UNIVERSIDADE FEDERAL DE PERNAMBUCO

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA

2008.2

ANDRÉ GUEDES LINHARES

INTERCEPTAÇÃO LEGAL DE CHAMADAS VOIP BASEADAS EM SIP

TRABALHO DE GRADUAÇÃO

ORIENTADOR: DJAMEL FAWZI HADJ SADOK

Recife, PE
2008

UNIVERSIDADE FEDERAL DE PERNAMBUCO
GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA

ANDRÉ GUEDES LINHARES

INTERCEPTAÇÃO LEGAL DE CHAMADAS VOIP BASEADAS EM SIP

Trabalho de Graduação apresentado ao Centro de Informática da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Bacharel em Ciência da Computação.

ORIENTADOR:

Prof. PhD. Djamel Fawzi Hadj Sadok

Recife, PE
2008

UNIVERSIDADE FEDERAL DE PERNAMBUCO
GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA

FOLHA DE APROVAÇÃO

ANDRÉ GUEDES LINHARES

INTERCEPTAÇÃO LEGAL DE CHAMADAS VOIP BASEADAS EM SIP

Aprovado em 01 de Dezembro de 2008.

Banca Examinadora:

_____

DJAMEL FAWZI HADJ SADOK
(ORIENTADOR)

_____

NELSON SOUTO ROSA
(AVALIADOR)

INTERCEPTAÇÃO LEGAL DE CHAMADAS VOIP BASEADAS EM SIP

RESUMO

Nos últimos anos a tecnologia VoIP tem sido utilizada como uma alternativa para realizar chamadas telefônicas. Entretanto, questões como segurança, tarifação, números de emergência e interceptação legal ainda estão sendo pesquisadas. A interceptação legal de chamadas telefônicas, conhecida popularmente como "grampo telefônico", vem sendo bastante utilizada pela justiça como instrumento de investigação. Embora exista um procedimento bem estruturado, maduro e eficaz para a realização de interceptações legais de chamadas no sistema de telefônico tradicional, na tecnologia VoIP procedimentos estão sendo propostos. Neste contexto, este trabalho propõe uma arquitetura capaz de realizar a interceptação legal de chamadas VoIP baseadas no protocolo SIP. Como prova dos conceitos desenvolvidos foram realizadas uma implementação e avaliação de desempenho do sistema proposto.

VOIP LAWFUL INTERCEPTION BASED ON SIP

ABSTRACT

In recent years, VoIP technology has been wide used as an alternative to perform telephone calls. However, issues like security, billing, emergency calls and lawful interception are still been researched. The lawful interception, also known as 'wiretapping', has been used by the justice as an investigation tool. Although there is a well structured, mature and effective procedure to perform a lawful interception on Public Switched Telephone Network, on VoIP procedures are being proposed. In this context, this work proposes an architecture to perform a lawful interception on VoIP calls based on SIP protocol. As a proof of concept, they were carried out an implementation and a performance evaluation of the proposed system.

AGRADECIMENTOS

Primeiramente, agradeço a Deus pela realização deste trabalho e aos meus Pais pelo apoio e incentivo que recebi durante a minha graduação. Gostaria de agradecer também a minha namorada, Thaisa Farias, responsável pelo desenho de todas as imagens presentes neste trabalho.

Agradeço ao Professor Djamel Sadok pela orientação e ao pessoal do Grupo de Pesquisas de Redes e Telecomunicações da UFPE por possibilitar o acesso a equipamentos para realização dos experimentos.

E, finalmente, gostaria de agradecer a todos os amigos que me acompanharam na graduação: Ademir José Carvalho Júnior (tangoso), Bruno Filipe de Oliveira Lins (pigmeu), David Levy Lucena Alves Aragão, Felipe Cavalcanti Ferreira (mozinho), Fernando Valente Kakimoto (o japa), Francisco Paulo Magalhães Simões (chicrito), Henrique Seabra Diniz, Rilter Tavares Nascimento (rilter seabra), Rodrigo Diego Melo Amorim (digão), Thiago de Barros Lacerda (xeroso), Jesus Sanchez-Palencia, Rebeka Gomes, e Filipe César (gravatá).

ÍNDICE

LISTAS DE FIGURAS

LISTAS DE TABELAS

LISTA DE SÍMBOLOS E SIGLAS

3GPP - 3rd Generation Partnership Project

AES - Advanced Encryption Standard

ARPANET - Advanced Research Projects Agency Network

CALEA - Communication Assistance for Law Enforcement Act

CCS - Common Channel Signaling

DRM - Digital Rights Management

DTMF - Dual-Tone Multi-Frequency

EPABX - Electronic Private Automatic Branch

IAX2 - Inter-Asterisk eXchange 2

IETF - Internet Engineering Task Force

IP - Internet Protocol

ITU - International Telecommunication Union

LSB - Least Significant Bit

MGC - Media Gateway Controller

MGCP - Media Gateway Control Protocol

MOS - Mean Opinion Score

NAT - Network Address Translation

NVP - Network Voice Protocol

PABX - Private Automatic Branch Exchange

PAMS - Perceptual Analysis Measurment System

PBX - Private Branch Exchange

PESQ - Perceptual Evaluation of Speech Quality

PSQM - Perceptual Speech Quality Measurement

PSTN - Public Switched Telephone Network

QoS - Quality of Service

RTP - Real-time Transport Protocol

SCCP - Skinny Call Control Protocol

SDP - Session Description Protocol

SIP - Session Initiation Protocol

SRTP - Secure Real-time Transport Protocol

SS7 - Signaling System #7

TCP - Transmission Control Protocol

UA - User Agent

UDP - User Datagram Protocol

VoIP - Voice over IP

1 INTRODUÇÃO

1.1 CONTEXTO

Acredita-se que a indústria de telecomunicações, impulsionada pelo sucesso da Internet, está em fase de convergência com a tecnologia IP [1] . Nesse contexto, a tecnologia de voz sobre o protocolo de Internet (VoIP) desenvolve um papel de destaque nessa tendência.

A tecnologia VoIP torna possível o estabelecimento de conversações telefônicas através de uma rede IP. Essa tecnologia apresenta grandes vantagens sobre a telefonia convencional, sendo a principal delas a redução de custos. O barateamento dos custos é possível, pois a ligação é realizada através da Internet, onde os dados de voz são transmitidos juntamente com outros dados (web, email, mensagem instantânea, vídeo, etc.). Desta forma, não há a necessidade de uma infra-estrutura adicional para prover o serviço de telefonia.

A convergência do serviço de voz com a rede de dados abre espaço para uma grande variedade de inovações. Através da convergência é possível agregar diversos serviços, como por exemplo, mensagem instantânea, telefonia e vídeo que podem revolucionar a maneira como atualmente as pessoas e empresas encaram a comunicação [2] .

1.2 MOTIVAÇÃO

Ao longo dos últimos anos a tecnologia VoIP tem se mostrado bastante interessante devido a uma série de razões, entre elas, o baixo custo de chamadas. Por este motivo, a utilização dessa tecnologia tem crescido rapidamente. Todavia, ainda há desafios em aberto tais como segurança, tarifação, números de emergência e interceptação legal.

A interceptação legal de chamadas, também conhecida, como “grampo telefônico”, tem sido um recurso bastante utilizado pela justiça como instrumento de investigação. As empresas de telecomunicações são responsáveis por implantar em sua infra-estrutura mecanismos capazes de realizar a interceptação de chamadas a fim de atender os requerimentos exigidos pelo Estado. Nos Estados Unidos, por exemplo, todas as empresas de telecomunicações estão sujeitas a lei CALEA *(Communication Assistance for Law Enforcement Act)* [3] . De acordo com essa lei,

as empresas que fornecem serviços de telecomunicações devem prover meios para a realização de uma interceptação. Inicialmente, essa lei era aplicada somente às empresas convencionais de telecomunicações (operadoras de telefonia fixa e celular), entretanto, devido ao crescimento notório da telefonia IP, esta lei foi estendida para os provedores de serviços de voz sobre IP [4] .

1.3 OBJETIVO

Este trabalho de graduação tem como objetivo primário desenvolver uma arquitetura capaz de realizar a interceptação legal de chamadas VoIP baseadas no protocolo SIP. Como objetivo secundário, foi desenvolvida uma implementação que prova os conceitos apresentados neste trabalho e realizada uma avaliação de desempenho do sistema de interceptação proposto.

1.4 TRABALHOS RELACIONADOS

Devido à relevância do tema, existem diversos trabalhos relacionados. Uma breve descrição dos principais trabalhos é apresentada a seguir.

Em [5] há uma introdução sobre a questão de interceptação legal de chamadas VoIP. Neste trabalho é discutido, brevemente, como é realizada a interceptação de chamadas no sistema telefônico, além de destacar as principais diferenças entre a interceptação na rede de telefonia pública e na Internet. Também são apresentados alguns desafios da interceptação de chamadas VoIP.

Em [6] [7] e [8] são apresentados modelos de interceptação legal em redes IP. Esses modelos utilizam a infra-estrutura do provedor de Internet para realizar a interceptação. Contudo, essa abordagem só é interessante se o indivíduo sob investigação for assinante do provedor. Caso contrário, os modelos não são capazes de realizar a interceptação.

Alguns métodos de interceptações de chamadas VoIP baseadas no protocolo H.323 são apresentados em [9] . Os autores propõem algumas formas de realizar a interceptação, como por exemplo, inserir ‘escutas’ no *gateway,* no *gatekeeper* ou em um roteador fixo por onde todas as chamadas passam obrigatoriamente. O *gateway* e o *gatekeeper* são elementos que compõem a arquitetura de um sistema H.323. Também é proposto um dispositivo que funciona em modo promíscuo, monitorando o tráfego da rede. Esse dispositivo é responsável por identificar e

monitorar as mensagens de sinalização e mídia da máquina que o indivíduo sob investigação utiliza.

Em [10] é proposto um sistema distribuído para interceptação de chamadas VoIP. Um protótipo foi implementado suportando apenas o protocolo H.323. Entretanto, não foi realizada uma avaliação de desempenho do protótipo. Além disso, para realizar a interceptação das chamadas, a solução proposta nesse trabalho deve estar implantada na rede local onde o indivíduo sob investigação está realizando a chamada.

Em [11] é apresentada uma arquitetura capaz de interceptar chamadas VoIP baseadas no protocolo SIP. Este artigo apresenta algumas idéias similares às apresentadas neste trabalho. Porém, não foi desenvolvido um mecanismo para prover a rastreabilidade da chamada conforme apresentado na seção 5 (Limitações). Além disso, também não foi realizada uma avaliação de desempenho do sistema proposto.

2 REDE DE TELEFONIA PÚBLICA

2.1 INTRODUÇÃO

O sistema telefônico, comumente conhecido por *Public Switched Telephone Network* ou, simplesmente, PSTN, tem sua origem datada do inicio do século XIX e, desde então, este sistema tem se tornado cada vez mais complexo devido ao seu próprio crescimento (maior número de assinantes), e também devido à adição de novos serviços.

No início do sistema telefônico, os usuários eram conectados diretamente em pares. Isso significa que um usuário qualquer que necessitava ligar para três telefones distintos deveria possuir três aparelhos e três linhas telefônicas. Com o aumento do número de assinantes, foi necessário encontrar uma maneira de realizar a interligação dos usuários. Então, surgiu um novo elemento na arquitetura do sistema telefônico: a central telefônica. Esse elemento é responsável pela gerência, distribuição, concentração, interligação e tarifação das chamadas realizadas pelos usuários. Ainda devido à expansão do sistema telefônico, houve a necessidade de interligar as centrais, a fim de possibilitar a realização de chamadas de longa distância.

No sistema telefônico há dois tipos básicos de informação: mídia e sinalização. As informações de mídia correspondem à voz, propriamente dita, que trafega entre os participantes de uma ligação. Informações de sinalização se referem às informações trocadas pelos dispositivos do sistema. A sinalização é imprescindível para prover e manter o serviço telefônico funcionando adequadamente. Alguns exemplos de sinalização são os tons de linha e ocupado.

Além da evolução estrutural e arquitetural, o sistema telefônico incorporou novas tecnologias para proporcionar melhorias na qualidade de seus serviços. Surgiram novos serviços os quais foram adicionados ao sistema, como por exemplo, acesso à Internet, serviços de conferências e caller-ID[1]. Porém, a partir da década de 60 começaram a surgir problemas de segurança graves. *Phone phreakers*[2] eram capazes, entre outras coisas, de estabelecer ligações de

[1] Serviço através do qual quem está recebendo uma ligação é capaz de identificar o número do telefone que originou a chamada.
[2] *Hackers* da telefônia

longa distância livres de tarifação. Devido a incidentes como esse, novos protocolos foram adotados com o objetivo de elevar a confiabilidade do sistema.

2.2 ARQUITETURA

A rede de telefonia é uma rede comutada por circuito. Isto significa que para estabelecer uma ligação a rede aloca um circuito que será de uso exclusivo dos participantes da ligação. Esse circuito permanecerá alocado até que os participantes terminem a ligação e o liberem.

O sistema telefônico é caracterizado por ter a sua inteligência concentrada nos equipamentos que compõem o núcleo da rede, enquanto que os pertencentes as extremidades são, basicamente, terminais "burros". A seguir serão apresentados os principais elementos que compõem a arquitetura do sistema telefônico.

2.2.1 TERMINAL

Os terminais correspondem aos aparelhos de telefone. A maioria dos terminais possui duas tecnologias de sinalização: *Dial Pulse* ou *Dual-Tone Multi-Frequency* (DTMF). Contudo, atualmente, alguns terminais, chamados de telefones IP, já possuem a tecnologia de voz sobre IP incorporada.

Na tecnologia *Dial Pulse*, a sinalização é feita por um disco que gera pulsos na linha telefônica a uma determinada freqüência. De acordo com o número selecionado no disco são gerados diferentes quantidade de pulsos que são compreendidos pela central telefônica. Apesar de esta ser a primeira tecnologia de sinalização, ela ainda é suportada pelo sistema telefônico.

A tecnologia DTMF sinaliza através de emissão de tons. Em cada linha e coluna do teclado do telefone há uma freqüência associada, ver Figura 1.

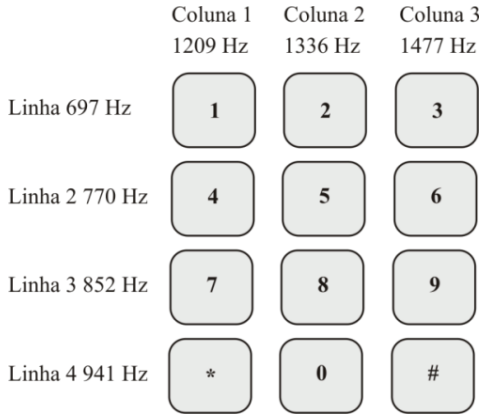


Figura 1 - Teclado DTMF

Dessa forma, quando, por exemplo, a tecla '1' é pressionada, é gerado um tom formado pela soma de uma senóide de freqüência 1209Hz e outra senóide de freqüência 697Hz. Este tom é, então, enviado para a central telefônica que através da análise do tom é capaz de descobrir que a tecla '1' foi pressionada.

2.2.2 *SWITCHBOARD*

Equipamento responsável por criar o circuito que será utilizado durante a ligação. Os primeiros *switchboards* eram operados manualmente por pessoas. Os operadores eram responsáveis por descobrir quais canais estavam disponíveis e estabelecer o circuito. Durante o estabelecimento de uma ligação, o terminal envia uma sinalização para o *switchboard*, que por sua vez, se comunica com outros *switchboards* para estabelecer um circuito até o terminal destino.

2.2.3 CENTRAL TELEFÔNICA

Centrais telefônicas são locais (*e.g.* prédios, casa, salas) onde se encontram diversos equipamentos de telecomunicação, como por exemplo, equipamentos de armazenamento, energia, tarifação e *switchboards*. As centrais telefônicas são administradas pela operadora de telefonia.

As centrais telefônicas estão organizadas de forma hierárquicas contento cinco níveis. Na periferia encontram-se as centrais classe 5 (*local exchange* ou *end office*) que estão ligadas diretamente aos terminais, normalmente centrais telefônicas de bairros ou regiões. A conexão

entre o terminal e a central classe 5 é conhecida como *local loop*. Em um nível acima se encontram as centrais classe 4 (*Toll Center*) que interligam duas *end offices* que por ventura não estejam diretamente ligadas, e assim, sucessivamente, até as centrais classe 1 que são responsáveis pelas ligações internacionais.

2.2.4 *PRIVATE BRANCH EXCHANGE*

As *Private Branch Exchange,* ou simplesmente PBXs, são equipamentos utilizados em empresas para interconectar os telefones, fax e *modems* entre si (ramais), além de conectá-los com a rede de telefonia pública. Esses equipamentos também oferecem diversos serviços como, por exemplo, IVR[3] (*Interactive Voice Response)* e conferência. É possível encontrar os termos PABX (*Private Automatic Branch Exchange*) e EPABX (*Electronic Private Automatic Branch Exchange)* substituindo o termo PBX.

2.2.5 LINHA E TRONCO

Uma linha representa a comunicação entre um terminal e uma central telefônica ou um PBX. Um tronco representa a comunicação entre centrais telefônicas ou entre uma central telefônica e um PBX. Ambos podem ser analógicos ou digitais.

Abaixo, a Figura 2 apresenta a arquitetura do sistema telefônico.

2.3 SINALIZAÇÃO

Conforme, já abordado anteriormente, a sinalização se refere à troca de informação entre os dispositivos do sistema telefônico para prover e manter o sistema de telefonia funcionando adequadamente.

[3] Um sistema IVR pode ser capaz de responder ou interagir com o interlocutor através de áudio pré-gravado permitindo, por exemplo, instruir o interlocutor das ações a tomar.

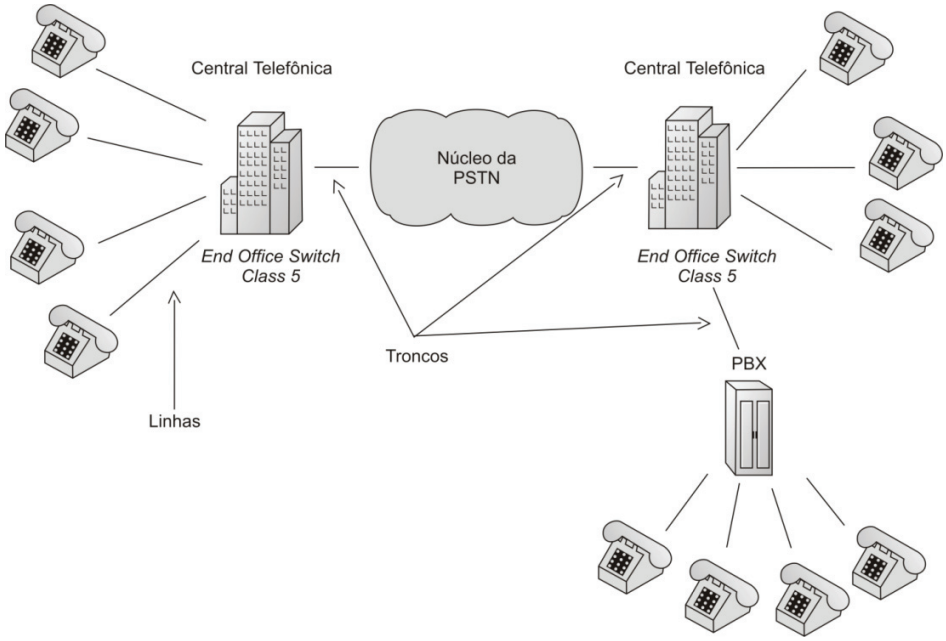


Figura 2 - Arquitetura do Sistema Telefônico

Todos os usuários do sistema telefônico trocam mensagens de sinalização entre si e com os componentes do sistema (*switchboard*, PBX, etc.). Como exemplos de mensagens de sinalização podem-se citar o tom de linha, os números digitados pelo usuário, o tom de ocupado, etc.

Existem três métodos de sinalização na PSTN:

- ***In-band Signaling:*** a sinalização é transmitida no mesmo circuito no qual a voz é transmitida. Também é utilizada a mesma faixa de freqüência da voz humana (300 a 3400 Hz).
- ***Out-of-Band Signaling:*** a sinalização continua sendo transmitida no mesmo circuito, porém fora da faixa de freqüência da voz humana (3400 a 3700 Hz).
- ***Common Channel Signaling (CCS):*** a sinalização é transmitida em outro circuito dedicado exclusivamente a sinalização.

Pelo fato dos métodos *In-band Signaling* e *Out-of-Band Signaling* utilizarem o mesmo circuito para transmitir voz e sinalização, eles são vulneráveis a fraude. Através de

equipamentos como o *blue box*[4], era possível realizar ligações de longa distância livres de tarifação.

O método *Common Channel Signaling* possui uma abordagem completamente diferente dos outros métodos, pois a sinalização é transmitida em um circuito dedicado. Como o usuário somente tem acesso ao circuito por onde a voz é transmitida, os tons de sinalização que ele transmitir não serão interpretados como sinalização, mas sim como voz. Portanto, as técnicas de fraude utilizadas em sistemas telefônicos do tipo *In-band Signaling* ou *Out-of-Band Signaling* são completamente ineficientes nos sistemas *Common Channel Signaling*.

Atualmente, a grande maioria dos sistemas telefônicos no mundo utiliza o método *Common Channel Signaling*. O protocolo de sinalização mais utilizado é o *Signaling System #7* (SS7), um padrão definito pela *International Telecommunication Union* (ITU).

2.4 SUMÁRIO

Neste capítulo, foram apresentados os principais elementos que compõem o sistema tradicional de telefonia fixa, juntamente com a sua arquitetura. Adicionalmente, também foram discutidas as tecnologias de sinalização existentes no sistema cuja evolução foi motivada, principalmente, por questões de segurança. Para maiores detalhes sobre o funcionamento do sistema telefônico, em [12] encontra-se uma ótima fonte de informações.

Ao longo dos anos, o sistema telefônico sofreu modificações em sua estrutura para incorporar novas tecnologias a fim de aumentar a confiabilidade e adicionar novos serviços, como por exemplo, o acesso a Internet. Acredita-se que, atualmente, esse sistema está passando por mais uma evolução impulsionada pela integração desse sistema com a tecnologia VoIP.

[4] Equipamento capaz de reproduzir tons nas freqüências utilizadas para a sinalização.

3 VOZ SOBRE IP

3.1 INTRODUÇÃO

O termo voz sobre IP, ou simplesmente VoIP (*Voice over IP*), se refere a um conjunto de tecnologias (protocolos de comunicação, *softwares* e *hardwares*) capazes de fornecer um serviço de telefonia em redes IP (*e.g.* a Internet) com qualidade bastante próxima das tradicionais redes de telefonia.

Apesar da tecnologia VoIP ter recebido bastante destaque nos últimos anos [13] [14] , o conceito de voz sobre IP é bastante antigo, datando da década de 1970. Os primeiros experimentos com transmissão de voz em redes não tradicionais ocorreram em 1973 com o protocolo NVP (*Network Voice Protocol)* [15] que utilizava a ARPANET[5] como rede de transmissão. Devido ao avanço das tecnologias de infra-estrutura de redes e, conseqüentemente, ao aumento das taxas de transmissão, somente por volta de meados da década de 1990, começaram a surgir aplicações, chamadas de *softphones*, capazes de realizar ligações IP com qualidade. Atualmente, com o grande avanço das tecnologias de infra-estrutura e com a adição de novas tecnologias, como o uso de *codecs* e protocolos de QoS (*Quality of Service*), a tecnologia VoIP é capaz de conferir qualidade de ligação bastante próxima a do sistema telefônico.

A tecnologia VoIP, basicamente, funciona da seguinte maneira: o sinal de voz é transformado em sinal digital, em seguida, estes dados de voz são comprimidos e divididos em pacotes. Os pacotes são enviados ao destino através de uma rede IP. Uma vez atingido o destino, esses dados são agrupados e transformados em sinal de voz novamente. Visto que, normalmente, os pacotes de voz são transmitidos através do protocolo UDP, que não garante a entrega do mesmo, é possível que haja perda de pacotes durante a transmissão devido a diversos motivos (*e.g.* congestionamento na rede).

Conforme ilustrado na Figura 3, a tecnologia VoIP pode ser utilizada em diferentes cenários. Chamadas VoIP podem ser realizadas entre dois terminais VoIP (*softphones* ou

[5] Desenvolvida nos Estados Unidos, a ARPANET (*Advanced Research Projects Agency Network)* foi a primeira rede de comunicação comutada por pacote. Alguns consideram a ARPANET a predecessora da Internet.

hardphones), cenário comumente utilizado, entre um terminal VoIP e um terminal do sistema telefônico e ainda entre dois terminais do sistema telefônico.

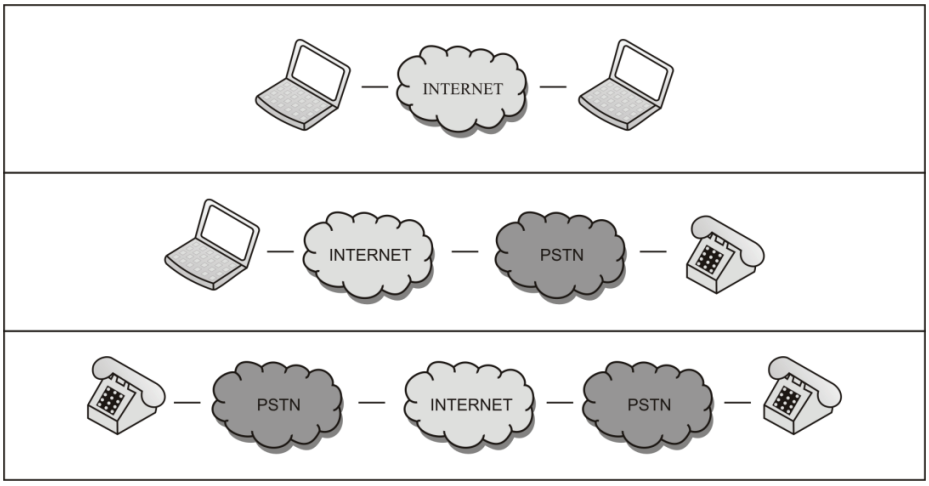


Figura 3 - Cenários da tecnologia de voz sobre IP

A tecnologia VoIP utiliza o *Real-time Transport Protocol* (RTP), que é um protocolo da camada de aplicação da pilha TCP/IP, para transmitir os dados relativos a voz. O RTP prover serviços de entrega fim-a-fim de áudio e vídeo em tempo real. Para garantir tais serviços, o protocolo prover identificação de conteúdo (*payload*), seqüenciamento, *timestamp*[6] e monitoramento de entrega de pacotes. Apesar de, normalmente, ser utilizado o protocolo de transporte UDP, o RTP também pode ser utilizado sobre o TCP (*e.g*. Skype).

Com o objetivo de prover uma boa qualidade de conversação três métricas de redes devem ser constantemente monitoradas: latência, *jitter* e perda de pacotes.

- **Latência**: A latência corresponde ao tempo que leva para um pacote ir de sua origem até o seu destino. No caso do VoIP, a latência corresponde ao tempo que a transmissão da voz leva até atingir o seu destino. Idealmente, a latência deve ser a mais baixa possível. A recomendação da ITU é que a latência deve

[6] Corresponde ao instante de tempo em que os dados de áudio pertencentes ao último pacote RTP foram enviados pela fonte.

ser mantida abaixo de 150 ms. Conversações com latências superiores a 150 ms tornam-se, praticamente, inteligíveis. Desta forma, as ligações VoIP tem que atingir esta restrição de tempo para emular com sucesso a qualidade de serviço provida pelo sistema telefônico.

- ***Jitter***: O *jitter* corresponde à variação de atraso dos pacotes que é, geralmente, causada por situações de enfileiramento de pacotes nos *buffers* dos roteadores. Variações de atraso podem ser piores para a qualidade de serviços do que o próprio atraso, pois o *jitter* pode causar a chegada dos pacotes fora de ordem. Apesar de o RTP possibilitar que aplicações façam o reordenamento usando as informações de número de seqüência e *timestamp,* o *overhead* em reordenar esses pacotes não é irrelevante, especialmente quando se está lidando com restrições de tempo estreitas como da tecnologia VoIP (150 ms). Uma solução genérica para controlar o *jitter* nos terminais VoIP é o uso de um *buffer* que libera os dados de voz, no máximo, a cada 150 ms. Assim, pequenas variações de atraso podem ser contornadas. O *jitter* também pode ser controlado através do uso de soluções com garantias de QoS nos equipamentos que compõem a rede VoIP.
- **Perda de Pacotes**: A tecnologia VoIP é muito intolerante quanto a perda de pacotes. Grandes latências podem gerar perda de pacotes. Quando pacotes chegam muito atrasados eles são descartados em favor de novos. Isso ocorre porque não faz sentido ”tocar” pacotes antigos depois que novos já foram “tocados”, tornando os pacotes antigos simplesmente inúteis. Além disso, o pacote pode ser perdido devido a problemas de redes comuns, como por exemplo, congestionamento. De um modo geral, os pacotes de voz são muito pequenos, normalmente eles contêm cerca de 12.5 a 62.5 ms de conversação somente. Portanto, a perda de um pacote destes não afetaria a qualidade da ligação como um todo. Porém, normalmente, a perda de pacote não ocorre de maneira isolada, pois os problemas que causam a perda de pacotes tendem a afetar todos os pacotes que estão sendo transmitidos naquele instante.

3.2 PROTOCOLOS

Nesta seção são apresentados os principais protocolos da tecnologia VoIP. Contudo, o protocolo SIP é apresentado em maiores detalhes, pois os conhecimentos sobre sua arquitetura e seu funcionamento são bastante relevantes para o sistema de interceptação legal de chamadas proposto neste trabalho.

3.2.1 *SESSION INITIATION PROTOCOL* (SIP)

Com a evolução da Internet e seus aplicativos surgiu a necessidade de um protocolo que capaz de criar e gerenciar sessões na Internet. Nesse contexto surgiu o protocolo SIP. Este protocolo tem como principal objetivo permitir que usuários da Internet sejam capazes de descobrir uns aos outros e gerenciar a criação, modificação e finalização de sessões multimídias. Além disso, através do SIP, também é possível negociar características da sessão que os usuários desejam compartilhar entre si. O SIP é um protocolo da camada de aplicação de propósito geral utilizado para gerenciamento de sessões na Internet capaz de operar sob vários tipos de protocolos de transporte (TCP, UDP e SCTP).

O SIP foi especificado pelo IETF, inicialmente, na RFC 2543 [16] em 1999, porém a sua versão mais atualizada, publicada em 2002, encontra-se na RFC 3261 [17] . Seguindo a tradição dos protocolos da camada de aplicação da Internet, ele é baseado em texto e suas mensagens se dividem em duas categorias: requisição e resposta. Por esse motivo, o SIP possui várias características do protocolo HTTP, inclusive possui alguns dos seus códigos de status, como por exemplo, o conhecido '*404 Not Found*'.

Este protocolo tem sido largamente utilizado em aplicações de conferência de áudio e vídeo, mensagens instantâneas e jogos *online*. Além disso, o SIP foi adotado como protocolo de sinalização pelo projeto 3GPP[7]. Na tecnologia VoIP, o protocolo SIP tem sido utilizado na sinalização de chamadas.

[7] 3rd Generation Partnership Project (3GPP) é um projeto criado em 1998 por grandes empresas de tecnologia móvel cujo objetivo é desenvolver especificações e relatórios técnicos para o sistema de telefonia móvel de terceira geração (3G).

Juntamente com o SIP, a tecnologia VoIP utiliza o *Session Description Protocol* (SDP) [18] para negociar as características da chamada. O SDP é um protocolo desenvolvido pelo IETF especificado na RFC 4566 [18] cujo objetivo é descrever sessões multimídias. Através deste protocolo é possível negociar informações sobre a sessão, como por exemplo, o *codec* a ser utilizado durante a chamada, chaves criptograficas, *time zone,* informações sobre largura de banda disponível, etc.

3.2.1.1 ARQUITETURA

Para prover as funcionalidades de gerenciamento de sessão apresentadas anteriormente, o SIP necessita de uma infra-estrutura de redes, composta por servidores, através da qual, os usuários se registram e tornam-se capazes de criar sessões, convidar outros usuários para uma sessão existente, finalizar sessões, etc. A arquitetura de um sistema SIP é composta por clientes e servidores, assim como um sistema HTTP (*browsers* e servidores web). Porém, na arquitetura SIP existem tipos de servidores diferentes com responsabilidades específicas. A seguir estão relacionados os elementos que compõem a arquitetura do sistema.

- *User Agent (UA):* São elementos que se encontram na periferia do sistema, assim como os terminais do sistema telefônico. De uma forma geral, os UAs representam as aplicações VoIP dos usuários (*softphones* e *hardphones*).
- *Registrar Server*: Servidor responsável por receber e armazenar informações de localização dos UAs. Inicialmente, o UA envia uma requisição de registro para este servidor. Quando algum elemento requisitar a localização de um determinado UA, o *registrar server* checa se possui tal informação. Em caso positivo, é enviada uma mensagem contendo o endereço IP do UA requisitado, e, em caso negativo, o servidor envia uma mensagem informando que ele não possui a localização do UA.
- *Proxy Server*: Servidor responsável por localizar o destino, através do *registrar server*, e encaminhar requisições. Quando um UA deseja estabelecer uma sessão, ele envia uma requisição para o *proxy server* que, por sua vez, encaminha esta requisição para o *proxy server* do destino ou para o próprio

UA destino, caso os UAs sejam do mesmo domínio. O *proxy server* também pode prover outras funções como autenticação, controle de acesso e segurança.

- *Redirect Server*: Desempenha as mesmas funções do *proxy server*, porém não realiza o encaminhamento de mensagens. Uma vez requisitado o estabelecimento de uma sessão, o *redirect server* responde ao UA a localização do destino. O UA, por sua vez, é responsável por enviar a mensagem de requisição de estabelecimento de sessão para o destino.

3.2.1.2 FORMATO DAS MENSAGENS

Todas as mensagens existentes no protocolo SIP possuem três parte: linha de início, cabeçalho e corpo.

Linha de início: Nesta parte é indicado o tipo de mensagem (requisição ou resposta) e a versão do protocolo. Em caso de requisição, é indicado o nome da requisição e, em caso de resposta, é indicado o código de resposta seguido de uma breve descrição. Os principais tipos de requisição são:

- *REGISTER*: mensagem utilizada para registrar um UA.
- *INVITE*: mensagem utilizada para inicializar uma sessão.
- *OPTION*: um UA pode obter informações sobre outro UA ou de um servidor SIP através deste tipo de requisição. Desta forma o UA pode descobrir informações, como por exemplo, os *codecs* suportados pelo UA destino antes de convidá-lo para participar de uma sessão.
- *ACK*: Utilizada para confirmação do recebimento de uma requisição de *INVITE*.
- *BYE*: Utilizada para terminar uma sessão.

Abaixo estão enumerados os tipos de respostas estabelecidas pelo protocolo SIP:

- 1xx Informativo: código utilizado para indicar informações gerais. Por exemplo, *'100 Trying'* indica que a requisição foi recebida e que está sendo

processada e *'180 Ringing'* indica que o UA destino recebeu a mensagem *INVITE* e está esperando uma ação do usuário (aceitar ou não a chamada).

- 2xx Sucesso: código utilizado para indicar sucesso na operação. Por exemplo, *'200 OK'*
- 3xx Redirecionamento: Código utilizado para indicar que houve mudança de endereço do servidor. Exemplo: *'302 Moved Temporarily'*.
- 4xx Falha do Cliente: Quando ocorre algum erro ou falha devido a uma requisição enviada pelo cliente. Exemplo: *'403 Forbidden'* e *'404 Not Found'*.
- 5xx Falha do Servidor: códigos utilizados quando ocorre algum erro no servidor. Exemplo: *'503 Service Unavailable'*.
- 6xx Falhas Globais: códigos utilizados para informar sobre falhas gerais do sistema.

Cabeçalho: Nesta parte da mensagem são apresentadas informações sobre os participantes da sessão. Os principais campos são:

- *To*: indica o endereço SIP do destino. Exemplo: alice@voip.provider.com.
- *From*: indica o endereço SIP do remetente.
- *Contact*: contém o endereço IP do UA para que ele seja contactado diretamente.
- *Via*: Neste campo é indicado o tipo de transporte que está sendo utilizado (TCP ou UDP), além de indicar os hosts por onde a mensagem passou até o momento. Quando enviada, UAs e servidores SIP alteram este campo, adicionando o seu endereço para informar que aquela mensagem passou por aquele host.
- *Call-ID:* este campo guarda um identificador único para a chamada em curso. Ele deve ser o mesmo para todas as requisições e respostas de uma mesma chamada.
- *Content-Type*: indica o tipo das informações contidas no corpo da mensagem.

- *Content-Length*: indica o tamanho do corpo da mensagem.

Corpo: Normalmente, no corpo da mensagem são colocados os dados referentes a sessão através do protocolo SDP.

3.2.1.3 ESTABELECIMENTO DE UMA CHAMADA SIP

Na Figura 4 são apresentadas as mensagens trocadas entre os elementos da arquitetura SIP durante o estabelecimento de uma chamada VoIP típica.

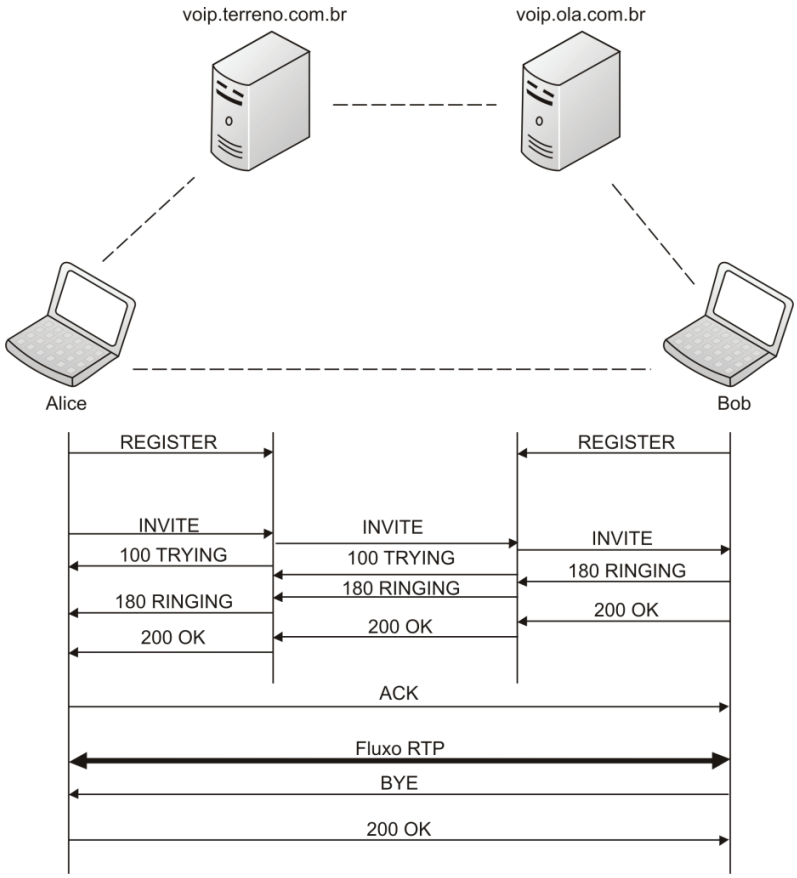


Figura 4 - Estabelecimento de uma chamada SIP

Inicialmente, quando Alice e Bob iniciam suas aplicações VoIP, ambos se registram no seus respectivos *Registrar Server* para que possam ser localizados por outros UAs. O processo de registro do UAs é feito através de requisições do tipo *REGISTER*. Uma vez registrado, o UA torna-se capaz de inicializar o processo de estabelecimento de uma chamada. Então, Alice envia uma mensagem do tipo *INVITE* para o seu servidor SIP endereçada a bob@ola.com.br

sinalizando que deseja estabelecer uma chamada com Bob. No corpo da mensagem *INVITE* são enviadas as informações do protocolo SDP para a negociação de sessão. Abaixo segue um exemplo da mensagem *INVITE*.

```
INVITE sip:bob@ola.com.br SIP/2.0
Via: SIP/2.0/UDP pc33.terreno.com.br;branch=z9hG4bK776asdhds
Max-Forwards: 70
To: Bob <sip:bob@ola.com.br.com>
From: Alice <sip:alice@terreno.com.br >;tag=1928301774
Call-ID: a84b4c76e66710@pc33.alicesprovider.com
CSeq: 314159 INVITE
Contact: <sip:alice@pc33.terreno.com.br>
Content-Type: application/sdp
Content-Length: 142

(Informações de Alice sobre a sessão são colocadas aqui)
```

O servidor SIP de Alice recebe a mensagem, localiza o servidor SIP de Bob através do protocolo DNS e o contacta, indicando que a sua cliente Alice deseja realizar uma chamada com Bob. O servidor SIP de Bob checa se ele está registrado, e a partir da informação de sua localização sinaliza para Bob, avisando-o que Alice deseja iniciar uma chamada com ele. A partir deste instante, a aplicação VoIP de Bob começa a "tocar". Neste momento, é enviada uma mensagem do tipo *'180 RINGING'*, sinalizando que Bob recebeu a mensagem de Alice e que o sistema está aguardando uma ação de Bob: aceitar ou não a chamada. Uma vez que Bob aceita, é enviada uma mensagem do tipo *'200 OK'*, significando que Bob aceitou a chamada e que está pronto para iniciar a conversação. Abaixo segue um exemplo da mensagem *'200 OK'*.

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP server10.ola.com.br
   ;branch=z9hG4bKnashds8;received=192.0.2.3
Via: SIP/2.0/UDP bigbox3.site3.terreno.com.br
   ;branch=z9hG4bK77ef4c2312983.1;received=192.0.2.2
Via: SIP/2.0/UDP pc33.terreno.com.br
   ;branch=z9hG4bK776asdhds ;received=192.0.2.1
To: Bob <sip:bob@ola.com.br>;tag=a6c85cf
From: Alice <sip:alice@terreno.com.br>;tag=1928301774
Call-ID: a84b4c76e66710@pc33.terreno.com.br
CSeq: 314159 INVITE
Contact: <sip:bob@192.0.2.4>
Content-Type: application/sdp
```

```
Content-Length: 131

(Informações de Bob sobre a sessão são colocadas aqui)
```

Alice responde com uma mensagem do tipo *ACK* diretamente para Bob (a partir deste momento todas as mensagens são enviadas diretamente para os UAs, não há mais a participação dos provedores VoIP) informando que ela está de acordo com os parâmetros da sessão e que também está pronta para o início da chamada. Uma vez que todos os participantes estão aptos a inicializar a chamada, os pacotes RTP, contendo os dados referentes à voz, começam a ser transmitidos. Quando um dos participantes deseja terminar a chamada ele envia uma mensagem do tipo *BYE* sinalizado que a chamada está sendo finalizada. O outro participante, então, envia uma mensagem do tipo *'200 OK'* sinalizando que a chamada está terminada.

3.2.1.4 INTEROPERABILIDADE COM A REDE DE TELEFONIA FIXA

Além das redes VoIP se comunicarem entre si através do protocolo SIP, também é possível que terminais SIP se comuniquem com terminais da rede de telefonia fixa. Diversos provedores VoIP têm disponibilizado esse serviço, o qual tem tido uma grande adoção nos últimos anos, principalmente, devido ao custo da chamada VoIP ser extremamente reduzido em relação ao custo de uma chamada de longa distância no sistema de telefonia tradicional. É possível encontrar algumas operadoras de telecomunicações que provêem parte do seu serviço de telefonia através de uma rede VoIP dentro da sua infra-estrutura.

Basicamente, para interoperar com a rede de telefonia pública é necessário adicionar à arquitetura SIP equipamentos como o *Media Gateway, Signaling Gateway* e *Media Gateway Controller* (MGC). O *Media Gateway* é o elemento responsável pela conversão de dados de voz contidos nos pacotes RTP para sinais de áudio utilizado no sistema telefônico e vice-versa. Da mesma forma, o *Signaling Gateway* é responsável por realizar a conversão das mensagens de sinalização contidas nas mensagens SIP (ou qualquer outro protocolo de sinalização VoIP) em sinalização do sistema telefônico e vice-versa. Já o MGC é responsável por controlar um ou mais *Media Gateway* e *Signaling Gateway*. Ele recebe mensagens de sinalização de um servidor SIP e configura o *Signaling Gateway* para repassar esta informação para o sistema telefônico. Além

disso, o MGC também configura o *Media Gateway* para prepará-lo para o início do fluxo de dados de voz.

A comunicação entre o MGC e os *gateways* é realizada através do *Media Gateway Control Protocol* (MGCP), definido pelo IETF, cuja especificação encontra-se na RFC 3435 [19] . O protocolo estabelece várias mensagens que são utilizadas para gerenciar (criar, modificar, finalizar) chamadas. Também existe o protocolo Megaco, que desempenha a mesma função do MGCP, o qual foi desenvolvido a partir de uma cooperação entre IETF (RFC 3525) e ITU (Recomendação H.248.1). A Figura 5 demonstra como esses equipamentos e protocolos estão dispostos na arquitetura SIP.

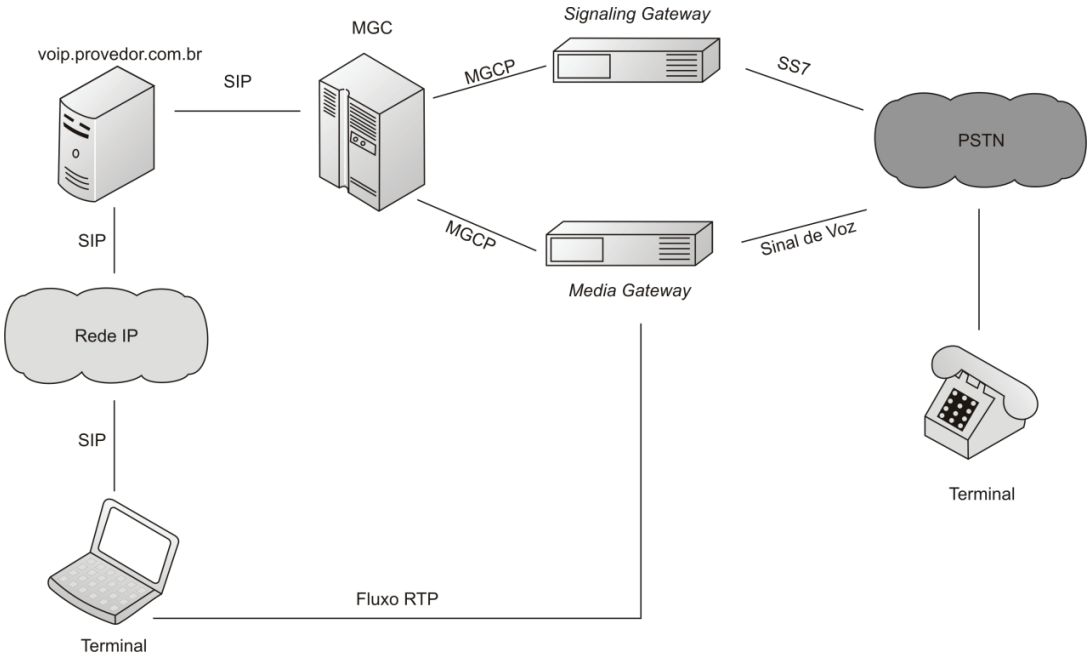


Figura 5 - Interoperabilidade

Quando um terminal VoIP inicializa uma chamada cujo destino é um terminar no sistema telefônico, a mensagem de *INVITE* é enviada para o servidor SIP (ver Figura 4). O servidor SIP detecta que o destino da chamada encontra-se no sistema telefônico, então, ele encaminha a mensagem *INVITE* para o MGC. O MGC de posse das informações contidas na

mensagem SIP configura o *Signaling Gateway* para que ele sinalize para o terminal destino. Logo depois, o MGC também configura o *Media Gateway* informando o endereço IP do terminal VoIP para onde o *gateway* deve enviar os pacotes RTP. Uma vez estabelecida a sinalização o MGC responde ao servidor SIP informando o endereço IP do *Media Gateway*. A partir deste momento, o terminal VoIP começa a enviar pacotes RTP para o *Media Gateway* que os converte em sinal de voz e os encaminham para o terminal destino vice-versa.

3.2.2 *REAL-TIME TRANSPORT* PROTOCOL

Devido a algumas características do protocolo TCP (*e.g.* controle de fluxo, reconhecimento e retransmissão de pacotes), este não se mostra adequado para transportar dados de tempo real[*]. O UDP possui algumas características interessantes, todavia, este protocolo também não se mostra adequado, pois não é confiável, não é orientado à conexão e o único serviço provido por ele é a verificação de erros. Por esse motivo, houve a necessidade de um protocolo que utilizasse o UDP como protocolo de transporte e, que fosse capaz de fornecer alguns serviços necessários para realizar entregas fim-a-fim eficazes. O protocolo RTP [20] foi desenvolvido com o objetivo de fornecer tais serviços para o transporte de dados de tempo real.

O RTP é um protocolo da camada de aplicação especificado pelo IETF, em 1996, na RFC 1889 [21] . Atualmente, a especificação do RTP possui uma nova versão, RFC 3550 [20] , lançada em 2003. Ele tem como principal objetivo prover serviços de entrega fim-a-fim de dados com características de tempo real. Os principais serviços são: identificação de dados *(payload)*, seqüenciamento de mensagem, *timestamp* e monitoramento de entrega.

É importante destacar que o protocolo RTP não fornece garantia de entrega de pacotes ou de ordem correta dos pacotes. Além disso, ele não garante o tempo de entrega e não provê qualidade de serviço. Conforme apresentado, o seu objetivo é fornecer informações sobre os pacotes para as aplicações, para que estas, de posse destas informações, possam modificar o seu comportamento, a fim de se adaptar as variações das características da rede.

[*] Apesar do protocolo TCP não parecer adequado para a transmissão de dados de tempo real, como dados de voz, o Skype utiliza o TCP para a transmissão destes dados quando os participantes da chamada encontram-se na mesma rede local.

A RFC 3550 [20] também especifica o protocolo RTCP, *RTP Control Protocol*. O RTCP tem como objetivo fornecer informações sobre o fluxo RTP em execução, como por exemplo, taxa de perda de pacotes e *jitter*, para os participantes da sessão em andamento. O seu funcionamento é baseado na transmissão periódica de pacotes de controle para os participantes da sessão.

Com o objetivo de fornecer características de segurança aos pacotes RTP, foi desenvolvido o protocolo *Secure Real-time Transport Protocol (SRTP)* [22] . O SRTP, cuja especificação encontra-se na RFC 3711 [22] , tem como objetivo prover confidencialidade, autenticação, integridade e proteção contra re-injeção de pacotes para tráfego RTP e RTCP. O *Advanced Encryption Standard* (AES) é utilizado para conferir serviços de criptografia aos pacotes e, para garantir autenticação e integridade é utilizado o algoritmo HMAC-SHA1, especificado na RFC 2104 [23] .

Na tecnologia VoIP, o protocolo RTP é utilizado para realizar o transporte dos dados de voz da chamada. Através das informações transmitidas pelo protocolo RTCP, as aplicações VoIP são capazes de renegociar o *codec* utilizado durante a execução da chamada, adaptando-a as novas condições da rede.

3.2.3 OUTROS PROTOCOLOS

Há diversos protocolos que podem ser utilizados para estabelecer e realizar chamadas telefônicas sobre redes IP. Além do SIP, RTP e RTCP, que foram discutidos nas seções anteriores, existem outros protocolos que merecem ser mencionados.

Inter-Asterisk eXchange [24] é o protocolo introduzido pelo Asterisk[8]. O Asterisk é um projeto *open source* de um PBX em software, inicialmente, desenvolvido pela empresa Digium[9]. O Asterisk foi desenvolvido para o sistema GNU/Linux x86, porém já foi portado para diversas plataformas como sistemas da família BSD (NetBSD, FreeBSD e OpenBSD), Mac OS X e Solaris. Além do *Inter-Asterisk eXchange,* o Asterisk suporta diversos protocolos como SIP, H.323, MGCP e SCCP. Atualmente, o Asterisk é mais do que um simples PBX, ele também pode

[8] http://www.asterisk.org/ (último acesso em 20/09/2008)
[9] http://www.digium.com/ (último acesso em 20/09/2008)

desenvolver funcões de *Gateway* (interfaceamento da rede VoIP com a sistema telefônico) e *Media Server.* O *Inter-Asterisk eXchange* já está na sua segunda versão conhecida como IAX2. Normalmente, o IAX2 utiliza o mesmo *stream* UDP (porta 4569) para transmitir sinalização e mídia. Por este motivo ele é conhecido como um protocolo *in-band,* já que transmite tanto informação de sinalização quanto de voz no mesmo fluxo. Já os outros protocolos apresentados neste trabalho são conhecidos como *out-of-band*, pois eles utilizam fluxos diferentes para transmitir mensagem de sinalização e voz. Com a crescente utilização do Asterisk, este protocolo tem sido freqüentemente utilizado.

O padrão H.323 [25] foi especificado em 1996 pela ITU. O H.323 estabelece padrões para codificação e decodificação de fluxos de dados, garantindo que produtos baseados no padrão H.323 de um determinado fabricante interopere com produtos de outros fabricantes. Atualmente, o H.323 se encontra na versão 6. Este padrão já foi bastante adotado em sistemas VoIP como por exemplo, o Microsoft NetMeeting, porém vem perdendo espaço para outros protocolos como o SIP e IAX2 devido a sua relativa complexidade. O padrão é composto por diversos outros padrões e protocolos que têm por finalidade possibilitar não só o estabelecimento de chamadas telefônicas, mas também o estabelecimento de vídeo conferências e transmissão de dados em geral. Na Tabela 1 são apresentados os principais padrões e protocolos do padrão H.323.

PADRÃO/PROTOCOLO	FINALIDADE
H.225	Sinalização
H.235	Segurança
H.450	Serviços adicionais como transferência (H.450.2), espera (H.450.6), identificação (H.450.8) de chamadas
H.460	Suporte a QoS
H.245	Interfaceamento com a rede de telefonia pública
H.510	Mobilidade para os usuários
H.530	Aspectos de segurança para o H.510
RTP	Encapsulamente de dados de voz

Tabela 1 - Padrões e Protocolos do H.323

O padrão H.323 é bastante extenso, porém os seus principais objetivos podem ser resumidos nos pontos a seguir:

- Interoperabilidade de equipamentos e aplicações;
- Suporte a conferências ponto-a-ponto e multiponto;
- Suporte a tarifação: o padrão H.323 possui mecanismos que permitem a controle do número de chamadas simultâneas e da largura de banda utilizada pelos terminais H.323. Dessa forma, possibilita a contabilização dos recursos de rede utilizados, podendo ser usada na tarifação dos serviços;
- Segurança: o padrão H.323 possui suporte à autenticação de usuários, à confidencialidade e ao não-repúdio das informações trocadas entre os participantes de uma chamada;
- Serviços suplementares: permite a elaboração e o desenvolvimento de serviços adicionais, como por exemplo, chamada em espera e identificação da chamada.

Concluindo esta seção onde foram apresentados os principais protocolos da tecnologia VoIP, observa-se na Figura 6 os protocolos agrupados de acordo com a sua finalidade (sinalização, mídia e controle de *gateway*). Além disso, também se observa como esses protocolos estão dispostos na pilha TCP/IP.

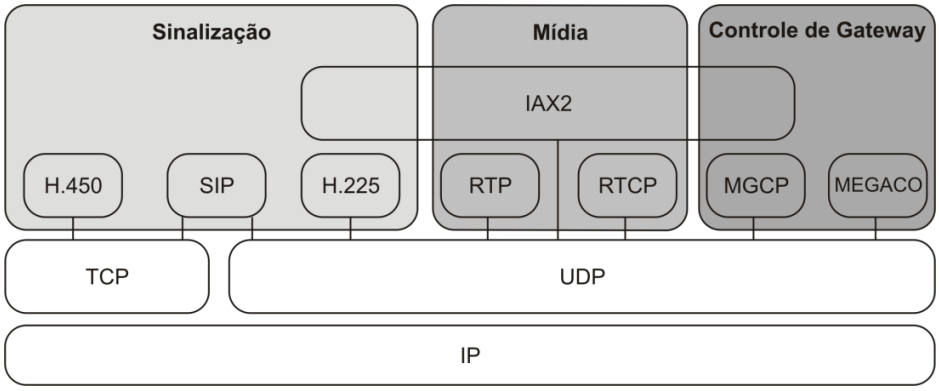


Figura 6 - Procolos da tecnologia VoIP

3.3 CODECS

Os *Codecs*[10] são implementações em *software* ou em *hardware* cuja finalidade é comprimir e descomprimir sinal de voz. Existem dois tipos de *codecs:* sem perdas (*lossless*) e com perdas (*lossy*). Conforme o próprio nome já diz, os *codecs* do tipo *lossless* são capazes de comprimir dados sem que haja perda de informação. Geralmente esse tipo de *codec* é utilizado em situações em que a perda de dados torna inutilizável o dado recuperado (*e.g.* compressão de arquivos) ou quando se deseja manter o máximo de qualidade possível (*e.g.* MPEG-4 ALS[11] e FLAC[12]). Por outro lado, os *codecs* do tipo *lossy* perdem informações durante o processo de compressão. Essa perda degrada a qualidade do dado recuperado, porém em algumas situações (a depender da aplicação) a perda de qualidade pode ser perfeitamente aceitável (*e.g.* mp3). Normalmente, esse tipo de *codec* apresenta uma maior taxa de compressão em relação aos *codecs* do tipo *lossless*.

O uso de *codecs* torna-se necessário na tecnologia VoIP, pois durante uma chamada um grande montante de dados de voz são transmitidos entre os participantes. Se esses dados forem transmitidos sem compressão eles não serão entregues de acordo com as restrições temporais de 150 ms, e, conseqüentemente, a chamada será totalmente inteligível. Dessa forma, os *codecs* são utilizados para comprimir dados de voz, fazendo com que menos dados sejam transmitidos, logo, menor taxa de transmissão será requerida da rede.

No sistema telefônico, uma vez alocado o circuito que será utilizado pela chamada, haverá uma vazão constante de 64 Kbps disponível para os participantes. Porém, nas redes IP, principalmente na Internet, o mesmo não ocorre. A vazão da rede pode variar bastante durante a chamada. Por este motivo, a tecnologia VoIP necessita de *codecs* para comprimir os dados de voz e, conseqüentemente, requerer menor largura de banda durante a chamada.

Algumas implementações mais robustas de terminais VoIP possuem mecanismos de adaptação de *codecs* para permitir que os *codecs* sejam negociados e substituídos por outros mais

[10] O termo *codec* é um acrônimo de codificador-decodificador.
[11] MPEG-4 Audio Lossless Coding
[12] Free Lossless Audio Codec

apropriados durante a chamada. Assim, é garantida a adaptabilidade das aplicações mediante a variações da rede durante a execução da chamada VoIP.

Um dos principais desafios no desenvolvimento de algoritmos de compressão de dados de voz é o fato de que além da preocupação com a perda de qualidade, o algoritmo deve ser rápido o suficiente para que o *overhead* de tempo inserido pelo próprio algoritmo durante a compressão e descompressão seja irrelevante em relação aos 150 ms da restrição temporal da tecnologia VoIP.

A Tabela 2 foi construída a partir das referências [26] [27] [28] [29] [30] . Ela apresenta os principais *codecs* e suas respectivas características.

CODEC	ALGORITMO	*TAXA DE TRANMISSÃO*	*MOS*	REQUER LICENÇA
G.711	PCM	64 Kbps	4.1	Não
G.722	PCM	48, 56 ou 64 Kbps	4.0	Não
G.726	ADPCM	32 Kbps	3.85	Não
G.728	LD-CELP	16 Kbps	3.6	Não
G.729A	CS-ACELP	8 Kbps	3.9	Sim
G.723.1	ACELP ou MPC-MLQ	5.3 ou 6.3 Kbps	3.65 ou 3.9	Sim
GSM	RPE-LTP	13 Kbps	3.7	Não
iLBC	Variação do CELP	13.3 ou 15.2 Kbps	4.14	Não
Speex	CELP	20.8 Kbps	4.1	Não

Tabela 2 – Codecs utilizados na tecnologia VoIP

O MOS, apresentado na Tabela 2, é a abreviação de *Mean Opinion Score.* Trata-se de uma métrica especificada pela ITU para avaliar a qualidade de conversações. A qualidade pode ser categorizada em cinco níveis: excelente (MOS = 5), bom (MOS = 4), regular (MOS = 3), pobre (MOS = 2), ruim (MOS = 1). O valor do MOS é avaliado de forma subjetiva, por este motivo, pode-se encontrar variações entre os valores de MOS para um determinado *codec.*

Nota-se que o MOS diminui de acordo com o a diminuição da taxa de transmissão mínima exigida pelo *codec*. Isso ocorre porque para alcançar taxas de transmissão cada vez menores os *codecs* precisam de algoritmos com taxas de compressão maiores, por conseqüência, ocorrem perdas de dados durante a compressão. Visto que os dados perdidos não podem ser reproduzidos pelo receptor, a qualidade da conversão diminui e conseqüentemente o MOS também.

Em contra partida, é possível notar que alguns *codecs*, como por exemplo o G.729A e o G.723.1, são capazes de manter um valor de MOS bastante elevado mesmo apresentando taxas de compressão elevadíssimas. Isso se deve, principalmente, as técnicas mais robustas de compressão que esses *codecs* utilizam.

A principal característica da tecnologia VoIP em relação aos *codecs* é que a tecnologia tem que ser capaz de detectar a variação das condições da rede (taxa de perda de pacotes, *jitter*, atraso, vazão, etc), durante a execução da chamada, e adaptar-se, escolhendo o *codec* mais apropriado para aquela situação.

3.4 SUMÁRIO

Neste capítulo, foram apresentadas diversas tecnologias que compõem uma rede VoIP. A tecnologia VoIP não se resume a apenas um protocolo de comunicação, mas sim, a um conjunto de tecnologias que, quando utilizadas em conjunto, tornam-se capazes de prover o serviço de telefonia através da Internet. Um sistema VoIP baseado em SIP está organizado da seguinte forma: o protocolo RTP é utilizado para o transporte de dados referentes a voz, os *codecs* são utilizados para comprimir esses dados para que seja necessário menor largura de banda da rede, o protocolo SIP é utilizado para a sinalização do sistema, o protocolo SDP é utilizado para descrever características da sessão e, finalmente, o protocolo MGCP é utilizado para controlar *gateways*, garantindo interoperabilidade entre o sistema VoIP e rede de telefonia pública.

4 INTERCEPTAÇÃO LEGAL DE CHAMADAS

4.1 INTRODUÇÃO

A interceptação legal de chamadas telefônicas, também conhecida popularmente como "grampo telefônico", vem sendo bastante utilizada pela justiça como instrumento de investigação. A interceptação legal tem sido uma poderosa ferramenta utilizada pelo Estado através das agências de investigação, não só para garantir evidências de crimes, mas também para identificar redes criminosas e os relacionamentos entre suspeitos.

A interceptação legal de chamadas, que está prevista na Lei 9.296, de 24 de julho de 1996, consiste no processo de interceptar, monitorar e rastrear a comunicação telefônica estabelecida por um indivíduo e seu(s) interlocutor(es). Essa ação é realizada por um órgão de monitoração, mediante autorização judicial, para fins de investigação criminal e instrução processual penal conforme disposto no artigo 1º da Lei Federal 9.296:

> *Art. 1º A interceptação de comunicações telefônicas, de qualquer natureza, para prova em investigação criminal e em instrução processual penal, observará o disposto nesta Lei e dependerá de ordem do juiz competente da ação principal, sob segredo de justiça.*
>
> *Parágrafo único. O disposto nesta Lei aplica-se à interceptação do fluxo de comunicações em sistemas de informática e telemática.*

Além do conteúdo de voz, dados importantes da chamada, como por exemplo, o número e a localização dos assinantes, a data e hora da chamada, a identificação dos canais de comunicação e a ocorrência de serviços suplementares, também devem ser monitoradas.

Maiores detalhes sobre questões jurídicas da interceptação legal de chamadas podem ser encontrados no artigo do Juiz Federal José Paulo Baltazar Junior [31] .

4.2 INTERCEPTAÇÃO LEGAL DE CHAMADAS NA REDE DE TELEFONIA PÚBLICA

Devido à arquitetura do sistema telefônico e ao fato de seu núcleo ser, completamente, controlado pelas operadoras de telecomunicações, a interceptação legal de

chamadas torna-se um procedimento, relativamente, trivial quando comparado com uma rede totalmente descentralizada como a Internet. Além disso, devido à maturidade do sistema telefônico, a problemática de interceptação de chamadas possui procedimentos bem estruturados, maduros e eficazes.

O fato do usuário sempre conectar ao sistema telefônico através do mesmo *local loop*, porção da rede telefônica que conecta a linha do assinante diretamente à rede da operadora, (ver seção 2.2.3), facilita bastante a interceptação. Além disso, a identificação dos terminais é estabelecida de maneira única, não podendo ser alterada pelo usuário. Por esse motivo, a rastreabilidade da ligação pode ser realizada através, apenas, do número do assinante.

Conforme ilustrado na Figura 7, a interceptação pode ser realizada diretamente no *local loop*. A maneira mais trivial de interceptação seria a agência de monitoração conectar um cabo extensor ao *local loop* do assinante. Dessa forma, a agência recebe todo a conteúdo de voz e dados que trafegam na linha telefônica, que em seguida são gravados e armazenados em equipamentos especializados.

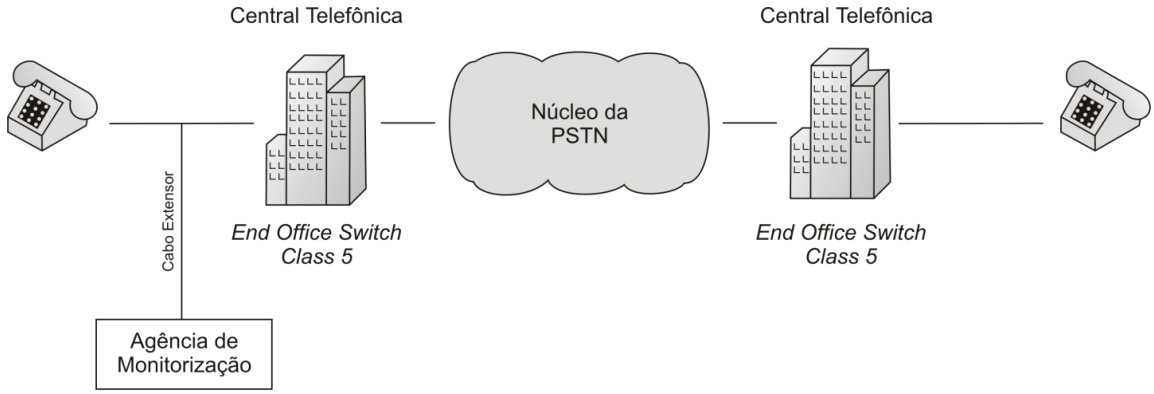


Figura 7 - Interceptação no sistema telefônico (1)

Entretanto, um significante número de serviços suplementares, aplicados à chamada telefônica básica, é oferecido pela central telefônica. Esses serviços incluem, por exemplo, redirecionamento de chamadas, transferência e conferência. Assim, quando a interceptação autorizada é conduzida através do *local loop*, muitas partes da chamada associada a estes serviços

são perdidas, já que a interceptação é realizada antes da comutação da chamada na central telefônica.

Uma forma um pouco mais sofisticada seria conectar equipamentos de monitoração, de modo não intrusivo, na entrada e saída da central na qual o assinante alvo da monitoração encontra-se conectado, conforme ilustrado na Figura 8.

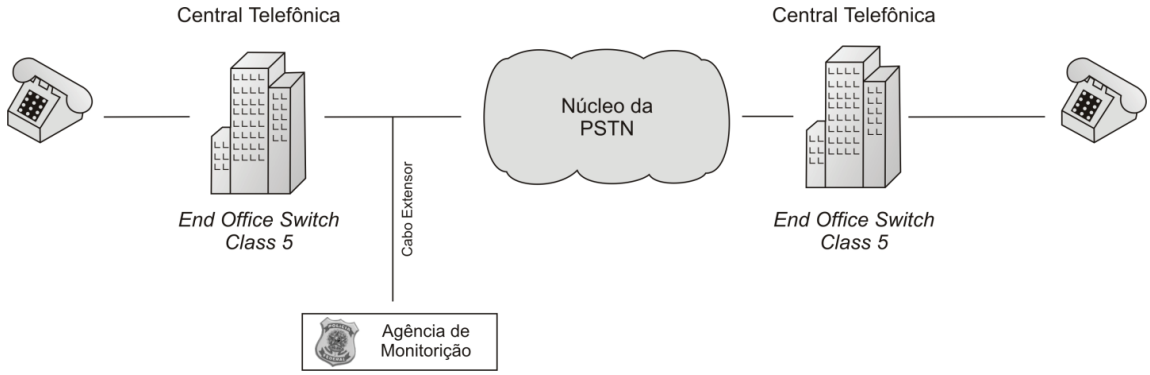


Figura 8 - Interceptação no sistema telefônico (2)

A interceptação é realizada pela monitoração dos canais de áudio e de sinalização. A partir dos dados derivados da sinalização telefônica, o equipamento identifica o número do assinante alvo e intercepta somente as informações pertinentes a este assinante.

Existem outras formas mais sofisticadas de realizar a interceptação legal de chamadas no sistema telefônico [32] , porém, o estudo destas técnicas não faz parte do escopo deste trabalho.

4.3 DESAFIOS DA INTERCEPTAÇÃO LEGAL DE CHAMADAS EM REDES VOIP

Ao longo dos últimos anos, a tecnologia VoIP tem se mostrado bastante interessante devido a uma série de razões, entre elas, o baixo custo de chamadas. Por este motivo, a adoção da tecnologia tem crescido rapidamente. Todavia, esta tecnologia ainda encontra-se em fase de desenvolvimento. Questões como segurança, tarifação, números de emergência e interceptação legal ainda estão sendo estudadas e propostas.

Para interceptar chamadas da rede de telefonia, há um procedimento bem estruturado, maduro e eficaz. Porém, para interceptar chamadas VoIP ainda não há nada consolidado. As técnicas de interceptação do sistema telefônico se baseiam no fato de que este sistema possui uma arquitetura centralizada. Por este motivo não é possível adaptar, diretamente, essas técnicas na interceptação de chamadas VoIP, visto que a Internet possui uma arquitetura totalmente descentralizada. No sistema telefônico, o usuário sempre conecta através do mesmo *local loop*, o que facilita bastante a interceptação, porém em redes VoIP nem sempre o usuário vai conectar com o mesmo IP. Portanto, novas técnicas necessitam ser desenvolvidas.

Existem diversos desafios que necessitam ser superados para que seja possível interceptar uma chamada VoIP de maneira eficaz. Uma chamada realizada de uma localização fixa com um endereço IP fixo conectado diretamente a um grande provedor de Internet é, relativamente, fácil de ser interceptada. Porém, devido a própria mobilidade oferecida pelas redes IP, é possível que um cliente VoIP migre entre várias redes IP (todas conectadas à Internet através de diferentes provedores de Internet) e que utilize endereços IPs diferentes. Nesta situação, a intercepção de chamada torna-se bastante complexa.

Diferentemente da rede de comutação de circuitos, que estabelece um circuito virtual entre a origem e o destino, a interceptação em uma rede de comutação de pacotes se torna mais complexa devido às características desse tipo de rede. A principal dificuldade da interceptação em redes IP está no fato desse protocolo não ser orientado à conexão, fazendo com que as informações de voz, divididas em vários pacotes, possam ser roteadas por diversos caminhos. Além disso, um roteador da Internet pode carregar tráfegos de tipos diferentes e de fontes diversas, de forma que para saber se um determinado pacote faz parte de uma comunicação interceptada, todos eles devem ser examinados.

Além de realizar a interceptação da chamada, é necessário que a mesma possa ser rastreada com o objetivo de garantir quem são os participantes da chamada. A ratreabilidade é necessária para que a prova criminal não possa ser refutada. Na rede de telefonia pública, a rastreabilidade é garantida através da identificação do número do assinante no *local loop,* todavia, não existe uma identificação análoga na rede de pacotes. O que mais se aproxima é o endereço

IP, entretanto, os endereços são, normalmente, atribuídos dinamicamente. Logo, é necessário desenvolver algum mecanismo de rastreamento de chamadas VoIP.

Portanto, para realizar interceptações legais de chamadas VoIP de forma eficaz, são necessárias pesquisas com o objetivo de desenvolver técnicas que sejam capazes de interceptar e rastrear chamadas de forma não intrusiva. Adicionalmente, essas técnicas não devem degradar expressivamente a qualidade da ligação, e a sua presença deve ser imperceptível para os participantes da chamada.

5 SISTEMA PROPOSTO

5.1 INTRODUÇÃO

Conforme pôde ser observado na seção 4.3, a problemática de interceptação legal de chamadas em redes VoIP apresenta uma considerável complexidade. Para construir uma arquitetura capaz de fornecer o serviço de interceptação legal foram necessários estudos sobre o sistema telefônico e a tecnologia VoIP. A arquitetura e os principais elementos do sistema telefônico foram estudados devido à necessidade de compreensão do funcionamento das técnicas de interceptação de chamadas existentes neste sistema. Da mesma forma, a tecnologia VoIP foi estudada para compreender o seu funcionamento, juntamente com os seus principais protocolos.

Conforme apresentado na seção 4.2, nos sistema telefônico, apenas a operadora de telecomunicações está diretamente envolvida no processo de interceptação legal de chamadas, facilitando o estabelecimento e gerenciamento da mesma. Contudo, na telefonia VoIP, existem duas entidades envolvidas: o provedor de Internet e o provedor do serviço de voz sobre IP.

Uma abordagem possível é utilizar, somente, o provedor de Internet para realizar a interceptação visto que o provedor tem controle sobre o roteador que está diretamente ligado ao usuário. Para tal, seria necessário equipar o roteador, através do qual o usuário sob investigação está conectado, com um *software* capaz de analisar os pacotes e detectar quais fazem parte da chamada VoIP e enviar uma cópia destes pacotes para a agência de monitoração. Esta abordagem é possível já que todos os pacotes enviados e recebidos pelo usuário nesta chamada, obrigatoriamente, trafegam por este roteador. A Figura 9 ilustra o funcionamento desta abordagem que é bastante similar ao mecanismo de interceptação utilizado no sistema telefônico (vide seção 4.2).

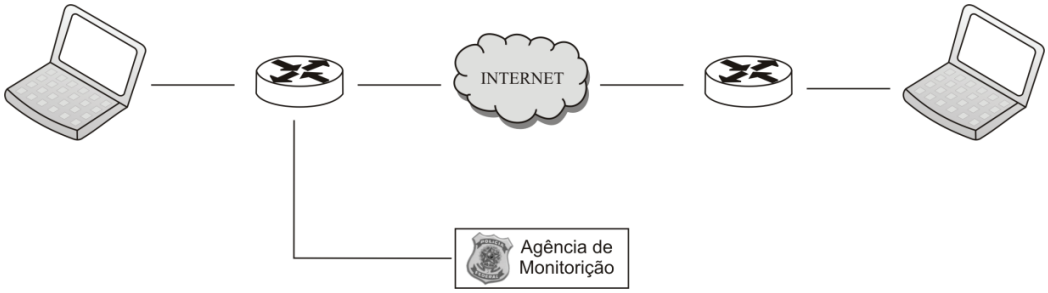


Figura 9 - Interceptação através do provedor de Internet

Porém, conforme discutido na seção 4.3, não há garantias de que o indivíduo sob investigação realizará chamadas a partir da mesma localidade, ou seja, não há garantias de que ele estará conectado através do mesmo provedor de Internet e diretamente ligado ao roteador em questão. Portanto, esta abordagem, sozinha, torna-se ineficaz para o propósito da interceptação legal de chamadas VoIP.

Outra abordagem está na utilização do provedor do serviço de telefonia IP. Esta abordagem apresenta-se bastante interessante, pois independentemente do provedor de Internet utilizado pelo indivíduo, a chamada sempre é estabelecida através do provedor VoIP. Portanto, todas as chamadas realizadas ou recebidas pelo indivíduo, obrigatoriamente, passam pelo provedor VoIP. Porém, esta abordagem sozinha não é capaz de realizar a rastreabilidade da chamada visto que é necessário ter acesso ao fluxo RTP trocado entre os participantes.

Este trabalho utiliza ambas as abordagens. Resumidamente, o provedor VoIP é utilizado para realizar a interceptação da chamada, enquanto o provedor de Internet é responsável por rastreá-la. Portanto é considerado que o indivíduo sob investigação possui mobilidade e pode realizar ou receber chamadas de qualquer lugar da Internet.

Para prover a rastreabilidade de chamadas foi utilizada uma técnica de esteganografia. A esteganografia é comumente definida como a ciência de ocultar mensagens. Diferentemente da criptografia, cujo objetivo é garantir que terceiros sejam incapazes de entender uma determinada mensagem, o objetivo da esteganografia é que terceiros não tenham conhecimento da existência da mensagem. Técnicas de esteganografia vêm sendo utilizadas como marca d'água digital. Mensagens são ocultadas em mídias (CD e DVD) para que a sua

origem possa ser rastreada e verificada a fim de combater a pirataria e auxiliar o gerenciamento de direitos autorais *(Digital Rights Management - DRM).*

5.2 ARQUITETURA

A arquitetura do sistema de interceptação legal proposto neste trabalho é baseada na arquitetura SIP apresentada na seção 3.2.1.1. Um elemento chamado de servidor de interceptação foi introduzido a fim de realizar a interceptação e rastreamento das chamadas. A Figura 10 ilustra a arquitetura do sistema proposto.

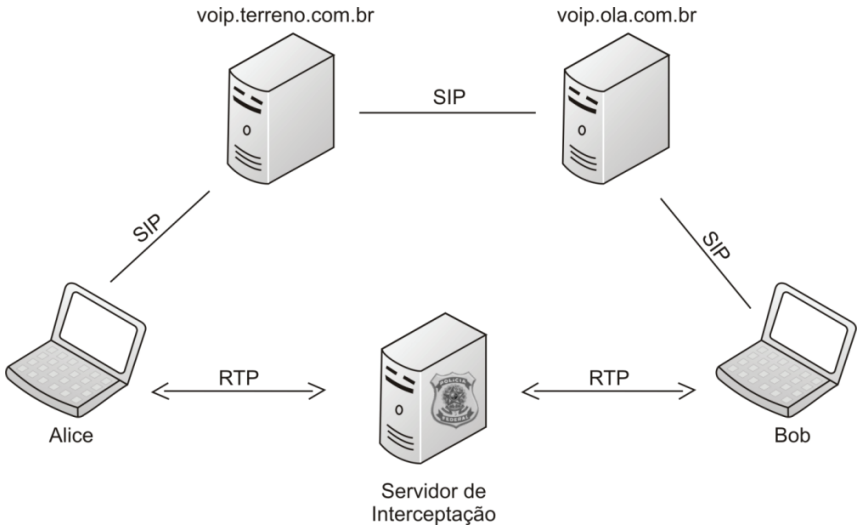


Figura 10 - Arquitetura do sistema de interceptação legal proposto

O servidor de interceptação é um elemento controlado pela agência de monitoração responsável por realizar a gravação da chamada, encaminhar os pacotes RTP para os seus respectivos destinos, e inserir a marca d'água nos pacotes para que a chamada possa ser rastreada.

5.3 MECANISMO DE FUNCIONAMENTO

O mecanismo de interceptação proposto neste trabalho baseia-se no ataque *man-in-the-middle*. Esse ataque é bastante conhecido e tem como objetivo permitir que o atacante seja

capaz de ler, inserir e modificar, mensagens entre duas entidades sem que estas tenham conhecimento da sua existência. Tipicamente, o atacante insere-se no meio da comunicação entre dois pontos, fazendo parte do canal de comunicação.

Em uma comunicação normal, os dois elementos envolvidos comunicam entre si diretamente (ver Figura 11). Porém, durante o ataque *man-in-the-middle* a comunicação é interceptada pelo atacante e retransmitida por este para a outra parte (ver Figura 11). No sistema proposto, o servidor de interceptação desenvolve o papel do atacante. Desta forma, toda a conversação da chamada passa pelo servidor de interceptação.

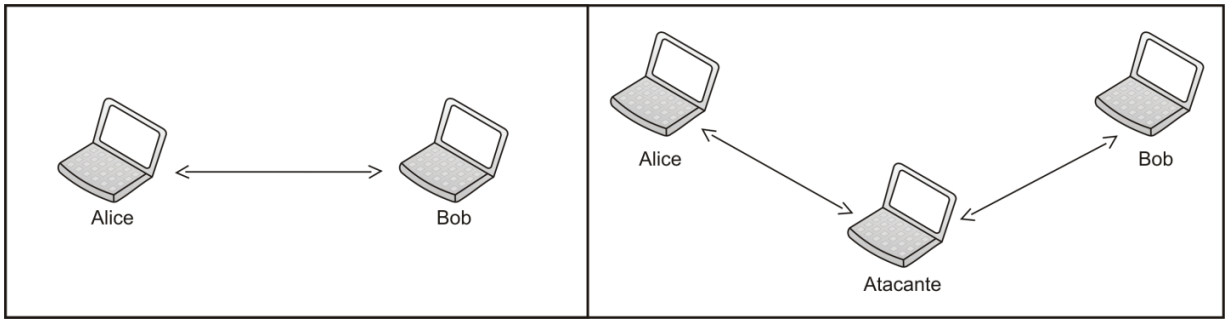


Figura 11 - Ataque *man-in-the-middle*

Contudo, para que o fluxo RTP da chamada seja direcionado para o servidor de interceptação, é necessário manipular a sinalização da chamada. Dado que todas as mensagens de sinalização obrigatoriamente passam pelo provedor VoIP, este é responsável por essa manipulação. Para realizar o direcionamento do fluxo RTP basta o provedor VoIP trocar o endereço IP contido no campo *'Contact'* (ver seção 3.2.1.2) do cabeçalho SIP pelo endereço IP do servidor de interceptação. A Figura 12 demonstra a manipulação das mensagens de sinalização trocadas durante o estabelecimento da chamada entre Alice e Bob para que o fluxo RTP seja direcionado para o servidor de interceptação.

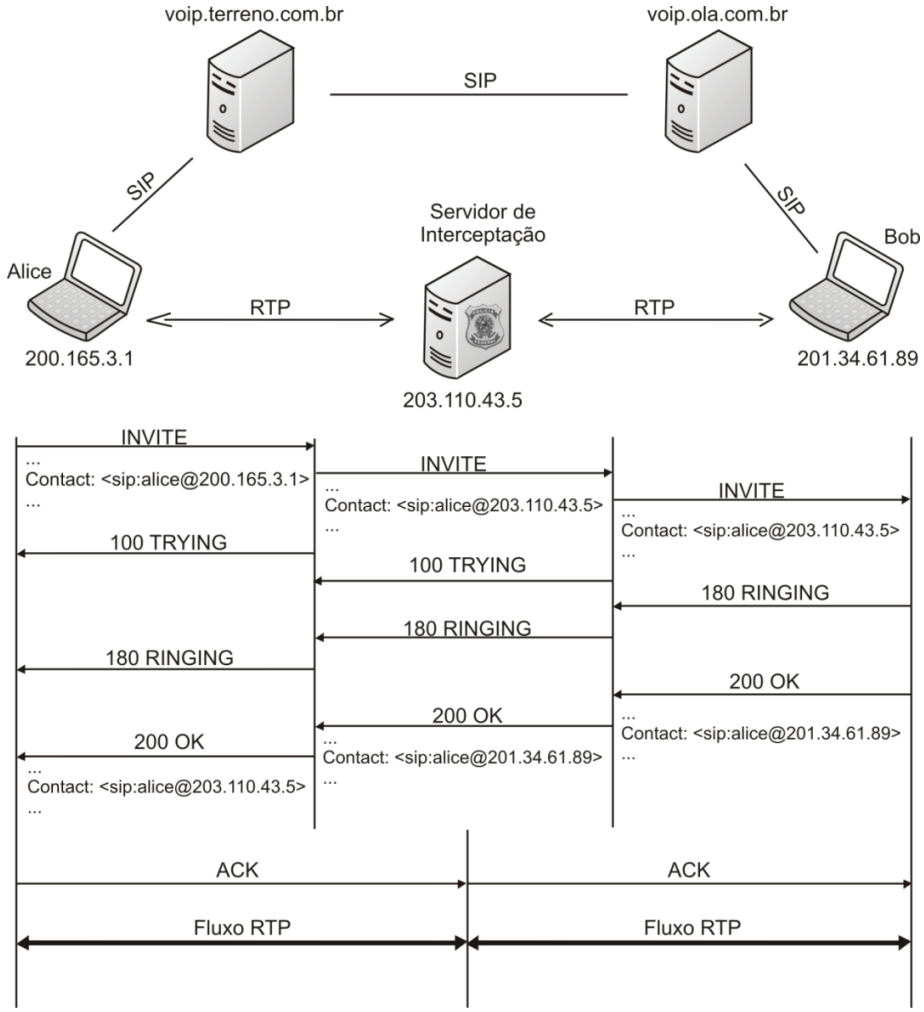


Figura 12 - Interceptação da sinalização SIP

Inicialmente Alice envia uma mensagem de *INVITE*. Ao receber a mensagem, o servidor SIP do provedor de Alice verifica se o usuário 'Alice' encontra-se na lista de usuário que estão sob investigação. Em caso positivo, o servidor SIP encaminha a mensagem de *INVITE* para o servidor SIP de Bob, porém o campo '*Contact*' da mensagem contém o endereço IP do servidor de interceptação. Ao receber a mensagem, o servidor SIP a encaminha para Bob. Considerando que Bob aceita o convite de Alice, ele envia um *'200 OK'* contendo no campo *'Contact'* o seu endereço IP. Neste momento, Bob acredita que Alice encontra-se no endereço IP do servidor de interceptação visto que no campo *'Contact'* da mensagem de *INVITE* enviada por Alice continha

o endereço 203.110.43.5. Quando a mensagem *'200 OK'* enviada por Bob chega ao servidor SIP de Alice, este altera o campo *'Contact'* inserindo o endereço IP do servidor de interceptação. Finalizando o estabelecimento da chamada, Alice envia um ACK para o servidor de interceptação que o encaminha para Bob.

Uma vez realizado o direcionamento do fluxo RTP para o servidor de interceptação, toda a conversação da chamada pode ser gravada. Ferramentas como o Wireshark[13] são capazes de decodificar o áudio contido nos pacotes RTP e reproduzi-lo. Portanto, é possível realizar a escuta da chamada, além de salvar os pacotes para utilizações posteriores.

Entretanto, para que haja conversação, os pacotes enviados para o servidor de interceptação precisam ser encaminhados para os seus destinos de fato. Para tal, é necessária uma ferramenta capaz de modificar os endereços de origem e destino dos pacotes que chegam ao servidor de interceptação (ver Figura 13). Uma vez realizado encaminhamento dos pacotes, a interceptação da chamada é realizada.

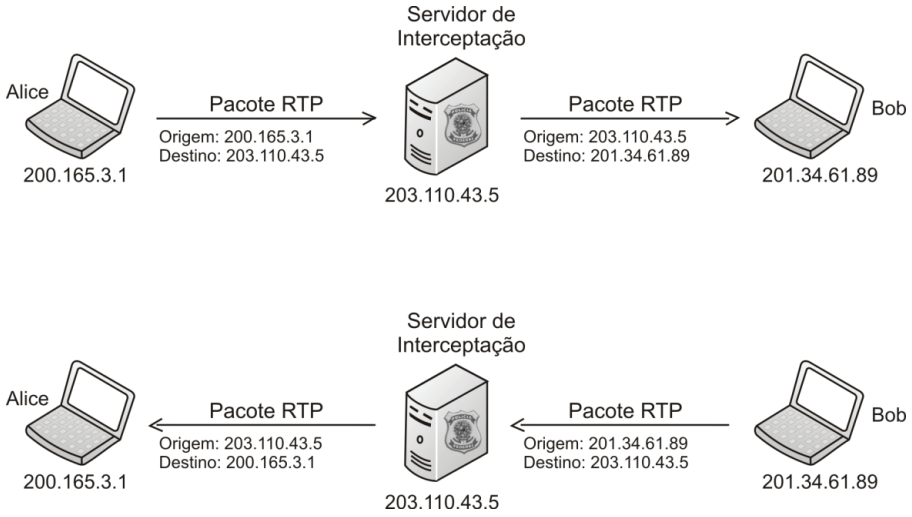


Figura 13 - Servidor de Interceptação - Encaminhamento de pacotes

[13] http://www.wireshark.org/ (último acesso em 17/11/08)

Contudo, ainda é necessário prover um mecanismo para garantir a rastreabilidade da chamada. Durante a fase de sinalização, o provedor VoIP do indivíduo sob investigação é capaz de identificar os endereços IP reais dos participantes. Isso é possível, pois, conforme apresentado na seção 3.2.1.2, o campo '*Contact'* das mensagens *INVITE* e *'200 OK'* enviadas durante a sinalização possui o endereço IP dos participantes. Portanto, para identificar os participantes do fluxo RTP basta o provedor de Internet, que controla o roteador diretamente ligado a Bob, checar se os pacotes do fluxo possuem o endereço IP do servidor de interceptação como origem. Desta forma, prova-se que Alice está realizando uma chamada com Bob.

Porém, um dos participantes pode utilizar algum mecanismo de anonimidade, como por exemplo, o TOR[14] ou o *findnot.com* para garantir o seu anonimato na Internet. Neste cenário, não é possível realizar a identificação dos participantes da chamada visto que o endereço IP de origem dos pacotes RTP não serão o endereço IP do servidor de interceptação, mas sim, o endereço IP de uma das máquinas do provedor do serviço de anonimato. Portanto, esta abordagem não é suficiente para prover a rastreabilidade da chamada.

Este trabalho propõe um mecanismo baseado em esteganografia para realizar o rastreamento da chamada de maneira eficaz. Quando o fluxo RTP passa pelo servidor de interceptação, este insere uma marca d'água nos pacotes para identificá-los (ver Figura 14). Logo, basta o provedor de Internet checar se o fluxo RTP que atravessa o seu roteador diretamente conectado a Bob possui a marca d'água. Desta forma, é possível garantir que o fluxo RTP recebido por Bob é o fluxo enviado por Alice, interceptado pela agência de monitoração. Portanto, através do mecanismo proposto, é possível identificar os participantes da chamada e garantir a rastreabilidade.

[14] http://www.torproject.org/ (último acesso em 21/11/08)

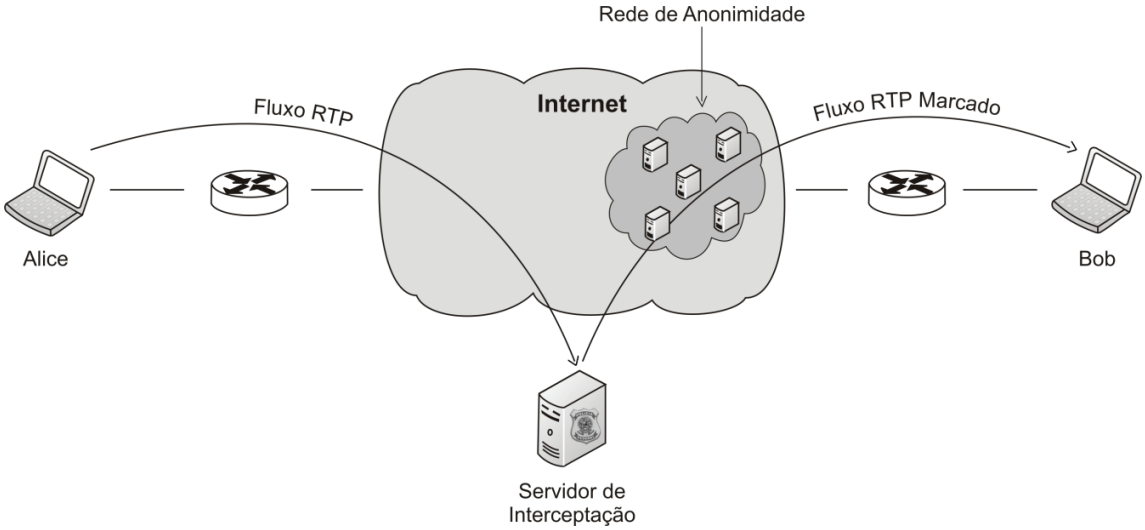


Figura 14 - Rastreamento de Chamada VoIP

A técnica de esteganografia utilizada neste trabalho chama-se *Least Significant Bit (LSB)*. Nesta técnica, o bit menos significativo de um determinado dado é utilizado para transportar informação. Essa técnica é comumente utilizada para ocultar mensagens em arquivos de imagem, como JPEG ou BMP. Geralmente, em uma imagem cada pixel é representado pelas cores do RGB. Cada cor contém um *byte* que representa a intensidade da componente. Considerando somente a cor vermelha, existem 2^8 tonalidades de vermelho. A diferença de tonalidade do vermelho representado por 11111111 e 11111110 é imperceptível para o olho humano, portanto, o bit menos significativo pode ser utilizado para outros fins. Da mesma forma, o bit menos significativo dos *bytes* dos dados de voz contidos no pacote RTP são utilizados para transportar a marca d'água que identifica a origem do pacote, provendo, assim, a rastreabilidade da chamada. Apesar de, possivelmente, impactar na qualidade da chamada, a marca d'água é inserida nos dados de voz do RTP para dificultar a sua detecção.

5.4 IMPLEMENTAÇÃO

Como resultado de implementação deste trabalho foram criados dois artefatos: um *software* para realizar o encaminhamento e um *software* para inserir a marca d'água nos pacotes.

Não foi necessário realizar uma implementação de fato de um *software* para realizar o encaminhamento dos pacotes, pois o próprio sistema operacional, através de suas tabelas de roteamento, pode fornecer esta funcionalidade. Porém, para que o sistema operacional realize o encaminhamento dos pacotes é necessário modificar o endereço de destino dos pacotes, visto que eles são endereçados ao servidor de interceptação. Para realizar a troca dos endereços também não foi necessária uma implementação, pois a ferramenta Iptables[15] é capaz de realizar esta tarefa. O Iptables, presente na maioria das distribuições Linux, é um ferramenta que permite a criação de regras de *firewall* e *NAT*. Através da inserção de algumas regras específicas de NAT é possível alterar os endereços IP de origem e destino dos pacotes interceptados. Um *shell script* foi criado com o objetivo de automatizar a inserção e remoção dessas regras.

Para prover a rastreabilidade dos pacotes interceptados foi escrito um módulo para o kernel do Linux versão 2.6.24, chamado *RTP Marker*. Uma vez carregado no kernel, o *RTP Marker* insere uma determinada marca d'água, configurada anteriormente, em todos os pacotes RTP da chamada interceptada. O *RTP Marker* utiliza o *Netfilter* que é uma parte do kernel do linux localizado na camada de rede da pilha TCP/IP do kernel. O *Netfilter* define cinco ganchos (*hooks*) localizados conforme ilustrado na Figura 15.

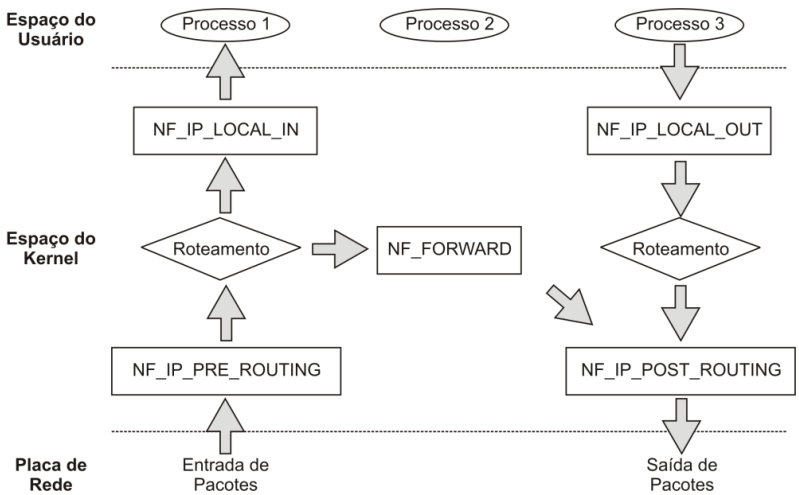


Figura 15 - Ganchos do Netfilter

[15] http://www.netfilter.org/projects/iptables/index.html (último acesso em 17/11/08)

Quando um pacote passa por um gancho, o *Netfilter* chama uma função do módulo do kernel registrada como função de *callback* para aquele gancho. Desta forma, o módulo *RTP Marker* registra uma função de *callback* no gancho NF_IP_POST_ROUTING para interceptar todos os pacotes que o servidor de interceptação está encaminhando. Neste ponto é inserida a marca d'água nos dados de voz contidos no pacote RTP.

A técnica utilizada para inserir a marca d'água nos pacotes deve ser apropriada para o *codec* utilizado na chamada. Os *codecs* possuem formas diferentes de codificar a informação, por este motivo, uma técnica de marcação desenvolvida para um determinado *codec*, não necessariamente apresentará uma boa performance quando utilizada em outro *codec*. Neste trabalho, foi desenvolvida uma técnica de marcação para o *codec* G.711 PCMU. A Figura 16 mostra o processo de codificação e decodificação do G.711 PCMU.

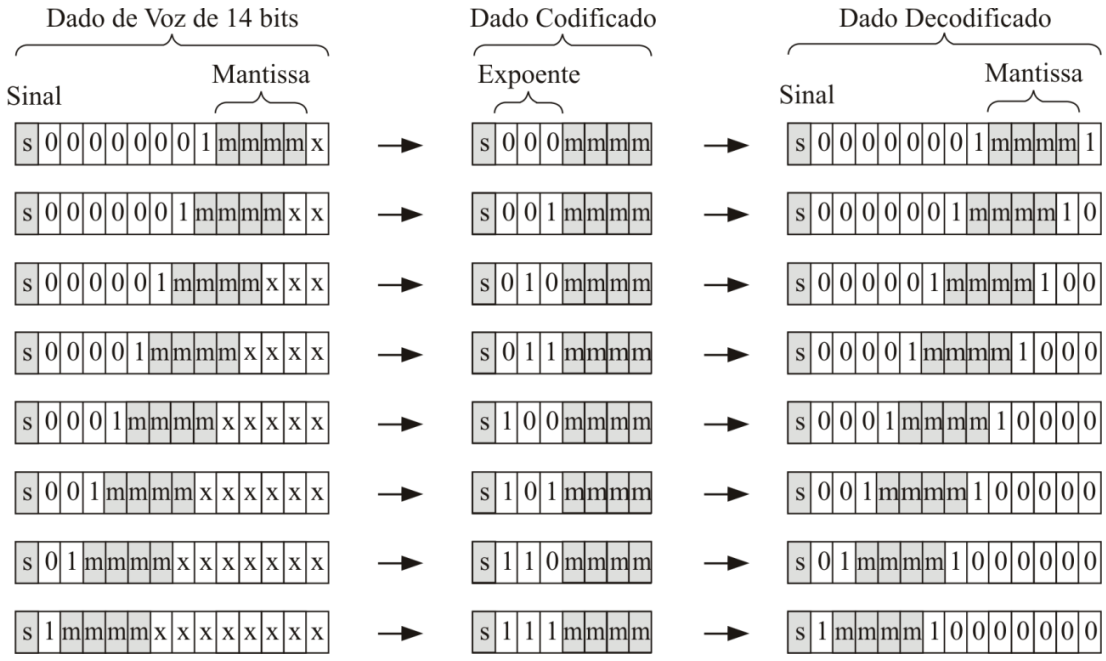


Figura 16 - Codificação/Decodificação G.711 PCMU

O G.711 codifica um dado de voz de quatorze bits em um dado de oito bits. Analisando o processo de codificação dos dados, concluiu-se que a melhor posição para inserir a marca d'água é no bit menos significativo da mantissa, pois alterar informação de sinal e expoente afetaria bastante o dado quando decodificado.

A marca d'água implementada neste trabalho é representada por um *byte*. Durante o processo de marcação, ela é inserida nos dados de voz do pacote RTP conforme ilustrado na Figura 17.

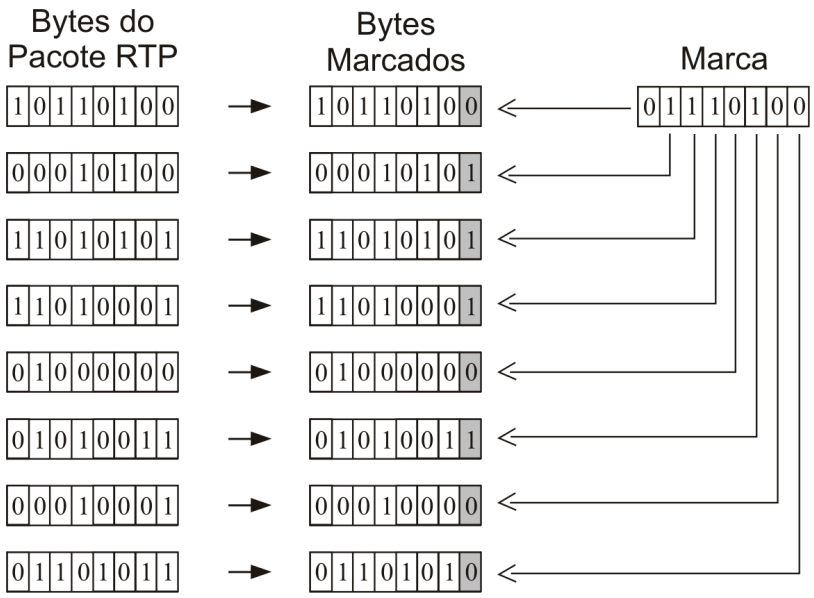


Figura 17 - Inserção da marca d'água

Troca-se o bit menos significativo dos últimos oito *bytes* do pacote RTP por um bit da marca d'água. Portanto, para determinar se um fluxo RTP foi interceptado, basta observar se os bits menos significativos dos últimos oito *bytes* de cada pacote pertencente ao fluxo formam a marca d'água.

Contudo a marca d'água não necessariamente ocorrerá nos últimos oito *bytes*, ela pode ser inserida em quaisquer *bytes* do campo de dados do pacote RTP. Antes da realização da chamada, a agência de monitoração define qual será o valor da marca d'água e em quais *bytes* ela

será inserida. Essas informações são conhecidas apenas pela agência de monitoração. Dessa forma, a detecção da existência da marcação e, conseqüentemente, a existência da interceptação torna-se complexa.

O controle sobre a sinalização não foi implementado, pois não faz parte do escopo deste trabalho. Porém, uma implementação possível seria modificar o *software* do servidor SIP para que antes de realizar o encaminhamento de uma mensagem *INVITE*, verificar se o usuário em questão pertence a uma lista, fornecida pela agência de monitoração, contento todos os assinantes do serviço VoIP que estão sob investigação. Em caso positivo, o servidor SIP realiza a manipulação da sinalização descrita na seção 5.3 para que a chamada seja direcionada para servidor de interceptação.

5.5 SUMÁRIO

Este capítulo apresentou um sistema para a realização de interceptação legal de chamadas VoIP baseadas no protocolo SIP. Foram discutidas duas abordagens de interceptação: uma baseada no roteador em que o indivíduo sob investigação conecta-se diretamente para acessar a Internet e uma baseada na interceptação da sinalização da chamada realizada pelo provedor do serviço de voz sobre IP. Constatou-se que ambas, isoladamente, são incapazes de realizar a interceptação legal com sucesso, porém, se utilizadas em conjunto, é possível realizá-la de forma eficaz.

Uma vez definida a abordagem a ser utilizada, foi proposta uma arquitetura capaz de realizar interceptações. Nesta arquitetura, apenas um elemento foi criado (servidor de interceptação) e adicionado à arquitetura SIP, mantendo a simplicidade do sistema e facilitando o gerenciamento das interceptações.

Foi desenvolvido um mecanismo para prover rastreabilidade de chamadas VoIP através de técnicas de esteganografia. Esse mecanismo é baseado na inserção de uma marca d'água nos pacotes do fluxo RTP da chamada. Além disso, foi desenvolvida uma técnica de esteganografia para fluxos RTP codificados através do *codec* G.711.

Como prova dos conceitos desenvolvidos neste capítulo, foi realizada a implementação do servidor de interceptação para o sistema operacional Linux.

6 METODOLOGIA E AVALIAÇÃO DE DESEMPENHO

Este trabalho tem com um dos objetivos avaliar o impacto do sistema de interceptação legal de chamadas proposto na qualidade da chamada. Para tal, é necessário apresentar o ambiente de execução dos experimentos, as métricas utilizadas e, finalmente, apresentar como os experimentos foram conduzidos. A metodologia deste trabalho foi baseada na metodologia utilizada em [33] .

6.1 AMBIENTE DE REALIZAÇÃO DE EXPERIMENTOS

O ambiente de execução dos experimentos foi elaborado de forma a permitir a replicação dos mesmos. Para a realização dos experimentos foram utilizadas três máquinas com as seguintes especificações:

- Máquina E (Emissor): PC, Windows XP, Pentium 4 - 2,8GHz, 1GB.
- Máquina R (Receptor): PC, Windows XP, Pentium 4 - 2,8GHz, 1 GB.
- Máquina SI (Servidor de Interceptação): Laptop, Ubuntu 8.04, Intel core 2 duo, 2 GB.

Os experimentos foram realizados em uma rede local. A Figura 18 ilustra o ambiente de execução dos experimentos.

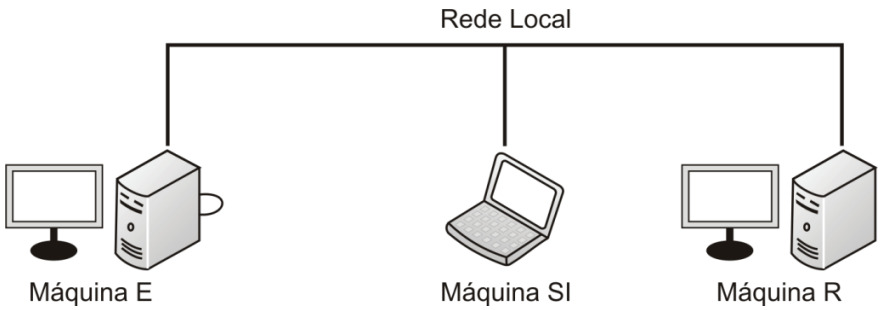


Figura 18 - Ambiente de realização dos experimentos

A máquina E é responsável por estabelecer uma chamada VoIP e transmitir um fluxo de áudio com destino final a máquina R. Para a realização da chamada foi utilizado o software

Ekiga[16]. O Ekiga é uma aplicação VoIP *open source*, através da qual é possível controlar o *codec* que será utilizado na chamada. A máquina E possui um cabo que transmite a saída de áudio para a entrada de áudio. Este cabo é necessário para capturar o som que é reproduzido pelo Winamp[17] e enviá-lo para a entrada de áudio para que este som seja transmitido pelo Ekiga. A máquina R também executa o Ekiga e grava o áudio recebido através de um *software* chamado Audacity[18]. A máquina SI é responsável por realizar a interceptação da chamada. Ela executa o Wireshark[19] para capturar os pacotes transmitidos pelas máquinas E e R, além de executar os artefatos apresentados na seção 5.4.

6.2 MÉTRICAS

Conforme apresentado na seção 3.1, a taxa de perda de pacotes, o *jitter* e o atraso podem ser considerados métricas para a avaliação de desempenho de sistemas VoIP. Porém, todas essas métricas refletem diretamente na qualidade da chamada. Portanto, é mais interessante utilizar a qualidade da chamada como métrica para avaliar o desempenho do sistema proposto.

O *Mean Opinion Score (MOS)* é uma métrica adotada mundialmente para a avaliação da qualidade de ligações telefônicas. O MOS foi padronizado pelo ITU através das recomendações P.800 [34] e P.830 [35] . O MOS é uma abordagem de avaliação subjetiva calculado como uma média de notas individuais atribuídas por um grande número de pessoas que ouvem o áudio resultante de um processo de codificação e decodificação, onde a nota varia de 1 (ruim) a 5 (excelente). Embora seu resultado seja bastante relevante, a dificuldade de realizar tal avaliação em larga escala motivou o desenvolvimento de técnicas objetivas para o cálculo do MOS.

Existem várias técnicas para o cálculo do MOS, como por exemplo o Modelo-E [36] , o *Perceptual Speech Quality Measurement (PSQM)* [37] , *Perceptual Analysis Measurment System (PAMS)* [38] e o *Perceptual Evaluation of Speech Quality (PESQ)* [39] . Neste trabalho, foi utilizada a técnica PESQ, pois ela combina características do PSQM e PAMS e apresenta

[16] http://www.gnomemeeting.org/ (último acesso em 17/11/08)
[17] http://www.winamp.com/ (último acesso em 17/11/08)
[18] http://audacity.sourceforge.net/ (último acesso em 17/11/08)
[19] http://www.wireshark.org/ (último acesso em 17/11/08)

melhorias nos algoritmos, atendendo a uma gama maior de cenários. Para medir a qualidade, o PESQ baseia-se na comparação do áudio original com o áudio recebido, possivelmente degradado pelo processo de codificação/decodificação e transmissão.

6.3 DESCRIÇÃO DOS EXPERIMENTOS

Para avaliar o impacto da solução proposta sobre a qualidade da chamada foram realizados experimentos em dois cenários:

- **Cenário de Ligação Direta**: neste cenário, o Emissor envia os pacotes de áudio diretamente para o Receptor, simulando uma chamada VoIP normal.
- **Cenário de Ligação Interceptada**: neste cenário, o Emissor envia os pacotes de áudio para o servidor de interceptação, que por sua vez insere a marca d'água e realiza o encaminhamento dos pacotes.

Uma vez estabelecida a chamada entre as máquinas E e R, a máquina E inicia a reprodução do arquivo de áudio, que por sua vez é transmitido para a máquina R utilizando o codec G.711 PCMU. A máquina R gravar o áudio que é reproduzido pelo Ekiga. No final do experimento, tem-se como produto o arquivo de áudio, provavelmente degradado devido ao processo de codificação/decodificação e transmissão, que foi gravado pela máquina R. Desta forma tem-se os arquivos de áudio original (tocado no Emissor) e o arquivo de áudio degradado (gravado pelo Receptor) necessário para a avaliação do MOS através do PESQ.

O arquivo de áudio utilizado nos experimentos possui 60 minutos de duração e foi construído utilizando duas vozes masculinas e duas vozes femininas, incluindo pausas, conforme recomendado pela ITU.

6.4 AVALIAÇÃO DE DESEMPENHO

De acordo com os critérios e cenários apresentados neste Capítulo foi realizada uma avaliação do desempenho do sistema de interceptação legal de chamadas VoIP proposto neste trabalho.

A Figura 19 apresenta a variação do MOS ao longo do tempo em ambos os experimentos. Nota-se que as curvas estão bastante próximas, logo, não houve degradação expressiva da qualidade da ligação ao se realizar a interceptação da chamada.

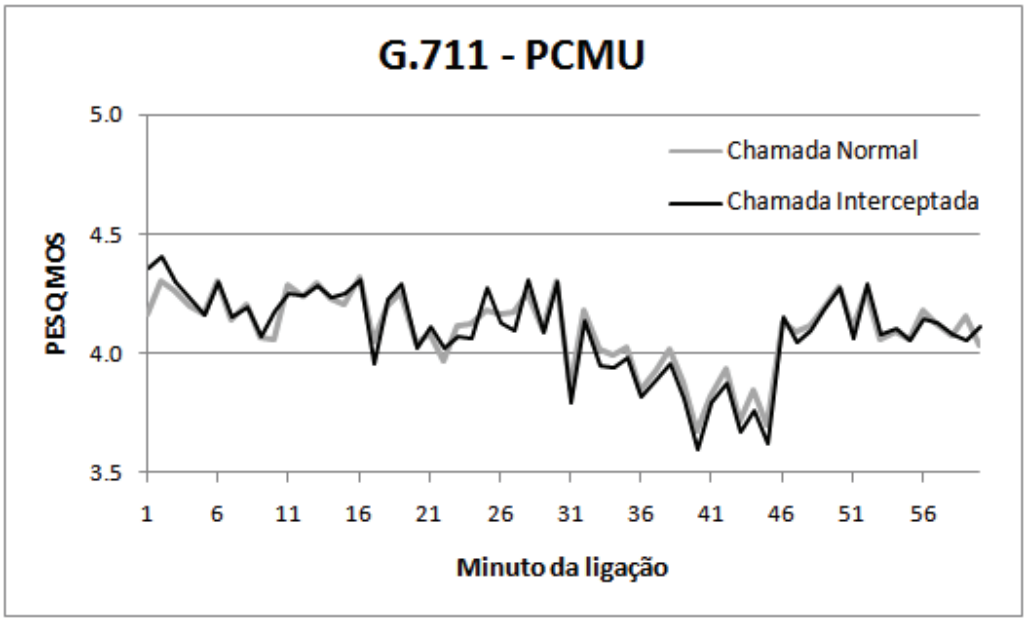
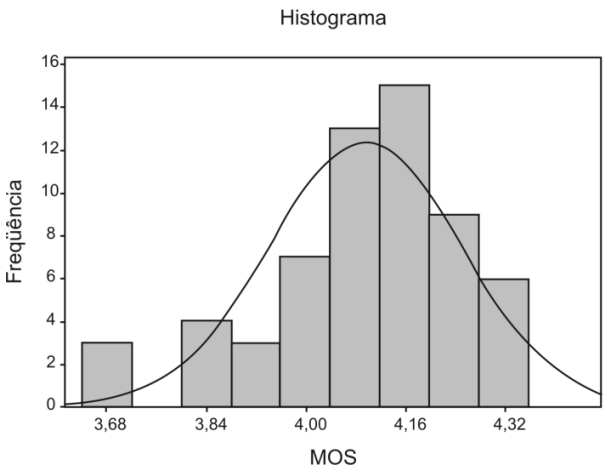


Figura 19 - Variação do MOS ao longo da chamada (G.711)

Em ambos os cenários, o MOS foi calculado para cada minuto da ligação. A Figura 20 apresenta um resumo sobre o experimento realizado no cenário que simula uma ligação normal.

O histograma mostra a distribuição dos valores de MOS calculados ao longo do experimento. Nota-se que o valor médio do MOS (4,0967) está de acordo com o valor encontrado na literatura para o *codec* G.711 (ver seção 3.3).

Chamada Normal



Média	4,0967
Desvio Padrão	0,1552
Variância	0,0241
Número de Amostras	60
Mínimo	3,6760
Máximo	4,3190
Intervalo de Confiança de 99.9% para a Média	
4,0273	4,1661

Figura 20 - Resumo do Experimento de Chamada Normal

Da mesma forma, a Figura 21 apresenta um resumo do experimento realizado no cenário em que a chamada é interceptada através do sistema proposto.

Nota-se que o experimento que simula uma chamada interceptada apresenta um valor médio de MOS bastante próximo do valor médio do MOS apresentado pelo experimento que simula uma chamada normal, e, conseqüentemente, também está de acordo com o valor encontrado na literatura. Visto que há interseção entre os valores pertencentes ao intervalo de confiança dos experimentos, pode-se afirmar que, estatisticamente, não há indícios que o sistema de interceptação proposto neste trabalho degrada, sistematicamente, a qualidade da chamada. Portanto, conclui-se que a arquitetura de interceptação legal de chamadas VoIP proposta é válida, pois realiza o seu objetivo com sucesso e eficiência.

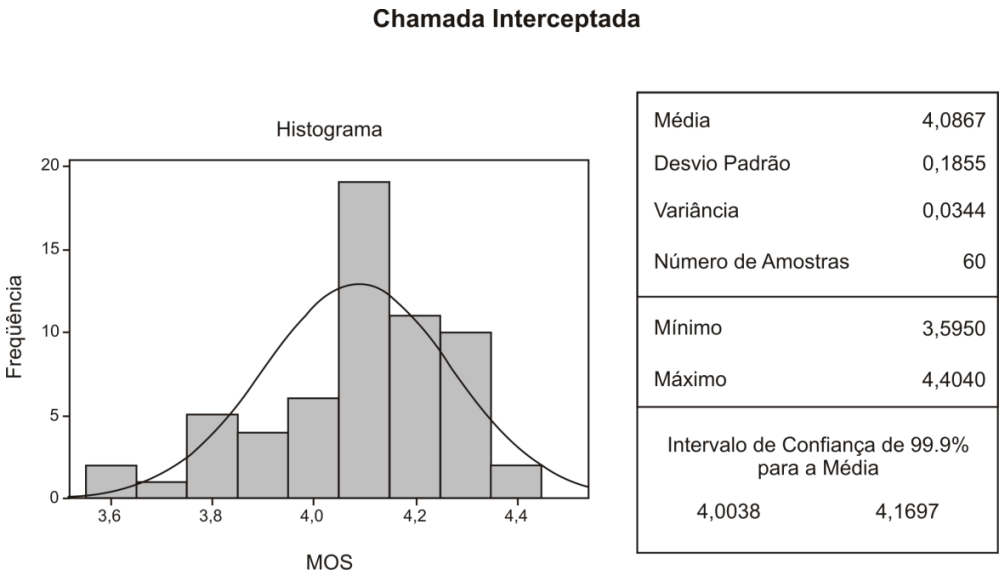


Figura 21 - Resumo do Experimento de Chamada Interceptada

Adicionalmente, foram realizados experimentos utilizando os *codecs* G.726 e GSM. Conforme esperado (ver seção 5.4), houve uma degradação considerável da qualidade da chamada. Na Figura 22 e Figura 23 é apresentada a variação do MOS das chamadas realizadas com os codecs G.726 e GSM, respectivamente.

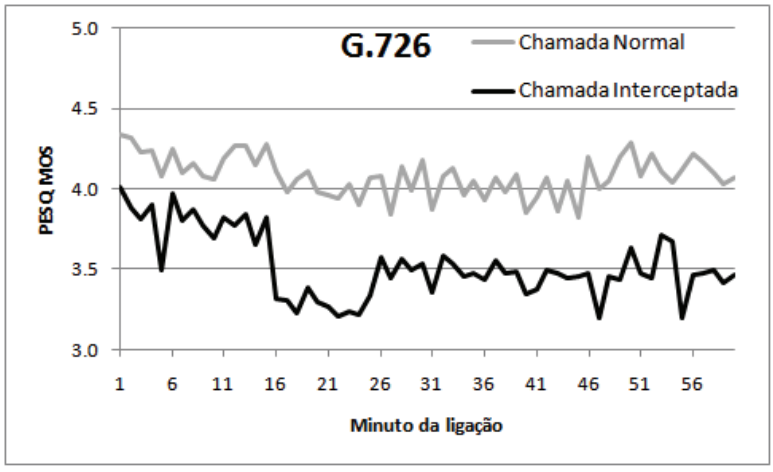


Figura 22 - Variação do MOS ao longo da chamada (G.726)

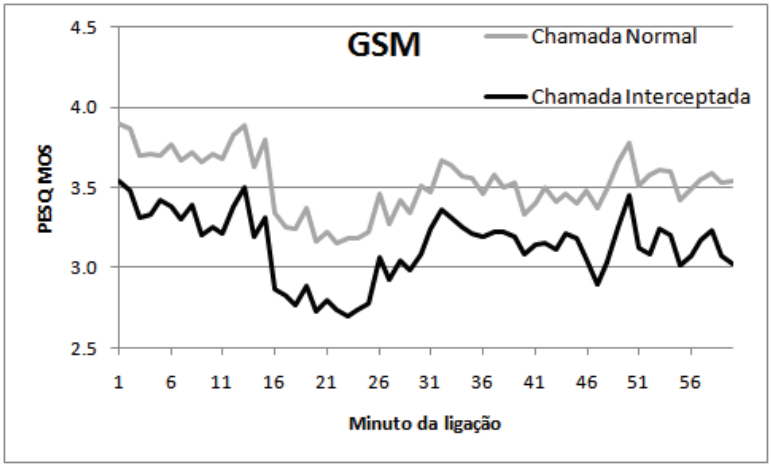


Figura 23 - Variação do MOS ao longo da chamada (GSM)

Porém, utilizando-se um artifício de taxa de amostragem (*sampling)* acredita-se que, como medida paliativa, a técnica de esteganografia para o *codec* G.711 desenvolvida neste trabalho pode ser utilizada. Dessa maneira, a degradação da qualidade da chamada será menor, visto que nem todos os pacotes serão marcados.

7 CONCLUSÃO

7.1 DISCUSSÕES

O sistema proposto neste trabalho é capaz de realizar a interceptação em cenários onde os participantes são terminais VoIP e em cenários onde a chamada é estabelecida entre um terminal VoIP e um terminal do sistema telefônico. O segundo cenário[1] é possível, pois de acordo com o mecanismo descrito na seção 5.3, o *Media Gateway* será configurado pelo MGC para enviar o seu fluxo RTP para o servidor de interceptação. Contudo, para prover a rastreabilidade da chamada será necessária a cooperação da operadora de telecomunicações responsável linha telefônica utilizada pelo *Media Gateway.*

O sistema proposto também é capaz de lidar com situações em que, devido à mobilidade oferecida pelas redes IP, um dos participantes migre entre várias redes IP durante a execução da chamada. Durante a migração, o participante envia uma mensagem *INVITE* informando o seu novo endereço IP no campo *'Contact',* este artifício é chamado de *RE-INVITE* [17] . Desta forma, o sistema de interceptação tem acesso à informação sobre novo endereço IP onde o participante se encontra.

Adicionalmente, o sistema proposto é capaz de controlar o *codec* que será utilizado na chamada. Conforme apresentado na seção 3.2.1, a informação sobre o *codec* utilizado durante a chamada é enviada através do protocolo SDP dentro da mensagem SIP. Uma vez que a mensagem SIP pode ser manipulada pelo sistema proposto, o *codec* pode ser controlado. Isso é bastante interessante, pois se a chamada for codificada com um *codec* não suportado pelo servidor de interceptação, não é possível decodificar os dados e realizar a escuta.

7.2 CONSIDERAÇÕES FINAIS

Neste trabalho foi proposto um sistema capaz de realizar a interceptação legal de chamadas VoIP. Como prova dos conceitos utilizados na construção do sistema proposto foram realizadas implementações e avaliação de desempenho.

Alguns trabalhos como [9] e [10] , apresentados na seção 1.4, não consideram o fato de que o indivíduo sob investigação pode conectar de qualquer lugar da Internet. Além disso, essas propostas necessitam ser implantadas na rede local do indivíduo, sendo algo bastante intrusivo.

Embora em [11] seja apresentada uma proposta de arquitetura capaz de interceptar chamadas VoIP baseadas no protocolo SIP, não foi desenvolvido um mecanismo para prover a rastreabilidade da chamada conforme apresentado na seção 5 do artigo (Limitações). Também não foi realizada uma avaliação de desempenho do sistema proposto.

7.3 CONTRIBUIÇÕES

Como principais contribuições deste trabalho, podem-se destacar:

- O desenvolvimento de uma arquitetura capaz de realizar interceptação legal de chamadas VoIP baseadas no protocolo SIP;
- Mecanismo de rastreamento de chamadas VoIP baseado em esteganografia;
- O desenvolvimento de uma técnica de esteganografia capaz de inserir informação em dados de voz codificados pelo *codec* G.711 PCMU;
- Elaboração de ferramentas capazes de interceptar e rastrear chamadas;
- Avaliação de desempenho do sistema proposto.

7.4 TRABALHOS FUTUROS

Como trabalhos futuros que visam estender e aprimorar o sistema de interceptação de chamadas VoIP proposto neste trabalho de graduação, podem-se citar:

- Desenvolvimento de técnicas de esteganografia para outros *codecs;*
- Avaliação da escalabilidade do sistema proposto para descobrir o número de interceptações simultaneamente suportado pelo servidor de interceptação sem que haja degradação da qualidade das chamadas;
- Adaptar o sistema para contemplar cenários em que são realizadas chamadas com mais de dois participantes (conferências);

- Adicionar suporte a outros protocolos de sinalização, como por exemplo, H.323 e IAX2.

REFERÊNCIAS

[1] Nokia, "VoIP: Convergência IP em Mobilidade", White Paper, 2006.

[2] D. Collins, *"Carrier Grade Voice Over IP",* McGraw-Hill, 2001.

[3] Communications Assistance for Law Enforcement Act, Pub. L. No. 103-414, 108 Stat. 4279, codified at 47 USC 1001-1010, 1994.

[4] Federal Communications Commission, "FCC Requires Certain Broadband and VoIP Providers to Accommodate Wiretaps", Maio, 2005.

[5] S. Bellovin et al., "Security Implications of Applying the Communications Assistance to Law Enforcement Act to Voice-over IP", ITAA, Junho, 2006.

[6] F. Baker, B. Foster e C. Sharp, *"Cisco Architecture for Lawful Interception in IP networks",* IETF, RFC 3294, Outubro, 2004.

[7] AQSACOM, *"Lawful Interception for IP Networks"*, White Paper, Novembro, 2005.

[8] N. Maloku, T. Aljaz e F. Dolenc, *"Legal call interception in next generation networks"*, Proceedings of ConTEL, Vol 1, pp. 47-50, 2003.

[9] A. Milanovic, et al., "Methods for Lawful Interception in IP Telephony Networks Based on H.323", IEEE EUROCON, vol.1, pp.198-202, Setembro, 2003.

[10] A. Milanovic, et al., "Distributed System for Lawful Interception in VoIP Networks", IEEE EUROCON, vol. 1, pp. 203-207, Setembro, 2003.

[11] B. Karpagavinayagam, R. State, e O. Festor. , *"Monitoring Architecture for Lawful Interception in VoIP Networks",* Second International Conference on Internet Monitoring and Protection - ICIMP 2007, Julho, 2007.

[12] T. Smith, *"Anatomy of Telecommunications"*, ABC TeleTraining, 1987.

[13] E. Sutherland, *"Enterprise VoIP Adoption? Gradual but Rapid, Say Experts"*, Wifi Planet, Março, 2005. http://www.wi-fiplanet.com/voip/article.php/3493136 (último acesso em 15/09/2008)

[14] A. Stone, *"While VoIP Adoption Explodes in Enterprise, Carrier Spending Lags",* VoIP Planet, Setembro, 2008. http://www.voipplanet.com/trends/article.php/3771266 (último acesso em 15/09/2008)

[15] D. Cohen, "Specifications for the Network Voice Protocol (NVP)", IETF RFC 741, Novembro, 1977.

[16] M. Handley et al., "SIP: Session Initiation Protocol", IETF RFC 2543, Março, 1999.

[17] J. Rosenberg et al., "SIP: Session Initiation Protocol", IETF RFC 3261, Junho, 2002.

[18] M. Handley et al., "SDP: Session Description Protocol", IETF RFC 4566, Julho, 2006.

[19] F. Andreasen e B. Foster, *"Media Gateway Control Protocol"*, IETF RFC 3435, Janeiro, 2003.

[20] H. Schulzrinne et al., "RTP: A Transport Protocol for Real-Time Applications", IETF RFC 3550, Julho 2003.

[21] H. Schulzrinne et al., "RTP: A Transport Protocol for Real-Time Applications", IETF *RFC* 1889, Janeiro, 1996.

[22] M. Baugher et al., "The Secure Real-time Transport Protocol (SRTP)", IETF RFC 3711, Março, 2004.

[23] H. Krawczyk, M. Bellare e R. Canetti, *"HMAC: Keyed-Hashing for Message Authentication"*, IETF RFC 2104, Fevereiro, 2004.

[24] M. Spencer et al., "IAX: Inter-Asterisk eXchange Version 2", Draft IETF, Outubro, 2008.

[25] International Telecommunication Union Telecommunication Standardization Sector ITU-T, "*H.323, Packet Based Multimedia Communications Systems"*, Junho, 2006.

[26] Global IP Sound, *"iLBC – Designed for the Future"*, White Paper, Outubro, 2004.

[27] L. Ding, A. Radwan, M. S. El-Hennawey, e R. A. Goubran, *"Performance Study of Objective Voice Quality Measures in VoIP,"* Proceedings of IEEE Symposium on Computers and Communications (ISCC'07), pp. 197 – 202, Julho, 2007.

[28] B. Goode, *"Voice Over Internet Protocol (VoIP)"*, Proceedings of IEEE, Vol. 90, No. 9, Setembro, 2002.

[29] R. Beuran, *"VoIP over Wireless LAN Survey"*, research report, IS-RR-2006-005, Japan Advanced Institute of Science and Technology (JAIST), Abril, 2006.

[30] N. Kitawaki, K. Nagai e T. Yamada, *"Objective Quality Assessment of Wideband Speech Coding"*, IEICE Transactions, Vol 88-B, pp.1111-1118 , 2005.

[31] J. P. B. Júnior, "Dez Anos da Lei da Interceptação Telefônica (Lei nº 9.296/96). Interpretação Jurisprudencial e Anteprojeto de Modificação". Revista Jurídica (Porto Alegre), v. 350, p. 240-271, 2006.

[32] L. Leite, *"Interceptação Autorizada de Chamadas Telefônicas",* Teleco, Janeiro, 2006. http://www.teleco.com.br/tutoriais/tutorialinterceptacao/pagina_1.asp (último acesso em 24/11/08)

[33] R. S. B. G. Barbosa, "Avaliação de Desempenho de Aplicações VoIP P2P", Dissertação de Mestrado, Centro de Informática, Universidade Federal de Pernambuco, 2007.

[34] International Telecommunication Union, "*Methods for subjective determination of transmission quality"*, Recomendation P.800, Agosto, 1996.

[35] International Telecommunication Union, "Subjective performance assessment of telephone-band and wideband digital codecs", Recomendation P.830, Fevereiro, 1996.

[36] International Telecommunication Union, "*The E-model, a computation model for use in transmission plainning"*, Recomendation G.107, Dezembro, 1998.

[37] International Telecommunication Union, *"Objective quality measurement of telephone-band (300-3400 Hz) speech codecs"*, Recomendation P.861, Fevereiro, 1998.

[38] A. W. Riz e M. P. Hollier, "The percentual analysis measurement system for robust end-to-end speech quality assessment", IEEE ICASSP, Junho, 2000.

[39] International Telecommunication Union, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone network and speech codecs", Recomendation P.862, Fevereiro, 2001.

[40] C. Benvenuti, *"Understanding Linux Network Internals"*, O'Reilly, Dezembro 2005.