INTERCEPTAÇÃO TELEFÔNICA EM REDES DE TELEFONIA IP

Maria Jane de Queiroz[1]
Aluízio Ferreira da Rocha Neto[2]

RESUMO

Observa-se que muito se fala em convergência de dados, voz e vídeo, conceito que levou ao desenvolvimento de novos protocolos e serviços de transmissão de pacotes multimídia utilizando-se a rede IP. Sabe-se que o protocolo IP não possui estruturas de segurança implementadas em seu cabeçalho. Afora isso, os demais protocolos criados especificamente para transmissões de pacotes em tempo real e que trafegam sobre o protocolo IP também não possuem campos e técnicas relacionadas à segurança. Com o surgimento do VOIP e da telefonia IP, a quantidade de riscos a que se expõem as conversações que utilizam esses protocolos aumenta gradativamente, pois na maioria das vezes, os responsáveis pela implantação e administração dos serviços não se preocupam em adicionar camadas de segurança, já que pouco se fala em ataques a redes de telefonia IP, embora eles existam. Considerando esses pontos, este artigo tem como finalidade apresentar as principais falhas de segurança, além dos protocolos e técnicas para corrigir essas falhas e erradicar a prática de escutas telefônicas ilegais na telefonia IP.

Palavras-chave: Telefonia IP. VOIP. Segurança. Redes de Computadores.

CALL EAVESDROPPING ON IP TELEPHONY NETWORKS

ABSTRACT

Nowadays, much is said about data, voice and video convergence. This concept led to the development of new protocols and services for multimedia packet transmission through IP network. It is known that the IP protocol has no security structures implemented in its header. Besides that, other protocols specifically designed for packet transmissions in real time,that travel over the IP protocol, also lack fields and techniques related to security. With the emergence of VOIP and IP telephony, the conversations that use these protocols present greater security risks because, in most cases, those responsible for the implementation and administration of the services do not bother adding security layers, since little is said about attacks on IP telephony networks, although they exist. Considering these points, this article aims at presenting major security flaws, in addition to the protocols and techniques to correct them and eradicate the eavesdropping practice on IP telephony.

Keywords: IP telephony. VOIP. Security. Computer Networks.

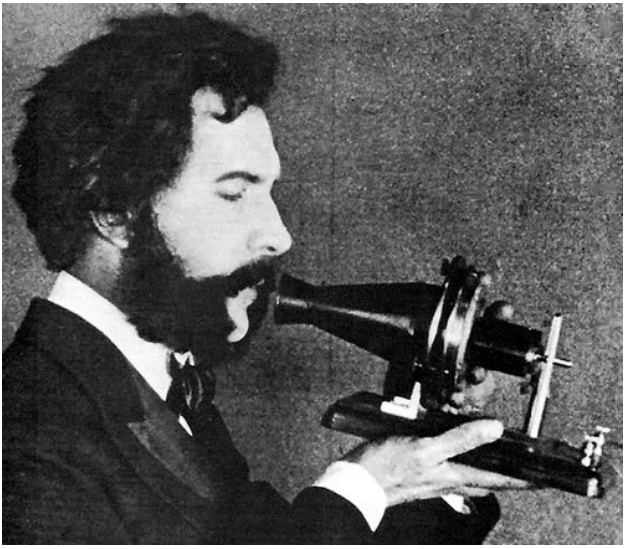
1 Docente em Redes de Computadores no IFRN, Campus de Caicó. Especialista em Redes de Computadores - UNI-RN. Graduada em Tecnologia em Redes de Computadores pelo IFRN. E-mail: jane.mjq@gmail.com. CV: http://lattes.cnpq.br/1114855519309156.

2 Docente Orientador do Curso de Especialização de Tecnologia em Redes de Computadores. E-mail: aluizio@imd.ufrn.br. CV: http://lattes.cnpq.br/5056619278818251.

1 INTRODUÇÃO

A cada marco histórico da evolução humana houveram meios de comunicação envolvidos, ou para registrar os fatos ou para repassar os conhecimentos adquiridos. Em épocas remotas, foram utilizados os sons guturais que se desenvolveram dando forma à fala humana como conhecemos e utilizamos hoje. Devido ao contínuo deslocamento de grupos humanos rumo a diferentes cidades, países e até continentes, além de inúmeros confrontos e guerras, com o passar do tempo apenas o desenvolvimento da voz humana não era suficiente para garantir a continuidade da comunicação. As grandes distâncias contribuíram para o desenvolvimento de aparelhos como o telégrafo, criado por Samuel Morse em 1844. Por sua vez, o sistema de telegrafia contribuiu para os estudos do cientista Alexander Graham Bell (figura 1), que patenteou o telefone no dia 14 de fevereiro de 1876. Segundo ele, o objetivo do aparelho era "... transmitir voz e outros sons [...] pelas variações da corrente elétrica, similares às variações do ar, acompanhando cada palavra pronunciada [...]" (COLCHER et al, 2005).

Figura 1 – Alexander Graham Bell e sua invenção

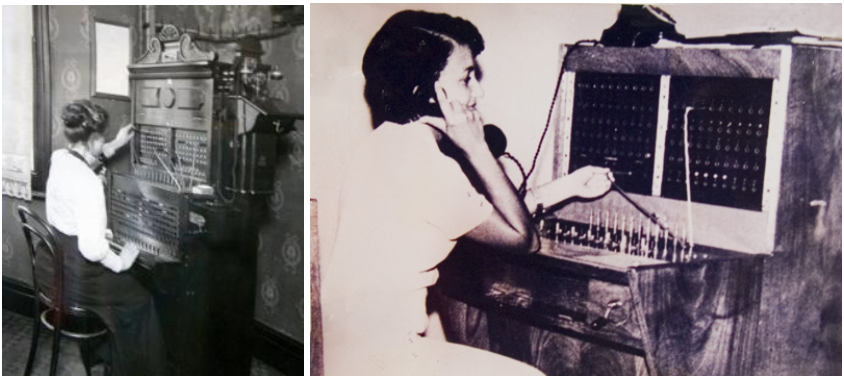


Fonte: Blog Roger Cinema (2011)

Desde então, surgiram as primeiras centrais telefônicas e órgãos reguladores, como o ITU (*International Telecommunications Union*), responsáveis pela padronização técnica e operações de sistemas de telecomunicações. Nascia então a Rede Telefônica Pública Comutada (RTPC), ou seja, a tradicional rede de telefonia fixa utilizada até hoje.

Nos primeiros sistemas telefônicos, as operações de estabelecimento de chamada consistiam em chaveamentos físicos, feitos manualmente pelos operadores, geralmente do sexo feminino, como mostram as imagens da figura 2.

Figura 2 – Antiga central telefônica utilizando chaveamento físico manual



Fontes: Blog Discípulos de Graham Bell; Blog Focado em Você

Com o crescente acesso ao sistema telefônico pela população, este evoluiu mais uma vez, tornando-se automático. Não eram mais necessários os chaveamentos físicos manuais. Houve uma notável evolução dos aparelhos e linhas telefônicas, o que aumentou ainda mais a quantidade de usuários da RTPC.

Após o desenvolvimento do sistema de telefonia, surgiram as redes de computadores, tendo início com a criação da ARPANET no ano de 1967 e de outras redes ao longo do tempo, que foram interligadas dando origem à grande rede mundial de computadores conhecida hoje como Internet.

A Internet consiste em um conjunto de redes interligadas, utilizando comutação por pacotes e não por circuitos, como acontece na rede de telefonia tradicional. A comutação por pacotes permite uma melhor utilização da largura de banda disponível para a transmissão de dados, pois podem

ocorrer diversas transmissões simultaneamente, sem que um circuito esteja dedicado durante todo o tempo de uso a apenas um usuário ou aplicação.

Devido ao crescimento e desenvolvimento da Internet, das ferramentas de desenvolvimento, das aplicações e consequentemente, dos mais variados serviços para transmissão de dados, voz e vídeo pela *web*, as tecnologias de telecomunicações e processamento de informações estão convergindo rumo à grande rede de computadores, dando origem a novas soluções de comunicação, utilizando a comutação de pacotes como princípio de funcionamento. Graças a esse processo de convergência, surgiram novos serviços, dentre eles o VoIP (*Voice over IP* ou Voz sobre IP).

2 VOZ SOBRE IP

Como citado anteriormente, a telefonia tradicional utiliza a comutação por circuitos. Isso significa que para realizar uma chamada, todos os recursos necessários (*buffers*, taxas de transmissão e circuitos físicos) serão reservados durante todo o período de comunicação. Há uma garantia quanto à entrega e taxa constante na transmissão dos dados, pois a largura de banda foi alocada fim a fim previamente para estas funções.

Já na comutação por pacotes, os recursos não são reservados, mas sim utilizados sob demanda. O protocolo IP (*Internet Protocol* ou Protocolo Internet) trabalha com base no conceito de melhor esforço, porém não garante a entrega confiável dos dados no seu devido tempo. Como consequência, podem ocorrer atrasos no envio de pacotes, perca de dados e falhas na conexão (KUROSE, 2010).

Por tais motivos, alguns autores defendem que a comutação por pacotes não é adequada a serviços de tempo real, como ligações telefônicas e videoconferências. Ligações telefônicas utilizando comutação por pacotes se tornaram possíveis graças à convergência de tecnologias de transmissão de voz e dados utilizando uma mesma infraestrutura de rede. Assim surgiu o VoIP (*Voice over IP*) que consiste na transmissão de voz utilizando a rede IP, sem a necessidade de se construir uma nova infraestrutura para a transmissão de voz.

Mesmo com os problemas citados anteriormente, a tecnologia VoIP vem se desenvolvendo por oferecer vantagens em relação ao uso do sistema telefônico tradicional, como a redução de custos com tarifas

telefônicas convencionais, melhor uso da largura de banda, além da possibilidade de se utilizar a mesma infraestrutura de cabeamento para a implantação da rede de voz.

ROSS (2007) defende que o padrão IPv4 (IP versão 4) usado nas redes atuais não é o mais indicado para trafegar voz por não prover mecanismos de controle de QoS (*Quality of Service* ou Qualidade de Serviço), como por exemplo, a priorização de tráfego. Para ele, a solução para trafegar voz em redes IP seria a adoção do protocolo IPv6 (IP versão 6) ou de protocolos de controle que possam garantir a qualidade necessária nesses tipos de transmissão.

3 SEGURANÇA EM REDES VOIP

Em meio a tantos desafios e controvérsias, a telefonia IP (forma mais comum de transmissão de Voz sobre IP) vem se desenvolvendo e sua adoção está em crescente expansão. Além dos desafios de qualidade de serviço, o que mais preocupa seus usuários é a questão da segurança.

Como todo serviço de rede, a telefonia IP não está imune a ataques. Pelo contrário, além dos cuidados comuns a quaisquer serviços de rede, os responsáveis pelos servidores PABX (*Private Automatic Branch Exchange* ou Troca Automática de Ramais Privados) devem se preocupar com todo o trajeto da informação, desde a origem até o destino, já que ataques de interceptação telefônica em redes IP, em alguns casos específicos, podem ser até um pouco mais simples que em redes telefônicas convencionais.

Embora existam diversas técnicas e dispositivos para grampear telefones fixos convencionais, esse tipo de interceptação exige que o espião instale um aparato no próprio telefone a ser grampeado ou no poste próximo à localização do equipamento a ser utilizado pela vítima. Qualquer uma das duas situações coloca em risco o descobrimento da ação. Mesmo as escutas telefônicas permitidas por lei (lei nº 9.296, art. 5°, inciso XII da Constituição Federal), correm o risco de serem descobertas.

No que se refere a escutas telefônicas realizadas em ambientes de redes, as técnicas utilizadas são difíceis de serem descobertas, pois em alguns casos não se faz necessária a instalação de *hardware* adicional para realizar a ação. Com um computador conectado à rede local, alguns *softwares* e técnicas adicionais, pode-se realizar a famosa *Call Eavesdropping*

(Escuta Telefônica) e até mesmo uma *VoIP Interception and Modification* (Interceptação e Modificação VoIP).

A principal diferença entre as duas técnicas supracitadas é que a primeira corresponde à escuta telefônica tradicional, em que o espião ou investigador de polícia apenas tem conhecimento sobre as conversas realizadas durante chamadas de voz, enquanto a segunda técnica consiste em interceptar e também alterar a conversação por meio de outras técnicas adicionais. Para fins de estudo, este artigo tem como finalidade abordar apenas a primeira técnica, ou seja, a escuta telefônica ou *Eavesdropping*.

4 CAPTURANDO O TRÁFEGO DE VOZ

Escuta telefônica ou *Call Eavesdropping*, conforme citado anteriormente, consiste em um método utilizado para monitorar a sinalização e o fluxo de dados entre dois ou mais *endpoints*, sem alterar a informação transmitida entre eles.

Esse tipo de ataque compromete a privacidade em ligações VoIP, onde o atacante descobre quem são os usuários do serviço, o que esses usuários estão conversando e quais teclas do *smartphone*, *softphone* ou telefone IP os usuários digitaram durante a conversação. Dessa forma, é possível descobrir senhas, números de documentos e demais informações sigilosas provenientes das partes comunicantes.

Para efetuar esse tipo de operação, o *cracker* precisa obter acesso a pontos chaves da rede a ser invadida e a *softwares* específicos para realizar a captura e reconstituição da mídia transmitida entre os usuários. Exemplos de *softwares* comumente utilizados nesses tipos de ataque são o *Wireshark* e o Voipong (*software* específico para aplicações VoIP).

4.1 SNIFFERS DE REDE

Um *sniffer* de rede é um programa que captura os pacotes trafegando em um segmento da rede e permite a visualização dos cabeçalhos dos protocolos utilizados na comunicação. Se a rede utilizar protocolos inseguros, o uso de ferramentas como essa pode facilitar ataques.

Na opinião de Frank Dzubeck, presidente da *Communications Ne-*

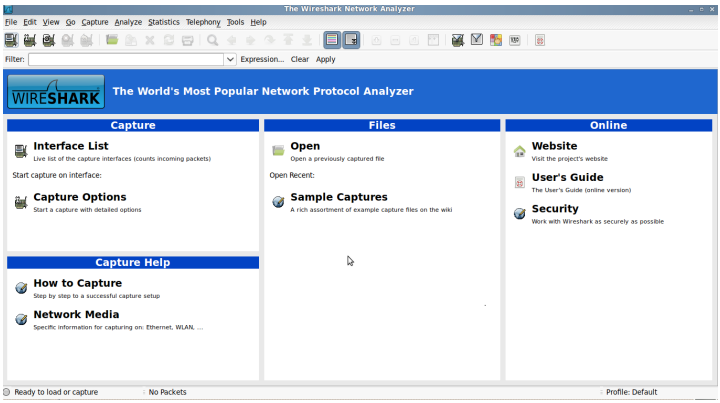
twork Architects, a principal falha de segurança em aplicações VoIP está no protocolo IP que, segundo ele, não foi construído com a segurança em pensamento. Se a aplicação VoIP possui vulnerabilidades, deveria haver uma camada de proteção que as minimizasse.

Além do protocolo IP, os protocolos utilizados na sinalização e transporte de chamadas telefônicas não são seguros, permitindo uma fácil reconstituição do áudio após a captura dos pacotes multimídia. O Voipong é um exemplo de *sniffer* de rede sofisticado e específico para detecção de chamadas VOIP, captura de pacotes e reconstituição de conversas em arquivos de áudio no formato WAV.

Aplicação leve e de fácil manipulação, o Voipong é uma ferramenta escrita em linguagem C que necessita apenas de 256 MB de RAM e utiliza em torno de 66 a 80% de processamento. Suporta os principais protocolos de sinalização, transporte e codificação relacionados a telefonia IP (como SIP, RTP, RTCP, G711) e foi desenvolvida para rodar em plataforma Linux, mas pode ser compilada e executada em outros sistemas operacionais também (MURAT, 2004).

Um outro *sniffer* de rede que não é específico para aplicações VOIP, mas contém funções que possibilitam capturar, filtrar e reconstituir conversas realizadas via telefonia IP é o *Wireshark*, visto na figura 3.

Figura 3 – O analisador de protocolos *Wireshark*

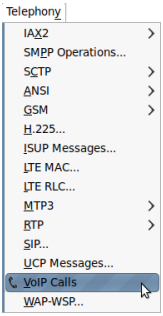


Fonte: print screen do Wireshark

Esse *sniffer* já é conhecido de longa data pela maioria dos *crackers* e administradores de rede. Ele possibilita capturar quaisquer tipos de mensagens

transmitidas em rede e, com a evolução das aplicações e convergência de dados, voz e vídeo para as redes de computadores, foi adicionado a ele um módulo (observe a opção *Telephony* na barra de menus do programa, figura 4) que permite trabalhar também com pacotes originados de transmissões multimídia.

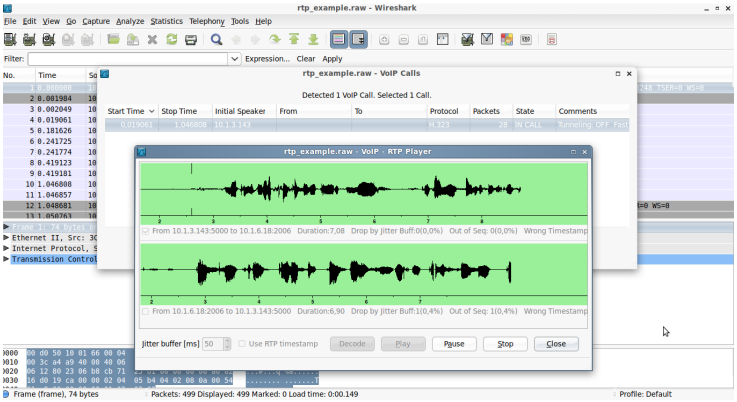
Figura 4 – Menu dedicado a análise de protocolos e pacotes multimídia



Fonte: print screen do menu telephony

Como se pode observar na figura 4, o *Wireshark* possui um menu dedicado à análise dos protocolos utilizados em transmissões multimídia. Na imagem foi destacada a opção *VoIP Call* porque através dela é possível detectar, ou seja, filtrar uma chamada VoIP no fluxo de pacotes capturados, selecionar essa chamada, decodificá-la e reproduzir a conversa capturada.

Figura 5 – Captura, detecção, decodificação e reprodução de chamada VoIP



Fonte: print screen do arquivo de chamada VoIP

A figura 5 apresenta a detecção, decodificação e reprodução de uma chamada VoIP disponibilizada em arquivo para *download* na *wiki* do *Wireshark* (*wiki.wireshark.org/VoIP_calls*). Qualquer pessoa pode fazer o *download* e realizar este mesmo teste em seu computador, utilizando o *Wireshark*. Com esse simples exemplo, dá para se ter uma ideia do quão rica é a ferramenta e o quanto as aplicações de telefonia IP estão vulneráveis a escutas telefônicas e outros tipos de ataques, dado o grande número de protocolos (SIP, H323, ISUP, MGCP, UNISTIM, RTP) suportados pela ferramenta.

A cada chamada VoIP, são listadas informações como início e término da chamada, endereço IP do *host* que iniciou a chamada, os campos *FROM* e *TO* dos protocolos de sinalização utilizados no estabelecimento e controle da ligação, o protocolo de sinalização utilizado, a quantidade de pacotes envolvida no diálogo, o estado da chamada (se está sendo realizada no momento, se foi cancelada, rejeitada, completada, etc.) e algum comentário adicional dependendo do protocolo utilizado.

É possível ainda gerar gráficos e verificar de forma simples os protocolos e as mensagens trocadas no estabelecimento da conexão e durante toda a comunicação, até que esta seja finalizada por alguma razão.

Além de *softwares*, existem telefones IP que possuem a função de captura de tráfego, como o modelo SNOM 320 (figura 6). Uma vez previamente configurado e inserido em uma rede local que utilize *hubs* em sua infraestrutura, esse equipamento de telefonia possibilita uma escuta telefônica semelhante às escutas telefônicas tradicionais.

Figura 6 – Telefone IP da marca SNOM, modelo 320

Fonte: site da empresa SNOM

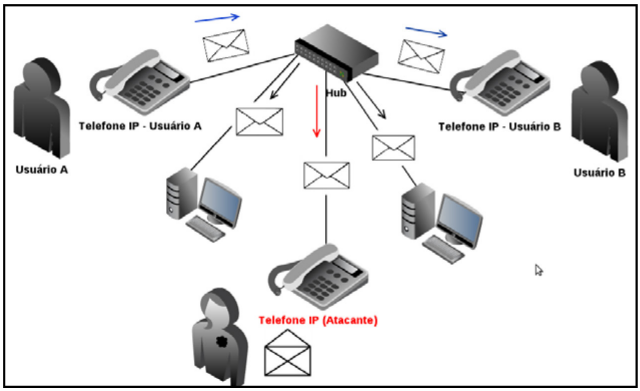
4.2 OBTENDO ACESSO A PONTOS ESTRATÉGICOS DA REDE LOCAL

Para que uma escuta telefônica seja possível, não é suficiente o uso de *sniffers* de rede. É necessário também o acesso a pontos chaves da rede local, para que o tráfego seja direcionado à máquina do atacante e em seguida, capturado, analisado e reconstituído através do uso de um dos *softwares* citados anteriormente.

Para tanto, o *cracker* mantém o foco nas peculiaridades dos equipamentos utilizados em redes locais, como *hubs* e *switches*, objetivando invadi-los. Existem várias técnicas para atacar esses dispositivos, a partir do estudo de sua estrutura interna de funcionamento e algumas falhas de configuração ou dos protocolos utilizados.

No caso dos *hubs*, por exemplo, que por padrão encaminham os pacotes em *broadcast* para a rede local e não apenas ao seu destinatário, a captura dos dados por um equipamento ou *software* que esteja na mesma rede física é extremamente simples. A figura 7 ilustra esse tipo de ataque.

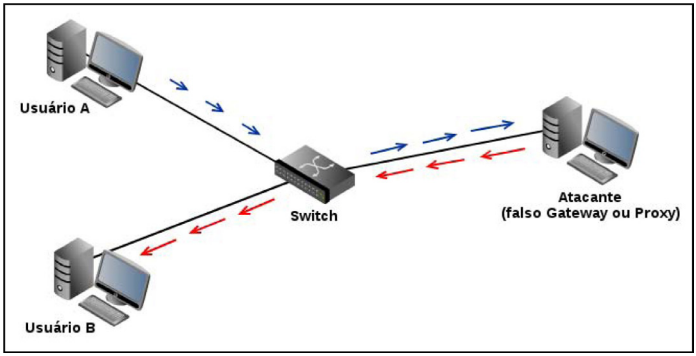
Figura 7 – *Call Eavesdropping* em redes utilizando *hubs*



Outra forma de ataque bem antiga, mas que pode ser utilizada em conjunto para possibilitar a escuta telefônica, consiste na técnica denominada *Man in the Middle*, em que o *cracker* envia diversos pacotes ARP alterados para a rede, informando que a máquina dele é o *gateway* ou servidor *proxy* da mesma. Assim, todos os pacotes serão encaminhados para o atacante antes

de serem enviados para o destinatário correto (como demonstra a figura 8), independente de se utilizar *switches* ou *hubs* nessa rede.

Figura 8 – O famoso ataque *Man in the Middle*



Com todo o tráfego de dados passando pela máquina do atacante antes de seguir seu destino, pode-se capturar os pacotes VoIP desejados a partir da análise dos campos *FROM* e *TO* do protocolo SIP, utilizando para isso um *sniffer* de rede como o *Wireshark*, por exemplo. Após a filtragem do tráfego e a captura dos pacotes desejados, pode-se reconstituir a conversa por meio de *softwares* específicos e finalmente realizar a escuta telefônica e/ou a gravação da conversa.

Como a maioria das redes atuais já não contém *hubs*, além do fato de alguns *switches* denominados *switches* gerenciáveis permitirem configurações para tornar a rede mais segura (FILIPPETTI, 2008), os *crackers* se concentraram em descobrir diferentes formas de atacar estes equipamentos.

4.2.1 VLAN e espelhamento de portas

Um exemplo de configuração avançada de *switches* consiste na criação de VLANs (*Virtual Local Area Networks*), de forma a dividir uma mesma rede física em diferentes redes lógicas utilizando *switches* gerenciáveis. Esse tipo de configuração permite segmentar a rede em diferentes domínios de *broadcast* sem a necessidade de um roteador, agrupando usuários e recursos em portas administrativamente definidas no equipamento. Assim,

diferentes conjuntos de dados trafegam apenas dentro das redes em que estão seus destinatários, melhorando a gerenciabilidade e aumentando a segurança da rede local.

Além da criação de VLANs, alguns *switches* possuem ainda um recurso chamado RSPAN (*Remote Switched Port Analyzer*) que permite espelhar todo o tráfego de portas ou VLANs determinadas em uma porta específica. Assim, ao contrário dos benefícios proporcionados pela criação de VLANs, um *switch* gerenciável que possua a opção de espelhamento de portas permite que qualquer equipamento ligado a essa porta possa capturar o tráfego das demais portas dos *switches* da rede, permitindo consequentemente a realização de uma escuta telefônica utilizando os mesmos métodos citados anteriormente (filtragem de tráfego, *sniffers* de rede e *softwares* específicos).

4.2.2 MAC Flooding Attack

Existem ainda outros tipos de ataques a *switches*, como o *MAC Flooding Attack*, onde o tráfego é gerado de forma aleatória e originado de vários endereços MACs falsos, preenchendo totalmente a tabela ARP do equipamento. Quando a tabela ARP é totalmente preenchida, o *switch* passa a atuar como *hub* e pode-se efetuar o ataque *Man in the Middle*, descrito anteriormente.

4.2.3 Spanning Tree Protocol Attack

O protocolo STP (*Spanning Tree Protocol*), presente em qualquer *switch* sendo ele gerenciável ou não, atua evitando *loops* nos *enlaces* entre *switches* interligados entre si e também é alvo de ataques.

Um *loop* pode ocorrer quando os *switches* são inicializados e suas tabelas MAC ainda não estão preenchidas ou quando ambos os *switches* não possuem uma entrada em sua tabela MAC para o *host* de destino, sendo necessário o envio de mensagens em *broadcast*. Tais mensagens podem causar *broadcast storm* (ou tempestade de *broadcast*).

O *broadcast storm* ocorre porque a mensagem é enviada em todas as portas (exceto a porta de origem) do primeiro *switch*. Considerando que esse *switch* está ligado a um segundo *switch*, esse segundo *switch* que possui uma segunda porta ligada ao primeiro *switch* (formando um *loop* físico),

irá enviar o pacote de *broadcast* de volta para o primeiro *switch*, mas agora por outro caminho/porta, causando o *loop* lógico. Se a máquina destino da mensagem estiver desligada, esse *loop* se repetirá indiscriminadamente, impossibilitando o funcionamento da rede local, pois nenhum dos *switches* encontrará a máquina destino e continuará a enviar os quadros em *broadcast*.

Embora possam causar *loops*, cenários em que dois ou mais *switches* são interconectados entre si possibilitam mais de um caminho para o tráfego de quadros *Ethernet*, proporcionando redundância, de forma que se um *switch* falhar, a rede continuará em funcionamento. Por isso, estruturas como essa são frequentemente utilizadas em redes locais e, para evitar *loop*, é utilizado o protocolo STP que também está na mira de ataques aos *switches*, sejam eles gerenciáveis ou não.

Antes de explicar o ataque em si, é importante conhecer um pouco sobre o princípio de funcionamento do protocolo *Spanning Tree* para entender melhor como tudo acontece. Para evitar um *loop* na rede, o protocolo STP elege um *switch root*, pelo qual todo o tráfego irá passar. Esse *switch root* é escolhido comparando-se o *Bridge ID* (campo específico dos quadros transmitidos na rede local pelos switches, os BPDUs) e escolhendo o *switch* que apresente o *Bridge ID* com menor prioridade. Caso os valores de prioridade sejam iguais em todos os *switches*, o *switch root* será escolhido de acordo com seu endereço MAC, sendo o menor o escolhido.

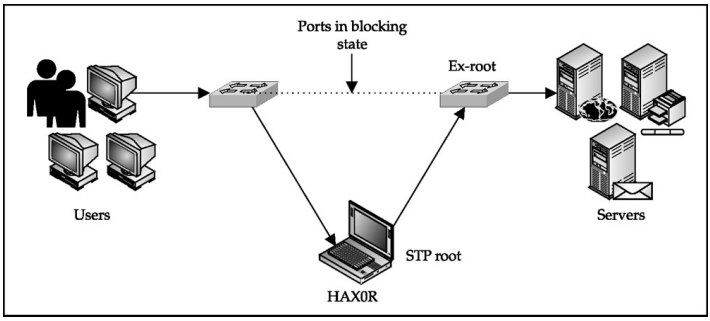
Uma vez escolhido o *switch root*, os demais *switches* da rede deverão eleger uma porta *root* (porta com o menor custo para chegar ao *switch root*). Após determinar as portas designadas, a árvore STP está criada. As demais portas que proporcionam a redundância entre os *switches* são bloqueadas, permitindo que apenas um caminho seja usado na transmissão de quadros entre *switches* e impedindo que ocorram *loops* na rede local. Essas portas bloqueadas só serão ativadas caso a porta designada ou algum *switch* da estrutura pare de funcionar, proporcionando resiliência à rede.

A partir do conhecimento sobre o funcionamento do STP, um tipo de ataque utilizado pelos *crackers*, conhecido como *Spanning Tree Protocol Attack*, é composto por um computador com duas placas de rede (chamado também de *multihomed*) que por sua vez estão ligadas em dois *switches* diferentes.

O invasor utiliza essa máquina para gerar BPDUs (*Bridge Protocol Data Units*) com baixa prioridade. Como o STP reconhece que o *switch* tem

mais de um caminho ou *enlace* físico para se comunicar com um nó da rede, esse protocolo determina o melhor caminho e bloqueia os demais, com base no valor *Bridge ID* das BPDUs. Nesse caso, a máquina atacante engana o protocolo STP fazendo com que os dois *switches* elejam-no como porta *root*, conforme ilustrado na figura 9.

Figura 9 – *Spanning Tree Protocol Attack*



Fonte: Hacking Exposed Cisco Networks – Cisco Security Secrets & Solutions

Assim, todo o tráfego passará pela máquina do atacante antes de ser encaminhado ao destino correto, tornando possível uma escuta telefônica.

Enfim, esses foram apenas alguns exemplos de ataques a redes locais. Uma vez que o atacante tenha acesso à rede e consiga direcionar o tráfego para sua máquina usando uma das técnicas descritas nesta seção, ele poderá filtrá-lo, capturar apenas os pacotes da conversa desejada e extrair a sinalização, a mídia e os dígitos DTMF (*Dual-Tone Multi-Frequency*, conhecidos em inglês como *touch tones* ou tons de discagem) dos pacotes capturados com o auxílio de um *sniffer* de rede, realizando assim a escuta telefônica.

5 PROTOCOLOS E SEGURANÇA NA TELEFONIA IP

Diante de tantas técnicas e ferramentas para invadir e realizar uma escuta telefônica fica claro que o maior problema está relacionado à falta de segurança dos protocolos utilizados na telefonia IP. Na maioria das comunicações VoIP são usados os protocolos padrão SIP e RTP/RTCP, além dos protocolos de codificação/decodificação de áudio e vídeo. Tais protocolos não foram construídos pensando-se na segurança.

No entanto, existem diversos protocolos desenvolvidos especificamente para prover segurança às transmissões multimídia, além dos protocolos pré-existentes ao surgimento do VoIP e que também podem ser aplicados com o intuito de tornar segura a comunicação por meio da telefonia IP. Nesta seção serão abordadas algumas peculiaridades dos protocolos padrão utilizados nesses tipos de transmissões e as respectivas soluções às suas falhas de segurança.

5.1 SIP

O SIP (*Session Initiation Protocol*) é um protocolo de sinalização responsável pela criação, modificação e finalização de sessões de comunicação multimídia. É usado tanto em conferências *multicast* quanto em chamadas telefônicas utilizando a telefonia IP.

Inicialmente especificado pela RFC 2543 (atualmente obsoleta), hoje é possível encontrar uma descrição completa do SIP na RFC 3261. Trata-se um protocolo de controle que atua na camada de aplicação do modelo OSI (*Open Systems Interconnection*) e utiliza o modelo cliente-servidor, onde o dispositivo solicitante é o cliente e o dispositivo requisitado é o servidor. O SIP não se preocupa com o tráfego de dados, mas sim com a interoperabilidade entre os dispositivos durante a transmissão de dados.

A arquitetura do protocolo SIP é baseada em quatro entidades lógicas: *User Agent* (UA), *Proxy Server*, *Redirect Server* e *Registrar*.

O primeiro, o UA, é dividido em duas partes: uma aplicação cliente (*User Agent Client* ou UAC) que inicia o pedido SIP e uma aplicação servidora (*User Agent Server* ou UAS) que localiza um usuário ao receber uma mensagem SIP e responde a ele.

O *Proxy Server* age simultaneamente como servidor e cliente, realizando pedidos em nome de um cliente. Ele lê, interpreta e se preciso, reescreve uma mensagem para depois enviá-la.

O *Redirect Server* possui a função de responder a um pedido redirecionando-o à nova localidade do usuário de destino da mensagem.

A última entidade lógica da arquitetura SIP é o *Registrar*. Ele trabalha em conjunto com o *Redirect Server*, sendo ele o responsável por atualizar o banco de dados que contém a localização dos usuários.

Além da arquitetura, o protocolo SIP possui características como a troca de mensagens em texto puro; o cabeçalho; mensagens de status; regras de codificação; requisições e respostas herdadas do protocolo HTTP (*HyperText Transfer Protocol*).

Um resumo dos seis métodos que podem ser utilizados no cabeçalho SIP estão listados na tabela abaixo.

Tabela 1 – Métodos SIP

Método	Descrição
INVITE	Solicita o início da comunicação
ACK	Confirma o início da comunicação
OPTIONS	Consulta os recursos do destinatário
REGISTER	Informa a localização atual do destinatário
CANCEL	Cancela uma requisição pendente
BYE	Termina a comunicação entre os UAs

Dentre os muitos campos encontrados em cabeçalhos de mensagens SIP, alguns campos que são encontrados tanto em mensagens de requisição quanto em mensagens de resposta, estão os seguintes:

- *FROM*: contém o remetente da mensagem (nome e endereço).
- *TO*: indica o destinatário da mensagem.
- *Call-ID*: identifica uma comunicação em particular ou todos os registros de um cliente particular. Por exemplo, em uma videoconferência um novo *Call-ID* será gerado para cada usuário convidado.
- *Cseq*: contém um número escolhido de forma aleatória e que é incrementado a cada nova requisição para organizar as transações em um diálogo, diferenciando novas solicitações de pedidos de retransmissão.
- *Via*: tem a função de indicar o caminho tomado pela mensagem de solicitação até o momento e o caminho que deve ser seguido no encaminhamento das respostas à tal solicitação.

Pelo fato de possuir uma estrutura de mensagem derivada do protocolo HTTP (*HyperText Transfer Protocol*); largamente utilizado na *web*; todos os mecanismos de segurança utilizados para plataformas HTTP podem ser aplicados

também às sessões SIP. Além disso, o SIP também herda algumas características do protocolo SMTP (*Simple Mail Transfer Protocol*), podendo também utilizar os mecanismos de segurança adotados na proteção de *e-mails*.

As formas de autenticação via HTTP podem ser usadas pelo SIP. No entanto, autenticação e não-repúdio não impedem que haja uma escuta telefônica. Para isso, pode-se utilizar o TLS (*Transporte Layer Security*), que adiciona uma camada de segurança entre a aplicação e a camada de transporte, realizando autenticação e criptografia simultaneamente.

O uso em conjunto de TLS e SIP recebe uma nova denominação: SIPS URI. Assim como páginas *web* seguras utilizam o protocolo HTTP com uma camada de segurança, nomeando-o HTTPS; com o protocolo SIP ocorre o mesmo: é utilizada a nomenclatura SIPS nas trocas de mensagens usando TLS.

No entanto, utilizar TLS é mais caro que enviar os pacotes usando apenas os protocolos de transporte comumente utilizados, pois com a adoção do TLS, o desempenho das máquinas cai devido ao processamento requerido por este protocolo. Além disso, o uso de TLS requer o uso do TCP (*Transport Control Protocol*) como protocolo de transporte e necessita de uma infraestrutura de chave pública para funcionar.

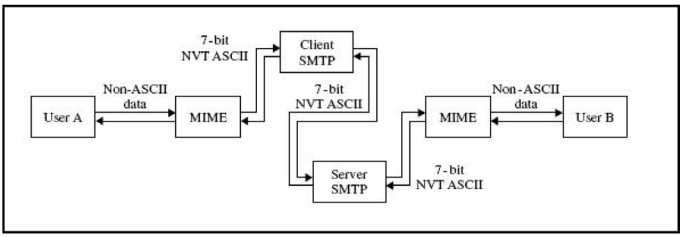
Por isso muitas empresas deixam de implementar essa camada de segurança visando à economia a curto prazo, sem imaginar que a longo prazo, essa falta de segurança pode lhes causar sérios prejuízos.

Outra solução com o objetivo de prover segurança a aplicações VoIP utilizando o protocolo SIP é aplicar os protocolos e técnicas de segurança utilizados na comunicação por correio eletrônico. Como o protocolo SIP utiliza o MIME (*Multipurpose Internet Mail Extensions*), é possível utilizar a versão segura desse protocolo na telefonia IP.

Quando o correio eletrônico foi criado, utilizando-se o protocolo SMTP (*Simple Mail Transfer Protocol*), não era possível enviar *e-mails* contendo imagens, vídeos, áudios e mensagens de texto em outras línguas que não a língua inglesa. Além disso, apenas o envio de textos contendo um número limitado de caracteres no formato ASCII era possível. O MIME surgiu em 1992 para extinguir essas limitações do sistema de *e-mail* e corresponde a um protocolo que padroniza a codificação e interpretação de arquivos em diversos formatos, como imagens, vídeos, arquivos binários e textos com conteúdo diferente do ASCII. O MIME trabalha em conjunto com o protocolo SMTP (figura 10),

convertendo conteúdos não-ASCII em dados ASCII para que eles possam ser transmitidos pelo SMTP.

Figura 10 – Uso do MIME em conjunto com o protocolo SMTP



Fonte: GTA/UFRJ – Grupo de Teleinformática e Automação da UFRJ

Da mesma forma que codifica e decodifica conteúdos enviados por *e-mail*, o MIME age codificando e decodificando mensagens enviadas por meio da telefonia IP.

A versão segura desse protocolo, o S/MIME (*Secure/ Multipurpose Internet Mail Extensions*) protege o conteúdo presente em comunicações via *e-mail* e/ou via telefonia IP de falsificações e da interceptação não autorizada de mensagens. O S/MIME possui as seguintes funções:

- Envelopamento de dados: permite criptografar qualquer tipo de dados e chaves, provendo privacidade e confidencialidade à comunicação.
- Assinatura de dados: consiste em assinar digitalmente uma mensagem. O conteúdo assinado é codificado, limitando sua visualização àqueles destinatários que também utilizam S/MIME.
- Assinatura de dados limpa: tem a mesma função da assinatura de dados especificada acima, porém a visualização do conteúdo da mensagem independe de o destinatário usar ou não o protocolo S/MIME. No entanto, ele não poderá verificar a assinatura da mensagem e consequentemente, não poderá constatar se o requisito não-repúdio foi assegurado.
- Assinatura e envelopamento de dados: é a combinação da criptografia de dados com a assinatura dos mesmos. Neste caso, tanto é possível encriptar dados assinados, como é possível assinar dados encriptados.

O S/MIME apresenta funcionalidades semelhantes ao protocolo PGP (*Pretty Good Privacy*), que também é usado como solução de segurança para correios eletrônicos e pode ser adotado no campo da telefonia IP. Ambos oferecem a possibilidade de autenticação e criptografia de dados. O protocolo *Pretty Good Privacy*, criado por Phil Zimmerman em 1991 e inicialmente gratuito, tornou-se proprietário após sua compra pela *Network Associates* em 1996. De acordo com SINGH (2009), atualmente existem versões comerciais e *freeware* disponíveis para uso.

O PGP é baseado na distribuição de chaves públicas a partir do próprio usuário, de um servidor ou de um terceiro confiável. Cada usuário é livre para decidir como irá distribuir sua chave e em que situações irá considerar confiáveis as chaves recebidas de outros usuários. Para a transmissão de mensagens seguras, o PGP atua da seguinte forma:

- Uma chave de sessão é criada de forma aleatória por um algoritmo simétrico;
- A mensagem é criptografada utilizando a chave de sessão criada (proporcionando confiabilidade à comunicação);
- A chave de sessão é criptografada utilizando-se a chave pública do destinatário;
- É gerado um resumo ou *hash* da mensagem. Esse *hash* é assinado utilizando-se a chave privada do remetente (proporcionando integridade e autenticidade);
- A chave de sessão encriptada é anexada à mensagem criptografada e ao *hash*;
- Finalmente, a mensagem é enviada ao destinatário.

Uma vez transmitida, é necessário realizar os passos inversos para recuperar e validar o conteúdo da mensagem no destinatário.

Existem outros protocolos e técnicas de segurança que podem ser adotadas para garantir a confidencialidade na telefonia IP, assim como existem outros protocolos que podem ser utilizados em substituição ao SIP. Porém como o SIP e o RTP (em conjunto com o RTCP) são os protocolos mais largamente utilizados em comunicações multimídia, eles são o foco deste estudo. Portanto, a próxima seção faz um breve relato sobre o funcionamento desse outro protocolo e traz soluções adicionais para garantir a segurança na transmissão de dados multimídia.

5.2 RTP E RTCP

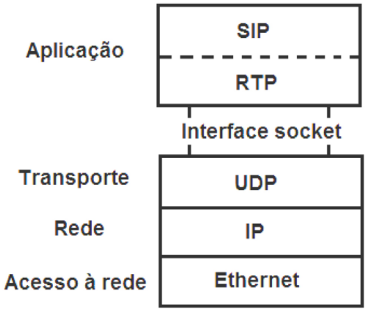
O RTP (*Real-time Transport Protocol*) é um protocolo utilizado em aplicações multimídia de em tempo real, sendo adotado nas transmissões de rádio pela Internet, telefonia IP, música sob demanda, videoconferências, vídeo sob demanda e outros serviços que estão em fase de convergência para a grande rede mundial de computadores.

Ele foi criado como um protocolo genérico para ser usado nas transmissões multimídia e resolver aspectos de compatibilidade, pois para cada aplicação os fabricantes estavam desenvolvendo um protocolo diferente, impossibilitando a interoperabilidade entre *softwares* e dispositivos, além de criar um grande número de protocolos diferentes que realizavam praticamente as mesmas funções (TANENBAUM, 2003).

Apesar de ser um protocolo dedicado ao transporte de informações em tempo real, o RTP não substitui o uso dos protocolos de transporte comumente utilizados, como o UDP (*User Datagram Protocol*) e o TCP (*Transmission Control Protocol*). Pelo contrário, o RTP trabalha em conjunto, principalmente com o UDP, já que aplicações de tempo real exigem uma maior velocidade nas transmissões em detrimento da confiabilidade.

Como exerce funções de transporte, mas está implementado na camada de aplicação, a posição do RTP na pilha de protocolos TCP/IP é um tanto confusa TANENBAUM (2003). Ele está na camada de aplicação, localizado entre esta e a camada de transporte, conforme apresenta a figura 11.

Figura 11 – Posição do RTP na pilha de protocolos TCP/IP



Como pode ser observado na figura 11, o protocolo UDP encapsula a mensagem RTP através do *socket* que possibilita a comunicação entre a camada de aplicação e a camada de transporte.

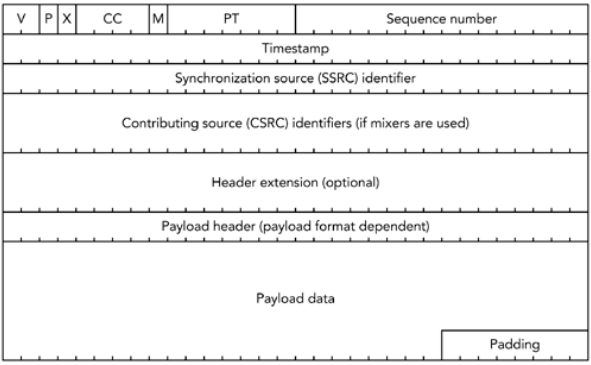
Ainda considerando a figura 11, pode-se descrever o funcionamento do RTP em conjunto com os protocolos apresentados na pilha TCP/IP acima da seguinte forma: o fluxo de mensagens multimídia gerado na camada de aplicação é armazenado na biblioteca RTP, que também está na camada de aplicação. Tal biblioteca realiza a multiplexação desses fluxos e os codifica em mensagens RTP que em seguida são colocadas no *socket*. Após atravessar o *socket*, o protocolo UDP, presente na camada de transporte, encapsula as mensagens RTP, transformando-as em *datagramas* UDP que serão encapsulados novamente na camada de rede e transformados em *datagramas* IP. Considerando que o dispositivo transmissor esteja em uma rede Ethernet, os *datagramas* IP serão inseridos em quadros *Ethernet* para transmissão (TANENBAUM, 2003).

No dispositivo receptor, é feito o processo de desencapsulamento normalmente realizado para qualquer informação transmitida utilizando a pilha de protocolos TCP/IP e a decodificação do fluxo de mensagens multimídia.

Dessa forma, a função básica do RTP é multiplexar os diversos fluxos de mensagens de tempo real em um único fluxo de *datagramas* UDP, enviando-o à um (unidifusão) ou vários (multidifusão) destinatários. Os pacotes não recebem tratamento especial pelos roteadores devido ao uso do protocolo UDP, não havendo também controle de fluxo, controle de erros, confirmação de entrega e consequentemente, retransmissões.

O cabeçalho RTP consiste em três palavras de 32 bits e possui a opção de adicionar extensões, como mostra a figura 12.

Figura 12 – Cabeçalho do protocolo RTP.



Fonte: LISHA - Software/Hardware Integration Lab, UFSC

- O primeiro campo, *Version* (V), indica a versão do protocolo. Atualmente utiliza-se a versão 2.
- O segundo campo é o *Padding* (P) que quando ativado, indica que o pacote foi preenchido com conteúdo adicional, apenas para completar o tamanho mínimo requerido em algumas situações.
- O campo *Extensions* (X) indica se existem extensões a serem adicionadas no cabeçalho do pacote.
- O quarto campo, designado *CSRC Count* (CC) apresenta o número de campos de identificadores CSRC existentes no cabeçalho. Esse campo é usado quando diferentes fluxos RTP são unificados, formando um só fluxo.
- O campo *Marker* (M) indica um nível de importância especial que determinado pacote têm para uma aplicação em uso, podendo indicar, por exemplo, um surto de fala após um período de silêncio.
- O *Payload Type* (PT) apresenta informações sobre o tipo de conteúdo transportado. Como o RTP é um protocolo de uso genérico, que pode transportar mídias codificadas em diversos formatos, é importante saber o formato utilizado por determinado conteúdo para que a aplicação utilize o codec adequado em sua reprodução.
- *Sequence Number* ou número de sequência: é um número incrementado em cada pacote RTP enviado a fim de detectar perda ou sequência de pacotes.

- *Timestamp*: indica quando o primeiro *byte* do *payload* do pacote foi amostrado, permitindo a reprodução da amostra no tempo adequado.
- *Synchronization Source Identifier* (SSRCI): número escolhido aleatoriamente que serve para identificar o emissor do pacote RTP.
- *CSRC Contributing Source Identifier*: pode existir de 0 a 15 campos como este em um pacote RTP, porém ele só estará presente se o pacote RTP for enviado por um *mixer*, que reúne pacotes RTP originados de diferentes emissores.

O RTP trabalha em conjunto com um outro protocolo, o RTCP (*RTP Control Protocol*), que possibilita o envio de informações relacionadas a qualidade dos serviços em uma sessão RTP. Dados estatísticos como número de pacotes enviados e recebidos, número de pacotes perdidos e *jitter* são enviados em forma de relatórios aos emissores e receptores presentes na comunicação.

Da análise do cabeçalho RTP, sua estrutura e funcionamento, além da breve análise acerca da principal função do RTCP, percebe-se que esses dois protocolos também não foram construídos pensando-se em segurança. Para tanto, foram criadas versões seguras de ambos os protocolos, o SRTP (*Secure Real-Time Transport Protocol*) e o SRTCP (*Secure Real-time Transport Control Protocol*) para proverem segurança ao RTP e ao RTCP respectivamente.

O SRTP (*Secure Real-time Transport Protocol*) é uma alternativa segura ao uso do RTP que oferece confidencialidade, autenticação de mensagens e proteção ao tráfego RTP. As funcionalidades do SRTP não causam grande *overhead* nos pacotes transmitidos e por isso não interferem nas taxas de transferência dos mesmos. Devido a isso, de acordo com a RFC 3711 o SRTP é uma solução de segurança adequada às redes heterogêneas (que possuem infraestrutura cabeada e sem fio).

Com o uso do SRTP, apenas a área de dados do pacote RTP é criptografada, junto com algum preenchimento, caso exista. São adicionados opcionalmente alguns campos como o *Master Key Identifier* (MKI) que identifica a partir da qual são geradas as demais chaves de sessão utilizadas. No SRTP, o número de sequência, presente de forma nativa no protocolo RTP é usado em conjunto com o *Rollover Counter* (CC ou contador de rolagem), usado nas sessões SRTP para evitar ataques de repetição. Há ainda o campo

Authentication tag que, como o próprio nome sugere, serve para prover autenticação, protegendo os dados de possíveis modificações não autorizadas.

No protocolo SRTCP, o uso de autenticação é obrigatório e não opcional, como acontece com o SRTP. O campo MKI também está presente nesse protocolo e possui a mesma função exercida no protocolo anterior. Além desses campos, há um novo campo denominado *SRTCP índex*, que é usado como um contador de sequência para evitar ataques de repetição, trabalhando de forma similar aos campos *Rollover Counter* e *Sequence Number* do protocolo SRTP.

Além das versões seguras dos protocolos utilizados em comunicações multimídia como a telefonia IP, pode-se adicionar ainda uma nova camada de segurança, adotando o uso do IPsec (*IP security*), uma versão segura do protocolo IP. Diferentemente dos protocolos abordados até o momento nesta seção, o IPsec não é um protocolo exclusivo de comunicações VOIP, mas sim uma estrutura antiga que provê sigilo, integridade dos dados e proteção contra ataques de reprodução (onde o atacante pode reproduzir uma conversa, por exemplo).

O IPsec atua na camada de rede, encapsulando partes do *datagrama* IP (modo de transporte) ou o *datagrama* IP por inteiro (modo de túnel). A criptografia é o forte desse modelo que proporciona autenticação, confiabilidade e confidencialidade. Nem mesmo o endereço IP original é conhecido pelos demais dispositivos da rede. Apenas na origem e no destino os pacotes são encapsulados e desencapsulados respectivamente, deixando à mostra seus endereços de origem e destino.

Apesar de prover segurança na camada de rede de forma transparente ao usuário, isso não significa que o IPsec substitui os protocolos de segurança de outras camadas, como a de transporte e aplicação. Ao contrário: combinar as diversas soluções de segurança proporciona redundância em requisitos como confiabilidade, autenticidade e integridade dos dados, além de outros aspectos de extrema importância para diminuir os riscos nas comunicações em rede.

6 CONSIDERAÇÕES FINAIS

Apesar de a telefonia IP possuir suas próprias especificidades, é evidente que pelo fato de ser uma aplicação sobre a rede IP, está suscetível a diversos tipos de ataques. Embora também relevantes, estes referidos ataques podem

ser proferidos a qualquer tipo de serviço e por isso optou-se por apresentar aqui um tipo de ataque em especial: o de escuta telefônica, o qual pode ser realizado tanto na rede de telefonia tradicional, quanto na rede de telefonia IP abordada neste trabalho.

Para tanto, apesar de algumas facilidades proporcionadas pela digitalização dos dados e avanços tecnológicos na área de telecomunicações, escutas telefônicas em redes comutadas por pacotes são mais fáceis de serem realizadas em redes locais, onde o atacante faça parte da mesma rede lógica e seja capaz de realizar esse tipo de ataque invadindo os dispositivos comutadores de pacotes da rede a qual pertence.

Uma vez transportados pela Internet, a interceptação de chamadas telefônicas se torna mais complicada, exigindo do atacante conhecimentos e técnicas mais elaboradas. Ainda assim, a comunicação não estará livre de intrusos pelo caminho.

A única forma garantir a segurança dos pacotes ao longo do caminho é utilizar protocolos de criptografia e autenticação de dados nas várias camadas do modelo OSI ou TCP/IP, como já mencionado anteriormente, além, é claro, de configurar corretamente os servidores de rede responsáveis por prover a comunicação multimídia através da rede IP. Assim, garante-se a segurança dos dados durante todo o processo de comunicação.

REFERÊNCIAS

BEZERRA, R. M. S. **Um estudo do protocolo SIP e sua utilização em redes de telefonia móvel**. UNIFACS.

BRASIL. **Lei n. 9.296**, de 24 de julho de 1996. Regulamenta o inciso XII, parte final, do art. 5° da Constituição Federal.

CARDOSO, L. C. M. **Central Telefônica, Nosso Passado**. Disponível em: <http://www.focadoemvoce.com/paramirim/curiosidades/centraltelefonica.php>. Acesso em: 13 ago 2012.

COLCHER, S., et al. **VoIP**: Voz sobre IP. Rio de Janeiro: Elsevier, 2005.

DEAN, T. **Network + Guide to Networks**. 5 ed. Boston: Course Technology. 2010.

DETTONI, F.; FLOR, J. P.; SOLDI, R. **RTP - Real-Time Transport Protocol**. 2009. Disponível em: <http://www.lisha.ufsc.br/teaching/dos/ine5424-2009-2/work/g2/rtp/rtp.html>. Acesso em: 15 ago 2012.

DISCÍPULOS de Graham Bell. **Profissões extintas**. 2011. Disponível em: <http://www.discipulosdegrahambell.tecnologia.ws/?p=42>. Acesso em: 13 ago 2012.

ENDERUNIX PROJECT. **What is VoIPong**. 2004. Disponível em: <http://www.enderunix.org/voipong/>. Acesso em: 04 jul 2012.

FARIAS, P. C. B. **STP (Spanning Tree Protocol)**. 2006. Disponível em: <http://www.juliobattisti.com.br/tutoriais/paulocfarias/redesbasico021.asp>. Acesso em: 11 ago 2012.

JOHNSTON, A. B. **SIP**: Understanding the Session Initiation Protocol. 2 ed. Norwood: Artech House. 2004.

KUROSE, J. F.; ROSS, K. **Redes de Computadores e a Internet**. 5 Ed., Editora Pearson, 2010.

MADEIRA, F. T. T. **Segurança em redes de voz sobre IP**. Olinda, 2007. Monografia de graduação - Associação de Ensino Superior de Olinda.

MANOJ, B. et al. **Fundamentos de VOIP**. 2 ed. Porto Alegre: Bookman Companhia ED. 2008.

NETWORK WORKING GROUP. **Request for Comments**: 3261. Disponível em: <http://tools.ietf.org/html/rfc3261>. Acesso em: 14 ago 2012.

NETWORK WORKING GROUP. **Request for Comments:** 3550. Disponível em: <http://www.ietf.org/rfc/rfc3550.txt>. Acesso em: 04 jul 2012.

NETWORK WORKING GROUP. **Request for Comments:** 3711. Disponível em: <http://www.ietf.org/rfc/rfc3711.txt>. Acesso em: 04 jul 2012.

NÓBREGA, J. **Segurança VoIP**. 2009. Disponível em: <http://www.computerworld.com.pt/>. Acesso em: 04 jul 2012.

ROSS, J. **VOIP**: Voz sobre IP. Rio de Janeiro: Antenna Edições Técnicas Ltda, 2007.

RZEUSNET. **História do Cinema** - Parte 4 (1881/1890). 2011. Disponível em: <http://rogercinema.blogspot.com.br/2011_08_01_archive.html>. Acesso em: 11 ago 2012.

SINGH, B. **Network Security and Management**. 2 ed. Nova Delhi: PHI. 2009.

SNOM VOIP PHONES. **snom 320 - SIP based IP phone**. Disponível em: <http://www.snom.com/en/products/ip-phones/snom-320/>. Acesso em: 11 ago 2012.

STEFFEN, A.; KAUFMANN, D.; STRICKER, A. **SIP Security**. E-Science und Grid, Ad-hoc-Netze, Medienintegration. 2004.

TANENBAUM, A. S. **Redes de Computadores**. 4 ed. Rio de Janeiro: Editora Campus. Tradução: Vandenberg D. de Souza.

VAZ. S. **MIME**. Disponível em: <http://www.gta.ufrj.br/>. Acesso em: 15 ago 2012.

VLADIMIROV, A. A. et al. **Hacking Exposed Cisco Networks**. 2006. Disponível em: <http://www.iphelp.ru/faq/35/0001.html>. Acesso em: 11 ago 2012.

VOIP Calls. Disponível em: <http://wiki.wireshark.org/VoIP_calls>. em: 14 ago 2012.