Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Mecánica Eléctrica

ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS

Paolo Renato Bonilla Duarte

Asesorado por la Inga. Ingrid Salomé Rodríguez de Loukota

Guatemala, abril de 2017

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

PAOLO RENATO BONILLA DUARTE
ASESORADO POR LA INGA. INGRID SALOMÉ RODRÍGUEZ DE LOUKOTA

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO EN ELECTRÓNICA

GUATEMALA, ABRIL DE 2017

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Pedro Antonio Aguilar Polanco
VOCAL I	Ing. Ángel Roberto Sic García
VOCAL II	Ing. Pablo Christian de León Rodríguez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Jurgen Andoni Ramírez Ramírez
VOCAL V	Br. Oscar Humberto Galicia Nuñez
SECRETARIA	Inga. Lesbia Magalí Herrera López

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. Carlos Eduardo Guzmán Salazar
EXAMINADORA	Inga. Ingrid Salomé Rodríguez de Loukota
EXAMINADOR	Ing. Romero Neftalí López Orozco
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería Mecánica Eléctrica, con fecha 15 mayo de 2014.

Paolo Renato Bonilla Duarte

Guatemala 21 de noviembre de 2016

Ingeniero
Carlos Eduardo Guzmán Salazar
Coordinador del Área de Electrónica
Escuela de Ingeniería Mecánica Eléctrica
Facultad de Ingeniería, USAC.

Estimado Ingeniero Guzmán.

Me permito dar aprobación al trabajo de graduación titular: ¨**Estudio sobre la vulnerabilidad de los servicios de telefonía sobre IP inteligentes a través de internet y su repercusión sobre las redes de telefonía por conmutación de circuitos**¨, del señor Paolo Renato Bonilla Duarte, por considerar que cumple con los requisitos establecidos.

Por tanto, el autor de este trabajo de graduación y, yo, como su asesora, nos hacemos responsables por el contenido y conclusiones del mismo.

Sin otro particular, me es grato saludarle.

Atentamente,

Inga. Ingrid Rodríguez de Loukota
Colegiada 5,356
Asesora

Ingrid Rodríguez de Loukota
Ingeniera en Electrónica
colegiado 5356

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA

FACULTAD DE INGENIERIA

REF. EIME 07. 2017.
Guatemala, 23 de ENERO 2017.

Señor Director
Ing. Francisco Javier González López
Director Escuela de Ingeniería Mecánica Eléctrica
Facultad de Ingeniería, USAC.

Señor Director:

Me permito dar aprobación al trabajo de Graduación titulado: ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS, del estudiante Paolo Renato Bonilla Duarte , que cumple con los requisitos establecidos para tal fin.

Sin otro particular, aprovecho la oportunidad para saludarle.

Atentamente,
ID Y ENSEÑAD A TODOS

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DIRECCION ESCUELA DE INGENIERIA MECANICA ELECTRICA
FACULTAD DE INGENIERIA

Ing. Carlos Eduardo Guzmán Salazar
Coordinador Área Electrónica

sro

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA

FACULTAD DE INGENIERIA

REF. EIME 07. 2017.

El Director de la Escuela de Ingeniería Mecánica Eléctrica, después de conocer el dictamen del Asesor, con el Visto Bueno del Coordinador de Área, al trabajo de Graduación del estudiante; PAOLO RENATO BONILLA DUARTE, titulado: ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS, procede a la autorización del mismo.

Ing. Francisco Javier González López

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DIRECCION ESCUELA DE INGENIERIA MECANICA ELECTRICA
FACULTAD DE INGENIERIA

GUATEMALA, 7 DE MARZO 2,017.

Universidad de San Carlos
de Guatemala

Facultad de Ingeniería
Decanato

DTG. 189.2017

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería Mecánica Eléctrica, al Trabajo de Graduación titulado: **ESTUDIO SOBRE LA VULNERABILIDAD DE LOS SERVICIOS DE TELEFONÍA SOBRE IP INTELIGENTES A TRAVÉS DE INTERNET Y SU REPERCUSIÓN SOBRE LAS REDES DE TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS,** presentado por el estudiante universitario: **Paolo Renato Bonilla Duarte,** y después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:

Ing. Pedro Antonio Aguilar Polanco
Decano

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DECANO
FACULTAD DE INGENIERIA

Guatemala, abril de 2017

/gdech

ACTO QUE DEDICO A:

Dios	Por guiar cada paso de mi vida y por bendecirme con este logro.
Mi madre	Blanca Leticia Duarte, por sacrificar todo por el bien de sus hijos. Por ser un ejemplo de vida y una madre espectacular.
Mi padre	Jorge Luis Bonilla, por apoyarme en todo momento y estar allí cuando más se necesita.
Mi hermana	Darlene Bonilla, por su apoyo incondicional y por ser un ejemplo de dedicación.
Mis abuelos	Por su amor y consejos, por ser una influencia importante en mi vida, por cuidarme y protegerme
Mis tíos	Por su cariño y por ser influencia en mi carrera.
Mis primos	Por su cariño y por una ser influencia en mi carrera, por estar ahí cuando se les necesita.

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por proveer un ambiente de estudios y desarrollo profesional.
Facultad de Ingeniería	Por ser una importante influencia en mi carrera y por proveer conocimientos y experiencias. Por su dedicación y profesionalismo.
Mis amigos de ingeniería electrónica	Por su amistad, por estar siempre en las buenas y en las malas y por toda su ayuda para finalizar la carrera.
Mis compañeros de ingeniería electrónica	Por compartir todos los conocimientos y por su apoyo para finalizar la carrera.
Mis amigos de la facultad	Por su cariño y por siempre apoyarme e incentivarme para cumplir mis metas.
Mi asesora de tesis	Inga. Ingrid Rodríguez de Loukota por su valioso tiempo y apoyo para finalizar mi trabajo de graduación.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

FIGURAS

TABLAS

GLOSARIO

1G	Abreviación de la primera generación de la telefonía móvil.
2G	Abreviación de la segunda generación de la telefonía móvil.
3G	Abreviación de la tercera generación de la telefonía móvil.
3GPP	Asociación que se encarga de definir estándares para que la integración de la tecnología móvil sea eficiente.
4G	Abreviación de la cuarta generación de la telefonía móvil.
AMPS	*Advanced mobile phone system,* por sus siglas en inglés, modelo análogo de transmisión de datos utilizada en la primera generación de la telefonía móvil.
BSC	*Base station controller,* estación base que controla a varias BTS de una red 2G

BTS	*Base transceiver station,* por sus siglas en inglés, es el equipo que es utilizado para la comunicación bidireccional entre uno o varias terminales móviles.
CAMEL	*Customized applications for mobile networks enhanced logic**,*** por sus siglas en inglés, es una serie de estándares basados en la arquitectura IN, este protocolo se utiliza para proveer servicios adicionales a los de la telefonía tradicional.
Carrier	Se le denomina así a una compañía autorizada para proveer servicios de telefonía.
CDMA	*Code division multiple access,* tecnología que permite la transmisión de paquetes simultáneos utilizando un mismo canal.
CDR	*Call detail record,* registros donde quedan almacenadas todas las llamadas de voz y datos.
DIAMETER	Protocolo de autenticación y autorización confiable en los protocolos de transporte TCP y UDP.
DNS	*Domain name system,* por sus siglas en inglés, es un sistema de nomenclatura jerárquico para dispositivos conectados a redes IP como Internet.

DoS	*Denial of service,* por sus siglas en inglés, es un tipo de ataque donde el perpetrador busca deshabilitar un equipo o servicio de datos o telefonía.
CSFB	*CS fallback,* por sus siglas en inglés, se le llama así al procedimiento de pasar de una red LTE hacia una red 3G cuando se utilizan servicios de la red conmutada por circuitos.
EDGE	*Enhanced data rates for GSM evolution,* por sus siglas en inglés, tecnología celular que permite improvisar velocidades de datos mayores y compatibles con GSM.
EIR	*Equipment identity register,* por sus siglas en inglés, es una base de datos que contiene la información de los subscriptores que pueden o no tener acceso a una red de telefonía o datos. .
ENUM	*E.164 number to URI mapping,* por sus siglas en inglés, es un protocolo el cual traduce números telefónicos a direcciones de Internet.
FDMA	*Frequency division multiple access,* tecnología basada en división de canales por frecuencia.
Firewall	Es un elemento de una red cuya función es proteger los bordes entre dos o más redes.

GPRS	*General packet radio service,* por sus siglas en inglés, sistema utilizado para la transmisión de datos mediante la conmutación de paquetes.
GSM	*Global system for mobile communications,* estándar desarrollado para describir protocolos de segunda generación .
Handover	En telecomunicaciones, transferencia de la responsabilidad de una llamada o una sesión de datos activa y establecida en una radio base hacia otra sin tener una caída del servicio.
HLR	*Home location register,* por sus siglas en inglés, equipo que almacena toda la información de los usuarios de una red.
HSS	*Home subscriber server,* por sus siglas en inglés, es un elemento de red que contiene la base de datos de usuarios de una red IMS o LTE.
HTTP	*Hypertext transfer protocol,* por sus siglas en inglés, protocolo de aplicación que se utiliza para la distribución de sistemas de información de hipermedia.
IMEI	*International mobile station equipment identity,* por sus siglas en inglés, es un código único que identifica a una terminal móvil a nivel mundial.

IMS	*IP multimedia subsystem,* por sus siglas en inglés, es una arquitectura de red integrada sobre sobre protocolo de Internet para ofrecer servicios multimedia.
IN	*Intelligent network,* por sus siglas en inglés, es un estándar de arquitectura de red que permite a los operadores proveer servicios de valor agregado.
IP	*Internet protocol,* protocolo de transporte encargado de entregar paquetes desde un equipo origen hacia un equipo destino utilizando solamente las direcciones IP que están en el encabezado.
IPRN	I*nternational premiun rate number,* por sus siglas en inglés, son números telefónicos en los cuales se ofrecen distintos tipos de servicios los cuales se cobran a una tasa más alta que una llamada habitual.
IPSEC	I*nternet protocol security,* por sus siglas en inglés, es un conjunto de protocolos cuya funcion es asegurar las comunicaciones sobre IP.
IRSF	*International revenue sharing fraud,* por sus siglas en inglés, es un tipo de fraude que se da sobre las comunicaciones VoIP, consiste en desviar tráfico ilegalmente hacia números IPRN.

LTE	*Long term evolution,* nombre con el que se le conoce a la red de datos 4G y está basado en un estándar para comunicación de alta velocidad en redes inalámbricas.
MAP	*Mobile application part,* por sus siglas en inglés, es un protocolo de señalización que permite interconectar varios elementos dentro de una red móvil, el principal es el HLR.
MITM	*Man in the midle,* por sus siglas en inglés, se le denomina así a un atacante que se coloca entre la red de acceso y la red del operador con fines fraudulentos.
MSC	*Mobile switching centre,* equipo que se encarga de controla la movilidad de la red y de direccionar las llamadas de voz y mensajería de texto de los usuarios.
MSISDN	*Mobile station integrated service digital network,* por sus siglas en inglés, nombre que se le da al compuesto del código de país más el número móvil del usuario.
NGN	*Next generation network,* por sus siglas en inglés, es un término que se refiere a la evolución de la infraestructura de redes de servicio y acceso telefónico hacia el mundo IP.

PBX	*Private branch exachange,* por sus siglas en inglés, se le denomina así a cualquier central telefónica que se conecta a la red pública telefónica.
RTP	*Real-time transport protocol,* por sus siglas en inglés, es un protocolo de la capa de sesión utilizado para la transmisión de información en tiempo real.
SBC	*Session boarder controller,* por sus siglas en inglés, son equipos que se conectan en el borde de una red, son capaces de derivar todo el tráfico VoIP proveyendo niveles de seguridad necesarios para evitar distintos tipos de ataques.
SIGTRAN	*Signaling transfer network,* por sus siglas en inglés, protocolo encargado de transferir señalización *SS7* sobre la red IP del operador.
SIM	*Subscriber identification module,* circuito integrado que guarda la identidad de un suscriptor móvil, este circuito es insertado en la terminal para tener servicio de la red del operador.
SIM Box	Se le denomina así al dispositivo que se utiliza para terminar llamadas internacionales fraudulentas haciéndolas pasar como llamadas nacionales.

SIP	*Session initiation protocol,* por sus siglas en inglés, es un protocolo utilizado para controlar sesiones multimedia como VoIP.
SMS	*Short message service, por sus siglas en inglés,* acrónimo que hace referencia al servicio de mensajería corta.
SPIT	*Spam over internet telephony*, es un tipo de ataque que sucede sobre VoIP en el cual se envian menjases multimedia indeseados a los usuarios.
SRTP	*Secure real-time transport protocol,* por sus siglas en inglés, es una mejora al protocolo RTP ya que le agrega cifrado de los paquetes de media.
SS7	*Signaling system núm. 7,* por sus siglas en inglés, conjunto de protocolos de señalización que son utilizados en la mayoría de redes de conmutación para establecer las llamadas de voz.
TCP	*Transmission control protocol,* se utiliza junto con el protocolo IP para el envío y rastreo de paquetes a través de una red de transporte.
TLS	*Transport layer security,* por sus siglas en inglés, es un protocolo de ciframiento utilizado en redes IP como internet para proveer seguridad.

UDP	*User datagram protocol,* parte del protocolo IP y al igual que *TCP* funciona sobre la capa de transporte.
UMTS	*Universal mobile telecommunications system,* por sus siglas en inglés, sistema basado en el estándar *GSM* y está orientado a la tecnología 3G de la telefonía móvil.
USIM	*Universal subscriber identity module,* aplicación de software para la telefonía UMTS análogo a lo que SIM es para *2G.*
UTRAN	*UMTS terrestrial radio access network,* denominación de la red de acceso para 3G.
Vishing	Es un tipo de ataque fraudulento en el cual se utiliza ingeniería social para obtener información personal no autorizada de un usuario para realizar fraudes.
VLR	*Visitor location register,* base de datos que contiene la información de suscriptores *roaming*.
VoIP	Voz sobre IP.
VoLTE	*Voice over LTE,* por sus siglas en inglés, es un conjunto de tecnologías utilizadas para mejorar la calidad de audio y tiempo de conexión de las llamadas telefónicas, las cuales se realizan sobre la red de datos y no sobre la red conmutada.

VPN	*Virtual private network,* por sus siglas en inglés, permite extender una red privada a través de una red de acceso público como Internet.
WCDMA	*Wideband code division multiple access,* por sus siglas en inglés, interfaz de aire estándar para 3G que permite la transmisión de paquetes de datos a una velocidad superior a la normal.

RESUMEN

La evolución de la telefonía convencional a la telefonía IP ha generado un negocio viable para los proveedores de servicios de telefonía móvil y fija, por su bajo costo de operación y por los múltiples servicios adicionales que ofrece, lo cual ha llevado a que cada vez más personas y empresas adquieran estos servicios.

Este crecimiento en los últimos años ha generado que sea más común que usuarios y operadores sean objeto de ataques que pueden generar pérdidas millonarias por la manipulación o interrupción de servicio y a la vez afectar la confianza del subscriptor en el servicio que ofrece el operador.

Debido a que la red de transmisión utilizada para los servicios de telefonía IP es Internet, los ataques son generados por personas que tienen acceso a una interfaz entre los subscriptores y dispositivos que ofrecen el acceso al servicio o en los bordes de las redes de transmisión de tráfico internacional, ambos a través de Internet.

Los atacantes pueden utilizan diversos métodos y aplicaciones para realizar ingresos no autorizados a la red de operador para realizar fraudes, generar ataques para denegar el servicio, atacar los dispositivos que ofrecen servicio y a la vez atacar las redes adyacentes de telefonía convencional.

Es entonces de sumo interés para operadores y usuarios proveer y exigir altas políticas de seguridad sobre cualquier servicio de telefonía inteligente con

el fin de proteger la identidad de los subscriptores y asegurar la continuidad del servicio por parte de los operadores.

El objetivo del presente trabajo de graduación es estudiar las vulnerabilidades de los protocolos por los cuales se rige la telefonía IP y exponer la repercusión que tienen las bajas o nulas políticas de seguridad en las redes de telefonía actuales. Exponer los métodos para la protección de sesiones SIP generadas por subscriptores y también políticas de seguridad que se deben aplicar a los elementos de red que sirven de interconexión de operadores a través de Internet para prevenir ataques de cualquier tipo. Otro de los objetivos es establecer métodos para la detección de ataques y protocolos para contrarrestar los mismos.

OBJETIVOS

General

Proporcionar el conocimiento necesario para contrarrestar intrusiones en las redes de los proveedores de telefonía que puedan generar fraudes o interrupción en servicio.

Específicos

1. Dar a conocer cuáles son los componentes de una red de telefonía y los protocolos por la cual se rigen los servicios que provee.

2. Evidenciar cómo se generan los fraudes o ataques sobre las redes de los proveedores de telefonía.

3. Proveer métodos para la prevención de fraudes y ataques sobre los servicios de telefonía.

4. Proporcionar criterios para la detección de fraudes y ataques sobre las redes y servicios de telefonía.

INTRODUCCIÓN

La comunicación multimedia sobre IP crece exponencialmente debido a que es más flexible y provee más servicios comparada con la telefonía tradicional. La telefonía IP conocida como voz sobre IP (VoIP) es una de las tecnologías comercialmente más emergentes de la comunicación multimedia sobre IP ya que el canal de comunicación se establece comúnmente sobre Internet.

Esta flexibilidad es gracias al protocolo de inicio de sesión SIP, muchas aplicaciones basadas en sesiones como VoIP y el *streaming* de video han presentado un aumento en el número de usuarios y organizaciones que los utilizan. La razón principal de este aumento sin lugar a duda es el costo más bajo de una llamada IP, esto se debe a que se utiliza la misma red para transmitir voz y datos.

El crecimiento exponencial de la telefonía IP crea la necesidad de integrar los servicios IP a las redes de telefonía convencionales y a su vez de garantizar la compatibilidad y la alta disponibilidad de los servicios para todos los usuarios por parte de los proveedores de estos servicios.

El desafío más grande es establecer un diseño y despliegue de una red IP eficiente la cual tenga un alto rendimiento y un funcionamiento robusto, este desafío se debe a la arquitectura abierta de Internet y a los sofisticados métodos de ataques en contra de la infraestructura que proveen los servicios VoIP, los cuales exponen las vulnerabilidades de los protocolos utilizados para establecer las sesiones IP.

En los primeros capítulos de este trabajo de graduación se hace una introducción a las redes de telefonía por conmutación de circuitos y a las redes de telefonía IP, se describen los protocolos, su evolución y adaptación con el fin de exponer sus vulnerabilidades. Luego, se describen cuáles son los tipos de ataques que pueden existir en una red de telefonía y a su vez que consecuencias se derivan de los mismos, tanto para los proveedores de telefonía como para los subscriptores.

Seguido de esto se enfoca en los métodos para la prevención de estos ataques, iniciando desde el diseño de la red IP que incluyen protocolos de seguridad, elementos de red capaces de detectar y bloquear ataques, renovación de hardware y software periódicamente, hasta las políticas de seguridad de acceso y la capacitación del personal.

Para finalizar, se exponen estrategias para detección y acción contra ataques, ya que no siempre se puede contar con todos los requerimientos necesarios para prevenir los ataques en todo momento. Entre las estrategias se destaca el análisis de tráfico, verificación de indicadores de desempeño, análisis de registros detallados de llamadas y acciones para contrarrestar un ataque al momento de ser detectado.

1. TELEFONÍA POR CONMUTACIÓN DE CIRCUITOS

La telefonía y los teléfonos han cambiado dramáticamente desde que Antonio Meucci inventó el teletrófono, luego patentado como teléfono por Alexander Graham Bell. Desde ese entonces estas tecnologías han mejorado constante y exponencialmente.

En un inicio, la comunicación era totalmente analógica, es decir, que estaba basada en señales eléctricas y elementos mecánicos. Un micrófono capaz de convertir vibraciones sonoras en señales eléctricas y un parlante capaz de decodificar estas señales para transformarlas en vibraciones sonoras, es decir, transductores electroacústicos.

Hoy en día la voz es transmitida digitalmente, es decir, se toma una muestra de las señales eléctricas analógicas y luego se transmiten en forma de pulsos utilizando diferentes técnicas que con el tiempo han evolucionado con el fin de aprovechar más eficientemente el medio de transmisión y asegurar la correcta recepción y confidencialidad de la información.

1.1. Origen de las telecomunicaciones y la telefonía

La primera red de telecomunicación aparece en Francia como la telegrafía óptica, que permite sustituir la comunicación por medios escritos, con esto se logró transmitir cualquier tipo de mensaje utilizando la luz. El inventor de este telégrafo fue Claude Chappe, quien diseñó la primera red óptica mecánica cuyos componentes consistían en una barra perpendicular con dos brazos móviles fijados.

Estos brazos se podían combinar en 196 distintas posiciones, es decir, 196 distintas figuras. En un inicio, cada una de estas combinaciones representaba una sílaba, luego este sistema evolucionó y se sustituyó por otro basado en un manual de 92 páginas y el mismo número de símbolos, se transmitían 2 posiciones, cada una de estas representaba un número de página y un número de posición, esta combinación representaba una palabra.

Gracias a los descubrimientos de Faraday y Ampere en relación a la corriente eléctrica y los campos magnéticos, nace el telégrafo eléctrico. Es un dispositivo capaz de enviar pulsos eléctricos a través de un circuito, estos pulsos son codificados como mensajes. El inventor del telégrafo eléctrico fue Joseph Henry en el año 1829. Con la ayuda del estadounidense Samuel Morse se le dio impulso luego de realizar la primera transmisión entre Baltimore y Washington en 1844.

El telégrafo eléctrico es alimentado por una batería, la cual interactúa con manipulador o conmutador para abrir o cerrar el circuito cuando este se presiona. El lado receptor tiene electroimán que atrae un punzón metálico con los pulsos eléctricos, dependiendo la longitud del pulso puede escribir una línea o un punto.

Una variación es colocar un dispositivo que emita algún sonido en vez de un punzón para escribir. Adicional a esto, Morse también invento un código basado en tres símbolos: un punto, una raya y un espacio, con los cuales se podían escribir palabras utilizando los primeros dos y el último para separarlas (código Morse).

1.1.1. Inicios del teléfono y las redes de telefonía

Alexander Graham Bell registró la patente del teléfono, el cual está basado en un principio denominado transmisión por resistencia variable, en un circuito eléctrico; el valor de la corriente depende de la resistencia, si la resistencia se hace variable la corriente que circula a través del mismo también variará. El mayor desafío fue lograr diseñar un transmisor y un receptor que unidos fueran capaces de transformar la voz en variaciones de corriente y viceversa.

El transmisor se basó en un tubo el cual conducía los sonidos hacia un disco que vibraba con la variación de los mismos, el disco a su vez tenía una parte móvil conectada a un compartimiento de gránulos de carbón, cuando el disco vibraba la parte móvil comprimía o descomprimía el compartimiento, se hacía pasar una corriente eléctrica a través de los gránulos la cual también variaba con los movimientos de la membrana, las variaciones de corriente simulan exactamente las variaciones de las ondas sonoras.

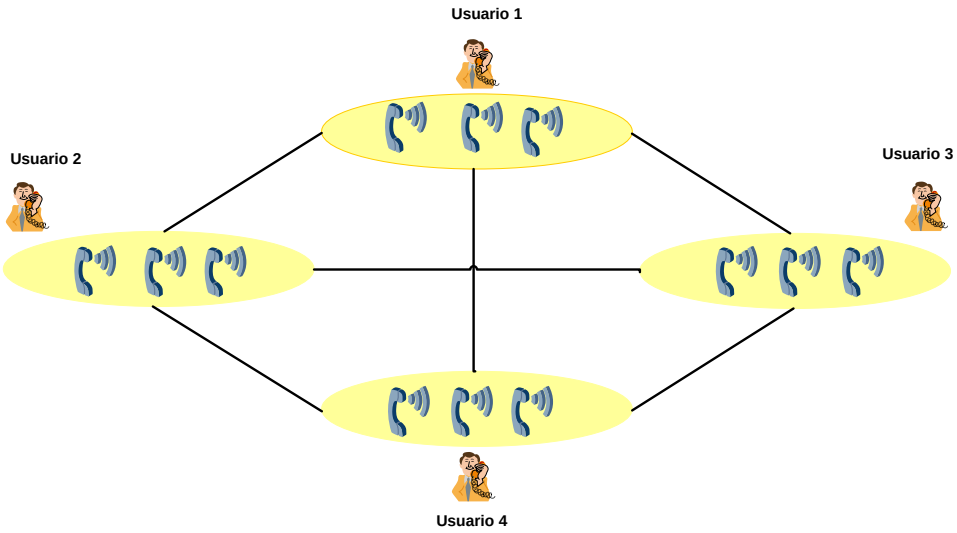
El receptor de igual forma lo integraba un tubo y un disco; el disco era atraído o repelido por las variaciones de corriente eléctrica que producían un campo magnético variable con el principio del electroimán, a diferencia de un imán que tiene un campo magnético que no puede ser variado, en el electroimán el campo magnético es inducido por la corriente eléctrica, al variar la misma también varía el campo magnético.

Hans Christian descubrió esta característica de los elementos conductores al ser recorridos por una corriente eléctrica en 1819. El disco al ser atraído o repelido producía las vibraciones para que el aire reprodujera el sonido transmitido en forma variaciones de corriente eléctrica.

En los inicios de la telefonía, la comunicación se realizaba uniendo los teléfonos directamente, es decir, con un teléfono se podía habla con una sola persona al ser unidos. Ahora bien, para poderse comunicar con tres personas eran necesarios seis cables y doce teléfonos en total, 3 por residencia.

Si se consideran X personas, el número de cables necesarios se obtienen interpolando, la fórmula será de $X \cdot (X-1)/2$; se observa que el aumento de cables es del orden de X^2. Asimismo, será necesario disponer de $N \cdot (N-1)$ aparatos telefónicos en total, por lo que los costos de instalación crecen de forma considerable a medida que se incrementa el número de usuarios, como se aprecia en la figura 1.

Figura 1. **Relación de la comunicación punto a punto**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

Para disminuir los costos de instalación, y considerando que en un momento dado solo se podía hablar con una persona a la vez, es posible

agregar un dispositivo conmutador con el cual se pueda seleccionar solo la línea con la cual se requiere comunicar, reduciendo a tan solo un teléfono y un conmutador por usuario.

Esta solución se puede mejorar considerablemente utilizando un solo conmutador. Se requiere entonces que todos los usuarios tengan una sola línea directa hacia el conmutador, de esta forma se reduce una relación lineal el número de líneas necesarias y de aparatos telefónicos, uno por residencia. Surge la necesidad de un nuevo elemento: un operador del conmutador; cada línea realizará una llamada al operador, y este será el encargado de conectar la llamada al destino solicitado

Con esta última solución se pasa de tener múltiples líneas residencias a una línea compartida entre varios usuarios lo cual aumenta considerablemente el rendimiento; hoy en día, aunque la tecnología sea diferente, se utiliza este mismo concepto. Los primeros tableros conmutadores tuvieron su origen en 1878 con una capacidad de 21 usuarios o abonados, los tableros de grande los de este tipo llegaron a tener una capacidad de 10 500 abonados.

Cuando el número de abonados crece y la distancia entre ellos empieza a considerarse un factor, es necesario unir diferentes tableros de conmutación. Estas líneas que unen distintos tableros son denominadas enlaces y a las líneas que unen a los teléfonos con sus tableros se les llama bucles de abonados.

1.2. Entidades de normalización

Con el crecimiento de la telefonía, aumentó también el número de compañías fabricantes de equipos las cuales fabricaban los mismos bajos sus

propios términos; es decir, 2 equipos de diferentes fabricantes podrían no poderse comunicar entre sí.

A raíz de esto surgen entidades cuyo único fin es normalizar la fabricación y definir estándares o recomendaciones de cómo estos deben funcionan, es decir, protocolos con el único fin de que exista compatibilidad en cualquier equipo de telefonía o telecomunicación.

1.2.1. Unión Internacional de Telecomunicaciones

Primero conocida como Unión Internacional de Telegrafía, fue fundada en París en al año 1865 con el fin de facilitar y regular las conexiones telegráficas internacionales y la interoperabilidad de las redes nacionales. En 1947 se declara como una agencia especializada de la Organización de las Naciones Unidas, para hacerse responsable de cualquier tema relacionado con la información y tecnologías de la comunicación.

La ITU, por sus siglas en inglés, International Telecommunication Union, está compuesta por tres sectores: radiocomunicación (ITU-R), normalización (ITU-T) y desarrollo de telecomunicaciones (ITU-D).

El sector de normalización divide cada recomendación por sectores, dentro las que son de interés para este trabajo de investigación:

- Serie E: operación general de redes, servicio de telefonía, operación del servicio y factores humanos.

- Serie G: medios y sistemas de transmisión, sistemas digitales y redes.

- Serie H: sistemas de media y audiovisuales.
- Serie Q: conmutación y señalización.
- Serie V: comunicación de datos sobre la red de telefonía.
- Serie X: redes de datos, sistemas abiertos de comunicación y seguridad.

Cada recomendación se identifica al inicio con la letra de la serie seguida de un número. Por ejemplo, la norma para el plan de numeración de las telecomunicaciones públicas es la E.164.

1.2.2. ETSI

El Instituto Europeo De Normas de Telecomunicación es una entidad normalizadora independiente con proyección en todo el mundo, con el fin de producir estándares relacionados con la tecnología de la información y comunicación (ICT). Uno de sus mayores logros fue la normalización del sistema de telefonía móvil GSM (*global system for mobile*). El proyecto de asociación de tercera generación o 3GPP es un organismo dependiente del ETSI.

Entre otros organismos dependientes del ETSI se puede mencionar a EMTEL (Emergency Telecommunications) encargado de ver los aspectos relacionados con proveer servicios de telecomunicaciones en situaciones de emergencias; ICANN (Internet Corporation for Assigned Names and Numbers), encargado de velar por preservar la operatividad y estabilidad de Internet principalmente, y GSC (Global Standard Collaboration) la cual busca mejorar la

cooperación entre diferentes entidades normalizadoras para facilitar el intercambio de información en la que se refiere al desarrollo de estándares.

El ETSI produce una gran variedad de normas, especificaciones y reportes con distintos propósitos como respuesta a las demandas del mercado que se identifican de la siguiente manera:

- EN: estándar europeo
- ES: estándar ETSI
- EG: guía ETSI
- TS: especificación técnica
- TR: reporte técnico
- SR: reporte especial del ETSI
- GS: especificaciones de grupo

Una especificación técnica se nombra por un número de documento, seguido de la versión del documento y su fecha de lanzamiento, por ejemplo, una especificación técnica de 3 GPP es la ETSI TS 124 327 V11.0.0 (2012-11).

1.2.3. IEEE

El Instituto de Ingenieros Eléctricos y Electrónicos se fundó en Estados Unidos en 1884, aunque adquirió su nombre actual en 1963, con el fin de servir a profesionales involucrados en aspectos eléctricos, electrónicos, computacionales y cualquier campo relacionados con ciencia y tecnología. El IEEE está dividido en 38 sociedades cuyos miembros son profesionales de todo el mundo.

El objetivo de la IEEE es establecer estándares basados en consensos, realizar conferencias y realizar publicaciones técnicas relacionadas con las 38 sociedades. El estándar IEEE 802 es probablemente el más conocido por ser el que define las capas físicas y de enlace de datos utilizado por las redes de área local inalámbricas o WLAN.

1.2.4. IETF

La Fuerza de Trabajo de Ingeniería de Internet es una entidad internacional de normalización, su objetivo principal es dar contribuciones para que el Internet funcione de mejor manera y a la vez ayudar a cualquier persona en diseño, uso y gestión. La IETF ha definido casi todos los estándares y protocolos que se utilizan hoy en día relacionados con la comunicación IP. Como lo es la familia de protocolos TCP/IP los cuales definen los protocolos de red en los cuales se basa Internet y la comunicación entre elementos de red IP.

Se les llama RFC (*request for comments*) a todos los protocolos, procedimientos, procesos, programas y conceptos que publica la IETF; cada documento posee un solo número el cual lo identifica siempre; aunque este deje de ser válido, el número jamás se reemplaza. Algunos de los RFC más conocidos son el RFC 2543 o protocolo de inicio de sesión (SIP), RFC 1889 o protocolo de transmisión en tiempo real (RTP), RFC 791 o protocolo de Internet.

1.3. Telefonía digital

Las redes de telefonía continuaron su desarrollo durante los años 1900 con la invención de las centrales de conmutación basadas en electromecánica; para eliminar la necesidad de los operadores u operadoras telefónicas fue necesaria la invención de la señalización. Al hablar de señalización se refiere de

mensajes de control que se transmiten entre uno o varios teléfonos hacia la central de conmutación.

Fue necesario, entonces, el desarrollo de un teléfono con el cual fuera posible elegir con que abonado se quería establecer la comunicación; así pues, surgen los teléfonos por marcación de tonos que aun en la actualidad son funcionales en las redes de telefonía modernas.

La primera tecnología se desarrolló a finales de 1800 e inicios de 1900, consistía en un disco giratorio con 10 agujeros dentados, dependiendo de cuando se giraba el disco producía diferente cantidad de pulsos para identificar la marcación de los dígitos del 0 al 9. Luego entre 1940 y 1960, se mejoró esta tecnología y se le dio el nombre de *dual-tone multi-frequency o* DTMF, por sus siglas en inglés.

Consiste en un teclado con botones con los dígitos del 0 al 9 y los símbolos asterisco y numeral comúnmente; en otros se incluían las letras de la a hasta la d, al pulsar alguno de estos botones se produce un tono de 2 frecuencias diferentes como lo muestra la tabla I.

Tabla I. **Rango de frecuencias tonos DTMF**

Frecuencia	**1209 Hz**	**1336 Hz**	**1477 Hz**	**1633 Hz**
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Fuente: elaboración propia, utilizando programa Microsoft Excel.

Tras el desarrollo de los primeros transistores a inicios de 1950, se logra reemplazar los relés de las centrales electromecánicas y da comienzo la digitalización que se extiende a las diferentes partes de la red telefónica.

Las ventajas en cuanto a la eficiencia que ofrece la digitalización son evidentes en el transporte de la voz, lo que permitió compartir el medio de transmisión entre miles de usuarios por la misma línea de cobre donde antes solo podía transmitirse una llamada telefónica analógica a la vez.

En un principio se digitalizó la transmisión en los enlaces de comunicación lo que fue normalizado por la ITU-T, entidad creada 1868, su objetivo es crear normar y recomendaciones internacionales en cuanto a temas de telecomunicaciones, en un principio era el ente normalizador del intercambio telegráfico internacional.

Dentro de estas normas están la jerarquía de multiplexión digital conocida como PDH, o jerarquía digital plesiócrona, con la cual se puede utilizar un solo medio para la transmisión de miles de canales telefónicos separándolos por intervalos de tiempo. De igual forma la señalización también se digitalizó, lo que originó el surgimiento de otros servicios que pueden convivir en una red de telefonía.

1.3.1. Digitalización de señales analógicas

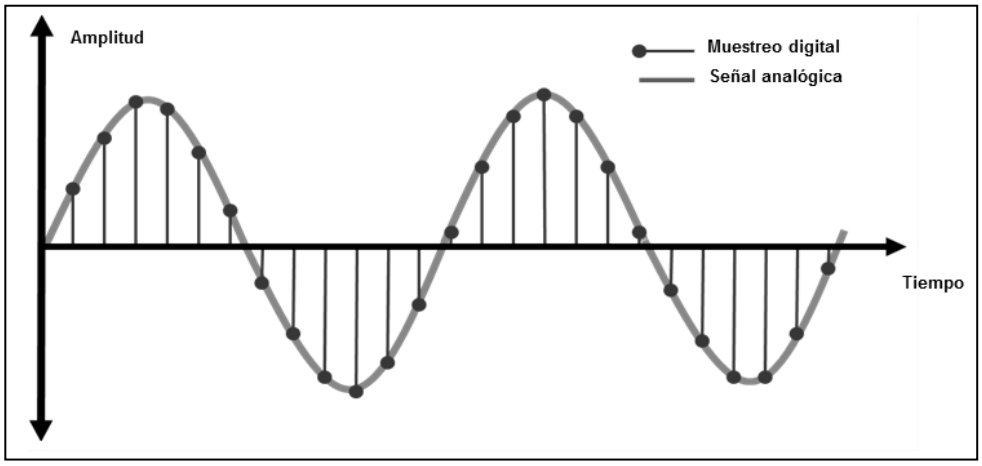
La digitalización es un proceso por medio del cual se convierte una señal analógica continua, como la voz, en una serie de valores puntuales numéricos en un periodo de tiempo determinado. Por ejemplo, una línea recta continua, luego de ser digitalizada, se tiene una serie de puntos que luego se unen entre sí y se obtiene una línea recta nuevamente.

1.3.1.1. Muestreo

Consiste en tomar una muestra en un periodo de tiempo dado de una señal analógica. La velocidad con que se toma la muestra es conocida como frecuencia de muestreo. El valor de la frecuencia de muestreo se relaciona directamente con la frecuencia máxima de la señal a la cual se quiere muestrear.

El rango de frecuencias audibles del oído humano va de entre los 20 Hz y 20 Khz; si se considera 0 el valor mínimo de frecuencias, se puede decir que se tiene un rango audible de 20 Khz, se introduce un nuevo término, el ancho de banda. El ancho de banda un canal de transmisión es el rango de frecuencias presentes en una señal que se transmite. A continuación, se ve un ejemplo de muestro de una señal analógica. Ver figura 2.

Figura 2. **Muestreo de una señal analógica**



Fuente: elaboración propia, utilizando programa Microsoft Excel.

Cualquier persona que ha utilizado un teléfono convencional ha notado que la voz que escucha tiene menos calidad si se compara con la que se escucha al hablar cara a cara con una persona; esto se debe a que cuando se escucha una conversación, para el oído, cualquier frecuencia más allá de los 10 Khz es irrelevante; en la telefonía convencional, la mayor frecuencia que se transmite es de 3.8 Khz que es lo necesario para identificar, en la mayoría de los casos, con quien se habla.

El teorema de Nyquist demuestra que para que una señal luego de ser muestreada pueda ser reconstruida, se necesita que la frecuencia de muestreo sea por los menos 2 veces mayor a la frecuencia máxima de la señal original si se toman en cuenta los 3,8 Khz que se utiliza en la telefonía, la frecuencia de muestreo mínima sería de 7,6 Khz, aunque realmente se aproxima a 8 Khz para la transmisión de voz en una red de telefonía.

1.3.1.2. Cuantificación

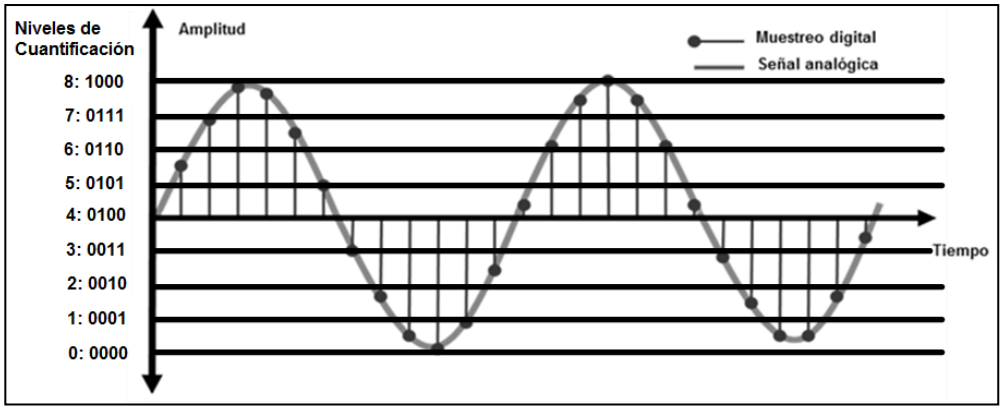
Luego de que una señal analógica es muestreada se obtiene una serie de valores puntuales en relación a la amplitud y al tiempo; sin embargo, los valores muestreados de la amplitud aún se encuentran en su forma análoga, es decir, pueden tomar cualquier valor: negativo, positivo, entero, decimal, entre otros.

Lo que se hace durante el proceso de cuantificación es definir niveles de valor finito o discreto en relación al valor mínimo y máximo de la amplitud de la señal muestreada. A mayor número de niveles de cuantificación más exacta será la señal reconstruida.

El valor de cada nivel de cuantificación está dado por un valor binario: una serie de ceros y unos; un valor unitario de esta serie se denomina bit, una serie de bits identifican a un nivel de cuantificación. Para calcular cuántos bits son necesarios para un número de niveles de cuantificación dado, se utiliza la fórmula $N_b = Log_2(N_c)$, donde N_c es el número de niveles de cuantificación.

Si se utilizan por ejemplo, 8 niveles de cuantificación, el número de bits sería 3. Aunque la ITU-T ha definido en la recomendación G.711 que para una frecuencia de muestreo de 8 Khz se deben utilizar 8 bits para los niveles de cuantificación; un ejemplo de cuantificación se muestra en la figura 3.

Figura 3. **Cuantificación de 9 niveles utilizando 4 bits**



Fuente: elaboración propia, utilizando programa Microsoft Excel.

1.3.1.3. Codificación

La codificación es el proceso por medio del cual se transforman todas las muestras cuantificadas en un tren de impulsos de unos y ceros, el cual es transmitidos a través de un canal con una velocidad o bit *rate* definido en bits por segundo.

Con el pasar del tiempo han surgido diferentes técnicas de codificación, con el fin de recortar el número de bits que identifican una muestra utilizando diferentes técnicas u algoritmos, todo para aprovechar de mejor manera el canal de transmisión y desarrollar métodos para corrección de errores ante la existencia de ruido en la decodificación.

La decodificación es únicamente el proceso inverso que se realiza para reconstruir una señal digital a la señal analógica inicial; este proceso es afectado por el ruido, el ruido en las telecomunicaciones se define como cualquier clase de error o distorsión sobre la información que se transmite, esto puede ser causado por interferencia de otras señales, calor, inducción eléctrica o descargas electromagnéticas.

1.4. Transmisión de señales digitales

Luego de finalizado el proceso de codificación de una señal única, hay que considerar que una señal digitalizada no puede ser transmitida directamente hacia el transmisor, debido a que la misma no podría ser diferenciada de otras señales en el medio por el cual se transmita, lo cual no sería un problema al considerar una comunicación punto a punto. Se busca entonces poder transmitir varias llamadas digitalizadas por un solo canal, separando cada una de ellas con alguna identificación; a esta acción se le llama multiplexación.

1.4.1. TDM

Es una de las primeras técnicas de multiplexación digital utilizada desde 1963 y vigente a la fecha, consiste en transmitir en un canal varias señales digitales separadas por intervalos de tiempo, de esta forma cada señal transmitida utiliza todo el ancho de banda disponible cuando transmite.

Una de las limitantes de esta técnica de multiplicación es que requiere que tanto el receptor como el transmisor estén sincronizados para poder procesar cada diferente señal de manera correcta. La mayor ventaja es que es una técnica muy simple en concepto.

Cada intervalo de tiempo se denomina *time slot* o TS, cada uno de estos se considera un canal a través del cual se puede establecer una llamada. Al conjunto de TS que representan un ciclo en la multiplicación se le denomina trama.

1.4.2. PDH

La jerarquía digital plesiócrona o PDH, por sus siglas en inglés, es una técnica de multiplicación TDM; esta jerarquía agrupa 32 TS de 64 kb/s cada uno, para obtener 2 048 kb/s, lo cual se denomina E1 para la jerarquía europea que se utiliza también en Guatemala. Luego se crean órdenes superiores en grupos de 4 para obtener velocidades de transmisión de 8 448 Kb/s (E2), 34 368 kb/s (E3) y 139 264 Kb/s (E4) Kb/s y 565 148 Kb/s (E5).

Los estándares de Estados Unidos y de Japón varían únicamente en los niveles de agrupación de cada jerarquía, aunque el TS siempre tiene el valor de 64kb/, las velocidades son distintas en niveles superiores.

El proceso de multiplexación PDH, por ser una multiplexación utilizando el tiempo como referencia, posee un elemento llamado reloj del cual se toma la referencia de tiempo; cuando se realiza la multiplexación cada nivel de la jerarquía. El reloj en TDM es una señal digital, un tren de impulsos de unos y ceros.

Se denomina jerarquía plesiócrona porque las referencias de reloj son independientes en cada nivel de jerarquía, es decir, es una jerarquía casi síncrona o plesiócrona. El sistema de transmisión solo es síncrono con el último nivel de jerarquía.

1.4.3. SDH

Las velocidades de transmisión de tecnología PDH son diferentes entre países como Estados Unidos, Japón y el resto del mundo que utiliza el estándar europeo, por lo que la interconexión de estas redes representaba costos muy altos. Otra desventaja de esta tecnología es que cada nivel jerárquico utiliza multiplexores y demultiplexores, por lo que es sumamente difícil localizar un canal de 64 Kb/s en órdenes superiores, ya que se debe demultiplexar toda la trama.

La jerarquía síncrona digital o SDH, por sus siglas en inglés, que surge a inicios de 1990, fue la solución para las desventajas de la tecnología PDH; además de normalizar las tramas mayores a 565 Mb/s y la transmisión por medio de fibra óptica. Esta tecnología permite la integración de tecnologías anteriores como PDH, ofrece también mecanismos de redundancia o protección.

A diferencia de la tecnología PDH, la jerarquía digital síncrona, como su nombre lo indica, utiliza una única fuente de reloj, la primera jerarquía se denomina STM-1 de 155 Mb/s. Las jerarquías de mayor nivel se denominan STM-N.

El contenedor es la unidad básica de la trama STM-1, un tributario, como se denomina a un conjunto de canales o TS, como lo es un E1, es

empaquetado dentro de un contenedor; durante este empaquetado se incluyen bits de relleno, si el E1 es síncrono o bits de justificación si el E1 es asíncrono, es decir, no se incluye ninguna información de control. Hay contenedores de diferentes dimensiones para adaptarse a los datos generados por diferentes tributarios PDH, por ejemplo, E2 o E3.

Luego, un contenededor se transforma en un contenedor virtual cuando se le agrega información de control y para monitoreo de la transmisión de información entre la fuente y el destino. También, se utiliza para identificar el contenido de un contenedor. Después, se convierte en una unidad tributaria en la cual se agrega un puntero indicando el inicio de un contenedor virtual. A partir de este momento comienza la multiplexación ya que se agrupan las unidades tributarias en grupos o TUG. Varios TUG hacen contenedores virtuales de orden superior en los cuales se agrega también información de monitoreo, control.

Varios contenedores de orden superior hacen unidades administrativas en a las cuales se agrega un puntero para diferenciar el inicio de un contenedor de orden superior. Un grupo de unidades administrativas o AUG se inserta en la trama de un STM-N donde N puede tomar los valores de 1, 4, 16, 64 y 256.

A pesar de que se intentaba normalizar la transmisión PDH surgen 2 estándares diferentes SONET y SDH. Las diferencias entre estas son mínimas, al igual que en PDH, lo que varía son los niveles jerárquicos; es decir, la velocidad de transmisión de datos, pero hay equivalencias entre estas para que sean compatibles. La tecnología SONET es utilizada en Estados Unidos, Canadá, Corea del Norte, Hong Kong y Taiwán, es definida por los estándares ANSI. La tecnología SDH se utiliza en Europa y el resto del mundo, la definen los estándares TU-T.

1.5. Inicio de la telefonía móvil

En un principio la comunicación entre teléfonos comenzó por medio de canales de transmisión alámbricos. Con el surgimiento de la radio en 1896, es posible la transmisión de señales de voz en forma de ondas electromagnéticas. Una onda electromagnética se produce cuando una carga eléctrica se acelera, al acelerarse produce un campo eléctrico variable y a su vez produce un campo magnético; las leyes de Maxwell deducen que estos campos pueden abandonar el medio en el cual se originan y, por lo tanto, se pueden transmitir en el aire.

Se dice entonces que los campos eléctricos y magnéticos, vibran u oscilan y esto hace que la onda electromagnética se mantenga mientras viaja en el aire, surge entonces el concepto de antenas, dispositivitos los cuales emiten y detecta ondas electromagnéticas las cuales se transmiten en una frecuencia u oscilación dada.

Por lo que el ancho de banda, definido anteriormente para conexiones alámbricas, aplica también para transmisiones inalámbricas para separar y poder interpretar las ondas electromagnéticas. Es necesario, entonces, que tanto el receptor emita ondas en un ancho de banda o rango de frecuencias dado y que el receptor filtre únicamente las ondas en ese ancho de banda.

En 1946 surge comercialmente la telefonía inalámbrica análoga, que consistía en un transmisor de alta potencia instalado en una ciudad y permitía interconectar a usuarios móviles con las líneas fijas en un rango de 80Km, para la transmisión de la voz se utilizaba la frecuencia modulada, en cambio a la amplitud modulada, que solo varía la potencia de la onda, varía la frecuencia de la onda original. Para conectar las llamadas también era necesario utilizar una operadora ya que aún no existía la marcación directa que llegó hasta 1960.

Con el incremento de las estaciones de transmisión y para aprovechar de mejor forma el ancho de banda, surge la necesidad de dividir las zonas geográficas en aéreas de cobertura, las cuales se denominan celdas o células; una celda es únicamente una estación base de transmisión y recepción de canales de audio en forma de ondas electromagnéticas.

1.5.1. Primera generación de la telefonía móvil

Antes de la primera generación, la mayoría de tecnologías de telefonía móvil se centralizaban únicamente en desarrollar bases de transmisión más poderosas con el fin de cubrir una mayor área. Y fue hasta que se introdujo el concepto de células que se considera la primera generación.

1.5.1.1. AMPS

El sistema avanzado de telefonía móvil fue el primer sistema de telefonía móvil implementado en los Estados Unidos durante 1980 y que permitió un área de cobertura nacional. Para la transmisión utilizaba frecuencia modulada y duplicación por división de tiempo, es decir, los canales de transmisión y recepción funcionan a frecuencias diferentes. Para poder soportar varios usuarios, una estación base utiliza el acceso múltiple por división de frecuencia, lo que hace es definir varios canales que corresponden a diferentes frecuencias y los cuales se asignan a distintos usuarios.

Se utilizaba el rango de frecuencias 800 Mhz y 900 Mhz para la transmisión entre los teléfonos móviles; la estación base utilizaba el rango de frecuencias entre 824 Mhz y 840 Mhz, para el camino inverso se utilizaba el rango de frecuencias entre 869 Mhz y 894 Mhz.

La tecnología AMPS permite a un usuario que está moviéndose de un área de cobertura a otra de mantener la llamada sin que la misma sea interrumpida, lo cual se denomina *handover*. Aunque AMPS representa un gran avance en la solución de la telefonía móvil aun no considera la compatibilidad con otras tecnologías.

1.5.1.2. TACS

El sistema de comunicaciones de acceso total es una variante del sistema AMPS; se utilizó durante la misma época, pero en Europa, Japón y Hong Kong. ETACS es una variación de este sistema y tenía más canales de comunicación.

Entre las diferencias entre el sistema TACS o ETACS y AMPS es que el primero tenía una capacidad de 1 278 canales de voz, mientras que AMPS únicamente 832. El espaciamiento entre cada canal también variaba en 5 Khz, 30 Khz que utilizaba AMPS y 25 Khz de TACS. Actualmente, ambas tecnologías se encuentran obsoletas.

1.5.2. Segunda generación de la telefonía móvil

La segunda generación de telefonía móvil es la sucesora de la primera generación ya que en la primera generación se utilizan señales analógicas y una baja tasa de transferencia, no es posible satisfacer las demandas ante los requerimientos de crecimiento de la red. La tecnología 2G entonces emplea multiplexación digital en el acceso y transmisión, hace que esta tecnología tenga una eficiencia 3 veces mayor a la primera generación.

Los protocolos de la segunda generación fueron diseñados por diferentes compañías y presenta el mismo problema que la primera generación: no hay compatibilidad entre las diferentes tecnologías.

1.5.2.1. CDMA

La tecnología de acceso múltiple por división de código es una técnica de transmisión y acceso, que se basa en una técnica de modulación denominada *Spread Spectrum* o espectro disperso, la motivación para desarrollar esta técnica fue proteger la transmisión de los datos de ser interceptados y decodificados.

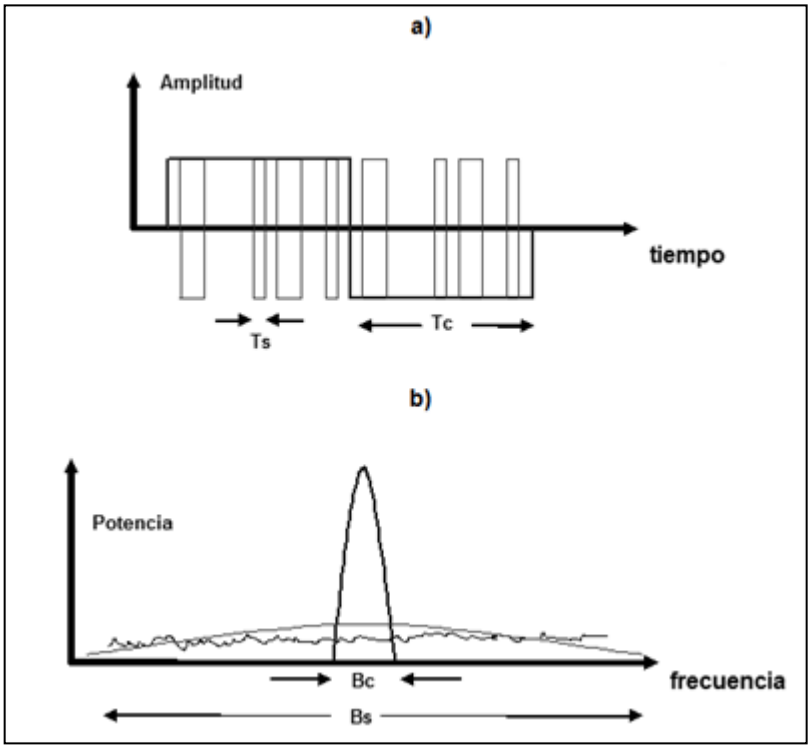
En CDMA todos los usuarios utilizan el mismo canal sin la necesidad de que el mismo deba ser liberado para poderlo utilizar. Para diferenciar a un usuario se le asigna un código único que luego se utiliza para cifrar la señal de información. El receptor debe conocer este código para recuperar la señal original.

Los sistemas CDMA proveen un código de control de errores, soporta también el *soft handover*, es decir, un usuario puede estar transmitiendo a través de varias celdas al mismo tiempo, un controlador se encarga de analizar la calidad de la señales de datos y elige la mejor; otra característica es el esparcimiento del espectro de frecuencia, que hace que la señal sea muy difícil de detectar y por esta razón es que al principio su uso era militar; otra característica es que no hay interferencia entre las señales de la misma banda ya que cada usuario tiene un código único de transmisión.

La señal transmitida por la técnica CDMA comparada con la señal original utiliza un ancho de banda mayor, por lo tanto, la técnica ensancha el espectro la

señal original. Esto se hace disminuyendo el período de la señal original como se muestra en la figura 4.

Figura 4. **Esparcimiento del espectro por secuencia directa**



Fuente: elaboración propia, utilizando programa Microsoft Excel.

Se considera Tc como el periodo de la señal base y Ts el periodo variable en función del código único de transmisión. El período de una señal es inversamente proporcional a la frecuencia, por lo que al disminuir el periodo aumenta la frecuencia. Al analizar ahora este cambio en función de la frecuencia, como en la figura 4b, se observa que lo que antes era una señal de banda estrecha Bc, ahora es una señal de banda ancha Bs, es decir, que cubre un rango de frecuencias más amplio gracias a las variaciones de período como se muestra en la figura 4a.

Como se observa también en la figura 4b, una señal esparcida en el espectro con ancho de banda Bs se asemeja a la forma de onda del ruido la cual es de potencia baja y está presente en todo el espectro de frecuencias.

La técnica de *frequency hopping spread spectrum* o FHSS, por sus siglas en inglés, es una técnica básica de esparcimiento de frecuencia. Es un sistema en el cual se modifica la frecuencia de la información transmitida; esta modificación se conoce como salto de frecuencias, la secuencia de salto puede ser elegida al azar o puede ser configurable; tanto receptor como transmisor deben conocer esta secuencia de saltos en frecuencia.

Las frecuencias se seleccionan de cierta forma que no haya interferencia entre cada salto. Hay dos tipos de técnicas de salto en frecuencias: saltos lentos y saltos rápidos; en la primera la rapidez de saltos es menos a la rapidez con que se transmiten los datos y en la segunda lo contrario.

THSS es otra técnica básica de esparcimiento en frecuencia, basada en saltos en el tiempo, *time hopping spread spectrum,* como lo indica su nombre en inglés. La frecuencia de igual forma varía, pero no en modo aleatorio, solo en múltiplos de la frecuencia base, pero la información es transmitida en diferentes *frames* o intervalos de tiempo los cuales varían según un código específico. El sistema THSS aprovecha eficientemente el ancho de banda a diferencia de otras técnicas de espectro esparcido, también es más fácil de implementar que la técnica FHSS.

1.5.2.2. GSM

El sistema global para las comunicaciones móviles, por sus siglas en inglés GSM, es un estándar aprobado mundialmente para las comunicaciones celulares digitales que fue desarrollado inicialmente para su uso en Europa pero que tuvo proyección mundial. El estándar fue desarrollado por la ETSI y fue lanzado comercialmente a mediados de 1992.

En GSM se ofrecen nuevos servicios; en las redes analógicas únicamente existía el servicio de voz, ahora existen otros servicios: mensajes cortos o SMS, servicio de fax y datos, correo de voz y otros servicios suplementarios de voz como desvío e identificación de llamadas.

La seguridad del sistema GSM comparada con la telefonía análoga es mejorada considerablemente ya que se incluye una tarjeta denominado módulo de identidad de subscriptor, SIM por sus siglas en inglés, no solo para la identificación o localización de un usuario sino también para la autenticación del mismo con la red móvil y así poder evitar que usuarios no autorizados utilicen la misma.

GSM utiliza una variación de TDMA y FDD, *frequency division duplex*, es una técnica en la cual las bandas para transmisión y recepción son diferentes. Comúnmente las bandas utilizadas en GSM son la 900 Mhz, con un máximo de 124 canales y 1 800 Mhz con un máximo de 374.

1.5.2.3. GPRS

General packet radio service o GPRS es un servicio de paquetes de datos para la comunicación inalámbrica; este servicio convive con las redes GSM y

CDMA descritas anteriormente. La diferencia entre las redes, las conmutadas por circuitos y las conmutadas por paquetes, es que las primeras es que existe un canal o recurso reservado en toda la red para la transmisión de la voz durante toda una llamada. En cambio, en la red conmutada por paquetes, la información es dividida en partes más pequeñas y enviada por diferentes rutas a través de la red.

La tecnología GPRS fue diseñada por la ETSI con el propósito de que no interfiriera con los servicios ya existentes de las redes 2G. Comúnmente se conoce como 2.5G. Se dice que un subscriptor móvil o MS es un usuario siempre en servicio del sistema GPRS ya que no necesita establecer una conexión por medio de marcación lo que hace que el servicio esté instantáneamente disponible para un usuario.

La máxima velocidad de la conexión oscila entre los 56 kbps y 118 kbps que representa un gran incremento comparados con los 9,6 kpbs que se obtenían de una conexión a través de la red conmutada por circuitos. Debido a estas velocidades más altas de conexión el sistema GPRS permite el uso de aplicaciones que dependen solo del uso de Internet, así como el uso video conferencias desde una terminal móvil.

1.5.2.4. EDGE

Esta tecnología, como sus siglas en inglés lo indican, es una evolución de la tasa de datos para GSM, también es conocida como EGPRS. Por medio de la cual se alcanza una velocidad más alta en la red GPRS, aunque requiere un cambio físico para implementar esta solución en la red de acceso GPRS existente, debido a que la técnica de modulación superior con la cual se puede alcanzar hasta 384 kbps.

1.5.3. Tercera generación de la telefonía móvil

3G es conocida como la tercera generación móvil, fue desarrollada a través del proyecto 3GPP que inicio en diciembre de 1998. Los principales objetivos de esta tecnología es el *roaming* a nivel global; se describe el *roaming* como la habilidad de un suscriptor de utilizar su terminal en otros países; otro objetivo es alcanzar velocidades de transmisión de datos más altas, en un principio hasta 2Mbps; también, una mayor capacidad del sistema de radio lo que significa mayor número de usuarios conectados al mismo tiempo.

El mayor cambio en las redes 3G es a nivel de acceso tanto para voz como datos ya que utilizan una modulación superior; también, se necesita una terminal móvil nueva para utilizar este servicio; estas nuevas terminales son compatibles con la tecnología GSM. A diferencia de que en las redes GSM la autenticación en las redes 3G no es más opcional.

Se introduce una nueva tarjeta de acceso denominada USIM, la cual tiene una quinteta de parámetros para autenticar; sin embargo, aún es posible el uso de una SIM para registrarse a la red 3G la cual solo utiliza una tripleta para la autenticación por lo cual muchos operadores continuaron utilizando la misma tarjeta de acceso.

Existieron diez propuestas para el desarrollo de la tecnología 3G; las que más auge tuvieron fueron la UMTS, iniciativa europea, y CDMA-2000, iniciativa estadounidense; se describe la UMTS que es la mayormente usada a nivel mundial.

La tecnología UMTS utiliza una técnica de acceso denominada WCDMA, a diferencia de la tecnología CDMA descrita anteriormente; WCDMA utiliza un

ancho de banda más grande: 5Mhz; mientras que la tecnología CDMA utiliza 1,25 Mhz. Debido a su mayor ancho de banda, las velocidades de transmisión de datos también son mayores. WCDMA también está diseñada para soportar los servicios de datos y voz simultáneamente.

UMTS puede utilizar FDD o TDD para los canales de transmisión o *uplink* y recepción o *downlink.* Para FDD se utilizan diferentes frecuencias para el *uplink* y el *downlink.* Mientras que para TDD se utiliza la misa frecuencia. Existen alrededor de 25 diferentes bandas de frecuencias para la tecnología UMTS, que van desde los 700 Mhz hasta los 3 500 Mhz.

1.5.3.1. HSDPA

La tecnología HSDPA o *high speed downlink packet access* es conocida también como 3.5G, es una mejora a la red de acceso UMTS. Esta tecnología es totalmente compatible con la tecnología WCDMA. Se puede alcanzar hasta una velocidad de transmisión de 14 Mbps.

La mayor capacidad de la tecnología HSDPA se debe a una modulación de mayor capacidad comparada con la utilizada en UMTS. HSPA+ es una evolución del estándar HSDPA, utiliza una técnica de múltiples antenas y alcanza velocidades de descarga de hasta 22 Mbps.

1.5.4. Cuarta generación de la telefonía móvil

LTE o *long term evolution* es el último estándar definido para la telefonía móvil, la principal motivación para el desarrollo de la red LTE es alcanzar tasas de datos más altas, proveer una mayor eficiencia del espectro electromagnético, simplificar la arquitectura de la red de datos, entre otros.

LTE utiliza la modulación OFDM, *orthogonal frequency division multiplexing;* esta técnica utiliza un gran número de portadoras de ancho de banda estrecho que se superponen entre sí, mediante una separación de frecuencia dada entre la frecuencia central de una portadora y las siguientes; esta separación está dada en el punto en que la señal anterior tiene una potencia de cero. A diferencia de CDMA para 3G que utiliza una sola portadora de ancho de banda amplio.

Diferente de la evolución de 2G a 3G, la evolución de 3G a 4G no considera ningún cambio en la parte de voz, únicamente representa un cambio a nivel de datos; se requiere terminales móviles que soporten 4G, aunque estas a la vez son compatibles con las generaciones anteriores. De debe realizar a la vez un cambio de hardware a nivel de acceso y control en la red de telefonía de datos para soportar la red LTE.

Debido a que la red 4G no presenta ningún cambio a nivel de voz, se necesita entonces utilizar el servicio de la red 3G o 2G para realizar o recibir llamadas de voz, entonces, cuando un dispositivo se encuentra en 4G y necesita utilizar servicios de voz este necesita desconectarse de la red 4G y conectarse a la red 3G; a este proceso se le llama CS *fallback* o CSFB abreviado.

Para evitar este inconveniente se puede realizar una llamada sobre IP, este proceso se llama VoLTE o voz sobre LTE, pero necesita involucrar a otra red llamada sistema multimedia IP o IMS por sus siglas en inglés. Con VoLTE se logra reducir el tiempo de conexión de las llamadas hasta 10 veces de lo que ofrece la opción CSFB; entre otras ventajas esta; también, el espectro de radio es también 4 veces más eficiente y la calidad de la voz es de alta definición.

Normalmente los operadores prefieren la opción de CSFB debido a que no requieren elementos adicionales de red y la implementación es más sencilla, así pueden reducir el tiempo de lanzamiento de la red 4G o LTE.

1.6. Redes de telefonía actuales

Por redes de telefonía actuales se refiere únicamente a las redes de última tecnología sino a las tecnologías que conviven en el periodo de tiempo y contexto de desarrollo de este trabajo de investigación. Normalmente los fabricantes de equipos de telecomunicación unen ciertas tecnologías o elementos de red en un solo, con el fin de simplificar la red de un operador; a su vez a lo largo de la evolución de las redes de telefonía también se da el caso inverso: la separación de ciertos elementos con el fin de simplificar en análisis de fallas y la expansión de capacidad de los equipos.

Es importante entonces definir cuáles son las subredes que componen una red de telefonía; también, qué elementos y de qué forma se comunican entre sí.

1.6.1. Redes de telefonía fija

Se entiende por telefonía fija todos aquellos servicios de voz y datos en los cuales el usuario que establece la comunicación se encuentra en una ubicación establecida y, por lo general, utiliza un medio cableado para establecer conectividad con la red de servicios.

Se describirán únicamente dos tipos de tecnologías de redes fijas ya que en ellas se incluyen otras tecnologías anteriores de telefonía fija.

1.6.1.1. Redes NGN

La tecnología NGN, o *next generation network,* es una red totalmente basada en el protocolo IP pero que también adapta otros protocolos basados en la tecnología TDM. La red NGN es una evolución y mejora de la red PSTN *public switched telephony network,* que está basada en el protocolo SS7, este protocolo define las comunicaciones basadas en TDM. El principio de la red PSTN es establecer llamadas de voz en tiempo real o la utilización de datos designando un circuito o canal específico para la comunicación entre usuarios estáticos o fijos.

Las principales motivaciones del desarrollo de la red NGN son el crecimiento de la comunicación de banda ancha, el decaimiento del negocio de las llamadas en la PSTN y, no menos importante, la reducción de costos para la integración de nuevos usuarios y servicios a las redes de telefonía.

1.6.1.2. Redes IMS

La tecnología IMS, IP *multimedia subsystem,* es una arquitectura para servicios de multimedia basados en tecnología IP, como servicios multimedia se tiene la transmisión de voz, imágenes, video u otro tipo de información, que puede ser simultáneamente entre uno o más usuarios.

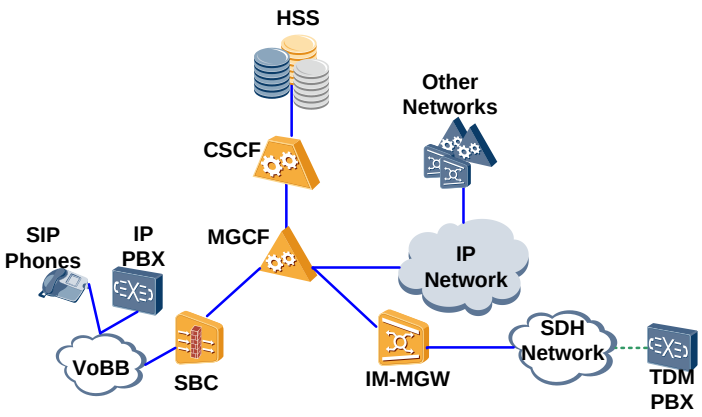
Estos servicios se caracterizan por requerir más estabilidad y disponibilidad sin dejar de lado la simplicidad de uso y de implementación, la movilidad de los subscriptores y una comunicación en tiempo real; esto se logra en gran parte por medio de la utilización de protocolos de inicio de sesión o registro y de transmisión de paquetes multimedia en los cuales se establece una calidad o prioridad en cada servicio para la transmisión, mismos a través de

la red de transporte, donde la prioridad de cada uno de estos servicios se define por el operador de la red.

1.6.1.3. Estructura de una red IMS-NGN

La arquitectura de la red IMS, aparte de proveer servicios puramente IP, también cuenta con elementos de red capaces de integrar la red NGN que es puramente TDM. En la figura 5 se observan los elementos que componen una red IMS.

Figura 5. **Topología de una red IMS**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

El HSS, *home subscriber server,* es el elemento de red que almacena la información de autenticación de los usuarios y sus perfiles.

CSCF *call session control function,* tiene tres distintos roles, el Proxy CSCF o P-CSCF que es el primer punto de contacto y de autenticación de los usuarios con la red IMS; una de sus funciones es proteger la conexión de los usuarios hacia la red y la red IMS como tal.

El *interrogating* o I-CSCF tiene como función la de consultar con el HSS la dirección del *serving* CSCF o S-CSCF que utilizará un usuario y el reenvió de solicitudes hacia este mismo elemento de red. El S-CSCF actual como registrador SIP es el responsable de registrar la ubicación y autenticación de cada usuario y el procesamiento y ruteo de las llamadas.

Otro elemento de la red IMS es el MGCF, *media gateway control function,* que puede cumplir dos funciones; primero ser el enlace de conexión para las redes TDM junto con e IM-MGW o *IMS media gateway* donde el MGCF controla la señalización y el IM-MGW provee canales de voz y media así como también conexión física para acceso; otra función es la de ser el punto de interconexión a otras redes no IMS como lo puede ser redes móviles o redes de otros operadores.

El SBC, *session boarder controller,* puede suplir la función de P-CSCF dentro del IMS; entre otras funciones esta la protección de la red IMS cuando esta interactúa directamente con otras redes a través de Internet.

1.6.1.4. Protocolos en una red IMS-NGN

Un protocolo es una norma establecida para estandarizar un proceso, en este caso para la red IMS, su función es normalizar la comunicación y funcionamiento de todos los elementos de red no importando quien las fabrique. Existe una gran variedad de protocolos, pero para este trabajo de graduación se analizarán únicamente los más importantes de una manera breve.

El protocolo NTP, *network time protocol* es de suma importancia para las redes IP ya que define los estándares para sincronización de tiempo en los

diferentes elementos de una red, ya sea móvil o fija, y los servidores que proveen la sincronía de tiempo.

DNS *domain name system,* es una base de datos que almacena los nombres de dominio utilizados en la red IMS y las IPs que corresponden a estos dominios. En la red IMS se utiliza entre CSCF y otros servidores de aplicación para simplificar y ocultar la conexión hacia la red.

Otro protocolo utilizado es el ENUM, el propósito es realizar un mapeo entre los números telefónicos utilizados en la PSTN que utilizan un formato E.164 y el Internet que utiliza el protocolo DNS. Al combinar ambos protocolos, un número telefónico dentro de la red IMS se convierte ya sea en una URI (*uniform resource identifier)* o en una dirección IP.

El protocolo *diameter* se deriva de su predecesor RADIUS, la función principal de estos protocolos es autenticar, autorizar y contabilizar a un usuario en la red IMS.

SIP *sesión initiation protocol,* es usado para controlar las llamadas mediante el inicio de sesiones multimedia de voz y de datos; este protocolo es utilizado entre el CSCF y otras entidades como las terminales, los servidores de aplicación y el MGCF.

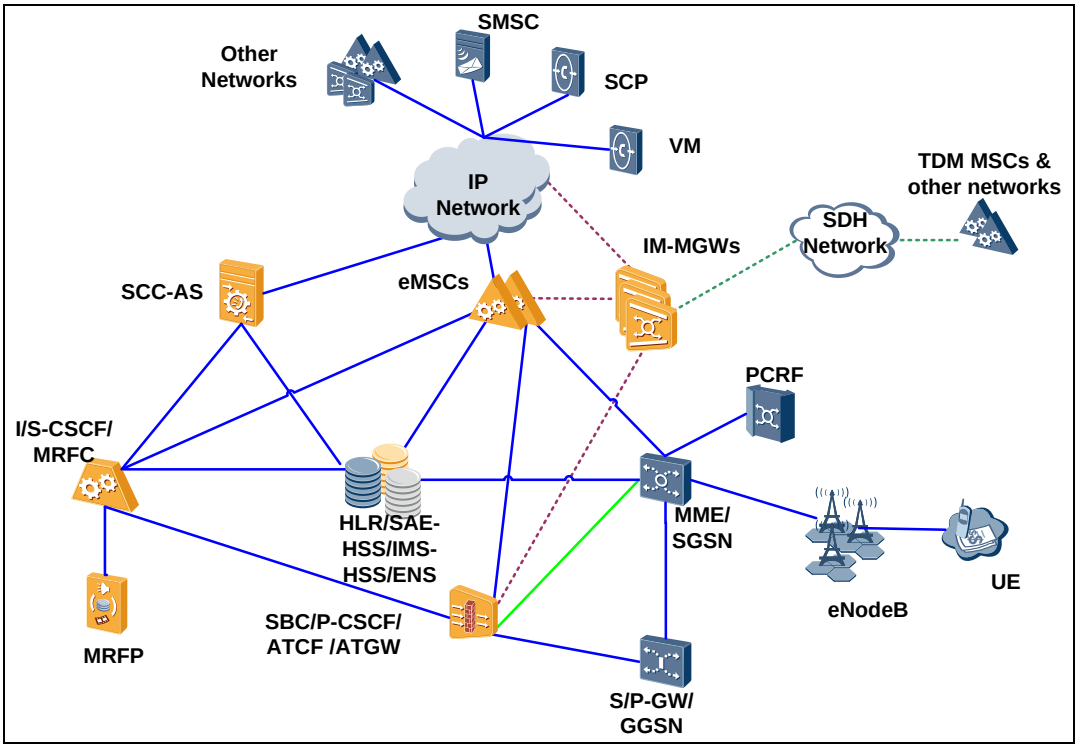
1.6.2. Redes de telefonía móvil

Anteriormente se indicaron los inicios de la telefonía móvil y en qué punto del mapa se encuentran actualmente. En esta parte se trata de describir una imagen global de qué elementos componen esta red y una descripción básica de su funcionalidad.

1.6.2.1. Estructura de red móvil

Las redes móviles proveen servicios de voz y datos a usuarios no importando cuál sea su ubicación. La arquitectura de la red móvil ha evolucionado de tal forma que varias tecnologías puedan convivir entre sí, como por ejemplo 2G, 3G, LTE y sus derivaciones. En la figura 6 se observa los elementos de red que componen una red móvil moderna la cual puede soportar VoLTE.

Figura 6. **Estructura de una red móvil moderna**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

No todas las redes de telefonía móvil tienen la misma estructura debido a que los servicios que se ofrecen son distintos, pero buena parte cuentan con elementos que se describirán en este enunciado.

Es de suma importancia describir los elementos de una red móvil moderna ya que son parte importante de esta investigación, porque son los vulnerables antes ataques o fraudes.

El HSS, descrito anteriormente para la red IMS, cumple una función similar para la red móvil: almacenar los datos de autenticación y los perfiles de voz y datos de los usuarios. Para una red moderna se pueden tener integradas las bases de datos de las diferentes redes, es decir, 2G, 3G, LTE y IMS en un solo elemento de red. Es decir, el HSS puede suplir la tarea del HLR, *home local register,* de las redes 2G y 3G, el HSS de la red LTE y el HSS de la red IMS.

La eMSC *enhanced mobile soft switch,* es una mejora del elemento de red MSC el cual es el encargado de conectar las llamadas de voz y también de ser el puente entre el UE, *user equipment,* es decir, un usuario y los servicios de valor agregado.

En una red moderna de telefonía la eMSC integra diferentes servicios de la red móvil y la red IMS para la red móvil integra el VLR, *visitor location register*, su función principal es almacenar una copia de la ubicación y del perfil de un subscriptor. Para el IMS integra el MGCF, mencionado anteriormente. Con la convergencia de estas dos redes, surge VoLTE y otro elemento integrado en el eMSC el SRVCC IWF, *single radio voice call continuity interworking function,* que no es más el nombre que toma el elemento de red encargado de realizar el *handover* de una llamada VoLTE hacia 3G o 2G.

Existe una gran cantidad de servicios de valor agregado: los mensajes de texto controlado por el SMSC, *short message service center,* los servicios IN, *intelligent network,* en los cuales se incluyen los usuarios prepago, los usuarios VPN, *virtual prívate number*, que ofrecen servicios de flotilla y de marcación corta, entre otros. A los servidores IN se les conoce como SCP, *server control point,* que utilizan el protocolo CAMEL para interactuar con la eMSC.

Entre otros servicios de valor agregado están el VM, *voice mail,* para poder depositar mensajes de voz en caso de que no se puede contactar a un usuario.

El IM-MGW también mencionado anteriormente, cumple la misma función para la red móvil la cual es establecer los canales de voz, así como también servir como punto de interconexión con la red legado TDM.

El SBC para la red IMS es otro elemento el cual puede integrar varios elementos de red como el P-CSCF. Para VoLTE, y para mejorar el proceso de *handover,* surgen dos nuevos elementos integrados en el SBC: el ATCF, *access transfer control function,* y el ATGW, *access transfer gateway* su función principal es la de reducir la trayectoria de conexión de una llamada cuando pasa de VoLTE hacia 3G o 2G y con este reduciendo el tiempo entre esta transición y mejorando la experiencia del usuario de VoLTE.

El MRFC, *media resource function controller,* y el MRFP, *media resource function proccesor,* proveen servicios de manipulación de media, por ejemplo, mezcla de flujos de audio y reproducción de tonos. El MRFC controla la parte de señalización proveniente de los servidores de aplicación y del S-CSCF y el MRFP es controlado por el MRFC y es el encargado de establecer los recursos de media.

Otro elemento importante para hacer compatibles la red VoLTE y los servicios 3G-2G es el SCC-AS, *service centralization and continuity application server,* entre sus funciones está permitir a un usuario VoLTE seguir utilizando servicios IN ya que este es el único elemento que soporta el protocolo CAMEL y también de anclar las sesiones de los usuarios VoLTE para poder ser transmitidas entre la red IMS y la red conmutada por paquetes.

Una red de telefonía ofrece servicios de datos por igual, es decir, el uso del Internet y aplicaciones que funcionan por medio del mismo, esta red se le denomina *packet swicth network,* red conmutada por paquetes. Aparecen entonces los elementos de red MME, *mobility management entity,* y el SGSN, *serving GPRS support node,* ambos encargados del manejo y transferencia de paquetes, así como del control de la movilidad de un usuario para una red de 4G o 3G-2G respectivamente.

El S/P-GW, *serving/packet data gateway,* y el GGSN, *gateway GPRS supporting node,* principalmente encargados de la interconectividad entre la red de datos y otras redes externas, comúnmente con Internet.

El PCRF, *policy and charging rules function,* define las reglas y políticas en tiempo real de cada usuario para el uso de multimedia y datos. Para la red VoLTE define las políticas de calidad de servicio para el canal de voz sobre IP, a su vez que controla la cuota de datos utilizada por un subscriptor en tiempo real entre otras funciones como establecer periodos de tiempo sin que se realice cobro por navegación o no realizar cobro cuando se utilizan ciertas aplicaciones sobre Internet como lo pueden ser las redes sociales.

Para finalizar el eNodeB es el elemento de acceso a la red LTE y que provee canales de radio para señalización, audio y datos, para varios usuarios

dentro de un área o cobertura. En la red 3G este elemento se denominada NodeB y en la red 2G BTS, *base transceiver station,* e iba acompañado de una RNC, *radio network controller* para 3G o una BSC, *base station controller,* encargados de la movilidad de los usuarios; tarea que es integrada en el MME para la red 4G.

1.6.2.2. Protocolos en una red móvil

El protocolo SIP es utilizado comúnmente para establecer sesiones de llamadas o servicios entre oficinas de diferentes redes que se conectan a través de Internet. El protocolo diameter es utilizado para el registro, autenticación de usuarios LTE entre el MME y el HSS.

Antes de continuar es necesario mencionar a la serie de protocolos SIGTRAN, *signaling transfer,* los cuales fueron desarrollados para adaptar la tecnología las capas SS7.

El protocolo MAP, *mobile application part,* pertenece al conjunto de protocolos del sistema de señalización 7 o SS7, este sistema fue desarrollado en 1975 para reemplazar los protocolos utilizados en la PSTN. El protocolo MAP se encuentra en la capa de aplicación y es utilizado para la comunicación a nivel de esta capa entre los equipos HLR, VLR, MSC SMSC y SGSN para el control de la movilidad, manejo de llamadas de voz y control de datos, entre otros.

CAMEL, *customized applications for mobile networks enhanced logic,* es un protocolo utilizado en la arquitectura de red inteligente o IN diseñada para ofrecer servicios adicionales a los de telefonía móvil tradicional como los

teléfonos prepago y grupos de teléfonos o flotillas mencionados anteriormente. El SCC-AS, SCPs y MSCs utilizan este protocolo para comunicarse entre sí.

El protocolo ISUP, *ISDN user part,* donde ISDN es la abreviación para *integrated services digital network,* pertenece a la familia de aplicación del SS7 y es utilizado en la red conmutada por circuitos para establecer, manejar y gestionar llamadas de voz y datos. Con el uso de este protocolo se pueden tener los servicios de identificación, redireccionamiento y puesta en espera de llamadas.

BICC, *bearer independent call control, es* un protocolo muy parecido al ISUP, tanto así que los mensajes de señalización son similares; este protocolo fue desarrollado para soportar servicios de banda angosta o N-ISUP sobre una red de banda ancha, en otras palabras, para soportar servicios de voz independientemente de cuál sea la tecnología que se utiliza para transmitir la media.

El protocolo GTP, *GPRS tunneling protocol,* es un grupo de protocolo utilizados para transmitir GPRS a través de las redes 2G, 3G y LTE. Dentro de este grupo de protocolos se pueden mencionar: el control GTP o GTP-C que se utiliza para intercambio de mensajes de señalización entre el MME y el SRVCC IWF, S-GW y el SGSN, también, para la comunicación entre el SGSN y el GGSN. El protocolo User GTP o GTP-U es utilizado para el intercambio de tráfico de datos entre el eNodeb y el S-GW; también, entre la RNC y la SGSN y por último entre la SGSN y GGSN.

Otro protocolo que vale la pena mencionar es el SGsAP, *SGs application part,* que comunica al MME y a la MSC para el envío de mensajes de texto

sobre LTE y también para disparar el servicio de CSFB el cual fue expuesto anteriormente, para las llamadas terminantes.

Para finalizar el protocolo H.248 es utilizado para el control de los elementos de red MGW, IM-MGW y MRFP por los nodos MSC, MGCF y MRFC respectivamente, con el fin de establecer canales de voz.

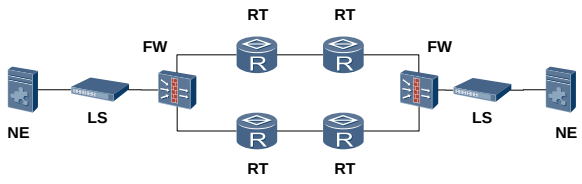
1.6.3. Redes de transporte IP

A lo largo de la evolución de las redes de telecomunicaciones han surgido distintas tecnologías en las redes de transporte, algunas ya mencionadas en enunciados anteriores; sin embargo, actualmente la más utilizada es la transmisión por IP.

1.6.3.1. Estructura red de transporte IP

Los elementos de una red de transporte IP operan en las capas inferiores a la tres del modelo TCP/IP. Este modelo básicamente establece las normas de conexión extremo a extremo dentro de una red; ofreciendo ventajas como alta fiabilidad y escalabilidad en redes de cualquier tamaño.

Figura 7. **Elementos de una red de transporte IP**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

En la figura 7 se pueden observar los elementos que componen una red de transporte IP.

A una red de transporte IP se le conoce también como *backbone* y su funcionalidad es la de proveer una o varias rutas para conectar distintos NE, *network elements,* que se encuentran en diferentes redes y los cuales a su vez utilizan diferentes protocolos para comunicarse.

Antes de continuar es necesario ahondar un poco en el modelo OSI, *open system interconnection*, ya que este define la estructura de muchos otros protocolos modernos que funcionan por IP. El modelo OSI se compone por 7 capas o niveles. Durante este enunciado se describirán las capas de la 1 a la 3 ya que son en las cuales funcionan los elementos de red de un *backbone.*

La capa 1 o física es la que establece las normas y características de la red material. Todos los elementos de red deben contar con las mismas características físicas para poder comunicarse entre sí.

En orden ascendente, la capa 2 o de enlace de datos, su función es la de proporcionar el servicio de acceso y envío de datos a través de la capa física, es decir, el direccionamiento físico. Los *lan switches* o LS, por sus siglas en inglés, trabajan en la capa 2, su función principal es la de proveer acceso a distintos elementos de red y de establecer comunicación directa entre varias redes.

La capa 3 o de red es la encargada de gestionar las conexiones y direccionamiento para las capas superiores, es decir, direccionamiento lógico. El *router* o enrutador, abreviado como RT en la figura 7, opera para conectar redes aisladas o en otro segmento de red.

El *firewall* o FW, como se hace referencia por igual en la figura 7, es otro elemento de red que funciona en la capa de red; su función principal es la de permitir o restringir el acceso no autorizado a la red por medio del filtrado de paquetes IP.

1.6.3.2. Protocolos de red de transporte IP

Así como los elementos de red trabajan en distintas capas, así también lo hacen los protocolos, que al final son los que permiten la comunicación correcta en una red IP; el modelo TCP/IP define todos los protocolos necesarios para este fin.

Tabla II. **Modelo TCP/IP**

Nombre de Capa	Protocolos
Capa de aplicación	HTTP - SIP
Capa de transporte	TCP - UDP - SCTP
Capa de nternet	IP - IPSEC
Capa de enlace	ARP - Ethernet - MAC

Fuente: elaboración propia, utilizando programa Microsoft Excel.

En la tabla número 2 se puede apreciar la pila de protocolos del modelo TCP/IP. El protocolo *Ethernet,* definido en la IEEE 802.3, define las características de cableado y señalización a nivel físico para la comunicación de dispositivos en una red de área local o LAN, a su vez establece el formato de las tramas de datos a ser transmitidas a capas superiores.

El protocolo MAC o control de acceso define una serie de métodos y algoritmos para regular el acceso a un medio físico compartido por varios

elementos de red, cada interfaz física de un elemento de red se identifica individualmente por una física, es decir, una dirección MAC.

El protocolo IP, *Internet protocolo,* es el principal protocolo en el modelo TCP/IP; su tarea principal es la de entregar paquetes desde un elemento de red origen hacia un destino basado en las direcciones IP. Una dirección IP es una etiqueta numérica que identifica individuamente a un elemento de red o a una interfaz de forma lógica y jerárquica en una red basada en el modelo TCP/IP.

IPSEC, *IP security,* es un conjunto protocolos cuya función principal es la de asegurar las comunicaciones cuando se utiliza el protocolo IP. El protocolo ARP, *address resolution protocol,* tiene como función principal encontrar la correspondencia de una dirección MAC y una dirección IP.

En la capa de transporte se encuentra el protocolo TCP, *transmission control protocol*, cuya función básica es asegurar una transmisión confiable entre un cliente y un servidor al establecer una conexión previa antes de empezar a transmitir la información. Además, permite la retransmisión de paquetes si en algún momento se llegara a interrumpir la comunicación entre dos elementos de red.

El protocolo UDP, *user datagram protocol,* a diferencia del protocolo TCP no está orientado a la conexión, debido a que un datagrama transmitido hacia su destino cuenta con la información necesaria para llegar al mismo. En este protocolo no existe la retransmisión ni la verificación de paquetes. Se utiliza UDP normalmente cuando se requiere un tiempo de respuesta corto o cuando se transmite audio o video en tiempo real, por lo cual no tiene sentido la retransmisión.

En la capa de aplicación tenemos al protocolo HTTP que está orientado a conexiones cliente servidor para la transmisión y presentación de texto o media. Es el protocolo utilizado en la *web wide web* (www) o red de informática mundial. El protocolo SIP, también en la capa de aplicación, fue desarrollado para el establecimiento y control de sesiones multimedia.

2. TELEFONÍA SOBRE IP

En las redes de telefonía moderna se utilizan los protocolos sobre IP ya que ofrecen una amplia gama de servicios en su capa de aplicación por su bajo costo de implementación y la facilidad de su uso a través de Internet lo cual reduce también el costo para el usuario final.

2.1. Bases de la telefonía IP

Una infraestructura de telefonía sobre IP consiste en diferentes componentes en esta sección se realiza una descripción de los componentes físicos y de los protocolos por medio de los cuales funciona.

2.1.1. Terminales IP

Una terminal es un punto final de comunicación donde terminan llamadas o flujos de media. Comúnmente puede ser un dispositivo físico o un programa para establecer llamadas de voz o video; normalmente estas terminales tienen la capacidad de utilizar servicios de datos que se proveen a través de la red móvil.

A una terminal de telefonía IP se le asigna por lo menos una dirección IP, puede que varias terminales utilicen la misma IP también, aunque cada una es tratada de manera independiente.

En la mayoría de casos, una terminal tiene asignadas una o varias direcciones IP que otras terminales utilizarán para marcar a esta. Si se utilizan servidores IP, la terminal se registra con su dirección IP en un servidor.

2.1.2. Servidores de telefonía IP

Para realizar una llamada IP se requieren por lo menos dos terminales y al menos una en el dominio de la red de telefonía IP. Adicional el conocimiento de la IP y puerto de la terminal a llamar. Por motivos obvios es poco práctico que un usuario utilice la dirección IP para llamar ya que esta consta de 4 números entre 0 y 255; por otro lado, es muy común que las direcciones IP se asignen dinámicamente.

Usualmente, una terminal se registra en un servidor el cual almacena un número telefónico tradicional y su IP correspondiente y así de esta forma se puede vincular lo que hace que la parte IP sea trasparente para el usuario final.

Cuando un usuario marca un número telefónico el servidor trata de resolver el mismo en una dirección IP para hacer esto el servidor puede interactuar con otros servidores o servicios dentro de la red de telefonía.

Un servidor puede ofrecer servicios adicionales como mecanismos de enrutamiento, por ejemplo, desviar una llamada a un grupo de contactos cuando se marca un número de soporte. Finalmente, el servidor es el encargado de autenticar, registrar y autorizar a una terminal para el uso de los servicios.

2.1.3. Puerta de salida en una red IP

Las puertas de salida o *gateways* son puntos en una red de telefonía IP que conectan usuarios de diferentes redes y que normalmente no tienen comunicación entre sí.

Normalmente un *gateway* no solo traduce un protocolo de señalización a otro, por ejemplo, de ISDN a SIP o viceversa, sino que también traslada direcciones IP en diferentes redes, por ejemplo IPv4 a IPv6, adicionalmente también convertir de un tipo de códec de audio a otro.

2.1.4. Direccionamiento

Para que un usuario pueda utilizar los servicios de telefonía sobre IP tiene que poseer un identificador para sí mismo y para el usuario con el cual desea comunicarse. Idealmente, este identificador es independiente de la ubicación física de los usuarios. La red de telefonía es la responsable de encontrar la ubicación actual del número de A o de quien quiere iniciar el servicio.

Regularmente los sistemas de telefonía utilizan el formato E.164, este identificador puede tener hasta de 15 dígitos liderado por un símbolo “+”, el formato es el que se muestra en la tabla a continuación.

Tabla III. **Formato E.164**

Country code	***National destination code***	***Subscriber number***
CC[1-3 dígitos]	NDC+ SN[12-14 dígitos]	

Fuente: elaboración propia, utilizando programa Microsoft Visio.

Normalmente cuando se realiza el marcado, el símbolo más se reemplaza por doble cero (00) o código internacional de llamadas, un ejemplo del formato E.164 para Guatemala seria el +50250001111, siendo 502 el código de país o *country code* y el 50001111 es el número de subscriptor, *subscriber number,* en este caso por la cantidad de usuarios no es necesario establecer un código nacional de destino, *national destination code*.

Actualmente los sistemas de telefonía IP utilizan dos tipos de identificadores: el URI definido por el RFC2396 y el E.164.

Un URI, *universal resource identifier*, utiliza un nombre registrado para describir un recurso de una manera independiente a la ubicación. Estos recursos están disponibles bajo una variedad de esquemas de nombramiento y métodos de acceso como por ejemplo direcciones de correo electrónico, identificadores SIP, identificadores H.323 o números telefónicos IPtel (RFC2806). Los identificadores tipo correo electrónico tienen varias ventajas, una de ellas es que son fácil de recordar ya que cualquier usuario que utilice el Internet posee una dirección de correo electrónico.

La ubicación de un subscriptor puede ser encontrada utilizando el sistema DNS. Si se desea integrar la telefónica habitual con el sistema de telefonía sobre IP, se debe tratar con los números de identificación telefónica aun del lado de la telefonía IP. Los números no son bien recibidos para el mundo del Internet debido a que en este se utilizan nombres de dominio. Pare resolver este problema el sistema ENUM fue creado, y de esta forma se pueden adaptar números telefónicos a nombres de dominio.

2.1.5. Protocolo H.323

El alcance del protocolo H.323 se enfatiza en la comunicación multimedia sobre redes basadas en IP, tomando en cuenta la comunicación hacia redes basadas en conmutación por circuitos.

El estándar H.323 define todo lo necesario para llevar a cabo la comunicación de audio y video en redes de área local, aunque posteriores revisiones ampliaron su uso en Internet y redes de mayor tamaño; H.323 consta de una serie de protocolos los cuales se describen en la tabla IV.

Tabla IV. **Protocolos H.323**

Protocolo	Descripción
H.225	Se utiliza para el control de las llamadas.
H.235	Se utiliza para la seguridad y cifrado.
H.245	Es utilizado para la señalización y control de los canales multimedia.
H.450	Su tarea es controlar los servicios suplementarios.
RTP	Utilizado para el transporte de contenido multimedia.
T.120	Se utiliza como protocolo de datos para conferencias multimedia.

Fuente: elaboración propia, utilizando programa Microsoft Excel.

Una red H.323 posee cuatro elementos principales, cuenta con una terminal la cual permite establecer conferencias de audio, datos o video bidireccionalmente. Cada terminal debe soporta la decodificación de formatos de audio utilizados en la telefonía tradicional, como el G.711, G.722, G.723.1, G.728 y G.729. Al igual que formatos de video como el H.261.

Otro elemento de red es el *gatekeeper* que es considerado el cerebro de la topología de una red H.323, este elemento provee control sobre el ancho de banda, limitación de conexiones simultaneas, conversión del estándar E.164 al formato H.323, control de admisión de terminales, restricción de llamadas y lista de llamadas en espera.

La MCU, *mutipoint control unit,* se encarga de soportar conferencias entre tres o más terminales H.323. Un MCU realiza la negociación de códec entre las terminales que desean establecer una conferencia.

El último elemento se denomina *gateway* y tiene como función conectar las redes H.323 a otras redes no H.323. Es decir, su función principal es traducir los distintos protocolos para establecer y finalizar una llamada

El protocolo H.323 está en desuso actualmente y es reemplazado por el protocolo SIP, sin embargo, vale la pena mencionarlo ya que es la base del inicio de las redes de telefonía IP.

2.1.6. Protocolo SIP

El estándar SIP, *sesión initiation protocol,* opera en la capa de aplicación del modelo TCP/IP. Este protocolo fue diseñado de tal forma que su implementación y escalabilidad son bastante flexibles.

El protocolo SIP es utilizado para crear, modificar y terminar sesiones de uno o más participantes. Una sesión se define como una serie de transmisores y receptores que comunican y mantienen su estado entre ellos mismos a lo largo de la comunicación.

SIP no es el único protocolo que las terminales van a necesitar para establecer la comunicación; su propósito principal es solo el de hacer la comunicación posible. Los protocolos más utilizados conjunto con SIP son el RTP y SDP.

El protocolo RTP, *real time protocol,* es utilizado para transmitir voz, audio o texto en tiempo real, lo hace posible ya que divide la información en paquetes los cuales envía a través de Internet u otra red basada en este protocolo. El protocolo SDP, *session description protocol,* es usado para establecer las capacidades de codificación de una sesión. Como por ejemplo el códec de audio a utilizar al igual que el protocolo de transporte.

SIP está basado en el protocolo HTTP, es el protocolo más usado en Internet, trata de combinar lo mejor de ambos. Las entidades SIP se identifican usando SIP URI, que consiste en dos partes: el nombre de usuario junto con el nombre del dominio unidos por el símbolo arroba o "@". Como se puede deducir las direcciones SIP URI son similares a la de los correos electrónicos.

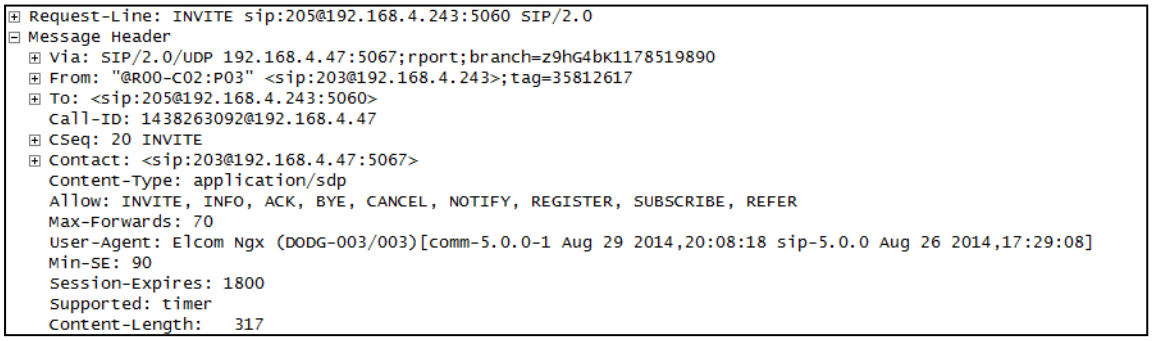
2.1.6.1. Mensajes SIP

La comunicación utilizando SIP está compuesta por una serie de mensajes que pueden ser transportados independientemente de la red. Usualmente cada mensaje es enviado en datagramas UDP distintos.

Existen dos tipos de mensajes SIP: las solicitudes y las respuestas; las solicitudes se utilizan usualmente para iniciar una acción o para informar al receptor algún evento; las respuestas se utilizan para confirmar que una solicitud fue recibida y procesada, también, puede contener el estado de la solicitud.

Entre los mensajes de solicitud se encuentran: INVITE, que es utilizado para establecer una sesión; en la figura 8 se puede ver el formato del encabezado de este mensaje.

Figura 8. **Encabezado mensaje INVITE**



Fuente: elaboración propia, utilizando programa Wireshark.

En la primera línea se indica el tipo de mensaje y la URI destino. El mensaje SIP puede tener una o más campos VIA, en el cual se guarda la URI solicitante. Este campo se utiliza para establecer la trayectoria para los mensajes de respuesta.

Los campos *from* y *to* identifican al número de A (203) y al número de B (205) respectivamente. El campo *call-ID* es utilizado como un identificador de diálogos y su propósito es identificar los mensajes que pertenecen a la misma llamada. El campo *contact* contiene la dirección IP y el puerto a el cual el número de A espera que el número de B le conteste las solicitudes. Los demás campos no son de tanta relevancia como los descritos anteriormente.

Otro mensaje de solicitud es el ACK, este mensaje reconoce como recibido a una respuesta final de un mensaje INVITE. El mensaje BYE se utiliza

para finalizar una sesión multimedia. Un mensaje CANCEL se utiliza para abortar una sesión que aún no ha sido totalmente establecida.

El propósito del mensaje REGISTER es el de hacer saber a la red cuál es la ubicación de un usuario. La información del puerto y la IP por el cual un usuario debe ser contactado se encuentra en este mensaje.

Cuando un usuario o un servidor reciben un mensaje de solicitud, este envía un mensaje de respuesta; todas las solicitudes son contestadas a excepción del mensaje ACK.

Un mensaje típico de respuesta se puede observar en la imagen a la figura 9.

Figura 9. **Mensaje SIP OK**

Fuente: elaboración propia, utilizando programa Wireshark.

Los mensajes de respuesta son similares a los mensajes de solicitud, a excepción de la primera línea. La primera línea de este mensaje contiene un código de respuesta y una frase relacionada con el código. El código de

respuesta es un número entero entre el 100 y el 699. Hay 6 clases de respuestas, las cuales se resumen en la tabla V.

Tabla V. **Códigos de respuesta SIP**

Código de respuesta	Descripción	Ejemplo
1XX	Son mensajes provisionales que indican una solicitud en proceso.	180 (ringing)
2XX	Son mensajes que indican una respuesta final positiva.	200 (OK)
3XX	Son mensajes finales que indican re direccionamiento.	301 (moved permanently)
4XX	Son mensajes que indican una respuesta final negativa.	400 (bad request)
5XX	Indica un problema del lado del servidor.	504 (server time-out)
6XX	Este mensaje indica que ningún servidor puede completar la solicitud.	603 (decline)

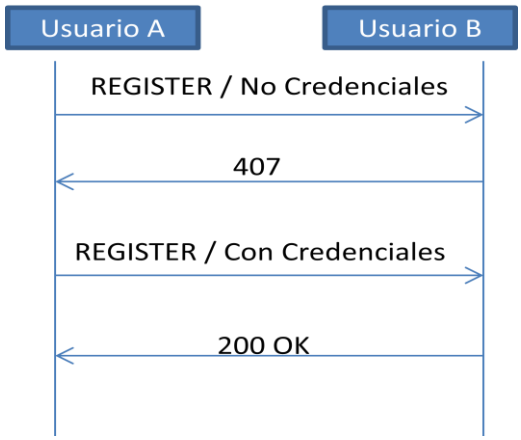
Fuente: elaboración propia, utilizando programa Microsoft Excel.

2.1.6.2. Escenarios típicos SIP

En este enunciado se realiza una descripción breve de los flujos básicos para establecer el tráfico SIP.

Un usuario debe registrarse con la red para poder ser alcanzado por otros usuarios, en la figura 10 se muestra el flujo de mensajes para un registro.

Figura 10. **Registro de un usuario**



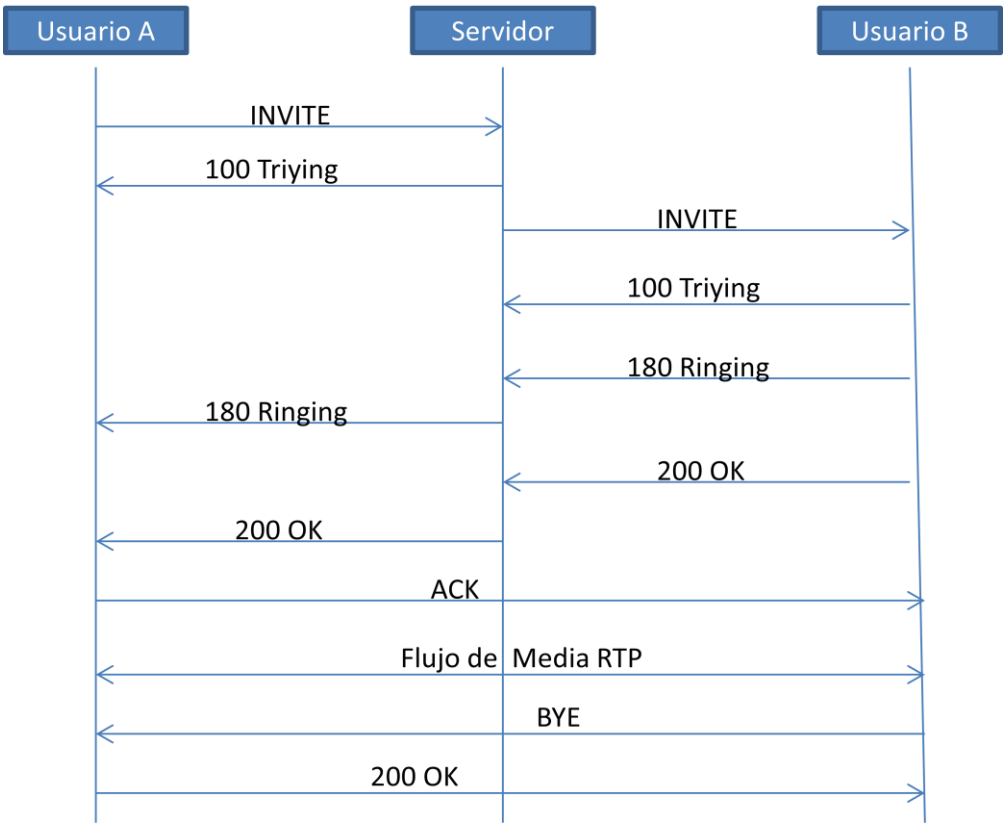
Fuente: elaboración propia, utilizando programa Microsoft PowerPoint

Para iniciar y terminar una sesión, debe generarse un mensaje INVITE, seguido de mensajes de procesamiento, el cual puede ser un mensaje 100 TRIYING que ayuda a detener las retransmisiones del mensaje INVITE indicando que la solicitud está siendo procesada.

Todas las respuestas provisionales generadas por el número de B son siempre enviadas al número de A.

Un mensaje 200 OK es generado una vez que el número de B contesta la llamada, este mensaje es retransmitido hacia el número de B hasta que B recibe un mensaje ACK, en este punto es donde se establece la llamada.

Figura 11. **Flujo de Inicio y finalización de una sesión SIP**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

Para terminar una sesión SIP se utiliza el mensaje BYE dentro del diálogo previamente establecido por un mensaje INVITE. Este mensaje es transmitido directamente de extremo a extremo. El mensaje 200 OK es la respuesta a un mensaje BYE, indicando así que la sesión ha sido terminada.

El protocolo RTP se utiliza para establecer el flujo de transmisión de voz, media o texto entre los usuarios finales. Este protocolo es descrito más a detalle en el siguiente enunciado.

2.1.6.3. Protocolos RTP y RTCP

Los protocolos RTP (*real time protocol*) y RTCP (*real time control protocol*) son los protocolos de transporte utilizados en los flujos de media para la telefonía sobre IP. Ambos son definidos en el RFC1889.

RTP fue desarrollado para transportar datos en tiempo real, mientras que el protocolo RTCP para monitorear la calidad del servicio y a la vez comunicar información sobre los participantes de una sesión en progreso.

Dentro de los servicios que ofrece el protocolo RTP están: identificación de la información a transportar; verificación de la entrega de paquetes en el orden establecido y reordenamiento de los mismos si es necesario; identificación del códec y decodificador para el transporte, y monitoreo de la información entregada.

Para manejar múltiples sesiones entre dos entidades distintas y para validar la integridad de los datos, el RTP utiliza el protocolo UDP de una manera subyacente. Vale mencionar que en ninguna parte del transporte este protocolo asegura la calidad del servicio durante un flujo de transmisión de información.

RTCP utiliza los mismos protocolos que RTP para enviar periódicamente paquetes de control a todos los participantes de una sesión. Entre los servicios que ofrece está la de retroalimentar la calidad de la distribución de la información, así como también de los códecs activos en una sesión.

Dentro de otras funcionalidades del RTCP está la de cuantificar el número de sesiones para ajustar la tasa de transmisión del RTP, asimismo el de controlar la información para identificar a los participantes de una sesión.

2.2. Escenarios de telefonía sobre IP

En este enunciado se intenta definir los escenarios más comunes en los cuales se utiliza la solución de telefonía sobre IP, también se hace ver las razones principales por la cual los operadores adoptan cada opción.

2.2.1. Enrutamiento por la ruta de menos costo

Es utilizado por operadores con grandes volúmenes de llamadas. Tradicionalmente, diferentes enlaces son utilizados para transmitir la voz conmutada por circuitos y la data IP (Internet) entre dos sitios, los precios se pueden reducir estableciendo cuentas con empresas que prestan el servicio de transporte de datos (*carriers*) por bajo costo para largas distancias.

Una alternativa a esta solución es la telefonía sobre IP ya que es posible que sobre la red existente de datos IP pueda enviarse el tráfico de voz a largas distancias, reduciendo así los costos totales de transporte. Esta opción requiere que las llamadas tradicionales se conviertan a la telefonía IP. Habitualmente para el transporte de datos se utilizan diferentes *carriers,* dependiendo de dos cosas: cuál es el destino final de la información y cuál es el costo del transporte.

Las tarifas de los *carriers* varían semana a semana y hasta día con día debido a la gran competencia que existe en el mercado, puede que dependiendo la hora o día exista un tipo de descuento dependiendo del volumen de tráfico; esto se hace para compensar un déficit del tráfico que el *carrier* pueda enfrentar ya que entre más tráfico trasporte un *carrier* más se abarata la tarifa en el mercado de mayoreo.

Para resolver este inconveniente se puede tener un Servidor LCR, *low cost route,* al cual se carga el listado de rutas y con su respectivo costo, así el LCR puede elegir cuál; *carrier* enviar el tráfico. Luego de elegir un *carrier* este servidor envía órdenes hacia las centrales de telefonía para modificar las rutas de salida.

Entre ventajas que ofrece esta solución es el enrutamiento de llamadas por periodo del día, por día y por semana, permitiendo la selección del *carrier*, basado en la ruta de menor costo, el mismo concepto aplica para el enrutamiento de llamadas por destino. Otra de las ventajas es la manipulación de los prefijos de marcación, permitiendo la modificación para facilitar el enrutamiento a los diferentes destinos y *carriers*.

2.2.2. Redes alternativas a los sistemas PBX

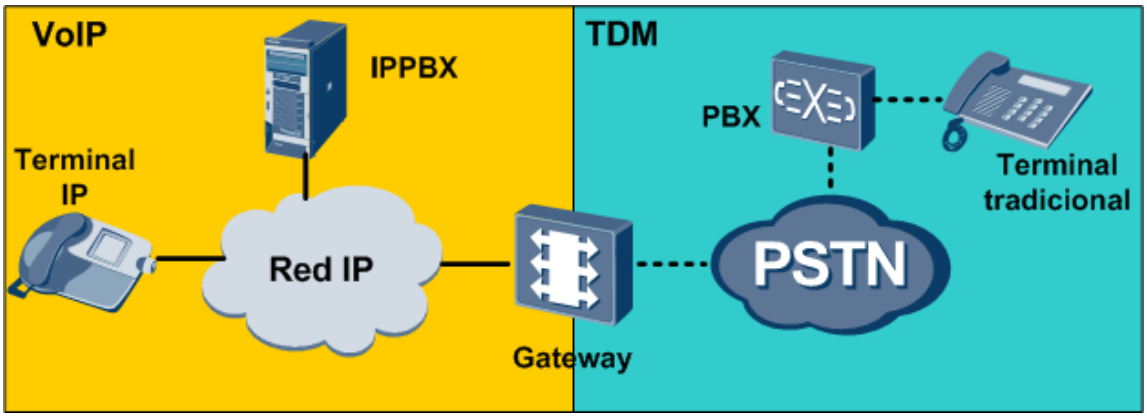
Un PBX, *private brand exchange,* es cualquier central telefónica que se conecta a la PSTN por medio de líneas troncales normalmente utilizando uno o varios E1s. Un PBX gestiona las llamadas internas o externas tanto entrantes como salientes. Normalmente es un servicio que los operadores prestan a empresas o compañías para la comunicación interna de los distintos departamentos.

Una de las soluciones más factibles es la de reemplazar los sistemas PBX, que corren sobre una PSTN, a los sistemas de voz sobre IP o VoIP. Esto requiere la instalación de terminales y centrales IP, denominadas IPPBX. La telefonía sobre IP permite también que ambas tecnologías puedan subsistir dentro de la misma red.

La solución más sencilla es la de establecer una conexión punto a punto entre los teléfonos IP. Para realizar una llamada se necesita del conocimiento de la IP de cada terminal. Esta solución es poco factible debido a que el método de marcación no es tradicional. Adicional, no existe la opción de establecer comunicación entre usuarios fuera de la red VoIP.

Debido a que la solución anterior tiene muchas desventajas, entonces es necesario establecer una solución en la cual el sistema legado de las redes PSTN pueda convivir con las redes VoIP. Este escenario se aplica cuando es necesario construir la red de telefonía sobre IP gradualmente; un ejemplo de esta solución se puede ver en la figura 12.

Figura 12. **Integración de la red VoIP con el sistema legado PBX**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

La mayoría de IPPBX puede soportar terminales análogas o IP, debido a que cuentan con interfaces dedicadas para soportar ambas tecnologías. Ofrece servicios como llamadas de emergencia, integración con la telefonía inalámbrica e integración de servicios de voz y data

Otra solución de telefonía sobre IP que está tomando mucho auge, como servicio suplementario de las redes móviles, es la voz sobre *wireless* o VoWLAN; es decir, VoIP sobre redes inalámbricas. Básicamente esta tecnología se utiliza para descongestionar la red móvil en lugares donde transita diariamente mucha gente como centros comerciales, estadios o iglesias.

Otra funcionalidad que ofrecen las redes VoWLAN es la de establecer llamadas, desde un dispositivo móvil, sin costo entre usuarios que pertenecen a una misma organización.

2.3. Establecimiento de los servicios sobre IP

En el enunciado anterior se expuso las soluciones en donde se utiliza la telefonía sobre IP es momento, entonces, de definir algunos conceptos necesarios para entender cómo implementar y cómo funciona la telefonía sobre IP o VoIP.

2.3.1. Planes de marcación

Un plan de marcación no es más que el rango de números telefónicos disponibles en una red de telefonía. Usualmente cuando se desea implementar VoIP ya existe un plan de marcación para la red PSTN. Entonces, surge el inconveniente de tener que distinguir los números de legado PSTN de los números VoIP.

Una solución para este inconveniente es el de asignar bloques numéricos dedicado para los teléfonos VoIP. Si dentro de los rangos de números PSTN existen números libres, entonces se pueden realizar bloques y reservarlos

únicamente para la telefonía sobre IP; entre más grande sea el rango que se pueda reservar, más fácil será la integración de VoIP.

El inconveniente de esta solución es que solo se pueden asignar rangos nuevos a los teléfonos IP, es decir, si un usuario desea migrar su servicio PSTN a VoIP necesita cambiar su numeración.

Otra solución a este inconveniente es definir algún prefijo que debe ser marcado antes del número al cual se desea llamar, con esta opción es posible migrar un número legado a uno de telefonía IP haciendo una pequeña modificación en el número original que es fácil de recordar. El problema de esta solución es que no es posible saber si el número al cual se desea llamar es un teléfono IP. Para evitar esto, se debe hacer que el prefijo sea invisible para el usuario final, por ejemplo, por medio de una tabla en un PBX que decida a que números les agrega o no un prefijo.

Sin embargo, la forma más transparente de hacer esto es identificar individualmente a un usuario PSTN de un VoIP. Por otro lado, es la configuración que requiere más esfuerzo, puesto que cada vez que se dé de alta o se migre un número a VoIP se debe hacer un enrutamiento de manera individual.

El enrutamiento se vuelve más complicado cuando la cantidad de protocolos de señalización y servicios crece. Por lo que es común utilizar combinaciones de las soluciones descritas anteriormente y tener un servidor designado como *gateway* el cual es el encargado de tomar la decisión de a donde transferir cada llamada. Todos estos problemas de marcación y su solución son dependientes de la marca de nuestros equipos de red, ya que no todos ofrecen las mismas funcionalidades.

2.3.2. Autenticación en SIP

En este enunciado se intenta exponer el mecanismo de autenticación utilizado en las redes que utilizan el protocolo SIP.

2.3.2.1. Autenticación de teléfonos IP

Digest es un tipo de autenticación que se utiliza en SIP, este sistema fue desarrollado originalmente para HTTP, el protocolo base de SIP es el HTTP. En el RFC2671 se describe la autenticación *digest.* Este tipo de autenticación busca que no se transfiera en texto plano el usuario y contraseña con el cual una terminal se registra en la red de telefonía IP como lo puede ser un IMS.

Cuando un servidor desea autenticar a un usuario y este genera un mensaje 401 (*unauthorized*) en respuesta al mensaje INVITE, en figura 13 se puede ver la estructura del mensaje.

Figura 13. **Mensaje SIP 401**

Fuente: elaboración propia, utilizando programa Wireshark.

La autenticación consiste en una serie de parámetros enviados a un usuario, el usuario entonces debe generar una respuesta correcta a partir de esta información.

Dentro de los parámetros está el campo *realm* que es un campo obligatorio y debe estar presente en todas las solicitudes de autenticación. Normalmente, el valor de este campo es el dominio al cual pertenece el servidor SIP.

El campo *nonce* es una serie de caracteres (*string*) únicos que se generan cada vez que un servidor requiere que un cliente se autentique. Para generar este *string* de datos se utiliza una frase secreta, una estampa de tiempo de expiración y un algoritmo de encriptación, por lo general se utiliza el MD5 (*message-digest algorithm 5)*.

Los clientes utilizan el campo *nonce* para generar una respuesta ante un mensaje de registro, por lo general un servidor verifica la validez del mensaje antes de procesar la información de usuario y contraseña.

Los mensajes de respuesta y solicitud de registro son similares entre sí, sin embargo, incluyen campos adicionales como el nombre de usuario, *username,* la URI a la cual el usuario quiere utilizar, el nivel de protección elegido por el cliente, *qop*; el valor hexadecimal *nc (nonce count)* en el cual se lleva el registro de las solicitudes que un cliente ha realizado con el mismo valor *nonce.* Otro campo a mencionar es el *response,* que es un *string* con el cual el cliente demuestra que conoce la contraseña para autenticarse.

2.3.2.2. Autenticación de telefonía sobre IP

Cuando se logra establecer el flujo de registro de un usuario en un sistema cliente servidor, sin embargo, como se estableció anteriormente, existen escenarios donde únicamente se utiliza la telefonía IP para transportar tráfico de llamadas de voz, como por ejemplo llamadas desde teléfonos móviles o desde una PBX legado, hacia destinos de larga distancia.

En este tipo de escenarios no existen las solicitudes de registro de usuario, solo múltiples sesiones SIP que se transfieren a lo largo de redes IP normalmente públicas.

Para poder transportar tráfico se deben establecer troncales o conexiones SIP entre las dos partes que desean intercambiar información. Cuando se transmite tráfico internacional por lo general se utilizan conexiones sobre Internet que, como se verá en el siguiente capítulo de este trabajo de investigación, suelen ser conexiones inseguras.

En este tipo de redes, más que autenticar, se utilizan elementos de red que filtran paquetes IPs. Estos dispositivos deciden si deben procesar o no las solicitudes, basados en las IPs y puertos lógicos origen.

Uno de los desafíos más grandes es la interoperabilidad de servicios en las comunicaciones de larga distancia, debido a las normas y políticas internacionales de otros países o de los *carriers* en sí; en algunos casos no existen tales y esto puede conducir a que se produzcan diferentes tipos fraudes de con el uso de este tipo de escenarios.

2.4. Integración de la telefonía global

En esta sección se describen los mecanismos y protocolos para proveer resolución de direcciones en diferentes dominios cuando se desean integrar redes sobre telefonía IP.

2.4.1. Enrutamiento de la telefonía sobre IP

TRIP, *telephony routing over IP,* es un protocolo definido en el RFC3219; este se utiliza para anunciar la información de accesibilidad de un número telefónico, por ejemplo E.164, entre dos distintos dominios administrativos. El protocolo TRIP es similar al protocolo BGP (*boarder gateway protocol*); el protocolo BGP se utiliza para intercambiar datos de enrutamiento capa 3 entre dos *routers.*

Dentro del protocolo TRIP se definen dominios administrativos de telefonía IP o ITAD (IP *telephony administrative domains*). Cada uno de estos dominios tiene un número único que lo identifica globalmente.

Un ITAD consiste en uno o más servidores de ubicación, de los cuales por lo menos uno tiene relación remota con otro servidor de otra ITAD. La comunicación entre distintas ITAD, que no están conectadas entre sí, se da de forma de salto a salto entre ITAD.

El objetivo del desarrollo de este protocolo es el de ser un mecanismo en la selección de *gateways* de egreso y es limitado únicamente al rute de números telefónicos, por lo que no es posible anunciar URIs o nombres, por lo cual hace que TRIP sea inadecuado para todos los tipos de enrutamiento inter dominio.

2.4.2. ENUM

Uno de los problemas principales de la telefonía IP es la integración con la PSTN, ya que más de un protocolo de señalización es utilizado para los distintos servicios, se deben incluir todos a la hora de realizar esta integración. Es necesario entonces establecer un identificador universal para unificar ambos mundos, para este objetivo se utiliza el E.164 definido por la ITU-T.

Entonces se necesita encontrar un sistema que pueda ser implementado a nivel mundial. Este sistema debería ser capaz de proveer información para traducir dicha información para la comunicación entre las redes PSTN y VoIP. Para esta tarea se consideró el protocolo DNS ya que es utilizado probablemente en todos los sistemas cliente servidor sobre Internet.

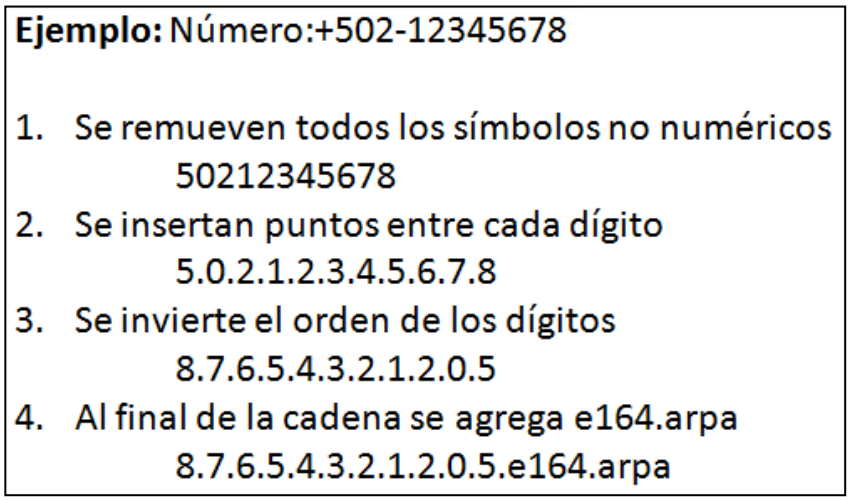
ENUM es una innovación del sistema DNS, su estructura comienza con nombre de dominios numéricos que son utilizados para consultar nombres de servidores DNS, los servidores responden con la información de la dirección encontrada en los registros DNS. La respuesta puede ser un número de teléfono, direcciones de correo electrónico, direcciones SIP u otra información. El concepto es el de usar un solo número para obtener toda la información de un contacto.

Los números ENUM son convertidos por servidores ENUM en nombres de dominio y luego son utilizados para consultar el nombre del dominio del sistema. Un servidor ENUM se encuentra en Internet, es decir, la consulta se realiza sobre Internet y la base de datos con toda la información al igual esta sobre esta red.

El grupo de trabajo de la IETF para ENUM decidió que un número ENUM debe tener el mismo valor que el número de teléfono de una persona, con esto se logra crear un estándar global de cómo deben asignarse estos números y debido a que los números ENUM, que son nombres de dominio, son numéricos, es más fácil marcarlos y muy similares a un número de teléfono tradicional.

La figura 14 muestra cómo se convierte un número con el formato E.164 a un nombre de dominio ENUM, para ser utilizado globalmente.

Figura 14. **Conversión de E.164 a ENUM**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

3. VULNERABILIDADES DE LAS REDES DE TELEFONÍA IP

Existen muchos beneficios con el uso de la telefonía sobre IP o VoIP, por ejemplo, el bajo costo para el usuario final y para el operador, movilidad y flexibilidad de la tecnología. Sin embargo, esta tecnología tiene algunas amenazas de seguridad que hay que tomar en consideración.

Uno de los principales problemas es la filtración o la falta de seguridad para almacenar o transmitir información relacionada con la topología, ubicación de nodos, cuentas de acceso, etc. Con la revelación de la información las vulnerabilidades aumentan.

Dentro de este capítulo se hace una compilación de las amenazas más conocidas dentro del mundo de las comunicaciones sobre IP.

3.1. Vulnerabilidades heredadas

En un principio un operador de telefonía no era ajeno a la relación con un gobierno o un país, debido a esto no hubo ningún inconveniente para interconectar todas las redes de diferentes países, puesto que permitía la comunicación con cualquier parte del mundo.

3.1.1. Vulnerabilidad de la red SS7 y SIGTRAN

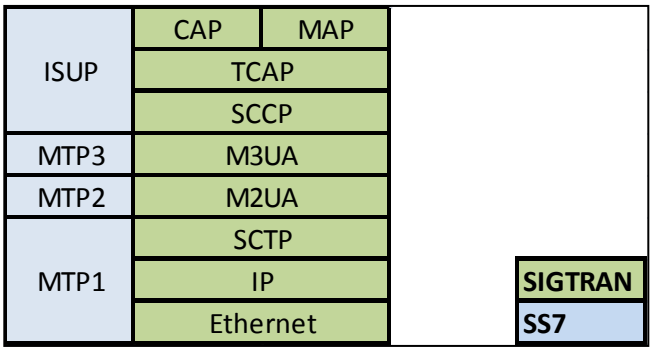
Las redes SS7 fueron diseñadas para topologías cerradas y en ningún momento se consideró su seguridad. La facilidad para conectarse a una red

SS7 es relativa ya que las regulaciones de cada país son distintas sobre quién, cómo y qué pasos hay que seguir para poder utilizar un servicio.

El sistema SIGTRAN fue desarrollado para adaptar el sistema SS7 a la tecnología IP, por lo cual también se encuentran las mismas vulnerabilidades que con SS7.

Antes de continuar más a detalle se requiere definir la pila de protocolos del sistema SS7 y SIGTRAN, este sistema se utiliza como base para la comunicación redes móviles.

Figura 15. **Pila de protocolos SS7 y SIGTAN**



Fuente: elaboración propia, utilizando programa Microsoft Excel.

Ninguno de los protocolos de la figura 15 es utilizado para establecer comunicación segura entre dos nodos de una red de telefonía legado, su función únicamente es la de establecer la comunicación de tal forma que la información pueda ser interpretada entre dos elementos de red distintos.

En SS7 y SIGTRAN los conceptos de autenticación y autorización son vanamente usados debido a que la comunicación entre nodos está basada exclusivamente en la confianza.

Considerada la red SS7 como una red cerrada, se han hecho pocos avances para evaluar los temas de seguridad. Los investigadores en el campo de la seguridad no tienen acceso a la red SS7 y los proveedores de servicios tienen poco interés en este tema.

Actualmente la red SS7 no es una red cerrada, los proveedores de servicios de red han abierto sus redes a terceros como parte de sus ofertas comerciales; algunos elementos de red se instalan en lugares poco confiables donde hackers pueden encontrar acceso en la red de un operador.

3.1.1.1. Reconocimiento de la red SS7

Actualmente existen herramientas capaces de realizar un escaneo de direcciones de IP utilizando el protocolo SCTP (*stream control transmission protocol*) como transporte; este protocolo es una mejora a los protocolos TCP y UDP ya que proporciona confiabilidad, control de flujo y secuenciación como TCP; a la vez, que permite el envío de mensajes fuera de orden puesto que está orientado el envío de mensajes.

Como se observa en la figura 15, este protocolo se utiliza en la tecnología SIGTRAN. SCTP es el estándar para establecer la conexión de señalización entre 2 nodos cuando se comunican por IP en una red telefónica.

El intercambio de la información de los subscriptores se realiza mediante mensajes SS7 a lo largo de las redes, un hacker puede encontrarse en

cualquier parte ya que los mensajes pueden ser enviados entre países. Un hacker puede adquirir acceso existente a la red SS7 de un operador en el mercado negro y obtener acceso para dar servicios de operadores en países que carecen de leyes.

Un hacker que trabaje como especialista técnico para un operador de telecomunicaciones puede ser capaz de conectar su sistema con la red SS7 del operador. Para realizar algunos tipos de ataques, conexiones legítimas deben ser utilizadas.

Con el acceso a una red legítima garantizada, es entonces posible realizar escaneos de redes en distintas partes del mundo con el fin de sustraer información de los suscriptores o interrumpir el servicio en alguna red.

3.1.1.2. Inyección de mensajes

Luego que un atacante logre identificar la topología e información de la red, le es posible asociar su propio nodo con el fin de poder enviar mensajes MAP; la mayoría de los ataques a nivel SS7 que suceden actualmente se dan sobre las redes móviles.

Como se mencionó anteriormente, la función del protocolo MAP es el control de la movilidad y la comunicación con el HLR donde se almacena toda la información de un usuario. Así que es posible enviar mensajes MAP, por ejemplo, para solicitar la ubicación de enrutamiento de un usuario o SRI (*send routing information*) con el cual se pueden obtener datos como IMSI, IMEI y VLR del usuario de B o evidenciar que el usuario no existe dentro de esa red.

El IMSI es utilizado para identificar un usuario globalmente en las redes móviles, consta de 15 dígitos los cuales incluyen el código de país de la red móvil, el código local móvil del operador y el número del suscriptor móvil o MSISDN (*mobile station international subscriber directory number*). Este número se encuentra grabado en una tarjeta SIM o USIM. Es como el nombre que utilizará el móvil para registrarse en la red.

El IMEI (*international mobile system equipment identity*) es el identificador que utiliza una terminal o equipo móvil, es decir, varía según el fabricante del teléfono. En una red móvil existe un elemento de red cuya función es permitir o no el acceso al equipo físico antes de que se autentique utilizando la SIM o la USIM, este equipo se llama EIR (*equipment identity register*).

Como se puede ver, con un simple mensaje se puede obtener información de un suscriptor. Dentro de la información que podría ser obtenida por este método se tiene la información de la ubicación, de los servicios que utiliza y hasta la manipulación de algunos servicios.

3.1.2. Vulnerabilidad sobre la red inalámbrica

Por diseño, todas las redes celulares se basan en el estándar 3GPP (GSM, GPRS, EDGE, UMTS y LTE). En un inicio la red GSM fue diseñada a propósito con niveles bajos de seguridad debido a que varios gobiernos europeos solicitaron la posibilidad de desactivar o atravesar la encriptación del canal de radio para interceptar y escuchar una llamada desde un teléfono móvil.

Desde la introducción de la red 2G fuera de Europa, los operadores se vieron obligados a utilizar una encriptación más débil que la utilizada en el viejo continente. Así nacieron dos tipos de encriptación ya que algunos estándares

de la red 2G se consideraron confidenciales y solo fueron compartidos bajo acuerdos para no revelar esta información.

Las terminales móviles 2G fueron diseñadas para manejar ambos tipos de encriptación, de esta manera podrían ser utilizadas en cualquier parte del mundo. Una desafortunada consecuencia de este diseño es la posibilidad de poner en servicio estaciones base falsas con las cuales un atacante puede engañar a una terminal para utilizar encriptación débil o no utilizarla en lo absoluto, incluso en países donde se utiliza una encriptación más fuerte.

Esto es posible debido a que la autenticación con la red no fue incluida en el estándar 2G, por lo cual las terminales móviles no saben si están conectadas a una red legal o a una red falsa puesta en servicio por un atacante. Esta vulnerabilidad permite que la encriptación de un usuario sea sobrepasada para obtener su clave de autenticación y posteriormente utilizada para poder realizar la escucha de una llamada.

La red 3G, subsecuentemente desarrollada a la tecnología 2G, igualmente estandarizada por 3GPP, intenta tomar en cuenta los problemas de seguridad de la red GSM; como resultado se logró diseñar una red con una seguridad relativamente más fuerte en la cual se incluyen algoritmos de encriptación más fuertes y autenticación con la red. Sin embargo, se queda corta en permitir la autenticación de la red móvil por parte de un usuario. La red LTE de igual forma aumenta la seguridad en la autenticación y encriptación de la información.

En las redes móviles actuales las tecnologías 2G, 3G y 4G se usan combinadamente en todo el mundo; los teléfonos móviles que se encuentran en el mercado están diseñados para poder conectarse a cualquiera de estas redes, con el objetivo de maximizar la cobertura en lugares donde alguna de estas

tecnologías no está disponible. Es obvio entonces que el punto más débil en la cadena de seguridad es la tecnología 2G.

Aunque un teléfono sea capaz de comunicarse de manera segura a través de las redes 3G y 4G, es posible obligarlo a no hacerlo. En este sentido un teléfono inteligente tiene la misma vulnerabilidad que cualquier otra terminal.

La vulnerabilidad de la red 2G y la posibilidad de obtener la llave de autenticación o Ki y la IMSI de un suscriptor abre un mercado para los denominados cazadores de IMSIs. Un cazador o atacante debe prevenir que un teléfono se conecte a una estación legítima 3G o 4G, para lo cual deben utilizar bloqueadores de señal, lo cual se denomina *jamming,* para los espectros de 3G y 4G. Para una terminal estas redes no están disponibles, entonces se conectarán a la estación base falsa, establecida por los cazadores.

Por medio de este proceso, se pueden conseguir varios números de IMSIs y claves, las cuales se pueden grabar en tarjetas SIM en blanco para luego ser vendida en el mercado negro. Normalmente este tipo de tarjetas únicamente se venden a organizaciones criminales, pero en algunos casos una persona individual puede adquirir una.

Por lo que mencionado anteriormente, la política en el estándar de 2G de poder sobrepasar la encriptación, tiene consecuencias, que hoy en día, las organizaciones criminales pueden utilizar.

Dentro de las estrategias que se pueden utilizar para contrarrestar estas prácticas criminales están, por ejemplo, la utilización de terminales que pueden detectar cuando se está trasmitiendo sin encriptación; también, la instalación de sensores que permitan escanear estaciones base falsas o también, se puede

incluir esta tecnología en las diferentes estaciones bases para detectar las estaciones falsas.

3.2. Interrupción del servicio de telefonía IP

Las amenazas de esta categoría tienen como objetivo afectar el ancho de banda de una red, la capacidad de un servicio o su calidad. Las amenazas de este tipo se denominan de denegación de servicio o DoS (*denial of service*) ataques cuyo fin es el de denegar o degradar el acceso de un usuario legitimo a un servicio o a un recurso de red, o el de derribar los servidores que ofrecen el servicio.

Los servicios VoIP son más susceptibles a los ataques DoS que otros servicios que corren sobre Internet debido a la transmisión en tiempo real de los mensajes.

Los ataques DoS pueden ser ejecutados en una forma distribuida y dirigida a diferentes entidades como servidores o usuarios finales, dependiendo de la intensión del atacante. En este enunciado se intenta describir los ataques DoS más conocidos, los cuales se basan en servicios VoIP que utilizan el protocolo SIP.

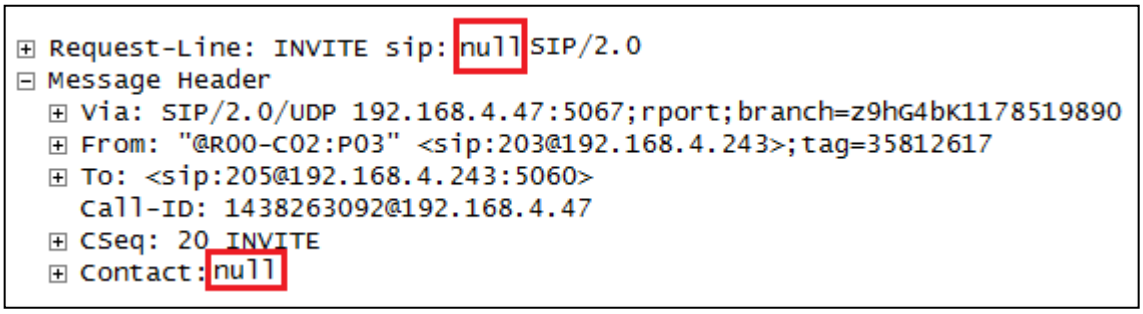
3.2.1. Mensajes SIP no válidos

Se consideran mensajes no válidos a todos aquellos que no provienen de una sesión correcta o legítima. Estos mensajes pueden ser generados cuando las aplicaciones o implementaciones del protocolo SIP no cumplen con los estándares o contienen errores en el código de desarrollo.

Los atacantes pueden manipular a placer mensajes SIP para tomar ventaja de los problemas de seguridad existentes para explotar los puntos débiles del protocolo SIP o de la red objetivo.

Debido al formato de texto plano que tienen los mensajes del protocolo SIP, los mensajes no válidos son fáciles de ser generados, de hecho existen infinidad de formas de manipular los mensajes SIP. Estos mensajes pueden ser simplemente mal formados en el sentido de que no cumplen con la gramática definida en el protocolo.

Figura 16. **Mensaje SIP con error de sintaxis**



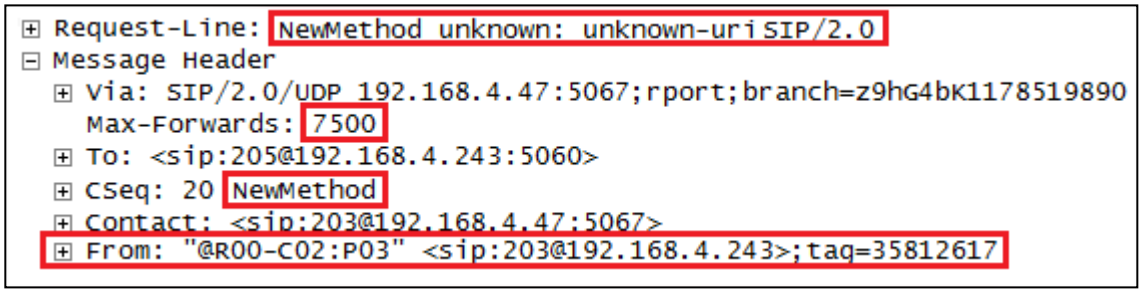
Fuente: elaboración propia, utilizando programa Wireshark.

La figura 16 muestra un ejemplo de un mensaje malformado donde la dirección del receptor no aparece. Normalmente el servidor que recibe este mensaje debería contestar con el mensaje 400 (*bad request*) indicando que el mensaje fue descartado, sin embargo, bajo ciertas circunstancias el servidor puede fallar en procesar mensajes consiguientes ya que puede quedar inhibido.

Los mensajes SIP, que pueden ser sintácticamente bien formados, pero pueden tener errores de semántica, pueden no carecer de significado, sin

identificados, ambiguos o llevar a un punto muerto al servidor. La figura 17 muestra un ejemplo de este tipo de mensajes.

Figura 17. **Mensaje SIP con error semántico**



Fuente: elaboración propia, utilizando programa Wireshark.

El mensaje de la figura 17 tiene un método desconocido de petición, un esquema URI desconocido, un campo mandatorio faltante (*call-id*), campos escalares demasiado grandes (*max-forwards*) y campos que deben ser establecidos solo una vez y en un orden específico (*from*).

El mensaje es sintácticamente válido pero un servidor al recibir este mensaje puede fallar en procesarlo y puede consumir mucho tiempo tratando de analizar el mismo para determinar el tipo de mensaje de petición y para adquirir la información que es necesaria para encontrar la ruta para el enrutamiento del mensaje.

Adicionalmente, los atacantes pueden manipular un mensaje SIP ingresando información sin ningún sentido o información errónea en varios campos de un mensaje para tomar ventaja de los problemas de seguridad en un sistema o simplemente para realizar diferentes tipos de ataques.

3.2.2. Ataques de inundación

Otro tipo común de ataque DoS en los escenarios VoIP sobre el protocolo SIP son los mensajes de inundación, donde algún usuario envía una gran cantidad de mensajes de petición INVITE o REGISTER hacia un servidor SIP o a una entidad SIP, en consecuencia, el elemento de red objetivo entra en un estado de ocupado al tratar de procesar todas las llamadas de usuarios maliciosos y es incapaz de procesar peticiones legítimas.

Con el envío masivo de mensajes, un servidor SIP puede ignorar o procesar los mensajes de una manera tan lenta que el servicio VoIP es prácticamente imposible de utilizar.

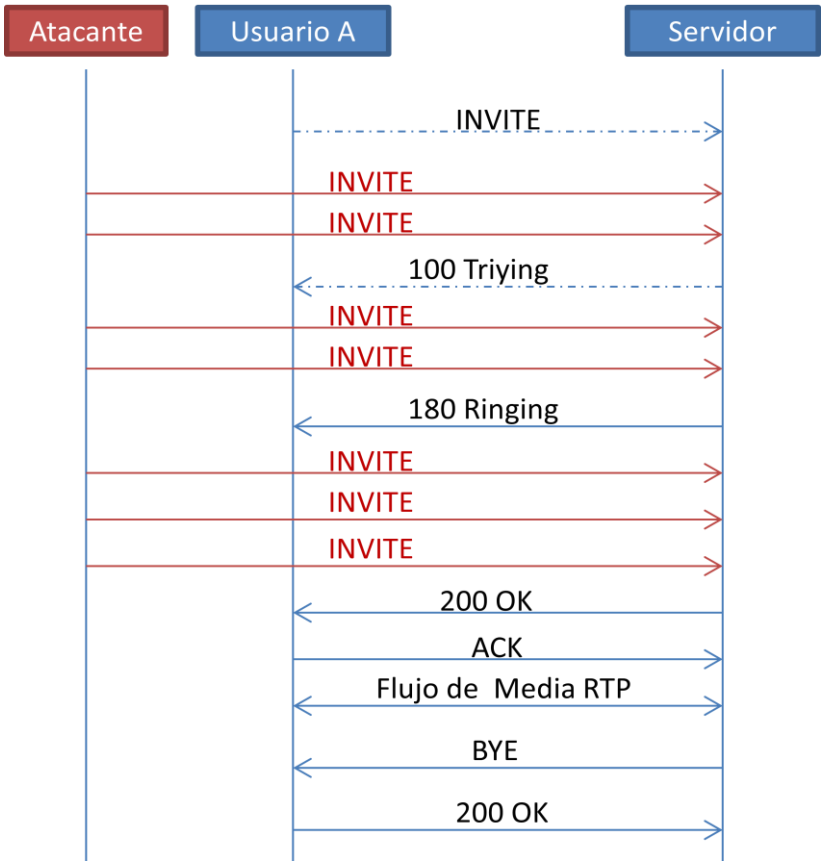
Los ataques por inundación, también, pueden ocurrir con la existencia de mecanismos de autenticación. En este escenario, luego de recibir mensajes de petición, el servidor o registrador responderá con un mensaje de solicitud de registro y esperará por el mensaje de respuesta por parte de un usuario malicioso. Los atacantes pueden enviar mensajes de petición para luego detener el proceso de solicitud y así mantener en espera varias sesiones en un servidor SIP.

Comúnmente, dos tipos de mensajes de inundación son los más vistos cuando se realizan este tipo de ataques: los mensajes de inundación por mensajes INVITE y los mensajes por inundación por mensajes REGISTER.

En el caso de los mensajes de inundación por mensajes INVITE, una gran cantidad de mensajes INVITE son enviados a un servidor SIP por parte de un atacante. Con este tipo de ataque, el servidor SIP se mantiene tan ocupado tratando de procesar los mensajes SIP falsos que no es capaz de procesar los

mensajes SIP legítimos. La figura 18 muestra el flujo de señalización de este tipo de ataque.

Figura 18. **Mensajes de inundación INVITE**



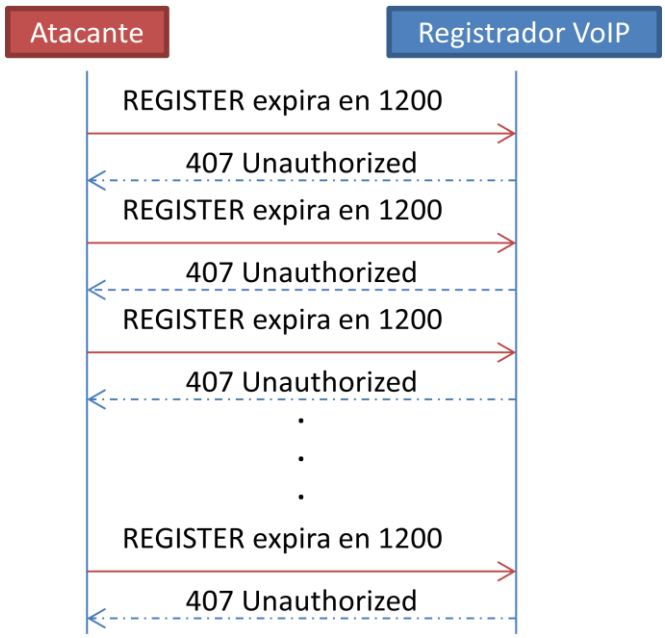
Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

Uno de los elementos principales en los sistemas VoIP basados en SIP es el registrador, que es un servidor que acepta las solicitudes de registro de los usuarios. Un mensaje SIP REGISTER solicita agregar un registro nuevo de una dirección de usuario SIP IP y una o más direcciones de contacto; normalmente estas direcciones son IP.

Cuando un atacante quiere paralizar el servicio de un servidor de registro, normalmente enviando una cantidad enorme de solicitudes de registro, lo hace por medio de un ataque DoS.

En el caso de los ataques de inundación por mensajes SIP REGISTER, el objetivo del atacante es el de degradar y retardar el proceso de registro en un servidor VoIP; en la figura 19 se puede ver el flujo de un ataque de este tipo.

Figura 19. **Mensajes de inundación REGISTER**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

El primer mensaje REGISTER que es enviado no lleva credenciales de autenticación, así que el primer mensaje 407 es para solicitar las mismas; los consiguientes mensajes falsos pueden o no llevar credenciales falsas o de tener acceso, pueden ser credenciales reales.

3.2.3. Secuestro de llamadas

El secuestro de llamadas ocurre cuando un atacante utiliza mensajes SIP para secuestrar una llamada existente hacia otro destino final, en este caso un servidor atacante, para este tipo de ataque es necesario que el atacante logre replicar.

El secuestro de llamadas es un ataque en la capa de aplicación que toma ventaja de algunos campos en los mensajes SIP. El diseño del protocolo SIP incluye mensajes de re direccionamiento o REDIRECT; esta funcionalidad le permite a un usuario SIP moverse de un ambiente a otro, por ejemplo, de una red inalámbrica de acceso a Internet residencial o una red móvil, mientras la llamada está en curso.

En un ataque de secuestro de llamadas el atacante envía un mensaje REINVITE con el cual puede realizar un nuevo enrutamiento del flujo de los paquetes RTP hacia la dirección IP del atacante. Con esto el atacante puede escuchar parte de la llamada.

3.2.4. Liberación de llamadas

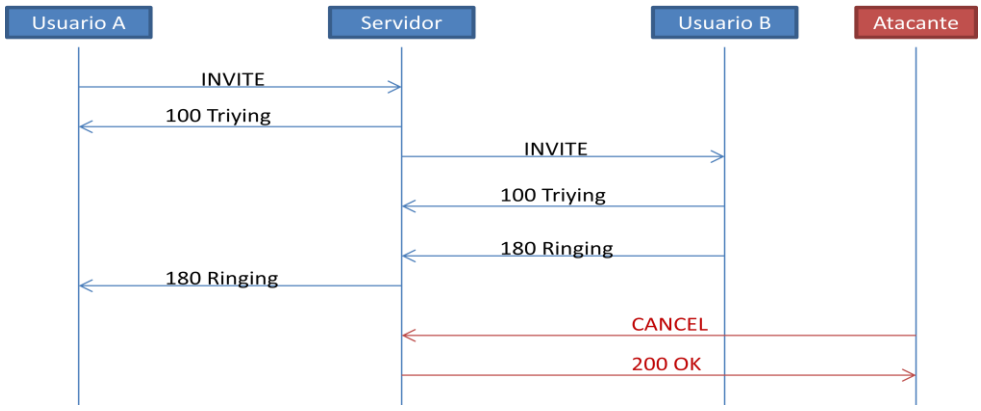
Los ataques para la liberación o terminación de llamadas ocurren cuando un atacante utiliza mensajes SIP para poder terminar una llamada existente. Esto se logra hacer replicando el encabezado SIP de la llamada en curso.

Existen dos tipos de ataque para la denegación de un servicio que corre sobre VoIP utilizando el protocolo SIP: utilizando mensajes SIP CANCEL y utilizando mensajes SIP BYE.

Un ataque realizado con mensajes CANCEL se refiere a una liberación no autorizada de una llamada que está en proceso de ser establecida entre dos entidades SIP. Este método de ataque se utiliza para cancelar procesamiento de búsquedas pendientes o intento de llamadas. Específicamente, le solicita al usuario SIP, que inicia o recibe una llamada, cesar la misma y genera una respuesta de error a la solicitud.

El mensaje de terminación de la llamada, también, puede ser dirigido al servidor VoIP. En la figura 20 se puede ver el flujo de este tipo de ataques.

Figura 20. **Ataque DoS con mensajes CANCEL**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

Para iniciar una nueva llamada, el usuario A envía un mensaje INVITE al usuario B; antes de que el usuario B reciba el mensaje, el servidor envía un mensaje 100 indicando que la llamada está siendo procesada; cuando el teléfono del usuario B empieza a anunciar la llamada, un mensaje 180 es enviado al teléfono del usuario A para indicarle que la llamada está en proceso de ser establecida.

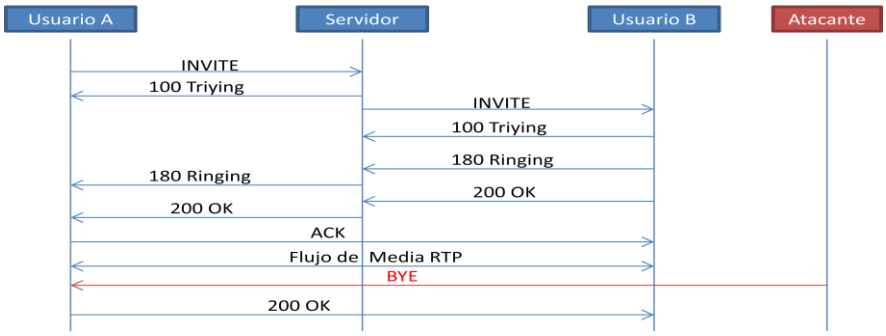
En este momento, un atacante puede fácilmente enviar un mensaje CANCEL falso al servidor VoIP. Sin la propia autenticación, un servidor no puede diferenciar entre un mensaje CANCEL falso de uno genuino y considera que el mensaje puede ser enviado ya sea por el usuario A o el usuario B.

Luego de que se recibe un mensaje CANCEL por parte del servidor, el proceso de establecer la llamada es terminado.

El atacante deber ser capaz de interceptar tráfico en la red de para poder identificar los mensajes SIP, este tipo de ataques serán tratados en otro enunciado. Los mensajes CANCEL también pueden ser enviados luego de que se recibe el mensaje INVITE.

Una sesión de media establecida entre dos usuarios es terminada cuando se recibe un mensaje BYE en cualquiera de los sentidos. Es un mensaje entre usuarios finales que es enviado por uno de los participantes. Un ataque por envío de estos mensajes se refiere a la terminación no autorizada de una llamada ya establecida.

Figura 21. **Ataque DoS con mensajes BYE**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

Como se puede notar en la figura 21, un flujo de media RTP se encuentra establecido, al igual que en el ataque DoS por mensajes CANCEL, el atacante debe ser capaz de identificar el flujo y contenido de los mensajes SIP para poder enviar un mensaje BYE falso válido y así poder dar por finalizada la sesión.

3.3. Abuso del servicio de telefonía IP

Este tipo de amenaza ocurre cuando se hace un uso inapropiado de los servicios VoIP, especialmente cuando el servicio se ofrece comercialmente. Entre estas amenazas se pueden encontrar fraudes o el uso de la telefonía VoIP sin intención de realizar un pago.

El sistema de telefonía de una compañía puede ser sujeto de fraude telefónico o *toll fraud.* Este fraude puede ser llevado a cabo por empleados de la compañía o por personas externas que se dedican a encontrar vulnerabilidades en el sistema. En el caso de los empleados simplemente ignoran las políticas de la empresa esperando que sus actividades no sean detectadas porque es difícil diferenciar entre una llamada por trabajo y una llamada privada únicamente basándose en el número marcado.

En el caso de las personas externas, tratan de encontrar vulnerabilidades en los dispositivos de red, incluyendo los sistemas de telefonía VoIP. Algunas veces, estas personas no buscan necesariamente sistemas de telefonía, buscan únicamente cualquier sistema del cual puedan tomar control.

La diferencia entre estos dos grupos es la forma en la cual se mitigan los ataques. En el caso de los atacantes externos, la idea es prevenir el acceso no autorizado a los sistemas y a sus elementos de red. Para los usuarios

autorizados, el administrador de la red VoIP tiene que ser cuidadoso al limitar las habilidades técnicas y características del sistema sin comprometer la flexibilidad y eficiencia de los usuarios.

3.3.1. Atacantes internos

En el caso de los fraudes internos, hay algunos servicios de la telefonía que pueden ser utilizados de manera incorrecta. Entre estos servicios están el desvío y transferencia de llamadas y el servicio de buzón de voz. Si estos servicios están bien protegidos, los atacantes podrían tratar de utilizar otros servicios. Por ejemplo, si un usuario no tiene permitido realizar transferencia de llamadas a números externos, el usuario podría tratar de establecer una conferencia entre dos usuarios y luego dejar la conferencia.

Un administrador de un sistema VoIP debe aceptar el hecho que el fraude telefónico no puede ser eliminado completamente. La única forma de alcanzar completamente la eliminación de esta amenaza seria con el bloqueo de todas las llamadas externas y también deshabilitar todos los servicios que permitirían a cualquier persona realizar llamadas fuera de la red local VoIP.

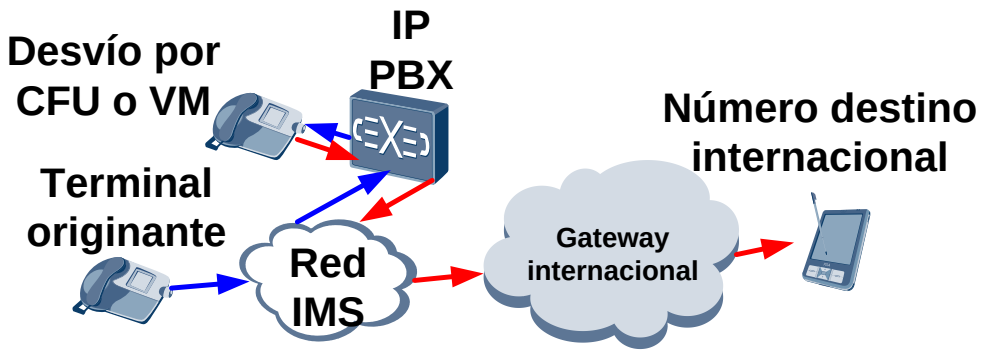
Esta técnica puede ser factible para redes locales y teléfonos con funciones básicas, por ejemplo, los que están colocados en la recepción de una compañía, pero no es deseable para teléfonos utilizados por empleados estándar.

En este caso, el administrador tiene que establecer permisos a cada terminal dependiendo de las necesidades de cada empleado, es decir, un gerente de una compañía transnacional podría no tener restricción alguna

debido a que necesita realizar llamadas internacionales o fuera de la red local. Sin embargo, un empleado promedio no necesita realizar este tipo de llamadas.

Otro tipo de restricción que puede realizar el administrador es el de establecer una sola terminal con la posibilidad de realizar todas las llamadas salientes. Para poder conectar las llamadas fuera de la red VoIP debe existir una operadora que sea la responsable de decidir cuáles llamadas pueden ser establecidas y cuáles no.

Figura 22. **Fraudes internos por desvío de llamada**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

La figura 22 describe dos casos posibles de fraude: mediante el uso del servicio desvío incondicional o CFU (*call forward unconditional*) o del servicio de buzón de voz o VM (*voice mail*). En el caso de que un empleado este de vacaciones fuera del país y adquiera un número móvil local para evitar pagar *roaming* por llamadas entrantes, si usase su teléfono personal.

El fraude se da cuando se realiza un desvío de llamadas desde la terminal VoIP hacia un destino internacional La persona puede solicitar a todos sus contactos que lo llamen al número de su trabajo, para que estos no incurran en

el gasto de una llamada internacional, sino que la red de empresa que absorba ese gasto.

3.3.2. Atacantes externos

El problema de los atacantes externos se define como un problema de autenticación: un hacker falsifica la identificación de un usuario y realiza llamadas desde el sistema VoIP del mismo para obtener ganancias financieras, el cual es su mayor incentivo para explorar las vulnerabilidades de los sistemas IP-PBX: ganar acceso y realizar llamadas internacionales.

Este tipo de ataques suceden en las organizaciones donde se instalan sistemas VoIP que carecen de seguridad. Debido a que el objetivo final de este tipo de ataques es el de realizar el enrutamiento de grandes cantidades de tráfico hacia redes fuera de la telefonía sobre IP, se puede considerar un ataque a redes adyacentes, por lo cual este tema será desarrollado ampliamente más adelante ya que este tipo de ataques son los que motivaron este trabajo de investigación.

3.4. Intercepción y modificación de servicios

Las amenazas por intercepción o modificación son aquellas por medio de las cuales, de manera ilegal y sin autorización, una persona puede escuchar o modificar la señalización o el contenido de una sesión VOIP.

Un hombre en el medio o MITM (*man in the middle*) se requiere para realizar la intercepción o modificación de los mensajes SIP. Los ataques MITM son la mayor amenaza para la seguridad y confianza de los protocolos y sistemas VoIP.

El MITM que está en la trayectoria de señalización o de media de un servicio de telefonía sobre IP puede fácilmente escuchar, desviar o secuestrar llamadas VoIP seleccionadas. Todos estos ataques MITM requieren que el atacante se encuentre en el camino de una ruta VoIP, por lo que se cree comúnmente que es poco viable que el atacante lo haga remotamente, debido a que no se encuentra en la trayectoria VoIP. Esto hace a la gente creer que asegurando todos los nodos a lo largo de del camino del tráfico VoIP es suficiente para prevenir los ataques MITM a los servicios de telefonía sobre IP.

El objetivo de este enunciado es demostrar que un atacante remoto que no se encuentra en la trayectoria del tráfico VoIP puede de hecho realizar todo tipo de ataques MITM aprovechando vulnerabilidades de implementación del servicio VoIP. Otro objetivo es definir que otros tipos de ataques puede realizar un MITM.

Uno de los principales problemas del protocolo SIP es la autenticación, no directamente, sino más bien que solo algunos campos en los mensajes que se envían entre una terminal SIP y un servidor son protegidos, dejando al resto vulnerables. Tomando ventaja de la vulnerabilidad del protocolo SIP y RTP, un MITM puede colectar información sensible como números de cuentas o contraseñas. También puede direccionar cualquier llamada VoIP hacia otro servidor SIP y manipular los servicios de desvío de llamadas y pretender ser una institución financiera o un representante de cuenta de un banco.

Otro tipo de fraude que puede realizar un MITM es el de lanzar ataques de facturación hacia un usuario VoIP en específico, con el fin de sobregirar las cuentas de estos por medio de cobros de llamadas que ellos no realizaron. Otro tipo de ataque, descrito anteriormente, es el de la denegación de servicio por medio del envío de mensajes BYE o BUSY.

Para descubrir cuál es la posibilidad de que una persona remotamente se convierta en un MITM, basta con sumir el rol del atacante que busca engañar a un teléfono VoIP con el objetivo de enviar todo su tráfico VoIP a través de este explotando las vulnerabilidades de la terminal SIP y los protocolos que utiliza.

Existen muchos métodos y herramientas que pueden permitir a un atacante remoto convertirse en un MITM. La seguridad de la información, por ejemplo de las conexiones y los elementos de red, debe ser tomada en cuenta muy drásticamente. Esta información debe ser únicamente transmitida por medios encriptados o seguros ya que una de las formas más comunes de que un atacante remoto se convierta en un MITM es porque pudo adquirir la información directamente de un facilitador, ya sea que la haya comprado o sea un descuido del mismo.

Sin embargo, esta no es la única forma, un atacante remoto puede olfatear (*sniff*) una red con el propósito de poder capturar el tráfico que pasa por un medio cableado o inalámbrico. Para que un atacante pueda ver y capturar otro tráfico que no sea el que está dirigido a su terminal, por ejemplo, una PC conectada a Internet o una red inalámbrica, se debe poner la misma en un modo promiscuo o de monitoreo. De este modo se puede capturar todo el tráfico, no solo es que está dirigido a las IP o MAC *address*.

En el caso de redes inalámbricas, esto puede ser realizado fácilmente con herramientas que se puede descargar en la red. En el caso de que sea un elemento de red de capa dos, un *switch*, puede ser un poco más complicado. Un *switch* está diseñado para reducir el tráfico de una red y la congestión, aislando el tráfico y solo enviando paquetes hacia una dirección o IP o MAC, la cual es su destino.

Para poder determinar a donde un *switch* envía el tráfico, este mantiene una tabla que esencialmente tiene un mapa de direcciones IP y MAC (puertos físicos). Es decir, por ejemplo, que cuando el tráfico tiene como destino la dirección IP1 se utilizará la dirección MAC1 para alcanzarlo. Para poder engañar a un *switch,* para que no cumpla con su tarea es necesario cambiar la información de esta tabla y así poder obtener el tráfico de alguien más, esta técnica se denomina ARP *spoofing*.

La tabla donde se almacena el mapeo entre direcciones IP y MAC en un *switch* se llama CAM (*content addressable memory*) y esta se actualiza por medio de mensajes ARP que son enviados fuera del *switch* para poder obtener la información de los elementos de red conectados a este.

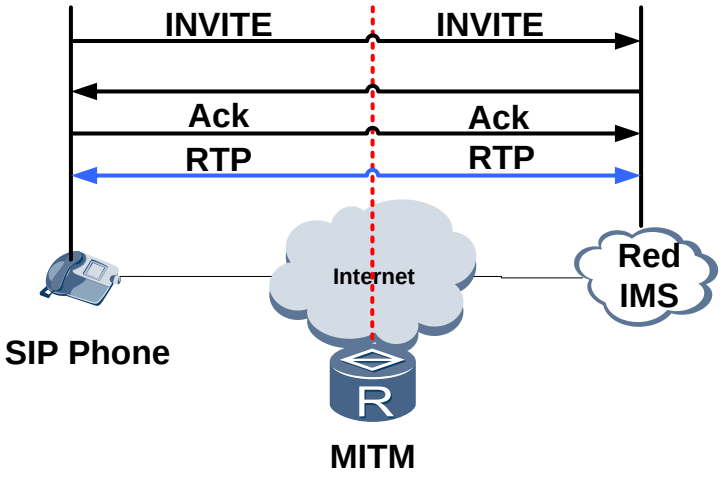
La idea de utilizar esta técnica es la de convertirse en un MITM, colocándose entre dos elementos de red, haciendo creer que el tráfico destino es nuestra MAC o IP, para poder olfatear el mismo. Es necesario también que el tráfico sea enviado a su destino final para que el ataque sea transparente para los elementos de red que se encuentran en el medio.

Existen también herramientas que ayudan a ver y a identificar el tráfico capturado, obtenido ya sea por medio de una red inalámbrica o el acceso a un *switch*. Con estas herramientas se puede realizar un filtrado de paquetes por protocolos, con lo cual se puede identificar fácilmente una sesión VoIP.

Para convertirse en un MITM de un servicio VoIP se requiere, aparte de de mucha habilidad, también paciencia, ya que muchas veces el tráfico capturado es puramente de sesiones de acceso a sitios web, en el caso de que el tráfico obtenido provenga de un equipo inalámbrico el cual provee servicio de Internet a cualquier persona que se conecte al mismo.

Luego de que un atacante logre identificar una sesión VoIP, este puede usar varios métodos para reiniciar la terminal y así capturar toda la mensajería de registro; por ejemplo, podría enviar paquetes malformados al cliente SIP para intentar hacer que la terminal se reinicie. Esto con el objetivo de que la terminal se registra en el servidor del atacante y así poder interceptar o modificar la mensajería SIP entre el cliente y el servidor SIP.

Figura 23. **Topología de un ataque MITM**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

3.4.1. SPIT

SPIT (*SPAM over internet telephony*) es un tipo de ataque de ataca MITM, en el cual se transmiten mensajes de vos no solicitados, con el objetivo de realizar telemercadeo o fraudes. La mayor parte de estos ataques son realizados por servidores SIP fraudulentos programados. Estos servidores pueden iniciar un número grande de llamadas en paralelo.

Si la llamada conecta, entonces, este servidor genera un mensaje ACK y procede a reproducir un audio pregrabado para luego termina la llamada. Este tipo de ataques se pueden construir completamente con software sobre computadoras de bajo perfil; además, no se requiere ser un gran experto para poder realizar este tipo de ataque.

3.4.2. VOIP *vishing*

El termino *vishing* se refiere a *phishing* sobre el dominio VOIP. En general *phishing* hace referencia al acto de hacerse pasar por una entidad confiable para poder adquirir información confidencial de las víctimas.

En el contexto de VoIP, un atacante MITM se presenta con credenciales falsas o engañosas obtenidas, por ejemplo, con el secuestro de llamadas previas, tratando de suplantar a un usuario o a una entidad válidas por medio de una llamada con el fin de obtener acceso a información personal o financiera.

Esto se puede hacer por medio del redireccionamiento de una llamada, al detectar que el usuario VoIP está intentando marcar un número conocido de una entidad bancaria, por ejemplo. El atacante puede conectar esta llamada hacia un IVR (*interactive voice responder*) o una contestadora automática para que la llamada sea potencialmente creíble.

3.4.3. Ataques de facturación

Los sistemas VoIP comerciales tienen o no limitaciones en cuanto al número de minutos de llamadas. Las tarifas de llamadas depende en el área

geográfica en la cual termina la llamada y la mayoría de los planes incluyen llamadas ilimitadas en ciertas a ciertos usuarios.

Teniendo en cuenta esto, un atacante remoto puede prolongar las llamadas o crear sesiones no autorizadas, provocando que el usuario VoIP tenga que pagar más por el servicio. Este tipo de ataque es de tipo MITM.

Este tipo de ataque puede ser perpetrado de varias formas, una es por medio de la repetición o reproducción de mensajes INVITE hacia diferentes destinos. Esta vulnerabilidad es un error en la implementación SIP de la función antirepetición. Este ataque no puede ser detenido aun con los mensajes de autenticación de un INVITE. Cuando un mensaje INVITE es interceptado, el atacante puede reenviarlo días después, modificando únicamente el campo SDP para poder recibir flujo de media RTP y así poder conectar la llamada y provocar un cobro no deseado a un usuario VoIP.

Otro tipo ataque de este tipo, es por medio del envío de mensajes de ocupado o también retrasando el mensaje de terminación de llamada; en ambos casos lo que se busca es extender la duración de las llamadas, por las cuales un suscriptor debe pagar.

3.5. Ataques a redes adyacentes

En los principios de la telecomunicación inalámbrica y la PSTN las redes eran cerradas ya que los mensajes de señalización se intercambiaban entre puntos privados aislados sobre las redes SS7. Estas redes no estaban conectadas a ninguna red pública y, por ende, los hackers no podían ganar acceso al mundo de las redes inalámbricas.

Los ataques que son posibles efectuar en estas redes aisladas dependen de la obtención de equipo de telecomunicación, el cual en la mayoría de casos no es muy fácil de adquirir, por cualquier persona u organización criminal, debido a su precio elevado.

Con la integración de las redes móviles, fijas y el Internet se abre un mundo nuevo de vulnerabilidades, a la vez que los atacantes pueden ganar fácil acceso a estas redes adyacentes a través de Internet. El Internet es una red abierta y accesible a cualquiera con equipo muy simple. También es muy fácil para los hackers tener acceso a los servidores en Internet debido a las muchas vulnerabilidades.

Si se obtiene acceso a un servidor sobre Internet que provee servicios a las redes adyacentes, se abre entonces la posibilidad para un atacante de tener acceso a los servidores de las móviles o fijas.

Si un atacante obtiene acceso a un servidor de la red móvil o PSTN este podría insertar, modificar o destruir data o recursos para que suscriptores maliciosos puedan cometer fraudes, colectar información confidencial o modificar la lógica del servicio para interrumpir la operación de la red o denegar el servicio, sin embargo, es muy difícil ganar acceso a una identidad de la red móvil o PSTN, a la vez se requiere alto grado de conocimiento en la lógica de configuración para poder realizar una operación como esta.

Considerando lo anterior, entonces para que un atacante pueda modificar los servicios o interrumpirlos, desde un nodo de la red de telefonía móvil o fija, es muy probable que este tenga familiaridad con la red y, por ende, acceso a la misma. Este problema se vuelve un asunto de seguridad y de control interno de un operador de servicios de telefonía.

Este enunciado trata de explicar entonces cuáles son los ataques o fraudes, en las redes adyacentes, como la red móvil y la red fija, desde las redes de servicios sobre IP; en específico aquellos que representan costos millonarios para los operadores. Este tipo de ataques no se basan en ganar acceso a los servidores de las redes adyacentes para alterar o denegar el servicio, sino de poder enviar tráfico de voz o datos, de una manera ilegal, utilizando como acceso a las redes VoIP para obtener beneficios monetarios.

3.5.1. Fraude bypass

El fraude por *bypass* ha plagado el rendimiento de los proveedores de servicio de telefonía a un grado considerable. Este tipo de fraude es el más perjudicial y costoso; también, se le conoce como enrutamiento gris u operadores de cajas SIM. Con este fraude, se explota los *gateways* entre las redes sobre VoIP.

Estas cajas SIM, *SIM boxes*, se utilizan para secuestrar llamadas entrantes internacionales, cuando estas son transmitidas utilizando el protocolo VoIP hacia una red móvil.

El objetivo final es inyectar de regreso la llamada hacia la red móvil para poder terminada como si fuera local; de esta forma, los operadores de telefonía pierden el cobro de las llamadas internacionales.

En algunos países o regiones en particular, el tipo de fraude por operador SIM *box* ocurre cuando las tarifas de terminación de llamadas internacionales son más altas de lo que son las tarifas de llamadas locales. El proveedor fraudulento de servicios asegura su ganancia proponiendo tarifas bajas para llamadas internacionales a otros operadores.

Este tipo de atacantes toman ventaja de las tasas de cobro por llamadas locales, ya que hacen que una llamada que debe ser terminada como internacional se cobre como una llamada nacional; como resultado las llamadas son producidas, en los usuarios finales, con una manipulación del número CLI (*calling line identity*) ya sea cambiándolo o borrándolo. Existen varios tipos de fraudes por *bypass* y se describirán en los siguientes enunciados.

3.5.1.1. Supervisión de contestación falsa

La supervisión de contestación falsa es un tipo de servicio ilegal que permite a los defraudadores o atacantes ganar una gran cantidad de dinero generando llamadas falsas que son terminadas en redes de telefonía móviles reales.

El servicio genera los mensajes de servicio de transporte necesarios y crea el tiempo de transmisión requerido para poder cobrar o facturar ese tráfico. Esto se hace por medio de canalizar llamadas a números que están fuera de servicio o fuera o también hacia teléfonos prepago que son adquiridos con el fin de terminar las llamadas.

Cuando se hace una llamada hacia un suscriptor que esta fuera de servicio, se reproduce un audio indicando que el número con que se desea comunica está disponible. Normalmente la llamada no es conectada entre el número de A y B, así que la llamada no debería ser cobrada. Si este es el caso, los defraudadores se encargan de supervisar estas llamadas para que puedan aparecer como llamadas completadas y así puedan ser cobradas.

Cuando un suscriptor posee el servicio de buzón de voz, las llamadas fuera de servicio son dirigidas hacia este, por lo cual son conectadas y se

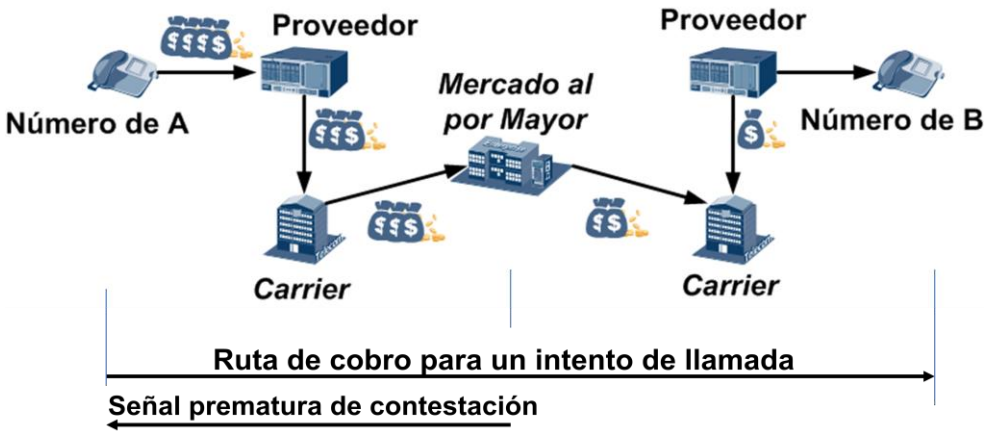
genera un cobro al suscriptor que realizó la llamada. En este caso los defraudadores no necesitan supervisar la contestación de las llamadas.

Los defraudadores obtienen dinero realizando tratos con los *carriers* de tráfico internacional ya que estos conectan el mundo de las redes VoIP con las redes móviles y fijas privadas. No todo el tiempo son tratos donde ambas partes se ponen de acuerdo para realizar fraudes, la mayoría de las veces los defraudadores ofrecen tarifas muy bajas para el transporte de llamadas y en la mayoría de los casos los *carriers* no validan la veracidad de todas las llamadas.

Para poder conectar una llamada de larga distancia internacional, un operador móvil o proveedor de servicios al por menor, envía las llamadas hacia un *carrier* internacional; en algunos casos estos *carriers* son proveedores de enrutamiento de bajo costo; estos a su vez envían el tráfico a los proveedores de transporte de tráfico por mayor; estos últimos se encargan de acumular y transportar tráfico desde otros proveedores de servicios al por menor o proveedores de enrutamiento de bajo costo.

Este fraude entonces toma a las personas que tienen un servicio legítimo de transporte de llamadas al por mayor y los convierte en defraudadores.

Figura 24. **Supervisión de contestación temprana**



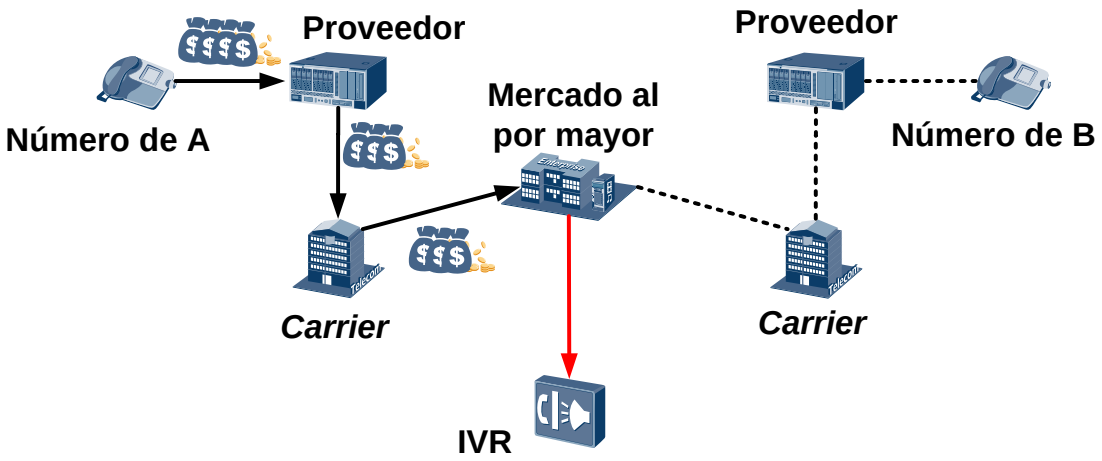
Fuente: elaboración propia, utilizando programa Microsoft Visio.

Una de las formas de ganar dinero utilizando la supervisión de contestación falsa o FAS, *false answer supervisión*, es la posibilidad de controlar la señal de conexión de una llamada. En la figura 24 se muestra este escenario, a la vez se puede apreciar en porcentaje los cobros que se disparan en cada punto de una llamada de larga distancia internacional. Entonces, si se lanza una señal de contestación temprana, desde los proveedores de transporte por mayoreo, antes de que la llamada llegue a su destino final, la misma puede ser cobrada antes a todas las entidades.

Para un usuario que realiza la llamada pueda que no signifique mucho dinero, pero para el volumen de tráfico de minutos de llamadas que cobran los *carriers* y los proveedores de servicio por mayoreo pueden llegar a representar millones de dólares mensuales. Hay variaciones del fraude FAS, por ejemplo, en vez de manipular la contestación se manipule la terminación de la llamada para poder retardarla y así poder cobrar ese tiempo extra.

Otro tipo de fraude FAS se puede dar cuando el proveedor de servicios al por mayor realiza el enrutamiento una llamada a un IVR con el fin de conectar la llamada antes de que la llamada llegue hasta el receptor final; en algunos de los casos, se reproducen tonos reales para hacer pensar al usuario que realiza la llamada que se está procesando el servicio; en otros casos, se pueden reproducir mensajes artificiales de voz para simula la contestación de una llamada. En otras ocasiones ni siquiera se conecta la llamada al usuario final, sino que solo al IVR.

Figura 25. **Fraude FAS por divergencia de llamada**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

El aumento en la conexión o terminación de una o varias llamadas no es la única forma del fraude FAS; existe otro tipo muy particular cuyo objetivo es realizar un volumen de llamadas falsas las cuales son terminadas de manera correcta en la red móvil de algún operador. El objetivo es aumentar la tasa de contestación de llamadas (*answer rate*) o ASR. En algunos casos pueden utilizarse las llamadas FAS por conexión o por terminación para aumentar el ASR, teniendo en cuenta que todas las llamadas FAS son forzadas a ser conectadas.

El ASR es una medida que utilizan los proveedores de servicio al por mayor para cobrar el tráfico enviado por los *carriers*; para el cálculo se utilizan el número de minutos promedio cursados por una ruta y el porcentaje del ASR multiplicado por el número de llamadas totales. Así que si en una ruta aumenta el número de llamadas FAS, también aumenta el porcentaje de las llamadas contestadas y con esto el monto de dinero utilizado por transportar el tráfico de voz.

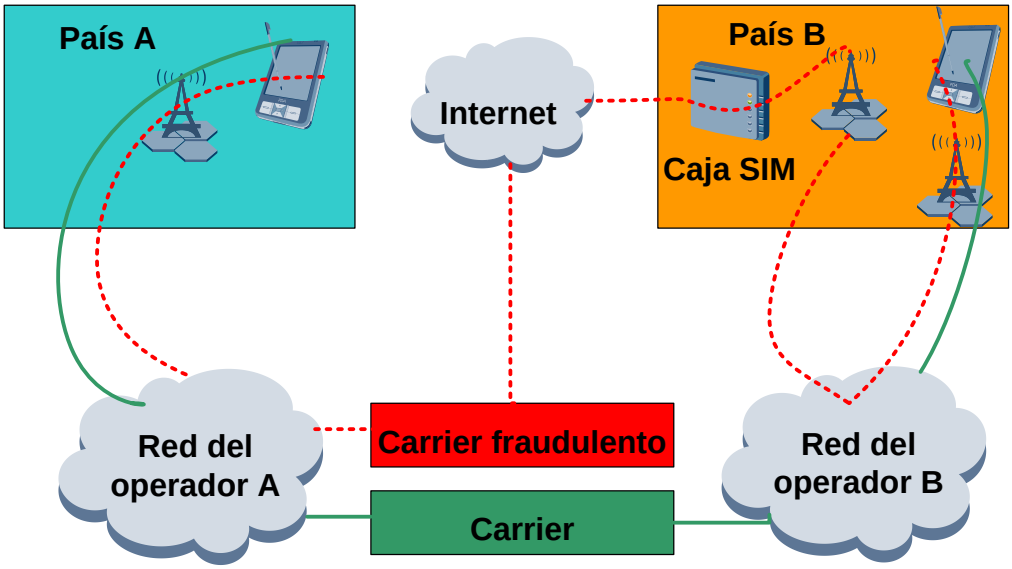
3.5.1.2. Bypass internacional

Este fraude también se conoce como terminación de tráfico internacional y comúnmente se utilizan cajas SIM para perpetrarlo. Las cajas SIM secuestran el tráfico de voz de llamadas internacionales y las transfieren de Internet hacia un dispositivo celular el cual inyecta el tráfico de vuelta a la red celular.

Como resultado las llamadas se vuelven locales en la red donde terminan las mismas, así los operadores celulares y los *carriers* internacionales no reciben el pago del enrutamiento de esta llamada secuestrada. En algunos países no solo se secuestra el tráfico internacional sino también algunas llamadas locales en las cuales la tasa de cobro varía de acuerdo al área con la cual se quiere conectar una llamada.

Las cajas SIM, además de causar pérdidas económicas para los operadores y *carriers*, también degradan el servicio local en las coberturas donde operan. A menudo las celdas celulares son sobrecargadas y las llamadas que son conectadas por medio de estas cajas tienen muy mala calidad lo que resulta en la inconformidad de los clientes.

Figura 26. ***Bypass* internacional utilizando caja SIM**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

Este tipo de fraude ocurre cuando el costo de terminación de las llamadas internacionales excede el costo de las llamadas locales, los defraudadores obtienen ganancias ofreciendo un bajo costo en llamadas de larga distancia a los operadores. Para poder realizar el fraude de *bypass* o enrutamiento ilegal de una llamada es necesario que los defraudadores obtengan una gran cantidad de tarjetas SIM ya sea comprando las mismas o realizando clonaciones, es decir, copiando la identidad de una tarjeta a otra, luego las instalan en un dispositivo electrónico el cual se convierte en una caja SIM.

En la figura 26 se puede apreciar un tipo de escenario en el cual se utiliza una caja SIM para realizar un *bypass* de llamadas internacionales. En color verde, se puede distinguir la ruta de una llamada legal, la cual es conectada entre el país A y el país B, pasando a través de los operadores de ambos

países y conectada por un *carrier* internacional, este último y el operador B generan un cobro por el uso de la red.

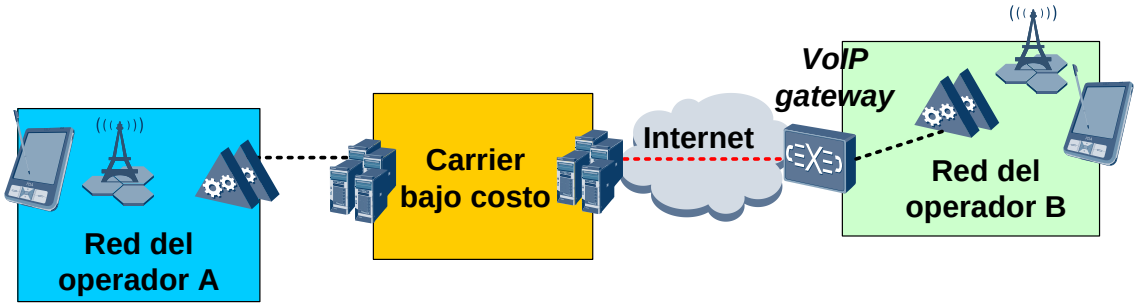
En el caso de una llamada ilegal, que se puede apreciar en la línea roja punteada, la misma llamada generada desde el país y operador A se transmite a través de un *carrier* internacional que ofrece bajo costo en la conexión, sin embargo, la llamada es finalizada de manera ilegal utilizando una ruta sobre Internet y una caja SIM. Así el *carrier* que ofrece precios bajos evita pagar al operador B el costo de la conexión de una llamada internacional ya que por medio de la caja SIM la inyecta en la red del operador B como una llamada local ya que se utilizan tarjetas SIM del operador donde termina la llamada.

Normalmente las rutas sobre Internet que se utilizan para conectar las llamadas hacia una caja SIM son VoIP. Un defraudador puede utilizar un servidor propio u obtener acceso ilegal un equipo VoIP para realizar el enrutamiento del tráfico internacional.

En el caso que un defraudador secuestre una ruta VoIP la cual permita la conexión de llamadas internacionales, no sería necesario que se cuente con una caja SIM, debido a que la llamada estaría siendo inyectada directamente en la red de algún operador.

El escenario en el cual se secuestra una ruta VoIP y es utilizada para la terminación de llamadas internacionales se ilustra en la figura 27.

Figura 27. ***Bypass* internacional utilizando VoIP *gateway***



Fuente: elaboración propia, utilizando programa Microsoft Excel.

Como se puede apreciar en la figura 27, un *carrier* que ofrece tarifas bajas para la transmisión de tráfico internacional, realiza un enrutamiento a través de Internet hacia un *gateway* VOIP; este equipo interconecta el mundo de la telefonía sobre IP con la telefonía fija o móvil de un operador. Un defraudador puede tener acceso al *gateway* por varios métodos, por ejemplo, MITM o comprando la información de acceso en el mercado negro.

Para evitar ser detectados, los defraudadores utilizan varios métodos para manipular la identificación del número de A, por ejemplo, insertando o borrando parte del número para que el número que recibe la llamada no pueda identificar de donde proviene la llamada. Otro método es borrando la identificación de la línea, de esta forma el número que origina no será mostrado en la llamada.

Hoy en día, los defraudadores utilizan herramientas de *software* que simulan comportamiento humano de marcación y patrones de tráfico con el fin de evitar ser detectados por los operadores.

Dentro la tecnología para manipular los patrones de marcación, se encuentran los servidores SIM los cuales se pueden instalar en cualquier parte

del mundo. Para evitar la detección se virtualizan las tarjetas SIM, de esta forma estas pueden ser rotadas entre llamadas.

Otro método para evitar ser detectados es mediante la modificación de la ubicación de las cajas SIM, estas normalmente se colocan en lugares donde existen varias celdas que dan cobertura, con el objetivo de no saturar solo una celda y ser detectados también, pueden movilizar las cajas SIM con algún vehículo.

3.5.2. IRSF

El IRSF *international revenue sharing fraud*, es uno de los fraudes más persistentes dentro de la industria de las telecomunicaciones. Los defraudadores utilizan recursos ilegales para obtener acceso a la red de un operador para poder generar tráfico hacia números telefónicos que son obtenidos de un IPRNP, *international premium rate number provider*.

Los IPRN *international* p*remium rate numbers*, son números hacia los cuales las llamadas terminantes se cobran a una tasa mucho más alta que a un número normal. Parte de ese cargo extra se le paga a un proveedor de un servicio, permitiendo que negocios sean fundados vía telefónica.

Ejemplos de estos servicios son las líneas calientes o de sexo telefónico, líneas humorísticas, servicios de psicólogos o cualquier servicio de atención al cliente vía telefónica. Estos tipos de servicios han existido desde el inicio de las comunicaciones telefónicas y son muy comunes en cualquier parte del mundo.

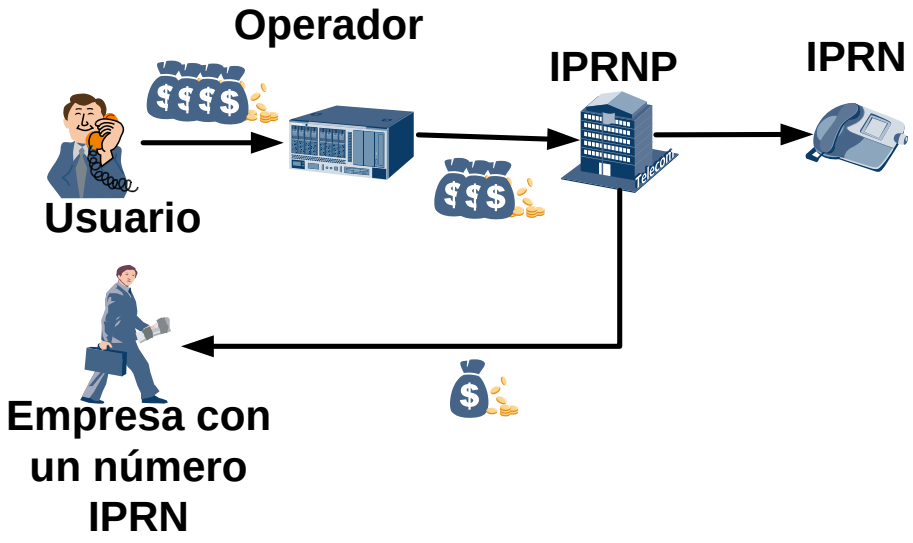
El problema de los IPRN es que es muy fácil realizar fraude sobre ellos, el dinero que es pagado a las empresas es un gran incentivo para generar tráfico

falso hacia estos números. La simulación del acceso a los PRN, es un esquema en el cual un defraudador inyecta tráfico hacia sus propios números PRN.

Es impactante la facilidad en que una persona puede empezar con un servicio IPRN, ya que existen varias empresas que pueden ser encontradas con una simple búsqueda en algún buscador de Internet. Un defraudador necesita únicamente contactar a estas empresas por Internet, las cuales poseen una lista de números IPRN.

Muchos de los servicios que se ofrecen son atendidos por un IVR o una contestadora automática, así que el defraudador únicamente tiene que especificar de qué manera desea que se atienda la llamada.

Figura 28. **Flujo de llamada IPRN**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

En la figura 28 se puede observar a un usuario utilizando un servicio IPRN; el usuario es el que paga la conexión de la llamada y el dinero es distribuido entre el operador, el proveedor de números IPRN y la empresa que adquiere un número IPRN.

Este tipo de fraude es muy difícil de eliminar debido a la complejidad de los sistemas de las redes móviles y la participación de múltiples operadores. Este fraude abarca grupos organizados que usan conexiones obtenidas ilegalmente para generar grandes volúmenes de tráfico de voz hacia los números IPRN y también capitalizando la habilidad de las tarjetas SIM de realizar *roaming*. Ya que algunas veces se necesita entre 24 y 36 horas para que esos registros de llamadas regresen a la red del subscriptor.

De esta forma los defraudadores tienen la oportunidad de marcar la cantidad de números IPRN que puedan antes de que la red local de un subscriptor detecte este tráfico y bloquee las tarjetas SIM.

4. MÉTODOS PARA LA DETECCIÓN Y PREVENCIÓN DE ATAQUES A TRAVÉS DE INTERNET

Considerado las amenazas y vulnerabilidades de las redes VoIP y de los elementos que las componen, en este capítulo se describen los métodos y las estrategias para la prevención de ataques o fraudes sobre cualquier red de un operador de telefonía.

Cabe mencionar que ninguna red es impenetrable, ya que en algunas ocasiones los defraudadores son operadores de las redes en las cuales se realizan ataques o fraudes; sin embargo, con la aplicación de los métodos que se describirán en los enunciados siguientes, es posible aumentar la seguridad de las redes de telefonía y a la vez estar preparado en el momento que ocurra una situación en el cual los servicios sean afectados o alguna red de telefonía sea utilizada de manera fraudulenta.

4.1. Dispositivos para la protección para una red IP

Los elementos de red que están más propensos a ataques no necesariamente son los que están en el interior de la red de una red móvil o fija, sino los elementos que se encuentran en la frontera de las mismas y que conectan otras redes públicas, como Internet, hacia las redes privadas de un operador.

Por otro lado, no necesariamente se obtiene acceso a un equipo que esté dentro de la frontera de dos redes al obtener la clave y usuario para entrar al mismo, debido a que es muy común que las redes de mantenimiento y servicios

estén separadas totalmente. En este enunciado se dan a conocer los equipos que deben instalarse dentro de las redes de telefonía y en sus fronteras para proteger ataques de denegación de servicio o envío de tráfico no deseado desde destinos desconocidos.

4.1.1. *Firewall*

Los *firewalls* o cortafuegos son elementos de una red IP que mejoran las políticas de seguridad. Un *firewall* es utilizado para filtrar el contenido de los paquetes IP, comúnmente en las capas de aplicación, transporte, red y enlace de datos del modelo OSI. Utiliza reglas de filtración o políticas para establecer qué paquetes pueden o no pasar a través de este dispositivo.

Debido a que los paquetes o la información deben pasar en ambos sentidos de este dispositivo, para que la red la cual se está protegiendo sea útil, no todos los ataques pueden ser detenidos por los *firewalls.*

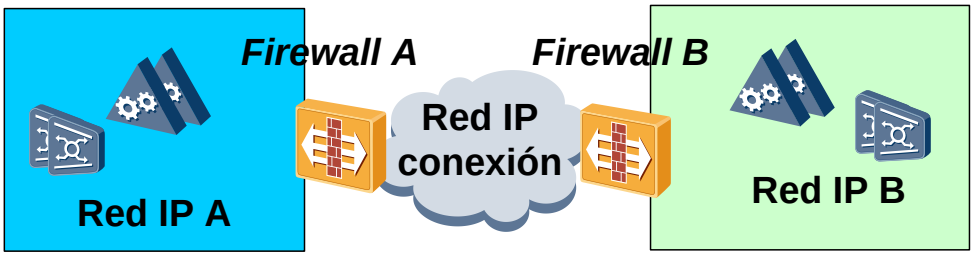
Con la evolución de las redes de telefonía a comunicarse por medio del protocolo TCP/IP, es necesario entonces incluir este elemento de red entre los límites de dos redes IP. Todo el tráfico que pasa entre dos o más redes debe pasar por los *firewalls.* Dentro de la telefonía, se establecen redes de comunicación IP según de la zona geográfica en la cual se instale un dispositivo ya sea con el objetivo de crear redundancia entre los nodos que proveen el mismo servicio o para dar servicio en esa zona.

El filtrado de paquetes se realiza por medio de la verificación de los encabezados de los paquetes de red y decidiendo que paquetes o no pueden pasar de acuerdo a las políticas establecidas en el *firewall*. La filtración de

paquetes es atractiva debido a que no depende la cooperación de ningún usuario, tampoco requiere ninguna acción especial.

Las políticas se pueden establecer dependiendo de la dirección IP fuente, la dirección IP destino, las opciones de un encabezado de un paquete, protocolo de transporte, puerto lógico fuente o destino entre otros, para que un *firewall* decida si dejar pasar o no un paquete.

Figura 29. **Posicionamiento de los *firewalls* en redes IP**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

Como se puede apreciar en la figura 29, cuando los dispositivos de una red de telefonía pertenecen al mismo segmento de red, la comunicación entre ellos se puede establecer sin la necesidad de la conexión hacia un *firewall,* debido a que se considera que la conexión es segura y que las redes se segmentan de tal forma que se ocupen en su totalidad los recursos disponibles. Mientras que si la conexión es entre dispositivos de distintas redes, es necesario permitir únicamente el paso de la información acordada en el planeamiento de dicha interconexión.

Debido a que los *firewalls* son dispositivos de control de la conexión entre dos o más elementos de red, es necesario que los administradores de las

plataformas de telefonía no tengan acceso a los mismos y viceversa. Esto con el objetivo de evitar que una sola persona pueda generar conexiones ilegales hacia otras redes con el fin de obtener algún beneficio económico. Haciendo esto se pueden proteger las redes de telefonía de ser manipuladas por un atacante interno.

Dependiendo de qué criterio utilice un *firewall* para examinar un paquete IP, por ejemplo el encabezado, el contenido del mismo o el patrón generado por una secuencia de paquetes, este se puede clasificar por tipos o técnicas, por lo general un mismo *firewall* puede soportar varias técnicas de filtración.

4.1.1.1. *Firewall* de filtrado de paquetes

Este tipo de *firewall* aplica una serie de reglas en los paquetes IP entrantes y salientes, para luego reenviar o descartar estos de acuerdo a estas reglas. La filtración de paquetes es construida en base a una lista de reglas que se basan en el contenido de los campos en los encabezados IP o TCP.

Si una regla se cumple, entonces el *firewall* deberá determinar si descartar o reenviar el paquete IP, en dado caso no exista ninguna regla que aplique al paquete analizado, entonces se debe establecer una regla predeterminada de cómo proceder.

La política predeterminada más conservadora, es la de descartar todos los paquetes, es decir, inicialmente toda la comunicación se encuentra bloqueada y los servicios que desean integrarse deben ser agregados uno por uno.

A continuación, se presenta en una tabla en la cual se enumera en base a qué información este tipo de *firewall* puede realizar el filtrado.

Tabla VI. **Información utilizada para la filtración de paquetes IP**

Campo	Descripción
IP fuente	Dirección IP que genera el paquete IP, por ejemplo 172.22.0.5.
IP festino	Dirección IP que el paquete IP quiere alcanzar, por ejemplo 172.28.0.2.
Puerto de transporte	El número de puerto a nivel de la capa de transporte que es utilizado para la comunicación, este define también que aplicación se está utilizando.
Protocolo IP	El protocolo de transporte utilizado, por ejemplo TCP o UDP.
Interfaz	Cuando un firewall tiene más de tres interfaces, porque interfaz el paquete entra o por cual interfaz el paquete sale.

Fuente: elaboración propia, utilizando programa Microsoft Excel.

Entre las ventajas de este tipo de *firewall* está la simplicidad para realizar el filtrado de paquetes y la rapidez con que se realiza esta acción, ya que el filtrado es transparente para los usuarios. Sin embargo, debido al reducido número de variables que se utilizan en las reglas, lo hace susceptible a problemas de seguridad debido a la mala configuración de las políticas.

También están otras desventajas ya que la filtración de paquetes no analiza las capas superiores del modelo OSI, no se pueden prevenir ataques que utilicen vulnerabilidades en la parte de aplicación, media vez se le permite el acceso a una conexión, todas las funciones y comandos de una aplicación en específico son permitidos. A la vez, no es soportada la función de técnicas avanzadas de autenticación, esto se debe también a la falta de funcionalidades en las capas superiores.

Existen algunas técnicas para poder burlar las políticas de un *firewall*, por ejemplo, un atacante puede transmitir paquetes desde fuera de una red, pero

con una IP fuente que existe dentro de la red que protege el *firewall.* El atacante espera que con este método pueda lograr penetrar el sistema, si este utiliza seguridad basada en direcciones fuente. Para evitar este tipo de ataques, se debe establecer dentro de las políticas de seguridad que los paquetes que lleguen desde una interfaz externa, con una IP local fuente, sean descartados.

Otro tipo de ataque es utilizado por los atacantes es el de la fragmentación de los paquetes IP, normalmente un *firewall* solo analiza el primer fragmento de un paquete IP, es decir, todos los fragmentos siguientes son rechazados o reenviados dependiendo de lo que suceda con el primero.

Un atacante entonces crea fragmentos extremadamente pequeños, con esto la información del encabezado TCP es separada también esperando que el *firewall* examine únicamente el contenido del primer fragmento y deje pasar a los fragmentos restantes. Este ataque puede ser evitado permitiendo únicamente pasar aquellos fragmentos de paquetes que incluyan suficiente información en el encabezado TCP.

4.1.1.2. *Firewall* con inspección de estado

Un *firewall* de inspección de estado pasa la mayor parte del tiempo examinando la información de un paquete en la capa de transporte y capas inferiores, a menudo también existen métodos más avanzados de inspección para la capa de aplicación. Lo que este tipo de *firewall* examina son aquellos paquetes que inician una conexión.

Si el paquete que es inspeccionado cumple con alguna regla establecida para reenviarlo, el paquete es dejado pasar y su entrada es agregada a una tabla de estado. Desde este punto, los paquetes de la misma conexión se

reenvían de acuerdo a la entrada ya existente en la tabla de estado, sin que se requiera análisis adicional. Esos paquetes únicamente deben tener información de la capa de red y de transporte, es decir, una dirección IP y un número de puerto lógico, para poder confirmar que estos paquetes pertenecen a la conexión que está almacenada en la tabla de estado.

Este método incrementa el rendimiento de los *firewalls* comparados con otros que analizan todos los paquetes, porque solo el paquete que inicia la conexión debe ser desencapsulado.

Uno de los conceptos más difíciles de comprender cuando se habla sobre los *firewalls* y el protocolo TCP/IP es el concepto de estado. La razón por la cual es un tema elusivo es porque puede significar varias cosas dependiendo de la situación. Básicamente, estado es la condición de ser o estar de una determinada sesión de comunicación. La definición de esta condición para un dispositivo o sesión difiere dependiendo de la aplicación con la cual dos entidades se comunican y también del protocolo que utilizan para intercambiar la información.

Los dispositivos que guardan el estado a menudo almacenan esa información en una tabla, esta tabla de estado retiene entradas que representan todas las sesiones de comunicación de que el dispositivo tiene conocimiento. Cada entrada de la tabla tiene información que identifica de forma única la sesión que representa. Dicha información puede incluir la IP fuente y destino, número de secuencia o de reconocimiento, etc.

Los protocolos de transporte pueden tener sus propios estados de conexión los cuales pueden ser rastreados en diferentes formas. Muchos de los atributos que son parte de una sesión de comunicación, como las direcciones

IP, números de secuencia, números de puerto, banderas de estado etc., pueden ser utilizados como la huella digital de una conexión individual. La combinación de esta información es la que se almacena en una tabla de estado.

Cuando el protocolo utilizado es el TCP, las entradas de la tabla de estado son removidas luego que la comunicación de una sesión es terminada para prevenir que alguna sesión sea cerrada inadecuadamente; las sesiones cuentan con un tiempo durante el cual son válidas.

Para los protocolos UDP e ICMP, que son protocolos que no estaban basados en conexiones, la comunicación se establece sin dejar rastro de un estado, por lo cual es más difícil analizar este tipo de comunicaciones.

Un *firewall* de inspección de estado se configura para analizar este tipo de comunicación considerando únicamente IPs o puertos lógicos, en el caso de una conexión UDP y en el caso del ICMP se guarda record del mensaje de solicitud y mensaje de respuesta. Para estos dos protocolos hay una indicación de que la comunicación ha sido terminada, debido a esto se debe configurar un tiempo para que el *firewall* borre las entradas UDP o ICMP, para que la tabla de estado no se llene.

La mayoría de los *firewall* de inspección de estado son capaces de analizar y filtrar tráfico a nivel de la capa de aplicación, es muy común que solo se analice el primer paquete con el cual se establece la comunicación; por lo mismo todos los paquetes siguientes son analizados de la capa de transporte a las capas inferiores.

Este es un método eficiente para analizar la comunicación, sin embargo, no tiene la capacidad de considerar todos los diálogos en una sesión de una

aplicación. Por esto cualquier comportamiento anómalo a nivel de aplicación que suceda luego del análisis del primer paquete no puede ser realizado. Sin embargo, cualquier ataque que ocurra deberá ser perpetrado utilizando las mismas IPs fuente y destino sobre el mismo puerto lógico, debido a la tabla de estado y dentro del periodo de tiempo de validez establecido para una sesión.

4.1.1.3. *Gateway* a nivel de aplicación

Un *gateway* a nivel de aplicación, también llamados *proxy* de aplicación, actúa como un relé para el tráfico de la capa de aplicación. Un usuario contacta al *gateway* usando una aplicación TCP/IP, como por ejemplo telnet, luego este le solicita al usuario el nombre de la entidad remota con la cual desea establecer la comunicación. Cuando el usuario contesta y provee un nombre de usuario y parámetros de autenticación, el *gateway* contacta la aplicación de la entidad remota y le entrega los paquetes TCP que contiene la información para la comunicación de las dos entidades.

Si el *gateway* no tiene el código *proxy* de una aplicación, el servicio no es soportado y no puede ser retrasmitido a través del *firewall*. Con esto, un *gateway* puede ser configurado para soportar únicamente servicios específicos de una aplicación, los cuales el administrador de una red considere que deben ser aceptados, denegando el resto de los servicios de la misma aplicación.

Este tipo de *firewalls* son más seguros que los que utilizan filtración de paquetes, porque en vez de tratar de lidiar con las múltiples posibles combinaciones para permitir y bloquear paquetes a nivel IP y TCP, únicamente validan unas pocas aplicaciones que son permitidas. Adicionalmente es más fácil auditar todo el tráfico a nivel de la capa de aplicación.

La desventaja principal de los *gateway* a nivel de aplicación es el procesamiento adicional que se realiza en cada conexión. En efecto, hay dos conexiones que se establecen entre el g*ateway* y dos usuarios finales por lo cual se debe examinar y reenviar el tráfico en dos direcciones.

4.1.1.4. *Gateway* a nivel de circuitos

También conocido como *proxy* a nivel de circuitos, así como el *gateway* a nivel de aplicación, este tipo de *firewall* no permite conexiones punto a punto, sino que establece dos conexiones TCP: una hacia la entidad local y otra hacia la entidad remota; luego de que se validan ambas conexiones, el *gateway re*envía la información de una conexión hacia la otra sin examinar el contenido de los paquetes. La seguridad del g*ateway* a nivel de circuitos consiste en determinar qué conexiones son permitidas y cuáles no.

Un uso típico de este tipo de *firewall* es en situaciones donde el administrador del sistema confía en los usuarios internos. Se puede realizar la configuración de tal forma que utilice para las conexiones entrantes un *gateway* nivel de aplicación y para las conexiones salientes la configuración como de *proxy.* Con esta solución el *gateway* puede analizar todo el tráfico entrante en busca de aplicaciones no permitidas, pero no perder mucho tiempo en analizar el tráfico hacia los servidores que controlan esas conexiones.

4.1.2. SBC

Como se vio anteriormente, un *firewall* es un elemento de red que permite o deniega el acceso de paquetes IP basándose en políticas que se aplican desde la capa de transporte hacia las capas inferiores. Aunque existen *firewalls* que también pueden realizar análisis en las capas superiores, su diseño está

más orientado al análisis de sesiones de datos pero no de sesiones de voz, por ejemplo, protocolo SIP utilizado en la comunicación VoIP.

Una de las razones por las cuales existen los SBCs, *session boarder controller*, viene *inicialmente* de un error en el desarrollo del protocolo SIP, ya que este no considera la existencia de la técnica de NAT (*network adress translator*) en las redes IP. Esta técnica se utiliza para limitar el número de IPs públicas que una red utiliza si, por razones de economía y seguridad. Normalmente esta técnica se utiliza en los *routers* o *firewalls.*

Como el SBC fue desarrollado principalmente para operar con el protocolo SIP, su diseño también consideración la seguridad y vulnerabilidad del protocolo SIP, lo cual lo hace un dispositivo indispensable para una red VoIP.

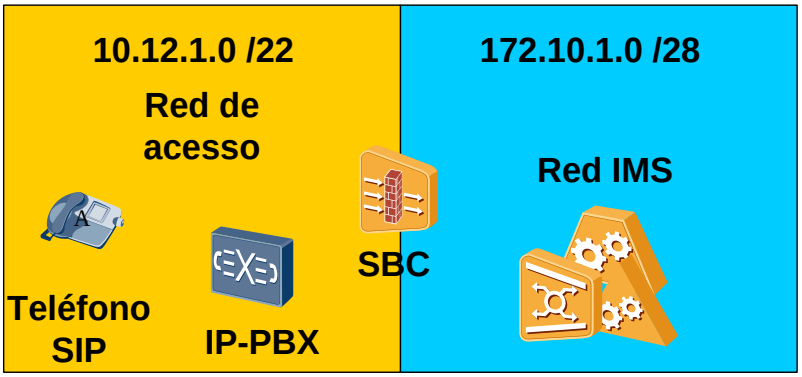
4.1.2.1. Ocultamiento de la topología de red

Como resultado del establecimiento de una sesión IP, las entidades finales involucradas en la conexión conocerán las direcciones IP por las cuales se envía el tráfico de media. Esto significa que un usuario SIP que intenta llamar a un usuario PSTN, también conocerá la IP del *gateway* PSTN que es responsable de unir la red VoIP con los servicios PSTN.

Un usuario malicioso podría usar esta información para obtener acceso al *gateway* PSTN o para realizar algún tipo de ataque para denegar o degradar el servicio. Al tener la habilidad de contactar al g*ateway* PSTN directamente el atacante, también, podría hacer mal uso de cualquier problema de seguridad que pueda existir en el *gateway* con el objetivo de reenviar llamadas a la red PSTN haciendo que el operador no obtenga beneficio por las mismas.

Para ocultar los elementos de red internos de un operador, cualquier mensaje que deje la red debe atravesar un SBC. El SBC reemplaza las direcciones IP de los elementos de red internos, con sus propias direcciones IPs asignadas. De esta forma los encabezados de los mensajes SIP que contienen información de direcciones IP incluyen únicamente direccionamiento IP del SBC. Este cambio de IP se hace de igual forma para los paquetes de media cambiando el direccionamiento de la información SDP.

Figura 30. **Ocultamiento de la red interna utilizando un SBC**



Fuente: elaboración propia, utilizando programa Microsoft Visio.

Como se puede apreciar en la imagen 30, el SBC se coloca en la frontera o borde de la red acceso SIP, en este caso en particular, y la red interna del operador, en este caso la red IMS. En el caso de la red de acceso se puede ver un direccionamiento IP utilizando la red 10.12.1.0 /22, la barra 22 es una abreviación de las máscara 255.255.252.0, donde se puede tener hasta 1,022 IPs asignables. Para la red IMS se tiene asignado un segmento de red distinto el 172.10.1.0 /28, con 14 IPs asignables por la máscara 255.255.255.240.

Se dice, entonces, que el SBC está conectado a ambas redes e internamente realiza el cambio de IPs en los encabezados de los mensajes SIP, tanto para el tráfico de entrada como de salida, de esta forma los usuarios que se conectan por la red de acceso no tienen visibilidad del direccionamiento IP de la red IMS.

4.1.2.2. NAT transversal

El NAT se desarrolló para contrarrestar la falta de direcciones IPv4 usando la habilidad de esconder la red de un operador detrás de una o pocas direcciones IP. Los elementos de red detrás de un NAT utilizan un direccionamiento privado y no pueden alcanzar directamente el Internet público.

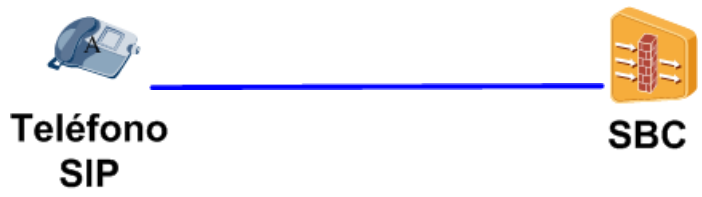
En el caso en que un usuario VoIP esteé localizado detrás de un NAT, este tendrá IPs privadas en el encabezado SIP y en la parte SDP. Con esta información será imposible que pueda ser contactado por otro usuario desde Internet.

Para que un usuario que está detrás de un NAT pueda ser localizado a través de las interfaces públicas de un SBC, este último debe manipular la información de registro del usuario. Un usuario VoIP incluye su IP privada como información de contacto en los mensajes REGISTER. Cualquier llamada hacia esta dirección fallaría ya que la IP privada no es alcanzable a través de una red pública.

El SBC reemplaza la información del campo de contacto en el encabezado de los mensajes SIP con su propia dirección IP. Y esta es la dirección IP que queda registrada en el registrador, que puede ser una red IMS. De igual forma

las llamadas destinadas para este mismo usuario serán enviadas antes al SBC, un ejemplo de NAT transversal se puede ver en la siguiente figura.

Figura 31. **NAT transversal**



IP origen: 10.10.10.2:1024
INVITE sip +50240170000@ejemplo.org SIP/2.0
From: <sip:user@prueba.com> To:<sip: +50240170000@ejemplo.org> Contact: <sip:user@10.10.10.2:1024>
SDP
c= IN P4 10.10.10.2

IP origen: 217.9.30.50:1024
INVITE sip +50240170000@ejemplo.org SIP/2.0
From: <sip:user@prueba.com> To:<sip: +50240170000@ejemplo.org> Contact: <sip:user@217.9.30.50:1024>
SDP
c= IN P4 217.9.30.50

Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

Para que se pueda saber a qué usuario se debe contactar, el SBC puede guardar una copia local de la información de registro, esta copia incluye la IP privada y la dirección URI así como también la IP pública que fue incluida en el encabezado SIP por la funcionalidad de NAT transversal.

Alternativamente el SBC puede insertar esta información en los mensajes SIP reenviados, de esta forma no se requiere almacenar en una tabla la información de cada usuario en el SBC. Esta es la mejor solución ya que almacenar una copia de los datos de registración aumenta los requerimientos de procesamiento y memoria del equipo.

En la figura 31 se pueden apreciar en rojo los cambios de direccionamiento IP que realiza el SBC con el propósito de ocultar la topología de red y para reducir el número de IPs públicas necesarias para interconectar un usuario privado con una red pública como Internet.

4.1.2.3. Protección contra ataques DoS

Como cualquier otro dispositivo que da servicios sobre Internet, el SBC puede ser el objetivo de ataques para la denegación de servicio. Este tipo de ataques puede ser disfrazado como tráfico VoIP legítimo, así que distinguir entre un ataque de denegación de servicio y un pico de tráfico no siempre es posible.

Por esta razón, los operadores de equipos VoIP tienen que incorporar mecanismos para monitorear la carga y el tráfico entrante a la red, identificar cuáles son los recursos sobrecargados y qué está causando el problema. Luego de detectar esto, debe reaccionar de una manera que pueda evitar la interrupción del servicio total.

Para poder evitar el tráfico malicioso y cualquier sobrecarga sobre los equipos detrás del SBC, se deben incluir algunos mecanismos de defensa en este último equipo.

Una de las características que debe poseer un SBC es la capacidad de limitar la taza de llamadas y registros de usuarios. Cuando los límites establecidos son alcanzados, el SBC comienza a descartar los mensajes. Los límites pueden ser aplicados de tal forma que no se acepte una cantidad determinada de mensajes de registro de una IP determinada o limitarlo a todos los suscriptores.

El bloqueo por lista negra también es otra característica que debe estar disponible, las llamadas de los números de subscriptores que se encuentren en esta lista son desechadas sin que exista un procesamiento previo de las mismas. Sin embargo, no todo el tiempo se pueden detectar los números que generan llamadas maliciosas; sin embargo, el SBC a menudo monitorea el tráfico entrante y si ciertas características son identificadas entonces los subscriptores VoIP son dinámicamente agregados a la lista negra.

Las características para realizar el bloqueo automático son: el número de mensajes enviados desde un usuario en un periodo de tiempo, el contenido de los mensajes SIP o la distribución de los números a los que se llama, por ejemplo, un mismo número que llama a varios destinos en periodos cortos de tiempo podría ser un indicador de que alguien está buscando alguna debilidad en la red. Cuando un número es agregado en la lista negra, todos los mensajes de esta fuente son desechados.

Un SBC también debe poseer un mecanismo de detección de mensajes con contenido malicioso, algunos de estos tipos de mensajes fueron descritos en el capítulo 3. Por medio del análisis del contenido de los mensajes SIP, un SBC puede detectar este tipo de mensajes malformados y proteger a los elementos de red que estén detrás de este.

Los usuarios del servicio VoIP esperan que el servicio esté siempre disponible, independientemente de que ocurra algún ataque, así que un SBC debe proveer priorización de las llamadas para ciertos tipos de usuarios incluso cuando hay una sobrecarga del sistema o un ataque en curso. Para lograr esto se debe identificar las llamadas generadas por los usuarios registrados en la red de un operador y guardar el registro en una base de datos local. Así en

escenarios de sobrecarga o ataque, únicamente son procesadas las llamadas de estos usuarios.

4.1.2.4. Control de acceso

Una de las características implícitas en el funcionamiento del SBC es la de controlar qué usuarios y qué mensajes pueden atravesar los bordes de una red VoIP y qué servicios VoIP pueden ser usados.

Un SBC debe tener una lista negra y también una lista blanca de los usuarios y fuentes de los cuales un SBC debe aceptar o desechar los mensajes.

Otra de las características que debe poseer es el control de la media; el SBC debe asegurarse que únicamente los usuarios que hayan establecido una sesión SIP sean capaces de recibir la media. De esta forma un SBC puede prevenir que un usuario malicioso contacte a un *gateway* PSTN o un servidor de aplicación directamente.

El precio de una tarifa plana a menudo es determinado por el comportamiento esperado de un usuario, sin embargo, un operador puede enfrentarse al caso en que algunos suscriptores de telefonía residencial revenden los minutos disponibles.

Este tipo de comportamiento causa pérdidas financieras para el operador y en algunos casos sobrecarga de la red. Para suprimir este tipo de fraude un SBC debe ser capaz de limitar el número de llamadas paralelas generadas por un usuario, así como su duración y frecuencia.

Para que un SBC pueda observar el número de llamadas paralelas que un usuario hace, se requiere que el comportamiento de los usuarios sea analizado a lo largo de un periodo de tiempo. Esta tarea puede ser delegada a otro tipo de equipos o soluciones desarrolladas especialmente para la detección de fraudes.

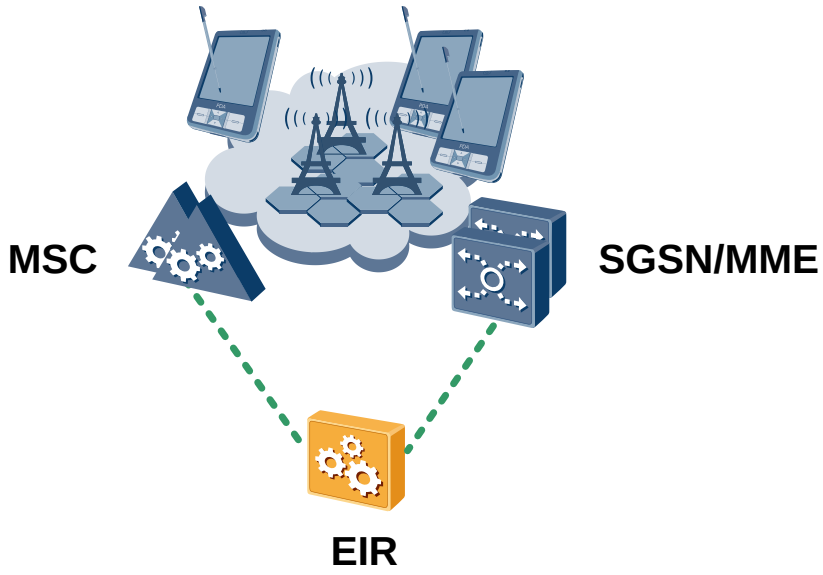
4.1.3. EIR

Un EIR, *equipment identity register*, almacena la información de estado de un IMEI, *international mobile equipment identity*. Dicho con otras palabras, un EIR es una base de datos que contiene el listado de las terminales móviles que pueden o no tener acceso a la red móvil.

Normalmente, cuando una terminal móvil es robada o extraviada, el operador realiza el bloqueo de la misma en el EIR para que cuando se intente usar cualquier tarjeta SIM la red no permita el registro. Para esto el EIR posee una lista negra y una lista blanca, todos los IMEI registrados en la lista negra no podrán tener acceso a la red del operador y todos los IMEI registrados en la lista blanca pueden tener acceso a la misma.

Normalmente se utiliza únicamente la funcionalidad de la lista negra, debido a que registrar a todos las IMEI que tienen permitido utilizar la red móvil de un operador representa que el EIR debe poseer características de *hardware* y procesamiento superiores a las que se requieren para manejar únicamente las terminales en la lista negra que representan un gasto adicional. A la vez debe considerarse el desarrollo para el aprovisionamiento en este equipo para terminales nuevas.

Figura 32. **Topología del EIR en una red móvil**

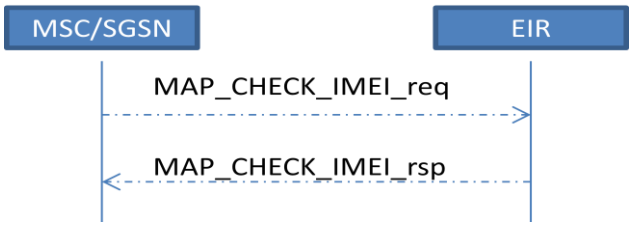


Fuente: elaboración propia, utilizando programa Microsoft Visio.

En la figura anterior se puede apreciar la ubicación del EIR en la topología de una red móvil. Según la configuración de los equipos MSC, SGSN o MME, se envía un mensaje CHECK IMEI hacia el EIR para validar si una terminal puede registrase en la red o no. Normalmente el disparo de este mensaje se configura para realizarse cuando lo usuarios se registran por primera vez en la red o cada vez que se registran con el HLR.

Para evitar que se afecte el registro de los usuarios en dado caso el EIR quede fuera de servicio, debe ser configurable un modo de emergencia, con el cual si el mensaje CHECK IMEI no tiene respuesta, se continúe con el proceso normal de registro y autenticación en la red móvil.

Figura 33. **Mensaje CHECK IMEI**



Fuente: elaboración propia, utilizando programa Microsoft PowerPoint.

La figura anterior muestra los mensajes de señalización que se intercambia la MSC y SGSN con el EIR para validar el estado de una IMEI. El EIR responde dependiendo de cuál sea el estado de la IMEI; lista blanca, lista negra, lista gris o desconocido. De los últimos dos estados, el operador decide qué hacer con las IMEI que se encuentren dentro esta clasificación.

Para evitar que las terminales que no están registradas en el EIR como lista blanca puedan utilizar la red móvil, se debe configurar que todas las IMEI que el EIR no conozca, no puedan tener acceso para utilizar los servicios móviles.

Otra de las características que debería poseer un EIR es la detección de IMEI clonadas, es decir, que una misma terminal se registre más de una vez en la red del operador con distinta tarjeta SIM. Cuando una terminal es robada y bloqueada para ser utilizada, la única forma de hacerla funcional otra vez es cambiándole de IMEI. Esta es una práctica ilegal ya que utiliza la identidad de otro equipo para poder utilizar los servicios de la red de un operador con una terminal robada. Esto se realiza asociando un terminal o una o varias tarjetas SIM.

Una de las técnicas utilizadas por los defraudadores, es la clonación de tarjetas SIM. Las cuales pueden ser utilizadas en las cajas SIM o vendidas en el mercado negro. Para que el EIR pueda detectar la clonación de tarjetas SIM, cada tarjeta SIM debe ser asociada a una o varias terminales. Si otra terminal distinta a las registradas intenta usar esta SIM, en el momento de realizar la consultar al EIR, este responderá el registro en lista negra.

4.2. Encriptación VOIP

El peligro de que una conversación VoIP o mensajería SIP sea interceptada para ser utilizada en algún tipo de los fraudes descritos en el capítulo 3 de esta investigación es realmente alto, ya que existen múltiples herramientas que pueden ser adquiridas de forma gratuita para poder leer e interpretar los mensajes SIP y el contenido de media en los paquetes RTP.

Considerando esto, es necesario entonces establecer algún tipo de encriptación, entre dos o más elementos de red de telefonía, con el fin de evitar que la información interceptada pueda ser interpretada por alguna persona con el fin de cometer un fraude.

4.2.1. TLS

El protocolo TLS, *transport layer security*, es un estándar para crear comunicaciones privadas a través de redes públicas, este provee seguridad e integridad de los datos durante la comunicación.

TLS está diseñado para las conexiones de tipo cliente/servidor con el objetivo de que la mensajería no pueda ser interpretada si es interceptada en

algún punto intermedio ya que provee autenticación y confidencialidad sobre redes como Internet.

Este protocolo utiliza encriptación como mecanismo para ocultar lo enviado entre el cliente y el servidor; la autenticación se utiliza para identificar una conexión válida y la integridad para identificar la manipulación o modificación de la información.

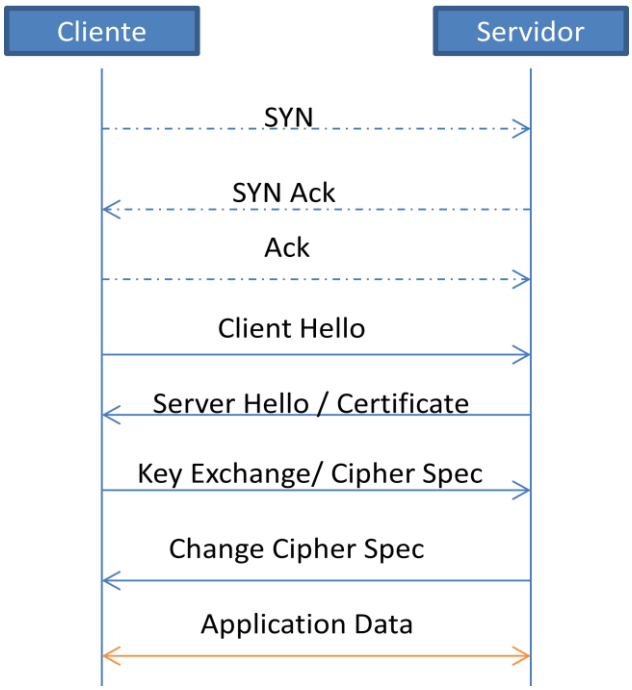
TSL es considerado como un protocolo de la capa de aplicación independiente, consiste en dos componentes principales; el protocolo de *handshake* que se utiliza para establecer sesiones y compartir las llaves privadas de la comunicación; y el protocolo *record* que se utiliza para transmitir de forma segura la información utilizando las llaves privadas.

4.2.1.1. Protocolo handshake

Antes de que dos elementos de red empiecen a intercambiar la información a nivel de capa de aplicación, un túnel con encriptación debe ser negociado, el cliente y el servidor deben acordar qué versión del protocolo TLS utilizar, qué tipo de cifrado y verificar si los certificados son necesarios. Desafortunadamente cada uno de los pasos anteriores requiere mensajes de ida y vuelta entre el cliente y el servidor, lo que agrega un poco de retardo en el establecimiento de las conexiones TLS.

TLS funciona sobre el protocolo TCP, lo que significa que primero se debe finalizar el *handshake* de tres vías del protocolo TCP antes de que que se inicie la negociación del túnel seguro.

Figura 34. **Procedimiento del *handshake* en TSL**



Fuente: elaboración propia, utilizando el programa Microsoft PowerPoint.

Como se puede apreciar en la figura 34, los primeros tres mensajes corresponden al *handshake* del protocolo TCP y los siguientes cuatro corresponden al *handhake* del protocolo TSL. La negociación del túnel seguro del protocolo TSL es un proceso complicado y hay muchas formas de hacerlo en la forma equivocada.

En el mensaje *client hello,* se envía la versión más alta del protocolo TSL que la entidad soporta, un número de 32 bytes, de los cuales 4 bytes son una estampa de tiempo y 29 bytes un número que se genera al azar y una combinación de parámetros de cifrado, en los cuales se incluye el algoritmo de la llave pública, los algoritmos de encriptación y los algoritmos de compresión.

El servidor informa al cliente cuáles son los algoritmos elegidos para el cifrado junto con un número al azar de 32 bytes generado de forma similar al generado por el cliente. El cliente luego de recibir este mensaje genera un número llamado *pre-master secret* que se genera utilizando el algoritmo de la llave pública en conjunto con las llaves públicas del certificado del servidor. Tanto cliente como servidor generan independientemente un número de 48 bytes llamado *long master secret* que sirve para obtener las llaves para intercambiar la información.

4.2.1.2. Protocolo record

Luego de que el cliente y el servidor están listos para comunicarse de manera segura, el protocolo *record* toma la información, la fragmenta en bloques, la comprime y aplica los parámetros de encriptación y cifrado para luego transmitir el resultado. De igual forma este protocolo se encarga de quitar la encriptación y frado, descomprimir y unir de nuevo la información luego que llega a su destino.

4.2.2. IPSEC

Una debilidad del protocolo de Internet es la falta de mecanismos para asegurar la autenticidad y privacidad de la información que se transmite. Hoy en día protocolo IP se utiliza en todas las redes de telecomunicaciones modernas y muchos servicios que los operadores proveen funcionan a través de Internet; debido a esto, es necesario incluir mejoras de seguridad en las redes sobre IP.

El protocolo de seguridad IP o IPSEC está compuesto por una serie de servicios y protocolos que proveen una solución completa de seguridad para una red IP. IPSEC funciona en la capa de red así que puede proveer protección

a las capas superiores del protocolo IP sin la necesidad de métodos de seguridad adicionales.

Cuando dos elementos de una red quieren emplear comunicación segura entre estos, se debe establecer un camino seguro entre los mismos, el cual puede atravesar varios sistemas que pueden ser inseguros. Para lograr esto se debe acordar el conjunto de protocolos de seguridad a utilizar para que la información que sea intercambiada pueda ser interpretada. Se debe también establecer un mecanismo de encriptación para codificar dicha información.

Los protocolos IPSEC son adiciones al protocolo IP que permiten enviar mensajes criptográficamente protegidos, esto se logra con el uso de dos encabezados IPSEC que se insertan inmediatamente después del encabezado IP de cada mensaje.

El encabezado ESP, *encapsulating security protocol*, provee privacidad y protección contra de modificaciones maliciosas y el encabezado AH, *authentication header*, protege contra modificaciones maliciosas, pero sin proveer privacidad.

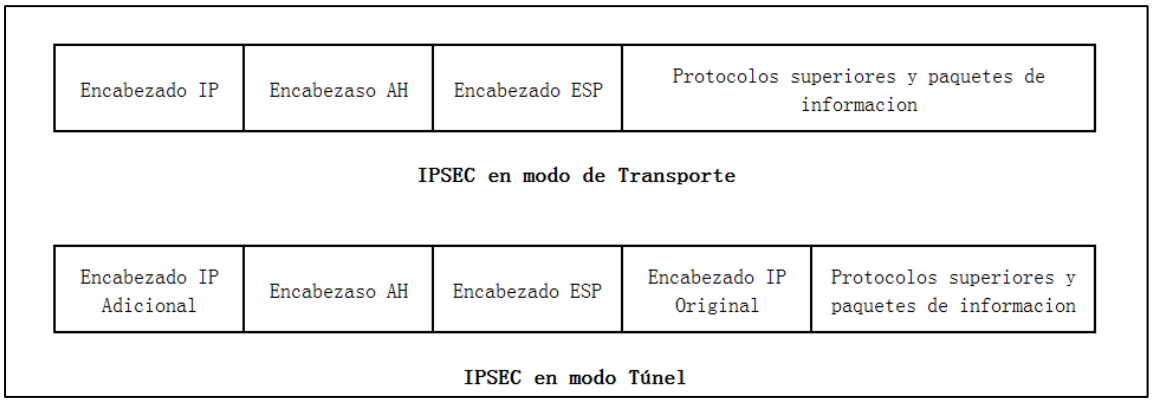
4.2.2.1. Encabezados AH y ESP

AH utiliza un algoritmo de autenticación de mensajes para proveer integridad y protección que no depende de la conectividad. Este tipo de protección cubre partes de los paquetes IP y también algunas partes del encabezado ya que existen algunos campos dentro de los encabezados IP que no pueden cambiar de una forma impredecible mientras cruzan una red IP.

El encabezado ESP puede proveer protección por medio de autenticación e integridad, adicionalmente este encabezado puede utilizar algoritmos de encriptación para proveer confidencialidad. La protección del ESP cubre la información de un paquete IP pero no su encabezado.

Los encabezados AH y ESP identifican cada uno la protección criptográfica aplicada en el paquete IP y a la vez incluyen otra información para decodificar y proteger la comunicación. Si AH o ESP son agregados sin la modificación del encabezado IP, se dice que se utiliza el modo de transporte. Si en cambio existe otro modo llamado de túnel, la diferencia es que el encabezado IP si es modificado, como se puede observar en la figura 35.

Figura 35. **Modo transporte y modo túnel IPSEC**



Fuente: elaboración propia, utilizando programa Microsoft Excel.

El modo de transporte en IPSEC está limitado a una comunicación punto a punto donde cada elemento de red es responsable de proveer las capacidades IPSEC. Por otro lado, el modo túnel en IPSEC, un y *gateway* seguro, puede proveer protección IPSEC a y uno o más elementos de red, incluso a las redes que se encuentren detrás del *gateway.*

Cuando se utiliza el modo túnel se provee protección al análisis de tráfico IP ya que los encabezados AH y ESP protegen al encabezado IP original y a los paquetes de información.

El encabezado ESP puede proveer la misma protección que el encabezado AH, adicionalmente provee privacidad. Existen dos tipos de encabezados por razones políticas ya que en algunos países se prohíbe la exportación de *software* que habilita o incorpora encriptación.

4.2.2.2. Algoritmos criptográficos

Todos los formatos de los paquetes de Internet son bien conocidos y de dominio público, por esta razón un paquete que es enviado sobre Internet puede ser capturado y su contenido interpretado y modificado.

El encabezado de un paquete IP incluye un campo que se llama *checksum* el cual se añade para proteger los paquetes de ser modificados o corrompidos. El *checksum* es un número de 16 bits que se calcula a partir de los valores hexadecimales del encabezado IP.

A pesar de que el campo *checksum* fue creado para proteger la modificación de los paquetes IP, este puede ser fácilmente recalculado por un atacante que modifica el encabezado de un paquete IP. Este problema se puede resolver estableciendo un código secreto. Si un paquete IP es ilegible luego de aplicar este código secreto, entonces, su contenido está seguro, aunque el paquete pueda ser capturado.

Los programas de computación que se utilizan para descifrar códigos o para analizar contenido criptográfico son capaces de romper la seguridad de

algunos códigos secretos complejos. Sin embargo, la información que es imposible de adivinar, incluso utilizando estos programas, debe ser parte integral de los códigos secretos. La información de un código secreto debe ser únicamente conocida por los equipos que participan en la comunicación.

En IPSEC se utiliza un MAC, *messag authenticacion code,* para autenticar un paquete IP; únicamente las entidades que conozcan la llave secreta son capaces de procesar la MAC. La MAC se utiliza para que los elementos de red que hablan IPSEC puedan validar que el mensaje que fue enviado sea exactamente el mismo que se recibe. Las MAC que se utilizan para los encabezados AH y ESP con HMAC-MD5 y HMAC-SHA-1.

Los algoritmos de encriptación del encabezado ESP son orientados a bloques. Cada bloque de texto es transformado en texto cifrado con el uso de un algoritmo de encriptación en conjunto con la llave secreta.

Si cada bloque fuera encriptado por separado, sería más fácil para un atacante descifrar el texto cifrado, debido a que algunas partes de los paquetes IP son conocidas. Cada bloque podría ser desencriptado por separado sin que exista necesidad de conocer el contenido de otro bloque. Una vez que la llave segura sea encontrada, los demás bloques podrían ser desencriptados.

Por esta razón, es mandatorio que cada algoritmo IPSEC incorpore un mecanismo por medio del cual el contenido criptográfico de cada bloque contenga alguna referencia del bloque anterior. El algoritmo que se utiliza en IPSEC es el DES, *data encription standard,* aunque este algoritmo se ha vuelto vulnerable a los ataques, así que en la mayoría de implementaciones de IPSEC se utilizan variantes más confiables como el triple DES.

4.2.3. Negociación IKE

Antes de que dos entidades puedan intercambiar información de una forma segura, ambas partes deben acordar la naturaleza de la seguridad que será aplicada a la comunicación, por ejemplo, el uso de los encabezados AH o ESP, el algoritmo criptográfico a utilizarse, las llaves secretas, la información que se desea proteger, etc.

Una asociación secreta o SA consiste en toda la información que es necesaria para caracterizar e intercambiar una comunicación protegida. El objetivo de la negociación IKE, *internet key exchange*, es el permitir a las entidades remotas acordar dinámicamente los parámetros IPSEC que serán aplicados en comunicaciones futuras.

Esto se logra a través de una negociación de dos fases, en la fase uno se establece una asociación segura ISAKMP, *internet security association and key management protocol*, que se describe como el canal seguro donde la negociación del SA IPSEC toma lugar. En la fase dos se establecen un par de asociaciones seguras IPSEC, la SA entrante y la SA saliente.

Los intercambios más comunes de la fase uno son el modo agresivo y el modo principal. El modo principal de intercambio consiste en seis mensajes mientras que el modo agresivo de intercambio consiste en tres mensajes. Los tres mensajes extras del modo principal agregan protección de identidad, de esta forma las entidades remotas de una conexión IPSEC pueden proteger su identidad de atacantes potenciales. Esto asegura que durante el curso de una negociación IKE las entidades remotas nunca son intercambiadas sin ser encriptadas antes.

La fase uno de la negociación IKE tiene como uno de sus objetivos la negociación de los parámetros de seguridad, los elementos de red que desean establecer una comunicación segura deben acordar los valores y ajustes de un número de parámetros que gobernarán el formato encriptado de los dos últimos mensajes de la fase uno y de todos los mensajes de la fase dos. También se debe negociar qué método se utilizará para la autenticación, el tiempo máximo de vida de la asociación segura de la fase uno y como se medirá ese tiempo.

Durante la fase uno se define también el método a ser utilizado para establecer el secreto compartido. Todos estos valores en conjunto hacen forman la ISAKMP SA.

Una vez que los elementos remotos han acordado que métodos y parámetros se usarán para generar el secreto compartido de la fase uno, un intercambio criptográfico se conduce utilizando el protocolo *diffie-hellman.* Este intercambio será utilizado para generar las llaves secretas.

La autenticación entre las entidades remotas se realiza con base en información adicional que es transmitida fuera de banda. Esta información puede ser una llave precompartida, una firma digital o un método de encriptación y desencriptación utilizando alguna llave pública o privada. La autenticación de los elementos remotos asegura que la asociación segura que se está estableciendo se realice con un elemento remoto identificable.

Una vez que la asociación segura ISAKMP es establecida, esta puede ser utilizada para proteger múltiples intercambios durante la fase dos de la negociación IKE, hasta que expire el tiempo de vida de la fase uno de la negociación.

Los intercambios más comunes de la fase dos de la negociación IKE son el modo rápido de intercambio y el intercambio informacional. El intercambio informacional utiliza la asociación segura de la fase uno para proteger y diagnosticar o para el envío de mensajes informativos.

El modo rápido de intercambio realiza una negociación de una asociación segura IPSEC, durante la fase dos también se definen los parámetros para generar el secreto compartido, el tiempo máximo de vida de la negociación SA, etc. Durante la fase dos se generan valores de autenticación aleatorios para asegurar que la negociación en curso no sea solo una respuesta de alguna negociación anterior.

Con el secreto compartido de la fase uno, se produce el material para obtener las llaves del IPSEC SA de la fase dos. En esta fase se asegura que solo una llave se genere para cada distinto intercambio. Si alguna clave es descubierta por algún atacante, solo la información contenida entre las dos entidades puede ser descubierta, pero no otro tráfico que pertenece a otros intercambios.

IPSEC presenta muchas ventajas en comparación con TLS ya que encripta todos los paquetes completamente, puede encriptar cualquier protocolo y actúa independientemente de la dirección IP.

4.2.4. SRTP

El protocolo SRTP, *secure real-time transport protocolo*, es un perfil del protocolo RTP para proveer confidencialidad, integridad y autenticación a los paquetes de media. SRTP es considerado como uno de los protocolos estándar para proteger la media en tiempo real en aplicaciones multimedia.

Además de proteger los paquetes de media, también provee protección a los mensajes que se transmiten con el protocolo RTCP, *real-time transport control protocol*, estos mensajes son utilizados principalmente para proveer retroalimentación de la calidad del servicio a los participantes de una sesión.

Los mensajes RTCP y RTP son transmitidos de forma separada y utilizando distintos protocolo y puertos lógicos. Por lo tanto, ambos tipos de mensajes deben ser protegidos durante una sesión multimedia. Si los mensajes RTCP no se protegen, un atacante puede manipular estos mensajes entre participantes de una sesión y así causar interrupción del servicio o realizar análisis de tráfico.

Los diseñadores del SRTP se enfocaron en desarrollar un protocolo que pudiera proveer protección adecuada a los flujos de media pero también que pudiera mantener las propiedades clave para soportar redes alámbricas y no alámbricas y donde puedan existir limitantes de ancho de banda o de transporte.

Dentro de las propiedades principales del protocolo SRTP se puede hacer mención del poco uso de ancho de banda que utiliza cuando se aplica, el código de implementación es relativamente conservativo en tamaño, lo cual lo hace perfecto para dispositivos que poseen memoria limitada, como los celulares. Además, provee independencia subyacente de transporte, incluyendo las capas de red y física que se puedan utilizar.

Un diseño similar posee el protocolo MIKEY, *multimedia Internet KEYing* que se utiliza para el intercambio de las llaves en conjunto con el SRTP para proveer una seguridad adecuada para las aplicaciones multimedia sobre Internet, por ejemplo, VoIP, video y conferencia.

La aplicación o elemento de red que implementa SRTP debe ser capaz de convertir los paquetes RTP en paquetes SRTP antes de que sean transmitidos a lo largo de la red de transporte. El mismo proceso es utilizado en forma inversa para desencriptar los paquetes SRTP.

Luego de que la media es capturada esta se codifica utilizando el estándar de codificación de audio negociado o por defecto, seguido de esto es necesario los paquetes RTP son encriptados utilizando el algoritmo de encriptación negociado. El algoritmo por defecto de encriptación que utiliza SRTP es AES, *advanced encryption standard*, que permite el procesamiento de los paquetes aun cuando estos son recibidos fuera de orden, que es deseable para aplicaciones en tiempo real.

Adicional a proveer encriptación de información, el estándar SRTP soporta autenticación da mensajes e integridad de los paquetes RTP. El código del mensaje de autenticación es producido incluyendo todo el mensaje RTP.

El formato del paquete SRTP es similar al del paquete RTP, con la excepción de que contiene dos encabezados adicionales y por supuesto que el contenido o información esta encriptada. Uno de los encabezados adicionales es el MKI, *master key identifier* que es utilizado por protocolo MIKEY para identificar la llave principal o para generar una nueva, de esta llave se derivan las llaves que usarán los elementos de red para desencriptar o para validar la autenticidad de la información asociada a un paquete SRTP.

El otro encabezado adicional es el de autenticación, es importante su uso ya que provee protección en contra de ataques de respuesta de mensajes. En soluciones VoIP, es recomendado que los mensajes de autenticación sean utilizados al mínimo si la encriptación no es opcional.

Todos los encabezados de un paquete SRTP son enviados sin encriptar, a excepción del contenido de media. Ya que el protocolo SRTP utiliza AES por defecto, este provee protección en contra de ataques de denegación de servicio. Estos ataques buscan corromper la media encriptada normalmente los flujos de media cifrados dependen de la desencriptación del bloque anterior para poder desencriptar el bloque siguiente. El algoritmo AES no tiene esta limitación porque puede desencriptar cada bloque si tener el conocimiento de ningún bloque anterior.

El uso de autenticación e integridad en los mensajes SRTP es una manera importante de protección en contra de ataques. Por ejemplo, un atacante podría modificar los mensajes SRTP para corromper el contenido de audio o video del paquete. Otro ataque podría ser el envío de mensajes SRTP falsos a los participantes de una sesión, de esta forma forzando a los dispositivos a intentar desencriptar estos mensajes, con esto afectando el desempeño de la sesión actual. Debido a lo anterior es recomendable que en implementaciones VoIP se utilice SRTP para proteger al sistema de este tipo de ataques.

4.3. Monitoreo de tráfico de señalización para la detección de fraudes

Recopilar la actividad de los suscriptores es un paso indispensable para la detección de fraudes, es poco práctico e imposible analizar todas las llamadas todo el tiempo en busca de fraudes. Un enfoque común es reducir todos los datos de llamadas en estadísticas que se generan por periodos.

Para esto es indispensable que los equipos que proveen los servicios de telefonía sea capaces de generar estadísticas de utilización, para que las mismas puedan ser analizadas para detectar distintos tipos de problemas en la red, incluyendo fraudes. Por lo general, estos equipos están integrados a un

NMS, *network management system*, que es un gestor para administrar los nodos de una red y a su vez de recopilar los datos o mediciones de tráfico de los distintos elementos de una red.

4.3.1. Análisis de tráfico nacional

Para que exista comunicación entre distintos operadores de telefonía dentro de un país, se requiere que exista un punto de interconexión físico y lógico para el intercambio de tráfico telefónico.

Es necesario, entonces, monitorear periódicamente el tráfico de entrada y salida sobre estas rutas para establecer patrones de comportamiento normal. Por ejemplo, un incremento abrupto en la completación de llamadas en alguna ruta de interconexión en un periodo determinado, podría significar que se puede estar utilizando esa ruta para completar el tráfico fraudulento.

La detección inicial del comportamiento anómalo puede ser por medio del análisis de las estadísticas generadas por el equipo. Ahora bien, para poder analizar más a fondo si se trata de un fraude o no, es necesario realizar una captura de paquetes sobre las interfaces físicas de señalización, para poder verificar si existe algún tipo de patrón en marcación desde o hacia ciertos destinos.

De encontrar algún patrón se debe proceder a bloquear las llamadas entrantes a los números con comportamiento anómalo y también se debe proceder a la revisión de las políticas de completación de llamadas. Por ejemplo, si una ruta es de tráfico nacional no se debe permitir que el número que marca o el número que recibe una llamada sean con formato internacional con un código país distinto al nacional. De esta forma se previene que el tráfico

indeseado salga o entre hacia una central de telefonía por una ruta no adecuada.

Otro punto a validar sobre las rutas nacionales son los prefijos numéricos que son permitidos sobre las rutas de interconexión, es necesario establecer una restricción para permitir pasar únicamente los prefijos de otros operadores y así evitar recibir tráfico fraudulento.

4.3.2. Análisis de tráfico internacional

Los fraudes que más perdidas generan a los operadores de telefonía a nivel mundial están relacionados con completación de o hacia destinos internacionales, como los fraudes IRSF y el *bypass* internacional. Ambos métodos de fraude se logran perpetrar por medio de la inyección de tráfico hacia o desde destinos internacionales, ya sea con el fin de completar llamadas hacia destinos de alto costo o completar llamadas internacionales como nacionales hacia algunos destinos.

La interconexión de las rutas internacionales de un operador se realiza hacia los *carriers* comúnmente a través de Internet, no siempre se utiliza IPSEC para encriptar el tráfico entre las interconexiones y aunque se utilizará, la mayoría de casos el tráfico fraudulento es inyectado por otros *carriers,* por lo cual estas rutas son las que más están expuestas a ser utilizadas para realizar fraudes.

Se deben monitorear las estadísticas de tráfico o KPI en estas rutas en busca de picos o comportamiento anómalos, por ejemplo, en la estadística de completación o caídas del ASR (*answer rate*), en las causas de liberación

globales, el tráfico global sobre las rutas SIP, la duración de las llamadas, etc. En dado caso se encuentre algún comportamiento extraño, sería necesario capturar el tráfico de señalización filtrándolo por las rutas donde se observa la anomalía. Y así pode identificar de qué destino provienen las llamadas para validar si no es tráfico fraudulento.

Si se detecta tráfico fraudulento se puede bloquear analizando los prefijos que generan las llamadas o solicitando al *carrier* internacional que no envíe ese tráfico. El bloque se debe realizar en las centrales de interconexión o GMSC.

Para evitar recibir tráfico de interconexiones no autorizadas es necesario definir correctamente las listas de acceso los equipos de interconexión, normalmente se utiliza uno o varios SBC, ese problema se puede detectar analizando la captura de tráfico de señalización.

4.4. Uso de CDR para la detección de fraudes

Los fraudes sobre los sistemas de telefónica evolucionan día a día, cada vez es más difícil la detección de los mismos a través del análisis de estadísticas de tráfico de señalización ya que los atacantes utilizan distintos métodos para hacer pasar el tráfico fraudulento como tráfico habitual de un operador para evitar ser detectados.

El bloqueo de los números que generan el tráfico fraudulento se vuelve sin sentido, ya que los números están siendo cambiados aleatoriamente y sin una secuencia correlativa por los atacantes.

Un CDR, *call detail record*, se genera en las centrales de telefonía cada vez que un usuario registrado en el VLR utiliza algún servicio, por ejemplo, una

llamada, envío de SMS, servicio USSD, etc. Los CDR se utilizan para poder realizar el cobro de estos servicios a los usuarios.

Dentro de los CDR se puede encontrar detalles de la llamada como duración de la llamada, número de a y b, hora de inicio y finalización, ubicación y ruta donde se generó, entre otros.

4.4.1. Análisis de los CDR

Los operadores de servicios de telefonía que dependen de otros proveedores para la detección de fraudes, están en riesgo por no conocer y entender la complejidad de los ataques, agregando que se pierden del conocimiento invaluable que podrían ganar sobre el comportamiento de su red.

Sin embargo, las compañías que proveen servicios de análisis contra fraudes son necesarias para la generación de llamadas de pruebas y para proveer el ambiente para identificar la ubicación física de los atacantes.

Una de las habilidades que cualquier operador puede adquirir para identificar fraudes sobre su red es el análisis de CDR. La información que esta detallada es poderosa si se estudia cuidadosamente y a tiempo.

Analizando los CDR se pueden identificar volúmenes de llamadas altos desde la misma MSISDN, volumen alto de llamadas desde la misma celda, promedio de número de llamadas salientes y entrantes por rutas de señalización, promedio de llamada a través de las rutas nacionales e internacionales, números que tienen una secuencia de llamadas, entre otros.

Sin embargo, estas técnicas por si solas pueden ser obsoletas y podrían producir muchos resultados como fraudes cuando no lo son.

La tendencia de cada tipo de fraude puede variar en cada país o en cada operador, incluso los patrones de números detectados en una celda podrían no ser los mismos en otro sitio debido a la diferencia de configuración de los equipos. Debido a todas estas variantes es necesario realizar un análisis profundo de toda la información contenida en los CDR.

4.4.2. Procesamiento de los CDR

Para poder detectar los distintos tipos de fraudes, es necesario procesar en tiempo real toda la información sobre las transacciones que realizan los subscriptores, contenidas en los CDR. Con esta información se puede construir una base de datos, utilizando un software y hardware poderosos, ya que la cantidad de transacciones que se generan y las consultas que se deben realizar, pueden incluir millones de datos a la vez.

Para detectar patrones dentro la base datos, se puede empezar analizando la cantidad de llamadas dentro de periodos de una o dos horas, el número de llamadas hacia números de b únicos, número de llamadas en secuencia dentro de un periodo determinado, número de llamadas desde la misma celda, etc.

El análisis de los CDR debe ir acompañado de múltiples pruebas desde diferentes redes utilizando un generador de llamadas, de esta forma se puede probar y mejorar la detección de fraudes en la red.

La automatización del proceso de detección y bloqueo de llamadas que se detectan como fraude es un componente clave para combatir a los atacantes, la intención es tratar de frustrar dichos ataques en un tiempo relativamente corto.

Para combatir las cajas SIM, es necesario establecer patrones de utilización del servicio móvil por usuario para poder realizar esto; primero se debe filtrar que usuarios serán los que estarán bajo análisis, empezando por los usuarios nuevos que se integran en la red, ya que los usuarios con un historial largo difícilmente se volverán defraudadores.

Para la detección de tarjetas SIM que pueden estar siendo utilizadas dentro de una caja SIM, se debe analizar la frecuencia de distribución de IMSI por IMEI. Habitualmente la mayoría de los usuarios de una red móvil, utilizan solo una terminal para una o pocas tarjetas SIM.

Otro valor a comparar es el número de llamadas originadas contra el número de llamadas recibidas, ya que una SIM que está siendo utilizada para este fraude recibe miles de llamadas y genera el mismo número de llamadas que normalmente son de larga duración. Un usuario normal genera y recibe el mismo número de llamadas, pero valores menores. Las distintas celdas desde las cuales una SIM genera llamadas también deben de ser analizadas ya que por lo general las llamadas que generan cajas SIM se realizan desde las mismas celdas, pero un usuario normal realizada llamadas desde distintas celdas.

CONCLUSIONES

1. Con la integración de las redes de telefonía a Internet se ha abierto una brecha enorme a través de la cual se pueden realizar fraudes telefónicos de distintos tipos desde cualquier parte del mundo sin la necesidad de que los atacantes tengan acceso directo a las redes donde perpetran los fraudes.

2. Los fraudes telefónicos que más pérdidas generan directamente a la industria de telefonía están relacionados con el acceso ilegal a centrales telefónicas PBX para la inyección de tráfico de voz hacia distintos destinos internacionales.

3. La mejor forma de reducir el acceso ilegal a cualquier elemento de red accesible desde Internet es el fortalecimiento de las políticas de seguridad y la protección de la información sensible de dichos elementos por parte de los operadores que ofrecen servicios de telefonía.

4. Es necesaria la implementación de IPSEC y SRPT para encriptar la señalización y la media en las troncales SIP que se conectan a través de Internet, para evitar que un atacante MITM pueda obtener información útil para realizar algún fraude telefónico o perpetuar un ataque para la inhabilitación del servicio.

5. Los ataques para la inhabilitación de servicio en las redes de telefonía pueden ser detectados por medio del análisis en tiempo real de las estadísticas de tráfico de todos los nodos que componen la red.

6. El método más efectivo para la detección de fraudes telefónicos es por medio del análisis de los CDR en tiempo real, con la información contenida en estos se pueden realizar modelos para la detección de los distintos tipos de fraudes.

7. Es imposible evitar en su totalidad los fraudes en las redes de telefonía, ya que los métodos que utilizan los defraudadores evolucionan constantemente para evitar la detección.

RECOMENDACIONES

1. Se deben establecer, actualizar e inculcar constantemente las políticas de seguridad para la protección de información sensible: ser usuarios, contraseñas, direcciones IP de gestión y servicio y puntos de acceso a los equipos, tanto al personal que labora para el proveedor de servicios de telefonía; también hacia los usuarios que utilizan los servicios para evitar el acceso no autorizado a cualquier elemento de la red de telefonía.

2. Es necesario considerar el uso de IPSEC en todas las troncales SIP que se conectan hacia los *carriers* internacionales ya que estas conexiones por lo general se realizan a través de redes públicas inseguras.

3. Antes de implementar un sistema para la detección y bloqueo automático de usuarios que son utilizados con fines fraudulentos, se debe realizar un estudio para establecer modelos de comportamiento para discernir este tipo de usuarios de los usuarios normales.

4. Se debe considerar en el dimensionamiento de la capacidad del sistema para la detección de fraudes en tiempo real por medio de los CDR, la cantidad de usuarios y rutas a ser analizados para la detección de fraudes.

5. Debido a que los métodos que utilizan los defraudadores evolucionan constantemente para evitar ser detectados, es necesario la constante actualización de los sistemas de detección y bloqueo, así como la

constante capacitación del personal encargado del departamento de fraudes.

BIBLIOGRAFÍA

1. ARXIV.ORG. *VoIP technology security issues analysis* [En línea]. <https: //arxiv.org/ftp/arxiv/papers/1312/1312.2225.pdf>. [Consulta: 10 de febrero de 2016].

2. BSWAN.ORG. *Nine Simple strategies for protecting an operator or MVNO from telecom fraud.* [En línea]. <http://bswan.org/ nine_strategies.asp>. [Consulta: 18 de mayo de 2016].

3. European Telecommunications Standards Institute, (ETSI). *TS 23.002: Digital cellular telecommunications system Network architecture.* [en línea]. http://www.etsi.org/deliver/etsi_ts/ 123000_123099/ 123002/09.02.00_60/ts_123002v090200p.pdf. [Consulta: 22 de enero de 2015].

4. Internet Engineering Task Force (IETF). *RFC 3261: Session initiation protocol.* [En línea]. <https://www.ietf.org/rfc/rfc3261.txt>. [Consulta: 10 de julio de 2015].

5. Internet Engineering Task Force (IETF). *RFC 3711: The secure real-time transport protocol.* [En línea]. <https://www.ietf.org/rfc/rfc3711.txt>. [Consulta: 9 de marzo de 2016].

6. Internet Engineering Task Force (IETF). *RFC 4301: Security architecture for the internet protocol.* [En línea]. <https://tools.ietf.org/ html/rfc4301>. [Consulta: 12 de febrero de 2016].

7. KAU.DIVA-PORTAL.ORG. *Unwanted traffic and information disclosure in voip networks.* [En línea]. <http://kau.diva-portal.org/smash/ get/ diva2:529268/FULLTEXT01.pdf>. [Consulta: 10 de mayo de 2016].

8. LINK.SPRINGER.COM. *VoIP-aware network attack detection based on statistics and behavior of SIP traffic.* [En línea]. <http://link.springer. com/article/10.1007/s12083-014-0289-8>. [Consulta: 16 de mayo de 2016].

9. SEARCHUNIFIEDCOMMUNICATIONS.TECHTARGET.COM. *Security in a SIP network.* [en línea]. <http://searchunifiedcommunications. te chtarget.com/feature/Security-in-a-SIP-network-Identifying-net wo rk-attacks>. [Consulta: 20 de enero de 2016].