como el *Wireshark* [9] que captura el tráfico de la red y presenta el porcentaje de paquetes perdidos pero no patrones o explosiones de tráfico.

2.5 FUNDAMENTOS BÁSICOS DE LA TECNOLOGÍA DE VOIP

En este apartado revisamos los conceptos básicos y generalidades de la VoIP.

2.5.1 Definición

VoIP es una tecnología que permite el envío de paquetes de voz sobre redes de datos basadas en IP. Actualmente todos los servicios están convergiendo a infraestructura

IP para proveer mayores y mejores experiencias al usuario final, la tecnología está en etapa de maduración en la actualidad.

2.5.2 Beneficios

- *Bajo Costo*. Para proveedores o usuarios finales, disponer de la tecnología VoIP es casi transparente a nivel de costos si poseen infraestructura IP. La inversión a nivel de proveedores se limita a aumentar la capacidad de sus equipos y enlaces para atender mayor demanda de tráfico pero no es una migración total o parcial de equipos para brindar este servicio. Adicionalmente, los costos de operación y mantenimiento se mantienen sobre una misma red y no serían dos rubros diferenciados en presupuestos empresariales. A nivel de usuario final, si dispone de internet en su hogar o lugar de trabajo, un teléfono IP o un softphone lo integra a la tecnología VoIP es con inversión mínima o cero.
- *Portabilidad*. En la telefonía tradicional estábamos sujetos a una posición geográfica definida, con un numero asignado limitadamente de acuerda al área, país o continente. En VoIP podemos atender nuestras llamadas desde cualquier parte del mundo en el cual nos conectemos a Internet y tengamos acceso a nuestro proveedor de cuenta telefónica y seremos la misma persona atendiendo los requerimientos de nuestros amigos y familiares con la misma identificación y perfil.

- *Crea valor agregado*. A diferencia de la telefonía tradicional, en el cual las características de la línea estaba sujeta a la facturación del servicio y su utilización estaba enmarcada a una llamada de voz, transferencia de llamada, identificador de llamada y en unos casos más sobresalientes la videoconferencia con altos costos. VoIP a más de proveer lo ofrecido por la telefonía tradicional permite la integración con otras aplicaciones generalizadas como mensajes instantáneos, video llamadas, transferencia de fotos y archivos bajo la misma infraestructura y por el mismo costo, sin dejar de lado otras aplicaciones ampliamente utilizadas como el correo, redes sociales, entre otros.
- *Flexibilidad*. Permite que el usuario final modele su aplicación y uso de acuerdo a sus necesidades, puede personalizar su ubicación, descripción, perfil, modo de contacto, entre otros. Adicionalmente, en el Mercado existen sin número de aplicaciones gratuitas o licenciadas que pueden ser elegidas por el cliente.
- *Expansión de IP*. El acceso a Internet cuenta con un gran despliegue de tecnologías que permiten llegar a casi cualquier parte del Mundo: *Asymmetric Digital Subscriber Line* (*ADSL*), *Very Small Aperture Terminal* (*VSAT*), *Gigabit-capable Passive Optical Network* (*GPON*), *Worldwide Interoperability for Microwave Access* (*WIMAX*), entre otros.

- *Escalable*. La VoIP está sujeta a la escalabilidad de Internet, por el despliegue de tecnología y abaratamiento de costos.

2.5.3 CODECs usados en VoIP

Como mencionamos en el apartado 2.2.1, para que la voz humana pueda ser empaquetada y transportada a través de la red IP, ésta debe ser digitalizada, es decir, tomar muestras continuas a intervalos regulares y asignar un valor en bits, comprimirlas y empaquetarlas. Si pasáramos toda la cantidad de datos producidos en el muestreo generaría un alto consumo de ancho de banda, para tal efecto existen mecanismos que permiten tomar muestras a intervalos fijos por segundo, comprimirlas y empaquetarlas utilizando algoritmos avanzados que permiten ahorro en uso de ancho de banda sin sacrificar calidad de audio, a estos mecanismos se los conocen como CODECs.

Un CODEC convierte una señal de audio analógico a una señal digital y en el destino realiza el proceso inverso, de digital a analógico proveyendo procesos de compresión. La elección del CODEC a utilizar en una llamada telefónica es vital para su desenvolvimiento, se debe considerar aspectos como la calidad del sonido, ancho de banda requerido y requisitos de computación.

En el mercado existen diferentes CODECs libres y propietarios que difieren en características y costos, entre ellos tenemos: G.711 [10], G726 [11], G.728 [12], G.729 [13], iLBc como los más conocidos. En la tabla 1 se presenta las principales características de ciertos CODEC utilizados. En [14] podemos encontrar las recomendaciones publicadas por la ITU-T sobre los CODEC mencionados.

CODEC	Técnica de Codificación	Duración Segmento de Voz (ms)	Tamaño del Segmento Codificado (bits)	Tasa de Datos (bits/s)
G.711	PCM (*Pulse-Code Modulation*)	0,125	8	64
G.723.1	MP-MPLQ (*Multi-Pulse Maximum Likelihood Quantization*)	30	189	6,3
G.723.1	ACELP (*Algebraic-code-excited linear prediction*)	30	158	5,3
G.729	CELP (*Code-excited linear prediction*)	10	80	8

Tabla 1. Resumen de CODEC.

2.5.4 Protocolos de señalización IP

Los protocolos mayormente utilizados en el entorno de las telecomunicaciones actualmente son *Session Initiation Protocol* (*SIP*) e IAX, protocolos que se describen en este apartado: sus características y funcionalidad más relevantes.

SIP

Es un estándar desarrollado por IETF basado en las mejores prácticas registradas en SIPv1 y *Simple Conference Invitation Protocol* (*SCIP*) el cual alcanzó la denominación de estándar en 1999 y fue publicado bajo el RCF 2543 que describe las operaciones básicas de su núcleo.

Este protocolo de señalización fue desarrollado con el enfoque de administración de sesiones bajo el modelo cliente/servidor y se caracteriza por su base en el protocolo *Hypertext Transfer Protocol* (*HTTP*) y *Simple Mail Transfer Protocolo* (*SNMP*), disponibilidad para trabajar con *Transmission Control Protocol* (*TCP*) en el nivel de transporte del modelo *OSI*, utiliza lenguaje humano en el intercambio de mensajes [15], es escalable e interoperable con nuevos estándares, dispositivos y servicios, provee políticas de autenticación y cifrado [16].

Por sus características y versatilidad, tiene la funcionalidad para determinar la localización, disponibilidad y capacidad del usuario o dispositivo terminal, y la administración de sesiones junto con su operación. Para cumplir con estos objetivos existen diferentes entidades en sus procesos, entre ellos tenemos: *User Agent* (*UA*), SIP Location Server, SIP Redirect Server, SIP Proxy Server, SIP Back to Back User Agent y el SIP Gateway [5]. En la figura 2.2 se muestra el flujo de mensajes utilizado entre las diferentes entidades de SIP diferenciando un INVITE con un SIP Location Server y un SIP Proxy Server (entiéndase por flujo al tráfico ininterrumpido perteneciente a un usuario o aplicación). La elección del CODEC a utilizar en una llamada telefónica es vital para su desenvolvimiento.

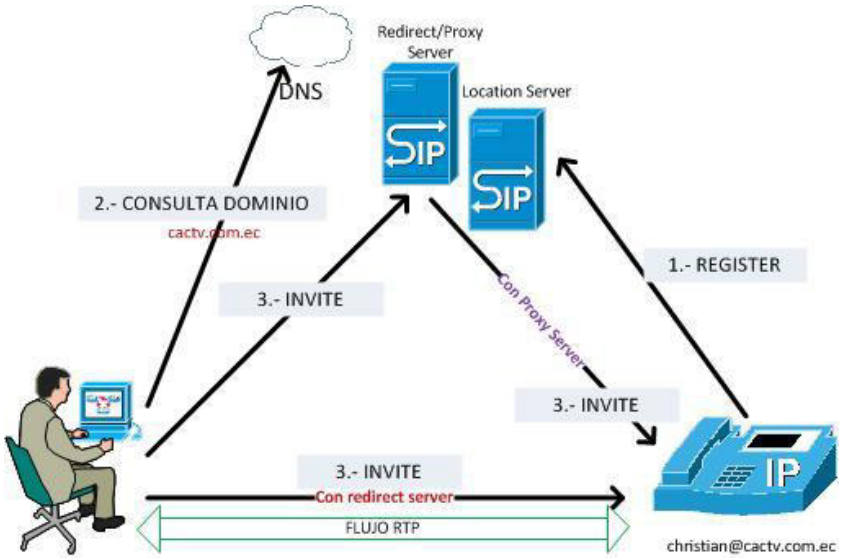


Figura 2.2. Flujo de llamada bajo esquema SIP.

Para el intercambio de mensajes entre los diferentes entidades de un escenario SIP, este protocolo ha establecido formatos de mensajes para cada proceso vital dentro de la administración de sus participantes y las comunicaciones recalcando que SIP únicamente se encarga de establecer las sesiones pero la descripción del mensaje de intercambio entre UA se lo puede realizar bajo cualquier estándar, como por ejemplo *Simple Description Protocol* (*SDP*). Los métodos o tipos de mensajes SIP son:

- *REGISTER*. El dispositivo final o UA se registra antes su servidor de localización.
- *INVITE*. Mensaje de petición de establecimiento de sesión en el cual se indica el tipo de flujo que serán intercambiados: voz, video, datos, entre otros.

- *ACK*. Acuse de recibo o confirmación de recepción de respuesta.
- *BYE*. Libera la sesión.
- *CANCEL*. Aborta el establecimiento de la llamada.
- *OPTIONS*. Pregunta parámetros para establecer flujo de llamada
- *INFO*. Mensaje informativo
- *REFER*. Transferencia de llamada
- *MESSAGE*. Mensaje instantáneo
- *SUBSCRIBY*/*NOTIFY.* Mensaje informativo.

Cada tipo de mensaje cumple una función determinada dentro del flujo de llamada, como se visualiza en la figura 2.3 Christian envía una invitación de llamada a Elisa, la cual requiere la comunicación entre dos servidores SIP Proxy de cada dominio al cual pertenecen Christian y Elisa, trabajo.com y hogar.com, una vez que llega el mensaje al destino, este le envía un tono de llamada (RINGING) y el origen confirma la comunicación (OK), una vez establecido los parámetros de la sesión empieza el flujo de audio y video, finalmente Elisa cuelga el teléfono (BYE) y Christian confirma la finalización de la llamada (OK).

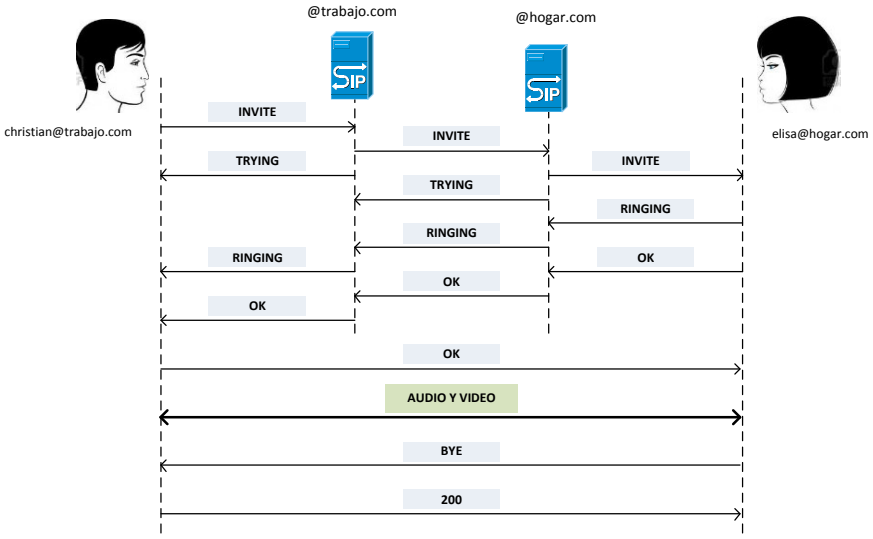


Figura 2.3. Flujo de mensajes en una llamada SIP.

IAX

Fue desarrollado por la comunidad de código abierto, *Asterisk Private Branch Exchange* (*PBX*) [17], para comunicaciones entre servidores Asterisk y con dispositivos terminales que hablen IAX.

A diferencia de SIP, IAX es un protocolo binario lo cual permite un consumo menor de ancho de banda que los basados en texto, esto permite que sea más robusto ante ataques de desbordamiento de buffer [18]. Utiliza únicamente el puerto UDP 4569 tanto en terminales como para servidores permitiendo el paso a través de firewalls y *Network Address Translation* (*NAT*) con simplicidad porque multiplexa la señalización y los datos a través del mismo puerto, a diferencia de SIP que usa dos puertos

diferenciados. Esto puede ser una fortaleza para IAX pero también una deficiencia porque lo hace vulnerable para ataques de denegación de servicio, bloqueando el tráfico en el único puerto utilizado [6]. A nivel de seguridad, IAX ofrece en el plano de autenticación bajo texto plano, *Message-Digest Algorithm 5* (*MD5*) y *Rivest, Shamir y Adleman* (*RSA*) [7].

Este protocolo maneja el *cuadro* como unidad de comunicación, el cual se encuentra en formato binario. Éste a su vez, maneja dos estructuras para su tráfico denominadas: *Full Frame* o *F Frame* y los *Meta Frame*, *Mini Frame* o *M Frame*. Los *Full Frame* son utilizados para llevar señalización y datos a la vez, y son conocidos como mensajes seguros porque el destinatario debe enviar una confirmación de recepción de mensaje en el momento de su arribo. El formato del *F Frame* se presenta en la figura 2.4 el cual describe cada campo y su longitud en Bytes: F, Source Call Number, R, Destination Call Number, Time-stamp, Oseqno, Iseqno, Frame Type, C, Subclass y Data. Cada campo tiene su código de representación en bits de los posibles valores.

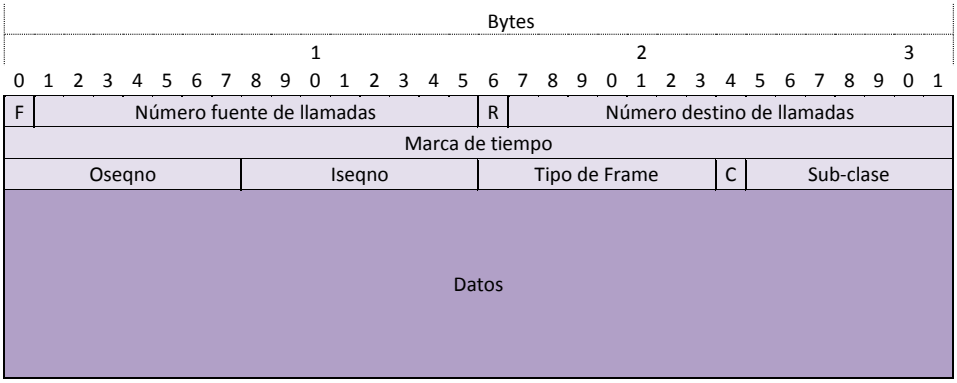


Figura 2.4. Formato de los *Full Frame*.

Los M Frames se utilizan únicamente para llevar flujo de datos y se conocen como mensajes no garantizados porque si un cuadro se pierde en el camino, éste no se retransmitiría. El formato del M Frame se presenta en la figura 2.5 el cual describe cada campo y su longitud en Bytes: F, Source Call Number, Time-stamp y Data. Cada campo tiene su código de representación en bits de los posibles valores.

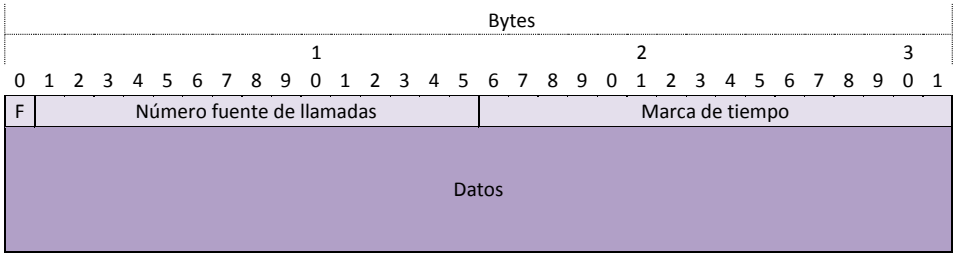


Figura 2.5. Formato de los *Meta Frames*.

En el flujo de una llamada bajo IAX están presentes los F Frame y M Frame, de acuerdo a la necesidad de información requerida y proceso en el que se encuentra, este flujo se muestra en la figura 2.6.

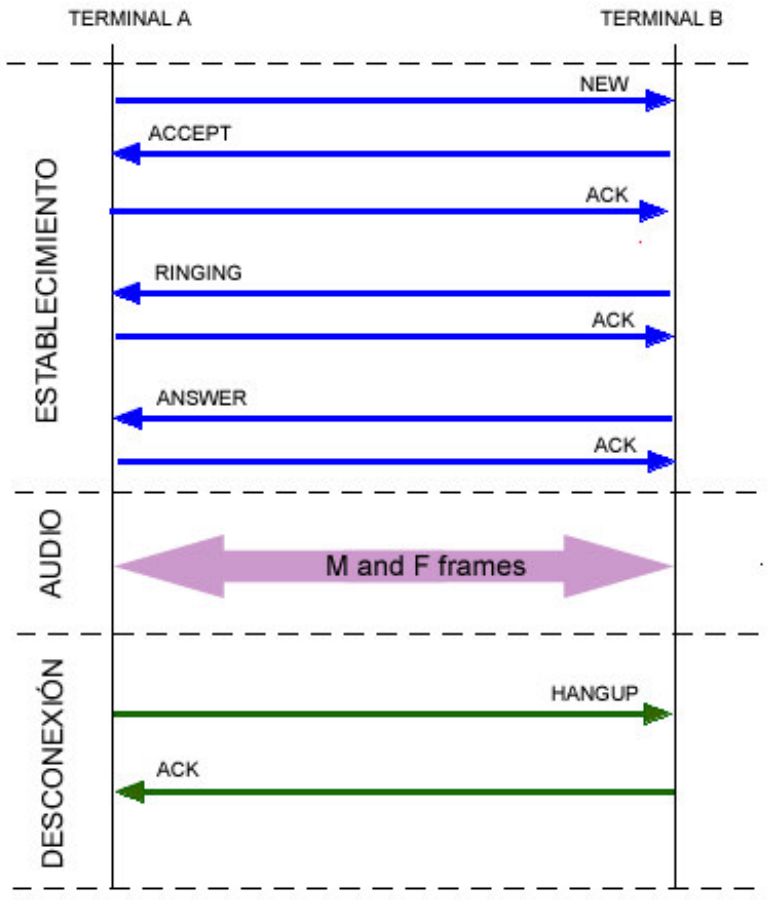


Figura 2.6. Flujo de mensajes en llamada IAX.

En definitiva, cada protocolo tiene sus fortalezas y debilidades pero debemos conocerlas para poder elegir uno u otro. En la Tabla 2 se muestra una comparación de las principales características mencionadas.

Característica	SIPv2	IAX2
Formato	Texto plano	Código binario
Consumo de ancho de banda		Menor ancho de banda que SIP
Traspaso de NAT y Firewall	Complejo	Simple
Estándar	Avalado por la IETF	En proceso de estandarización
Puertos a utilizar	5060 para señalización otros para flujo datos	Único puerto 4569
Tipo de datos	Cualquier información	Audio y Video

Tabla 2. Comparación de SIP e IAX.

CAPÍTULO 3

RTP: PROTOCOLO DE TRANSPORTE PARA APLICACIONES DE TIEMPO REAL

3.1 ANTECEDENTES

El desafío de los diseñadores de RTP fue construir un mecanismo robusto para enviar paquetes en tiempo real sobre una plataforma no confiable. Es de propósito generalista y permite gran flexibilidad. Estas decisiones están tomadas en base a varios puntos:

- La aplicación es responsable de hacer la toma de decisión sobre cómo tratar y comunicar la información multimedia (application level framing). Esto es debido a que el grado de heterogeneidad de las aplicaciones multimedia es muy elevado (suelen ser muy diferentes entre ellas). Por ello no es posible definir un mecanismo

concreto que se adapte a todas ellas de forma eficiente, porque solo la aplicación tiene el suficiente conocimiento de cómo hacer el tratamiento de sus datos y la manera de empaquetarlos para comunicarlos.

- La inteligencia en Internet suele estar en los terminales (end to end principle). Esto significa que en los encaminadores no se toman decisiones importantes sobre el re-envío de la información multimedia. Por ello, se debe definir un protocolo a nivel de transporte que permita ser usado por los extremos de la comunicación. Pero el protocolo de transporte debe aceptar "unidades de datos de aplicación" [7].
- Perfiles de comportamiento. Dado que pueden existir conjuntos de aplicaciones que sean parecidas, es posible adaptar la comunicación a un conjunto de ellas mediante los denominados perfiles de RTP.

3.2 DEFINICIÓN

RTP es un protocolo diseñado para el envío de datos en tiempo real sobre redes no confiables. Dentro del modelo OSI, no hay una ubicación específica del protocolo porque realiza funciones de la capa de transporte, sesiones y aplicaciones. Es un protocolo arbitrariamente incompleto, es decir, deja en potestad del diseñador de la aplicación mecanismos para sincronización, seguridad, reproducción de flujo, entre otros. Además, es eligible el tipo de carga útil y el formato a seguir.

3.3 RTP: PROTOCOLO DE TRANSPORTE PARA APLICACIÓN DE TIEMPO REAL

La transferencia de datos del RTP se encarga de administrar el envío de los datos en tiempo real. Incorpora número de secuencia para detectar paquetes perdidos, sellado de tiempo para recuperación, tipo de carga útil y flujo de datos. Paquetes de datos son enviados en la escala de milisegundos.

El formato de los paquetes se esquematiza en la figura 3.1.

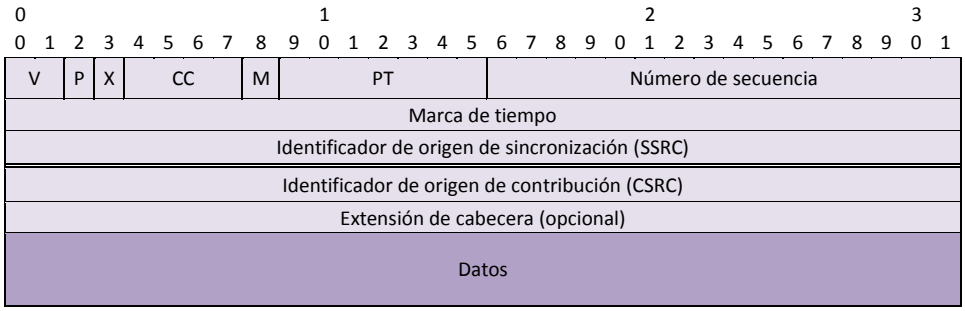


Figura 3.1. Formato del paquete RTP.

A continuación se detalla cada campo del paquete RTP:

- Versión (V). Indica la versión del RTP. Actualmente se trabaja con la versión 2 del protocolo descrita en el RFC3550.
- Relleno (P, del inglés Padding). Determina la existencia de relleno que no es parte de la data. Si el bit está activado indica que contiene uno o más byte de relleno que no es parte de la carga útil y la cantidad está registrada en el último byte del paquete.

- Extensión (X). Bandera activada a uno cuando existe una extensión del encabezado del paquete RTP.
- Contador de Fuente de Contribución (CC, del inglés CSRC Count). Campo que indica el número de CSCR.
- Marcador (M). Bandera utilizada para determinar eventos de interés, su definición completa depende del perfil a utilizar.
- Tipo de carga útil (PT, del inglés Payload Type). Identifica el tipo de tráfico que se transporta en el paquete de datos para que la aplicación destino conozca como manipularlo. El PT determina el formato de los datos y su posterior tratamiento y está ligado al perfil utilizado, por ejemplo, el RFC 3551 denominado "RTP Profile for Audio and Video Conferences with Minimal Control" aprobado en 2003 por la Internet Society, determina una tabla de asignaciones estáticas sobre el formato de la información y el numero en el campo PT del paquete RTP [6]. En la Tabla 3 se registran las asignaciones más utilizadas.

Tráfico	PT	Nombre de codificación	Tipo de carga útil	Nombre de codificación
Audio	0	PCMU	12	TPSO
Audio	1	1016	13	VSC
Audio	2	G.721	14	MPA
Audio	3	GSM	15	G.728
Audio	4	G.723	16	DVI4
Audio	5	DVI4 8kHZ)	17	DVI4
Audio	6	DVI4 16kHZ)	18	G.725
Audio	7	LPC	19	CN
Audio	8	PCMA	20	Sin asignar
Audio	9	G.722	21	Sin asignar
Audio	10	L16 estéreo	22	Sin asignar
Audio	11	L16 mono	23	Sin asignar
Video	24	HDCC	32	MPV
Video	25	CEL IB	33	MP2T
Video	26	JPEG	34	H.263
Video	27	CUSM	35-71	Sin asignar
Video	28	NV	72-76	Reservado
Video	29	PicW	77-95	Sin asignar
Video	30	CPV	96-127	Dinamico
Video	31	H.261		

Tabla 3. Formato del paquete RTP.

- Numero de secuencia (Sequence Number). Es un número positivo de 16 bits utilizado como identificador de paquetes, es decir, al destino le permite ordenarlos en forma ascendente en caso de arribo en desorden o detectar si existen paquetes perdidos en el camino. El número inicial es aleatorio en un flujo RTP y se incrementa en uno por cada paquete enviado y al llegar al máximo número permitido, 65535, se encera el contador y comienza una nueva asignación. Según [7] recomienda utilizar un número de secuencia de mayor tamaño, es decir, de 32 bits o más porque el proceso de enceramiento se produce en un tiempo reducido

y puede ocasionar pérdida de control de los paquetes recibidos. Por ejemplo, en VoIP el p-time o tiempo de empaquetamiento es de 20 milisegundos, entonces cada 21,85 minutos se reinicia el contador.

$$\mathbf{65535}\ \boldsymbol{paquetes}\ \boldsymbol{x}\frac{\mathbf{1}\boldsymbol{s}}{\mathbf{50}\boldsymbol{paquetes}}\boldsymbol{x}\frac{\mathbf{1}\boldsymbol{min}}{\mathbf{60}\boldsymbol{s}} = \mathbf{21,85}\boldsymbol{min}$$

- Marca de tiempo (T, del inglés Timestamp). Es un número de 32 bits que determina el instante en el cual fue muestreado el primer octeto de los datos en un paquete [6]. Este número incrementa dependiendo de la tasa de muestreo del tipo de dato determinado por el PT. Por ejemplo para RTP/AVP, típicamente el video se muestrea cada 90KHz y el audio cada 8KHz. Este valor es el que toma como referencia la aplicación destino para ordenar los paquetes para su posterior reproducción. Al igual que el número de secuencia, el timestamp empieza con un número aleatorio para evitar ataques sobre el flujo RTP, se incrementa gradualmente en pasos de uno y se reinicia cuando llega a su máximo valor. Las aplicaciones deben ser diseñadas para tolerar que el timestamp no inicia en cero y deben estar preparados para el enceramiento.
- Fuente de sincronización (SSRC, del inglés Syncronization Source). Número de 32 bits que identifica los participantes en una sesión RTP. Dicho número es elegido aleatoriamente y puede suceder que dos fuentes seleccionen el mismo número, la aplicación debe estar preparada con algoritmos de colisión y generación de nuevo SSRC. Si es cero, la fuente de sincronización es el generador de la carga útil. Por ejemplo, en una transmisión de video con múltiples cámaras, la sesión RTP

manejara diferentes SSCR por cada una para que el destino identifique la fuente de cada cuadro de video y pueda reproducirla en un punto específico.

- Fuente de contribución (CSRC, del inglés Contributing Sources). Es un campo que contiene una lista de SSCR y cada uno tiene un espacio de 32 bits para registrarse. El número total de fuentes está indicado por el campo CC que puede ser máximo hasta 16 fuentes. Se utilizar sobre todo en los Mixer, el cual es una aplicación intermedia que recibe flujos de datos de diferentes SSCR y genera un único flujo saliente y que tiene como CSRC todos los SSCR de las fuentes contribuyentes y modifica el SSCR al identificador del mezclador.
- Extensión del encabezado. Ofrece flexibilidad en el formato del paquete en caso de aplicaciones o experimentos que requieran otro tipo de información diferente a la registrada en los doce primeros octetos del encabezado fijo de un paquete RTP. Cuando la bandera X tiene valor igual a uno, indica que existe una extensión del encabezado que se encuentra después del CSRC y está compuesta por dos campos de 16 bits cada uno: tipo de campo y su longitud. Generalmente no es muy utilizada este encabezado, se registra que su uso es reemplazado por el desarrollo de un nuevo perfil en el cual se describan las nuevas necesidades y significado de los campos ligados al mismo.

3.4 RTCP: PROTOCOLO DE CONTROL DE LA COMUNICACIÓN DE VOZ DIGITAL EN TIEMPO REAL

RTCP es el encargado de enviar paquetes de control a los participante de una sesión RTP y tiene como principal objetivo informar sobre la calidad de servicio, además provee identificación de participantes y sincronización entre flujo de datos [6]. El envío de paquetes RTCP opera en la escala de segundos y por sí solo no transporta ningún dato.

Existen tres componentes básicos para comprender el mundo RTCP, estos son: el formato de los diferentes tipos de paquetes, las reglas de tiempo y la base de datos de los participantes en una sesión RTP.

TIPOS DE PAQUETES RTCP

Existen cinco tipos de paquetes RTP que cumplen un objetivo específico junto a su formato particular, entre ellos tenemos: Informe del receptor (RR, del inglés *Receiver Report*) y Informe del emisor (SR, del inglés *Sender Report*).

RR

Este paquete permite conocer la calidad de recepción de los datos y es emitido por todos los participantes de una sesión que solo reciben datos pero que no los emiten. Está compuesto por un encabezado y varios bloques de reportes con un máximo de 31 bloques, en caso de requerir enviar más reportes se debe generar múltiples paquetes RR. La retroalimentación recibida por estos paquetes pertenecientes al receptor es útil porque permite adaptar su tasa de transferencia, adicionalmente

facilita a los otros participantes de la sesión tener una visión más general de la red para determinar si un problema es local o global. El formato de paquetes RR esta detallado en la Figura 4.2 donde se indica todos los campos presentes y sus longitudes a nivel de byte.

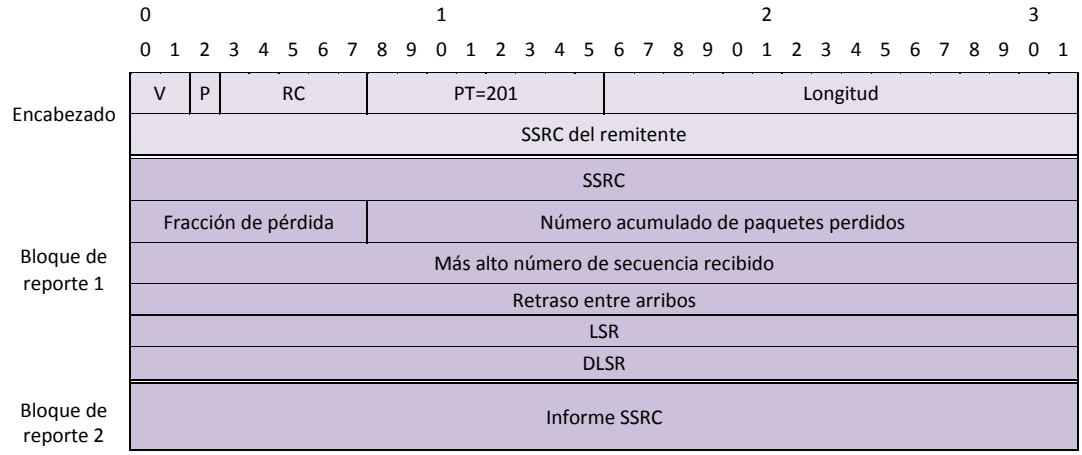


Figura 3.2. Formato del paquete RR.

Los campos V y P son similares a los paquetes RTP, versión del RTP y bandera de relleno del paquete respectivamente.

- Contador de informes de recepción (RC). Indica el número de bloques de reportes contenidos en el paquete.
- Tipo de paquete (PT). Campo definido por el valor 201 correspondiente a paquetes RR.
- Longitud. Indica la longitud total del paquete.

- SSRC del remitente. Describe el identificador SSRC de la fuente de sincronización del autor del paquete RR.

Los siguientes campos corresponden a los diferentes bloques de reportes definidos en el campo RC y su formato se repite en cada bloque de reporte que contiene el paquete.

- SSRC. Describe el identificador SSRC del participante a quien le pertenece la información descrita en el bloque de reporte.
- Fracción de pérdida. Es el cociente de la división del número de paquetes perdidos y los paquetes esperados contados a partir del último paquete RR o SR enviado, dicho valor es multiplicado por 256. Por ejemplo 1/4 x 256 es 64. Los paquetes que arriban tarde no son contados como perdidos y si el número de paquetes recibidos en más alto al esperado es debido a los paquetes duplicados el cual da como resultado una fracción igual a cero. La contabilidad de los paquetes se lo realiza en relación a los números de secuencia.
- Número acumulado de paquetes perdidos. Es la suma de todos los paquetes perdidos desde el inicio de la recepción de datos.
- Más alto número de secuencia recibido. Es el número de secuencia del último paquete recibido por el receptor.
- Retraso entre llegadas. Es una estimado de la varianza estadística del tiempo entre arribo de paquetes.

- Marca de tiempo de último SR (LSR, del inglés Last SR Timestamp). Si no se ha recibido anteriormente paquetes SR entonces LSR es cero.
- Retraso desde el último SR (DLSR, del inglés Delay Since Last SR). Si no se ha recibido anteriormente paquetes SR entonces LSR es cero.

SR

Este paquete permite notificar a los participantes receptores la sincronización de diferentes flujos emitidos (son enviados por los participantes que recientemente han enviado datos). En la figura 3.3 se detalla el formato del paquete con un encabezado similar a los anteriormente descritos.

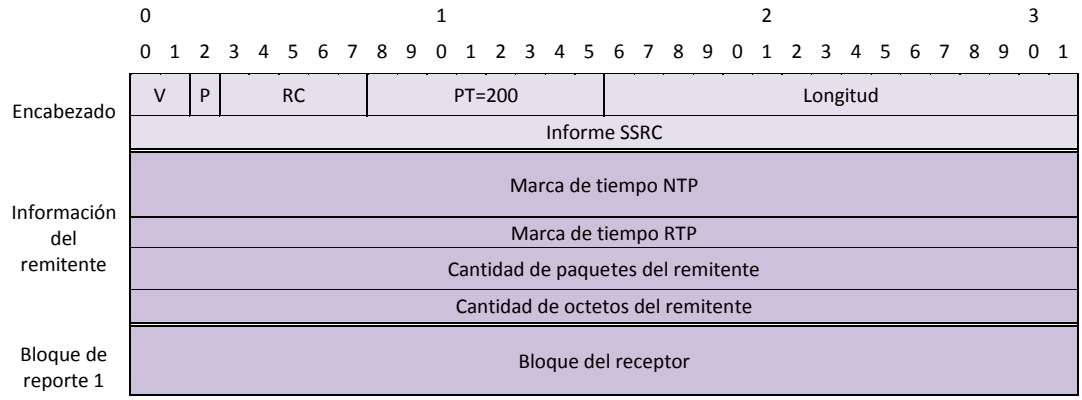


Figura 3.3. Formato del paquete SR.

- Marca de tiempo NTP. Es el instante de tiempo en el cual el paquete SR fue enviado en formato *Network Time Protocol* (*NTP*). Los 32 bits más significativos

de la fecha corresponden a la parte entera, segundos medidos a partir de las 0h del 1 de enero de 1900, y los 32 últimos bits corresponden a la parte fraccionaria.

- Marca de tiempo RTP. Es el mismo sello de tiempo que el indicado por NTP pero registrado en unidades del reloj del RTP.
- Cantidad de paquetes del remitente. Es el número de paquetes de datos enviados por el remitente desde el inicio de la sesión.
- Cantidad de octetos del remitente. Es el número de octetos enviados por el remitente dentro de los paquetes de datos desde el inicio de la sesión sin contar el encabezado ni el relleno.

A partir de este campo se repite los bloques de reporte similares a los revisados en el paquete RR porque un participante puede ser remitente y receptor a la vez.

El paquete de descripción de la fuente

Este paquete denominado en inglés *Source Description* (*SDES*) es útil para las aplicaciones porque a través de ellos se puede transmitir información suplementaria que generalmente es provista por el usuario final a través de una interfaz gráfica. Está compuesto por el encabezado básico de los paquetes RTCP y los bloques de descripción de fuentes. En la figura 3.4 se detalla el esquema del paquete SDES.

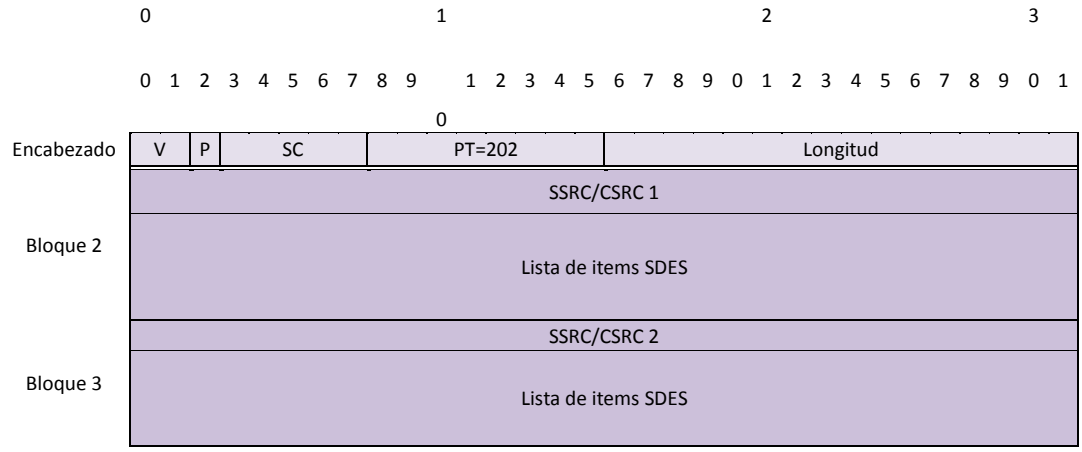


Figura 3.4. Formato del paquete SR.

Los campos V, P y longitud son similares a los paquetes RTP, versión del RTP, bandera de relleno del paquete y longitud del paquete.

- *Tipo de paquetes* (*PT*). Contiene la constante 202 que identifica a los paquetes SDES.
- *SSRC/CSRC*. Indica el SSRC de la fuente, la cual va a ser descrita en los campos siguientes. El número de SSRC o bloques a describir está dado por el campo *SC*.
- *Lista de items SDES*. Es una serie de parámetros que transmiten información ocasional del SSRC/CSRC, entre ellos tenemos los descritos en la tabla 4.

Parámetro	Tipo	Descripción
NULL	0	Fin de lista de parámetros
CNAME	1	Identificador único y permanente durante la sesión
NAME	2	Nombre real del usuario
EMAIL	3	Dirección de correo del usuario
PHONE	4	Número telefónico en formato internacional
LOC	5	Localización del participante
TOOL	6	Nombre y versión de la aplicación utilizada por el participante
NOTE	7	Mensajes temporales
PRIV	8	Para extensiones o aplicaciones en desarrollo

Tabla 4. Descripción de parámetros de paquetes SDES.

Estos items utilizan el formato descrito en la figura 3.5 en el cual consta un campo tipo de 8 bits, un campo longitud de 8 bit y un campo donde se registra el valor del parámetro en formato UTF-8 especificado en el RFC 2279 [19].

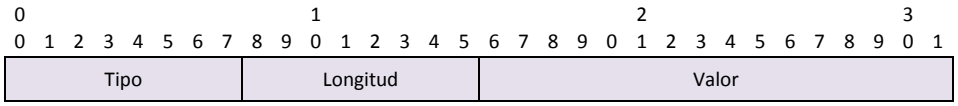


Figura 3.5. Formato de cada item en paquetes SDES.

Administración de miembros

Este paquete (denominado BYE) es generado para indicar que un participante ha dejado la sesión, esto puede ser por finalizar su actividad dentro del grupo o por

cambio del SSRC a causa de una colisión. En la figura 3.6 se describe el formato del paquete BYE con PT igual a 203. Luego del encabezado, se registran todos los SSRC de los participantes que abandonan la sesión, el número total de SSRC está dado por el campo RC. Al final del paquete, de forma opcional, se puede incluir un texto indicando la razón por la cual se produce el abandono.

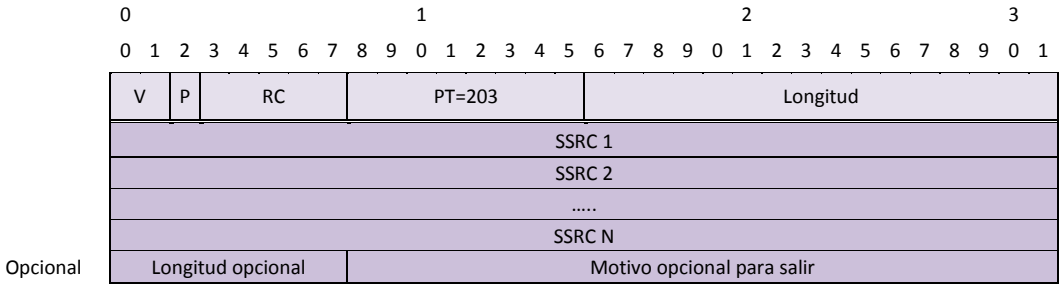


Figura 3.6. Formato de paquetes BYE.

Definidas por la aplicación

El último tipo de paquete definido en RTCP es APP, paquete creado propiamente por las aplicaciones. Generalmente lo utilizan aplicaciones en desarrollo (en verificación de su funcionamiento), para realizar el análisis del alcance de su diseño y proyectar su registro como tipo de paquete propio de RTP. Su formato está definido en la figura 3.7, en el cual los campos V, P y longitud son similares a los descritos anteriormente. La constante 204 se debe definir en el campo PT que identifica el tipo de paquete de la aplicación. El campo SSRC/CRCS es el identificativo de la fuente; NAME es el nombre elegido por la aplicación que debe ser único a nivel de todos los paquetes

que recibe dicha aplicación y la sección de datos es de libre diseño para que sea ajustable a las necesidades del desarrollador.

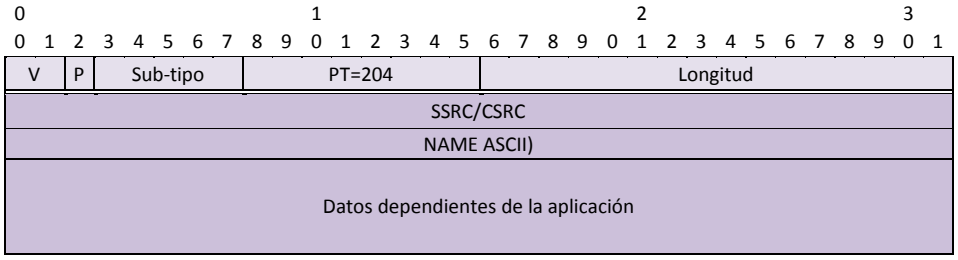


Figura 3.7. Formato de paquete APP.

3.5 TRADUCTORES Y MEZCLADORES

Los traductores y mezcladores son sistemas intermedios dentro un flujo RTP que permiten el establecimiento de sesiones entre diferentes escenarios, es decir paquetes: producidos en IPv4 para transportarlos a IPv6, comunicados mediante difusión múltiple controlada (multicast) a difusión particularizada (unicast), traspasados desde IP a ATM... Para instalar un mezclador y/o un traductor dentro de la red se debe asegurar la interconexión de dos tipos de protocolo diferentes, red o puerto para evitar la creación de lazos y adicionalmente evitar la participación simultánea de dos o más sistemas intermedios.

Un traductor, mayormente conocido como *translator*, es una componente invisible a los sistemas finales que operan sobre los datos en una sesión RTP pero sin modificar el SSRC de la fuente, es decir, ingresa un flujo de datos y sale el mismo flujo con el

SSRC de la fuente original. Un traductor no participa en las sesiones pero permite descifrar diferentes formatos, protocolos o eliminan el cifrado de los datos sin cambiar su esencia. A nivel de RTCP, un traductor no generan este tipo de paquetes, solamente re-envían los paquetes recibidos de sus fuentes originales y si es necesario, modifican algún campo interno del paquete que es producido por el cambio generado en el paquete original, es decir, si se cambia la codificación del paquete debe actualizar el nuevo dato en el campo contador de byte.

Un mezclador, o mix es un participante en una sesión RTP que combina los paquetes de entrada y producen un solo flujo de salida entre dos escenarios diferentes. Estas cajas sellan el nuevo paquete con su propio SSRC y guarda en el campo CSRC, el listado de los SSRC de las fuentes originales de los paquetes recibidos. A nivel de paquetes RTCP, un mezclador genera nuevos paquetes RR, SR y BYE. El paquete SDES generalmente se transmite sin cambios. En la figura 3.8 se revisa el esquema de funcionamiento y su injerencia en los paquetes RTP.

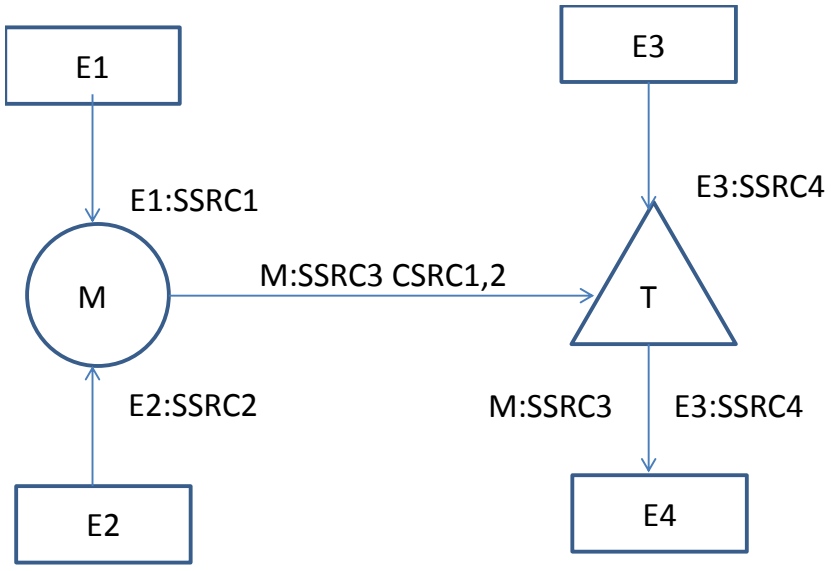


Figura 3.8. Esquema de funcionalidad del traductor y mezclador.

3.6 ASPECTOS BÁSICOS DE SEGURIDAD

En las especificaciones técnicas de RTP se define un método simple para proveer confidencialidad al paquete, es decir, únicamente sea legible a su destinatario, mas no a terceros. La confidencialidad se la provee a través del cifrado de la carga útil mediante el uso del algoritmo de cifrado *Data Encryption Standard* (*DES*) en modo de encadenamiento de bloques. Dicho algoritmo en la actualidad no se considera seguro porque puede romperse la clave en menos de 24 horas. Características de autenticación, no están disponibles en el RFC.

El perfil *Secure Real-Time Transport Protocolo* (*SRTP*) es un estándar definido en el RFC 3711 [20], el cual fue desarrollado por expertos de las empresas Cisco y Ericsson y posteriormente publicado por la IETF en el año 2004. Este protocolo nos permite agregarle al RTP y RTCP características de seguridad como son la autenticación de mensajes, confidencialidad de información y protección contra reenvío de paquetes al establecer llamadas VoIP.

El formato del paquete SRTP está definido en la figura 3.9 en el cual se visualizan los campos agregados a un paquete RTP que son *Master Key Identifier* (*MKI*) y la etiqueta de autenticación. El campo Payload es la porción del paquete SRTP que se encuentra encriptado en conjunto con sus referencias de relleno y contador. El algoritmo de encriptación utilizado es *Advanced Encryption Standard* (*AES*) en modo F8 [21]; la cabecera no se cifra.

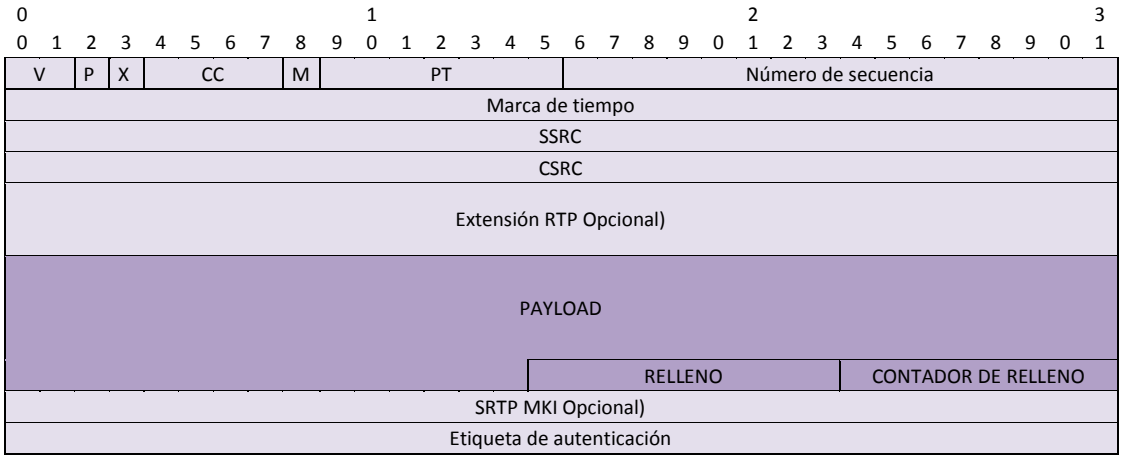


Figura 3.9. Esquema de paquete SRTP.

SRTP provee mecanismos de autenticación que comprende la integridad de los datos y la autenticación de la fuente. Para la integridad de los datos se utiliza el campo etiqueta de autenticación y se utiliza el algoritmo *Hash Message Authentication Code* (*HMAC*) [22] con la función hash SHA-1 [23].

Otro método de proveer seguridad es mediante el uso de IPSec [24], el cual provee seguridad a nivel de la capa IP y no es utilizado únicamente en envío de paquetes RTP sino en todo tipo de tráfico. Utiliza un conjunto de protocolos y mecanismos que permitan un túnel seguro extremo a extremo para el envío de mensajes.

CAPÍTULO 4

3. PARAMETROS DE CALIDAD DE LA VOZ

4.1 CONCEPTOS BÁSICOS DE QOS

La Recomendación UIT-T E.800 [25] aprobada en el 2008, define “calidad de servicio” como “*la totalidad de las características de un servicio de telecomunicaciones que determinan su capacidad para satisfacer las necesidades explícitas e implícitas del usuario del servicio*”, e identifica los componentes tradicionales extremo a extremo de una red, es decir la QoS depende de cada uno de los componentes y sus parámetros parciales de calidad dentro de la red, estos son retraso, paquetes perdidos, jitter, tiempos de procesamiento y propagación que afectan directamente a la QoS (figura 4.1).

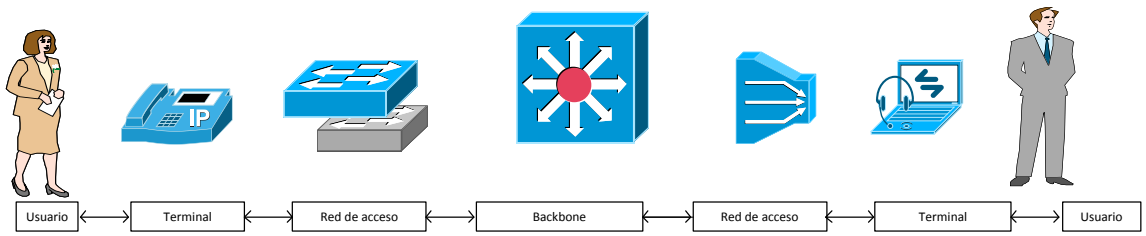


Figura 4.1. Esquema de contribuciones a la QoS extremo a extremo.

Como se revisó en la sección 2.4, los paquetes perdidos, el retraso y jitter son factores que influyen en los servicios de voz y determinan su calidad de recepción en el destino. Los paquetes se pierden perdidos por diferentes factores, entre ellos porque viajan sobre UDP, protocolo no orientado a conexión que en caso de pérdida de un paquete no realiza la retransmisión del mismo a diferencia de TCP, además por ser una comunicación en tiempo real si un paquete llega con mucho retraso, el receptor lo descarta. Otra consideración muy importante es la plataforma sobre la cual trabaja VoIP, esta es Internet. IP realiza el mejor esfuerzo y no garantiza la entrega del paquete a su destino, cada paquete puede seguir un camino diferente generándose retraso en la llegada y/o jitter en el flujo de paquetes.

En escenarios reales, cuando en medio de una conversación se escucha la voz robotizada es por presencia de jitter en el envío de paquetes, lo máximo permitido es 50 ms. A diferencia de cuando se escucha la voz entrecortada, se debe al retraso y/o pérdidas en el arribo de los paquetes, mientras mayor es el porcentaje de pérdida y

mayor el tiempo de retraso, la conversación se vuelve menos fluida e incomprensible, lo máximo permitido a nivel de latencia es 200 ms y mínimo 2% de paquetes perdidos.

Existen diferentes mecanismos para poder minimizar el impacto o la presencia de dichos parámetros que degradan la voz, entre ellos tenemos servicios integrados (*IntServ*) y servicios diferenciados (*DiffServ*).

Servicios integrados

Modelo que implementa reserva de recursos de extremo a extremo brindando prioridad por flujo (cada paquete de un flujo requiere el mismo tratamiento de calidad). IntServ define 3 tipos de clases de servicios, estos son servicios garantizados, de carga controlada o de mejor esfuerzo. La primera clase asegura una tasa de transferencia sin paquetes perdidos por encolamiento, clase idónea para aplicaciones en tiempo real. La segunda clase o de carga controlada es similar a una red basada en el mejor esfuerzo pero sin sobrecarga y finalmente, la clase de mejor esfuerzo no garantiza ningún servicio.

El protocolo más representativo de esta modalidad es *ReServation Protocol* (*RSVP*), encargado de gestionar o administrar la reserva de recursos de extremo a extremo. Una vez definida la clase del flujo a enviar y el ancho de banda necesario, el origen envía mensajes de señalización al próximo salto solicitando los recursos necesarios para comenzar el envío de datos, en caso de que el encaminador siguiente no cuente

con dichos recursos, niega la petición. RSVP se ayuda de los protocolos de encaminamiento implementados en la red para conocer diferentes caminos disponibles para llegar a su destino final, en caso de una petición denegada redefine el camino para que al final de la negociación se establezca un camino concreto que garantice el ancho de banda solicitado y empiece del flujo.

Servicios diferenciados

Modelo que implementa la priorización de paquetes dependiendo del tipo de servicio que desea brindar. Cada paquete es marcado con una prioridad de acuerdo a la sensibilidad a parámetros como el porcentaje de paquetes perdidos, latencia y jitter, esto define su jerarquía de acceso a los recursos de la red. En el formato del paquete IP definido en el RFC 791 se encuentra el campo de 8 bits denominado Tipo de Servicio (*ToS*) visible en la figura 4.1, este campo se subdivide en dos partes: Punto de código de servicios diferenciados (*DSCP*, del inglés *Differentiated Services Code Point*) y Actualmente en Desuso (*CU*, del inglés *Currently Unused*).

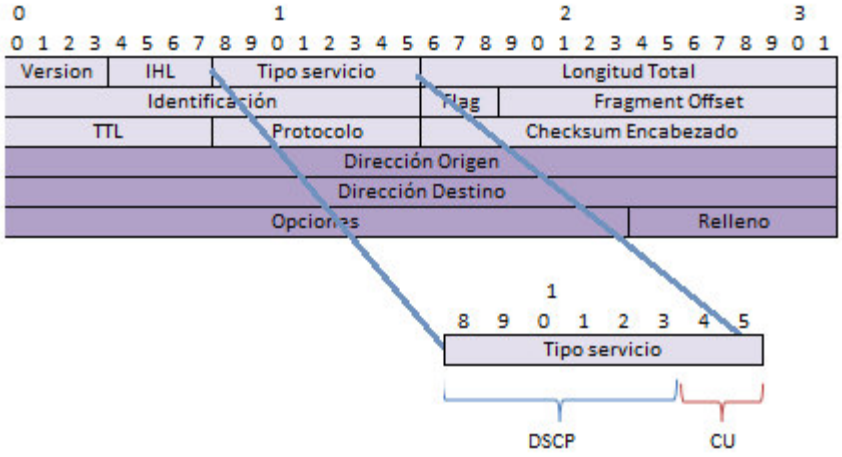


Figura 4.2. Campo tipo de servicio dentro del formato IP.

Los posibles valores de DSCP utilizados actualmente, de acuerdo al RFC 2474, se registran en la tabla 6 de acuerdo a la prioridad que se debe de dar a cierto flujo de datos. Este modelo es compatible con la definición de *ToS* anterior en el cual los 3 bits más significativos determinaban la precedencia del flujo.

BIT DS	CODIGO BINARIO	DESCRIPCION DE CAMPO
7-6-5	111	Control de la red
	110	Control de red interna
	101	Crítico/ECP
	100	Señal continua
	011	Destello
	010	Inmediata
	001	Prioridad
	000	Rutina
4	1	Retraso bajo
	0	Retraso Normal
3	1	Rendimiento alto
	0	Rendimiento normal
2	1	Fiabilidad alta
	0	Fiabilidad normal
1	1	Costo monetario mínimo
0	0	Cero

Tabla 5. Categorías de tipo de tráfico para DS.

4.2 CONCEPTOS BÁSICOS DE QOE

Según la recomendación UIT-T P.10/G.100 [26], define QoE como "*la aceptabilidad general de una aplicación o de un servicio, conforme a la percepción subjetiva del usuario extremo*", que incluye todos los componentes de la red graficados en la figura 4.1 pero estos son transparentes para el usuario evaluador.

QoE es un término ampliamente usado en los últimos años, los proveedores de servicios de telecomunicaciones no solo desean conocer el nivel de rendimiento de sus infraestructuras, equipamientos o métodos y procedimientos del negocio, sino también cómo percibe el usuario final el servicio que está brindando, si el rendimiento

de su red y la inversión hecha es a la medida de las exigencias del usuario y es suficiente para poder tener una cartera de clientes holgadas con buenas referencias. Con estos antecedentes, es prioritario para los ISP conocer el grado de percepción de sus clientes sobre los servicios de datos, voz y/o video en escalas tan simples como excelente, buena, regular o mediocre, entre otras. Estas mediciones generalmente se lo realizan siguiendo estándares que garanticen un resultado lo más apegado a la realidad como el P.800 (método de determinación subjetiva de la calidad de transmisión); pero el cumplimiento de dichos parámetros significaría una gran inversión económica y tiempo para cumplir las exigencias, por tal motivo se desarrollan constantemente diferentes estudios para determinar la QoE basados en modelos o leyes físicas, por ejemplo en la ley de Weber-Fechner, la cual establece una relación entre la magnitud del estímulo físico y cómo es percibido por el ser humano [21] o la generación de diferentes modelos del funcionamiento del cuerpo humano y sus sentidos. En la figura 4.1 se detalla los diferentes métodos desarrollados para medir la QoE, métodos que han dado lugar a nuevas formas de concepción y medición descubriendo las carencias y fortalezas de cada uno para proponer una mejora en su desarrollo.

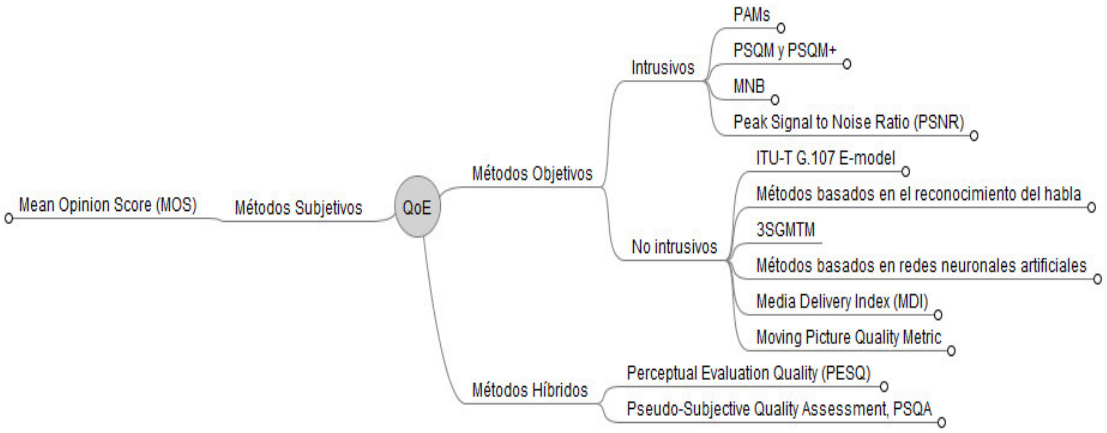


Figura 4.3. Métodos para obtener QoE.

Los métodos existentes para el cálculo de la QoE pueden ser métodos objetivos y subjetivos, dependiendo del esquema utilizado para realizar la medición.

Métodos Objetivos

Estos métodos utilizan esquemas matemáticos y/o algoritmos que se ayudan de modelos de percepción, audición, reconocimiento cognitivo que emulan funciones propias del ser humano a fin de emitir un resultado del nivel de QoE percibida.

Medida perceptual de la calidad vocal

La Medida perceptual de la calidad vocal (*PSQM*, del inglés *Perceptual Speech Quality Measure*), esta descrito en la recomendación P.861 [27] que fue aprobada en 1996, en el cual se describe el método para evaluar la calidad de un códec en banda telefónica (300 - 400Hz) en pruebas de escucha asumiendo que la señal de voz es libre de ruido y no existe degradación en el canal. Para esto, simula la percepción del sonido registrado mentalmente por los humanos a través de la conversión de señales físicas al dominio psicofísico a través de tres operaciones: correspondencia tiempo-frecuencia, transposición de la frecuencia y transposición de la intensidad que representa la interpretación mental del estímulo del sonido causado en el cerebro humano. Este método genera un valor entre 0 y 6.5 que indica el nivel de degradación de la calidad de voz al realizar el proceso de codificación [28], mientras más cercano el valor se encuentre de 6.5, la calidad del audio se deteriora considerablemente. No existe una fórmula pre-establecida para transformar dicho valor a la escala MOS.

Posteriormente, este algoritmo fue reemplazado por el *Perceptual Evaluation of Evaluation of Speech Quality* (*PESQ*) descrito en la recomendación P.862 [29] porque era muy limitado en su aplicación, únicamente en ambientes de laboratorios y sin aplicación en nuevas redes de datos, además solo podía utilizarse en evaluación de CODEC y esta es solo una pequeña parte en toda la cadena de transmisión de audio de extremo a extremo.

Sistema de medición mediante análisis por percepción

Método objetivo denominado *Perceptual Analysis Measurement System* (*PAMs*) desarrollado por British Telecom en 1998, tomando como referencia a PSQM pero corrigiendo ciertos procesos que no permitían una alta correlación con pruebas subjetivas, esto es, filtrado lineal para interfaces analógicas, soporte para perdida de paquetes y retraso variable presentes en redes IP [30]. PAMs, al igual que PSQM, compara la señal original con la señal degradada y registra las distorsiones presentes, analiza no solo la cantidad sino la distribución de la distorsión audible lo que permite hacer predicciones de la calidad más precisas.

E-Model

Definido bajo la recomendación ITU-T G.107 en el cual se describe un modelo computacional que utiliza diversos parámetros como es el retraso de transmisión, ruido en el circuito, distorsión de cuantización, nivel vocal, entre otros, para calcular un factor de evaluación de la transmisión denominado R. El valor del parámetro R está dentro del rango de 0 a 100 donde 0 es la peor calidad y 100 la mejor calidad percibida por el usuario final. E-model está dirigida para diseñadores o arquitectos de red que definen un nivel de calidad de voz y tratan de esquematizan la red bajo estos parámetros, los lineamientos para la planificación del diseño de una red basada en este modelo lo encontramos en [31].

La fórmula general del E-Model es:

$$R = Ro - Is - Id - Ie + A \quad (4.1)$$

Donde Ro es la relación básica señal ruido de la conexión, al cual se resta valores como el deterioro de la señal de voz en la transmisión por factores simultáneos (Is), retrasos en la transmisión (Id), factor de deterioro del equipo más la robustez del códec utilizado frente a la pérdida de paquetes (Ie) [32] y finalmente el valor A el cual depende del diseñador. Existen propuestas de simplificar la formula general como lo describen [33] [34].

En [35] se registra tabuladas las categorías de la calidad de transmisión vocal en términos de "satisfacción del usuario", dicho valor puede ser traducido a una escala MOS para determinar el nivel de calidad subjetiva del escenario a evaluar mediante la aplicación de una simple formula detallada en el anexo B de [36]. Estos valores están detallados en la tabla 5.

Valor R	Nivel de Satisfacción	MOS
90-100	Muy satisfecho	4.3+
80-90	Satisfecho	4.0-4.3
70-80	Algunos usuarios insatisfechos	3.6-4.0
60-70	Muchos usuarios insatisfechos	3.1-3.6
50-60	Casi todos los usuarios insatisfechos	2.6-3.1
0-50	No recomendado	1.0-2.6

Tabla 6. Correspondencia entre valor R y MOS.

Es un modelo bastante interesante y muy complejo de aplicar se debe tomar en cuenta que está bajo estudios de funcionamiento y pruebas rigurosas porque existen

muchos parámetros o comportamientos que no son contemplados en la ecuación o son muy generalizados de tal forma que afectaría el resultado final como lo describe en el Anexo A de la recomendación.

PESQ

Método objetivo para la evaluación de la calidad vocal por percepción (*PESQ*, del inglés *Perceptual Evaluation of Evaluation of Speech Quality*) descrito en la recomendación IUT-T P.862 como un algoritmo que permite determinar la calidad subjetiva de microteléfonos de 3.1kHz (banda estrecha) y CODEC vocales de banda estrecha, además de mediciones de extremo a extremo de redes telefónicas y aplicaciones VoIP [29].

Este método es considerado intrusivo porque compara la señal original X(t) con una señal degradada Y(t), que es la misma señal original pero luego de ser trasmitida por el dispositivo o sistema sometido a prueba, como se muestra en la figura 4.2. Estas dos señales ingresan al algoritmo PESQ que calcula los retrasos entre ambas, luego realiza su alineación mediante el resultado del cálculo de los retrasos y un modelo por percepción, en el cual, radica el proceso relevante de este método, crear un modelo basado en las audiofrecuencias y la sonoridad perceptibles para el oído humano.

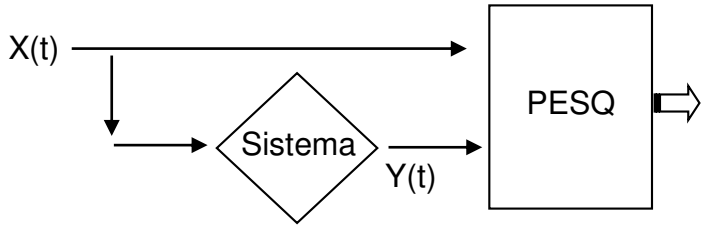


Figura 4.3. Esquema general de PESQ.

Los resultados de PESQ están dentro de la escala de -0,5 a 4,5 y se considera que tienen una buena correlación con los resultados de pruebas subjetivas para aplicaciones VoIP [37], pero se debe aplicar la función de correspondencia para traducir a una escala MOS [38].

Esta recomendación es ampliamente utilizada como referencia para nuevas investigaciones [39] [40] pero no es utilizada para mediciones de calidad de voz en tiempo real debido a que necesita una señal de referencia, aunque existen propuestas, por la alta correlación de PESQ con pruebas subjetivas, de suprimir la señal de referencia [37] o también generar herramientas de monitoreo de la red [41].

Métodos Subjetivos

Estos métodos están normalizados por la *Unión Internacional de las Telecomunicaciones* (*UIT*) en las recomendaciones P.800 para audio y BT.500 para

video. Esta metodología está ligada con los sentimientos, vivencias o intereses del sujeto evaluador, es decir, se necesita la intervención directa del usuario final y su calificación sobre la experiencia en el uso de una tecnología, sistema, dispositivo o servicio final. Entre los métodos subjetivos tenemos MOS.

Nota Media de Opinión

Media Opinión Score (*MOS*) está definido como un método subjetivo por excelencia que mide la calidad de funcionamiento del sistema de transmisión telefónica y esta descrito en la recomendación P.800 la cual fue emitida por la Sector de la Normalización de las Telecomunicaciones de la UIT en el año de 1996 [42], aunque existen estudios que el valor de MOS no es suficiente, se debe tomar en cuenta la desviación estándar [43]. En dicha recomendación se detalla todo lo relaciones con las instalaciones físicas para las pruebas, diseño del experimento, conversaciones y procedimientos a seguir para tener un escenario controlado donde varios sujetos puedan escuchar audios específicos y luego determinan una puntuación numérica dependiente del grado de satisfacción de la calidad percibida. Existen dos variantes para estas mediciones: pruebas de opinión de conversación o pruebas de opinión de solo escucha descritos en el anexo A y B de la recomendación respectivamente.

Las pruebas de conversación se realizan entre dos participantes que no hayan realizado o contribuido en métodos de evaluación de calidad o codificación de

señales ni participado en una prueba similar como mínimo dentro de los 6 meses anteriores y es indistinto el género de cada participante, esto debe ser aleatorio. La conversación debe ser lo más natural y comprensible posible y debe tener un principio y fin. Los resultados de la prueba pueden ser registrados en diferentes escalas de opinión que puede ser por conversación (escala: Excelente, Buena, Regular, Mediocre y Mala) o por dificultad (respuesta binaria de presencia de dificultad SI/NO) entre las más populares.

La prueba de solo escucha cuenta con el método de determinación de índices de categoría absoluta (ACR, del inglés Absolute Category Rating) que consta de dos partes diferenciadas: grabación de audio y el proceso de escucha. La grabación del audio debe ser en un ambiente controlado a nivel de ruido (menor de 30dBA), volumen (entre 30-120m3), reverberación (menor a 500ms) y los participantes deben utilizar frases sencillas sin mostrar diferencias en la entonación vocal tanto para hombre como para mujeres. Se utiliza un nivel del ruido menor de 30dBa dado que permite un correcto entendimiento de las palabras y frases emitidas por las voces a intervenir en el audio. Los participantes deben oír el material de audio y asignar un puntaje dependiendo de la escala de opinión utilizada: escala de calidad de escucha (Excelente, Buena, Regular, Mediocre y Mala), escalada de esfuerzo de escucha y escala de sonoridad preferida.

La unidad de medida dBA utilizada para determinar el nivel de ruido ambiental permitido en las pruebas es una ponderación frecuencial empleada en mediciones

acústicas dentro de los rangos de frecuencias audibles de 20Hz (sonidos graves) hasta los 20KHz (sonidos agudos) que una persona normal puede percibir. En base al Nivel de Presión Sonora (NPS), podemos conocer la percepción del oído humano al ruido ambiental a través de la ponderación A porque nos permite aproximar la sensación subjetiva del nivel sonoro dado que trabaja a niveles bajos, útil para la medición de ruidos de fondo. También existen las ponderaciones B, C, D, Z pero no son aplicables dado que trabajan a niveles más altos de presión sonora.

Los métodos subjetivos son considerados muy costosos porque involucra un ambiente controlado, cuartos aislados acústicamente, la participación de muchos usuarios como sujetos de prueba, reproducción de material de audio con parámetros fijos… que determinan una alta inversión o sitios especializados como laboratorios o universidades con apoyo a la investigación. Por tal motivo la medición de QoE se hace preferentemente mediante los métodos objetivos que son de menor costo.

4.3 RELACIÓN ENTRE QOS Y QOE

Como se ha revisado en los apartados anteriores existe un estrecho vínculo entre QoE y QoS, parámetros de medición sobre la calidad que experimenta el cliente final. Podemos decir que un alto puntaje en la evaluación de QoE en un servicio, la plataforma de red es considerablemente eficiente pero no necesariamente si los parámetros de QoS son buenos genera una excelente experiencia a los clientes; en

términos de lógica matemática podríamos decir que existe una relación condicional entre QoE y QoS. Esto es fácilmente identificable si consideramos que la QoS es un factor incluido dentro de la evaluación realizada por la QoE, esto es, porque interviene un actor adicional que agrega dentro del proceso de evaluación, cualidades subjetivas propias de ser humano (figura 4.3).

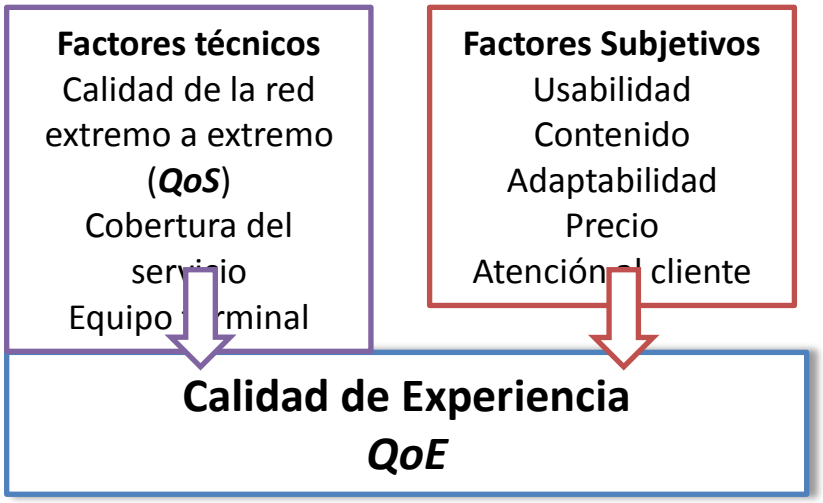


Figura 4.4. Relación entre QoS y QoE.

Existen estudios que permiten visualizar claramente las diferentes concepciones de la relación entre estos dos parámetros a través del tiempo, algoritmos que determinan la relación existente entre la QoE y el QoS para modelar el comportamiento del QoE basado en la capacidad de rendimiento de la red y la disponibilidad de recursos [44]

[45] o generando diferente tipos de relaciones como exponencial a través de una formula basado en parámetros de QoS [46].

CAPÍTULO 5

4. PRUEBAS DE CAMPO

5.1 COMPONENTES DEL ESCENARIO DE PRUEBA

El entorno de pruebas del presente proyecto comprende diversas partes que detallamos en la figura 5.1. El locutor A y locutor B están conectados a Internet con el mismo aplicativo VoIP a evaluar. El Locutor A envía una conversación pre-grabada y el locutor B se limita a escuchar y evaluar la calidad del audio. Finalmente el locutor B califica la calidad de la transmisión de audio percibida a través de la encuesta descrita en el Anexo B.

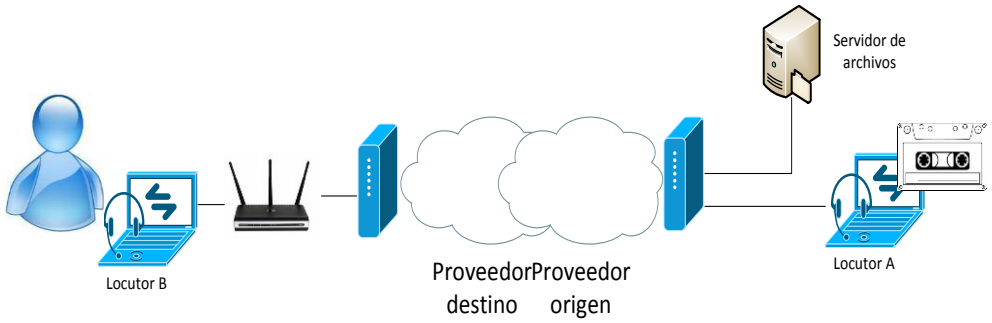


Figura 5.1. Escenario de pruebas.

5.1.1 Software

Se propuso en el anteproyecto evaluar las aplicaciones MSN Messenger, Skype y Gtalk, pero debido a que Microsoft adquirió Skype y retiro del mercado a MSN Messenger, obligando a sus usuarios el uso de Skype y el aplicativo Gtalk fue reemplazado por Google Hangouts, únicamente se evaluaran las dos aplicaciones mencionadas: Skype y Hangout.

A continuación mencionamos las características más relevantes de este software de VoIP:

- Skype. Aplicación privada de mensajería instantánea bajo tecnología P2P desarrollado por Kazaa en el 2003. El código y el protocolo utilizado son propietarios pero la aplicación se puede descargar de forma gratuita desde su página oficial permitiendo conexión entre todos sus usuarios. Este software permite chat, videoconferencia, compartición de archivos; además se puede llamar

desde y hacia teléfonos convencionales a través de SkypeOut y SkypeIn respectivamente. En el 2011, Microsoft adquirió Skype y anuncio la salida de MSN Messenger en el 2013 a los usuarios de Windows, induciéndolos a instalar Skype. Según estudios realizados sobre la arquitectura de red utilizada por Skype [47] [48] [49] [50], se identifican tres tipos de participantes: los terminales, los supernodos y el servidor de autenticación. Los terminales son las computadores o dispositivos que contiene la aplicación Skype localmente instalada, los supernodos son servidores públicos robustos distribuidos geográficamente, con alta capacidad de procesamiento y memoria; finalmente los servidores de autenticación son una granja de equipos manejados directamente por Skype donde se almacena y administra los usuarios de Skype y su autenticación.

- Google Hangouts. Aplicación web basada en una topología cliente-servidor [47] que es instalado sobre un navegador como un plugin o complemento. Fue desarrollado por Google Inc. en reemplazo de Google Talk, Google+ Messenger y Google+ Hangouts, plataforma que fue lanzada a inicios del 2013. Es compatible con todo tipo de computadores, teléfonos móviles o tabletas Android, iPhone, iPad y iPod touch dando gran cobertura de sus servicios de mensajería instantánea en pareja o grupo hasta 10 integrantes, y de forma similar para video llamadas en grupo en donde se puede compartir archivos, visualizar videos de forma conjunta, además se integra a nivel de Web con Gmail. Hangouts no utiliza el Protocolo extensible de mensajería y comunicación de presencia (XMPP, del inglés

Extensible Messaging and Presence Protocol) como lo hacía GTalk sino fue desarrollado un nuevo protocolo propietario del cual se conoce muy poco. Hangout es muy apropiado para ser usado por usuarios de la red social Google+.

Para hacer las medidas de tráfico usamos Wireshark, una aplicación que captura todos los paquetes que circulan por la red y los presenta en toda su estructura, tal como se describe en el modelo OSI. Plataforma muy recomendada para uso educativo y a nivel profesional porque permite diferenciar los protocolos presentes en la red, el envío/recepción de paquetes desde su origen y destino, los diferentes campos y su estructura.

5.1.2 Proveedores de Internet

Para la presente prueba se escogieron dos proveedores de Internet populares en la provincia del Guayas con gran participación en el Mercado masivo como son: la Corporación Nacional de Telecomunicaciones Empresa Pública (CNT E.P.) y Tv Cable.

En el extremo transmisor tenemos como proveedor de Internet a TV Cable, empresa privada fundada en 1986 la cual está formada por las empresas SATNET, SURATEL, SETEL y TV CABLE. Dispone de redes de acceso en las principales ciudades del Ecuador y de una cartera de servicios que está compuesta por Internet, Datos y

telefonía fija. Su principal mercado yace en el sector masivo con su producto estrella, televisión pagada.

En el extremo receptor tenemos como proveedor a CNT EP, empresa pública encargada de las Telecomunicaciones a nivel nacional. Dispone de la red de fibra que cubre todo el territorio nacional con más 10.000Km de distancia e interconecta todas las provincias del Ecuador por medio de las distintas tecnologías de acceso como son: ADSL, Wimax, GPON, Metroethernet, VSAT entre otras, a través de su red de transporte robusta Multiprotocol Label Switching (MPLS) y Dense Wavelength Division Multiplexing (DWDM). Es un proveedor de Internet de categoría TIER II y posee cinco salidas internacionales para interconectarse con el resto del mundo. Su mercado comprende clientes corporativos y gubernamentales, PYMES y masivos por su completa cartera de servicios como son internet, datos y telefonía tanto fijo como móvil, además de las nuevas tendencias tecnológicas, Tecnologías de la Información y la Comunicación (TICs).

5.1.3 Participantes

En base a lo establecido en P.800 sección A.4.1, los usuarios elegidos no han participado en ninguna prueba anterior relacionados a la evaluación de calidad de audio ni en ninguna prueba subjetiva en los últimos seis meses.

Los participantes fueron elegidos dentro de la provincia del Guayas para obtener una visión general de la percepción de dicha población sobre las aplicaciones a analizar.

Según el último censo poblacional realizado en la provincia del Guayas en el año 2010, realizada por el Instituto Nacional de Estadísticas y Censos (INEC) se obtuvo una distribución general de edades registradas en el grafico 5.2, de las cuales el 49.8% corresponde a hombres y el 50.2% a mujeres [51]. Para la presente prueba se tomó un universo de 60 personas distribuidas entre las edades de 15 a 74 años.

Rango de edad	2001	%	2010	%
De 95 y más años	9.743	0,3%	2.281	0,1%
De 90 a 94 años	11.995	0,4%	5.712	0,2%
De 85 a 89años	17.350	0,5%	13.655	0,4%
De 80 a 84 años	25.477	0,8%	25.924	0,7%
De 75 a 79 años	37.182	1,1%	37.219	1,0%
De 70 a 74 años	51.412	1,6%	53.901	1,5%
De 65 a 69 años	45.703	1,4%	56.752	1,6%
De 60 a 64 años	94.293	2,8%	118.685	3,3%
De 55 a 59 años	91.994	2,8%	138.010	3,8%
De 50 a 54 años	130.270	3,9%	166.684	4,6%
De 45 a 49años	158.124	4,8%	204.345	5,6%
De 40 a 44 años	200.728	6,1%	220.145	6,0%
De 35 a 39 años	229.555	6,9%	249.779	6,9%
De 30 a 34 años	255.593	7,7%	289.594	7,9%
De 25 a 29 años	276.926	8,4%	307.034	8,4%
De 20 a 24 años	336.609	10,2%	321.308	8,8%
De 15 a 19 años	321.456	9,7%	338.370	9,3%
De 10 a 14 años	332.561	10,1%	373.511	10,2%
De 5 a 9 años	341.476	10,3%	362.896	10,0%
De 0 a 4 años	340.587	10,3%	359.678	9,9%
Total	3.309.034	100,0%	3.645.483	100,0%

Figura 5.2. Distribución de edades de los guayasenses.

Con el objetivo de cumplir la distribución de participantes masculinos y femeninos determinados en el censo poblacional 2010, se equilibró el número de participantes de ambos géneros. En la tabla 7 se visualiza la cantidad de participantes por rango de edades.

Grupo	Rango de edad	N° Participantes
12	De 70 a 74 años	2
11	De 65 a 69 años	1
10	De 60 a 64 años	3
9	De 55 a 59 años	3
8	De 50 a 54 años	4
7	De 45 a 49años	5
6	De 40 a 44 años	5
5	De 35 a 39 años	6
4	De 30 a 34 años	7
3	De 25 a 29 años	8
2	De 20 a 24 años	8
1	De 15 a 19 años	8
	TOTAL	60

Tabla 7. Rango de edades de la muestra de población Guayas.

5.2 MEDICIÓN DE QoS

El escenario de prueba para la medición de QoS es el determinado en la figura 5.1. A continuación detallares sus componentes adicionales.

5.2.1 Escenario de pruebas para medir QoS

A continuación se describe el proceso de grabación del material de conversación utilizados en las pruebas de campo.

Se generó un archivo de audio con la conversación descrita en el Anexo A de acuerdo a lo especificado en el estándar P.800 para que sea escuchada por los participantes en las diferentes pruebas. Es una conversación normal entre dos personas con inicio y fin natural, y de una duración de 53 segundos.

El entorno de grabación comprende una sala con un volumen de 42 m3, un tiempo de reverberación de 300ms y ruido inferior a 30dBA. En el sistema de grabación se utilizaron los siguientes componentes: un micrófono de condensador AKG PERCEPTION 420 para la recepción de la voz [52], la interfaz Project Mix I/O Fireware [53] para la grabación y mezcla de audio digital basado en computadora y finalmente el programa de audio digital Pro Tools licenciado para el tratamiento del audio digital [54].

En el proceso de grabación intervinieron cuatro participantes, dos mujeres y dos hombres. Para equilibrar el material de grabación a nivel de voces, se realizaron tres grabaciones diferentes mezclando a los participantes, esto es: mujer-mujer, hombre-hombre y mujer-hombre, audios que son alternados en las pruebas finales para contrarrestar la posibilidad de que los resultados sean inclinados a las características de las voces escuchadas por los evaluadores.

El sistema emisor comprende un computador portátil en el cual se encuentra instalados los aplicativos Skype y Hangouts. En dichas aplicaciones se reproduce los audios previamente grabados por medio del dispositivo de entrada Mezcla Estéreo, mecanismo propio de la aplicación que permite la reproducción directa del archivo al receptor sin presencia de ruidos externos.

Las características del equipo emisor se describen en la figura 5.3.

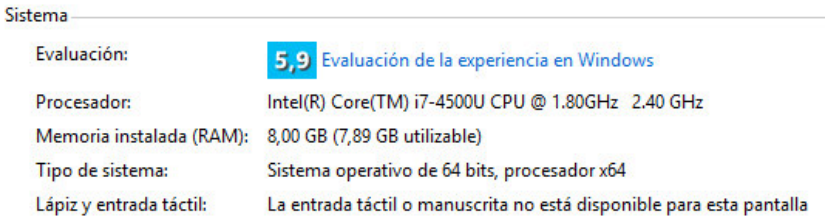


Figura 5.3. Característica de equipo terminal.

5.2.2 Proceso de pruebas

Para llevar a cabo las pruebas experimentales se planificaron las siguientes actividades:

1. Ejecutar la misma aplicación de voz en ambos puntos. El locutor A es permanente y los participantes evalúan desde el extremo B.

2. Levantar el aplicativo Wireshark como analizador de paquetes a fin de capturar el tráfico generado por la conversación en ambos extremos.

3. El locutor A genera la llamada hacia el locutor B a través del aplicativo, cuando el locutor B conteste, A reproduce el audio pre-grabado.

4. El locutor B cierra la llamada y se detiene el analizador de paquetes en ambos extremos.

5. El locutor B llena la encuesta, registrando el nivel de calidad percibida y las dificultades encontradas a lo largo del audio.

6. Los archivos generados por Wireshark son guardados bajo el siguiente formato de nombre:

   grupo(n°)-(identificador_encuestado)-(aplicación_utilizada)-(Audio escuchado)_extremo

   Donde:

   - *N°* = grupo asignado de acuerdo a la edad del encuestado, según distribución de tabla 7.
   - *Identificador encuestado* = primera letra del nombre más el primer apellido: Sara Beltrán, el usuario es sbeltran
   - *Aplicación utilizada* = HG si es Hangouts y SK si es Skype
   - *Audio escuchado* = de acuerdo a la voz de los locutores, participantes en el audio. HH para hombre-hombre, MM para mujer-mujer y HM para hombre-mujer.
   - *Extremo* = *rx* si es recepción y *tx* si es transmisión.

   Un ejemplo de utilización es:

   grupo4-sbeltran-sk-hh-rx.pcap

5.2.2 Toma de resultados

Para realizar la toma de resultados, se realizaron varios pasos:

Para el cálculo de la latencia y jitter, se considera el último salto del extremo receptor, tal como se identifica en la figura 5.4 y para el cálculo de paquetes perdidos se registran los paquetes en ambos extremos.

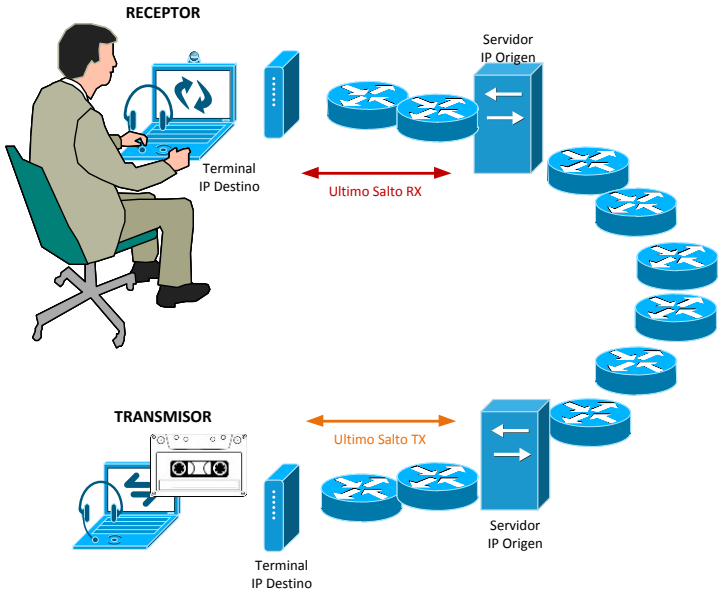


Figura 5.4.Extremos para el cálculo de la latencia y jitter.

En el extremo receptor, luego de capturar el tráfico de las conversaciones de prueba con Wireshark, se exporta el archivo a texto plano desde el aplicativo para poder ejecutar los scripts desarrollados en AWK de la figura 5.4. AWK es un lenguaje de programación utilizado para el procesamiento de datos en texto plano.

El script calcula la latencia obteniendo el promedio de las diferencias entre los intervalos de tiempo de los paquetes enviados desde el servidor público que envía los paquetes UDP hacia el terminal, en el último salto. Para el extremo receptor, la IP destino corresponde a la dirección IP configurada en el terminal donde corre el aplicativo a evaluar y la dirección IP origen es la dirección del servidor público desde el cual se envía los paquetes de voz durante la llamada.

Para el script descrito en la figura 5.5 tenemos dos variables importantes: *ip_destino* e *ip_origen*, parámetros que deben ser actualizados en cada archivo a procesar antes de ejecutar el script. En primera instancia, el script filtra los paquetes UDP, dado que la voz se transmite únicamente bajo este protocolo tanto en Skype como Hangouts; seguidamente filtra los paquetes entrantes y salientes desde y hacia las direcciones *ip_destino* e *ip_origen* descritas como variables globales. Una vez identificado el paquete, luego de estos dos filtros, se suma el jitter global acumulado en la variable *jitterTotal* al jitter del paquete en cuestión obtenido de la diferencia entre el intervalo de tiempo del paquete actual menos el paquete anterior. Si es el primer paquete, la variable *tiempoAnterior* es cero. En cada iteración realizada por el script luego de los dos filtros descritos anteriormente también se cuentan el número de paquetes en el receptor.

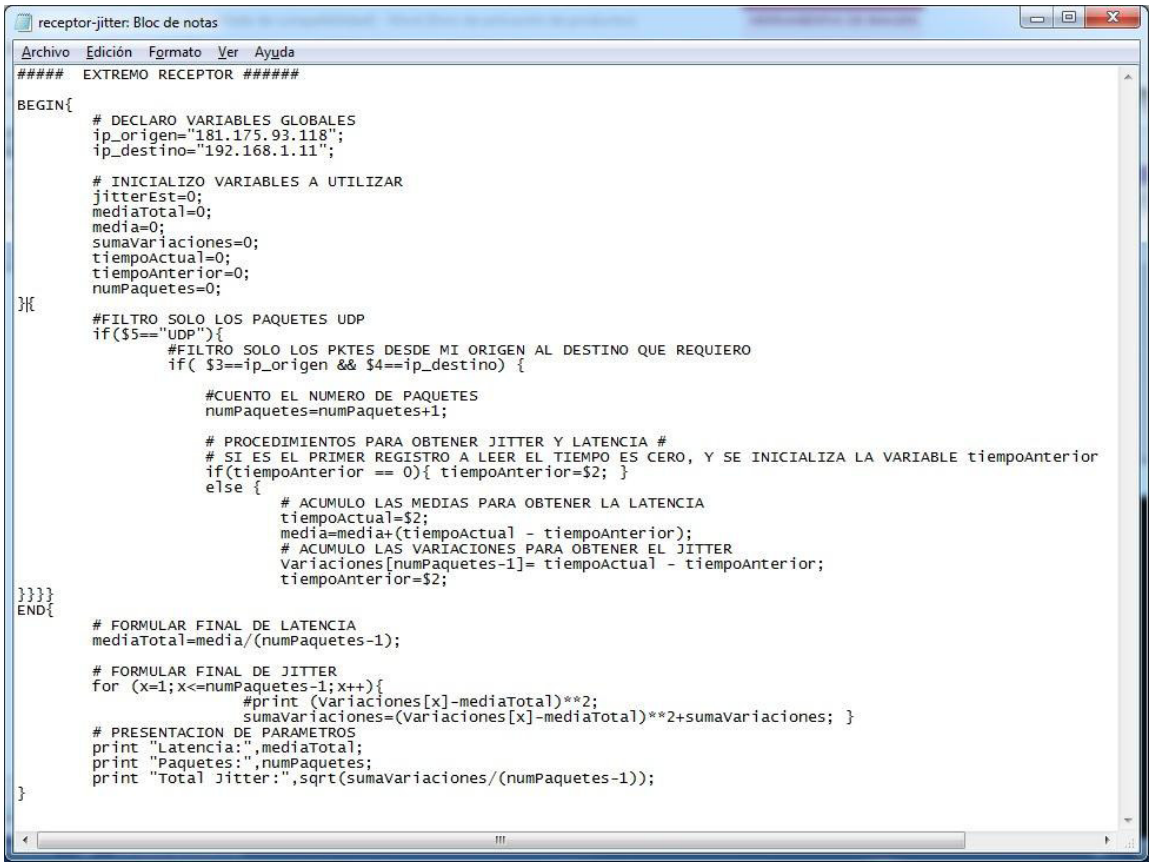


Figura 5.5. Script en AWK del extremo receptor.

De igual manera, en el extremo transmisor, se captura el tráfico con el Wireshark y al archivo exportado a texto plano se ejecuta el script de la figura 5.6 para obtener los paquetes recibidos.

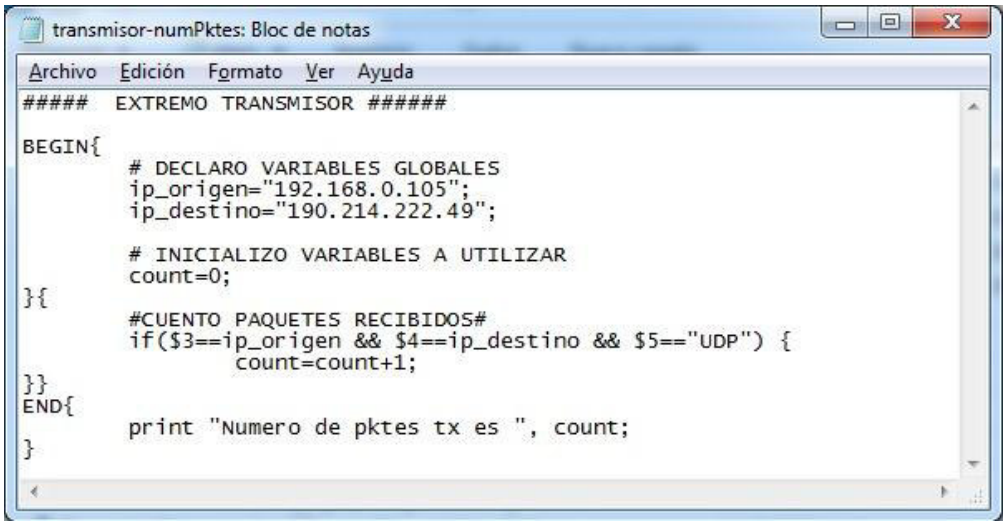


Figura 5.6. Script en AWK del extremo receptor.

De igual manera, la *ip_origen* e *ip_destino* son las direcciones obtenidas en el último salto del lado del transmisor, tal como se describe en la figura 5.4. El script de este extremo es mucho más sencillo, únicamente cuenta los paquetes UDP enviados desde la *Ip_origen* hasta la *Ip_destino*.

Todos los resultados obtenidos de ambos script son tabulados en una tabla de Excel.

5.3 MEDICIÓN DE QOE

Para la medición cualitativa de las pruebas, se realiza una encuesta, para registrar los resultados de la calidad de audio percibidos por el usuario. La encuesta está dividida en dos secciones. La primera sección contiene dos preguntas generales

útiles para obtener información sobre la experiencia del usuario con Internet y aplicaciones VoIP, y en la segunda sección se encuentran las preguntas relacionadas con la evaluación subjetiva.

La primera sección se llena luego de que se explica al participante el objetivo de las pruebas y la metodología a seguir. La segunda sección la describimos en el apartado 5.3.2.

5.3.1 Escenario de pruebas para medir QoE

El escenario de pruebas es compartido con el proceso para la medición de QoS.

Adicionalmente, antes de iniciar las pruebas, se conversa con cada participante explicando el objetivo de la prueba, el registro de la información en la encuesta y el impacto de su participación dentro de la muestra. Una vez que el participante comprende el proceso a seguir, se procede a colocar los audífonos al participante y se realiza una llamada de verificación para determinar el nivel de volumen y la calidad del audio percibido, a fin de ajustarlo de acuerdo a su necesidad y comodidad.

Una vez realizados los pasos descritos en la sección 5.2.1, se solicita la firma y número de cédula del participante para su registro.

5.3.2 Proceso de pruebas

Una vez que el participante haya escuchado el audio de la primera aplicación VoIP, procede a llenar la segunda sección de la encuesta entregada. Esta sección está

compuesta por dos preguntas. La primera pregunta está basada en la escala de opinión sobre la conversación descrita en la sección A.4.2.1 del P.800 [42] en donde el participante selecciona una escala dentro de las categorías indicadas sobre la evaluación de la calidad de audio percibida: excelente, buena, regular, mediocre y mala.

La segunda pregunta está basada en la Escala de dificultad descrita en la sección A.4.2.2 del P.800 donde se asigna una respuesta binaria sobre el grado de dificultad para escuchar la conversación a través de la conexión. Adicionalmente, a esta segunda pregunta se agregó varios items donde el participante puede seleccionar de entre las opciones las dificultades percibidas durante el audio, como son: volumen bajo, ruido o zumbido, distorsión, interrupción, eco, voz robotizada u otro.

5.3.3 Toma de resultados

Las encuestas se recopilaron y tabularon de la siguiente manera:

La primera y segunda pregunta, de la primera sección, son tabuladas de acuerdo al puntaje asignado a cada una de las categorías sobre el uso del Internet y aplicaciones VoIP: Muy a menudo (=5), A menudo (=4), Regular (=3), Alguna vez (= 2) y Nunca (=1).

La primera pregunta de la segunda sección califica la calidad del audio percibido en cada una de las aplicaciones. El puntaje es registrado por cada participante a cada

una de las aplicaciones: Excelente (=5), Buena (=4), Regular (=3), Mediocre (=2) y Mala (=1).

La segunda pregunta es una respuesta binaria sobre la dificultad de escuchar el audio sobre la conexión establecida, los parámetros de evaluación son Sí (=1) y No (=0).

Finalmente, los items registrados sobre las dificultades percibidas son sumados por cada aplicación.

5.4 ANÁLISIS DE LOS RESULTADOS

En esta sección mencionamos a breves rasgos los conceptos básicos de los métodos estadísticos utilizados para el análisis de los resultados y el comportamiento de las dos aplicaciones de VoIP utilizadas.

METODOS ESTADÍSTICOS

ANOVA

El análisis de la varianza es una técnica estadística que permite comparar las medias de más de dos poblaciones; para esto se determina una variable nominal u ordinal que define los grupos a comparar conocida como variable independiente y variables de tipo cuantitativo que cambian a razón o intervalos dados por la variable independiente denominadas variables dependientes, bajo estas variables delimitadas

se analiza si las medias de los grupos son iguales. Este método es importante para comparar tratamientos o experimentos realizados.

Se plantean dos hipótesis para el análisis de los resultados:

Ho: $\mu_1 = \mu_2 = \mu_3 = ... = \mu_n$ versus *Ha*: no todas las medias poblacionales son iguales.

La hipótesis nula (*Ho*) plantea que la media de todos los grupos son iguales y la hipótesis alternativa (*Ha*) que por lo menos una media difiere del resto. Para aceptar la hipótesis nula debemos obtener un valor estadístico llamado *F* calculado de la división entre la varianza población obtenida entre las medias de cada grupo δ_1^2 y la varianza población dentro de cada grupo δ_2^2:

$$F = \frac{\delta_1^2}{\delta_2^2}$$

Si *F* refleja un valor cercano a 1, se acepta la hipótesis nula, caso contrario la hipótesis alternativa.

REGRESION LINEAL

La regresión lineal es un método matemático que permite relacionar una variable dependiente denominada Y con una o más variables independientes $X_1, X_2, X_3, ...X_k$ mediante una relación de dependencia lineal $X_1 = 1$:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + ... + \beta_k X_k + e$$

Donde $\beta_1, \beta_2, +\beta_3, ...\beta_k$ son valores que determinan la influencia de la variable independiente sobre la variable dependiente y *e* es el error que recoge los factores arbitrarios o al azar de nuestro experimento. Si existe una única variable independiente, el método se denomina regresión lineal simple, caso contrario regresión lineal múltiple.

DIAGRAMA DE DISPERSIÓN

Es una representación cartesiana que permite identificar una posible relación entre dos variables. Este diagrama permite visualizar de forma gráfica la tendencia de dicha relación, si la correlación de ambas variables es directa, la recta obtenida tiene una pendiente positiva, caso contrario se tendrá una correlación inversa con una pendiente negativa. En caso de que ambas variables sean independientes entre sí se dice que tienen una correlación nula y la nube de puntos tiene una forma redonda (convexa).

ANALISIS DE LAS APLICACIONES UTILIZADAS

Para la realización de las pruebas de campo, se capturó el tráfico de cada aplicación para luego analizarlo.

SKYPE

Mediante la captura del tráfico con Wireshark, no se registró el envío o recepción de paquetes RTP para poder visualizar su estructura y los mecanismos que usa para brindar QoS, esto es plenamente comprensible dado que Skype utiliza algoritmos y protocolos privados con técnicas de cifrado de datos y un esquema de transferencia de paquetes P2P poco conocido, el cual es actualizado en cada nueva versión. Este hecho ha generado numerosos estudios sobre su identificación, clasificación y tratamiento del tráfico generado dentro de la red [43] [44] [45] [46].

En base a las pruebas realizadas, una conexión hacia cualquier servidor se inicia por medio del envío de una solicitud de conexión para sincronizar los números de secuencia (SYNC, del inglés Synchronize) a varias direcciones IP públicas e iniciar el acuerdo de tres vías (three-way handshake) mediante TCP. El flujo conocido como "apretón de manos" se describe en la figura 5.7. Dicho proceso lo inicia el terminal hacia varios servidores durante todo el ciclo de la llamada.

Terminal	Mensajes	Servidor Público
CLOSED		LISTEN
SYNC-SENT →	<SEQ=0><CTL=SYN>	→ SYNC-RECEIVED
ESTABLISHED ←	<SEQ=0><ACK=1><CTL=SYN,ACK>	← SYNC-RECEIVED
ESTABLISHED →	<SEQ=1><ACK=1><CTL=ACK>	→ ESTABLISHED
ESTABLISHED →	<SEQ=1><ACK=1><CTL=PSH,ACK><DATOS>	→ ESTABLISHED

Figura 5.7. Acuerdo en tres pasos entre terminal y servidor nodo

Posteriormente, se generan paquetes bajo el protocolo de Seguridad en la capa de transporte (TLS, del inglés Transport Layer Security) versión 1.2 [55]. Bajo este protocolo se establece un canal seguro hacia todos los servidores con los cuales se realizó un acuerdo de tres vías previamente.

El flujo de paquetes TLS consta de varios tipos de mensajes que mostramos en la figura 5.8. El terminal inicia enviando un mensaje Client Hello exponiendo todos los conjuntos de cifrados que soporta, ordenados por preferencia. Cada conjunto de cifrado describe el método de clave secreta, método de cifrado y el algoritmo de Hash. Posteriormente, el servidor responde con un Server Hello escogiendo los métodos a utilizar de los propuestos por el terminal, luego se autentica enviando su certificado digital en el cual consta su clave pública mediante un mensaje Certificate, realiza el intercambio de claves mediante el mensaje Change Key Exchange y finalmente cierra la negociación con el mensaje Server Hello Done. En las negociaciones TLS registradas, los servidores públicos no solicitan un certificado digital al terminal, por tal motivo una vez que el terminal recibe el Server Hello Done verifica el certificado digital enviado por el servidor mediante el uso de la clave pública para descifrar el

mensaje recibido y envía un mensaje Finished. Establecido el canal seguro, se realiza un intercambio de información de forma cifrada entre ambas partes.

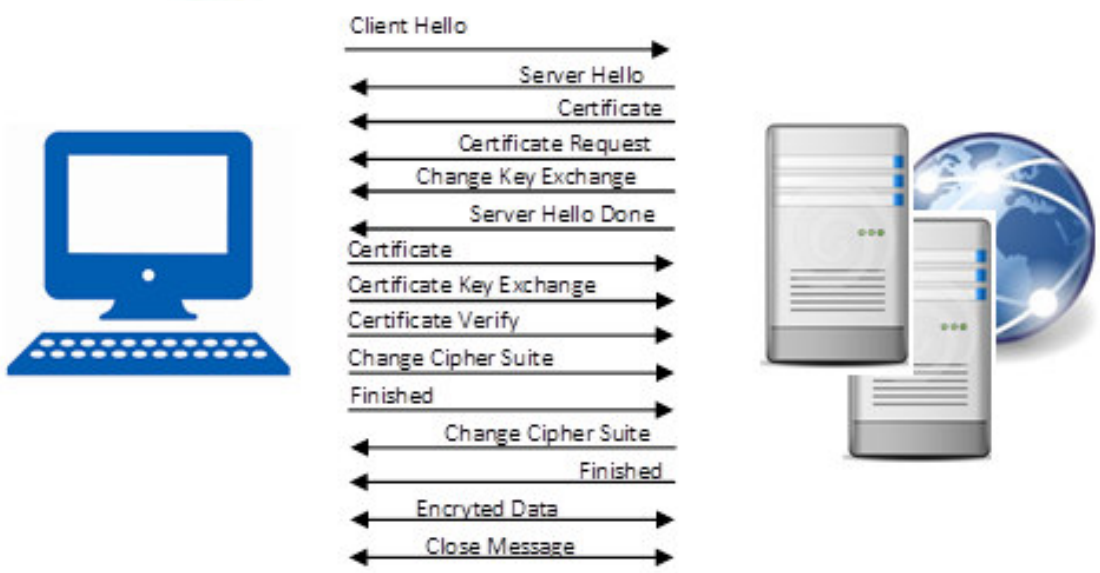


Figura 5.8. Flujo de paquetes para establecer conexión segura.

El presente análisis se divide en 3 partes: inicialización de la aplicación, autenticación e inicio y fin de llamada.

En la primera fase, inicialización de la aplicación, se generan tres canales seguros hacia servidores en redes de Microsoft [56] y Akamai [57], destinos registrados en la tabla 7 donde se visualiza las direcciones IP, la organización propietaria de la red y su país origen, los puertos de conexión de origen y destino y finalmente las entidades certificadoras, de tres llamadas de pruebas realizadas. Durante esta fase, los certificados digitales enviados por los servidores públicos corresponden a dos

subdominios de Microsoft y su entidad certificadora dentro de Microsoft, denominada Microsoft Internet Authority. El tercer certificado corresponde al emitido por una entidad externa, GeoTrust [58], es la empresa que emite y avala el certificado digital de Microsoft Corporation.

Prueba	IP Origen	Propietario de IP	País	Puerto Origen	Puerto Destino	Detalle del certificado
1	191.237.169.142	Microsoft Informatica Ltda	BR	443	51461	*.trap.skype.net
						Microsoft Internet Authority
	172.229.173.137	Akamai Technologies	US	443	51463	www.skypeassets.com
						GeoTrust SSL CA - G4
	23.99.221.17	Microsoft Corporation	US	443	51464	*.pipe.skype.com
						MSIT Machine Auth CA 2
						Microsoft Internet Authority
2	191.238.101.181	Microsoft Informatica Ltda	BR	443	49466	zuul-test.cloudapp.net
						Microsoft IT SSL SHA2
	191.237.169.142	Akamai Technologies	BR	443	49468	*.trap.skype.net
						Microsoft IT SSL SHA2
	23.1.121.83	Microsoft Informatica Ltda	US	443	49469	www.skypeassets.com
						GeoTrust SSL CA - G4
3	191.238.101.181	Microsoft Informatica Ltda	BR	443	49466	zuul-test.cloudapp.net
						Microsoft IT SSL SHA2
	191.237.169.142	Microsoft Informatica Ltda	BR	443	49468	*.trap.skype.net
						Microsoft IT SSL SHA2
	23.1.121.83	Microsoft Informatica Ltda	US	443	49469	www.skypeassets.com
						GeoTrust SSL CA - G4

Tabla 7. Certificados digitales enviados al ejecutar Skype.

Posteriormente se muestra en la figura 5.9 que todas las negociaciones seguras son realizadas a través del conjunto de cifrado TLS_EDCHE_RSA_WITH _AES_256_CBC_SHA384 [59] que corresponde al código 0X28 del conjunto de cifrado determinado en el estándar RFC5289. Este comprende el uso del protocolo de intercambio de claves denominado Intercambio Diffie-Hellman con curva elíptica (ECDHE, del inglés Elliptic Curve Diffie–Hellman Exchange) [60], el algoritmo de

clave pública Rivest-Shamir-Adleman (RSA) [61], el algoritmo de cifrado Advanced Encryption Standard (AES) con una clave de 256 bits de longitud [62] y finalmente el algoritmo Hash de encriptación denominado Algoritmo Hash de Seguridad 384 (SHA, del inglés Secure Hash Algoritm) [63] y según [59].

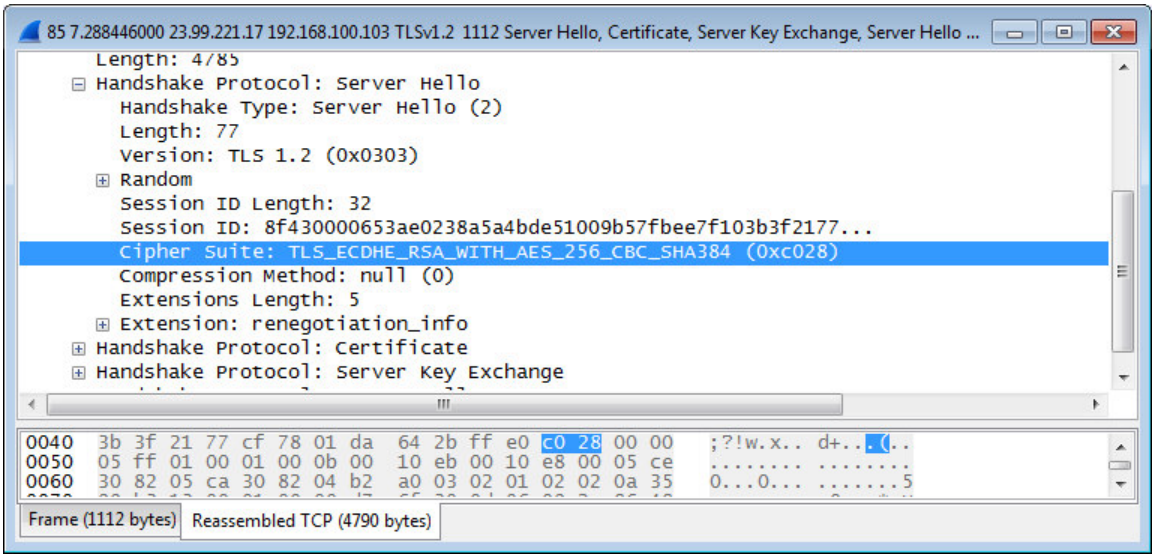


Figura 5.9. Conjunto de cifrado utilizado en el canal seguro en Skype.

La segunda fase comienza al digitar el usuario y la contraseña y pulsar el botón Iniciar Sesión para ingresar a la aplicación. La información de autenticación se envía cifrada a los servidores con los cuales previamente se estableció un contacto seguro. En caso de que las credenciales de acceso sean incorrectas, un mensaje de error se muestra al usuario y no se generan conexiones adicionales.

Una vez iniciado la sesión exitosamente, al instante se despliegan conexiones simultáneas a distintos servidores como MSN, Skype, Hotmail, Bing, Facebook… porque dentro de la aplicación se empiezan a descargar información del perfil y los contactos. Las conexiones hacia estos servidores es una política de migración de contactos una vez que Microsoft adquirió a Skype, este determinó que los usuarios podrían mantener sus libretas de direcciones de ambas aplicaciones, MSN Messenger y Skype, sin perjuicio alguno.

Luego, se generan conexiones seguras hacia Facebook, en base a una nueva integración de Skype hacia dicha plataforma, permitiéndole al usuario enviar mensajes instantáneos hacia los contactos de Facebook directamente a través de una opción dentro de una barra de *Bing*, entidad con la cual también se establece un canal seguro. Los certificados de ambas entidades son visibles en las negociaciones TLS. En la tabla 8 se puede visualizar todas las conexiones seguras establecidas en la fase dos.

IP Origen	Propietario	País	Puerto Origen	Puerto destino	Detalle del certificado
91.190.218.17	Skype	IE	443	50373	api.skype.com Microsoft IT
134.170.24.65	Microsoft	US	443	50373	*.gateway.messenger.live.com Microsoft IT
137.116.80.67	Microsoft	US	443	50378	*.dc.trouter.io Microsoft IT
168.61.160.96	Microsoft	US	443	50379	*.config.skype.com Microsoft IT
137.116.195.37	Microsoft	US	443	50377	api.mcr.skype.com Microsoft IT
23.202.114.161	AKAMAI	US	443	50382	*.skype.com Verizon Akamai SureServer Baltimore CyberTrust Root
191.239.20.186	Microsoft	US	443	50380	prod.tpc.skype.com MSIT Machine Auth CA 2 Microsoft IT
23.202.95.139	AKAMAI	US	443	50388	connect.facebook.net Verizon Akamai SureServer Baltimore CyberTrust Root
93.184.215.200	EDGECAST	EU	443	50387	*.vo.msecnd.net Microsoft IT Baltimore CyberTrust Root
193.149.68.30	MarkMonitor	US	443	50390	*.dc.trouter.io Microsoft IT
172.229.173.134	TUCOWS	US	443	50392	secure.skypeassets.com Verizon Akamai SureServer Baltimore CyberTrust Root
23.202.82.110	TUCOWS	US	443	50394	s-static.ak.facebook.com Verizon Akamai SureServer Baltimore CyberTrust Root
134.70.0.200	DESKTALK	US	443	50393	*.hotmail.com GlobalSign Organization Validation
204.79.197.200	MarkMonitor	US	443	50396	www.bing.com Microsoft IT
131.253.14.192	Microsoft	US	443	50395	c.msn.com Microsoft IT
31.13.73.1	MarkMonitor	US	443	50397	*.facebook.com Digicert High Assurance

Tabla 8. Certificados digitales posterior a la autenticación en Skype.

La tercera fase es el inicio de una llamada hacia un contacto de Skype, además de generarse los procesos conocidos descritos anteriormente, existen cuatro paquetes del protocolo *Quick UDP Internet Connections* (*QUIC*) que se intercambian entre

clientes y servidor. El terminal envía dos paquetes con una carga útil de tamaño fijo en ambos mensajes (figura 5.10), pero la respuesta del servidor es siempre distinta, el tamaño es variable generalmente de mayor longitud que las enviadas por el terminal. Esto es constante en todas las llamadas de prueba realizadas, pero tomando en consideración que el tamaño fijo del paquete del cliente es diferente en cada llamada realizada.

Se desconoce el contenido y la funcionalidad de dicha información pero se presume son los datos del contacto con el cual se establecería la llamada por el comportamiento de las trazas.

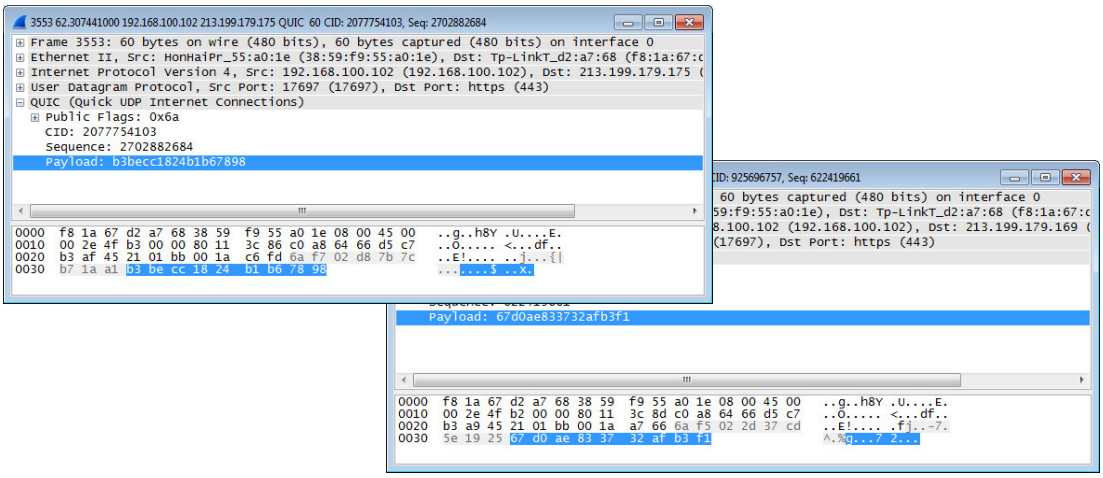


Figura 5.10. Vista interior a paquetes del protocolo QUIC.

Una vez establecida la conexión con los supernodos y servidores de autenticación, se da paso al flujo de voz únicamente sobre paquetes UDP entre el origen y el

destino, confirmado por el estudio realizado en [64]. Este flujo de datos se realiza con el supernodo más cercano para el presente caso es con la dirección IP 190.10.154.149 perteneciente a SATNET, proveedor de Internet ecuatoriano según se muestra en la figura 5.11 y 5.12. Esto se comprueba en las 60 trazas tomadas en las pruebas de campo realizadas, se conecta a un servidor dentro de la red del proveedor por el cual está conectado el terminal para el envío de los paquetes UDP.

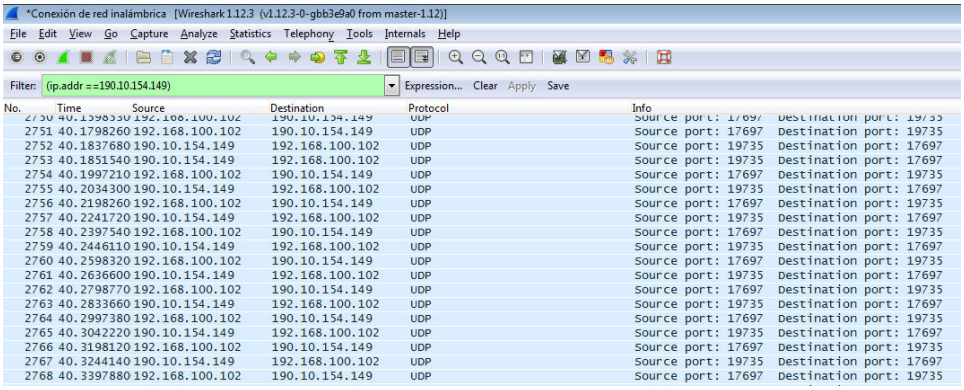


Figura 5.11. Flujo de voz sobre paquetes UDP de Skype.

En resumen, Skype establece generalmente conexiones TCP con los supernodos y servidores de autenticación y utiliza TLS 1.2 para establecer un canal cifrado. Las conexiones UDP son utilizadas para la transmisión de la voz entre los terminales origen y destino y toda la información se envía de forma cifrada. El esquema que maneja Microsoft para sus usuarios de Skype es altamente distribuido y escalable.

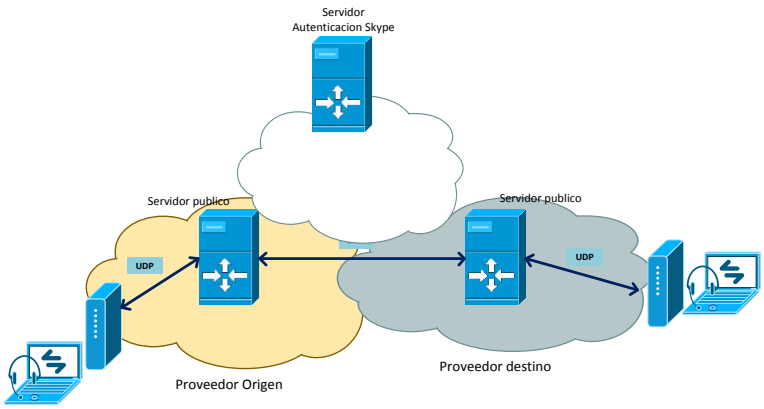


Figura 5.12. Esquema de flujo de paquetes de Skype.

GOOGLE HANGOUTS

Similar al análisis seguido para Skype, se capturó el tráfico por medio de Wireshark y se determina que Hangouts establece la misma secuencia de procesos descritos en la figura 5.7 y 5.8, esto es, el apretón de manos y la conexión segura.

Dado que Hangouts es un plugin dentro de un navegador, inicialmente se ejecuta el navegador que pueden ser Firefox, Chrome, Iexplorer u otro. Posteriormente se procede a ejecutar el plugin pero nos solicita autenticación, paso en el cual se debe autenticar con una cuenta de gmail. Una vez dentro de la cuenta de correo visualizamos el despliegue de nuestros contactos. Para este proceso inicial se establecen diferentes canales seguro y se registra la transmisión de los certificados digitales hacia los servidores de autenticación como se muestra en la tabla 9. En

primera instancia se visualizan certificados correspondientes al navegador utilizado, datos que se escapan del presente análisis. Posteriormente se establecen canales seguros hacia diferentes direcciones IP de Google pero únicamente se transmiten dos certificados: el de la página web de google y el *googleusercontent*. Además, se registra el envío de los certificados digitales más de una vez a ciertas direcciones IP, esto es un escenario que se repite de forma constante tanto al momento de autenticarse y de las llamadas. Se presume este comportamiento por perdida de paquetes, pero el dato curioso es que no se registra el mismo comportamiento en Skype.

IP Origen	Propietario de IP	País	Puerto Origen	Puerto Destino	Detalle del certificado
64.233.185.104 (2)	GOOGLE	US	443	49614	www.google.com
					Google Internet Authority G2
					GeoTrust Global CA
216.58.219.110	GOOGLE	US	443	49625	www.google.com
					Google Internet Authority G2
					GeoTrust Global CA
216.58.219.142 (8)	GOOGLE	US	443	49626	www.google.com
					Google Internet Authority G2
					GeoTrust Global CA
216.58.219.131 (4)	GOOGLE	US	443	49627	www.google.com
					Google Internet Authority G2
					GeoTrust Global CA
216.58.219.129	GOOGLE	US	443	49636	*.googleusercontent.com
					Google Internet Authority G2
					GeoTrust Global CA

Tabla 9. Certificados digitales posterior a la autenticación en Hangouts.

Posteriormente, como se muestra en la figura 5.13, todas las negociaciones seguras son realizadas a través del conjunto de cifrado TLS_EDCHE_RSA_WITH _AES_128_GSM_SHA256 que corresponde al código 0X2B del conjunto de cifrado

determinado en el estándar RFC5289. Este comprende el uso del protocolo de intercambio de claves denominado Intercambio Diffie-Hellman con curva elíptica ECDHE, el algoritmo de llave pública RSA, el algoritmo de cifrado AES con una clave de 128 bits de longitud (menor que en Skype) y finalmente el algoritmo Hash de cifrado denominado Algoritmo Hash GCM (Galois/Counter Mode) [65] de 256 bits.

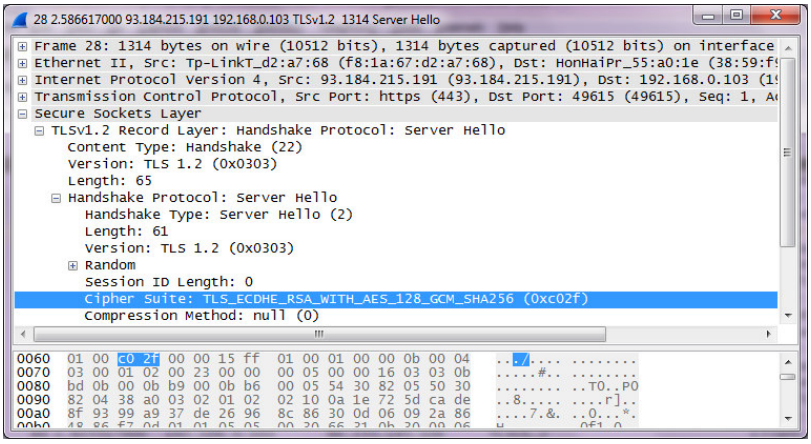


Figura 5.13. Conjunto de cifrado utilizado en el canal seguro en Hangouts.

Al igual que en Skype, el flujo de paquetes de voz se hace con UDP pero con la diferencia que el servidor que recibe o transmite está ubicado en redes correspondientes a Estados Unidos. Escenario que se repite en las 48 pruebas de campo realizadas.

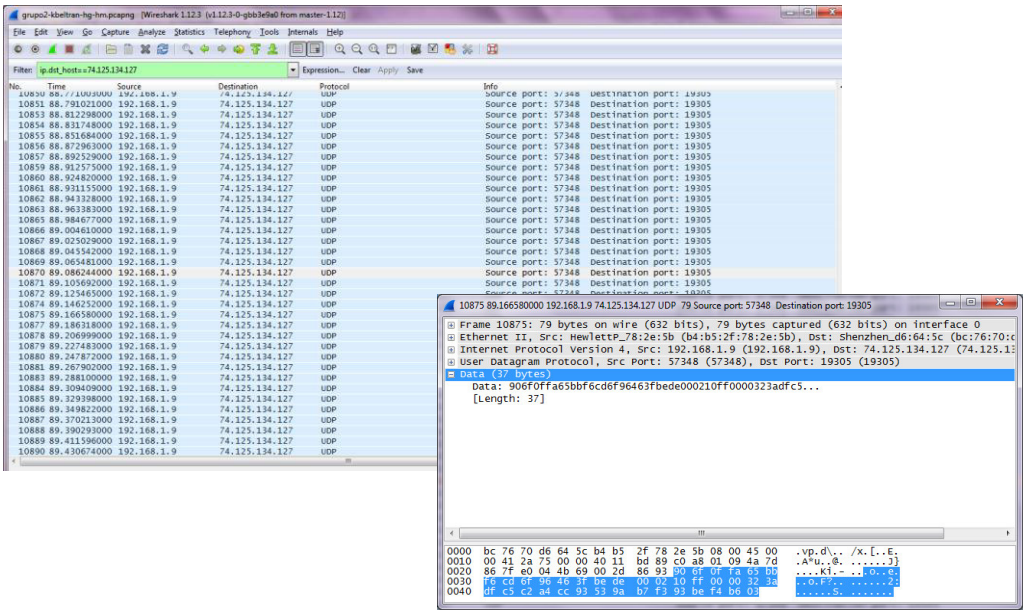


Figura 5.14. Flujo de voz sobre paquetes UDP de Hangout.

En resumen, Hangouts establece generalmente conexiones TCP con los servidores de Google y utiliza TLS 1.2 para establecer un canal cifrado. Las conexiones UDP son utilizadas para la transmisión de la voz entre los terminales origen y destino y toda la información se envía de forma cifrada. El esquema que maneja Google para sus usuarios se concentra en sus servidores de USA tal como se describe en la figura 5.15.

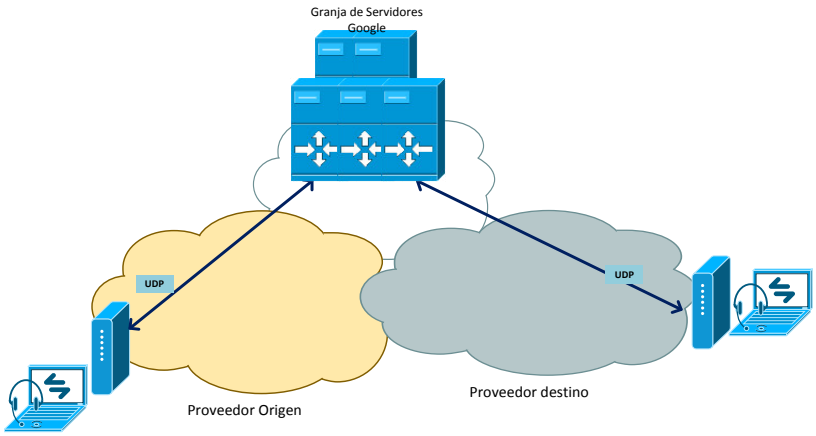


Figura 5.15. Esquema de flujo de voz sobre paquetes UDP de Hangout.

5.4.1 Resultados obtenidos de parámetros QoS

A fin de comprender el escenario topológico de las pruebas, se toma en consideración las figuras 5.9 y 5.12 correspondientes al esquema de flujo de paquetes de voz en las llamadas realizadas en Skype y Hangouts.

En resumen, como mostramos en la tabla 10, las direcciones IP utilizadas por Skype corresponden a servidores de proveedores locales, a diferencia de Hangouts, todos los servidores son de Estados Unidos. A nivel de transporte, el puerto 6523 y 42974 son los utilizados por Skype a transmisión y recepción de paquetes respectivamente, a nivel de Hangouts el puerto que se utilizaba mayoritariamente en ambos extremos es el 19305.

EXTREMO	APLICACIÓN	DIRECCIÓN IP	PUERTO ORIGEN
TRANSMISIÓN	HANGOUTS	173.194.73.127	19307
		64.233.177.127	19305
			19306
		74.125.137.127	19305
		74.125.196.127	19305
	SKYPE	181.113.137.9	6523
		181.113.143.247	6523
		181.113.250.202	6523
		181.196.154.105	6523
		186.178.219.113	6523
		186.178.242.53	6523
		186.178.252.225	53375
		186.178.252.73	1024
		190.214.222.49	6523
		190.214.236.25	6523
RECEPCIÓN	HANGOUTS	173.194.73.127	19305
		74.125.134.127	19305
	SKYPE	181.175.93.118	6523
			42974
		190.10.154.149	42974

Tabla 10. Direcciones IP y puertos utilizados por las aplicaciones de prueba.

Luego de tabulado los resultados, se obtienen los valores de latencia, paquetes perdidos y jitter de ambas aplicaciones que detallamos a continuación:

A nivel de latencia, para garantizar que el oído humano no perciba degradación en la calidad de la voz, los paquetes deben arribar máximo a 20ms. Se registra en la figura 5.16 que el promedio de tiempo de envío/recepción de paquetes a nivel de Hangouts es de 24.2ms a diferencia de Skype que es de 22ms.

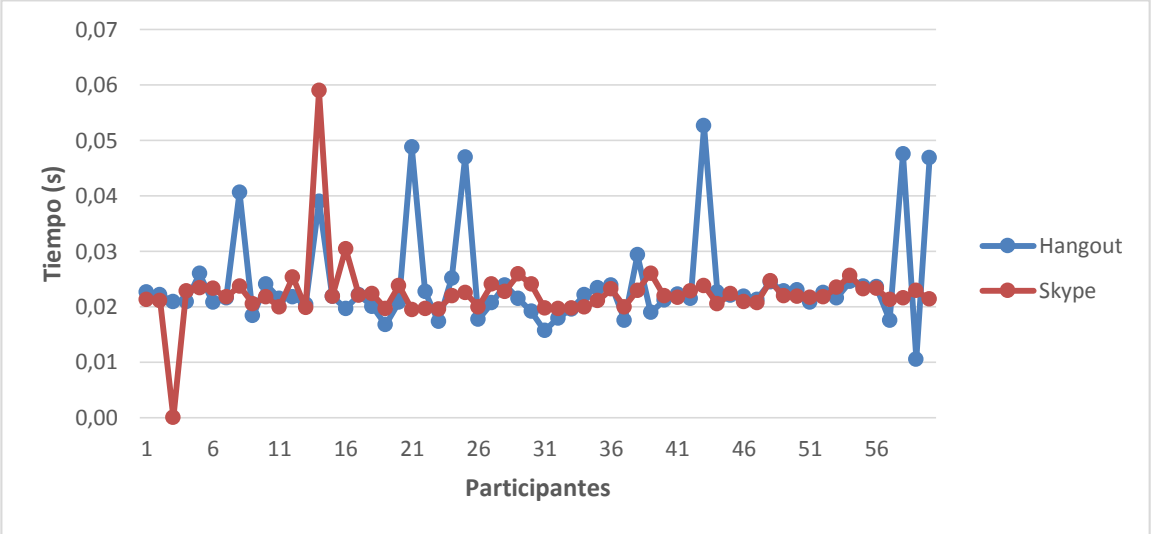


Figura 5.16. Latencia de las 60 pruebas realizadas.

En la estadística del porcentaje de paquetes perdidos de la figura 5.17, se registra que Skype tiene un tendencia a cero de paquetes perdidos a diferencia de Hangouts, el cual presenta un promedio de 28% de paquetes perdidos. Estos es comprensible considerando que Skype dispone de servidores dentro de los dos operadores de telecomunicaciones utilizados para las presentes pruebas, a diferencia de Hangouts, todos los paquetes de voz se envían a los servidores de Google, ubicados en Estados Unidos generando una larga ruta de ida y regreso para los paquetes UDP.

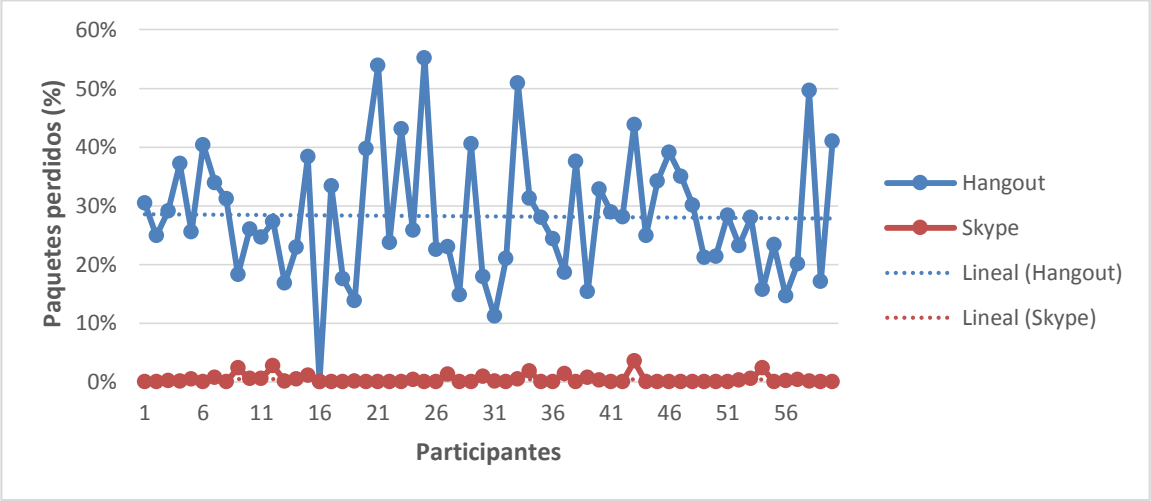


Figura 5.17. Porcentaje de paquetes perdidos presente en las pruebas.

El jitter por ser considerado la media de la variación del tiempo de llegada de paquetes, fue calculado desde el archivo de captura del lado del receptor, es decir se calcula la diferencia de tiempo en que el servidor público envía los paquetes UDP al terminal. Se visualiza que el jitter presente en Hangout es más estable y con un rango de variación menor comparando con Skype como muestra la figura 5.18.

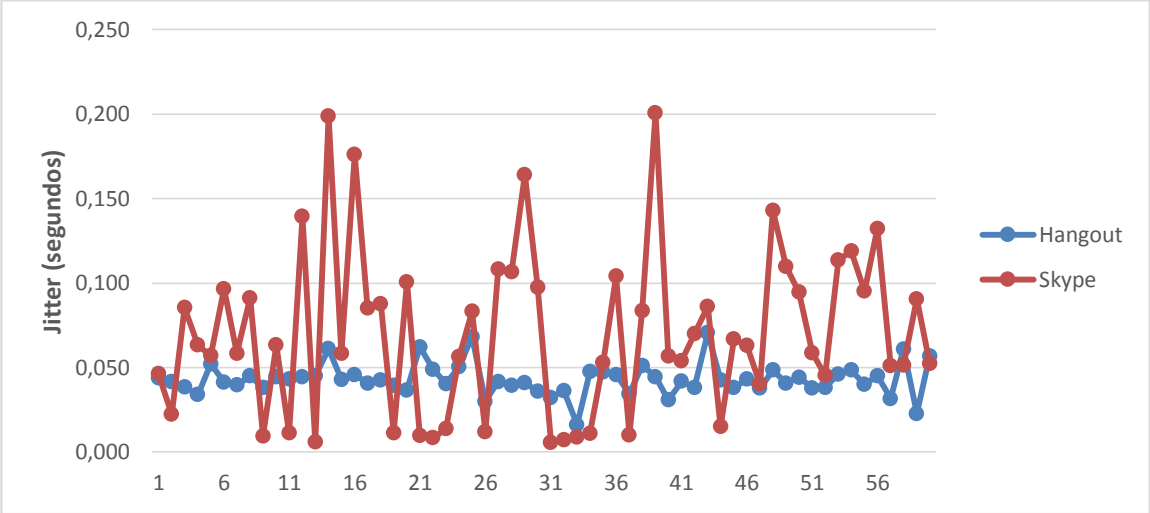


Figura 5.18. Jitter presente en las pruebas.

5.4.2 Resultados obtenidos de parámetros QoE

De la muestra tomada, se establece un equilibrio entre los participantes. El número de hombres y mujeres permanece en equilibrio para mantener el resultado de la muestra poblacional guayasense como se puede visualizar en la figura 5.19.

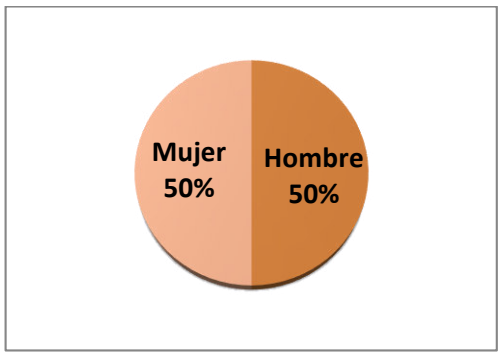


Figura 5.19. Porcentaje de género de los encuestados.

A nivel informativo, se consulta a los participantes sobre la frecuencia de uso del Internet y de aplicaciones VoIP para determinar el grado de conocimiento con el que el participante evalúa, a fin de realizar análisis posteriores y contrastar resultados. Los datos fueron registrados en la figura 5.20. Se visualiza claramente que le juventud son los cibernautas por excelencia, a medida que edad avanza el nivel de uso disminuye gradualmente llegando a cero en los grupos 11 y 12 (edades entre 65 a 74 años).

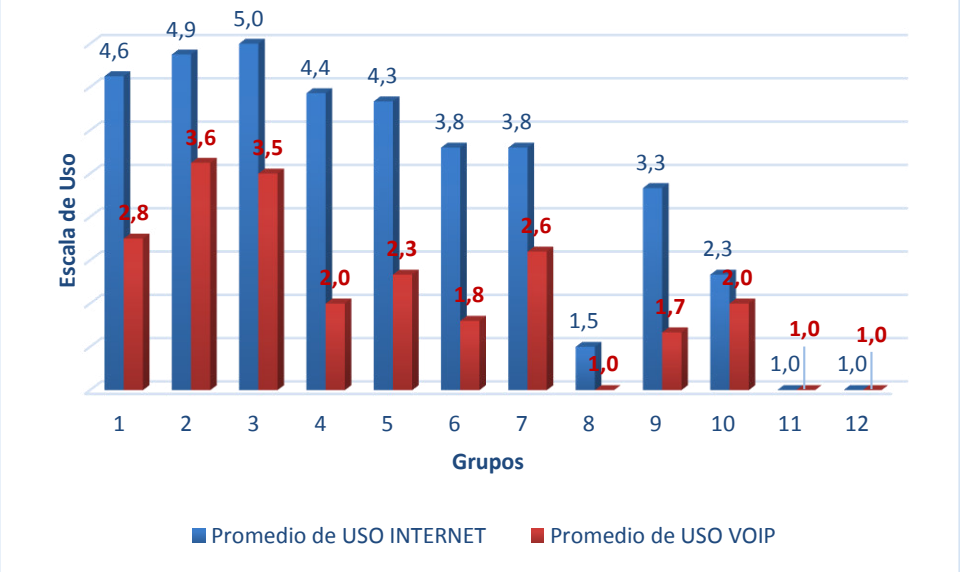


Figura 5.20. Uso de Internet y VoIP.

A nivel de novedades registradas en las conversaciones escuchadas por los participantes durante la llamada realizada para las dos aplicaciones a evaluar se obtiene que Hangouts presenta un 61.2% de dificultad frente a un 38.1% con respecto

a lo evaluado en Skype. Bajo estos resultados, Hangouts presenta un deterioro significativo en la calidad del audio percibida por el usuario final. Ver la figura 5.21 En la figura 5.22 y 5.23, registramos el tipo de dificultad percibida por los participantes en las dos aplicaciones.

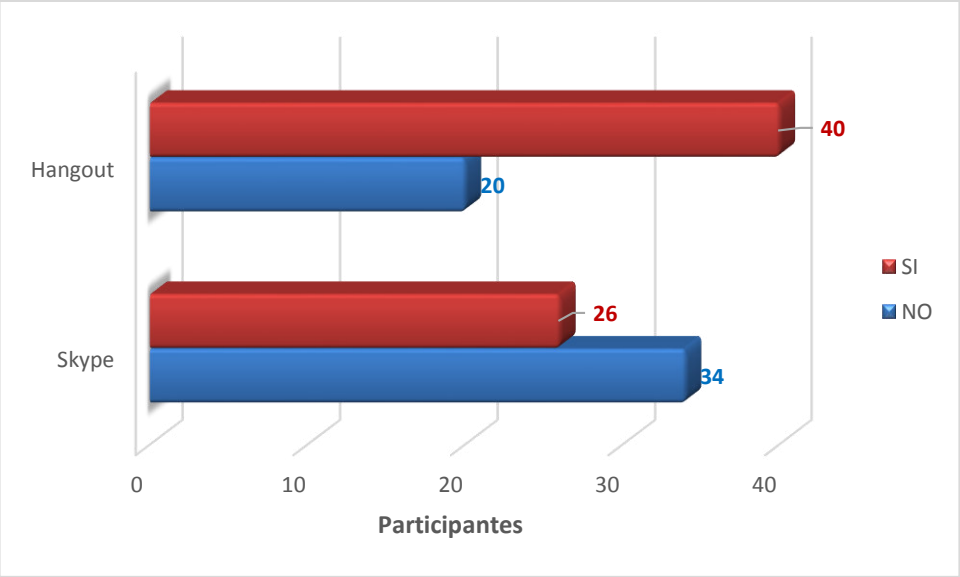


Figura 5.21. Dificultad al oír en Hangouts y Skype.

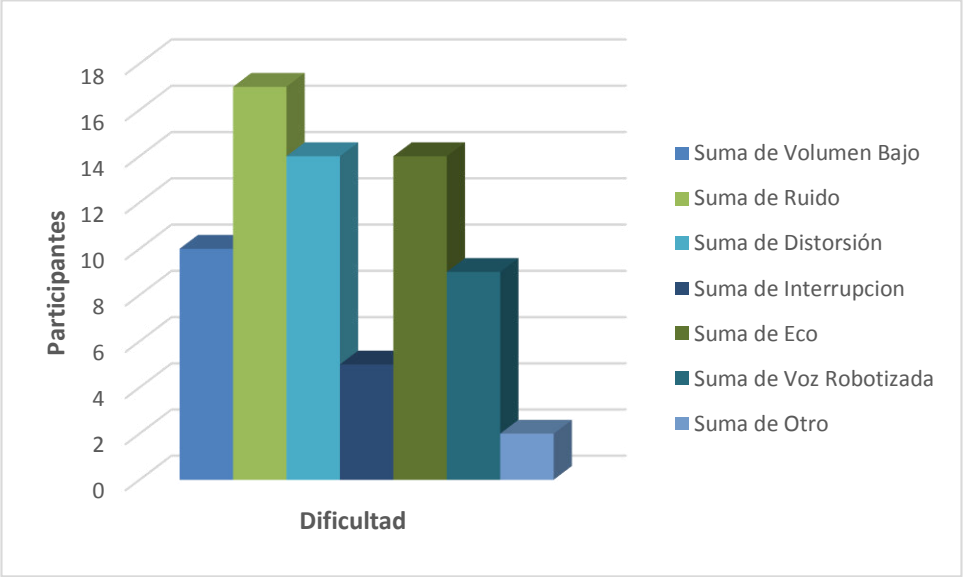


Figura 5.22. Tipo de dificultad al oír en Hangout.

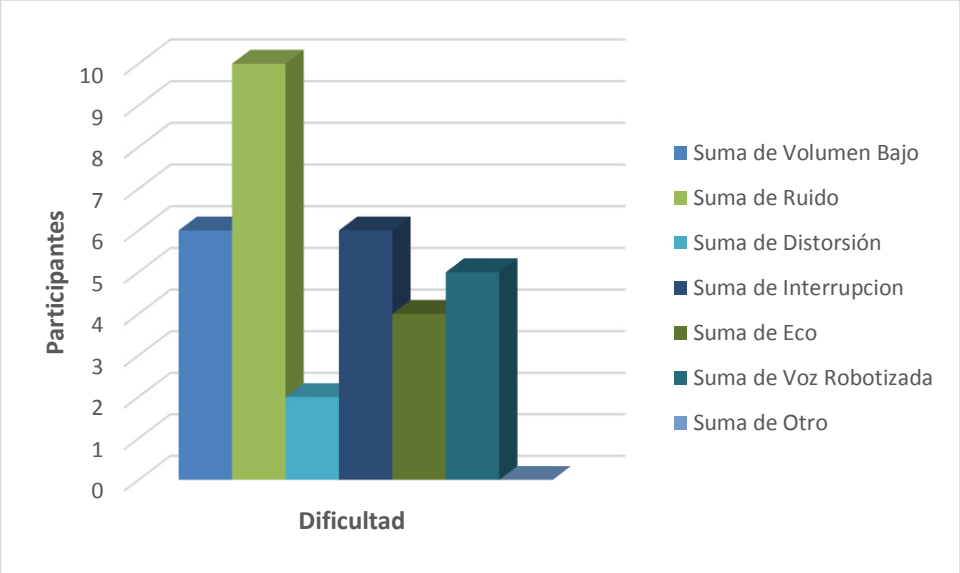


Figura 5.23. Tipo de dificultad al oír en Skype.

Finalmente, el resultado subjetivo sobre la calidad de audio de ambas aplicaciones, Hangouts y Skype, es registrado en la figura 5.24 con un valor de MOS de 4.6 de frente a un 3.6 de Hangouts.

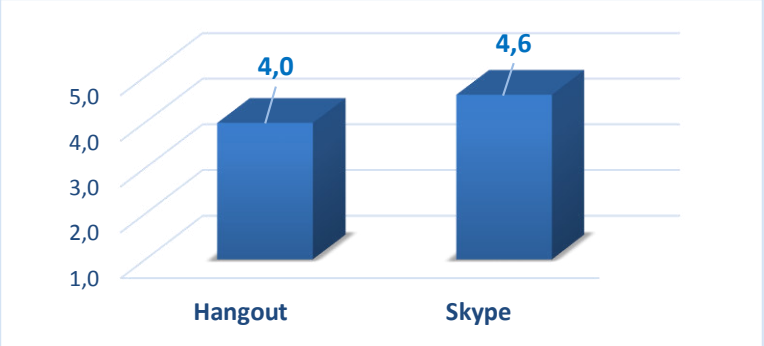


Figura 5.24. Valor de MOS en Hangouts y Skype.

5.4.1 Análisis comparativo de resultados QoS y QoE

A continuación se detallaran los diferentes análisis realizados de los resultados tabulados en base a las pruebas realizadas.

EXPERIENCIA DEL USUARIO

En las preguntas realizadas en la encuesta, sección I, se consultó el nivel de experiencia que el participante tenía con respecto al uso del Internet y aplicaciones VoIP similares a las utilizadas en la presente prueba. En la tabla 11 se visualizan los resultados del MOS obtenido en base a dichos parámetros.

APLICACIÓN	MOS	Promedio de USO INTERNET	Promedio de USO VOIP
Hangout	1	4,0	4,0
	2	5,0	3,3
	3	4,6	2,5
	4	3,8	2,5
	5	3,5	2,2
Skype	3	4,0	4,0
	4	4,3	2,8
	5	3,8	2,2

Tabla 11. MOS en base a la experiencia del usuario.

TIMBRE DE VOZ

La hipótesis sobre el timbre de voz y su afectación en la calidad de audio percibida por el evaluador es confirmada en base a la prueba Anova de un factor resultante de los datos obtenidos en la tabla 12.

AUDIO	HH	MH	MM
Hangout	3,8	4,1	4,0
Skype	4,7	4,5	4,5

Tabla 12. MOS por tipo de audio.

Luego del análisis estadístico, en la tabla 13 se registra un valor de *F*=0,00819672, valor muy pequeño con respecto al límite crítico para F con una probabilidad de 0,99185891, por lo tanto se acepta la hipótesis nula de igualdad de medias y se confirma que el puntaje de MOS no depende del timbre de voz de los participantes de la grabación.

Origen de las variaciones	*Suma de cuadrados*	*Grados de libertad*	*Promedio de los cuadrados*	*F*	*Probabilidad*	*Valor crítico para F*
Entre grupos	0,00333333	2	0,00166667	0,00819672	0,99185891	9,5520945
Dentro de los grupos	0,61	3	0,20333333			
Total	0,61333333	5				

Tabla 13. Resultado del análisis de varianza de la tabla 9.

RELACIÓN DE VARIABLES Y MOS

A fin de encontrar la influencia de las tres variables medidas en las diferentes observaciones realizadas, se verifico su impacto en las dos aplicaciones.

HANGOUTS

En Hangouts, aplicando regresión lineal para obtener la relación que existe entre el MOS obtenido y los valores de latencia, jitter y paquetes perdidos tenemos un factor R^2 de 0,100959244 según la tabla 14, valor cercano a cero por lo que se concluye que el modelo no es confiable para realizar pronósticos de MOS en base a los tres parámetros mencionados.

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0,317740844
Coeficiente de determinación R^2	0,100959244
R^2 ajustado	0,052796347
Error típico	0,956466143
Observaciones	60

Tabla 14. Resultado de la regresión lineal múltiple para Hangouts.

Modelando la ecuación, tenemos:

$$Y = a + bX1 + cX2 + dX3$$

$$Y= 4,8 + 62,9\ X1 - 54,26\ X2 - 0,05\ X3$$

Donde X_1 es la latencia, X_2 corresponde al jitter y X_3 al porcentaje de paquetes perdidos.

Dado que la regresión lineal múltiple no fue exitosa, se comprobó la regresión lineal simple entre el MOS y cada una de las tres variables de forma individual. Los resultados son mostrados en la Tabla 15, el valor de R^2 es muy cercano a cero en los tres casos y por lo tanto las ecuaciones no son confiables.

PARÁMETRO	ECUACIÓN	VALOR DE R^2
LATENCIA	y = 0,0011x + 0,0199	0,0149
JITTER	y=-0,0006x + 0,0455	0,0043
PAQUETES PERDIDOS	y = 0,0217x + 0,1954	0,0381

Tabla 15. Regresión lineal simple entre variables para Hangouts.

En resumen, no se encontró una forma, a través de los métodos estadísticos utilizados, de pronosticar el resultado de MOS por medio de los parámetros de latencia, jitter y paquetes perdidos obtenidos en las presentes pruebas, resultados plenamente comprensibles debido a:

1. Hangout establece comunicación con servidores ubicados en Estados Unidos, el número de saltos y las rutas por el cual los paquetes IP recorren pueden ser variadas, afectando los parámetros de latencia y jitter del extremo final.

2. La latencia y el jitter son medidos entre el terminal y el último salto de recepción en base al grafico 5.4, es decir, la visión de ambos parámetros es limitada dado que no conocemos el trayecto completo por el cual circula la voz. Las capturas de tráfico tomadas a través del sniffer únicamente muestran los paquetes recibidos por el extremo receptor.

3. El valor de paquetes perdidos equivale al número de paquetes enviados en el transmisor menos los recibidos en el extremo receptor, pero se desconoce si en el trayecto el tratamiento de los paquetes en los saltos intermedios es únicamente reenviarlos o son multiplexados, comprimidos, agregado información, etc, procesos que pueden afectar el número de paquetes recibidos.

4. El proceso de manejo de la voz y los protocolos usados son invisibles para el presente análisis por ser propietarios.

5. El valor de MOS obtenido evalúa la calidad del audio de escucha del tráfico que circula de extremo a extremo a diferencia de los parámetros de latencia y jitter.

SKYPE

Realizando un análisis estadístico similar para Skype, exponemos los resultados a continuación.

A nivel de relación entre las variables de latencia, jitter y paquetes perdidos mediante regresión lineal múltiple obtenemos un factor R^2 de 0, 03849484, resultado que se muestra en la tabla 16, valor muy cercano a cero que permite concluir que el modelo no es confiable para pronósticos.

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0,180623751
Coeficiente de determinación R^2	0,032624939
R^2 ajustado	-0,019198725
Error típico	0,532261868
Observaciones	60

Tabla 16. Resultado de la regresión lineal múltiple para Skype.

La ecuación obtenida es:

$$Y= 4,7 + 2,12\ X_1 - 1,98\ X_2 – 3,28\ X_3$$

Dado que la regresión lineal múltiple no fue exitosa, se comprobó la regresión lineal simple entre el MOS y cada una de las tres variables de forma independiente. Los resultados son mostrados en la Tabla 17, el valor de R^2 es muy cercano a cero en los tres casos y por lo tanto las ecuaciones no son confiables.

PARÁMETRO	ECUACIÓN	VALOR DE R^2
LATENCIA	y = -0,0009x + 0,0266	0,0064
JITTER	y = -0,0051x + 0,0906	0,0104
PAQUETES PERDIDOS	y = -0,0013x + 0,0093	0,0274

Tabla 17. Regresión lineal simple entre variables para Skype.

En resumen, no se encontró una forma, a través de los métodos estadísticos utilizados, de pronosticar el resultado de MOS por medio de los parámetros de latencia, jitter y paquetes perdidos obtenidos en las presentes pruebas, resultados plenamente comprensibles debido a:

1. Skype a pesar de tener servidores locales para el manejo de la voz, el número de saltos y las rutas por el cual los paquetes IP recorren pueden ser variadas, afectando los parámetros de latencia y jitter del extremo final.

2. La latencia y el jitter son medidos entre el terminal y el último salto de recepción en base al grafico 5.4, es decir, la visión de ambos parámetros es limitada dado que no conocemos el trayecto completo por el cual circula la voz. Las capturas de tráfico tomadas a través del sniffer únicamente muestran los paquetes recibidos por el extremo receptor.

3. El valor de paquetes perdidos equivale al número de paquetes enviados en el transmisor menos los recibidos en el extremo receptor, pero se desconoce si en el trayecto el tratamiento de los paquetes en los saltos intermedios es únicamente reenviarlos o son multiplexados, comprimidos, agregado información, etc, procesos que pueden afectar el número de paquetes recibidos.

4. El proceso de manejo de la voz y los protocolos usados son invisibles para el presente análisis por ser propietario.

5. El valor de MOS obtenido evalúa la calidad del audio de escucha del tráfico que circula de extremo a extremo a diferencia de los parámetros de latencia y jitter.

Finalmente concluimos que Skype tuvo un mejor desempeño para la muestra poblacional obtenida de la provincia del Guayas con un MOS de 4.6 frente a un 4.0 de Hangouts, esto es comprensible dado que Skype mantiene un servidor local en la red de ambos proveedores de prueba pero con los datos tomados no podemos pronosticar la calidad de audio por medio de los parámetros de latencia, jitter y paquetes perdidos.

CONCLUSIONES

Las principales conclusiones alcanzadas son las siguientes:

1. En esta tesis hemos realizado un estudio detallado de RTP y RTCP para conocer su funcionalidad y comportamiento.
2. En las pruebas de campo, las capturas realizadas de los paquetes no fueron visualizadas a nivel de estructura y mecanismos utilizado por RTP y RTCP de las aplicaciones seleccionadas dado que utilizan protocolos privados y encriptados. Las aplicaciones seleccionadas para el análisis del RTP y RTCP, Skype y Hangouts, no fueron las apropiadas.
3. El timbre de la voz de los interlocutores del audio pre-grabado no influye en los resultados de la evaluación de MOS para los usuarios.

4. El nivel de experiencia del usuario influye directamente en la percepción de la calidad de las aplicaciones, a mayor experiencia más exigencia para las aplicaciones sobre Internet.
5. A través de los resultados obtenidos en las pruebas realizadas no se pudo generar una ecuación que permita deducir el valor de MOS por medio de los parámetros de latencia, jitter y paquetes perdidos
6. La aplicación VoIP de propiedad de Microsoft, Skype, tiene un mejor rendimiento que la aplicación desarrollada por Google, Hangouts, según la encuesta realizada a la muestra tomada de la población perteneciente a la provincia de Guayas, Ecuador en el rango de edades de 15 a 74 años.

RECOMENDACIONES

En esta tesis se ha abierto varias posibles líneas de investigación que consideramos muy interesantes:

1. Profundizar en el estudio de los protocolos RTP y RTCP utilizando aplicaciones VoIP con licencia GNU.
2. Realizar pruebas de funcionamiento con una muestra a nivel nacional para determinar si la población ecuatoriana prefiere Skype frente a Hangouts.
3. Realizar pruebas similares con otros proveedores locales o desde otro país para determinar si el comportamiento de Hangouts es similar a nivel de paquetes perdidos, jitter y latencia en el tráfico de voz.

ANEXO A

CONVERSACION

Locutor A: ¡Buenos días! <SILENCIO 1 s> Empresa de productos agrícolas ARCOS. Un gusto saludarle ¿en qué le puedo ayudar?

Locutor B: Buenos días. <SILENCIO 1 s> ¿Me puede comunicar con Mauricio Villegas?

Locutor A: El Sr. Villegas no se encuentra en oficina. ¿Desea dejar un mensaje?

Locutor B: <SILENCIO 2 s> Sí, mi nombre es Elisa Mendoza. Llamo para confirmar un pedido que le hice al Sr. Villegas.

Locutor A: ¿Tiene el número de pedido?

Locutor B: Sí, <SILENCIO 1 s> es el 156283736-99.

Locutor A: ¿156283736-99?

Locutor B: Sí, correcto.

Locutor A: ¿Cuál es el pedido?

Locutor B: Dos toneladas de arroz Flor, 10 cajas de aceite La Favorita, 3 cajas de margarina Bonita y 10 libras de canela en polvo.

Locutor A: Estimado <SILENCIO 1 s> su pedido fue despachado el día de ayer, en el transcurso de la tarde estaría llegando a sus instalaciones.

Locutor B: Excelente. Muchas gracias. ¡Que tenga un Buen día!

ANEXO B

ENCUESTA

SECCIÓN I

¿Con qué frecuencia utiliza UD Internet?

Muy a menudo _____
A menudo _____
De vez en cuando _____
Alguna vez _____
Nunca _____

Indique con qué frecuencia utiliza UD aplicaciones VoIP (similares a las utilizadas en esta prueba)

Muy a menudo _____
A menudo _____
De vez en cuando _____
Alguna vez _____
Nunca _____

SECCIÓN II

Determine la calidad del audio en esta prueba

	SKYPE	***HANGOUTS***
Excelente	_____	_____
Buena	_____	_____
Regular	_____	_____
Mediocre	_____	_____
Mala	_____	_____

Tuvo dificultad al oír algún tramo de la conversación:

En caso afirmativo indique cual dificulto percibió:	***SKYPE***	***HANGOUTS***
Volumen bajo	_____	_____
Ruido o zumbido	_____	_____
Distorsión	_____	_____
Interrupción	_____	_____
Eco	_____	_____
Voz robotizada	_____	_____
Otro. Especifique	_____	_____

ANEXO C

INSTALACION DE LOS PROGRAMAS

C.1 Instalación Skype

Ingresa a la página oficial de Skype: www.skype.com/, hacer click en la opción *Descargar* y escoge el equipo sobre el cual se realizara la instalación puede ser un móvil, tablet, pc de escritorio u otros. Para el presente ejemplo se realiza la instalación de Skype Versión 6.20.0.104 sobre una portátil con Windows 7. Una vez seleccionado el equipo, guarda el archivo de SkypeSetup.exe y luego ejecútalo.

La primera ventana del proceso de instalación contiene el enlace hacia las Condiciones de Uso de Skype y la Política de Privacidad de Skype, si deseas leerlo debes hacer click en el link. Adicionalmente te solicita el idioma sobre el cual deseas trabajar, escoge el idioma preferido y presiona Acepto Siguiente tal como se muestra en la figura C.1.

Figura C.1. Selección de Idioma de instalación de Skype.

Luego muestra una ventana consultando si deseas instalar Clic, el cual es un complemento para tu navegador Web que te permite llamar directamente a los números registrados en la Web a través de Skype. Para el desarrollo del presente proyecto omitimos la instalación de dicho complemente, quitamos el visto en dicha opción y presionamos el botón Continuar.

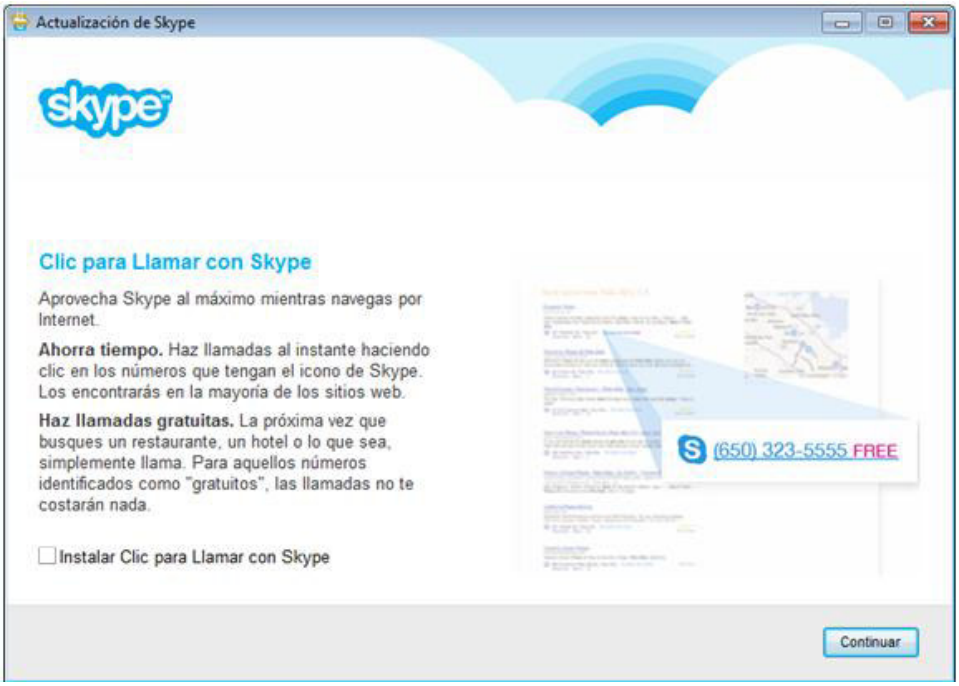


Figura C.2. Solicitud de instalación de clic para llamar con Skype.

Seguidamente aparece una ventana como en la figura C.3 solicitando la aceptación de registrar Bing como buscador web y MSN como página principal de tu navegador. Deseleccionar ambos campos y presionar el botón Continuar.

Figura C.3. Solicitud de registrar Bing y MSN en tu navegador.

Finalmente inicia el proceso de instalación de Skype como se muestra en la figura C.4 el cual puede tardar unos minutos de acuerdo a las características de tu equipo.

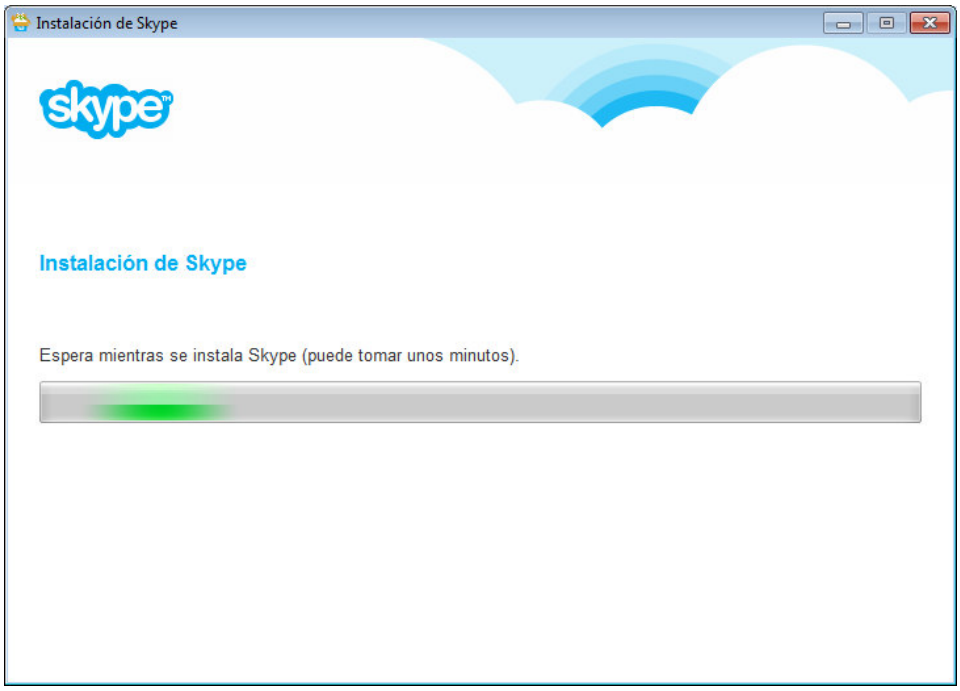


Figura C.4. Proceso de instalación de Skype.

Una vez instalado se muestra la pantalla de inicio de sesión de Skype como en la figura C.5. Ingresa con tu cuenta de Hotmail, Facebook o Skype que tengas creada o genera una nueva cuenta.

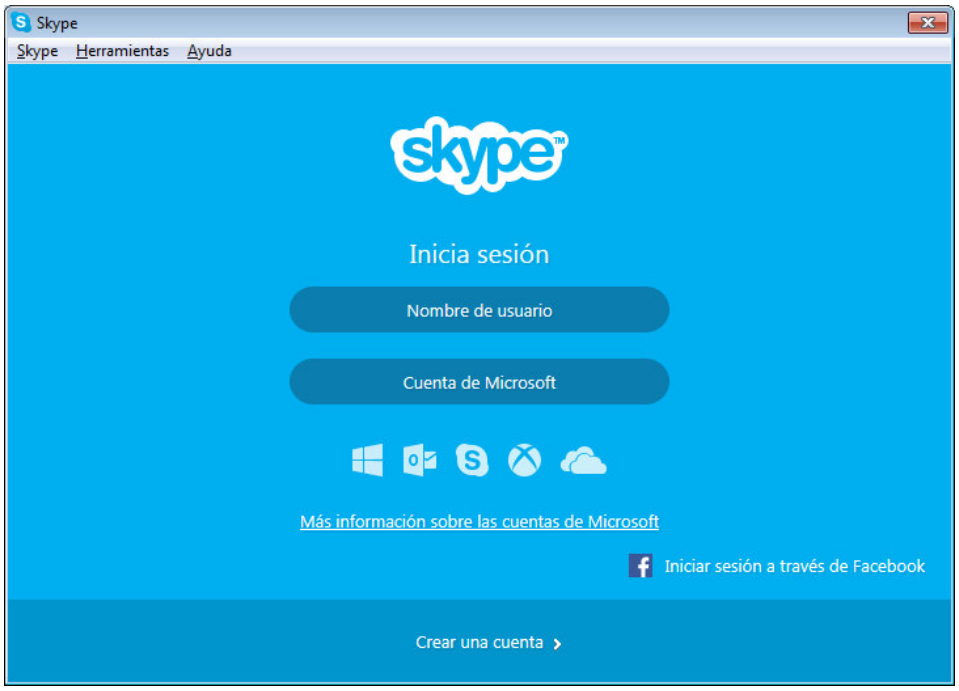


Figura C.5. Pantalla de Inicio de Sesión de Skype.

Para las presentes pruebas se debió realizar una configuración especial a nivel de entrada de audio en la aplicación como se registra en la figura C.6. Luego de haber iniciado sesión, nos dirigimos al menú principal y damos click en Herramientas y luego en Opciones. En el menú lateral izquierdo escogemos la opción Configuración de sonido y en el campo Entrada de Audio seleccionamos "Grabar reproducción". Presionamos el botón Guardar.

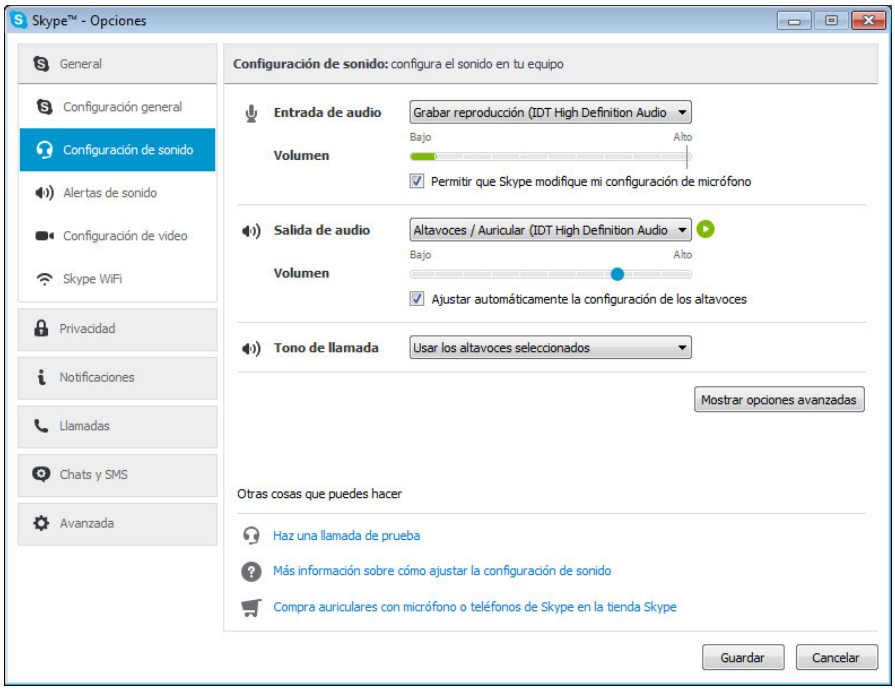


Figura C.6. Selección de entrada de Audio para pruebas.

C.2 Instalación Hangouts

Ingresa al siguiente enlace: http://www.google.com/intl/es/+/learnmore/hangouts/ donde puedes descargar la aplicación para móviles, equipos Apple o computadoras. Para el presente ejemplo se instalara en una portátil con Windows 7. Damos click en la opción “Disponible para tu computadora” y nos mostrar una pantalla consultando si añadimos al navegador Chrome y presionamos el botón Añadir tal como se muestra en la figura C.6.

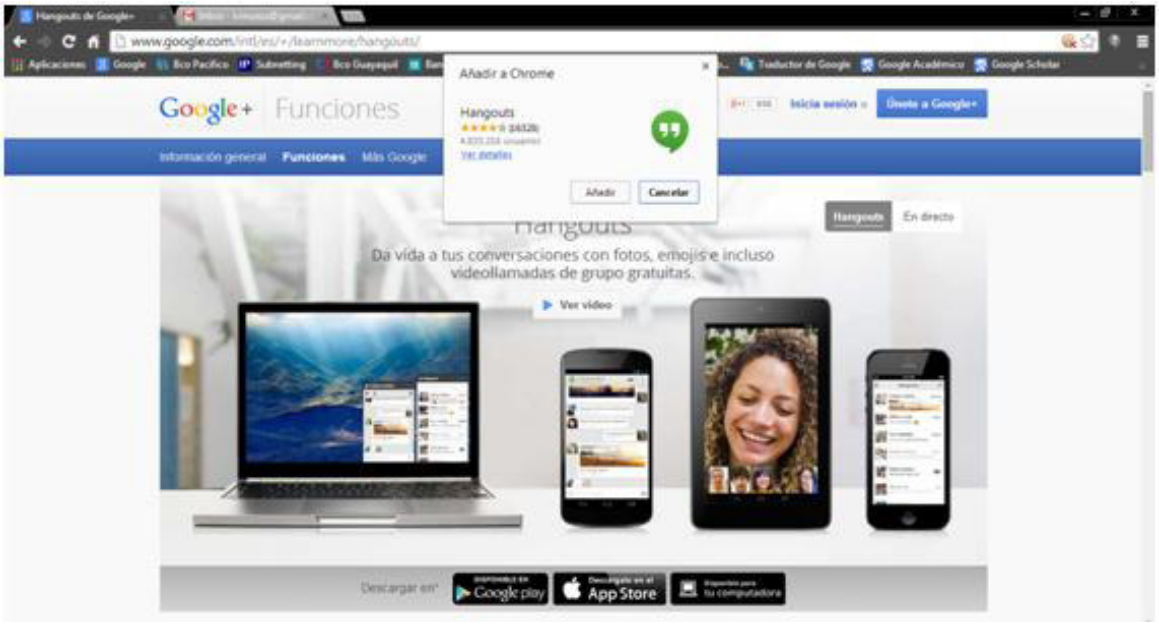


Figura C.7. Solicitud de instalación de Hangouts en Chrome.

Para verificar su correcta instalación nos dirigimos a la sección Complementos del navegador como se muestra en la figura C.8.

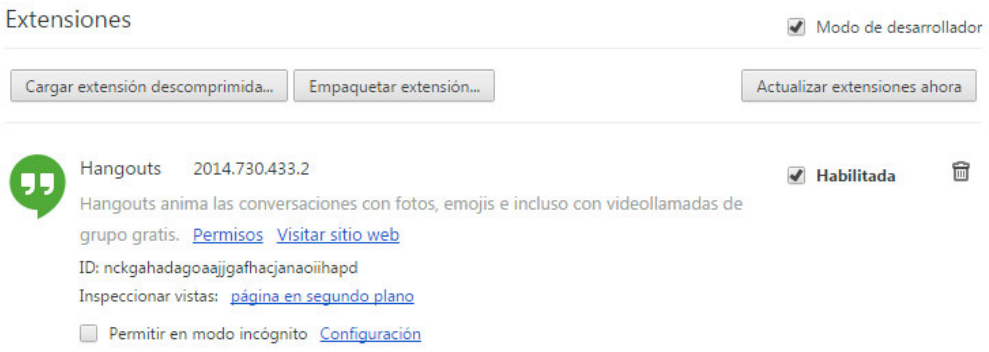


Figura C.8. Instalación de complemento Hangouts en Chrome.

Para iniciar sesión en la aplicación, se debe dar clic en el icono de Hangouts que se registra en la parte superior del navegador y aparece la ventana mostrada en la figura C.9. Dado que Hangouts es una aplicación propia de Google, se debe tener una cuenta registrada en cualquier aplicación de Google para poder acceder a este servicio.

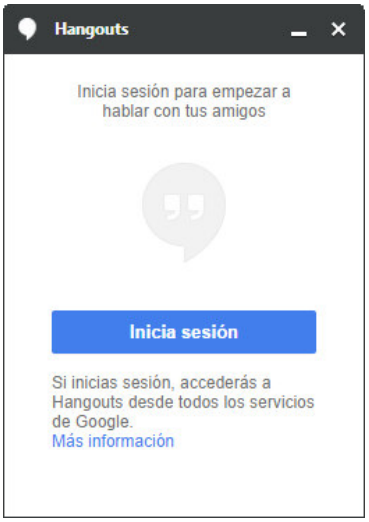


Figura C.9. Inicio de Sesión de Hangouts.

C.3 Instalación Wireshark

Ingresa a la página oficial de Wireshark: https://www.wireshark.org/ y luego a la sección de Descargas, escoge las características del equipo sobre el cual vas a trabajar y descarga el archivo Wireshark-winX-1.12.0.exe. Abre el archivo y ejecútalo directamente.

La primera ventana mostrada es la bienvenida por parte del fabricante, presionamos el botón Next de acuerdo a lo mostrado en la figura C.10.

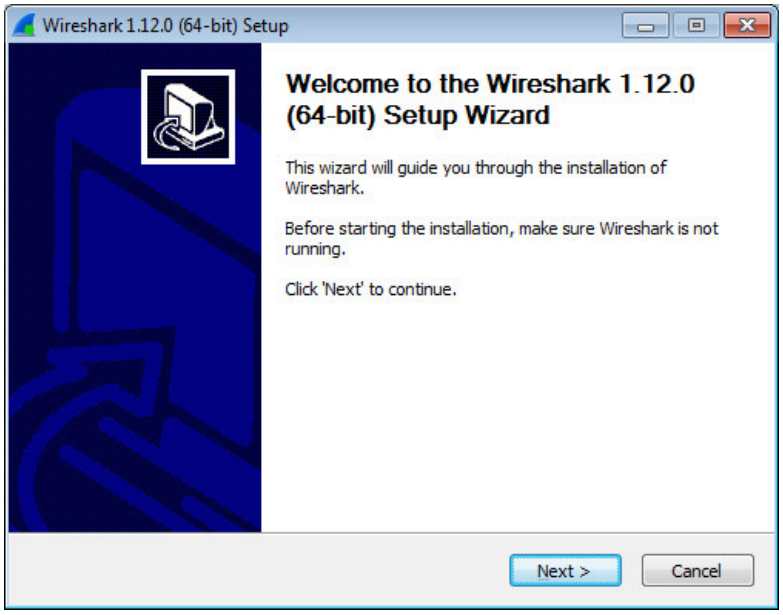


Figura C.10. Inicio de Wizard para instalar Wireshark.

Posteriormente nos aparece el Acuerdo de Licencia, hacemos click en I Agree.

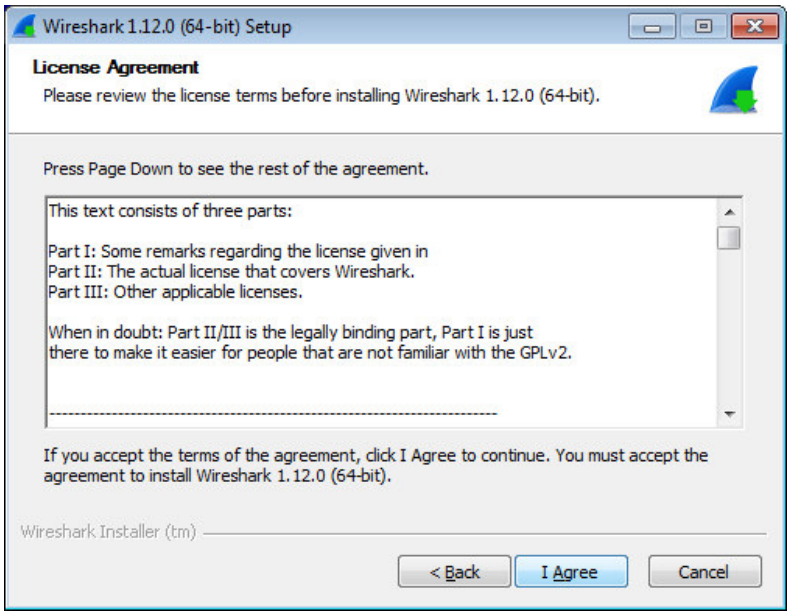


Figura C.11. Aceptación del acuerdo de licencia de Wireshark.

Luego nos solicita la elección de componentes y la selección de iconos y extensiones a instalar tal como se muestra en las figuras C.12 y C.13 respectivamente.

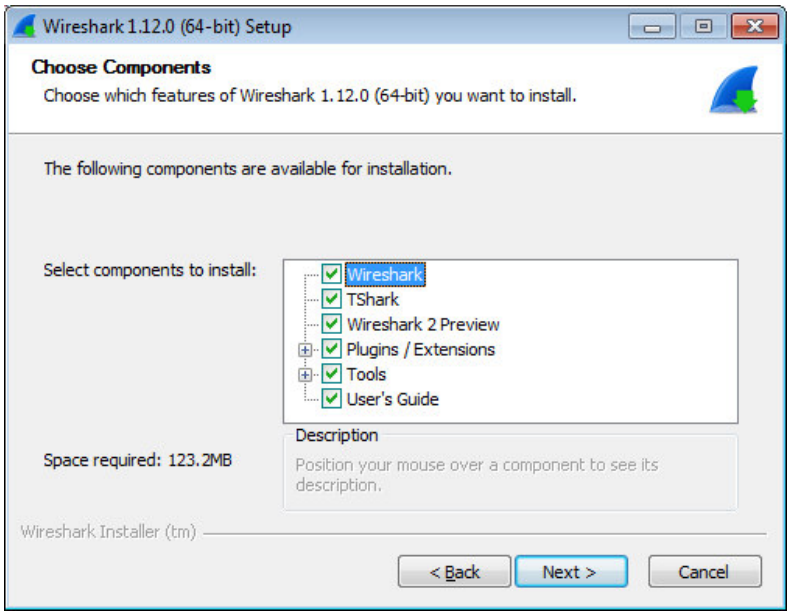


Figura C.12. Elección de componentes a instalar de Wireshark.

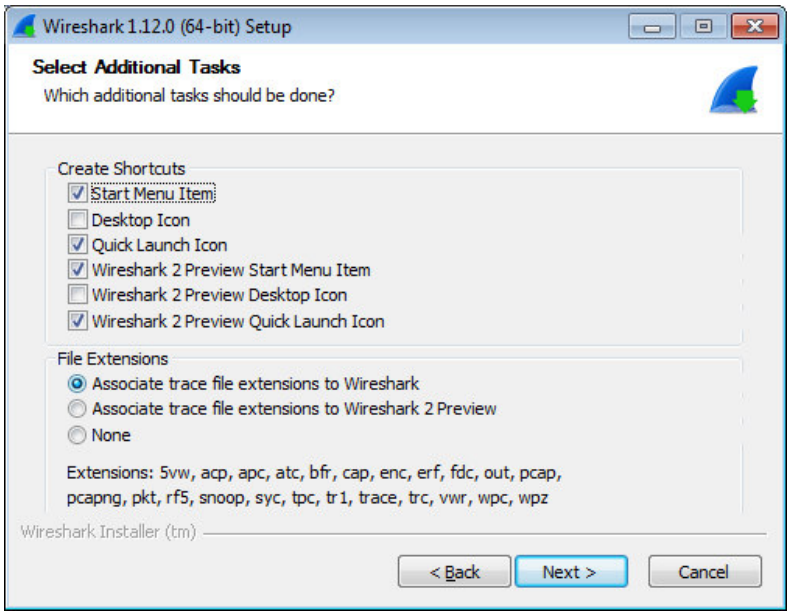


Figura C.13. Elección de extensiones e iconos de Wireshark.

En la figura C.14, nos solicita la selección del directorio donde deseamos la instalación del programa. Por defecto se instala en la ruta C:\Program Files\Wireshark. Presionamos el botón Next.

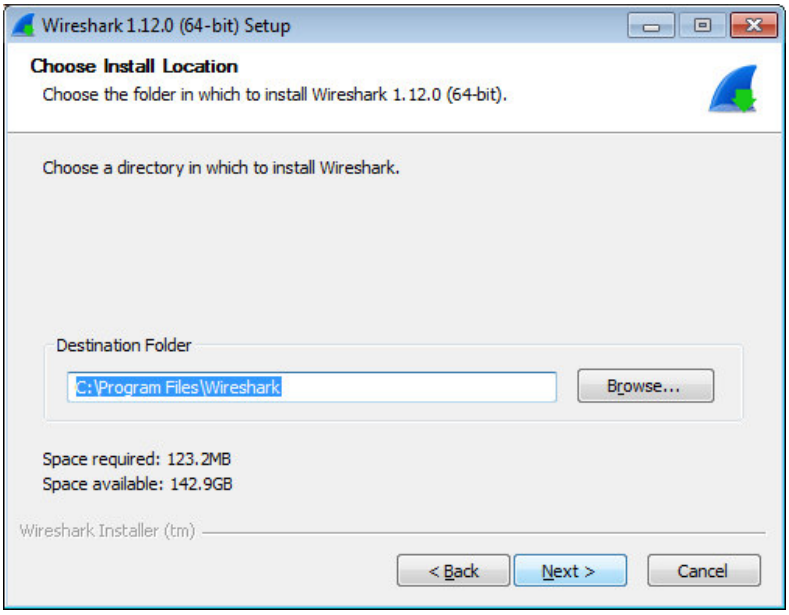


Figura C.14. Elección de directorio de instalación de Wireshark.

Posteriormente, nos solicita la confirmación de instalar Wincap como se muestra en la figura C.15. Wincap es la librería libpcap para Windows e incluye un controlador útil para la captura de paquetes en tiempo real. Damos clic en Install y se procederá a descargar todos los archivos necesarios para su instalación. Ver las figuras C.15, C.16.

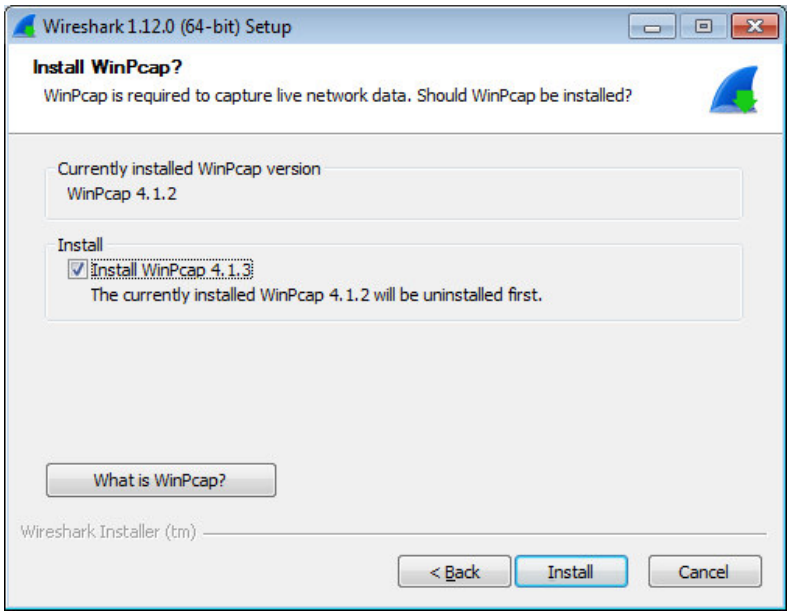


Figura C.15. Elección de instalación de Winpcap.

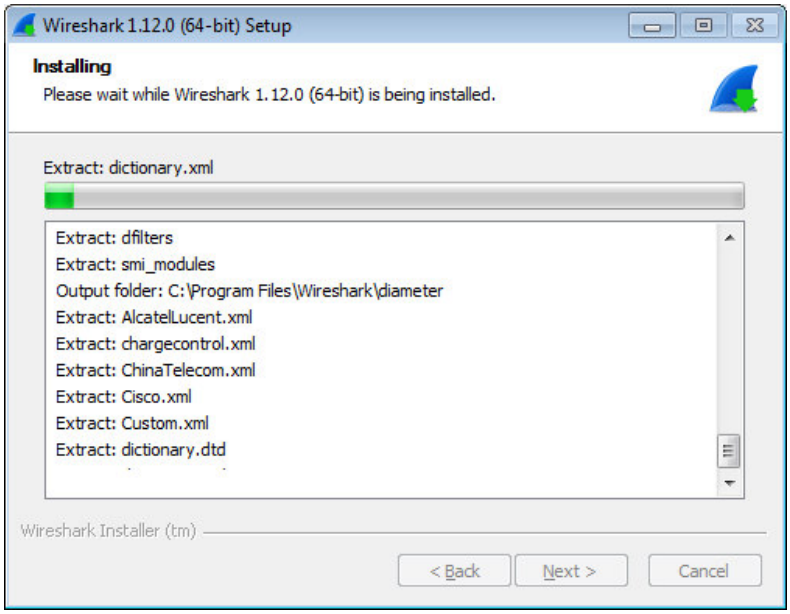


Figura C.16. Proceso de instalación de Winpcap.

Una vez descargado todos los archivos necesarios para su instalación, nos muestra la pantalla de Bienvenida de Winpcap, presionamos el botón Next como se visualiza en la figura C.17.

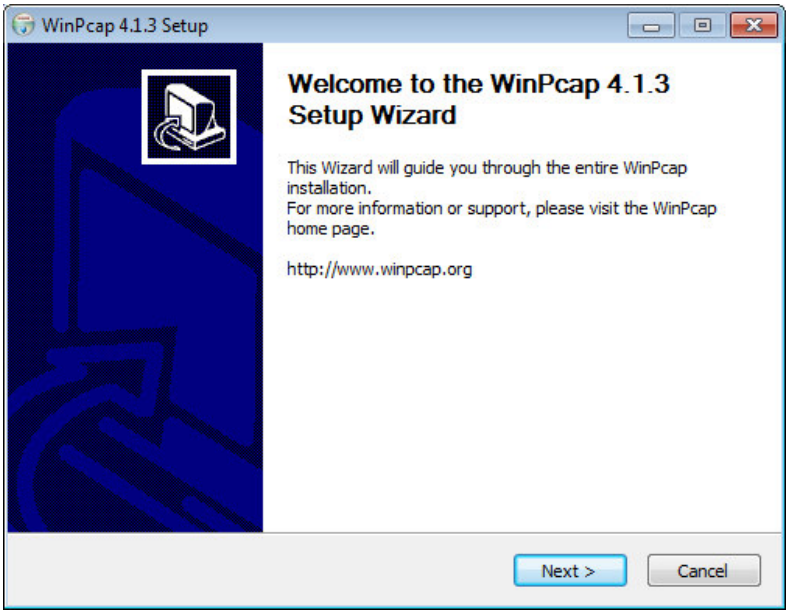


Figura C.17. Bienvenida de instalación de Winpcap.

Seguidamente Aceptamos el Acuerdo de Licencia de Winpcap dando clic en el botón I Agree. Finalmente nos aparece informa la culminación de instalación de Wincap. Ver figura C.18 y C.19.

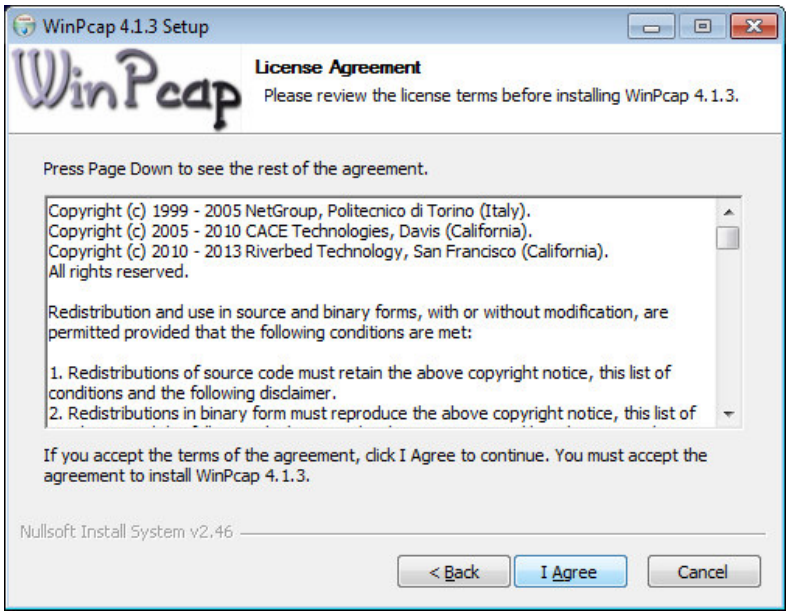


Figura C.18. Aceptación de acuerdo de licencia de Winpcap.

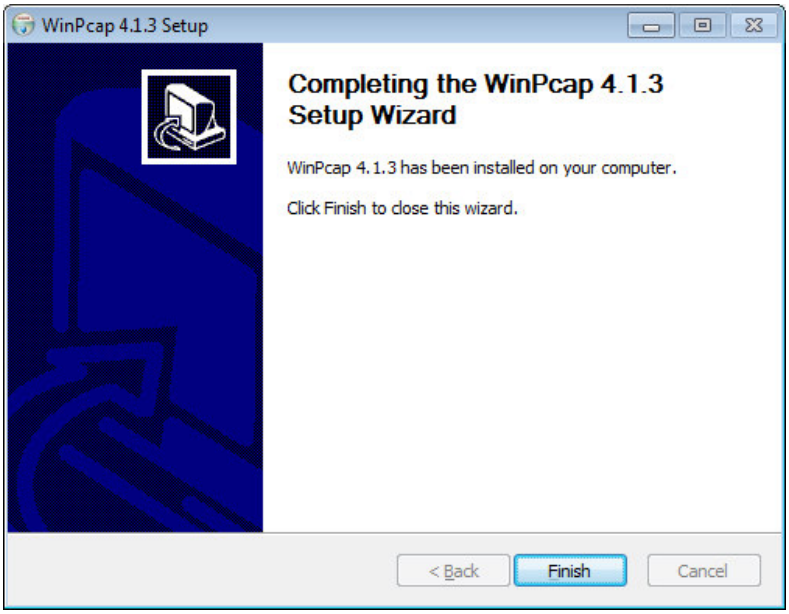


Figura C.19. Finalización de instalación de Winpcap.

Una vez culminado el proceso de instalación de Wincap, nuevamente redirecciona a la instalación de Wireshark, damos clic en Next y finaliza exitosamente la instalación. Presionamos el boton Finish.

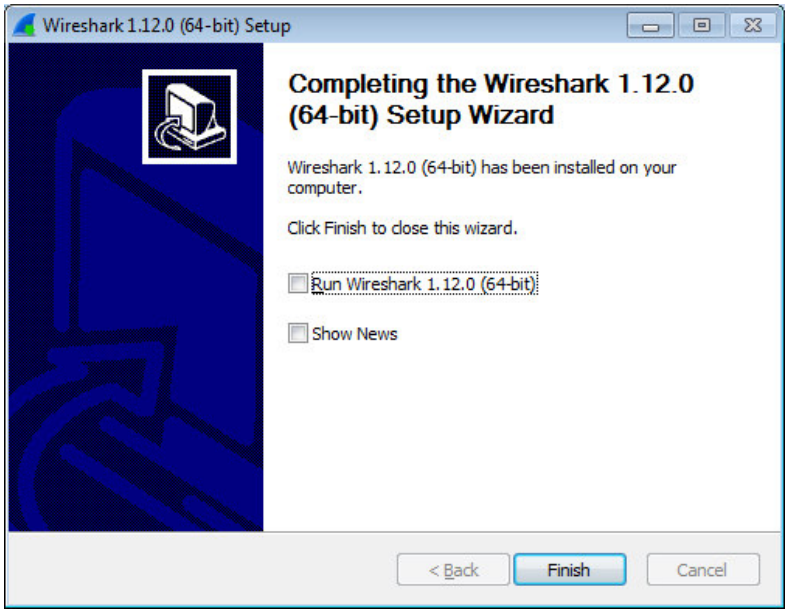


Figura C.20. Finalización de instalación de Wireshark.

BIBLIOGRAFIA

[1] F. Moreno, *Comunicación en tiempo real sobre Internet,* Buran, 2000.

[2] Advameg, Inc, «Comp.realtime: Frequently Asked Question (FAQS) (versio 3.5),» 1998. [En línea]. Available: http://www.faqs.org/faqs/realtime-computing/faq/. [Último acceso: 3 Abril 2015].

[3] D. Liu, X. Hu, M. Lemmon y L. Qiang, «Firm real-time system scheduling based on a novel QoS constraint,» Computers, IEEE Transactions on (Volume:55 , Issue: 3), 2006.

[4] T. McGraw−Hill, Broadband Telecommunication Handbook, The McGraw−Hill Companies, 2002.

[5] R. Horak, Telecommnunications and data communications handbok, New Jersey: John Wiley & Sons, Inc., Hoboken,, 2007.

[6] H. Schulzrinne, S. Casner, R. Frederick y V. Jacobson, «RTP A Transport Protocol for Real-Time Applications,» The Internet Society, 2003.

[7] C. Perkins, RTP Audio and Video for the Internet, Boston: Pearson Education, Inc., 2003.

[8] «Cacti - The Complete RRDT Tool-based Graphing Solution,» The Cacti Group, Inc, 2012. [En línea]. Available: http://www.cacti.net/. [Último acceso: 3 Abril 2015].

[9] «Wireshark. Go Deep,» Wireshark Foundation, [En línea]. Available: https://www.wireshark.org/. [Último acceso: 3 Abril 2015].

[10] Unión Internacional de Telecomunicaciones, «G.711,» Unión Internacional de Telecomunicaciones, Ginebra, 1993.

[11] Unión Internacional de Telecomunicaciones, «G.726,» Unión Internacional de Telecomunicaciones, Ginebra, 1990.

[12] Unión Internacional de Telecomunicaciones, «G.728,» 2013.

[13] Unión Internacional de Telecomunicaciones, «G.729,» 2012.

[14] I. T. Union, «ITU-T Publications and Recommedations,» ITU, 2008. [En línea]. Available: http://www.itu.int/rec/T-REC-G/en. [Último acceso: 15 09 2013].

[15] G. Camarillo, SIP Demystified, New York: McGraw-Hill, 2002.

[16] S. Kingham, «VoIP Workshop,» de *SIP Tutorial*, Terema, 2005.

[17] M. Spencer, B. Capouch, E. Guy, F. Miller y K. Shumard, «RFC 5456 IAX: Inter-Asterisk eXchange Version 2,» IETF, 2010.

[18] L. Zane y D. Himanshu, «IAX Voice Over-IP Security,» 2013. [En línea]. Available: http://www.isecpartners.com/. [Último acceso: Agosto 2013].

[19] F. Yergeau, «RFC 2279: UTF-8, a transformation format of ISO 10646,» 1998.

[20] M. Baugher, D. McGrew, M. Naslund, E. Carrara y K. Norrman, «The Secure Real-time Transport Protocol (SRTP),» The Internet Society, 2004.

[21] «Advanced Encryptions Standard (AES),» Federal Information Processing Standards, 2001.

[22] H. Krawczyk, M. Bellare y R. Canetti, «HMAC: Keyed-Hashing for Message Authentication,» The Internet Society, 1997.

[23] D. Eastlake y P. Jones, «US Secure Hash Algorithm 1 (SHA1),» The Internet Society, 2001.

[24] S. Kent y R. Atkinson, «Security Architecture for the Internet Protocol,» The Internet Society , 1998.

[25] ITU-Recomendación E.800, «Definiciones de términos relativos a la calidad,» 2008.

[26] ITU-T Recomendación P.10/G.100 Enmienda 2, «Enmienda 2: Nuevas definiciones para incluir en la Recomendación ITU-T P.10/G.100,» 2008.

[27] ITU-T Recomendación P.861, «Medición objetiva de la calidad de los códecs vocales de banda telefónica (300 - 3400 Hz),» 1996.

[28] R. Dai, «A Technical White Paper on Sage's PSQM Test,» 2000.

[29] ITU-T Recomendación P.862, «Evaluación de la calidad vocal por percepción: un método objetivo para la evaluación de la calidad volca de extremo a extremos de redes telefónicas de banda estrecha y códecs vocales.,» 2001.

[30] A. Rix y M. Hollier, «The perceptual analysis measurement system for robust end-to-end speech quality assessment,» Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on (Volume:3), Istanbul, 2000.

[31] ITU-T Recomendación G.108, «Aplicación del modelo E: Directrices para la planificación,» 1999.

[32] ITU-T Recommendation G.113, «Transmission impairments due to speech processing,» 2007.

[33] J. Chunlei y H. Peng , «Research of Monitoring VoIP Voice QoS,» International Conference on Internet Computing & Information Services (ICICIS), Hong Kong, 2011.

[34] H. Assem, D. Malone, J. Dunne y P. O'Sullivan, «Monitoring VoIP Call Quality Using Improved Simplified E-model,» International Conference on Computing, Networking and Communications, San Diego, CA, 2013.

[35] ITU-T Recomendacion G.109, «Definición de las categorías de calidad de transmisión vocal,» 1999.

[36] ITU-T Recommendation G.107, «The E-model: a computational model for use in transmission plannig,» Geneva, 2011.

[37] S. Basterrech, G. Rubino y M. Varela, «Single-sided Real-time PESQ Score Estimation,» 2012.

[38] ITU-T Recomendación P.862.1, «Función de correspondencia para convertir los resultados brutos de la prueba P.862 en nota media de opinión de la calidad de escucha objetiva,» 2003.

[39] P. Partilla, M. Kohut, M. Voznak, M. Mikulec, J. Safarik y K. Tomala, «A methodology for measuring voice quality using PESQ and interactive voice response in the GSM channel designed by OpenBTS,» 2010.

[40] T. Fukumori, M. Nakayama, T. Nishiuri y Y. Yamashita, «Estimation of speech recognition performance in noisy and reverberant environments using PESQ score and acoustic parameters,» 2010.

[41] M. Voznak y J. Rozhon, «Automated Speech Quality Monitoring Tool based on Perceptual Evaluation,» Ostrava, 2011.

[42] ITU-T Recomendación P.800, «Metodos de determinación subjetiva de la calidad de transmisión,» 1996.

[43] T. Hobfeld, R. Schatz y S. Egger, «SOS: The MOS is not enough!,» Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on, Mechelen, 2011.

[44] K. Hyun-Jong, D. H. Lee, J. M. Lee, K.-H. Lee, W. Lyu y S.-G. Choi, «The QoE Evaluation Method through the QoS-QoE Correlation Model,» de

Networked Computing and Advanced Information Management, 2008. NCM '08. Fourth International Conference on (Volume:2), Gyeongju, 2008.

[45] F. Aqboma y A. Liotta, «QoE-aware QoS management,» de *MoMM '08 Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, 2008.

[46] T. Hossfeld, P. Tran-Gia y M. Fiedler, «A generic quantitative relationship between quality of experience and quality of service,» Network, IEEE (Volume:24 , Issue: 2), 2010.

[47] D. Adami, C. Callegari, S. Giorda, M. Pagano y T. Pepe, «Skype-Hunter: A real-time system for the detection and classification of Skype traffic,» International Journal of Communication Systems, 2011.

[48] D. Bonfiglio, M. Mellia, M. Meo y D. Rossi, «Detailed Analysis of Skype Traffic,» IEEE, 2009.

[49] S. Del Rio, J. Ramos, J. Garcia-Dorado, J. Aracil, A. Cuadra-Sanchez y M. Cutanda-Rodriguez, «On the processing time for detection of Skype traffic,» IEEE, 2011.

[50] S. Del Río, J. Aracil, D. Corral y J. Garcia-Dorado, «On the impact of packet sampling pn Skype traffic classification,» Universidad Autonoma de Madrid, Spain, 2012.

[51] Instituto Nacional de Estadisticas y Censos INEC, «Resultados del Censo 2010 de población y vivienda en el Ecuador. Fasciculo provincial Guayas.,» Guayaquil, 2010.

[52] Harman International Company, *PERCEPTION 420 - User Instructions,* Harman International Company, 2015.

[53] Avid Technology, *ProjectMix I/O. Conrol Surface with Motorized Faders and 18x14 Audio Interface. User Guide,* Avid Technology, 2005.

[54] Avid Technology, *Pro Tools 10 Software. Professional audio recording and music creation software,* Avid Technology Inc, 2011.

[55] T. Dierks y E. Rescorla, «The Transport Layer Security (TLS) Protocol Version 1.2,» IETF, 2008.

[56] Microsoft, «Microsoft,» Microsoft, 2015. [En línea]. Available: http://www.microsoft.com/es-EC/default.aspx. [Último acceso: 3 abril 2015].

[57] «Akamai Technologies,» Akamai Technologies, 2015. [En línea]. Available: http://spanish.akamai.com/enes/. [Último acceso: 3 Abril 2015].

[58] «SSL Certificate, EV SSL, Wildcard SSL, and Code Signing - GeoTrust,» GeoTrust, Inc, 2015. [En línea]. Available: https://www.geotrust.com/. [Último acceso: 3 Abril 2015].

[59] E. Rescorla, «TLS Elliptic Curve Cipher Suites with SHA-256/384 and AES Galois Counter Mode (GCM),» IETF, 2008.

[60] S. Blake-Wilson, N. Bolyard, V. Gupta, C. Hawk y B. Moeller, «Elliptic Curve Cryptography (ECC) Cipher Suites for Transport Layer Security (TLS),» IETF, 2006.

[61] B. Kaliski y J. Staddon, «PKCS #1: RSA Cryptography Specifications Version 2.0,» IETF, 1998.

[62] P. Chown, «Advanced Encryption Standard (AES) Ciphersuites for Transport Layer Security (TLS),» IETF, 2002.

[63] S. Kelly y S. Frankel, «Using HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512 with IPsec,» IETF, 2007.

[64] X. Yang, Y. Chenguang, L. Jingjiang y L. Young, «Video Telephony for End-Consumers: Measurement Study of Googlw+,iChat, and Skype,» IEEE, 2014.

[65] D. McGrew y J. Viega, «The Galois/Counter Mode of Operation (GCM),» 2013.

1