Seguretat en serveis de veu sobre IP i missatgeria instantània

Antoni Martínez Ballesté

PID_00191683

Universitat Oberta de Catalunya

Índex

Introducció

Internet, a part d'esdevenir una font immensa de cerca i compartició d'informació, és un entorn que permet la comunicació entre usuaris. Tant és així, que el correu electrònic ha acabat essent una forma habitual de comunicació i enviament d'informació entre persones i ha relegat el correu postal a propòsits més concrets o que no tenen implementació possible sobre un sistema telemàtic. A més, la tecnologia IP (i per extensió Internet) s'ha fet servir per a implantar aplicacions de comunicació síncrona, o en temps real. Des de les primeres aplicacions de xat o missatgeria instantània, fins a la integració de la videoconferència en les xarxes socials i en els telèfons mòbils, han anat sorgint aplicacions de tota mena que enriqueixen les possibilitats de comunicació entre persones i abarateixen costos destinats a la tramesa d'informació.

Si bé aquests sistemes basats en IP i Internet han suposat moltes millores en la productivitat o en la comunicació interpersonal, el cert és que també representen un focus d'atacs informàtics. Cal tenir present que, per mitjà de la missatgeria instantània o veu sobre IP, es transmet informació que pot ser considerada privada i sensible: per exemple, es poden tancar acords comercials, es pot proporcionar informació sobre un projecte futur, es pot comprar tot interactuant amb sistemes automàtics de reconeixement de veu, etc. Així doncs, és convenient comprendre quins són els problemes de seguretat als quals aquestes tecnologies s'exposen, i també tenir en compte quines mesures i tecnologies es poden utilitzar per a mitigar els efectes d'atacs potencials.

Més enllà de protegir els sistemes i les xarxes que sustenten aquests sistemes de comunicació, serà important conèixer com protegir la informació que hi circula i, també, vetllar per l'autenticació correcta de les parts que intervenen en les comunicacions.

Aquest mòdul descriu els problemes de seguretat informàtica que poden presentar els sistemes de comunicacions síncrones, en concret la veu sobre IP i la missatgeria instantània. També recull el seguit de tècniques que convé aplicar per a protegir els sistemes i la informació que hi circula.

Per a la comprensió dels conceptes d'aquest mòdul, és convenient que els estudiants tinguin coneixements fonamentals de les xarxes de computadors, concretament de tot allò que envolta la comunicació amb TCP/IP. També és important que hagin estudiat amb anterioritat les tècniques criptogràfiques i els productes de seguretat àmpliament usats en la transmissió segura de la informació.

Objectius

Els objectius que han d'haver assolit els estudiants un cop finalitzat el mòdul són els següents:

1. Comprendre com funcionen els serveis de veu sobre IP i de missatgeria instantània, centrant-se en els elements de l'arquitectura i els protocols que els fan possibles.

2. Conèixer com afecten els atacs sobre els equips i les xarxes al bon funcionament d'aquests serveis.

3. Avaluar l'impacte dels atacs sobre la confidencialitat de la informació que poden experimentar aquests sistemes.

4. Avaluar l'impacte dels atacs sobre l'autenticació de la informació i els participants que poden experimentar aquests sistemes.

5. Aplicar eines de seguretat als protocols i components de l'arquitectura d'aquests serveis per tal d'evitar problemes de seguretat o minimitzar-ne l'impacte en cas que es produeixin.

1. Serveis de comunicació síncrona

En aquest apartat introduïm dos entorns, dels quals analitzarem els problemes de seguretat: la veu sobre IP i la missatgeria instantània. Cal dir que altres sistemes de comunicació síncrona com ara la videoconferència poden experimentar atacs similars, i es poden usar les tècniques descrites en aquest mòdul per a mitigar-los. Així doncs, amb la intenció de no estendre massa el mòdul amb conceptes repetitius, només ens centrem en aquests dos entorns.

Cronològicament parlant, la primera aplicació de comunicació síncrona és la missatgeria instantània. L'origen es podria remuntar a la instrucció *talk* disponible als sistemes Unix, amb la qual un usuari del sistema podia enviar missatges curts de text a un altre usuari del mateix sistema de manera immediata. Aquest servei ha anat evolucionant i donant desenes d'aplicacions diferents.

La missatgeria instantània (IM) consisteix en la comunicació entre dos o més usuaris fent servir missatges curts de text que s'envien en temps real.

IM

Les sigles IM, que denominen la missatgeria instantània, vénen de l'anglès *instant messaging*.

La veu sobre IP (VoIP) permet aprofitar la infraestructura de xarxa de dades per a l'establiment de converses telefòniques, tant punt a punt com en grup. Aquesta tecnologia permet l'estalvi en recursos TIC, ja que no hi ha una xarxa telefònica paral·lela que s'hagi de mantenir. En la mateixa línia, per a unes noves instal·lacions, només cal pensar en un únic cablejat de dades que també servirà per a la veu.

La veu sobre IP (VoIP) és una tecnologia que permet l'establiment de trucades telefòniques fent servir els datagrames IP com a mitjà de transport.

VoIP

L'abreviatura VoIP, que denomina la veu sobre IP, ve de l'anglès *voice over the Internet protocol*.

En el fons, hi ha diferents variants que es poden englobar dins la VoIP. D'una banda, quan les trucades es poden establir cap a la xarxa telefònica convencional, s'anomena *telefonia IP*. De l'altra, si la trucada es transmet més enllà de la xarxa interna, és a dir, es transmet per Internet, parlem de veu sobre Internet o telefonia sobre Internet. D'aquesta manera, l'ús d'IP per al trànsit de veu dins una xarxa tancada i controlada (per exemple, una LAN) es correspondria purament amb la VoIP. De totes maneres, s'utilitza VoIP per a referir-se a totes aquestes variants i, a més, els protocols emprats no són diferents.

Cisco i Skype

Com a exemple, l'empresa Cisco desenvolupa solucions de VoIP per a grans corporacions. D'altra banda, l'aplicació Skype ha esdevingut un estàndard en la telefonia i la videoconferència sobre Internet (fins i tot la seva tecnologia és utilitzada per altres proveïdors de serveis d'Internet).

La integració d'aquestes tecnologies en el web ha permès que la missatgeria instantània, la veu sobre Internet i la videoconferència es puguin gestionar des d'un navegador sense la necessitat d'instal·lar programari específic. Aquest seria el cas, per exemple, de les eines de Google i Facebook per a comunicacions síncrones.

Finalment, és conegut que algunes eines d'IM han incorporat tecnologies de veu sobre Internet (com és el cas de Messenger) o, per contra, algunes aplicacions de veu sobre Internet també incorporen IM. També cal afegir la possibilitat de la videoconferència en molts d'aquests sistemes populars, com ara l'esmentat Skype.

En aquest mòdul estudiem la seguretat relacionada amb la VoIP i la IM i, per fer-ho, començarem en primer lloc per veure com funcionen aquestes dues tecnologies.

1.1. Funcionament de la veu sobre IP

En aquest subapartat introduirem la tecnologia sobre la qual es desenvolupa la VoIP. D'aquesta manera, podrem entendre els problemes de seguretat i comprendre com es poden evitar o solucionar. En primer lloc, descriurem l'arquitectura d'un sistema VoIP en una xarxa, en dispositius i en servidors. Després, introduirem els protocols que s'usen en aquesta tecnologia: els proposats per l'UIT i els de la IETF.

UIT

UIT és la Unió Internacional de Telecomunicacions, en anglès International Telecomunications Union. És la responsable de definir molts dels protocols usats en telefonia fixa i mòbil, sistemes de vídeo conferència, cablejats, etc.

1.1.1. Arquitectura

Un sistema VoIP funciona sobre una xarxa IP. En aquesta xarxa, s'estableix la comunicació entre dos o més agents d'usuari per mitjà de l'enviament de paquets de senyalització, els quals descriuen qui és el destinatari, especifiquen quina és la localització (adreça IP) de l'originador de la trucada, envien senyals de trucada a l'agent d'usuari del destinatari, etc. Un cop s'ha establert la comunicació, s'inicia un enviament de paquets de veu per a transmetre la conversa entre els participants. En general la senyalització es pot establir sobre un protocol de transport orientat a connexions (és a dir, TCP), mentre que el transport de veu s'esdevé d'usuari a usuari per mitjà de datagrames i UDP.

IETF

La Internet Engineering Task Force (IETF) és un dels organismes més influents pel que fa a estàndards d'Internet. Els estàndards es publiquen en uns documents anomenats RFC (Request For Comments).

El més habitual avui en dia és que el sistema funcioni sobre una xarxa d'àrea local (LAN) o bé amb una comunicació entre dos o més usuaris connectats a Internet. Així doncs, per posar un exemple de LAN, suposarem un entorn d'oficines. En les instal·lacions, hi ha un commutador Fast Ethernet que centralitza les connexions de la xarxa de dades formada per diversos ordinadors que usen IP connectats a aquest commutador. També suposem que hi ha un encaminador que connecta la LAN a Internet.

Els elements que formen part de l'arquitectura VoIP són els telèfons IP, les centraletes IP i les passarel·les de telefonia. Addicionalment, hi pot haver altres servidors complementaris.

Els telèfons IP són els nodes des dels quals sovint s'inicien o es reben les comunicacions (és a dir, agents d'usuari). Hi ha diferents variants d'aquests telèfons: d'una banda, hi ha telèfons semblants als de la telefonia convencional però que realment usen VoIP; de l'altra, hi ha telèfons de programari que estan instal·lats als ordinadors i, per tant, necessiten un micròfon i una sortida d'àudio per a poder fer efectives les trucades. Addicionalment, es poden usar adaptadors a VoIP endollats a la línia de telèfons convencionals, o bé telèfons USB que, en el fons, són dispositius que interactuen amb un programari instal·lat a l'ordinador. Cada telèfon IP té un número d'extensió que l'identifica i que es fa servir a l'hora d'especificar el destinatari de la trucada. Alguns sistemes de VoIP permeten especificar el destinatari de la trucada fent servir el seu nom d'usuari, com si es tractés d'una adreça de correu electrònic.

Telèfons de programari

Els telèfons de programari reben, en anglès, el nom de *softphones*.

TCP

Recordeu que TCP (*transmission control protocol*) estableix una sessió entre els comunicants, de manera que el protocol s'encarrega que tots els datagrames arribin correctament i s'entreguin a l'aplicació en el mateix ordre en què l'aplicació origen els ha emès.

Figura 1. A l'esquerra, telèfon IP; a la dreta, un telèfon IP de programari

En els telèfons IP té lloc una de les etapes més importants de la VoIP: la digitalització i compressió o descompressió de la veu. Per a dur a terme aquestes tasques, es poden fer servir diferents còdecs, en funció dels requisits d'amplada de banda, qualitat de veu i retards màxims. En aquest sentit, un còdec que mantingui una bona qualitat de veu per a poca amplada de banda, sol implicar la introducció d'un retard en la comunicació. Com s'ha apuntat anteriorment, la veu comprimida es transmet en paquets IP que van sobre UDP, ja que l'ús de TCP (que garanteix el lliurament en ordre de tots els paquets emesos) introduiria retards que podrien arribar a degradar en excés la qualitat de servei de la trucada.

UDP

Recordeu que UDP (*user datagram protocol*) és l'enviament de datagrames de manera independent, sense establir una connexió entre els comunicants.

La centraleta de VoIP, o IP-PBX[1], és un ordinador amb un programari que fa les funcions de centraleta telefònica. Les centraletes són un element clau a l'hora d'establir i gestionar les trucades. Per qüestions de disponibilitat i eficiència, hi

(1)Abreviatura en anglès de *IP private branch exchange*

pot haver diverses centraletes actuant de manera coordinada. Una altra opció és que la centraleta sigui un dispositiu externament similar a un commutador o encaminador, més que un ordinador on corre un programari.

Veu IP amb molts usuaris

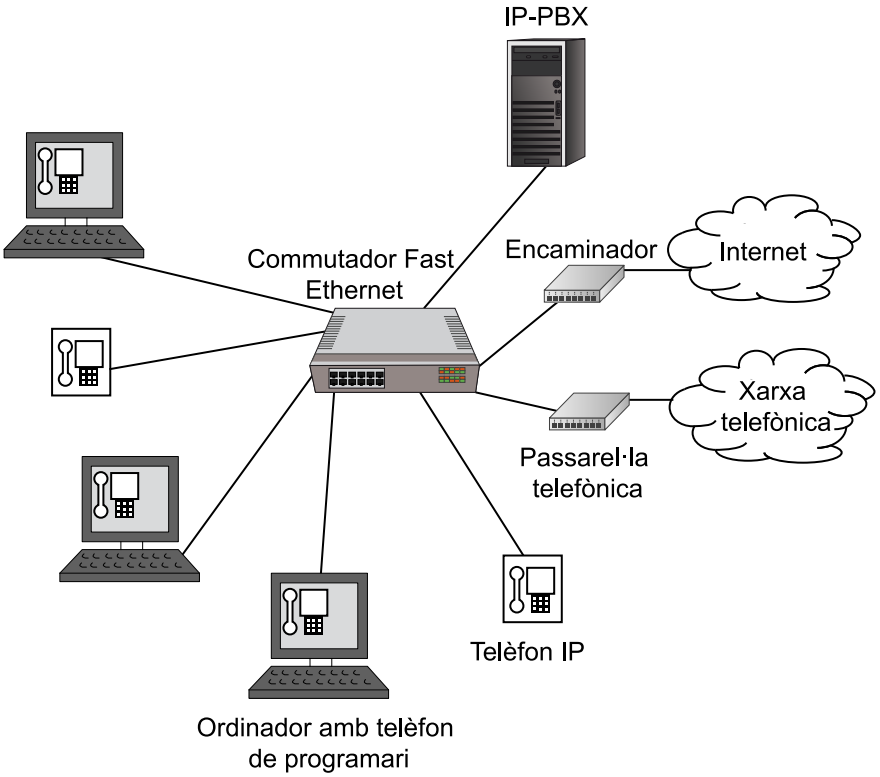
Un cas il·lustratiu extrem de diverses centraletes actuant de manera coordinada seria el d'un sistema de veu sobre Internet amb centenars de milers d'usuaris, el qual requeriria centenars de servidors distribuïts geogràficament.

Finalment, la passarel·la telefònica o *gateway* és un dispositiu que connecta la xarxa de dades per on circula el trànsit de VoIP amb la xarxa telefònica convencional, ja sigui a través de línies analògiques o a través de línies XDSI[2].

[2]XDSI significa *xarxa digital de serveis integrats*.

A banda dels elements anteriors, en un sistema de VoIP hi sol haver altres servidors. Per exemple, servidors amb funcionalitat de bústia de veu, servidors de comptabilitat i facturació, servidors de directori i control d'agents d'usuari (*gatekeepers*), servidors de configuració, etc. Tot i no ser obligatoris, solen ser el focus d'alguns atacs. La figura 2 mostra un escenari d'exemple amb tots aquests elements.

Figura 2. Escenari d'exemple d'una xarxa amb VoIP



Un cop definits els elements que formen una xarxa de VoIP, introduirem les tendències principals pel que fa a protocols que permeten l'establiment i el control de trucades.

1.1.2. Protocols UIT

En primer lloc, introduïm els protocols definits per la UIT com a estàndards per a la VoIP, aprofitant alguns dels estàndards que s'usen àmpliament en la telefonia convencional.

La UIT defineix, dins l'estàndard H.323, un conjunt de protocols per a l'establiment i el control de trucades a través de la VoIP.

Per exemple, dins el protocol H.323 es defineixen els protocols següents en relació amb la senyalització i el control de la trucada:

- H.225, que defineix l'ús del protocol Q.931 en l'establiment de trucades telefòniques en l'XDSI, enviant aquests paquets Q.391 sobre connexions TCP en comptes de fer-ho sobre el canal de control de l'XDSI. També inclou les especificacions per al registre i l'admissió d'agents d'usuari en les trucades (RAS, *registration, admission and status*), fent servir UDP.

- H.245, que defineix missatges per a concretar termes relacionats amb la compressió de la veu, els ports que s'han d'emprar, capacitats dels telèfons IP, etc., sobre la connexió TCP que ha iniciat l'H.225/Q.931.

El protocol H.323 també especifica que, per al transport multimèdia de la veu, cal utilitzar una capa RTP (protocol definit per la IETF) per a inserir als paquets de veu informació de seqüència i de temps real, la qual és essencial per a garantir una bona comunicació. Com veurem més endavant, alterar aquesta informació és un dels atacs més freqüents en la VoIP. Addicionalment, la UIT defineix el protocol H.248 per a la conversió entre VoIP i la xarxa telefònica convencional (tindria relació, doncs, amb les passarel·les telefòniques).

RTP

RTP significa *real-time transport protocol*. Està definit en l'RFC 3550.

En el protocol H.323 els paquets es codifiquen en forma binària amb la tècnica PER[3], per tal de transmetre'ls eficientment en les xarxes de comunicacions.

[3]PER significa *packet encoding rules*.

Tot i que la UIT es considera el referent quant a estàndards per a telefonia, el cert és que la tendència la marquen altres alternatives. D'una banda, alguns dels grans fabricants de tecnologia per a VoIP (per exemple, Cisco) empren protocols propis, de manera que cal utilitzar passarel·les entre arquitectures H.323 i els seus sistemes. De l'altra, moltes de les aplicacions de VoIP basades en telèfons de programari (com ara Skype, o Google Voice), s'han decantat per l'ús del protocol SIP, que introduïm a continuació. Els protocols H.323 es recullen en la taula 1.

Taula 1. Alternativa H.323 per a VoIP

<table>
<tr><th colspan="2">Control i establiment de trucada</th><th>Registre i admissió</th><th>Control de l'estat de la xarxa</th><th>Veu comprimida, G.711, G.729,...</th></tr>
<tr><td>H.245</td><td>H.225/Q.931</td><td>H.225/RAS</td><td>RTCP</td><td>RTP</td></tr>
<tr><td colspan="3">TCP</td><td colspan="2">UDP</td></tr>
<tr><td colspan="5">IP</td></tr>
</table>

1.1.3. Protocol IETF

El protocol SIP ha acabat esdevenint l'estàndard de facto en aplicacions de VoIP i també en IM.

SIP

SIP significa *protocol d'iniciació de sessions* (en anglès, *session initiation protocol*). Està definit en les RFC 2543 i 3261.

El protocol SIP es va dissenyar per a poder establir sessions multimèdia sobre una xarxa IP amb l'objectiu que servís per a un espectre d'aplicacions ampli.

A diferència del protocol H.323, SIP fa servir missatges de text en les peticions i respostes (senyalització i control de trucades), d'una manera similar al que fan FTP i HTTP. Els identificadors d'usuari dels sistemes VoIP usen l'arrova, igual que els identificadors (o adreces) de correu electrònic.

Les funcionalitats de SIP són ben diverses, i per a implementar-les es fan servir diferents peticions o respostes. Les funcionalitats que recull SIP són les següents:

- Ubicació d'usuaris, que permet conèixer la localització d'un agent d'usuari o d'un usuari dins una xarxa de VoIP.
- Disponibilitat d'usuaris, que permet saber si l'usuari es troba disponible.
- Capacitats d'usuari, per a poder saber quins paràmetres convé usar per a la trucada (ample de banda, còdecs acceptats, etc.).
- Establiment de sessió, per a establir la trucada en el cas de VoIP, o la conversa en cas d'IM.
- Gestió de la sessió, per a modificar els paràmetres durant la connexió, afegir usuaris, acabar la trucada, etc.

Per a implementar aquestes funcionalitats se solen usar protocols auxiliars com ara l'RTP, l'RTSP[4], l'MGCP[5] i l'SDP[6]. El primer, tal com hem vist per al cas del protocol H.323, afegeix informació sobre temps real als paquets d'informació. L'RTSP serveix per a controlar l'enviament d'informació multimèdia (per exemple, per a aturar l'enviament de vídeo en una conferència de VoIP). L'MGCP és el protocol encarregat de connectar un sistema SIP amb la xarxa telefònica convencional per mitjà de la passarel·la, tal com ho fa el protocol H.248. Finalment, l'SDP permet descriure la informació necessària per a iniciar una sessió (per exemple, la localització del destinatari, l'adreça de qui inicia la conversa, ports que s'usaran, etc.).

(4)RTSP significa *real-time streaming protocol*. El defineix l'RFC 2326.

(5)MGCP significa *media gateway control protocol*. El defineix l'RFC 3435.

(6)SDP significa *session description protocol*. El defineix l'RFC 4566.

En un escenari basat en SIP, la centraleta (o altres servidors complementaris) poden fer les funcionalitats de registre d'usuaris, intermediaris de trucada, etc. donant un ventall ampli de possibilitats a l'hora d'implantar solucions centralitzades, distribuïdes, amb possibilitats de balanceig de càrrega, etc.

Els missatges més habituals en una sessió SIP són el REGISTER i l'INVITE:

- Quan un usuari es valida al sistema, envia un missatge REGISTER al servidor de registres (en anglès, *registrar server*).

- Quan un usuari vol iniciar una conversa, envia un missatge INVITE a un dels servidors intermediaris de trucada.

A continuació, mostrem un exemple de missatge INVITE, on podem apreciar que s'usen identificadors similars als del correu electrònic (tipus URI[7]), tant per als usuaris com per als identificadors de trucada.

(7)URI significa *uniform resource identifier*.

```
INVITE sip:toni@voip.uoc.edu SIP/2.0
Via: SIP/2.0/UDP toni.intranet.uoc.edu;branch=z9hG4bK776asdhds
Max-Forwards: 70
To: Jordi <sip:jordi@voip.uoc.edu>
From: Toni <sip:toni@voip.uoc.edu >;tag=1928301774
Call-ID: a84b4c76e66710@toni.intranet.uoc.edu
CSeq: 314159 INVITE
Contact: <sip:toni@toni.intranet.uoc.edu >
Content-Type: application/sdp
Content-Length: 142
...
```

Els protocols relacionats amb el sistema SIP es recullen en la taula 2.

Taula 2. Alternativa SIP per a VoIP

Control de fluxos multimèdia	Control i establiment de trucada (SDP)	Control de l'estat de la xarxa	Veu comprimida, G.711, G.729,...
RTSP	SIP	RTCP	RTP
TCP	UDP o TCP	UDP	
IP			

1.2. Funcionament de la missatgeria instantània

Les aplicacions d'IM permeten converses entre usuaris fent servir missatges curts de text que s'envien en temps real. En general, les converses es poden establir potencialment entre un grup de contactes, al qual es poden anar afegint usuaris: un usuari es registra al servei IM i després va afegint (o convidant) altres usuaris del servei, que passaran a formar part de la seva llista de contactes.

Des dels inicis de la popularització d'Internet (mitjan anys 1990) han anat sorgint desenes d'aplicacions d'IM. Cadascuna de les aplicacions que ha anat apareixent, ha anat definint-ne els protocols i les implementacions de mane-

ra que, en general, la interoperabilitat entre diferents aplicacions era pràcticament inviable. Tanmateix, algunes aplicacions s'han popularitzat molt en els darrers anys i han esdevingut eines quotidianes per a molts usuaris.

Com hem comentat, aquestes eines incorporen veu sobre Internet o possibilitat de transmetre vídeo, tant des d'aplicacions específiques com des del navegador web mateix. Aquests sistemes també permeten enviar fitxers (documents, executables, fotografies, etc.) o bé enllaços web.

1.2.1. Arquitectura

L'arquitectura bàsica que definim està composta per agents d'usuari (programari per a establir converses i trametre missatges) i per diferents servidors. Per al cas dels servidors, en la nostra arquitectura en definim tres tipus:

- **Servidor de despatx**. Aquest és el servidor amb el qual contacta un agent d'usuari quan es connecta al sistema.

- **Servidor de notificacions**. Aquest servidor conté informació sobre els usuaris que estan actius, és a dir, en disposició d'efectuar o rebre trucades.

- **Servidor de centraleta**. Aquest servidor centralitza les comunicacions entre dos o més participants, tot recollint els missatges que generen i enviant-los cap als altres participants.

S'entén que si el sistema té un nombre elevat d'usuaris, hi haurà diversos servidors que actuaran de manera coordinada, com en el cas que hem comentat en la VoIP.

1.2.2. Funcionament d'una sessió

Pel que fa als protocols, abunden els protocols propietari, tot i que també és habitual trobar sistemes basats en SIP. Això implica que la IM pot patir problemes similars als que pateix la VoIP, clarament en el cas que ambdues implementacions utilitzin SIP. Consegüentment, moltes de les solucions als problemes de seguretat seran comunes a VoIP i a IM.

Exemple

En el nostre escenari, suposem un sistema basat en missatges de text tipus SIP per a l'establiment de trucades i altres gestions. Aquestes operacions requeriran l'enviament de paquets de senyalització, on es descriurà quin és el contacte destinatari, quina és la seva IP, etc.

Una sessió s'esdevé amb un seguit de passos, que s'il·lustren en la figura 3. Suposem que l'usuari 1 vol iniciar una conversa amb un dels seus contactes, l'usuari 7:

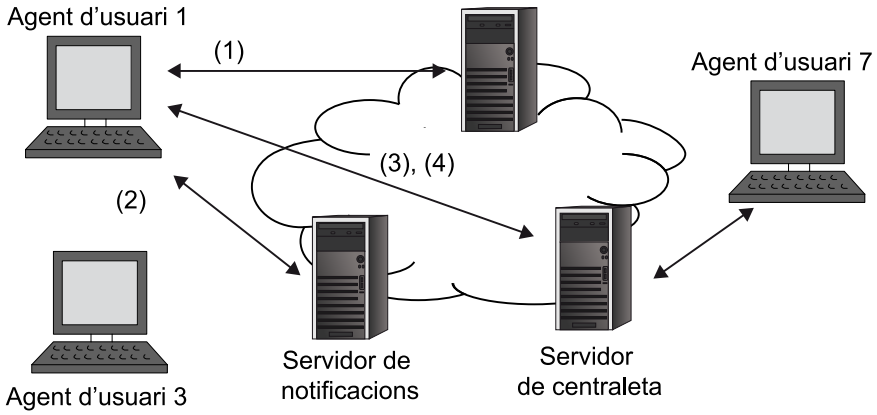
1) En primer lloc, l'usuari 1 es valida contra el servidor de despatx usant un missatge REGISTER. Si la validació és correcta, el servidor de despatx envia a l'agent d'usuari de l'usuari 1 una clau i una adreça IP on hi ha el servidor de notificacions que ha d'usar.

2) L'agent d'usuari es connecta al servidor de notificacions corresponent i, a canvi de la clau que li ha proporcionat el servidor de despatx i el seu identificador, obté la llista de contactes {usuari 2, ..., usuari 20} i quins d'aquests contactes estan actius. També se li assigna la IP del servidor de centraleta. El servidor de notificacions ha d'avisar, als agents d'usuari dels contactes que estan actius, que l'usuari 1 s'acaba d'incorporar.

3) Quan l'usuari 1 vol establir una connexió amb l'usuari 7, l'agent d'usuari del primer envia un missatge INVITE cap al servidor de centraleta. El missatge conté l'identificador del destinatari (usuari 7), amb la qual cosa el servidor de centraleta contactarà amb l'agent d'usuari de l'usuari 7 per establir la conversa.

4) Es generen missatges i es transmeten en paquets UDP.

Figura 3. Exemple d'establiment de sessió en IM



En el cas que presentem aquí, tots els missatges de text que s'envien els participants passen pel servidor de centraleta. Tot i això, alguns sistemes de missatgeria instantània aposten perquè les comunicacions entre usuaris es facin directament entre els seus agents d'usuari, és a dir, sense centralitzar-se en un servidor.

En general, de manera similar al que succeeix en VoIP, les connexions i el control de sessions s'esdevindran mitjançant una connexió TCP, mentre que el contingut (text) es podrà enviar per mitjà d'UDP. De fet, hi ha tants escenaris com implementacions possibles, i més si tenim en compte que les aplicacions de missatgeria instantània també ofereixen funcionalitats d'enviament d'arxius, per exemple. Això darrer pot comportar l'establiment de connexions TCP per a l'enviament de dades.

Un cop hem vist l'arquitectura i els protocols en què es basen dos dels serveis de comunicacions síncrones més utilitzats, és el moment d'estudiar els problemes de seguretat que presenten.

2. Denegació i degradació del servei

Els sistemes presentats funcionen amb protocols de transmissió de dades que són executats per ordinadors. Alguns dels sistemes efectuen les funcionalitats de servidor, amb tot el que això implica. Per aquests motius, els serveis de VoIP i d'IM són susceptibles de rebre atacs similars als que pot rebre qualsevol altre sistema de transmissió d'informació per mitjà de xarxes de computadors, com ara el servei web, el correu electrònic, etc.

Per a resoldre els problemes de seguretat, o intentar evitar-los, s'utilitzen tècniques que ja es fan servir en els sistemes de transmissió d'informació. Així doncs, trobarem atacs relacionats amb els equips (que n'afecten la disponibilitat) i atacs relacionats amb la seguretat de la informació (que afecten la confidencialitat i la integritat de la informació). Dins els atacs relacionats amb els equips, hi ha atacs específics contra els servidors i atacs específics contra els agents d'usuari (és a dir, telèfons VoIP i programari d'IM). Finalment, farem un cop d'ull a quines són les recomanacions de fabricants i organismes per a donar seguretat als equips.

2.1. Atacs contra els servidors

Els atacants poden col·lapsar els servidors que prenen part en un servei de VoIP o d'IM amb l'objectiu de denegar el servei als usuaris autoritzats. L'efecte d'aquests atacs va des d'una pèrdua en la qualitat de servei (comunicacions que triguen més de l'habitual a establir-se, retards i talls en les converses, etc.) fins a la impossibilitat total d'usar el servei.

Un dels mètodes habituals d'aconseguir una denegació de servei és per mitjà de la inundació de peticions cap als servidors. En general, aquests atacs se solen originar des de milers de màquines que inicien l'atac a petició de l'atacant.

Ara bé, com en tots els servidors, una de les maneres d'aconseguir que deixi de prestar el servei és per mitjà de la intrusió al sistema. L'atacant aconsegueix entrar al sistema, ja sigui explotant vulnerabilitats o bé usant enginyeria social (per exemple, trucant a l'administrador i demanant-li les credencials d'autenticació). Tot seguit, l'atacant pren el control de la màquina i atura els processos que hi presten servei.

També és possible bloquejar el programari que resideix als servidors per mitjà de l'enviament de peticions mal formades, les quals penjaran el servidor. Aquest mètode forma part de l'aprofitament de vulnerabilitats del programari.

Per tal de mitigar aquests possibles problemes, és evident que cal aplicar totes les politiques de protecció de sistemes que aplicaríem a qualsevol altre servidor. En aquest sentit, apuntem les accions següents:

- **Actualització del sistema operatiu i del programari en execució**. Els sistemes operatius presenten vulnerabilitats i, a mesura que es van descobrint, els fabricants van proporcionant "pedaços" per a corregir-los. El mateix sol succeir amb els diferents programes que hi pot haver instal·lats al servidor. Cal, per tant, prendre consciència de la importància de tenir els sistemes actualitzats.

- **Protecció dels servidors amb tallafocs**. Un tallafoc és un programari que gestiona quin trànsit pot sortir cap a la xarxa o bé entrar cap a l'equip o una altra xarxa. Els tallafocs que filtren paquets, per mitjà de la inspecció dels ports o adreces de xarxa, són més que suficients per a controlar els accessos al servidor, o bé a la xarxa on s'ubiqui el servidor. Com que en un sistema moderadament gran de VoIP hi haurà diversos servidors relacionats amb el servei VoIP per a donar cobertura a una organització, convé que aquests estiguin ubicats dins una zona desmilitaritzada[8]. Tot i això, l'ús de tallafocs per a protegir servidors de VoIP pot comportar alguns problemes que cal tenir en compte, i que detallem més endavant.

[8]El concepte de zona desmilitaritzada (DMZ, de l'anglès *demilitarized zone*) s'estudia en el mòdul "Sistema de tallafocs".

- **Instal·lació d'un sistema detector d'intrusions**. Els sistemes detectors d'intrusions (IDS[9]) analitzen l'activitat en una màquina o xarxa per tal de poder detectar comportaments anòmals per part d'usuaris i processos, amb l'objectiu de notificar que hi ha una intrusió al sistema. Un IDS és el complement idoni del tallafocs per a garantir al màxim que, si hi ha alguna intrusió que ha saltat el sistema de tallafocs, podrà ser detectat.

[9]IDS ve de l'anglès *intrusion detection system*. Aquest concepte es veu en el mòdul "Sistemes de detecció d'intrusos en xarxa".

- **Autenticació forta d'usuaris administradors**. La majoria de programari i dispositius actuals permeten la gestió des d'interfícies web on l'usuari administrador s'ha de validar. Enfortir aquesta validació per mitjà de certificats digitals ajuda a evitar que atacants entrin al sistema i trenquin els sistemes d'autenticació per contrasenya.

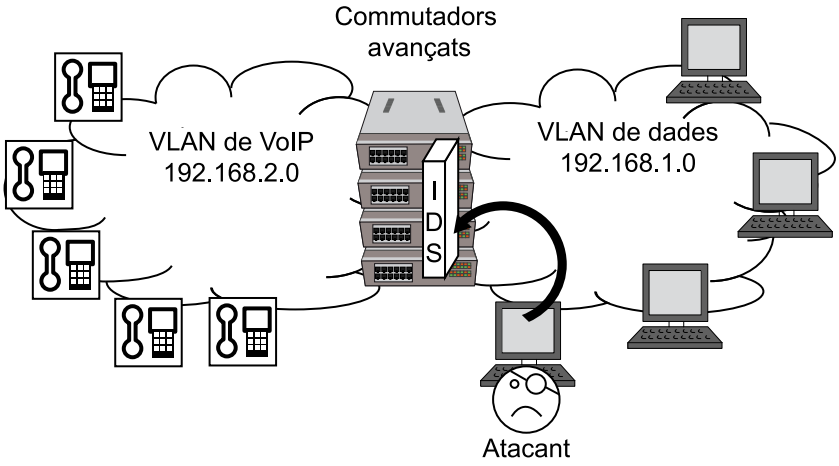
Per al cas dels servidors de VoIP dins una organització, els fabricants recomanen, si és possible, usar tecnologia VLAN per a poder definir un segment de xarxa exclusiu per a la VoIP. D'aquesta manera, el trànsit de la VoIP aniria per uns ports en concret i un atacant connectat a la VLAN de dades on es connecten els equips tindria més complicat accedir al segment reservat a la veu. Addicionalment, es recomana que en el segment de VLAN dedicat a veu resideixi un IDS, fins i tot, si és possible, implementat dins mateix del commutador,

VLAN

La tecnologia VLAN permet organitzar màquines connectades a un commutador com si estiguessin connectades a diferents xarxes, independentment de la seva connexió física a la xarxa.

per a detectar possibles intents d'intrusió. De fet, els commutadors avançats amb tecnologia VLAN solen tenir aquesta mena de característiques, entre les quals destaquem la detecció d'intrusions o la protecció envers alguns atacs de denegació de servei. La figura 4 il·lustra l'ús de VLAN i un IDS en un sistema de VoIP.

Figura 4. Ús de VLAN en un sistema de VoIP



Tot i que les propostes anteriors maximitzen el fracàs dels atacs més habituals contra els servidors, la varietat de sistemes i aplicacions és tan àmplia que sempre hi ha algun nou atac que hi podria tenir èxit.

2.2. Atacs contra els telèfons IP

Els telèfons IP també poden rebre atacs de denegació (o degradació) de servei, en el sentit que poden ser inundats amb peticions SIP, o bé amb paquets de veu que interrompin constantment el bon flux d'una conversa telefònica. Al capdavall, es tracta de nodes accessibles dins una xarxa IP. D'aquests atacs específics en trobem dues variants:

- **Enviament de paquets de finalització de trucada**. L'atacant envia cap a agents d'usuari que estiguin en una trucada un paquet de finalització de trucada que sigui conforme al protocol usat en el sistema.

- **Enviament de paquets amb informació RTP falsa**. El protocol RTP proporciona control del temps real en un flux multimèdia per mitjà de la utilització de números de seqüència (per tal que el receptor en pugui determinar l'ordre d'enviament correcte cap al descompressor), i el marcatge de paquets amb informació de l'instant de temps (es coneix com a *timestamping*). Clarament, si un atacant envia paquets de veu amb informació d'RTP falsa, el descompressor produirà un àudio ple de talls, errors i soroll.

En aquest cas, si es produeixen els atacs pot ser per dos motius. Un d'ells, que l'atacant hagi entrat al servei de VoIP i l'estigui atacant "des de dins". S'entén que l'atacant hi pot entrar si les propostes de l'apartat anterior per a evitar

atacs als servidors no han funcionat. L'altre motiu podria ser que un usuari té l'equip infectat per un codi maliciós i aquest sigui el verdader responsable d'anar enviant atacs per a degradar la qualitat del servei d'altres usuaris del servei de VoIP.

Els fabricants principals de tecnologia per a VoIP recullen, en les documentacions, un seguit de recomanacions per a dotar de seguretat el servei dels telèfons IP. Tot seguit les reproduïm:

- **Recomanacions sobre el maquinari**. Alguns fabricants de tecnologia VoIP es decanten per emprar exclusivament telèfons IP i no pas telèfons de programari. Els primers són menys susceptibles de rebre atacs i infeccions per part de programari maliciós. A més, és més probable que un usuari inexpert es descarregui qualsevol telèfon de programari que, en realitat, aprofiti per a piratejar el sistema de VoIP, ja sigui per a enviar converses a un tercer, o per a atacar els servidors de VoIP des de dins la xarxa.

- **Recomanacions sobre l'adreçament IP**. L'assignació d'adreces IP als telèfons hauria de ser manual, per a minimitzar el risc de patir atacs. Ara bé, si l'assignació només pot ser automàtica per la dimensió del sistema de VoIP, les IP només s'haurien d'assignar a adreces físiques[10] conegudes. Com que les adreces físiques es poden falsejar, els fabricants recomanen que l'usuari posi al telèfon una contrasenya per tal que l'aparell obtingui una adreça IP. Es recomana usar exclusivament adreces privades[11] per als telèfons IP. Una adreça privada correspon a una sèrie de rangs reservats amb aquest objectiu, i no es poden assignar a interfícies de xarxa connectades a la "part pública" de les xarxes. Finalment, els servidors amb informació de configuració, directori, etc. tan sols haurien de donar informació als telèfons IP que pertanyin a una llista controlada per l'administrador del servei de VoIP.

[10] Les adreces físiques es coneixen com a adreces MAC (*medium access control*).

[11] Un exemple d'adreça privada és 192.168.0.7.

2.3. Altres atacs contra veu sobre IP

Els sistemes de VoIP poden ser el focus d'atacs específics relacionats amb les característiques pròpies del servei. Tot seguit en presentem alguns:

- **Manipulació de la configuració dels agents d'usuari**. En algunes implementacions, els telèfons IP utilitzen serveis de transferència de fitxers contra un servidor de configuració per a determinar paràmetres com permisos d'usuari, obtenir entrades de directoris telefònics, o fins i tot permetre l'actualització de programari. Si un atacant té accés a aquests servidors de configuració, o bé és capaç de suplantar el servidor original, els telèfons IP es configurarien erròniament per tal que l'atacant pogués fer efectius diversos atacs.

- **Manipulació dels registres de comptabilitat**. Si un atacant pot tenir accés als servidors de comptabilitat i facturació, podrà modificar els registres corresponents al seu usuari per tal d'abaratir considerablement el cost que pugui pagar pel servei. Amb la mateixa facilitat, podria alterar els comptes d'altres usuaris i incrementar-los la factura tot afegint serveis que realment no s'han prestat.

- **Manipulació de l'equip d'usuari**. En determinats casos, quan els usuaris fan servir un telèfon basat en programari, un programa maliciós podria entrar a l'ordinador de l'usuari amb finalitats d'espia. Així com existeix programari espia que recull informació sobre contrasenyes i altres dades que pugui teclejar l'usuari per enviar-les a un atacant, es pot donar el cas d'un programari que enviï les converses de VoIP cap a un atacant.

- **Enviament de trucades publicitàries no sol·licitades**. De la mateixa manera que els qui envien correus publicitaris massius fan servir servidors SMTP víctima com a plataforma, un atacant podria entrar en un servidor d'un servei de VoIP i usar-ne els recursos per a efectuar trucades publicitàries. Aquestes trucades degradarien la qualitat del servei (per la disminució de recursos disponibles), alhora que causarien una molèstia evident als usuaris del sistema i receptors eventuals de la publicitat massiva. Els qui envien correus massius publicitaris sovint utilitzen ordinadors personals infectats amb programari maliciós amb l'objectiu que enviï també missatges d'aquesta mena. Així doncs, es podria pensar en atacar telèfons IP per a efectuar trucades publicitàries en nom d'un usuari d'un servei de VoIP.

Els dos darrers atacs també tenen l'equivalent en la IM. La solució o prevenció dels problemes anteriors de seguretat dependrà de cadascun dels escenaris concrets. Tanmateix, en gran mesura, la protecció de xarxes, servidors i agents d'usuari en un servei haurien de minimitzar les possibilitats que els atacs anteriors tinguin èxit.

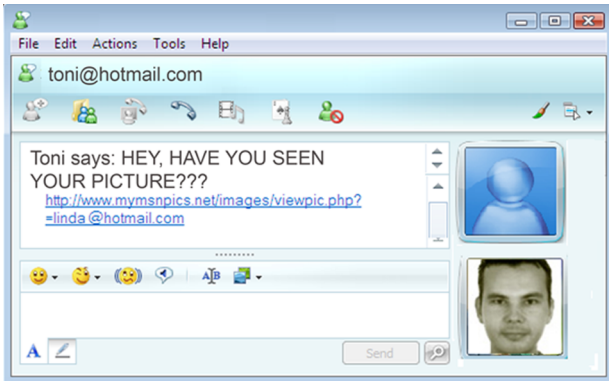
2.4. Atacs contra el programari de missatgeria instantània

El programari d'IM també pot ser víctima dels atacs contra els telèfons IP. En aquest sentit, els atacants poden enviar missatges de finalització de trucada o bé missatges de text falsos o que continguin programari maliciós. Hem comentat que una de les utilitats afegides als serveis d'IM és la transferència de fitxers. Addicionalment, també es poden enviar enllaços web per a poder compartir informació amb els participants d'una conversa. Doncs bé, aprofitant aquest fet es pot transferir programari maliciós mitjançant serveis d'IM.

Més concretament, per a fer més efectiu aquest atac convé que prèviament s'hagi tingut èxit en un atac de suplantació d'identitat. Si un contacte obre una conversa i ens demana que descarreguem un fitxer, en principi no hauríem de pensar que es tracta d'un programari maliciós. Ara bé, si hi ha hagut una suplantació d'identitat i el nostre contacte està utilitzant un idioma que no

és l'habitual, hi haurà motius per a malfiar-se. En la figura 5 es mostra un enviament que ha efectuat, previsiblement, un atacant. El detectem perquè la llengua que utilitza el contacte és l'anglès, que no és la seva habitual. Fent clic a l'enllaç mostrat, es descarregaria alguna mena de programari maliciós.

Figura 5. Exemple d'un enviament per part d'un atacant



A banda de la distribució de programari maliciós per mitjà de la suplantació d'identitat, alguns atacants utilitzen les deficiències de programació de les aplicacions d'usuari. En aquests casos, la distribució del programari maliciós a gran escala pot ser qüestió de minuts: els contactes que es tenen al servei d'IM formen, en el fons, una teranyina de connexions on milers d'usuaris estan indirectament connectats. El programa maliciós s'escampa cap a tots els contactes dels usuaris que estiguin connectats.

Per tal de minimitzar l'impacte d'aquests atacs, l'única solució és desconnectar de la xarxa els usuaris amb més contactes, o com a solució més dràstica, tancar temporalment els servidors.

Ara bé, si el que es vol és prevenir que "robots" enviïn programari maliciós, es proposa que l'agent d'usuari necessiti una resolució de *captcha* per tal de poder enviar fitxers i enllaços.

Captches

Un captcha és un test que es fa servir per a comprovar que l'usuari és humà. El mot ve de les sigles en anglès *completely automated public turing test to tell computers and humans apart*. En general, es tracta de teclejar els mots que apareixen en una imatge deformada.

Per concloure aquest apartat, diem que les aplicacions d'IM presenten cada cop menys problemes de seguretat. Avui dia són poques les aplicacions que han acabat acaparant grans quotes de mercat. Gràcies al fet que són aplicacions àmpliament conegudes i utilitzades, constantment són un banc de proves de milers d'usuaris *hacker*. Així doncs, si troben problemes de seguretat, s'avisarà la comunitat d'usuaris o els fabricants del programari per tal que hi posin remei.

3. Problemes de seguretat en la comunicació

Un cop hem estudiat els aspectes de seguretat relacionats amb servidors i agents d'usuari, analitzarem els problemes relacionats amb la seguretat de la comunicació. En concret, veurem la confidencialitat en les comunicacions, la integritat del que es transmet i l'autenticació dels participants.

3.1. Confidencialitat de la comunicació

Els serveis de VoIP i IM sovint serveixen per a organitzar reunions a distància entre diferents participants. En aquestes converses es poden tractar temes confidencials.

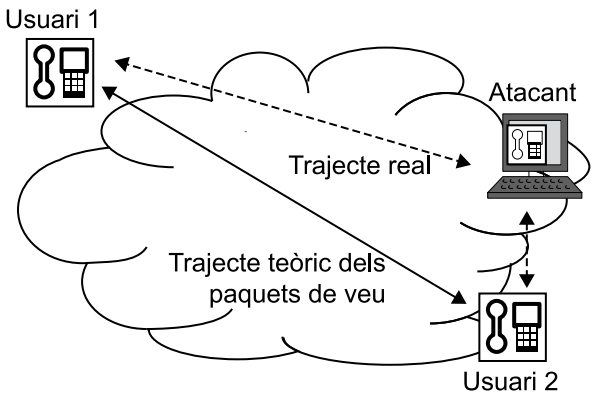
De la mateixa manera que els sistemes de telefonia tradicional poden ser "punxats" per tal d'enviar la conversa a un atacant (en aquest cas espia), les converses que circulen per sistemes de VoIP i IM poden ser interceptades.

Evidentment, les LAN actuals, basades en commutadors (*switches*) i no en concentradors (*hubs*), intenten evitar per si soles que usuaris no destinataris escoltin converses. Tanmateix, cal recordar l'efectivitat d'atacs basats en el protocol ARP en la intercepció de paquets de dades: de la mateixa manera que és possible que tots els datagrames IP siguin enviats cap a l'espia i després reenviats cap al destinatari, és perfectament viable que un atacant rebi paquets que continguin la veu o els missatges de text.

ARP

ARP significa *address resolution protocol*, i és el protocol encarregat de traduir adreces IP en adreces físiques de dispositiu de xarxa.

Figura 6. Escolta d'una conversa de VoIP



En el cas de la veu, tan sols caldrà descomprimir la informació i escoltar la conversa (recordeu que els còdecs de veu solen ser estàndard i per tant la seva interpretació i descompressió és coneguda). Aquest atac, per al cas de la VoIP, es mostra en la figura 6.

Per a garantir la confidencialitat, caldrà aplicar algun sistema de xifratge a la comunicació. En el cas de la VoIP, cal tenir en ment que el sistema emprat no ha d'introduir un retard significatiu que, afegit al retard mateix de la compressió de veu i la transmissió de paquets, degradi la qualitat de la conversa.

Els serveis d'IM també presenten problemes de seguretat relacionats amb la confidencialitat i la integritat de la informació.

En la VoIP, els atacs relacionats amb la seguretat de la informació tenen unes característiques específiques relacionades amb el fet que la informació que es transmet és veu digitalitzada. En canvi, la transmissió de text fa menys complexos la captura i el processament de paquets i, en conseqüència, l'èxit dels atacs.

Més concretament, és relativament senzill per a un atacant capturar de manera automàtica els paquets, processar-los per a seguir la conversa i filtrar els paquets que puguin contenir informació sensible. Per exemple, és més senzill filtrar tots els paquets que continguin el text *PIN* o *contrasenya*, que no pas aplicar un sistema de reconeixement de veu per a seleccionar aquells talls d'una conversa de VoIP que poder resultar interessants per a un atacant.

Finalment, però no menys important, cal tenir en compte que tots els paquets que es transmeten en entorns SIP, com ara els paquets de descripció de trucada del protocol SDP, són paquets en clar. Espiant aquests paquets es pot obtenir informació diversa, amb la qual un atacant podria fàcilment fer-se passar pel destinatari inicial i verdader de la trucada.

3.2. Integritat de la comunicació

La informació que es transmet per un sistema de VoIP és, en essència, veu humana digitalitzada i codificada. La modificació de paquets no tindria sentit, ja que la informació que es transmet no és textual, sinó un àudio.

Tanmateix, sota l'escenari d'atacs basats en ARP, es podria plantejar que l'atacant no solament escolta la conversa, sinó que la modifica. Un atac interessant seria canviar la freqüència i la modulació de la veu, per tal de "feminitzar" una veu masculina o a l'inrevés, o bé anar inserint silencis en el flux de veu. Aquest atac tindria incidència, en el fons, en la qualitat de la trucada.

Sembla més interessant que l'atacant pugui substituir un *sí* per un *no* amb diverses finalitats, com ara canviar la intenció de tota una conversa. Tot i això, aquest seria un escenari poc probable atesa la complexitat que té.

També és ben senzill canviar el contingut dels paquets de text en IM. Un atacant pot canviar text o fins i tot inserir-ne, ja que en principi l'aplicació no està tan lligada a mides màximes i especificitats dels còdecs de compressió com sí que ho està la VoIP. Només cal pensar en la complexitat tècnica que representaria substituir una paraula *sí* per un *no*, tal com hem suggerit abans per a la VoIP.

Caldrà, doncs, assegurar que els telèfons IP només processin paquets de veu que no s'han modificat. Per a garantir la integritat es poden fer servir codis d'autenticació de missatge a cada paquet, tenint en compte, per a la VoIP, els mateixos requisits quant a retards que s'han assenyalat en el cas de garantir la confidencialitat.

3.3. Autenticació dels participants

Els participants que prenen part en una conversa d'IM o VoIP han d'estar autenticats convenientment, és a dir, cal comprovar-ne la identitat. Una fallada en l'autenticació dels comunicants podria resultar en la suplantació d'identitat. Per exemple, en un escenari de VoIP l'atacant podria fer ús de paraules i expressions capturades durant una conversa de l'usuari suplantat per generar noves converses. Aquesta mena d'atac es podria usar amb un servei automàtic de contractació de serveis per telèfon.

Interessa més, però, descriure un atac basat en l'enverinament dels sistemes de localització IP d'usuaris: la redirecció de trucades o converses. Suposem que un atacant enverina el directori que emmagatzema la relació entre usuaris i l'adreça corresponent en un sistema VoIP. Aleshores serà possible que un usuari víctima truqui a un telèfon determinat i, en realitat, acabi trucant a l'atacant. Per exemple, un usuari truca al telèfon de l'entitat bancària i el sistema de VoIP el remet al telèfon d'un atacant. Aquest escenari sembla més senzill d'aplicar en trucades telefòniques que no pas en els casos de pesca (*phishing*), que emulen webs que suplanten la identitat (en general, d'entitats bancàries).

Un punt clau en l'autenticació dels comunicants és el moment en què els usuaris es validen al sistema. Si el sistema només empra un sistema de nom d'usuari i contrasenya, hi podrà haver problemes derivats de l'enginyeria social. A més, els sistemes d'identificació autenticats per contrasenya poden presentar els problemes inherents a l'ús de contrasenyes. Per exemple, l'ús d'atacs de diccionari o atacs de força bruta per a aconseguir trobar contrasenyes considerades febles. Tanmateix, tot i tenir una contrasenya robusta, els anomenats *atacs d'enginyeria social* poden assolir amb èxit l'objectiu d'aconseguir una contrasenya, per robusta que sigui.

Atacs d'enginyeria social

No es tracta només de trucar a un usuari i demanar-li la contrasenya d'IM o VoIP, fent creure que som el proveïdor de servei, sinó que es pot anar molt més enllà. Per exemple, en algunes pàgines web es podria trobar algun anunci o petit joc en línia (*online*) que ens fes creure que per a accedir-hi cal que ho fem per mitjà del servei d'IM. El web en qüestió presentaria un quadre de diàleg amb una interfície similar al de l'agent d'usuari d'IM, on la víctima posaria el nom d'usuari i la contrasenya. De seguida aquesta informació faria cap als servidors de l'atacant, el qual podria usar les credencials "robades" per a entrar al servei d'IM l'objectiu d'atacar.

Clarament, l'ús d'autenticació dels participants per mitjà de certificats digitals hauria de garantir que es podrà comprovar l'autenticitat dels participants en la comunicació i, per tant, hauria de fer més difícil la suplantació d'usuaris. Tanmateix, també és important la cultura de la precaució a l'hora de navegar per Internet i facilitar dades confidencials.

4. Eines per a comunicacions segures

Un cop hem vist quins són els problemes de seguretat relacionats amb les comunicacions, analitzarem les tècniques de seguretat que s'hi poden aplicar. Amb aquest objectiu, veurem protocols concrets que es poden fer servir.

Primerament, tractem com protegir la senyalització i la gestió de la trucada per mitjà de la protecció amb tècniques i protocols criptogràfics dels paquets que contenen la informació d'establiment de trucada, gestió dels participants, etc. En segon lloc, descriurem els sistemes estàndard que permeten garantir la seguretat en el transport de la veu.

Per a assegurar sistemes basats en SIP, la IETF va proposar l'ús de protocols ja existents. Pel que fa les recomanacions UIT per a VoIP, són similars però no s'entra en tant de detall. Aquest organisme recull les recomanacions de seguretat per a l'H.323 dins el protocol H.235.

4.1. Seguretat en la senyalització

Per a la senyalització (establiment i gestió de trucades, bàsicament), convé garantir l'autenticació en l'origen dels paquets, i també que la informació no es modifica (és a dir, la integritat dels paquets). Addicionalment i, com que SIP utilitza missatges en format de text i en conseqüència no massa difícil d'espiar, també s'assegurarà la confidencialitat de la informació que es transmeti durant la senyalització.

4.1.1. Sistemes SIP

La IETF proposa utilitzar, per als sistemes de VoIP, IM i videoconferència, el format S/MIME per a l'autenticació de la informació corresponent a SDP. Tot i ser un format inicialment pensat per al correu electrònic segur, el cert és que permet la tramesa de signatures digitals, xifratges i la gestió dels certificats necessaris.

S/MIME

S/MIME significa *secure multipurpose Internet mail extensions*, i es troben especificats en les RFC 3850 i 3851.

Ara bé, cal tenir present que la informació corresponent, per exemple, al destinatari, no pot estar xifrada: els diferents servidors per on s'anirà encaminant la trucada o possibles passarel·les han de disposar d'aquesta informació. Els missatges S/MIME contenen una signatura digital d'aquesta informació, de manera que el destinatari pot comprovar, si més no, que el missatge SDP no s'ha modificat durant el trajecte. La figura 7 mostra un exemple d'ús de S/MIME per a autenticar la informació SDP que conté un paquet SIP.

Figura 7. Estructura d'un paquet SIP que usa S/MIME per a autenticar la informació SDP

```
INVITE sip:jordi@uoc.edu SIP/2.0
Via: SIP/2.0/udp ...
From: <toni@urv.cat> ...
...
...
Content-Type: multipart/signed;boundary=e4ef8847482d240d0
Accept: application/sdp, multipart/mixed
Content-Length: 3381
--e4ef8847482d240d0
Content-Type: application/pkcs7-mime
smime-type=envelopeddata; name=smime.p7m
Content-Disposition: attachment;handling=required;filename=smime.p7m
Content-Transfer-Encoding: binary
*** envelopedData object containing encrypted SDP body ***
* v=0
* o=- 0 0 IN IP6 2001:db8::27:2
* m=audio 49170 RTP/AVP 112 113
...
**********************************************************
--e4ef8847482d240d0
Content-Type: application/pkcs7-signature;name=smime.p7s
Content-Disposition: attachment;handling=required;filename=smime.p7s
Content-Transfer-Encoding: binary
```

Missatge SIP

Disposició en clar SDP

Signatura SDP

A més d'assegurar certa informació per mitjà de l'ús d'S/MIME, per als sistemes SIP es recomana usar TLS[12] per a enviar paquets de senyalització de manera segura. Aquest protocol permet l'autenticació dels participants en una comunicació, alhora que assegura la integritat i la confidencialitat de les dades que hi circulen. Tanmateix, és necessària la participació d'una infraestructura de clau pública[13].

(12)TLS significa *transport layer security*. Està basat en SSL (*secure socket layers*) i es troba definit en l'RFC 2246.

(13)En anglès, *public key infraestructure*, PKI.

Vegeu també

TLS i SSL es descriuen en el mòdul "Seguretat en xarxes WLAN" d'aquesta assignatura.

Qualsevol protocol de VoIP que utilitzi TCP com a canal de transport en l'establiment i el control de la trucada pot usar de manera senzilla el TLS per a assegurar la comunicació.

TLS proporciona una capa de transport segur fent servir un conjunt ampli d'algorismes i paràmetres possibles, els quals es defineixen per mitjà d'un protocol de negociació entre els participants[14]. Val a dir que TLS proporciona seguretat basant-se, en el fons, en la confiança que es tingui envers qui genera el certificat dels comunicants. Si un comunicant vol establir una trucada VoIP amb un altre, convé que el primer confiï en l'emissor del certificat del segon participant.

(14)També anomenat *handshake* o *encaixada de mans*.

4.1.2. Recomanacions de la UIT

L'estàndard H.235 especifica un conjunt de mecanismes de seguretat d'aplicació en un entorn de VoIP basat en protocols de l'H.323. Aquestes recomanacions són en certa manera similars al que s'especifica per als serveis basats en SIP. Pel que fa als protocols relacionats amb l'establiment i la gestió de les trucades, la UIT recomana el següent:

- Els missatges del protocol H.225/RAS poden transportar informació sobre quins protocols i paràmetres de seguretat es faran servir en la resta de protocols implicats.

- El protocol de control de trucada H.245 es pot protegir per mitjà de l'ús de TLS. Així doncs, els participants en una comunicació es poden autenticar en l'execució pròpia del TLS. Dins aquest protocol de control també es pot especificar amb quins algorismes i paràmetres s'asseguraran els paquets de veu.

- Els missatges del protocol H.225/Q.931 es poden protegir fent servir TLS.

El protocol H.235 defineix una sèrie de perfils de seguretat, cadascun dels quals regula o bé diferents aspectes de la protecció de VoIP (per exemple, autenticació en el protocol RAS usant claus compartides, intercanvi de claus per al xifratge de la veu, etc.) o bé diferents nivells de protecció.

Un dels nivells més usats és el protocol H.235.1 (*baseline security profile*), que garanteix autenticació i integritat, però no confidencialitat. En concret, protegeix envers els atacs de tipus "home al mig" (*man-in-the-middle*) i "segrest de sessió" (*session hijacking*), entre d'altres.

4.2. Seguretat en l'enviament de la conversa

Pel que fa les dades corresponents a la veu, hem vist que cal assegurar-ne la integritat (que no s'hagi modificat la veu o els missatges de text). També cal evitar atacs d'enviament de paquets fraudulents de veu (que implicarien talls i soroll en les converses). També cal garantir l'origen dels paquets de veu i text.

4.2.1. Sistemes SIP

Els sistemes SIP, i com veurem també en l'estàndard UIT, proposen l'ús d'SRTP per a donar seguretat als paquets que contenen veu. SRTP xifra les dades que transporta un paquet RTP, autentica el paquet RTP i a més protegeix envers atacs de *replay*:

SRTP

SRTP significa *secure real time transport protocol* i està definit en l'RFC 3711. Aquest protocol també es pot aplicar als paquets de vídeo, en el cas de la videoconferència.

- La confidencialitat s'assoleix per mitjà de xifratge AES.

- L'autenticació també utilitza el xifratge AES.

- Un atac de replay consisteix en el fet que un atacant intercepti paquets vàlids i els enviï de manera repetida, o bé amb retard, per malmetre la composició del missatge en l'aplicació destinatària. Els atacs de replay s'eviten amb l'ús de números de seqüència.

Ara bé, el problema més important és com es comparteix la clau que s'usarà en AES. Un dels sistemes emprats en aquest cas és el protocol MIKEY, un protocol de gestió de claus proposat amb la intenció de ser utilitzat en aplicacions multimèdia de temps real. Aquest protocol presenta tres maneres de compartir l'anomenada TEK (*traffic encryption key*):

MIKEY

MIKEY significa *multimedia Internet keying*. Està definit en l'RFC 3830.

- Clau precompartida, que és un mètode eficient perquè usa la criptografia simètrica. Malgrat això, cal intercanviar una clau única TGK (*TEK generation key*) per a cada participant possible en la comunicació i això pot implicar problemes d'escalabilitat.

- Clau pública, on la TEK es transmet per mitjà de criptografia de clau pública, com el seu nom indica. Implica l'ús d'una PKI.

- Diffie-Hellman, on s'empra aquest sistema d'intercanvi. Malgrat que és el mètode que més recursos consumeix, tant d'esforç computacional com d'amplada de banda, es considera el més segur ja que proporciona l'anomenada *perfect forward secrecy*.

Addicionalment, el missatge SDP conté un atribut anomenat *crypto* que especifica els protocols usats i les claus. Cal recordar que, com que SIP envia els missatges en clar, se suposa que l'enviament de l'SDP es fa emprant una comunicació segura, com ara la proposada amb S/MIME + TLS. Un exemple de l'atribut seria el següent, on a inline s'especifica la clau:

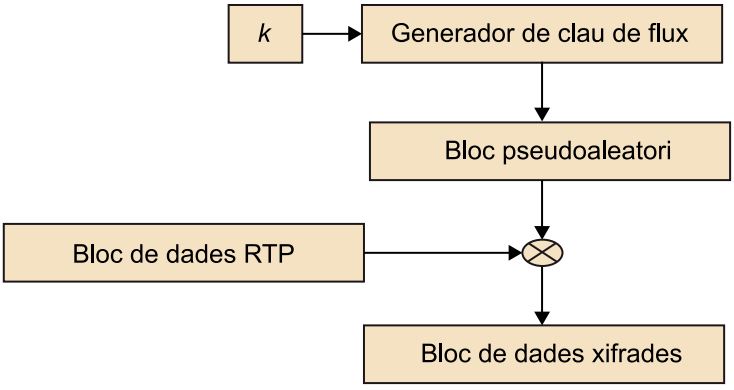
```
a=crypto:1 AES_CM_128_HMAC_SHA1_32
inline:NzB4d1BINUAvLEw6UzF3WSJ+PSdFcGdUJShpX1Zj|2^20
```

En l'exemple anterior, se sol utilitzar el protocol AES en mode comptador (*counter mode*, CM) per a obtenir confidencialitat. Aquest protocol es fa servir per a generar un flux de bits de xifratge a partir de la clau de xifratge. Aquests s'apliquen en blocs de 128 bits (el nombre de bits que s'especifica a l'entrada *crypto*) al camp de dades dels paquets RTP, tot fent una XOR bit a bit. Aquest procés es mostra en la figura 8.

Enllaç d'interès

En aquest subapartat, per raons didàctiques i de simplicitat, hem fet una descripció breu del procés SRTP. Com a lectura recomanada, podeu llegir *Securing Internet Telephony Media with SRTP and SDP*.

Figura 8. Xifratge d'un paquet RTP per mitjà d'SRTP



Per a l'autenticació de paquets, SRTP crea un codi d'autenticació de missatge emprant la funció hash especificada en l'entrada *crypto*, en el cas de l'exemple SHA-1, tenint com a entrada la capçalera i les dades del paquet RTP que s'ha d'autenticar i la clau.

Cal tenir present que aquestes operacions criptogràfiques no han d'alentir l'enviament dels paquets de veu, o si més no garantir un retard màxim acceptable per a mantenir una conversa sense dificultats. En VoIP, tenint present que la major part de les dades que genera el servei es correspon amb la veu, i la necessitat de tenir el retard afegit mínim possible, convé prestar especial atenció a l'eficiència de les tècniques i als protocols de seguretat.

Per a la IM, observeu que afegir autenticació implica en determinats casos una pèrdua d'eficiència. En les converses d'IM tot sovint s'envien com a resposta missatges ben curts (*ok*, *adéu*, etc.). En canvi, en VoIP, la digitalització i la compressió de la veu implica una successió de paquets d'una mida prou gran perquè surti a compte afegir-hi informació d'autenticació. Tot i això, en IM els requisits de temps real són més baixos, ja que el servei es pot permetre uns retards màxims superiors als estipulats per a VoIP.

4.2.2. Recomanacions de la UIT per a veu sobre IP

Dins l'H.235 hi ha tres perfils de seguretat que tenen relació amb la protecció de l'enviament d'àudio:

- H.235.6. Perfil de xifratge de veu amb gestió de claus per mitjà del protocol H.235/H.245
- H.235.7. Ús del sistema MIKEY per a la gestió de claus per a SRTP
- H.235.8. Intercanvi de claus per a SRTP usant canals segurs de senyalització

D'aquesta manera, usant el protocol H.235.1 més H.235.6 i H.235.8 per a intercanviar claus es podrien garantir les propietats d'autenticació i integritat (que proporciona el protocol H.235.1) i confidencialitat (proporcionada per SRTP).

4.2.3. Ús d'IPsec

Cal preveure que en organitzacions amb seus distribuïdes geogràficament, s'efectuaran trucades que utilitzaran Internet o bé la xarxa de trànsit de l'operadora de telecomunicacions. En aquestes xarxes, els paquets de VoIP s'han de protegir adequadament, sobretot si la informació circula per Internet o bé l'operadora no garanteix l'ús de canals segurs del tipus de xarxa privada virtual.

Vegeu també

IPsec és un sistema que proporciona confidencialitat i autenticació en xarxes IP. Es pot obtenir més informació d'aquest sistema en el mòdul "Sistemes de tallafocs" d'aquesta assignatura.

Així doncs, els fabricants recomanen que, en aquests escenaris, els encaminadors que uneixen la xarxa interna amb la xarxa externa puguin utilitzar el protocol IPsec per a assegurar la informació de VoIP que circula entre seus.

4.3. Implicacions en els sistemes tallafoc

Hem comentat que l'ús de tallafocs que controlin la circulació de paquets entre les diferents xarxes i els servidors és essencial per a prevenir atacs de denegació i degradació del servei. A més de controlar amb diferents regles el trànsit de paquets, caldrà preveure la detecció d'atacs contra aquests tallafocs. En el cas concret de VoIP, alguns fabricants recomanen que els sistemes de veu (tant servidors com agents d'usuari) es posin dins una zona segura, creada a partir de tallafocs. Ara bé, cal tenir algunes consideracions en l'ús dels tallafocs i VoIP o IM, que tot seguit recollim.

Els tallafocs solen fer funcionalitats de traducció d'adreces i ports en els paquets que entren i surten, amb l'objectiu que amb una mateixa adreça "pública" es faciliti l'accés a un conjunt de dispositius de xarxa "interns", cadascun amb una adreça de rang privat. Per mitjà de la monitorització dels ports i de les adreces indicats als paquets que travessen el tallafoc, es pot determinar si un paquet correspon a una sessió ja iniciada.

IPtables

En l'eina IPtables es poden permetre connexions noves a partir de connexions ja permeses per mitjà d'*ESTABLISHED* i *RELATED*.

Això permet que un servidor de VoIP o IM intern pugui establir diverses connexions amb l'exterior, ja que per a la senyalització de la trucada s'utilitzen protocols que usen TCP i, clarament, els tallafocs no tenen cap mena de problemes a controlar els paquets corresponents a una sessió TCP.

Ara bé, sí que hi pot haver problemes per a intentar controlar diversos fluxos de paquets UDP, que és el cas de les dades, la veu i el vídeo, si escau. Així com els ports TCP que s'empren per a la senyalització de trucades estan ben definits i són coneguts, no passa el mateix amb els ports corresponents a UDP i el trànsit de trucades. Una solució possible seria que el tallafocs tingués obert, per defecte, un rang de ports UDP, però aquesta solució clarament obre un forat de seguretat i fa que el tallafoc pugui perdre eficàcia. De fet, ni el nombre de trucades que es poden establir pot ser previsible: sempre hi hauria ports per obrir o sempre quedarien ports oberts.

L'ideal seria que el tallafoc mateix, en detectar que una sessió SIP s'ha fet efectiva, obrís un determinat rang de ports UDP per tal que el canal de veu també es pogués materialitzar. El tallafoc podria analitzar immediatament quin és el port que es fa servir de manera efectiva per al trànsit de veu, dades o vídeo, i tancar-ne la resta. Tot i això, aquesta anàlisi hauria d'implicar usar tallafocs amb gran capacitat de computació si no volem que la seva aplicació degradi l'eficiència del tallafoc i afegeixi un coll d'ampolla considerable.

Els tallafocs amb inspecció d'aplicació tenen un clar avantatge respecte dels que només es limiten a analitzar breument les capes de xarxa i transport dels paquets que reben. Aquesta mena de tallafocs analitzen els paquets (per exemple, el camp del missatge SIP on s'especifica quins ports UDP s'usaran) i ho tenen menys complicat per a saber a quina trucada o conversa pertany cada paquet UDP i, per tant, què es pot obrir i cap a quina màquina s'ha d'enviar.

Malgrat això, hem vist que en les connexions per a sistemes de seguretat que fan servir protocols de xifratge i autenticació, si la informació està xifrada, els tallafocs d'aplicació ho tenen difícil per a gestionar connexions (a no ser que se'ls permeti desxifrar la informació, cosa que implicaria gestió de claus...).

De tots aquests aspectes es desprèn que l'ús de tallafocs en els sistemes de VoIP i IM segurs té implicacions en la gestió de connexions i, de retruc, en l'eficiència del tallafoc mateix. És per això que moltes aplicacions empren sistemes de túnel, amb els quals tot el trànsit de veu UDP passa a la xarxa per mitjà d'un port determinat i usant connexions TCP.

Resum

En aquest mòdul hem estudiat els problemes de seguretat que presenten les aplicacions de veu sobre IP i missatgeria instantània. Molts dels problemes de seguretat que presenten són, en certa forma, comuns i, per tant, les solucions a aplicar-hi solen ser similars.

Hem començat veient com funcionen els serveis de veu sobre IP i missatgeria instantània. Hem vist clarament que funcionen mitjançant dos tipus de comunicació: comunicació de senyalització, en general sobre un canal TCP, i comunicació de la informació (veu digitalitzada o bé missatges curts de text). Aquesta darrera comunicació se sol dur a terme en forma d'un flux de paquets UDP. El protocol SIP utilitza missatges de text també per a la senyalització (és a dir, establir i gestionar trucades). El protocol H.323 de la UIT fa servir paquets binaris. Tot i això, els estàndards no especifiquen mai que la informació s'ha de xifrar.

Com a tendències de protecció en veu sobre IP, hem vist que s'ha de senyalitzar sobre un transport segur TLS, mentre que la veu pot fer ús de l'SRTP, un afegitó al protocol RTP que permet confidencialitat dels paquets i autenticació. Aquesta tendència s'usa tant en sistemes SIP com en sistemes de la UIT. Per a l'enviament de paquets sobre Internet es recomana, si és possible, l'ús d'IPsec.

Un altre aspecte fonamental dels serveis estudiats és que se sustenten en una sèrie de servidors amb diferents funcionalitats. Aquests servidors solen ser el blanc d'atacs d'usuaris maliciosos, que també poden tenir com a objectiu atacar els terminals de veu sobre IP o el programari d'usuaris de missatgeria instantània. Així doncs, cal protegir aquests servidors tenint en compte, però, que l'ús de tallafocs no és tan simple com es podria pensar, sobretot si s'utilitza comunicació xifrada.

Pel que fa a missatgeria instantània, després d'indicar que és menys complex atacs contra la confidencialitat en missatges de text que en paquets de veu, o bé és més senzill modificar missatges de text que converses de veu, hem enfocat el problema principal de seguretat de la IM: la distribució de programari maliciós.

Activitats

1. Esbrineu quins són els mecanismes de seguretat que es fan servir en les eines més populars de VoIP i missatgeria instantània. Amb aquest objectiu, usareu un detector (*sniffer*) per a capturar els detalls de sessions en aquests programes. Caldrà que analitzeu si hi ha una protecció de transport, per exemple utilitzant SSL o TLS, o bé si s'utilitzen tècniques d'autenticació d'establiment de sessió. Finalment, digueu si s'estableixen diferents comunicacions i mesures de seguretat per a la senyalització i per al transport de missatges i veu.

2. La missatgeria instantània és una eina que s'ha integrat al web, sigui des dels seus inicis amb les pàgines web que permeten xats, o sigui des dels portals de xarxes socials. Analitzeu si l'establiment de converses entre membres en xarxes socials, per exemple, s'esdevé de manera segura segons el que s'ha vist en el mòdul.

3. La Voice over IP Security Alliance (VOIPSA) és una aliança de fabricants i desenvolupadors de productes per a VoIP que vetlla per promoure les tècniques de seguretat per a VoIP. Aneu al web i feu un cop d'ull a alguns dels articles més recents en matèria de seguretat en VoIP.

4. Dins el mateix web, accediu a l'apartat sobre eines de seguretat en VoIP. Intenteu provar alguns dels detectors (*sniffers*) i analitzadors de SIP o, fins i tot, de veu.

5. Consulteu Internet i cerqueu possibles atacs orientats als telèfons IP que s'hi hagin descrit.

Glossari

agent d'usuari *m* Element de l'arquitectura VoIP o de missatgeria instantània per mitjà del qual l'usuari interactua amb el servei. En el cas de VoIP es tracta dels telèfons IP.

contacte *m* Usuari d'un servei de missatgeria instantània amb el qual un usuari pot establir una conversa.

H.323 *m* Conjunt de protocols proposat de la UIT sobre l'establiment de trucades de veu sobre xarxes IP. Utilitza protocols de senyalització de la telefonia tradicional, com ara el Q.931.

IETF *f* Força operacional de l'enginyeria per a Internet. Organització que s'encarrega de l'estandardització de protocols que intervenen en el funcionament d'Internet, com ara el TCP, l'HTTP, el SIP, etc.
en Internet engineering task force

IM *f* Vegeu **missatgeria instantània**.

International Telecommunications Union *f* Vegeu UIT.

Internet engineering task force *f* Vegeu IETF.

IP-PBX *m* Servidor de l'arquitectura VoIP encarregat de gestionar l'establiment de trucades.

malware *m* Vegeu **programari maliciós**.

media gateway control protocol *m* Vegeu **MGCP**.

MGCP *m* Protocol de control de la passarel·la de mitjans que s'encarrega de la conversió de mitjà entre telefonia tradicional i VoIP i l'adaptació de senyals d'establiment de trucada.
en media gateway control protocol

MIKEY *m* Distribució de claus per Internet en multimèdia. Protocol senzill que permet compartir i generar claus de sessió per al xifratge de contingut en sessions multimèdia.
en multimedia Internet keying

missatgeria *f* Tecnologia que permet mantenir converses fent servir missatges curts de text que es transmeten en temps real.
sigla **IM**

multimedia Internet keying *m* Vegeu MIKEY.

passarel·la telefònica *f* Servidor de l'arquitectura VoIP encarregat de connectar la xarxa de VoIP a la xarxa telefònica convencional.

programari maliciós *m* Programari divers l'objectiu del qual és causar alguna molèstia, lleu (per exemple, aparició de publicitat massiva) o greu (pèrdua de dades del sistema), als usuaris.
en malware

real-time transport protocol *m* Vegeu RTP.

RTP *m* Protocol de transport en temps real que afegeix als datagrames transportats en UDP informació sobre l'instant de temps i amb quin número de seqüència s'han generat.
en real-time transport protocol

SDP *m* Protocol que, dins un entorn SIP, s'encarrega de descriure paràmetres diversos sobre la sessió multimèdia que s'ha d'iniciar.
en session description protocol

session description protocol *m* Vegeu SDP.

session initiation protocol *m* Vegeu SIP.

SIP *m* Protocol d'inici de sessions que utilitza peticions en missatges de text per a establir connexions multimèdia en entorns IP.
en session initiation protocol

softphone *m* Vegeu **telèfon IP**.

telèfon *m* Dispositiu que fa les funcions de digitalització i reproducció de veu i que permet l'establiment de trucades en un entorn de VoIP. Programari que fa les funcions de telèfon IP.

en softphone

UIT *m* Unió Internacional de Telecomunicacions. Organisme que s'encarrega de l'estandardització de protocols i sistemes que intervenen en les telecomunicacions.
en International Telecommunications Union

veu sobre IP *f* Tecnologia que permet establir converses telefòniques usant datagrames IP com a mitjà de transport.

VoIP *f* Vegeu **veu sobre IP**.

Bibliografia

Baugher, M. i altres. *Securing Internet Telephony Media with SRTP and SDP* (en línia). Cisco.

Butcher, D. i altres. (2007). "Security Challenge and Defense in VoIP Infrastructures". *IEEE Transacions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* (vol. 37, núm. 6, pàg. 1152-1162).

Endler, D. i altres. (2006). *Hacking Exposed VoIP: Voice Over IP Security Secrets & Solutions*. McGraw-Hill Professional Publishing.

Gollmann, D. (2005). *Computer Security* (2a. ed.). Wiley.

Herrera Joancomartí, J. (2006). *Aspectos avanzados de seguridad en redes*. Editorial UOC.

Mannan, M. i altres. (2005). *On Instant Messaging Worms, Analysis and Countermeasures, Proceedings of the 2005 ACM workshop on Rapid Malcode*.

Sisalem, D. i altres. (2009). *SIP Security*. Wiley.

Symantec. *Securing Instant Messaging*.

Zhua, Y. i altres. (2011). "Traffic analysis attacks on Skype VoIP calls". *Computer Communications* (vol. 34, núm. 10, pàg. 1202-1212).