UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

IMPLEMENTAÇÃO DE UM SISTEMA ESTEGANOGRÁFICO PARA INSERÇÃO DE TEXTOS EM SINAIS DE ÁUDIO

por

JINNETT PAMELA CARRIÓN CASIERRA

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do grau de Mestre em Engenharia Elétrica - Telecomunicações.

Orientador: Hélio Magalhães de Oliveira, Docteur
Co-Orientador: Ricardo Menezes Campello de Souza, PhD

Recife, Março de 2009

C339i **Casierra, Jinnett Pamela Carrión.**

Implementação de um sistema esteganográfico para inserção de textos em sinais de áudio / Jinnett Pamela Carrión Casierra. – Recife: O Autor, 2009.

vii, 134 folhas, il : grafs., tabs., figs.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2009.

Inclui Bibliografia e Anexos

1. Engenharia Elétrica. 2. Ocultação de Informação. 3. Esteganografia. 4. Dados. 5. Sinais de Áudio. 6. *Wavelet*. I. Título.

UFPE

621.3 **CDD (22. ed.)** **BCTG/2009-091**

Universidade Federal de Pernambuco

Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE DISSERTAÇÃO DO MESTRADO ACADÊMICO DE

JINNETT PAMELA CARRIÓN CASIERRA

TÍTULO

"IMPLEMENTAÇÃO DE UM SISTEMA ESTEGANOGRÁFICO PARA INSERÇÃO DE TEXTOS EM SINAIS DE ÁUDIO"

A comissão examinadora composta pelos professores: HÉLIO MAGALHÃES DE OLIVEIRA, DES/UFPE, RAFAEL DUEIRE LINS, DES/UFPE e BENJAMÍN RENÉ CALLEJAS BEDREGAL, DIMA/UFRN sob a presidência do primeiro, consideram a candidata **JINNETT PAMELA CARRIÓN CASIERRA APROVADA.**

Recife, 23 de março de 2009.

EDUARDO FONTANA
Coordenador do PPGEE

HÉLIO MAGALHÃES DE OLIVEIRA
Orientador e Membro Titular Interno

BENJAMÍN RENÉ CALLEJAS BEDREGAL
Membro Titular Externo

RAFAEL DUEIRE LINS
Membro Titular Interno

"A MEUS SERES AMADOS"

AGRADECIMENTOS

Quando acontecem mudanças em nossas vidas surgem expectativas, novos desafios e duvidas. Foram justamente novas experiências enriquecedoras que surgiram no recorrer destes dois anos de Mestrado. Outra cultura, outro idioma, novas metodologias de ensino e novas pessoas, que acompanharam minha estadia. A saudade pelo meu país foi muita, mas as vivências, a acolhida e as oportunidades que me deu o Brasil me fizeram sentir como em casa; assim quisera reiterar minha gratidão com meus mais sinceros agradecimentos:

No primeiro lugar a *Deus*, que é a luz que guia meu caminho dia a dia, a força que me permite continuar sempre em frente e que faz que as melhores coisas aconteçam na minha vida.

Um agradecimento especial ao meu orientador, *Professor Dr. Hélio Magalhães de Oliveira* por todo o seu apoio, seu otimismo na realização do projeto e o interesse na pesquisa, por seu tempo para aclarar minhas dúvidas: obrigada por compartilhar a cultura e costumes do Brasil. Além de ser um orientador solícito é um excelente professor; consegue com o seu dinamismo transmitir e compartilhar seus conhecimentos com os estudantes.

Ao Professor Ricardo Campello, por o aporte de idéias e contribuições no desenvolvimento do trabalho e pelas valiosas observações realizadas.

A minha Mãe, que é minha melhor conselheira, obrigada pelas largas conversas, por seus conselhos, seus ânimos e seu amor.

A minha Irmã, obrigada por ser minha amiga incondicional.

A minha tia Salma, obrigada pelo recebimento no seu lar, por seu acolhimento, sua paciência, seus conselhos e por facilitar a convivência, que permitiram minha dedicação por completo à elaboração da dissertação.

A Jorge, obrigada pelos seus ensinamentos, pelas aulas em casa, porque sua dedicação e seu entusiasmo por aprender sempre coisas novas é tal, que termina contagiando as pessoas que se encontram perto.

A meu Pai, *familiares* e *amigos* que desde a distância me transmitiam o seu carinho, ânimo e palavras de conforto.

Tenho que dar também graças pela presencia de uma acompanhante, *Cleo*, sobretudo nas noites de trabalho. Obrigada Salma e Jorge por adaptar-se a presencia e convivência com Cleo.

Aos Professores: Ronaldo Aquino, Hélio Magalhães, Ricardo Campello, Valdemar Rocha, Fernando Campello, por seus ensinamentos, porque tive a oportunidade de assistir suas aulas.

Aos professores da área da coordenação da pós-graduação: Joaquim Martins e Eduardo Fontana.

Aos colegas da turma de Inteligência Artificial, a primeira turma com a qual compartilhei; muito obrigada pela sua aceitação e seu entusiasmo, porque fizeram dos primeiros meses de adaptação mais amenos, porque o grupo conformado tanto pelo professor como pelos estudantes me deram um cordial recebimento.

Às pessoas que sempre estiveram prestes a brindar sua ajuda e estenderam uma mão amiga, meu agradecimento para: *Caroline Farias*, *William Guterres*, *Giovanna Angelis A. de Araújo*.

A todos os colegas, que amavelmente interagiam, colaboravam, e compartilhavam na sala de aulas.

Ao CNPq pelo apoio econômico que permitiu minha sustentabilidade e permanência no Brasil.

E a todas e cada uma das pessoas que colaboraram de forma direta ou indireta com o desenvolvimento deste trabalho.

Muito Obrigada.

Resumo da Dissertação apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica.

IMPLEMENTAÇÃO DE UM SISTEMA ESTEGANOGRÁFICO PARA INSERÇÃO DE TEXTOS EM SINAIS DE ÁUDIO

Jinnett Pamela Carrión Casierra

Março/2009

Orientador: Hélio Magalhães de Oliveira, Docteur.

Co-Orientador: Ricardo Menezes Campello de Souza, PhD.

Área de Concentração: Comunicações.

Palavras-chave: Ocultação de Informação, Esteganografia, Dados, Sinais de Áudio, *Wavelet*.

Número de Páginas: 150.

RESUMO: A arte de ocultar uma mensagem dentro de outro objeto é conhecida como Esteganografia. Detalham-se técnicas convencionais para ocultação de mensagens e propõe-se uma nova abordagem. Este novo método de esteganografia em dois passos combina a cifragem do texto-pleno através de um criptossistema padrão, seguido pela imersão dos dados cifrados no arquivo de áudio. O trabalho enfoca-se na inserção de textos curtos em arquivos com formato wav – a entrada dos dados é realizada nas componentes que resultam da transformação do sinal mediante as transformadas de wavelet. O objetivo é introduzir dados de forma quase transparente, de tal maneira que a detecção por terceiros seja pouco provável, como também para garantir a recuperação praticamente inalterável dos dados. O áudio é decomposto em doze níveis mediante a escolha de uma wavelet-mãe, os dados são codificados e ocultados nos diferentes níveis segundo o critério do usuário. Para um melhor espalhamento dos dados em cada nível são utilizadas senhas alfanuméricas de tamanho proporcional à quantidade de caracteres ingressados em cada um dos níveis. A implementação computacional foi realizada no Matlab® e simulações com arquivos de áudio de diferentes tamanhos foram realizadas. Mudanças nos arquivos de áudio após a inserção dos dados foram medidas. Baseadas no esquema da Esteganografia, aplicações comerciais podem ser desenvolvidas para garantir a autenticidade dos arquivos, assim como a proteção de direitos autorais em arquivos digitais.

Abstract of Dissertation presented to UFPE as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering.

STEGANOGRAPHIC SYSTEM IMPLEMENTATION TO INSERT TEXTS IN AUDIO SIGNALS

Jinnett Pamela Carrión Casierra

March/2009

Supervisor(s): Hélio Magalhães de Oliveira, Docteur.

Ricardo Menezes Campello de Souza, PhD.

Area of Concentration: Communications.

Keywords: Information hiding, Steganography, Data, Audio signals, Wavelet.

Number of Pages: 150.

ABSTRACT: The art and science of communicating in a way that hides the existence of a plaintext is known as steganography. As technology progresses, new models of steganography have been proposed that exploits the features and weakness of particular digital formats. Standard techniques for data hiding are presented, besides the wavelet-based approach proposed in this thesis. This novel two-step steganography approach combines plaintext encryption by a standard cryptosystem followed by the embedding of the encrypted data in an audio file. The study is focused on inserting short texts into wav files– the data input is performed in the wavelet transform components of the audio. The goal is to perform an almost transparent data inserting in such a way that the detecting by potential eavesdroppers is very unlikely. This low-throughput stego-method was conceived not only to hide a message in an audio file, but also to guarantee an unchanged data retrieval. The audio signal is decomposed in twelve levels after a mother-wavelet choice; data are encoded and embedded at different decomposing levels according to the user choice. In order to guarantee a better data scattering in each level, alphanumeric passwords of length proportional to the number of inserted characters in each sub-block are used. This novel stego-tool was fully implemented using Matlab™. Computer simulations with several audio files with different lengths were performed, by inserting texts varying the number of characters, thereby obtaining the percent change rate of the sound file. Changes in audio files due to the text embedding can be assessed by some dedicated system or algorithm that is able to distinguish minor sound alterations. Many commercial steganography-based schemes have been developed to offer protection and security for digital multimedia, so as to testify the authenticity of files, or to protect copyrighted digital files.

SUMÁRIO

Lista de Figuras

Lista de Tabelas

Lista de Abreviações

Sigla	*Significado em Inglês*
A/D	Analog / Digital.
AIFF	Audio Interchange File Format.
ARPA Net	Advanced Research Projects Agency Network.
ASCII	American Standard Code for Information Interchange.
AU	Unix Audio.
CD-Rom	Compact Disc, read only memory.
CWT	Continuous Wavelet Transform.
dB	Decibel.
DCT	Discrete Cosine Transform.
DFT	Discrete Fourier Transform.
DVD	Digital Video Disc.
DWT	Discrete Wavelet Transform.
HAS	Human Auditory System.
Hz	Hertz.
IDEA	International Data Encryption Algorithm.
IDWT	Inverse Discrete Wavelet Transform.
ISO	International Standard Organization.
ITU	International Telecommunication Union.
kHz	Kilohertz.
LP	Long Play.
LSB	Least Significant Bit.
Midi	Musical Instrument Digital Interface.
MP3	MPEG Audio Layer 3.
PCM	Pulse Code Modulation.
SAFER	Secure And Fast Encryption Routine.
STFT	Short Time Fourier Transform.
TDES	Triple Data Encryption Standard.
TIC	Technology Information and Communication.
WAV	Windows Audio Visual.
WMA	Windows Media Audio.

Capítulo 1

INTRODUÇÃO

1.1. Introdução

O desejo do homem de deixar divulgado e armazenado seus conhecimentos, pensamentos, descobrimentos e experiências, o levam à busca de métodos que permitam comunicar-se com o mundo. Observam-se os fatos da humanidade e pode-se ver o ímpeto do homem por deixar um vestígio, transcender limites, eternizando assim suas idéias. Por exemplo, os túmulos nos quais se registram hieróglifos, papiros que contém informação, papel com manuscritos, fitas magnéticas, CD-ROM, DVDs, computadores, entre outros, estes são alguns dos meios que o homem empregou para registrar a informação.

A esteganografia permite atingir uma comunicação mediante o anonimato – ela se encarrega de ocultar o envio e a presença de alguma informação. Métodos e técnicas desenvolvidas disfarçam mensagens em arquivos digitais, na rede, ou em diversos objetos, com a finalidade de transmitir informações sigilosas. O procedimento da ocultação pode variar desde técnicas mais rudimentares e manuais até técnicas digitais bastante sofisticadas.

Existem diversos motivos pelos quais pessoas (ou entidades) poderiam desejar manter uma comunicação em segredo ou mesmo enviar informação oculta, seja simplesmente por manter uma conversação privada, para que ninguém tenha conhecimento do conteúdo da informação que está sendo transmitida, como também ocultar o teor de uma mensagem sigilosa, ou ainda, para que ninguém suspeite de que alguma comunicação está ocorrendo, tudo isso é factível mediante a Esteganografia.

Conforme a tecnologia avança, existem maiores recursos à disposição, nas diferentes áreas e nos diferentes meios. Uma que cresceu rapidamente e se tornou conhecida mundialmente por suas vantagens é a Internet, disponibilizando múltiplos serviços. Hoje, com maior facilidade pode-se realizar transações bancárias, e-comércio, verificação de agendas culturais, publicações de livros e jornais, envio de e-mails, assistir vídeos e televisão, escutar música, descarregar informação e outros múltiplos serviços que nos brinda o advento da Internet; e são cada vez mais as pessoas envolvidas no mundo digital, que possuem acesso à Internet e outros meios.

A esteganografia se popularizou pelas várias aplicações possíveis em torno de seu princípio básico que é o de ocultar informação sem haver percepção. Com a evolução da Internet, uma enorme quantidade de informação disponível circula pela rede e a capacidade de copiar e movimentar dados é crescente; com isso, surge um maior interesse e necessidade de utilizar a esteganografia como meio de segurança, inserindo dados que permitem identificar o autor do arquivo, códigos ou número de série, etc., os quais usam alguns poucos bits de informação, mas que são de grande utilidade no rastreamento de arquivos pela rede, permitindo um maior controle sobre a sua utilização, assim como uma averiguação da existência de cópias não autorizadas.

Possuir o domínio total sobre o tipo de informação enviada através da rede e controlar o uso impróprio da esteganografia é uma tarefa quase impossível, assim como também é inevitável o envio de informação maliciosa inserida em arquivos "inofensivos", tais como uma canção, um desenho ou um simples texto. Presume-se que na mídia existam inumeráveis arquivos aparentemente inofensivos como áudio, imagens, textos, etc., que contém dados inseridos.

1.2. Objetivos

Os objetivos da dissertação especificam-se a seguir:

1.2.1. Objetivo Geral

Conceber e implementar em aplicativo um algoritmo que insira mensagens de texto curtos utilizando arquivos de áudio com formato wav para a ocultação da mensagem sem ocasionar modificações auditivas severas no arquivo hóspede; podendo o algoritmo ser empregado como uma ferramenta para a realização de diversas experiências.

1.2.2. Objetivos Específicos

1. Abordar o desenvolvimento da esteganografia.

2. Conceber um novo sistema de esteganografia com base em decomposições de *wavelet* do sinal de áudio.

3. Utilizar e aproveitar o melhor desempenho da ferramenta Matlab® para o desenvolvimento do programa de esteganografia com base no novo sistema (item 2).

4. Realizar a aplicação do programa desenvolvido com áudios predefinidos e gravação de voz (simulações, validação etc.).

5. Fazer testes de variação percentual do áudio ao inserir dados mediante a técnica desenvolvida.

6. Medir a qualidade de um conjunto de áudios mediante um teste auditivo (subjetivo), realizada num grupo de pessoas.

Capítulo 2

ESTEGANOGRAFIA

2.1. Introdução à Informação Oculta

Antes de iniciar a abordagem no contexto da esteganografia é necessário ter conhecimento de alguns termos que serão empregados neste e nos capítulos a seguir:

- **Arquivo encobridor.** Também pode ser denominado arquivo hospedeiro, é o arquivo portador ou arquivo base que vai ocultar a informação ou mensagem.

- **Mensagem.** São os dados que vão ser inseridos dentro do arquivo hóspede, ou seja, é a informação que se deseja enviar de forma secreta.

- **Arquivo-estego.** É o arquivo encobridor contendo a mensagem.

- **Senha-estego.** São informações que controlam o processo tanto de ocultar como de recuperar a mensagem inserida, isto é, uma senha(s) para proteger a informação oculta.

- **Atacante.** Entidade que monitora a transmissão de dados na rede. Podem existir dois tipos de atacantes ou intrusos:

    - *Intruso Passivo*: Não realiza nenhum tipo de modificação nos arquivos.
    - *Intruso Ativo*: Realiza modificações nos arquivos.

- **Capacidade.** Quantidade de informação que pode ser inserida num arquivo.

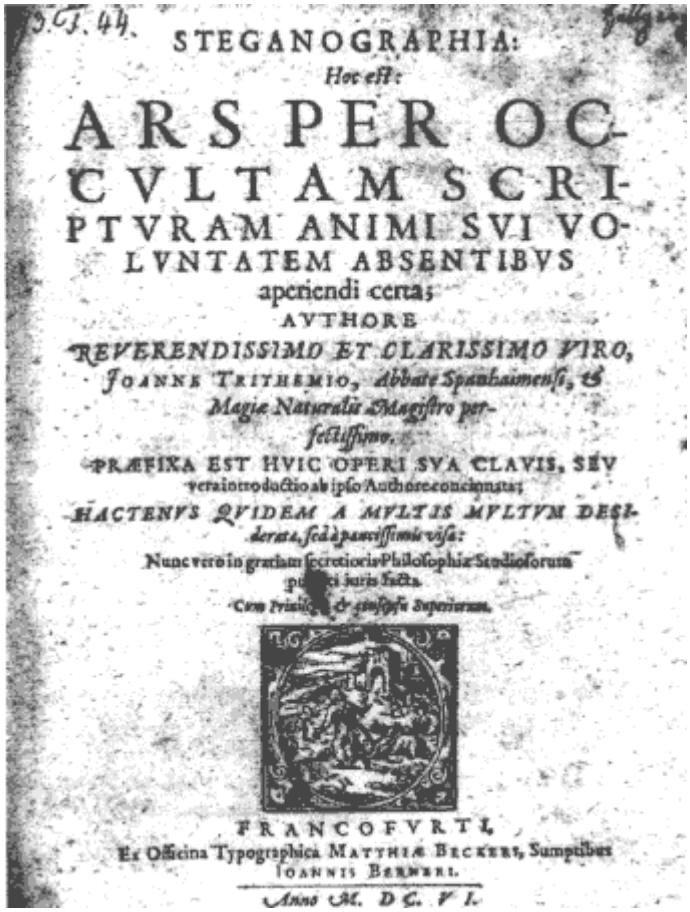
- **Robustez.** Quantidade de processamentos que um arquivo-estego suporta antes que a informação inserida se perca completamente.

- **Segurança.** Dificuldade para detectar dados ocultos num determinado objeto.

2.2. História da Esteganografia

A Esteganografia pode ser definida como a ocultação de uma mensagem dentro de outra mensagem [1]. O objetivo da esteganografia é de não levantar suspeita de que uma informação está sendo enviada; e foi Johannes von Heidelberg - Jean Trithemius (cf. Figura 1) - nascido na Alemanha em 1462, um dos primeiros a escrever sobre mensagens ocultas. Seu interesse pelos livros e manuscritos o leva a escrever várias obras, entre as quais "Esteganografia" (que forma parte de uma trilogia), ver (Figura 2) na qual especificava formas de escrever mensagens secretas sem provocar suspeita, e é por esse escrito que Jean Trithemius foi julgado de ser mágico e sua obra proibida sob a alegação de que utiliza os espíritos para enviar mensagens secretas. Anos depois de sua morte em 1516, um livro intitulado Esteganografia foi publicado em 1608 (se diz que não era o original). Uma só desvantagem possuía a técnica de Jean Trithemius, que tanto o receptor como o transmissor deveriam ter conhecimento do método esteganográfico utilizado para decifrar a mensagem. Sendo poucos os ensinamentos deixados por Trithemius, jamais se saberá exatamente em que consistia a técnica de enviar mensagens "a mais de cem léguas de distância" e "para o lugar mais profundo que se possa imaginar" [2].

Figura 1. Johannes von Heidelberg (Jean Trithemius).



STEGANOGRAPHIA:
Hoc eſt:
ARS PER OC-
CVLTAM SCRI-
PTVRAM ANIMI SVI VO-
LVNTATEM ABSENTIBVS
aperiendi certa;
AVTHORE
REVERENDISSIMO ET CLARISSIMO VIRO,
IOANNE TRITHEMIO, Abbate Spanhaimenſi, &
Magiæ Naturalis Magiſtro per-
fectiſſimo.
PRÆFIXA EST HVIC OPERI SVA CLAVIS, SEV
veratotiusductio ab ipſo Authore concinnata;
HACTENVS QVIDEM A MVLTIS MVLTVM DESI-
derata, ſed à paucissimis viſa:
Nunc vero in gratiam ſecretioris Philoſophiæ Studioſorum
publici juris facta.
Cum Priuilegio & gratioſa Superiorum.
FRANCOFVRTI,
Ex Officina Typographica MATTHIÆ BECKERI, Sumptibus
IOANNIS BERNERI.
Anno M. DC. VI.

Figura 2. Página do livro de Esteganografia de Jean Trithemius.

Quando se trata de ocultar informação existem várias estratégias, até mesmo a entonação utilizada ao pronunciar as palavras de uma frase pode transmitir uma mensagem. Mas a esteganografia não é uma matéria recente, já desde a antiguidade se realizavam processos para esconder informação.

A história envolve muitos acontecimentos que serão brevemente detalhados no único intuito de se ter noção histórica dos fatos. Existe um relato da antiga Grécia, aproximadamente do século V AC., mencionando como um escravo chegou a ser utilizado no envio de uma mensagem oculta: o seu cabelo foi raspado, a mensagem tatuada, quando o cabelo cresceu a mensagem ficou oculta, podendo atravessar pela fronteira inimiga sem levantar suspeita. Na China no século XI eram transmitidas mensagens escritas numa fina seda, o pedaço de lenço era enrolado e engolido pelo transportador. No século XV um método foi descoberto para esconder mensagens dentro de um ovo cozido, a mensagem era escrita na casca do ovo com uma solução especial, esta atravessava a casca e se impregnava no interior do ovo cozido

podendo ser lida a mensagem quando o ovo era descascado, vários destes tipos de métodos são ilustrados na Figura 3. Em época de guerra inumeráveis estratégias apareciam com o propósito de transmitir informação oculta, por exemplo, um sistema de "micro pontos" foi desenvolvido na Alemanha (cf. Figura 4), na época nazista, um aparelho fotográfico que permitia gerar negativo de fotos do tamanho de um alfinete, com a finalidade de preservar livros escritos por Judeus em forma de filme fotográfico, que pudessem ser transportados sem perigo. Na segunda guerra mundial, eram muito comum o uso de métodos como mensagens escritas com tinta especial que, ao reagir com outros agentes, tornava visível a mensagem. Pode ser encontrada informação inserida dentro de um texto, por exemplo, em acrósticos, outra tática é a ordem em que se encontram as letras ou palavras, pequenas perfurações feitas acima de letras num texto, a utilização de uma máscara que contem orifícios aleatórios e que colocada sobre algum texto formam uma mensagem. Em 1550 o matemático italiano, Girolamu Cardano (1501-1576) propõe um método para escrever mensagens ocultas empregando uma grade que leva o nome de *Cardan grille* ou *grelha de Cardano* (cf. Figura 5).

Figura 3. Técnicas Esteganográficas.

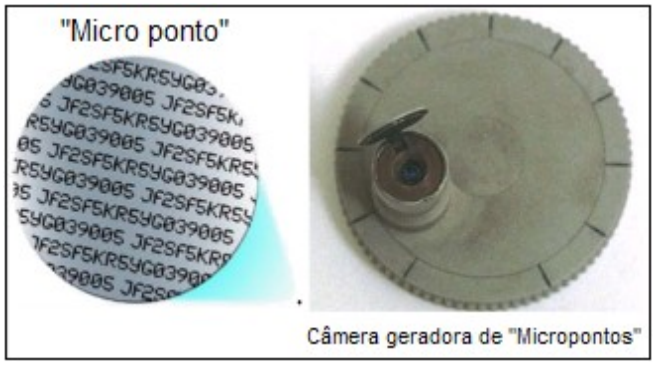


Figura 4. O "Microdot".

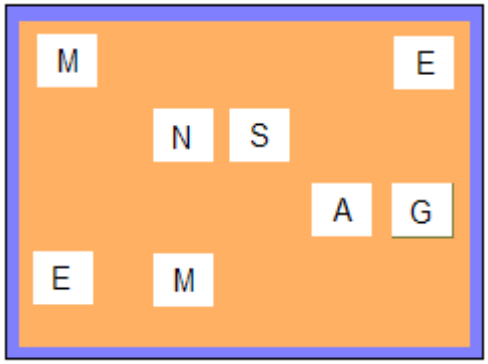
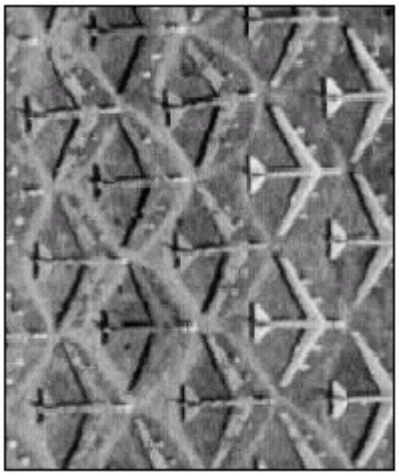


Figura 5. Exemplo da grelha de Cardano.

Ao final da década de 40, com a aparição dos discos de vinil ou LP (*Long Play*), surgiu outra expectativa; acreditava-se que algumas canções tinham no seu conteúdo mensagens subliminares, as quais poderiam ser escutadas ao colocar o disco no sentido contrário. Com essa perspectiva, na década de 50 surgiram questionamentos sobre as imagens, propagandas, programas televisivos e muitos outros meios de difusão que poderiam conter algum tipo de mensagem subliminar.

A esteganografia atravessou a barreira do tempo e hoje segue buscando técnicas para encobrir informação, utilizando-se de novas tecnologias. E através da informação digital (representada por bits), vislumbraram-se novas formas de inserir dados ocultos em arquivos digitais (cf. Figura 6, a). Mediante diversas técnicas, modificam-se poucos bits sem modificar a essência do arquivo encobridor. O sinal cobertor pode ser um texto, imagem, áudio, vídeo, etc., enquanto que a mensagem a inserir pode ser um texto, imagem ou som.

Há tanta "controvérsia" sobre o envio de mensagens ocultas que até mesmo filmes foram produzidos, os quais abordam o tema das mensagens ocultas: poder-se-ia citar como exemplo o filme o "Código Da Vinci" (cf. Figura 6, b).



(a)

(b)

Figura 6. Imagens que ocultam informação. a) Imagem disfarçada. b) Mona Lisa.

2.3. O Problema do Prisioneiro

O modelo clássico de esteganografia foi abordado mediante o "Problema do prisioneiro" proposto por Simmons [3]:

Duas entidades, Alice e Beto, encontram-se presos, cada um numa cela diferente; eles precisam comunicar-se para armar um plano e poder escapar, mas todas as comunicações entre ambos passam através de uma guarda, como mostra a Figura 7. Se a guarda suspeitar que eles estejam tramando algo, cancelará a troca das mensagens e os colocará em isolamento. Então, eles deverão buscar uma forma de comunicar-se sem levantar nenhum tipo de suspeita.

Este problema pode ser abordado mediante a Esteganografia, onde se efetua a troca de mensagens inócuas, podendo esconder-se informações sem serem descobertos.

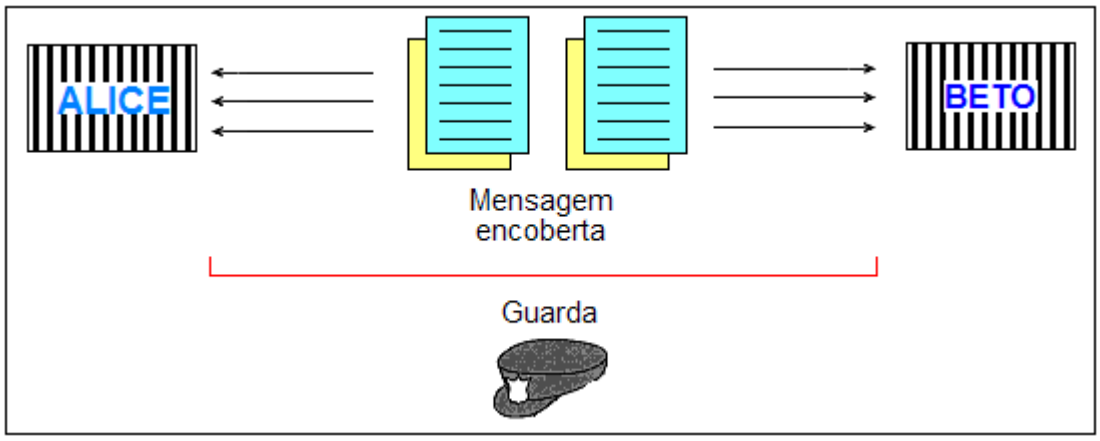


Figura 7. O problema do prisioneiro.

Já que a guarda monitora todas as informações, deve-se ter em consideração que ela como um atacante ativo, pode manipular os dados enviados; atuando como um atacante passivo, ela simplesmente se limita a olhar a troca de mensagens.

2.4. Esteganografia

Esteganografia é a arte e a ciência da comunicação em segredo. Esteganografia provém do grego *Steganos* que significa "segredo" e *Graphy* "escrita", ou seja, é a escrita em segredo [4-5], a esteganografia oculta a correspondência entre duas entidades. Para realizar a inserção de dados necessita-se um sinal cobertor, uma mensagem, que pode ou não estar cifrada, e a senha ou grupo de informações que dão maior segurança aos dados inseridos, tal informação é empregada no momento da recuperação da mensagem; todos estes requerimentos originam um arquivo com uma mensagem oculta que pode ser chamado de arquivo-estego como se ilustra na Figura 8.

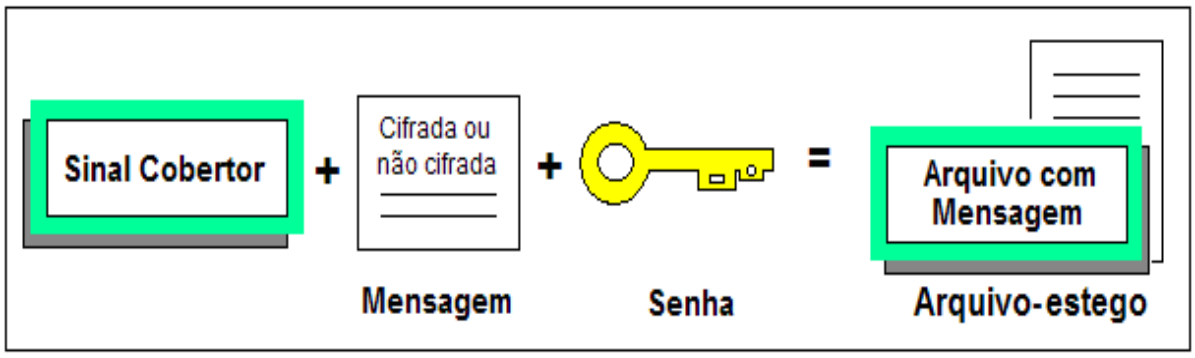


Figura 8. Geração de um arquivo-estego.

A idealização de um sistema esteganográfico seria o de inserir mensagens secretas com tanta cautela que os únicos entes em ter conhecimento e conseguir recuperar a informação oculta seriam o transmissor e o receptor. O sistema não poderá realizar modificações perceptíveis no arquivo encobridor, ou seja, deve existir um balanço entre:

- A escolha da técnica a empregar para a ocultação da mensagem.
- A porção de dados a ser inseridos.
- A escolha do sinal cobertor,

i.e., uma técnica mal empregada pode não ser robusta o suficiente para evitar perda dos dados inseridos ou originar modificações abruptas no arquivo; o tamanho da mensagem é diretamente proporcional às alterações que possa sofrer o sinal e por outro lado, arquivos muito grandes circulando pela rede podem levantar algum tipo de suspeita.

Softwares que permitem a ocultação e detecção de mensagens podem ser encontrados na mídia, alguns deles estão detalhados na Tabela 1 a seguir:

Tabela 1: *Softwares* de ocultação e detecção disponíveis como ferramenta na mídia.

Aplicativos disponíveis como Ferramenta
EZStego [6-8]
Hide and Seek [6, 8]
Hide4PGP [8]
Jpeg-Jsteg [6, 9]
MP3Stego [8]
Mandelsteg [10]
Steganos [8, 11]
StegoDos [8, 12]
S-Tools [6, 8, 12]
SysCop [13]
Stegdetect [6]
WbStego [8]
White Noise Storm [8]

2.5. Tipos de Esteganografia

Teoricamente podem existir três tipos de Esteganografia: Esteganografia pura, Esteganografia de chave secreta e Esteganografia de chave pública. Uma breve descrição de cada uma delas se realiza a seguir:

- **Esteganografia pura**. Conhece-se como esteganografia pura, ao sistema que não precisa de informações secretas (senhas) para o intercâmbio de mensagens. Os dados inseridos são distribuídos dentro do objeto e a recuperação se realiza através de uma sondagem. Esta técnica só depende do princípio da esteganografia, que a ocultação da informação não seja detectável.

- **Esteganografia de chave secreta**. Precisa de um grupo de informações secretas ou senha(s) para poder realizar o intercâmbio de mensagens, e ninguém que não possua a senha-estego deve ser capaz de recuperar a mensagem. Este tipo de esteganografia deverá cumprir com o sigilo da informação oculta e também deve realizar o intercâmbio da senha de forma confidencial.

- **Esteganografia de chave pública**. É similar a criptografia de chave pública [3]. Emprega duas chaves: uma privada e outra pública. A chave pública pode ser usada no processo de ocultar a mensagem, enquanto à chave privada permite a recuperação dos dados. Devido a que a guarda tem acesso a todas as informações transmitidas, poderia manipular a senha pública, neste caso a certificação da chave pública é necessária [3].

2.6. Aplicações da Esteganografia

Existem diversas aplicações que ocultam informação dentro de arquivos, e dependendo de sua utilidade e a função que cumprem possuem diferentes características, algumas das mais conhecidas se especificam na Tabela 2:

Tabela 2. Aplicações mediante Esteganografia.

Aplicações práticas mediante Esteganografia
▪ Monitoramento de arquivos.
▪ Proteção de Direitos Autorais.
▪ Autenticação de arquivos.
▪ Marcas de água.
▪ Controle de copias não autorizada.
▪ Impressões Digitais.
▪ Legendas.
▪ Número de série.
▪ Comunicação secreta.
▪ Organização de arquivos (imagens médicas).
▪ Modificações de arquivos.

2.7. Métodos Empregados para Ocultar Informação

Cada vez são mais os métodos disponíveis para ocultar informação. Geralmente o que se procura para inserir dados são lugares pouco empregados dentro do arquivo cobertor ou bits redundantes, e que a inserção dos dados seja invisível diante dos sentidos, mas nem sempre é factível ocultar informação em lugares pouco utilizados porque poderiam se perder os dados inseridos ao momento de realizar algum tipo de processamento ou manipulação no sinal. Pode-se sugerir na escolha do arquivo encobridor fotos ou desenhos que possuam redundância, assim também, em arquivos de áudio se pode aproveitar a existência de ruído para ser substituído pela informação desejada, ou a utilização de frequências inaudíveis ao sistema auditivo humano. Mas se deve ter cautela para não comprometer o arquivo, já que pode ocorrer que demasiada informação inserida ou a técnica incorreta empregada para a inserção mudem o arquivo base, fazendo com que o sinal cobertor sofra uma variação perceptível chamando a atenção do usuário(s). Por isso é necessário tomar em conta aspectos como:

- O arquivo que se vai empregar como base.
- A técnica que se vai utilizar para a inserção de dados.
- O meio onde vai circular o arquivo com a informação oculta.
- Quantidade de dados a inserir.
- Segurança da informação.

- O arquivo empregado como base não poderá ser aproveitado mais de uma vez para ocultar dados.

Detalham-se algumas das técnicas empregadas que podem ser aplicadas como referência ao momento de realizar a ocultação de dados em arquivos:

- **Permutação de dados.** Troca de caracteres.

- **Bit menos Significativo (LSB).** Substituindo o bit menos significativo pela informação que se deseja inserir, esta é uma das técnicas mais empregadas para a ocultação de mensagens por ser uma das mais simples e que permite a inserção de uma quantidade significativa de dados, mas também muito vulnerável quando se realiza processamentos no sinal, a informação tende a perder-se.

- **Transformação do sinal cobertor.** Realiza-se a transformação do sinal mediante o uso de transformadas como: Transformada Discreta de Fourier (DFT), Transformada Discreta do Cosseno (DCT), Transformada *Wavelet*, etc.; uma vez realizada a transformação do sinal os coeficientes são manipulados para proceder com a inserção dos dados.

Estas técnicas são usualmente mencionadas quando se trata de ocultar uma mensagem, seja porque algumas permitem fácil manipulação dos dados e simplificam a ocultação da mensagem (a informação é mais bem distribuída), seja porque dão maior robustez aos dados inseridos. Alguma delas são técnicas precursoras porque foram muito exploradas até nossa época. Mas também se deve levar em conta contramedidas que possuem cada uma das técnicas como se mostra na Tabela 3:

Tabela 3: Técnicas usuais empregadas para a ocultação de dados e as correspondentes contramedidas.

TÉCNICA	CONTRAMEDIDAS
▪ Permutação de dados.	Apesar de que não modifica o tamanho do arquivo é fácil de detectar.
▪ Bit menos significativo (LSB).	Fácil de detectar e corromper a informação inserida.
▪ Transformação do sinal cobertor.	Pode ocasionar mudanças perceptíveis no arquivo encobridor.

2.8. Precauções que devem ser tomadas em Esteganografia

A inserção de uma mensagem dentro de um arquivo ocasiona modificações na característica do sinal, sendo a quantidade de informação considerada como um fator degradante do arquivo cobertor, então, para obter garantia de que não se poderá detectar a presença da mensagem, a quantidade de dados a ser inserida deverá ser mínima.

Uma das primeiras táticas a ser adotadas e que contribui notavelmente é a encriptação dos dados; a informação original ou texto claro pode ser codificado antes da inserção usando técnicas conhecidas (Triple DES, IDEA, SAFER+, etc.), tais técnicas levam tempo e esforço até serem decifradas, e o atacante agora vai ter dois desafios: detectar a mensagem e decifra-la.

Para aumentar a resistência é possível utilizar chaves ou informações como códigos secretos que podem ser denominados como senha-estego. Para lograr robustez no sistema, a redundância da informação inserida é factível, então um compromisso existe entre a robustez e a segurança, assim, entre mais quantidade de dados inseridos, a detecção aumenta e o arquivo esteganografado se torna menos seguro. Técnicas adicionais são úteis quando a mensagem foi descoberta, ou seja, quando o princípio da esteganografia, que é o de manter oculto o envio de informação, é destruído mediante algum processo de estego-análise ou ataque.

2.9. Estego-análise: Na procura de Informação Oculta

Estego-análise é a arte de descobrir mensagens ocultas. Quando se deseja procurar informação oculta, é valido salientar a utilização dos sentidos, já que podem existir dentro de arquivos informações escondidas, pelo qual, é importante prestar atenção a todas as características, sobretudo quando se trata de imagens. A inserção de informação em imagens é muito comum; a internet pode abarcar milhares de fotos ou desenhos que podem conter informação oculta. Imagens com algum tipo de distorção como modificação da cor ou perda da resolução podem conter informação além do que podemos observar. Algo similar acontece com arquivos de áudio, um pequeno ruído pode ter sido gerado por inserção adicional de dados. A percepção de algum ruído é subjetiva, já que depende do ouvinte e de fatores externos. Na análise de arquivos de texto se deve considerar a frequência com que aparecem as letras, palavras, digramas, trigramas, etc., que dependerá de cada idioma. Por exemplo, em Português (Brasil) a letra mais utilizada é a letra A e seguindo uma ordem descendente do uso das letras se tem a Figura 9 que indica tal sequência [14]. Outro fator importante que se deve tomar em conta é a entropia no sinal encobridor, que é a quantidade de desordem que pode ser encontrado no sinal ou a incerteza num texto.

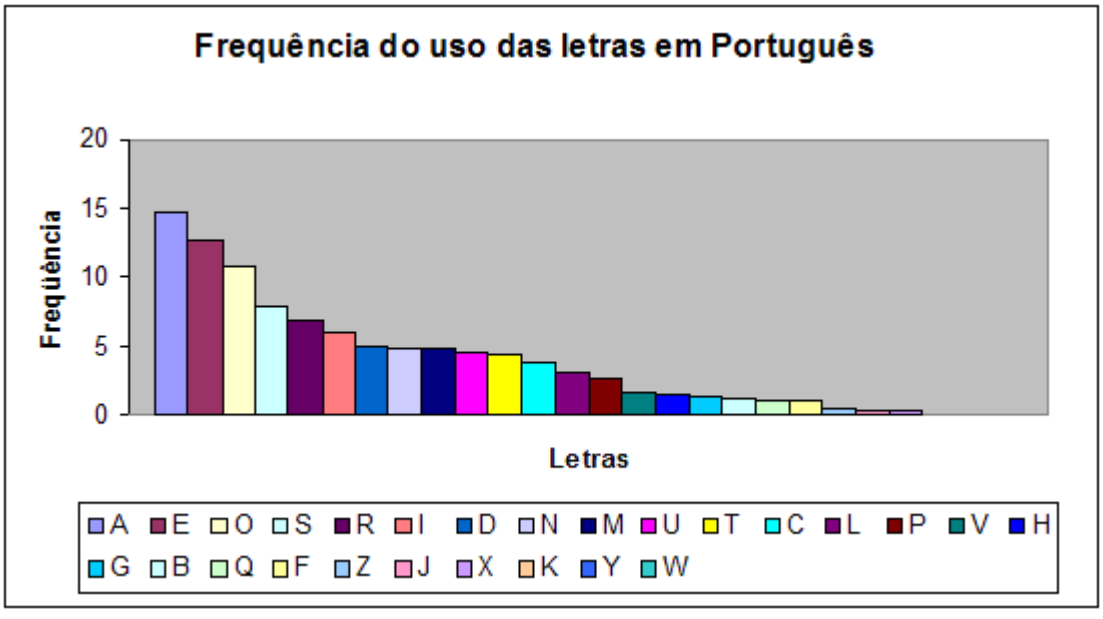


Figura 9. Histograma de frequência do uso das letras em Português.

A situação seria melhor caso fosse possível comparar o arquivo original com o arquivo que contem os dados inseridos, mas é muito difícil de encontrar disponibilizados ambos os

arquivos, já que o mensageiro unicamente libera o arquivo que contém a mensagem para evitar qualquer tipo de ataque.

Textos ou imagens aparentemente inofensivas podem conter informação oculta sem se quer ter percepção, tal como se mostra na Figura 10. As duas imagens visualmente iguais diferem no seu conteúdo; o gráfico (a) mostra à imagem original junto ao texto que vai ser inserido, enquanto a imagem (b) é a imagem esteganografada, a qual possui a mensagem de texto oculta. Para realizar a inserção dos dados na imagem, foi empregada a técnica de transformação do sinal mediante a transformada de *wavelet*, manipulando os coeficientes e introduzindo a mensagem.

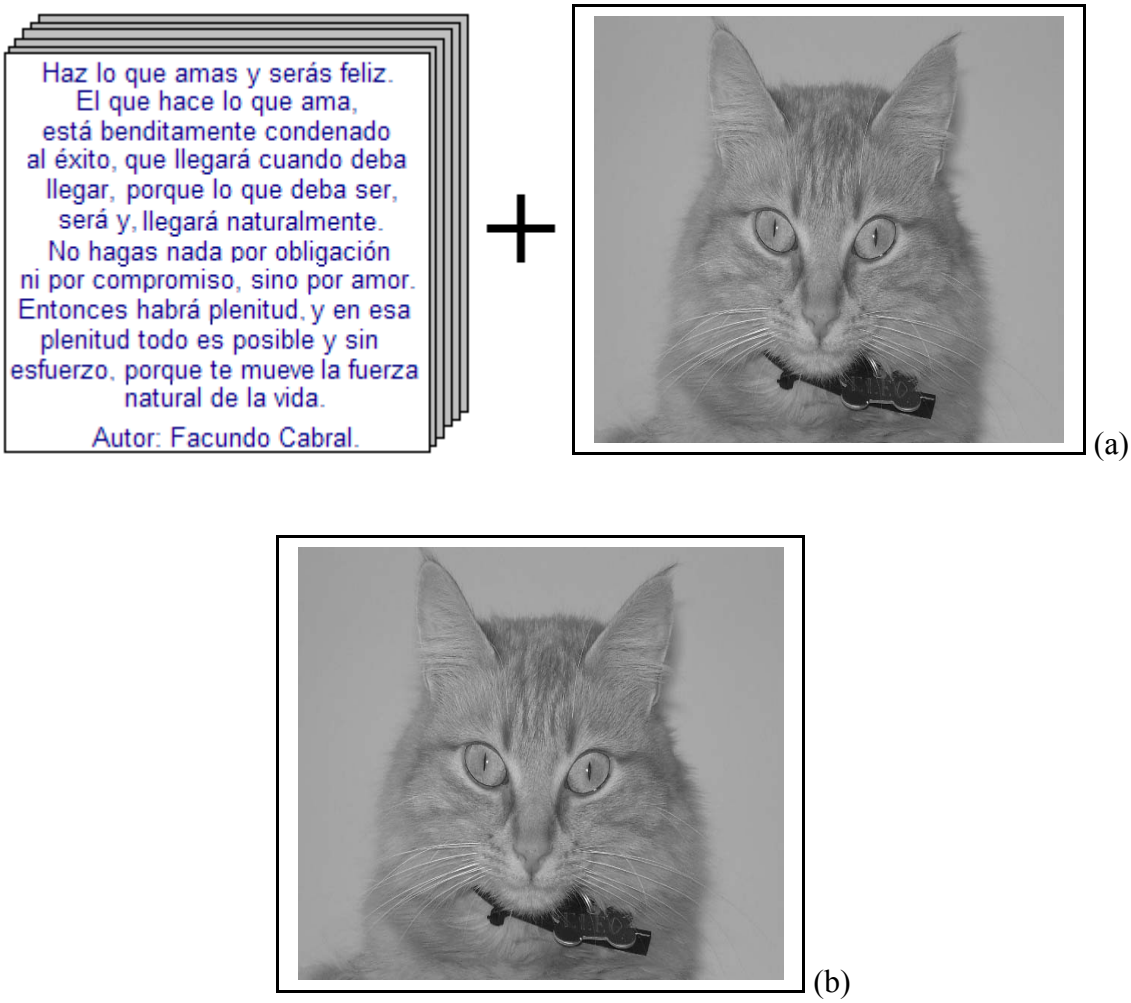


Figura 10. Aplicação de esteganografia numa imagem.
(a) Imagem sem dados e texto a ser oculto. (b) Imagem com dados.

Mediante estego-análise pode-se verificar a probabilidade da existência de informação oculta dentro de arquivos. Estego-análise é a arte de descobrir e tornar inútil uma mensagem. Um sistema ideal seria capaz de realizar os seguintes passos:

1. Detectar uma mensagem oculta dentro de um arquivo.
2. Extrair a mensagem oculta.
3. Decifrar a mensagem (conteúdo).

Mediante programas estatísticos pode-se verificar a presença ou ausência de mensagens ocultas, para as quais a probabilidade de detecção dependerá da precisão do programa e do tamanho da mensagem oculta; se diz que um sistema esteganográfico não é seguro quando um atacante é capaz de provar a existência de uma mensagem secreta [3].

Um fato interessante de mencionar que pode acontecer com testes de verificação são os *falso-negativos*, que são arquivos que alojam conteúdo no seu interior, mas não podem ser detectados pelo sistema, e os *falso-positivos,* que são arquivos que não possuem informação, mas são detectados como um estego-objeto. Destas duas possibilidades, os *falsos positivos* devem ser os mais evitados, já que o sistema não contaria com a precisão adequada para a detecção se existisse um grande número de *falso-positivos*.

Em situações nas quais se aplica estego-análise e não se consegue chegar a decifrar a mensagem, a destruição ou alteração da informação também é considerada válida. O fato de que se tenha conhecimento de que num certo arquivo existe uma mensagem oculta, se diz que a esteganografia não alcançou o seu propósito.

2.10. Outro ponto de vista para a Esteganografia

Em ocasiões a aplicação da esteganografia é mal utilizada por pessoas que infringem a lei, fazendo uso indevido desta ferramenta para transmitir mensagens violentas pela Internet sem qualquer restrição. Apesar de não se ter nenhum tipo de prova contundente da utilização da esteganografia como meio de transmissão de mensagens ocultas entre grupos terroristas, se presume que este meio pode estar sendo empregado para realizar comunicações entre membros desses grupos. Por exemplo, há uma inquietação maior sobre a catástrofe

acontecida o dia 11 de Setembro de 2001, sendo novamente julgada a esteganografia como provável meio de propagação dos planos e da informação - segundo uma análise realizada, existiu ineficiência na vigilância eletrônica [15].

2.11. Esteganografia *vs.* Criptografia

Apesar da Esteganografia e a Criptografia serem fortes aliadas em sistemas esteganográficos e ambas cumprem com o objetivo de "resguardar o segredo", existem características que as diferenciam. Enquanto a Esteganografia *oculta a presença de uma mensagem*, a Criptografia *torna a mensagem incompreensível*. Para abordar essa diferença, especificar-se-á a finalidade de cada uma delas mediante a Figura 11 e Figura 12 esboçadas a seguir:

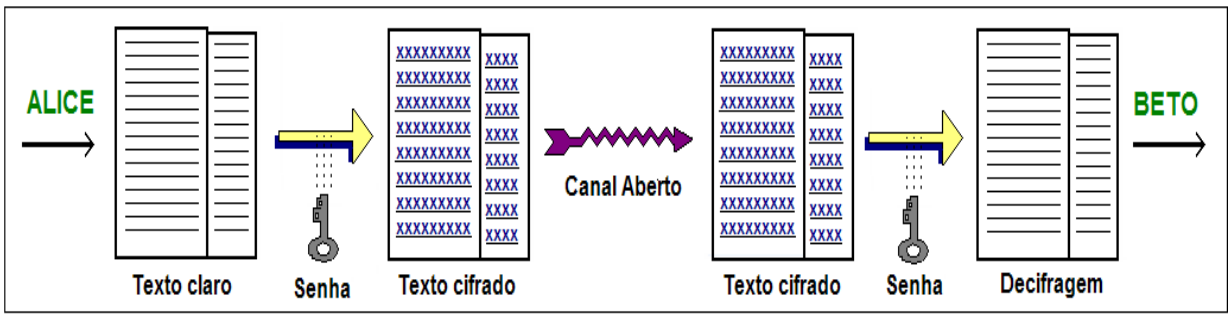


Figura 11. Processo de criptografar uma mensagem.

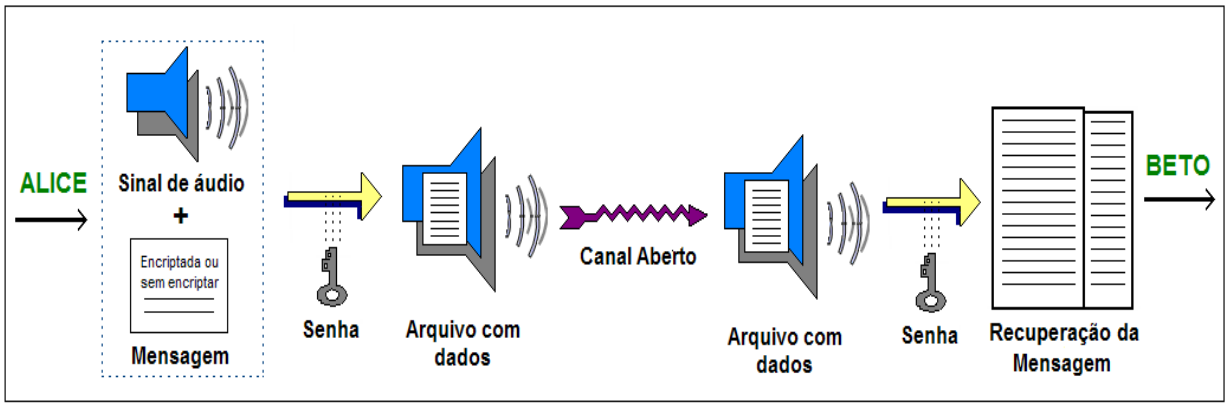


Figura 12. Processo de esteganografar uma mensagem.

Criptografia deriva-se das palavras gregas: *kryptós* que significa "secreta" e *gráphein* "escrita", ou seja, é a arte da escrita secreta [16]. De forma geral, a criptografia transforma um "texto claro" num "texto cifrado" ou incompreensível empregando uma chave(s),

mediante a utilização de um algoritmo (cf. Figura 11); de forma diferente, a esteganografia esconde dados num outro arquivo, fazendo com que a percepção desses dados seja indetectável (cf. Figura 12).

Semelhante à esteganografia, a criptografia também possui sua própria história e evolução, desde antigas práticas envolvendo substituição de letras por símbolos, alteração da ordem do alfabeto, códigos, etc., até os processos mais complexos, e mais seguros de hoje em dia. Ambas as ciências são muito antigas e com o progresso da tecnologia e o surgimento do computador, seu desenvolvimento tem avançado rapidamente.

Métodos que pretendem descobrir ou danificar a mensagem secreta têm sido criados, e em esteganografia, a ciência que se encarrega de detectar ou corromper uma mensagem oculta é denominada por estego-análise. Em criptografia, a ciência que se encarrega de decifrar informação e quebrar códigos é conhecida como criptoanálise. Quando o inimigo consegue detectar a presença de informação oculta dentro de um sinal (um arquivo "esteganografado"), se a mensagem se encontra encriptada, existirá um desafio adicional para o invasor, é por isso que a combinação das duas em um algoritmo esteganográfico aumenta o potencial do sistema.

- **Diferenças entre Esteganografia e Criptografia.**

Para especificar algumas diferenças entre estas duas ciências apresenta-se a Tabela 4:

Tabela 4: Diferenças entre Esteganografia e Criptografia.

ESTEGANOGRAFIA	CRIPTOGRAFIA
▪ Esconde a presença da mensagem.	▪ Altera o conteúdo da mensagem.
▪ A maior segurança é que a mensagem não pode ser percebida.	▪ A segurança se encontra na senha do algoritmo.
▪ Não existe a possibilidade de saber que se está realizando o envio de mensagens ocultas.	▪ Pode-se ter conhecimento que se está realizando o envio de uma mensagem cifrada.
▪ O inimigo não sabe que existe uma mensagem esteganografada.	▪ O inimigo sabe que existe uma mensagem cifrada.
▪ Uma pequena descoberta não é permitida (princípio da esteganografia).	▪ Uma pequena quantidade de informação pode ser descoberta ou decifrada.
▪ Como resultado de um algoritmo esteganográfico se obtêm um estego-objeto.	▪ Como resultado de um algoritmo criptográfico se obtêm um texto cifrado.
▪ O desafio esta em detectar se uma mensagem foi esteganografada em um arquivo.	▪ O desafio não esta em detectar se algo foi cifrado, e sim em decifrar a mensagem.

Capítulo 3

WAVELETS EMPREGADAS COMO FERRAMENTA

3.1. As Wavelets

A teoria das *wavelets* foi formulada nos meados dos anos 80 na França. A "definição teórica" do que se conhece com *wavelets* foi proposta por Jean Morlet e por um grupo de pesquisadores, incluindo Alex Grossman e Yves Meyer [17]. A palavra "*wavelet*" pode ser traduzida como "*Ondaleta"* e consiste numa função "curta" capaz de decompor outras funções.

A decomposição do sinal é conhecida como transformada de *wavelet* e se realiza no domínio da frequência de forma a poder analisar as funções em diferentes escalas de frequência e tempo, isso é uma grande vantagem já que na prática são empregados sinais que apresentam o espectro variante no tempo e não estacionários.

A análise mediante a transformada de *wavelet* apresenta grande flexibilidade, permitindo mudanças da base, ajustando-se desta forma ao tipo de sinal em estudo. As *wavelets* permitem analisar e revelar características dos dados que outras técnicas não conseguem descobrir, como por exemplo: ruído, detecção de descontinuidade [18-19], comportamentos de curto e longo prazo (processos de memória longa), etc.

3.2. Introdução a Teoria das Wavelets

Diversas áreas como matemática, física e engenharia, ambicionavam o surgimento de uma técnica que permitisse o estudo de sinais em diferentes resoluções e escalas; foi justamente em torno dessa visualização que surgiram as *wavelets*. Uma breve introdução da evolução das *wavelets* é realizada a seguir.

No que se tem referência, em 1909, Alfred Haar refere-se sobre as *wavelets* na sua tese de doutorado. Um ano depois constrói as *Wavelets de Haar* - que levam o seu nome, sendo estas à primeira família de *wavelets* de suporte compacto (a despeito de não usar este nome na época). No princípio dos anos 80, em 1982 Strömberg desenvolve as primeiras *wavelets* ortogonais, mas elas não foram muito mencionadas em aquela época [20]. Alex Grossman e Jean P. Morlet introduziram o termo *wavelet* em análise de sinais geofísicos [21], com a possibilidade de melhorar os resultados, já que a técnica de Fourier não oferecia muita

eficiência nesse tipo de estudo. Em 1987, Stéphane Mallat relaciona a teoria das *wavelets* com o processamento digital de sinais [18]. S. Mallat e Y. Meyer desenvolvem a teoria de análise em multirresolução, a qual disponibilizou uma ferramenta para a construção de outras bases [20, 22]. Yves Meyer construiu uma das primeiras *wavelets* não trivial. Ingrid Daubechies possibilita uma análise muito mais eficiente, construiu as *wavelet*s ortogonais de suporte compacto - sendo um dos conjuntos mais usado em aplicações.

Em 1995, surge uma nova geração de *wavelets* conhecida como: *wavelets de segunda geração*, as quais empregam novas técnicas de implementação. No mesmo ano Sweldens introduz o algoritmo de *lifting*, uma construção diferente para as *wavelets*, independente da transformada de Fourier [23].

O desenvolvimento, criação e utilização das *wavelets* no estudo de sinais vêm crescendo rapidamente sendo empregadas atualmente em análise e processamento de imagens, detecção de descontinuidade e pontos de quebra, supressão de ruído e compactação, síntese e processamento de sinais de voz, análise de multirresolução para visão artificial em computadores, compactação de sinais, localização das áreas de maior concentração de energia, edições de curva, manipulação de superfícies, análise de textura, compressão de imagens, representação de curvas, análise de superfícies, representação de fluxo de luz, modelamento geométrico, reconhecimento e extração de padrões, interpretação de imagens e sinais biomédicos, análise de voz, representações auditivas, em sinais sísmicos, acústica, música, geofísica, etc. [19, 24-26].

3.3. Características das Wavelets

Algumas das características mais importantes das *wavelets* são:

1. A área total sobre a curva da função é nula:

$$\int_{-\infty}^{\infty} \psi(t).dt = 0 \tag{1}$$

2. A energia da função *wavelet* é finita:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \tag{2}$$

Isto é, a *wavelet* é um sinal com formato de onda que tem duração limitada e valor médio nulo no domínio do tempo [27]. Para tornar possível o emprego das *wavelets* em aplicações de processamento de sinais devem ser observados certos requisitos:

- Devem possuir energia finita.
- Anular-se no infinito.
- Possuir certo número de momentos nulos.
- Apresentar suporte compacto, no tempo e frequência.
- Possuir certo grau de regularidade – suavidade.

Além das características e requerimentos mencionados precisa-se da condição de admissibilidade que garante a existência da transformada inversa.

3.4. Vantagens das Wavelets

Comparada com outras técnicas de processamento, as *wavelets* apresentam várias vantagens que vêm popularizando sua utilização no processamento e estudo de sinais; seu aparecimento permite abordar o problema no plano tempo-frequência. Algumas vantagens das *wavelets* com respeito a outras técnicas como é o caso da transformada de Fourier, são apresentadas na Tabela 5:

Tabela 5: Vantagens das *wavelets* com relação a outras técnicas de processamento de sinais.

WAVELETS	OUTRAS TÉCNICAS
▪ As bases das *wavelets* são variáveis.	▪ A base permanece constante em todo o estudo do sinal.
▪ Resolução frequêncial e temporal variável.	▪ Resolução frequêncial e temporal constante.
▪ Permite analisar descontinuidades em sinais não periódicos.	▪ Geralmente a análise é realizada em sinais periódicos e estacionários.
▪ Emprega menor número de coeficientes para representar um sinal.	▪ Precisa maior número de coeficientes para a representação precisa do sinal.
▪ Baixo custo computacional.	▪ Alto custo computacional.

3.5. Decomposição de Sinais Mediante as Wavelets

Quando se realiza a decomposição de sinais mediante as *wavelets*, obtém-se como resultado *Aproximações* e *Detalhes*, como ilustrado na Figura 13.

- *Aproximações*: Maior escala; *wavelet* dilatada; mudanças suaves; baixa frequência.

- *Detalhes*: Baixa escala; *wavelet* comprimida; mudanças rápidas; alta frequência.

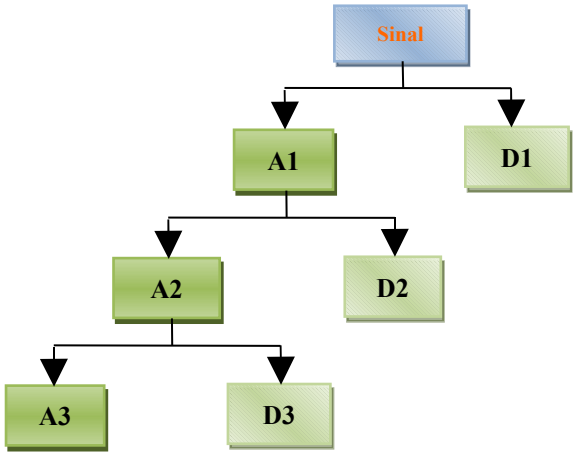


Figura 13. Decomposição de um sinal em Aproximações e Detalhes.

em que:

A = Aproximação

D = Detalhe

O processo é repetido em cada nível, a aproximação é subdividida numa nova aproximação e detalhe, então, para *n* níveis de decomposição se têm *n*+1 formas de decompor o sinal. Se um sinal é decomposto em três níveis, conforme apresentado na Figura 13, o sinal poderá ser reconstruído da aproximação (A3) e os detalhes (D3, D2, D1), de acordo com a seguinte equação [17]:

$$\text{Sinal} = A3 + D3 + D2 + D1 \qquad (3)$$

Por ser o método iterativo, poder-se-ia continuá-lo indefinidamente, porém, na realidade é adequado decompor o sinal até que o detalhe seja constituído de apenas uma amostra ou pixel [17]. Na prática de acordo com a aplicação que se estiver realizando e a natureza do sinal, será selecionado o número de níveis apropriado para a decomposição.

Para realizar a decomposição do sinal, emprega-se uma ***wavelet*-mãe** ou *wavelet*-padrão - todas as outras são versões escalonadas dela, a qual se comporta como uma "função base" para descrever outras funções. A *wavelet*-mãe pode apresentar-se: expandida, de baixa frequência ou comprimida, de alta frequência de acordo à análise a efetuar-se, seja esta frequêncial ou temporal.

A Figura 14 mostra as *wavelets* expandidas, comprimidas e deslocadas.

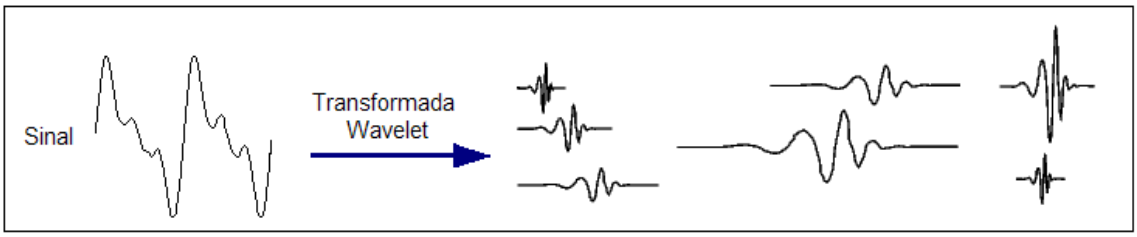


Figura 14. Componente *wavelet* de diferente escala e posição.

A reconstrução do sinal é factível e é conhecido como *Transformada Wavelet Discreta Inversa* [17].

3.6. O Comportamento das Wavelets

As *wavelets* possuem uma estratégia de janelamento variável, isto é, utiliza funções de base curta em altas frequências e funções de base longa em baixas frequências, ver Figura 15.

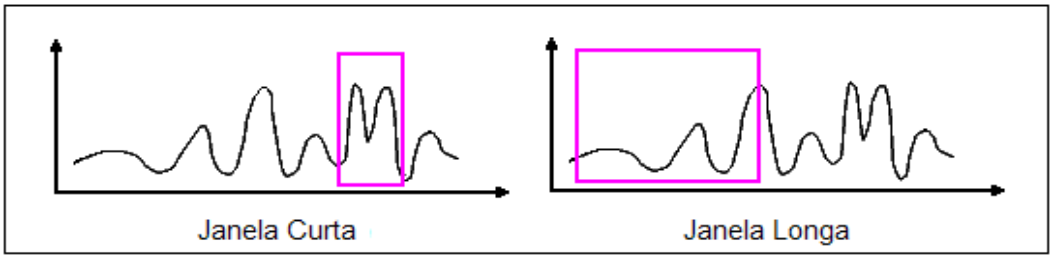


Figura 15. Janelamento variável das *wavelets*.

Existem dois fatores importantes a considerar:

- Escalonamento.
- Deslocamento.

3.6.1. Escala

A escala depende da frequência, e permite dilatar ou comprimir um sinal, (cf. Figura 16). Uma mudança de escala permite:

- *Escala maior:* Visão mais global (baixa frequência).

- *Escala menor*: Detalhes (alta frequência).

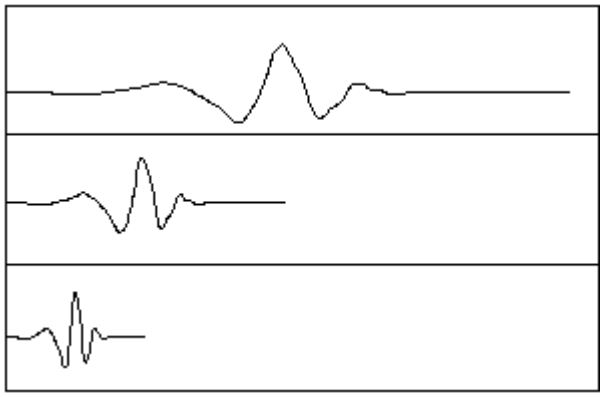


Figura 16. Fator escala nas *wavelets*.

A escala pode ser comparada analogamente com um mapa geográfico, assim, para uma visão geral da paisagem é adotado um fator de escala maior (frequências baixas), pelo contrario, se o objetivo é ter uma visão específica da paisagem é empregada uma escala menor (frequências altas).

3.6.2. Deslocamento no Tempo das Wavelets

É a translação no eixo do tempo. Matematicamente atrasar uma função *f(t)* por *k* é:

$$f(t-k) \tag{4}$$

A Figura 17 (a) mostra a função *wavelet* e a Figura 17 (b) mostra a *wavelet* deslocada.

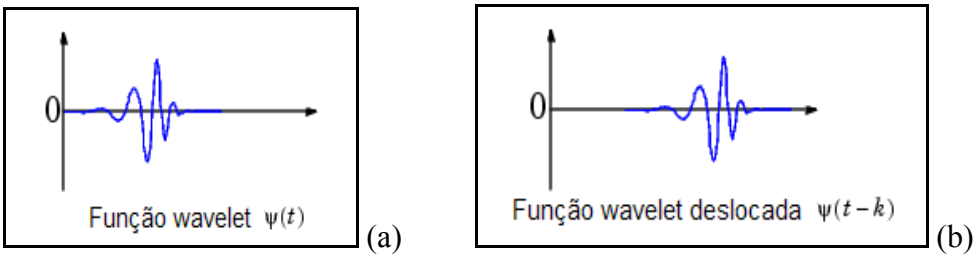


Figura 17. Fator deslocamento da *wavelet*. (a) Função *wavelet*. (b) Função *wavelet* deslocada.

3.7. Análise via Wavelets

A análise por *wavelets* inclui [17]:

- A transformada de *wavelet* contínua.
- A transformada de *wavelet* discreta.
- A transformada de *wavelet* de corpo finito.

3.7.1. Transformada de Wavelet Contínua: CWT

O desenvolvimento da Transformada de *wavelet* constitui uma evolução à restrição que apresenta a Transformada de Fourier de tempo curto (STFT), que emprega tanto na resolução no tempo como na frequência uma janela fixa, permanecendo constante em todo o intervalo, tal como mostra a Figura 18.

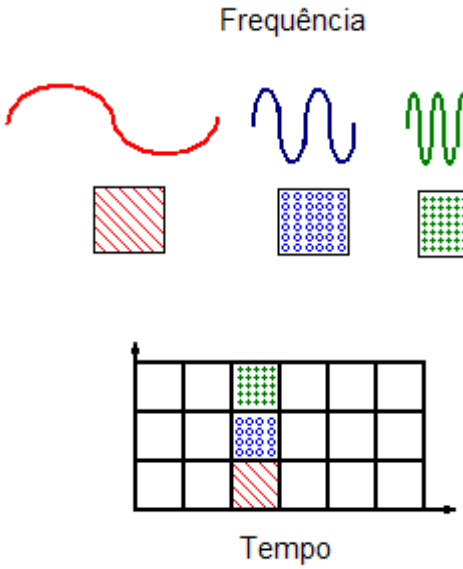


Figura 18. Janelamento da Transformada de Fourier em tempo curto.

Pelo contrário, a transformada de *wavelet* analisa o sinal em diferentes frequências com resoluções diferentes em cada uma delas (cf. Figura 19).

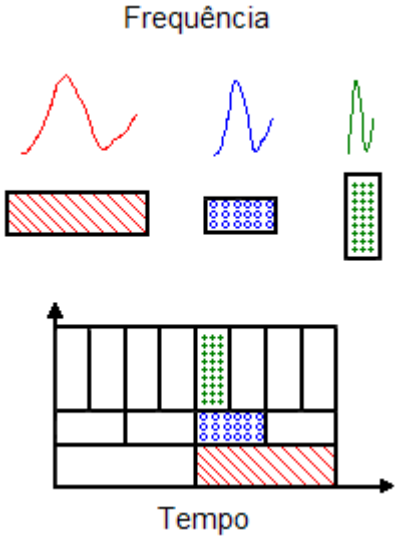


Figura 19. Janelamento da Transformada de *Wavelet*.

Para frequências altas, possui alta resolução no tempo e baixa resolução na frequência; para frequências baixas, tem alta resolução na frequência e baixa resolução no tempo.

A transformada de *wavelet* contínua é a soma em todo instante de tempo do sinal *f(t)* multiplicado pelas versões escalonadas e transladadas da função *wavelet*-mãe $\psi(t)$.

$$C(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{ab}(t)dt \tag{5}$$

em que:

a = escala

b = posição

Como resultado, obtém-se os coeficientes de *wavelet C*, que são função da escala e posição.

Diversas operações podem ser realizadas com a *wavelet*-mãe:

- Escalonamento: $\psi_a(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t}{a}\right),\ a \neq 0$ (6)
- Deslocamento: $\psi_b(t) = \psi(t-b)$ (7)
- Deslocamento com Escalonamento: $\psi_{a,b}(t) = \psi_a(t-b) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right)$, (8)

$$\{\psi(t)\} \rightarrow \{\psi_{a,b}(t)\} \quad (\forall a, a \neq 0)(\forall b \in \Re)$$

3.7.2. Transformada de Wavelet Discreta: DWT

A transformada discreta de *wavelet* aqui faz referência ao algoritmo do Matlab® [17] - sem ser o único disponível. A escala e posição baseiam-se em intervalos discretos geralmente em potência de 2. O sinal é decomposto mediante dois filtros, um filtro passa - baixa no qual são obtidas as *aproximações* e um filtro passa - alta no qual se obtém os *detalhes,* (cf. Figura 20).

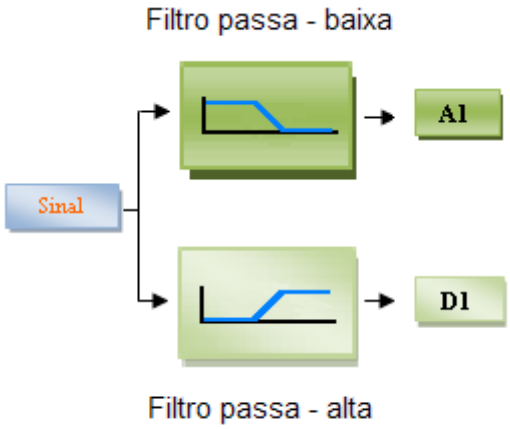


Figura 20. Decomposição do sinal mediante filtros.

A reconstrução do sinal original é possível mediante as componentes de aproximações (cA) e componentes de detalhes (cD), (cf. Figura 21).

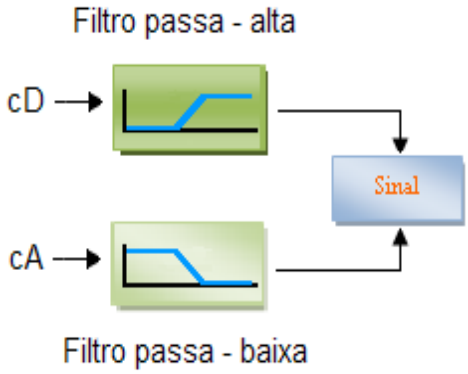


Figura 21. Reconstrução de um sinal mediante filtros.

Geralmente quando se realiza o estudo da forma discreta, aparecem parâmetros variáveis do sinal: escala (a) e deslocamento (b). O valor de a (escala) toma valores inteiros positivos ou negativos que resulta de um valor fixo a_0 elevado a uma potência:

$$a = a_0^{-j}, \tag{9}$$

em que:

$$a_0 > 1 \quad e \quad j \in Z .$$

Normalmente a_0 e b_0 assumem valores inteiros. O menor passo inteiro para a escala é $a_0 = 2$; tal fator é conhecido como escalonamento diádico [27].

O valor de *b* (deslocamento) dependerá de *j*, de forma que *wavelets* curtas (alta frequência) sejam deslocadas em passos pequenos, e *wavelets* largas (baixa frequência), sejam deslocadas em passos maiores:

$$b = k b_0 a_0^{-j} , \tag{10}$$

em que :

$b_0 > 0$ *e* $j,k \in Z$.

Então, a *wavelet* discreta será expressa por:

$$\psi_{j,k}(t) = a_0^{j/2} \psi (a_0^j t - k b_0) \tag{11}$$

Substituindo o valor de $a_0 = 2$ e fazendo $b_0 = 1$, a equação 11 se torna a *wavelet* diádica como se mostra a seguir:

$$\psi_{j,k}(t) = 2^{j/2} \psi (2^j t - k) \tag{12}$$

A transformada de *wavelet* discreta possui sua correspondente inversa expressa da seguinte maneira:

- *Transformada Wavelet Discreta Inversa (IDWT):*

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_{j,k}(t) d_{j,k} , \tag{13}$$

em que:

$d_{j,k}$ são os coeficientes *wavelets* (representação homogênea).

3.7.3. Transformada de Wavelet de Corpo Finito: TWCF

As transformadas discretas clássicas como a Transformada Discreta de Fourier - DFT, a Transformada Discreta do Cosseno – DCT e a Transformada Discreta de *Wavelet* – DWT, podem ser consideradas como "Transformadas analógicas", algo análogo, por exemplo, ao sistema de Pulsos Modulados em Amplitude (PAM). As transformadas definidas sobre

corpos finitos são discretas e têm coeficientes que assumem valores num alfabeto finito, de modo que podem ser interpretadas como "Transformadas Digitais".

A definição das *wavelets* sobre corpos finitos será especificada a continuação. A *wavelet*-mãe, será definida como um vetor de comprimento *N*:

$$\underline{\psi}_{1,0} = (\psi_{1,0}(0), \psi_{1,0}(1), \psi_{1,0}(2), ..., \psi_{1,0}(N-1)) \tag{14}$$

em que cada componente de *y* pertence ao corpo de extensão GF(p^s) onde *s* é um inteiro, $s \geq 1$.

As *wavelets* sobre corpos finitos primos serão abordadas a seguir. Seja *N* um inteiro e D(*N*) o conjunto dos divisores de *N*. O escalonamento sobre corpos finitos não pode ser tomado como um número real, como usualmente se faz, mas sim como um inteiro divisor do comprimento (*N*). As seguintes operações podem ser realizadas:

- Escalomento: $(\underline{\psi}_{j,0}) : \psi_{j,0}(i) = \psi_{1,0}(ji), \quad \forall j \in D(N/2) := \{j \text{ tal que } j \mid N/2\}$ (15)
- Translações: $(\underline{\psi}_{j,k}) : \psi_{j,k}(i) = \psi_{j,0}(i + \frac{Nk(\text{mod } N)}{j}), \quad \forall k = 0,1,...N-1$ (16)

A seguinte expressão da função *wavelet* mostra a versão escalonada e transladada da *wavelet*-mãe $\underline{\psi}_{1,0}$:

$$\underline{\psi}_{j,k} = (\psi_{j,k}(0), \psi_{j,k}(1), \psi_{j,k}(2), ..., \psi_{j,k}(N-1)) \tag{17}$$

Propriedade 1: $\sum_{i=0}^{N-1} \psi_{j,k}(i) \equiv 0(\text{mod } p), \quad \forall j,k$ (18)

Definição: Seja $\underline{V} = (V_0, V_1, ..., V_{N-1})$ um sinal-vetor de comprimento *N* sobre um Campo de Galois GF(*p*), de característica $p \neq 2$. $\underline{\psi}_{j,k}$ são as funções *wavelet* sobre GF(p^s), $s \geq 1$. A transformada *Wavelet* sobre Corpos Finitos (TWCF) do sinal $\underline{V}$ será definida como:

$$TWCF(j,k) := \sum_{i=0}^{N-1} v_i \psi_{j,k}(i) (\text{mod } p), \tag{19}$$

o que é denotado por:

$$TWCF(j,k) = \left\langle \underline{v}, \underline{\psi}_{j,k} \right\rangle \tag{20}$$

3.8. Tipos de Wavelets

Na ferramenta do Matlab® encontram-se disponibilizadas uma série de *wavelets*, e sua utilização depende da aplicação a se realizar. A seguir se detalha de forma breve alguns tipos de *wavelets*; informação e ilustrações foram obtidas da caixa de ferramenta do Matlab® [17]:

- **Wavelet Haar ('haar').** Do grupo das *wavelets* é a mais simples: este tipo de *wavelet* é descontinua, com um único momento nulo, e se assemelha à "função passo", também pode ser considerada como uma Daubechies 'db1'; a Figura 22 ilustra a *wavelet* de Haar.

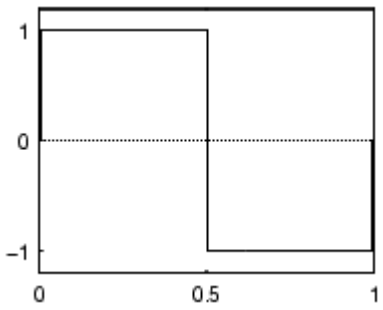


Figura 22. *Wavelet* Haar.

A expressão da *wavelet* de Haar se detalha a seguir:

$$\psi^{(H)}(t) := \begin{cases} 1 & 0 < t \le 0,5 \\ -1 & 0,5 < t \le 1 \\ 0 & \text{c.c.} \end{cases} \tag{21}$$

- **Wavelet de Daubechies ('db').** Foi criada por Ingrid Daubechies [17], são *wavelets* ortogonais, de suporte compacto e suavidade regulável, permitem uma análise mais efetiva que as *wavelets* de Haar. No conjunto de *wavelets* de Daubechies, as *wavelets* são numeradas de acordo ao número de momentos nulos que possuem. Se especificam com '*dbN*', em que, *db* é o sobrenome da *wavelet* e *N* é a ordem. A regularidade deste grupo

de *wavelets* aumenta linearmente com a ordem – parâmetro *N*, a Figura 23 mostra o conjunto das *wavelets* Daubechies.

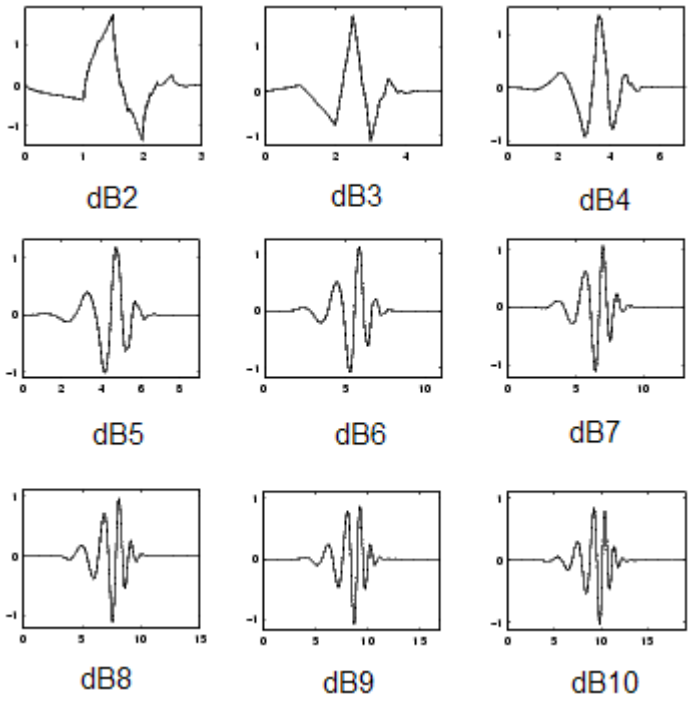


Figura 23. Família das *wavelets* Daubechies.

O suporte das *wavelets* Daubechies $\psi_{2N}^{(D)}(t) = dbN$ é o intervalo fechado dado por [1-N, N]. As *wavelets* Daubechies não possuem uma expressão específica, com exceção da *wavelets* db1, que é a *wavelet* de Haar. No entanto, o valor do módulo ao quadrado da função de transferência *h* é explícita e simples:

$$P(y) = \sum_{k=0}^{N-1} C_k^{N-1+k} y^k \tag{22}$$

em que:

C_k^{N-1+k} denota os coeficientes binomiais.

Então:

$$|m_0(w)|^2 = \left(\cos^2\left(\frac{w}{2}\right)\right)^N P\left(\sin^2\left(\frac{w}{2}\right)\right) \tag{23}$$

em que:

$$m_0(w) = \frac{1}{\sqrt{2}} \sum_{k=0}^{2N-1} h_k e^{-ikw} \tag{24}$$

- **Wavelet Symlets ('sym').** São *wavelets* quase simétricas, modificadas da família das *wavelets* '*db*', possuem de 4 a 10 números de momentos nulos, a Figura 24, mostra o grupo de *wavelet* Symlets.

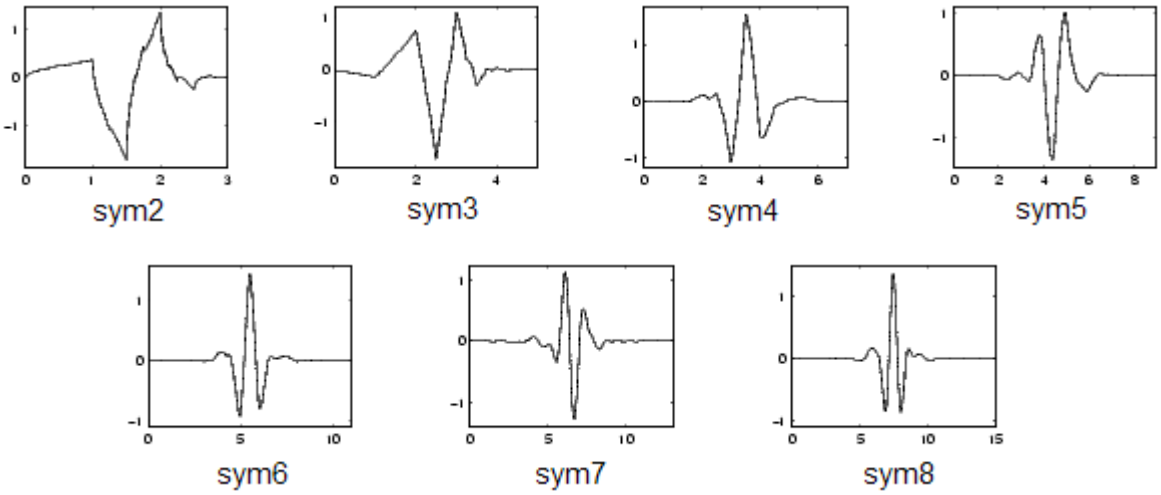


Figura 24. Família das *wavelets* Symlets.

Daubechies propôs a modificação da suas *wavelets* para incrementar sua simetria enquanto mantem grande simplicidade. A idéia consiste de reutilizar a função m_0 introduzida na *dbN*, considerando a $|m_0(w)|^2$ como uma função W de $z = e^{iw}$. Então se pode fatorar W de diversas maneiras na forma $W(z) = U(z)\overline{U\left(\frac{1}{z}\right)}$ porque a raiz de W com modulo diferente de 1 aparecem em pares. Se uma das raízes é z_1, então $\frac{1}{z_1}$ é também uma raiz.

Selecionando U de modo que o módulo de todas as suas raízes seja estritamente menor que 1, construí-se as *wavelets* de Daubechies dbN. Selecionando-se de outra forma, obtêm-se filtros mais simétricos; estes são os Symlets.

- **Wavelet Coiflets ('coif').** Sob petição de R. Coifman, Ingrid Daubechies construiu este tipo de *wavelets*, projetadas para satisfazer certo número de momentos nulos. A função *wavelet* possui 2*N* momentos nulos e a função escala tem 2*N*-1 momentos igual a 0 (cf. Figura 25). As duas funções possuem suporte de comprimento 6N-1.

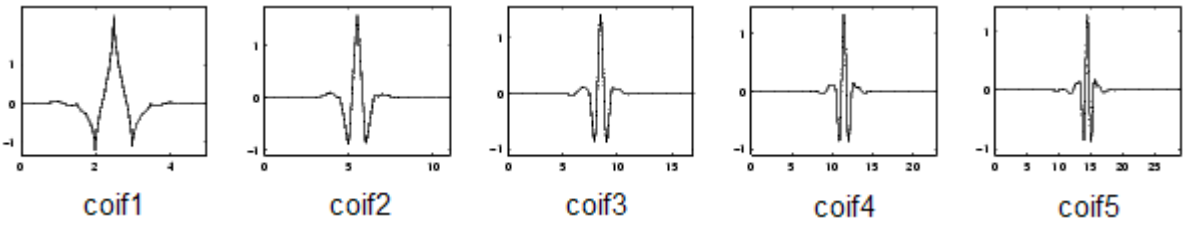


Figura 25. Família das *wavelets* Coiflets.

A *coifN* ψ e ϕ são muito mais simétricas que as *dbNs*. Com relação ao suporte de comprimento, a *coifN* pode ser comparada à *db3N* ou *sym3N*. Com relação ao número de momentos nulos de ψ, a *coifN* pode ser comparada à *db2N* ou *sym2N*.

- **Wavelet Biorthogonal ('bior').** Esta família de *wavelets* exibe a propriedade de fase linear, que é requerida para a reconstrução do sinal e da imagem. Emprega dois tipos de *wavelets*, uma para decomposição e outra para reconstrução:

  - Uma *wavelet*, $\tilde{\psi}$ é usada na análise, e os coeficientes do sinal *s* são:

$$\tilde{c}_{j,k} = \int s(x)\tilde{\psi}_{j,k}(x)dx \qquad (25)$$

  - A outra *wavelet*, ψ, é usada na síntese:

$$s = \sum_{j,k} \tilde{c}_{j,k}\psi_{j,k} \qquad (26)$$

Algumas *wavelets* desse grupo são apresentadas na Figura 26.

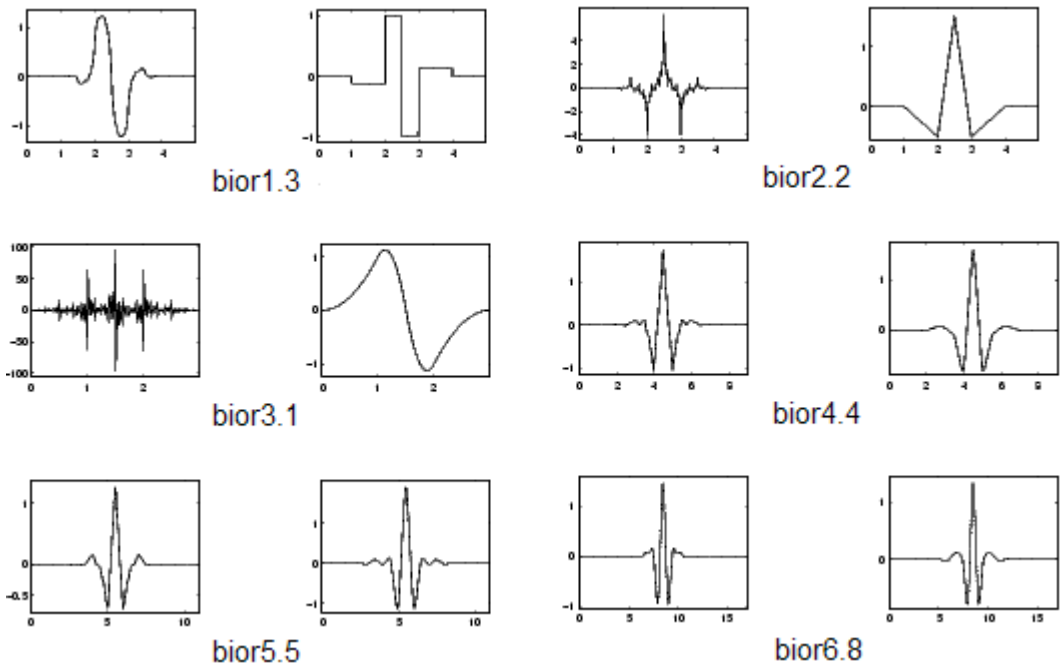


Figura 26. Grupo das *wavelets* Biorthogonal.

- **Wavelet Morlet ('morl').** Este tipo de *wavelet* não tem função escala, mas é específica, sendo a primeira *wavelet* contínua construída (cf. Figura 27).

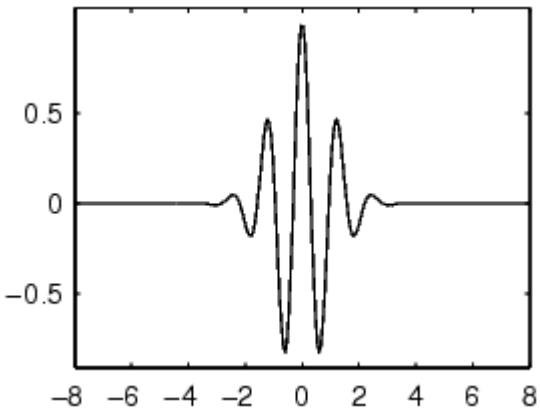


Figura 27. *Wavelet* Morlet.

A função da *wavelet* Morlet é dada pela seguinte equação:

$$\psi(t) = e^{jw_0 t} e^{-t^2/2} \tag{27}$$

- **Wavelet Chapéu Mexicano ('mexh').** Este tipo de *wavelet* não tem função escala, e se deriva de uma função que é proporcional à segunda derivada da função de probabilidade Gaussiana da função densidade, (cf. Figura 28).

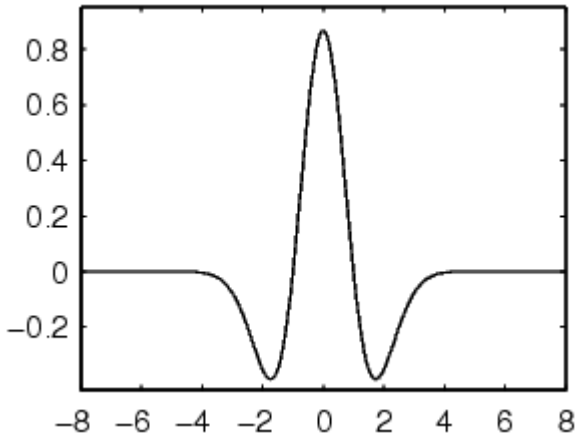


Figura 28. *Wavelet* Chapéu Mexicano.

A *wavelet* chapéu mexicano é representada mediante a seguinte equação:

$$\psi^{(Mhat)}(t) = \frac{2(t^2 - 1)e^{-t^2/2}}{\pi^{1/4}\sqrt{3}} \tag{28}$$

- **Wavelet Meyer ('meyr').** A *wavelet* e a função escala são ambas definidas no domínio da frequência, (cf. Figura 29). Este tipo de *wavelet* deriva-se da *wavelet* de Shannon, que são suavemente "enjaneladas" na frequência, de forma que o decaimento no tempo (t) possa ser mais rápido que qualquer potência de t.

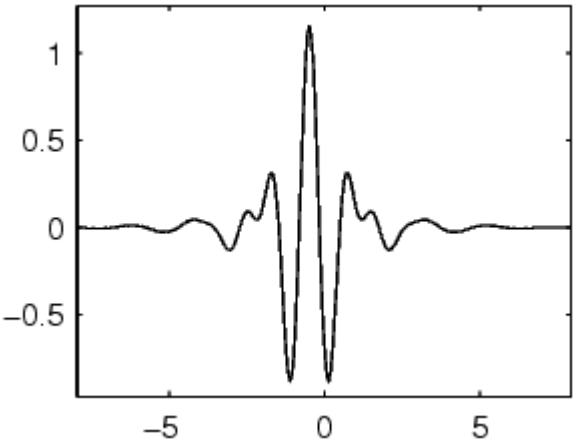


Figura 29. *Wavelet* Meyer.

A *wavelet* de Meyer é definida no domínio frequencial como se mostra [27]:

$$\psi(w) = \begin{cases} \frac{1}{\sqrt{2\pi}} sen\left[\frac{\pi}{2} v\left(\frac{3|w|}{2\pi} - 1\right)\right] e^{jw/2} & 2\pi/3 \le |w| \le 4\pi/3 \\ \frac{1}{\sqrt{2\pi}} \cos\left[\frac{\pi}{2} v\left(\frac{3|w|}{4\pi} - 1\right)\right] e^{jw/2} & 4\pi/3 \le |w| \le 8\pi/3 \\ 0 & \text{c.c.} \end{cases} \tag{29}$$

Capítulo 4

TRANSFORMADA DE WAVELETS EM ESTEGANOGRAFIA: OCULTANDO TEXTOS SIMPLES EM ARQUIVOS DE ÁUDIO

4.1. Introdução

A aplicação da esteganografia transcende o conceito de simplesmente ocultar uma mensagem. De fato, consiste em que a informação oculta dentro de um arquivo seja imperceptível aos usuários que não dispõem da chave secreta (senha-estego), diante de manipulações. Quando um arquivo no qual foi aplicado esteganografia for analisado, a informação oculta deve estar disfarçada de tal forma que não se tenha nenhum indício que nesse arquivo existe uma mensagem armazenada, podendo circular livremente sem levantar suspeitas ou questionamentos. Para a inserção de dados em áudio, podem-se mencionar diversas técnicas de codificação como: codificação em baixos bits, codificação de fase, codificação em eco [28], entre outras.

A esteganografia e a criptografia podem-se complementar. Enquanto à criptografia torna uma mensagem incompreensível a esteganografia oculta a mensagem [29-33]. O usuário poderia como primeira instância selecionar um criptosistema conhecido [34] e o comprimento da chave (128 ou 256 bits), e logo aplicar a técnica esteganográfica proposta nos dados criptografados.

Existe um compromisso entre a quantidade de informação a ser inserida e a degradação do sinal hóspede. A quantidade de texto a ser escondida dentro do sinal de áudio deve ser escolhida de tal forma que no arquivo esteganografado não se consiga detectar a presença de informação oculta nem pelo homem nem pelo computador, mediante algum tipo de sistema estatístico. Um estudo recente interessante, lista aproximadamente 32 diferentes ferramentas esteganográficas para áudio (.wav 50%, .mp3 28%, midi 6%, outros 16%) [35]. Aplicações STEGO tais como marcas de água digital e impressão digital são possíveis.

Neste capítulo, expõe-se a ocultação de dados nos bits mais significativos de uma decomposição *wavelet*. A técnica é usada de forma a garantir que o texto inserido permaneça praticamente inalterado no momento da recuperação, e passe despercebido diante manipulações, levando-se em consideração o tamanho da mensagem, do som e os níveis nos quais a mensagem é adicionada.

4.2. Garçonete, por favor: Tem um texto no meu Áudio!

Desenvolver uma técnica para "esconder" informação dentro de um arquivo de áudio é uma tarefa particularmente complicada, porque o sistema auditivo humano funciona sobre uma ampla gama dinâmica [36-37]. A sensibilidade ao ruído aditivo é incrivelmente aguda e perturbações num arquivo de som podem ser detectadas mesmo abaixo de 60 dB, o que é inferior ao nível típico do ambiente. Entretanto, existem algumas distorções ambientais tão comuns que são praticamente ignorados pelo ouvinte, na maioria dos casos. Além disso, sons mais fortes tendem a mascarar os sons fracos. Ao incorporar mensagens secretas em um arquivo de áudio, dois aspectos devem ser levados em conta: o formato digital do áudio e o provável ambiente por onde viajará o sinal entre a codificação e decodificação.

O formato mais popular para representar amostras de alta qualidade de áudio digital é com 16 bits de quantização linear [38], por exemplo, *Windows Audio-Visual* (WAV) e *Audio Interchange File Format* (AIFF). Esses formatos de áudio digital proporcionam uma boa cobertura para esteganografia de baixa inserção de dados. A taxa de amostragem do áudio determina a ocultação de dados por impor um limite superior sobre a porção funcional do espectro de frequências (se um sinal é amostrado a ~ 8 kHz, nenhuma modificação pode ter componentes de frequências além de ~ 4 kHz). Com 16 bits por amostra e uma taxa de amostragem de 44,1 kHz, o áudio digital tem uma taxa de bits para comportar mensagens extensas.

Possíveis estratégias para inserir dados no interior do áudio incluem o uso de frequências inaudíveis para os seres humanos [39-41], codificação de fase [42] (o ouvido é incapaz de perceber a fase absoluta, apenas a fase relativa), ocultação em eco [40], espalhamento espectral [43], inserção de dados empregando o LSB (*Least Significant Bit*) [29, 44-45], inserção mediante técnicas da transformação do sinal [46] (uma variante desse método é precisamente a abordagem descrita neste trabalho), e codificação em tons musicais [42]. Este capítulo apresenta os resultados prévios deste projeto esteganográfico e introduz a idéia de combinar duas chaves secretas na operação, como mostrado na Figura 30. A primeira chave secreta encripta o texto usando um sistema criptográfico padrão (por exemplo, IDEA, SAFER +, Triple DES, etc.) antes da decomposição via *wavelet* do áudio [47]. A maneira

como o texto cifrado é embutido no arquivo requer outra chave, a saber, a chave STEGO, a qual está associada com as características da análise *wavelet* do áudio.

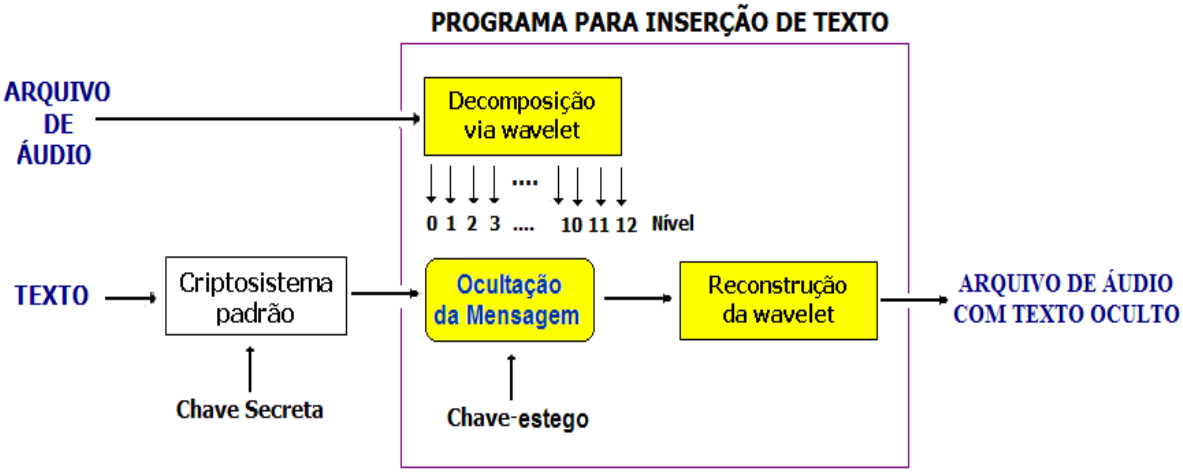


Figura 30. Método para inserir texto em um sinal de áudio mediante a técnica da transformada de *wavelet*.

4.3. As Wavelets como Ferramenta de Análise

As vantagens e características que apresentam as *wavelets* permitem a inserção dos dados em diferentes escalas, distribuindo assim a informação a ser ocultada de tal forma a não levantar nenhuma suspeita.

Trata-se de uma ferramenta poderosa no estudo e análise de sinais, possuindo características de oscilação e curta duração. Mediante a decomposição do sinal de áudio via *wavelet* têm-se aproximações e detalhes - o processo é repetitivo, e as aproximações são subdivididas para cada nível. A Figura 31 ilustra a decomposição de um sinal através das *wavelets*.

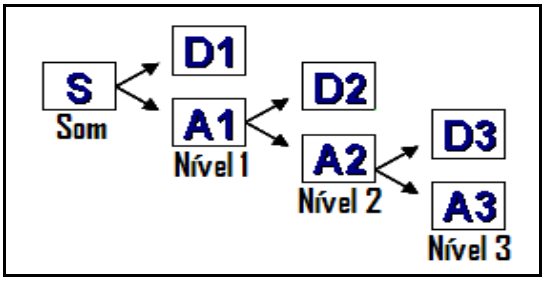


Figura 31. Decomposição de um sinal de áudio via *wavelets*.

O processo inverso é usado na reconstrução, como ilustrado na Figura 32. Observa-se como gerar o sinal original a partir de componentes de aproximação (A3) e detalhes (D3, D2, D1), de acordo com:

$$S = A3 + D3 + D2 + D1 \tag{30}$$

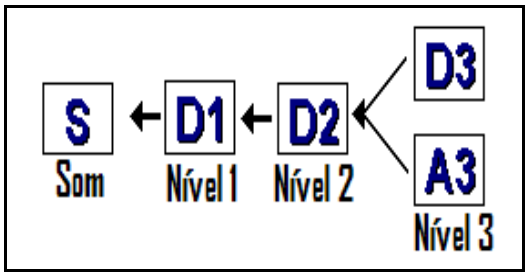


Figura 32. Reconstrução do sinal de áudio decomposto em três níveis.

Existem diversas aplicações que se podem realizar com o uso das *wavelets*, uma delas é a análise de sinais de áudio. Graças às características que apresentam, foi possível levar a cabo a inserção dos dados em diferentes níveis, distribuindo a informação de forma a não levantar suspeita.

4.4. Implementação do estego-sistema para Áudio

A aplicação foi desenvolvida no Matlab® [48] para inserir informação secreta no som. A idéia baseia-se na ocultação de uma mensagem dentro de um arquivo de som, realizando a decomposição do sinal via *wavelets* e introduzindo os dados em casas decimais (previamente estabelecidas por uma chave secreta) da matriz gerada.

Na decomposição do áudio, geram-se duas matrizes; a primeira, contem informação da decomposição e é de ordem $n \times 1$, em que a coluna especifica o número de canais do som (neste caso monofônico), e a segunda matriz define o comprimento empregado em cada nível da decomposição.

A mensagem a ser introduzida é convertida previamente em código ASCII. Diferentemente de todos os esquemas mencionados anteriormente [49-50], a inserção dos dados se realiza nas três primeiras casas decimais da matriz resultante da decomposição *wavelet* do som, eliminando estes elementos e substituindo-os pelos dados a inserir. Sendo estes os três bits mais significantes, deve-se ter cautela com o tamanho da mensagem e do som. Restrições

foram assumidas no momento da realização do programa, introduzindo mensagens de tamanho 1.000 vezes menor que o do som original, evitando assim, uma modificação facilmente detectável no arquivo de áudio.

Na decomposição do sinal são necessários três elementos:

- A matriz do arquivo de áudio no qual se vai inserir a mensagem.
- O número de níveis no qual vai ser decomposto o sinal de som.
- O tipo de *wavelet* empregada.

Dentro deste aplicativo, disponibilizou-se um grupo de *wavelets*-padrão tomadas da caixa de ferramentas do Matlab®, podendo o usuário escolher uma entre cinquenta e quatro possibilidades. O valor do número máximo de níveis foi fixado em 12, devido a que se esta trabalhando com arquivos de áudio de tamanhos relativamente pequenos (com o objetivo que passem despercebidos na rede), o valor de 12 considerou-se apropriado para que o número de amostras nos últimos níveis não seja mínimo (1 amostra) e permita a inserção da informação. Os níveis que resultam da decomposição permitem ao usuário a opção de inserir, ou não, dados, fragmentando assim a mensagem. Dependendo do tamanho do som, há aproximadamente entre 2^{10} e 2^{12} alternativas para eleger a chave, ou seja, entre 1024 e 4096 possibilidades distintas para inserção de dados. Isto acrescenta pelo menos 10 bits para a chave secreta.

Para um melhor espalhamento dos dados dentro de cada nível são utilizadas senhas alfanuméricas de tamanho proporcional à quantidade de caracteres inseridos em cada um dos níveis ou sublocos. A senha – transformada previamente em código binário – posiciona cada elemento dos dados nos locais onde se encontra um 1. Para especificar onde se inicia a entrada do texto dentro de cada subloco, tem-se que eleger a posição (restringida pelo tamanho do comprimento de cada nível). Depois de realizados os passos supramencionados, o sinal é reconstruído armazenando a mensagem contendo o texto escondido.
Para o processo inverso, i.e., a recuperação dos dados, é preciso conhecer um grupo de informações necessárias que permitirão decifrar a mensagem oculta de maneira correta:

• Tipo de *Wavelet*.

- Tamanho da mensagem.
- Número de níveis.
- Nível(s) onde o texto foi inserido.
- Tamanho dos sublocos de texto escolhidos em cada nível.
- Senha(s): alfanuméricas, uma senha em cada nível.
- Local de inserção do texto.

4.5. Linha de Raciocínio para Inserção de Dados Ocultos

Inserção de dados em um som:

1) Ingresso do texto com caracteres alfanuméricos.
2) Leitura de um arquivo de som.
3) Escolha da *wavelet* para decomposição do áudio.
4) Decomposição do som, o número de níveis foi predeterminado em 12.
5) Divisão do texto em níveis para alocação da informação em diferentes escalas (cf. Figura 34).
6) Ingresso da(s) senha(s).
7) Local de inserção do texto.
8) Inserção dos dados nas matrizes de coeficientes da decomposição *wavelet*.
9) Reconstrução e gravação do sinal contendo a informação de texto.

O texto a inserir é previamente convertido em código ASCII como mostra a Figura 33, para poder realizar a inserção dentro do arquivo de áudio selecionado.

Figura 33. Conversão do texto alfanumérico em código ASCII.

Junto com a seleção da *wavelet* é feita a decomposição do som em 12 níveis; então o texto é dividido a critério em sublocos para ser alocado em diferentes escalas (cf. Figura 34).

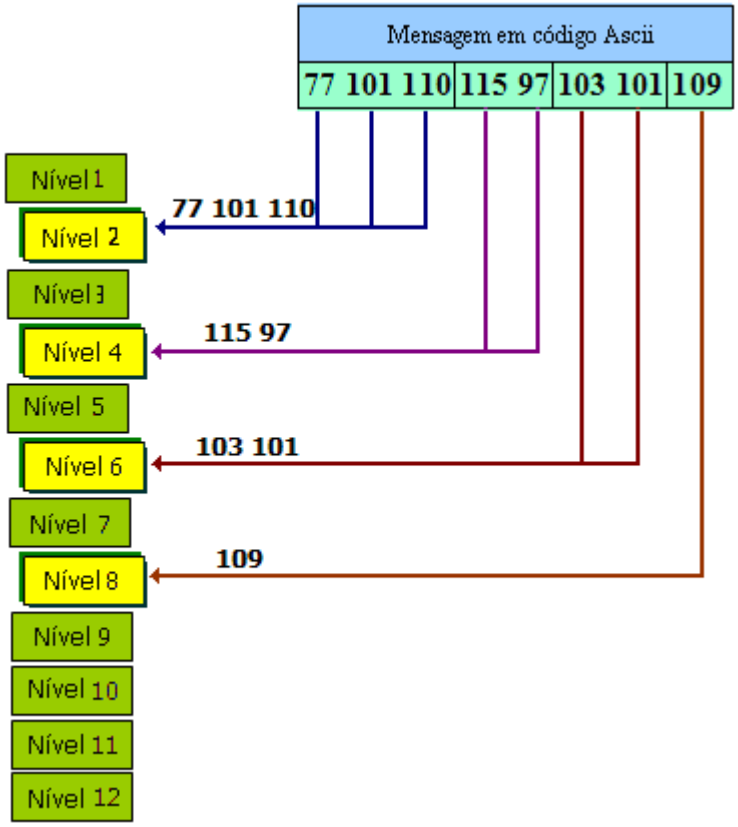


Figura 34. Subdivisão da mensagem para ser inserida nos níveis especificados (2, 4, 6 e 8).

Uma vez subdividida a mensagem, ingressam-se senhas alfanuméricas para um melhor espalhamento da informação dentro do arquivo de áudio, como ilustrado na Figura 35. Estas senhas são mapeadas em código ASCII e posteriormente em código binário, as quais são proporcionais ao tamanho do bloco de texto inserido, já que cada número 1 representa uma posição dentro da matriz de decomposição para armazenar um dado.

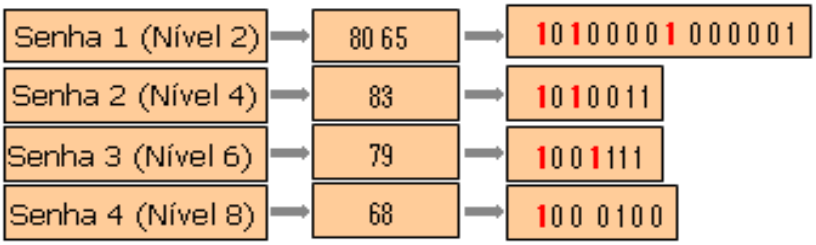


Figura 35. Senhas em código binário correspondentes a cada nível de decomposição.

O local de inserção do texto indica o início da colocação da mensagem dentro de cada nível. A Figura 34 mostra quatro níveis selecionados (2, 4, 6 e 8) nos quais os dados serão inseridos. O esboço a seguir, Figura 36, mostra um exemplo da inserção dos dados no

segundo nível, o ingresso nos outros níveis é similar, de acordo com os dados fornecidos pelo usuário.

A inserção do texto é realizada por substituição, ou seja, as três primeiras casas decimais são eliminadas e trocadas pela mensagem.

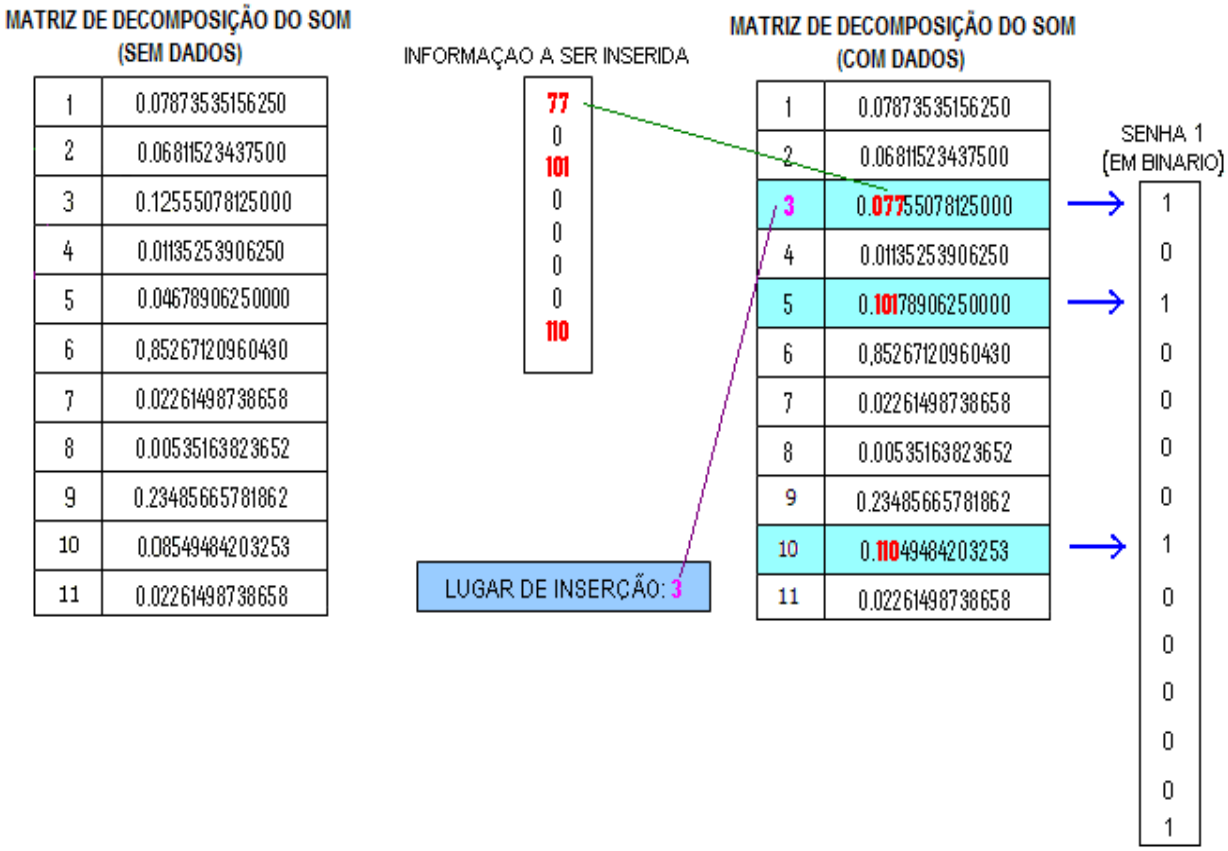


Figura 36. Inserção de dados na matriz de decomposição via *wavelet*.

Por último, o arquivo de som é reconstruído contendo a informação (escondida) de texto.

A Figura 37 esboça de forma geral a inserção de dados no nível 2 (o processo é repetido para os outros níveis 4, 6 e 8) num arquivo de áudio.

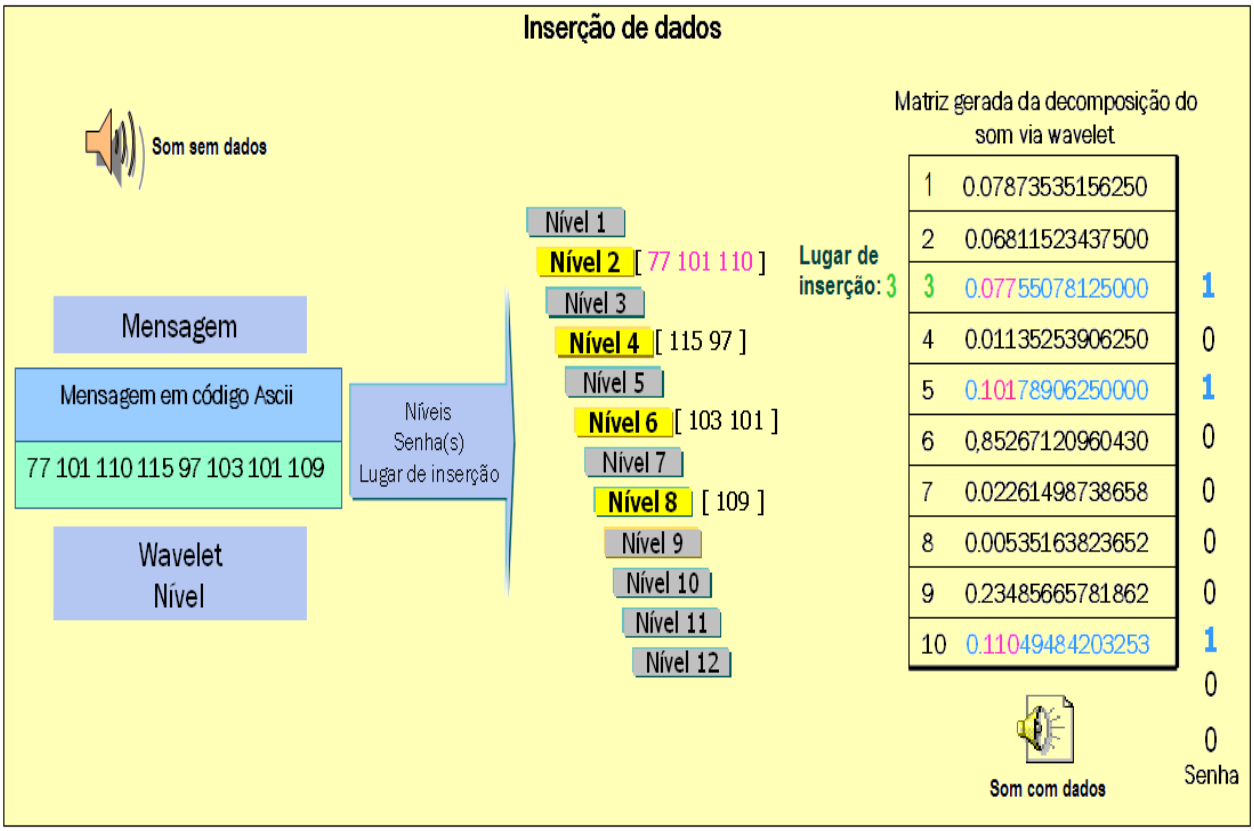


Figura 37. Inserção de dados num som mediante a técnica da transformada de *wavelet*.

Decodificação dos dados:

1) Leitura do arquivo de som contendo a mensagem "escondida".
2) Ingresso dos parâmetros:
   - Tipo de *wavelet*.
   - Tamanho da mensagem.
   - Número de níveis contendo informação do texto.
   - Especificação dos níveis.
   - Tamanho do texto inserido.
   - Senha(s).
   - Lugar(s) de inserção do texto.
3) Decomposição do som usando a *wavelet* selecionada.
4) Filtragem do texto no(s) nível(is) especificado.
5) Recuperação da mensagem.

A Figura 38 esboça o gráfico geral da recuperação dos dados num arquivo de áudio.

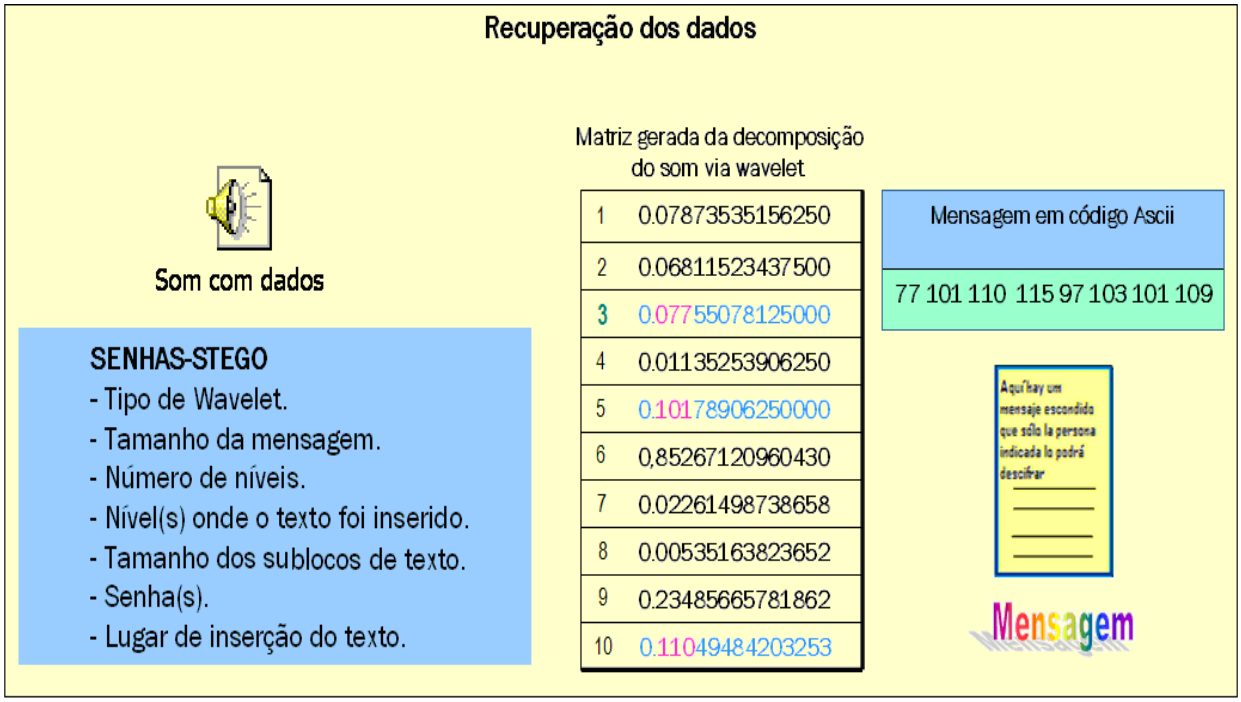


Figura 38. Decodificação dos dados num som esteganografado mediante a técnica da transformada de *wavelet*.

4.6. Testes de Validação

O tamanho da mensagem e o som se encontram restringidos em um fator de 1.000, ou seja, o número de elementos da matriz representativa do som deverá ser pelo menos mil vezes superior ao tamanho da mensagem, visando manter uma variação apropriada do áudio. Realizando uma análise estatística da variação do som, verificou-se que existem três fatores que influenciam a modificação percentual do áudio no momento da inserção dos dados:

• *Quantidade e localização dos níveis nos quais são inseridos os dados.*

Além do número de níveis nos quais os dados são inseridos, é necessário levar em conta o nível de inserção, quando os dados são inseridos nos primeiros níveis a modificação é menor. O contrário acontece, informação inserida nos últimos níveis ocasiona uma modificação marcante.

• *Tamanho da mensagem.*

É diretamente proporcional com a variação: quanto maior o tamanho da mensagem, maior a porcentagem de modificação do áudio.

• *Tamanho do arquivo de som.*

É inversamente proporcional à variação da porcentagem: quanto maior o tamanho do arquivo de áudio, menor a variação.

Se os dados são inseridos unicamente nos primeiros níveis, a variação é relativamente pequena; dados ingressados nos últimos níveis acarretam maior alteração; uma distribuição de dados em todos os níveis produz uma maior modificação. As tabelas a seguir mostram a variação percentual no áudio conforme variam os parâmetros supramencionados. Para cada tabela se fixou o tamanho do som, realizando a variação dos outros parâmetros.

Tabela 6: Variação percentual para um arquivo de som de 10 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5	5										-	-	-	0.1089799
5			5								-	-	-	0.3487358
5					5						-	-	-	1.394943
5							5				-	-	-	4.184830
7	2	2	2	1							-	-	-	0.4359198
7	2	2	1	1	1						-	-	-	0.2179599
9		9									-	-	-	0.3923278
9						9					-	-	-	4.184830
9	5			4							-	-	-	0.4577158
9					6			3			-	-	-	4.533566
9							8			1	-	-	-	5.579773
9	2		2		2		2		1		-	-	-	8.936356
TAMANHO DO ARQUIVO DE ÁUDIO: 10 kB														

Tabela 7: Variação percentual para um arquivo de som de 25 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5	5											-	-	0.08108985
5			5									-	-	0.3243594
5					5							-	-	1.297438
5							5					-	-	5.189750
9		9										-	-	0.2919235
9						9						-	-	4.670775
9	5			4								-	-	0.6000649
9					6			3				-	-	7.006163
9							8			1		-	-	16.60720
9	2		2		2		2		1			-	-	6.908855
12	12											-	-	0.1946156
12				12								-	-	1.556925
12		4			4			4				-	-	8.303600
12						6	3		2	1		-	-	17.64515
12	2	1	1	1	1	1	1	1	1	1	1	-	-	32.67921
TAMANHO DO ARQUIVO DE AUDIO: 25 kB														

Tabela 8: Variação percentual para um arquivo de som de 50 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)												VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5		5												0.07951970
5					5									0.6361576
5								3			1		1	24.42845
10			10											0.3180788
10						7				3				12.46869
10	2		2				3				1	1	1	42.25677
15	15													0.1192796
15				15										0.9542364
15							14						1	23.41060
20		20												0.3180788
20					10			10						11.45084
20			6			7			7					15.96756
20			5		5		5		5					13.42293
25				25										1.590394
25						23				2				13.99547
25		9		7		3		3		2			1	25.01690
TAMANHO DO ARQUIVO DE ÁUDIO: 50 kB														

Tabela 9: Variação percentual para um arquivo de som de 100 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5		5												0.03706930
5					5									0.2965544
5								3			1		1	12.81115
10			10											0.1482772
10						7				3				6.524197
10	2		2				3				1	1	1	15.93239
20		20												0.1482772
20					10			10						5.337979
20			6			7			7					7.562137
20			5		5		5		5					6.301781
30				30										0.8896632
30						28				2				7.117306
30		15						14					1	13.39685
40	20						20							4.819009
40				17		12			11					12.36632
53					53									3.143477
53		16				30				7				16.96291
53	5	10	6		7	3	8		9		3		2	26.54533
TAMANHO DO ARQUIVO DE ÁUDIO: 100 kB														

Tabela 10: Variação percentual para um arquivo de som de 200 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
20		20												0.03859700
20					10			10						1.018961
20			6			7			7					0.6715878
20			5		5		5		5					0.8221161
60	30	15	15											0.1157910
60				30		30								0.9958026
100								100						7.410624
100	50								50					3.753558
150							150							6.175520
150					75			75						7.364307
150		40		66					30		14			3.261446
207	207													0.1997395
207					100			100				7		7.812033
207								93	57	29	14	7	7	27.17229
TAMANHO DO ARQUIVO DE AUDIO: 200 kB														

Tabela 11: Variação percentual para um arquivo de som de 300 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
20		20												0.02573133
20					10			10						0.5969669
20			6			7			7					0.6329908
20			5		5		5		5					0.6715878
60	30	15	15											0.07719400
60				30		30								0.7101848
100								100						3.952333
100	50								50					2.173011
150							150							3.869992
150					75			75						3.983210
150		40		66					30		14			3.911163
200	200													0.1286567
200					100			100						5.115389
200								93	57	29	14	7		8.686898
310		310												0.3924028
310						110		160		40				9.695566
310	90		60		50		40		40		20		10	17.15411
TAMANHO DO ARQUIVO DE AUDIO: 300 kB														

Tabela 12: Variação percentual para um arquivo de som de 600 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
20					10			10						0.2436022
20			6			7			7					0.3291221
20			5		5		5		5					0.2500810
60	30	15	15											0.03887269
60				30		30								0.3109815
100								100						1.824425
100	50								50					2.155329
150							150							2.529316
150					75			75						2.306122
150		40		66					30		14			2.003725
200	200													0.06187237
200					100			100						2.979754
200								93	57	29	14	7		5.319080
300		300												0.1937156
300						100		160		40				6.473599
617	300			317										0.8545513
617			174				400				43			7.410755
617	123		123		200		86		43		21		21	17.70943
TAMANHO DO ARQUIVO DE AUDIO: 600 kB														

Tabela 13: Variação percentual para um arquivo de som de 930 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
20					10			10						0.2155245
20			6			7			7					0.2597212
20			5		5		5		5					0.1587302
60	30	15	15											0.02435542
60				30		30								0.1721676
100								100						1.719367
100	50								50					1.192996
150					75			75						1.622575
150		40		66					30		14			1.089485
200	200													0.04115226
200					100			100						1.753485
200								93	57	29	14	7		3.434954
300		300												0.1217771
300						100		160		40				3.982741
617	300			317										0.5416982
617			174				400				43			5.720270
952							786			133		33		11.06366
952		300		352		150		150						3.417003
952	222		200		265		133		66		33		33	15.79176
TAMANHO DO ARQUIVO DE AUDIO: 930 kB														

Existe um compromisso (*trade-off*) entre a distribuição de dados nos níveis, e a possibilidade da informação oculta ser desvendada por terceiros. Enquanto os dados estão mais dispersos, existe menor possibilidade da informação completa ser desvendada, porém a variação percentual será significativa; por outro lado se os dados fossem ingressados em um único dos primeiros níveis, a variação seria sutil, no entanto a probabilidade de que detectores obtenham a mensagem completa cresce.

Para poder determinar a variação percentual que sofreu um arquivo de áudio quando foi inserida uma mensagem, foi realizado um processo de comparação. Tal procedimento consiste em comparar as amostras dos sons (original e esteganografado), ou seja, cada linha da matriz gerada da leitura do áudio original é comparada com a linha correspondente da matriz do áudio com dados. Dessa forma, verifica-se quanto pode ter mudado um áudio com relação ao outro, podendo-se obter uma estimativa do valor percentual; o algoritmo empregado foi desenvolvido no Matlab® e se detalha a seguir:

```
% ANÁLISE ESTATÍSTICA DO SOM
   clear all
   clc

% Carregando som
   format long

   prompt = {'Digite o nome do arquivo de Som Original:'};
   dlg_title = 'ANÁLISE ESTATÍSTICA DO SOM';
   num_lines = [1];
   entrada = inputdlg(prompt,dlg_title,num_lines);
   prompt1 = {'Digite o nome do arquivo de Som com Dados'};
   entrada1 = inputdlg(prompt1,dlg_title,num_lines);

   A1 = wavread(entrada{1});
   A2 = wavread(entrada1{1});

% Comparando as matrizes dos Sons
   compara = (A1(1:end) == A2(1:end));
   diferente = find(compara == 0);
   format short
   porcentagem = (numel(diferente)/numel(compara))*100;
   sprintf('O som mudou %u porcento do Som Original',porcentagem)
```

Para poder definir um valor aceitável da variação percentual que pode sofrer um arquivo de áudio, se aplicou testes de escuta nos sons que apresentavam maior variação percentual quando inserida uma mensagem. Percebeu-se que quando a variação percentual era menor que 25%, nos testes de escuta não se identificaram modificações marcantes (considerando que a detecção de algum ruído é subjetiva), para os arquivos de áudio que apresentam valores percentuais acima de 25%, há detecção de algum tipo de ruído. Apesar destes fatos, nenhum dos arquivos de áudio nos quais se ocultaram mensagens delata auditivamente a suspeita de informação oculta. De acordo com a informação acima mencionada poder-se-ia estimar um valor percentual tolerável de 25% para que a inserção de dados passe despercebida, não sendo este valor fixo, já que varia devido a subjetividade da detecção de um ruído por parte do ouvinte.

A condição para a introdução dos dados no som foi estabelecida pela seguinte desigualdade:

$$x.f \leq M\,, \tag{31}$$

em que:

x é o tamanho da mensagem introduzida.

M é o tamanho do arquivo de áudio.

f é um fator condicional.

A interpretação da desigualdade é clara, quanto maior o fator condicional, menor o tamanho da mensagem que se pode introduzir num dado arquivo de áudio fixo. Pode-se expressar o tamanho da mensagem introduzida (x), como sendo a quantidade de caracteres na mensagem (n) multiplicada pelo tamanho médio de cada caractere inserido ($\overline{c}$):

$$x = n.\overline{c} \tag{32}$$

A desigualdade (31) fica:

$$n.\overline{c}.f \leq M \tag{33}$$

Se quisermos o máximo de caracteres que podem ser inseridos no áudio, a desigualdade torna-se uma igualdade e então:

$$n_{max} \approx \frac{M}{\overline{c}.f} \tag{34}$$

Com os dados obtidos experimentalmente se pôde estimar o tamanho médio ocupado pelo caractere. Baseando-se na Tabela 6, foi constatado que com o fator de 1000 se pode ingressar no máximo 9 caracteres em um som de 10kB:

$$\overline{c} \approx \frac{M}{n_{\max}.f} = \frac{10\,KB}{9*1000} \approx 1\ byte\,, \tag{35}$$

sendo 1 byte o valor médio que ocupa cada caractere inserido no áudio.

4.7. Conclusões do Capítulo

Mediante as transformadas de *wavelets* se realizou a inserção de dados em arquivos de sons, onde o objetivo principal consiste em que o áudio modificado passe despercebido ante a percepção humana. Alguns métodos inserem dados nos bits menos significativos para garantir a qualidade do som. Pôde-se evidenciar que é válido ingressar dados nos bits mais significativos, tendo em conta as considerações apropriadas como quantidade de dados a inserir e tamanho do som, de tal forma que não se tenha uma degradação auditiva inaceitável no sinal.

Os resultados obtidos neste trabalho são resultados preliminares, e a ocultação dos dados é feita de tal forma a não inspirar suspeita de que determinado arquivo de áudio contém alguma informação oculta.

Nesta versão inicial do software, a estego-senha é fornecida por etapas, como descrito no texto deste capítulo. Porém, para facilitar a distribuição da senha e tornar prático o procedimento de inserção de textos, disponibiliza-se (sob demanda a jpcc1@hotmail.com) uma versão do tipo "senha integral", na qual a (enorme) estego-senha é fornecida ao programa em um só passo. A idéia é desenvolver futuramente um procedimento que a partir de um texto selecionado (e.g. poesia, pagina de livro), gere automaticamente uma estego-senha compatível como o procedimento. Isso tornaria viável e prático o uso do programa desenvolvido.

Capítulo 5

IMPLEMENTAÇÃO DO PROGRAMA NO MATLAB

5.1. A Implementação do Programa

Usando os comandos do Matlab® foi desenvolvido o programa tanto para a inserção quanto para a recuperação dos dados em arquivos de áudio com formato wav. A seguir se disponibiliza o código fonte (*freeware*).

5.2. Programa de Inserção de Dados

```
% Esteganografia: Ingresso de dados em um som em 12 níveis.
    clear all
    clc

% Carregando som
    format long
    prompt1 = {'Digite o nome do arquivo de audio a ser inserido
                dados:'};
    dlg_title1 = 'LEITURA DO ARQUIVO DE AUDIO';
    num_lines = [1];
    entrada1 = inputdlg(prompt1,dlg_title1,num_lines);
    Nome = cell2mat(entrada1);
    [A,Fs,bits] = wavread(Nome);

% Texto em código ascii
    prompt = {'Digite o texto a esconder:'};
    dlg_title = 'ROTINA PARA ESTENOGRAFIA';
    num_lines = [1];
    entrada = inputdlg(prompt,dlg_title,num_lines);
    S = cell2mat(entrada);
    save arquivo S -ascii;
    load arquivo S -ascii;

% Ingressando o tipo de wavelet
    prompt = {'Ingresse o tipo de Wavelet:'};
    dlg_title = 'Nome da Wavelet';
    num_lines = [1];
    ingresso = inputdlg(prompt,dlg_title,num_lines);
    w = lower(cell2mat(ingresso));

    Permitidas = {'haar','db1','db2','db3','db4','db5','db6', 'db7',
                  'db8', 'db9','db10','sym2','sym3','sym4', 'sym5',
                  'sym6','sym7','sym8','coif1','coif2','coif3',
                  'coif4','coif5','bior1.1','bior1.3','bior1.5',
```

```
            'bior2.2','bior2.4','bior2.6','bior2.8','bior3.1',
            'bior3.3','bior3.5','bior3.7','bior3.9','bior4.4',
            'bior5.5','bior6.8','rbio1.1','rbio1.3','rbio1.5',
            'rbio2.2','rbio2.4','rbio2.6','rbio2.8','rbio3.1',
            'rbio3.3', 'rbio3.5','rbio3.7', 'rbio3.9','rbio4.4',
            'rbio5.5','rbio6.8','dmey'};

if strcmp(w,Permitidas) == 0
   errordlg('Essa wavelet não está cadastrada!'), break
else
 % Decomposição do som
   [C,L] = wavedec(A,12,w);

   % Condição do texto e som
     if numel(C) <= numel(arquivo)*1000    %fator de 1000
        errordlg('Diminua o texto ou troque por um arquivo de
                  áudio maior'),break
     else

         % Divisão da matriz C
           for k0 = 1:13
               C1{k0} = C(sum(L(1:(k0-1)))+1:sum(L(1:k0)));
           end

         % Inversão da matriz C
           C11 = C1(13:-1:1);
           acumulado = 0;

         % Inserção do texto nos diferentes níveis
           valor_inserir = floor(L/7);
           niveis_inserir = find(valor_inserir);
           element_niveis_inserir = numel(niveis_inserir)-1;
           niveis = valor_inserir(niveis_inserir);
           button = questdlg(sprintf('A inserção de dados pode ser
                    realizada em %u níveis. Deseja continuar....',
                    element_niveis_inserir),'Níveis Disponíveis',
                    'Sim','Não','default');
           if button == 'Sim'
             for k = 1:element_niveis_inserir
                 if acumulado < numel(arquivo)
                    dlg_title = (sprintf('%u é o tamanho do texto.
                                Divida o texto',numel(arquivo)-
                                acumulado));
                    if valor_inserir(14-k) > numel(arquivo)-
                       acumulado
```

```
                prompt = (sprintf('Ingresse o tamanho do
                                  texto no nível %u',k));
            else
                prompt = (sprintf('Ingresse o tamanho do
                          texto no nível %u, no máximo %u
                          caracteres:',k,valor_inserir
                          (14-k)));
            end
                num_lines = [1];
                ingresso(k) = inputdlg(prompt,dlg_title,
                num_lines);
                acumulado = acumulado + str2num(cell2mat
                (ingresso(k)));
            if ((str2num(cell2mat(ingresso(k))) < 0) ||
               (str2num(cell2mat(ingresso(k)))>
                valor_inserir(14-k)))
                errordlg(sprintf('O número ingressado esta
                         fora do limite, é menor que 0 o
                         maior que %u',valor_inserir
                         (14- k))),break
            else
                num(k) = str2num(cell2mat(ingresso(k)));
            end
        end

    end

     soma = sum(num);
     if soma <= 0
        errordlg('Não foram digitados dados em nenhum
                  nível'),break
     elseif soma > numel(arquivo)
            errordlg('Excedeu o tamanho do texto'),break
     end

     ingre = find(num);
     lugar = niveis(element_niveis_inserir:-1:1);
     L11 = L(niveis_inserir); L1 =
           L11(element_niveis_inserir:-1:1);
     lu = {}; C2 = {};
  end

% Senha e lugar de inserção do texto
  for k1 = 1:numel(ingre)
      s{k1} = input(sprintf('Ingresse a senha %u:',k1),
```

```
                        's');
senha{k1} = (char(s(k1)));

% Condição da senha
  se_bi{k1} = de2bi(double(cell2mat(s(k1))));
  se1_bi{k1} = reshape(se_bi{k1}',1,numel
               (se_bi{k1}));
  u{k1} = sum(se1_bi{k1}); se_m = cell2mat(u);
  ingre1 = num(ingre);
  if se_m(k1) < ingre1(k1)
     errordlg(sprintf('A senha %u precisa ter mais
                      caracteres',k1)),break
  else

      % Matriz senha
        senha_l{k1} = find(cell2mat(se1_bi(k1))
                           ==1);
        senha_lu(k1) = senha_l{k1}(ingre1(k1));
        senha_F{k1} = se1_bi{k1}(1:senha_lu(k1));

      % Lugar de inserção do texto
        posi(k1) = L1(ingre(k1))- (numel
                                   (senha_F{k1}));
        lug(k1) = input(sprintf('Ingresse o lugar
                        de inserção do texto %u [1-
                        %u]:',k1,posi(k1)));
        disp(' ');

      % Condição de lugar de inserção do texto
        if (lug(k1) <= 0) || (lug(k1) > posi(k1))
           errordlg(sprintf('Eleja um lugar entre 1
           e %u',posi(k1))), break
        else
            C2(k1) = C11(ingre(k1));
            mi{k1} = [zeros(1,lug(k1)-1),
            senha_F{k1}, zeros(1,numel(C2{k1})-
            (lug(k1)-1)-numel(senha_F{k1}))];
        end

      % Divisão do arquivo de dados
        ingre2{k1} = arquivo(sum(ingre1(1:(k1-
                     1)))+1:sum(ingre1(1:k1)));
  end
  % Inserção de Dados
    p{k1} = find((mi{k1})==1);
```

```
                    arq1{k1} = (ingre2{k1})*10^-3;
                    c{k1} = C2{k1}';
                    c1{k1} = c{k1}(p{k1});
                    c2{k1} = find((c1{k1})<0);
                    c3{k1} = (c1{k1}(c2{k1}))*-1;
                    c4{k1} = c1{k1};
                    c4{k1}(c2{k1}) = c3{k1};
                    c5{k1} = (floor((c4{k1})*10^3))*10^-3;
                    dados1{k1} = (c4{k1} - c5{k1}) + arq1{k1};

                    sig{k1} = (dados1{k1}(c2{k1}))*-1;
                    dados2{k1} = dados1{k1};
                    dados2{k1}(c2{k1}) = sig{k1};
                    C3{k1} = C2{k1}; C3{k1}(p{k1}) = dados2{k1};
                end

                C4 = C11;
                for k2 = 1:numel(ingre)
                    C4{ingre(k2)} = C3{k2};
                end

                C5 = C4(13:-1:1);
            end
        end

% Recuperação do som
  C6 = cell2mat(C5(:));
  A1 = waverec(C6,L,w);

% Salvando som
  prompt2 = {'Digite o nome do arquivo a salvar:'};
  dlg_title2 = 'SALVANDO O ARQUIVO DE AUDIO:';
  entrada2 = inputdlg(prompt2,dlg_title2,num_lines);
  N = cell2mat(entrada2);
  wavwrite(A1,Fs,N);
  sound(A1,Fs,bits);
```

5.3. Programa de Recuperação de Dados

```
% Decripta os dados de um som
    clear all
    clc

% Carregando o som
```

```
    format long
    prompt = {'Digite o nome do arquivo de audio para recuperar a
              mensagem:'};
    dlg_title = 'LEITURA DO ARQUIVO DE AUDIO COM MENSAGEM';
    num_lines = [1];
    entrada = inputdlg(prompt,dlg_title,num_lines);
    Nome = cell2mat(entrada);
    A1 = wavread(Nome);

% Decomposição do som e obtenção da senha
    w = input('Ingresse o tipo de wavelet:','s');
    disp(' ');
    N = input('Ingresse o tamanho da mensagem:');
    disp(' ');
    n = input('Em quantos níveis ingresso dados:');
    disp(' ');
    for k = 1:n
        ni(k) = input(sprintf('Ingresse o nível em que ingresso o
                              texto %u:',k));
        t(k) = input(sprintf('Ingresse o tamanho do texto %u:',k));
        s{k} = input(sprintf('Ingresse a senha %u:',k),'s');
        lu(k) = input(sprintf('Ingresse o lugar de inserção do texto
                              %u:',k));
        disp(' ');
    end

% Decompondo o som
    [C,L] = wavedec(A1,12,w);

% Divisão em níveis
    for k1 = 1:13
        C1{k1} = C(sum(L(1:(k1-1)))+1:sum(L(1:k1)));
    end

    C2 = C1(13:-1:1);
    C3 = C2(ni);
% Lugar e Criação da matriz senha
    senha2 = {};
    si = {};
    for k2 = 1:n
        senha{k2} = (char(s(k2)));
        se_bi{k2} = de2bi(double(cell2mat(senha(k2))));
        se1_bi{k2} = reshape(se_bi{k2}',1,numel(se_bi{k2}));
        b{k2} = find(se1_bi{k2} == 1);
        senha1{k2} = b{k2}(1:t(k2));
```

```
        ultimo{k2} =  senha1{k2}(end);
        ultimo1 = cell2mat(ultimo);
        senha2{k2} = se1_bi{k2}(1:ultimo1(k2));
        si{k2} = [zeros(1,lu(k2)-1),senha2{k2},zeros(1,numel(C3{k2})-
                 (lu(k2)-1)-numel(senha2{k2}))];
    end

% Filtração de dados
  C4 = cell2mat(C3(:));
  si1 = cell2mat(si(1:end));
  p = C4'.*si1;
  q = abs(p);
  indices = find(q);
  p1 = q(indices);

% Recuperação da mensagem
  recu = p1*10^3;
  recu1 = floor(recu);
  mensagem = char(recu1)
```

5.4. Interface do Programa

A aplicação conta com uma série de janelas que permitem a interação do usuário com o programa, realizando a inserção dos dados necessários para a ocultação e recuperação da mensagem dentro do arquivo de som.

De forma detalhada o processo de ocultação e recuperação de uma mensagem inserida em um arquivo de áudio de 10 kB é descrito a seguir.

5.4.1. Esquema Gráfico da Inserção de Dados num Arquivo de Áudio

A inserção dos dados é realizada mediante a decomposição do sinal, para o qual se precisam especificar três parâmetros que permitirão tal decomposição:

- O arquivo encobridor (o áudio).
- A mensagem.
- O tipo de *wavelet*.

Como primeiro passo se tem a seleção do arquivo de áudio em formato wav no qual vão ser inseridos os dados ou mensagem, Figura 39, o áudio deve-se encontrar gravado na pasta de trabalho do Matlab® para ser oportunamente "chamado".

A escolha do arquivo de áudio é fundamental já que o tamanho da mensagem depende diretamente do tamanho do som.

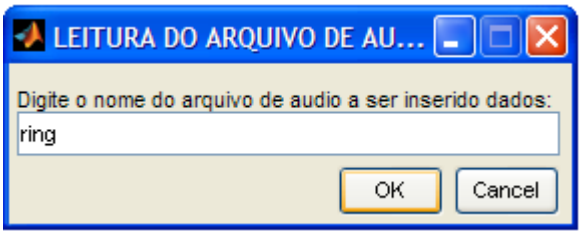


Figura 39. Janela de leitura do arquivo de áudio.

Insere-se a mensagem que vai ser oculta, Figura 40.

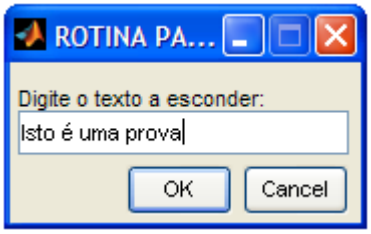


Figura 40. Janela de inserção da mensagem.

O último dado necessário para a decomposição do sinal é o tipo de *wavelet*, Figura 41. O grupo das *wavelets* disponibilizadas no programa foi tomado a partir da caixa de ferramentas do Matlab®.

Figura 41. Janela de inserção do tipo de *wavelet*.

Uma restrição foi realizada no programa em relação com o tamanho da mensagem e o tamanho do som para garantir que o arquivo de áudio no sofra modificações marcantes, isto é, o tamanho da mensagem deverá ser pelo menos 1000 vezes menor que o tamanho do som.

Quando se insere uma quantidade de dados maior aos permitidos, o programa interrompe para que se diminua o número de caracteres a inserir ou se utilize um arquivo de áudio maior, tal como mostra a Figura 42.

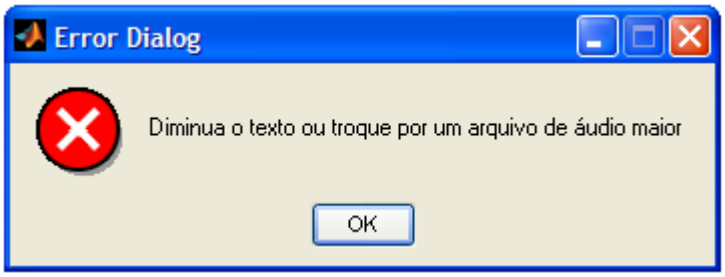


Figura 42. Janela de interrupção do programa para diminuir a quantidade de dados ou aumentar o tamanho do áudio.

Com um arquivo de áudio de 10 kB podem ser inseridos no máximo 9 caracteres, então a mensagem a ser inserida será a palavra **PROVA** como especifica a Figura 43.

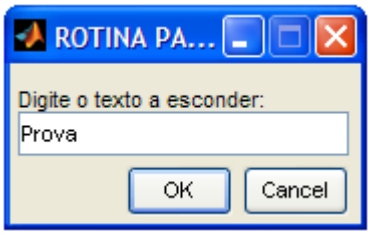


Figura 43. Janela com a mensagem a ser inserida.

Concluída a inserção da informação para a decomposição do áudio se disponibiliza uma janela que especifica o número de níveis que se dispõe para a inserção dos dados, Figura 44.

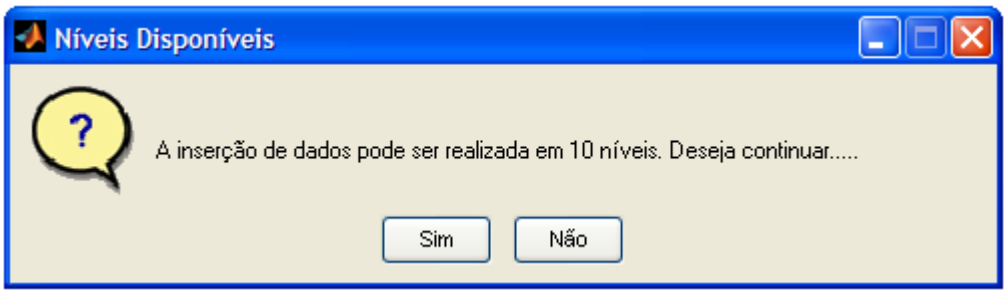


Figura 44. Janela com especificação do número de níveis para a inserção dos dados.

A mensagem pode ser subdividida em blocos e distribuída em diferentes níveis de acordo a critério do usuário. O tamanho da mensagem é de 5 caracteres que poderão ser divididos nos 10 níveis, então, podem ser ingressados 2 caracteres no primeiro nível, Figura 45.

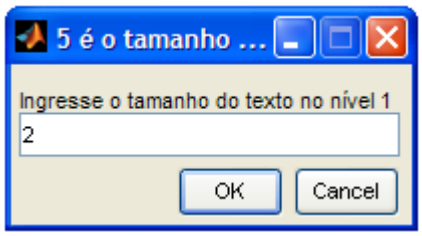


Figura 45. Janela de distribuição de dados - nível 1.

Assim como também nenhum dado pode ser inserido no nível 2 e nível 3, Figura 46 e Figura 47 respectivamente.

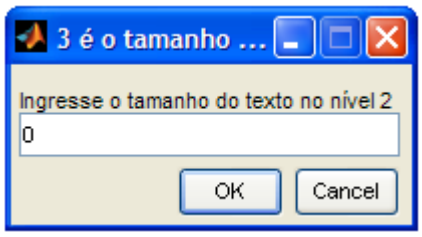


Figura 46. Janela de distribuição de dados - nível 2.

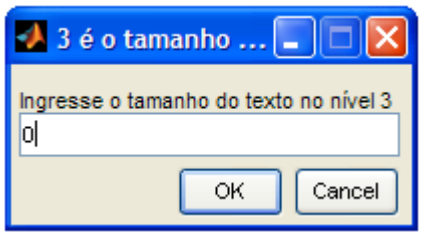


Figura 47. Janela de distribuição de dados - nível 3.

Um caractere inserido no nível 4, Figura 48.

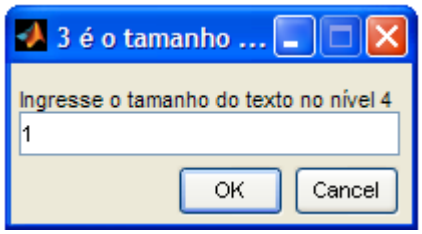


Figura 48. Janela de distribuição de dados – nível 4.

Nenhum caractere inserido no nível 5 e nível 6, Figura 49 e Figura 50.

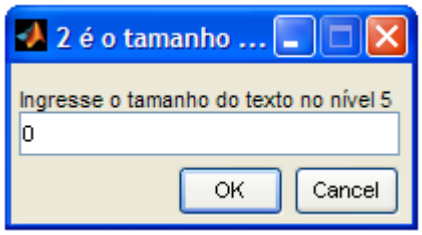


Figura 49. Janela de distribuição de dados – nível 5.

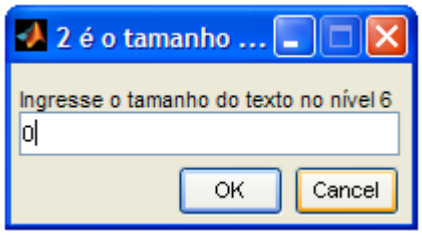


Figura 50. Janela de distribuição de dados – nível 6.

Um caractere inserido no nível 7, Figura 51.

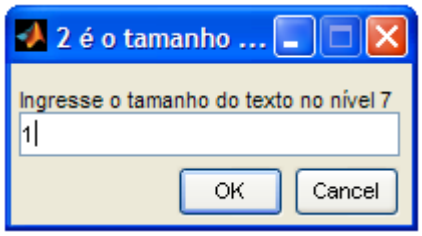


Figura 51. Janela de distribuição de dados – nível 7.

Nenhum caractere inserido no nível 8 e nível 9, Figura 52 e Figura 53.

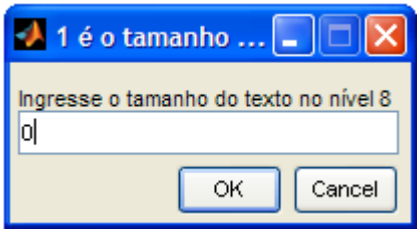


Figura 52. Janela de distribuição de dados – nível 8.

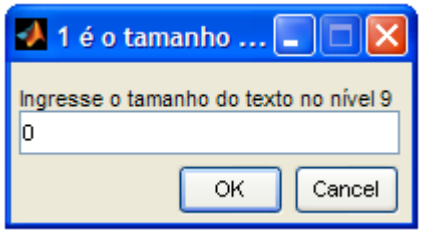


Figura 53. Janela de distribuição de dados – nível 9.

O último dado será inserido no nível 10, Figura 54.

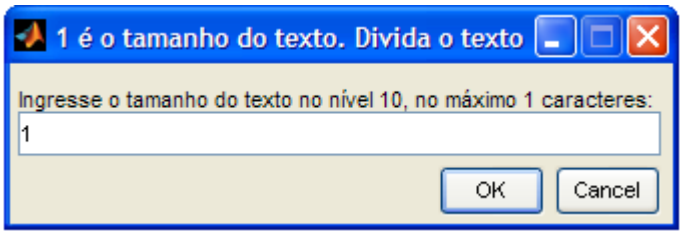


Figura 54. Janela de distribuição de dados – nível 10.

Uma vez distribuídos os dados nos níveis elegidos são solicitadas informações: senha(s) e lugar de inserção do texto (para cada nível onde foram inseridos dados) que formarão parte do que se conhece como chave-estego junto com a *wavelet* que foi selecionada.

As Figuras 55, 56, 57 e 58 mostram o ingresso da(s) senha(s) e lugar de inserção do texto para cada nível elegido (nível 1, 4, 7, 10).

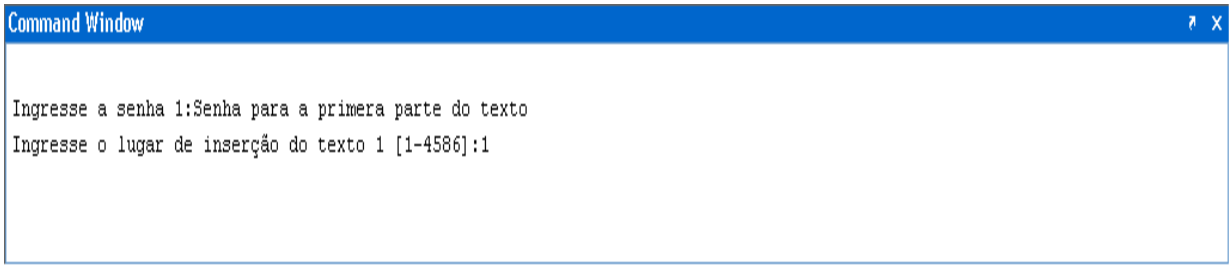


Figura 55. Janela de ingresso da primeira senha e lugar de inserção do texto para o nível 1.

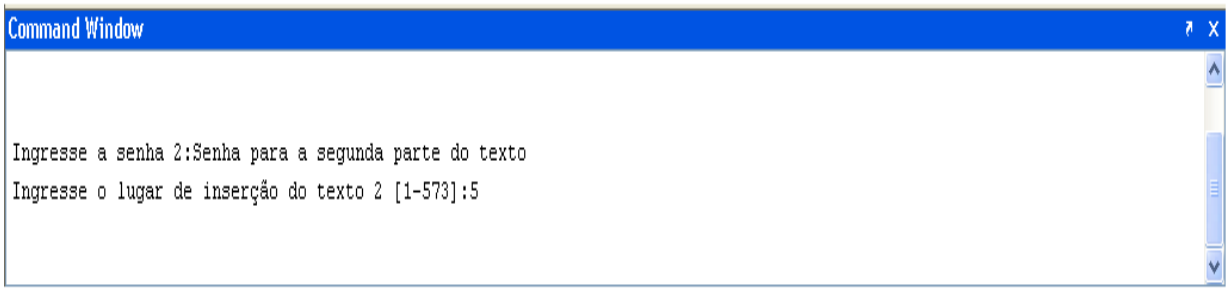


Figura 56. Janela de ingresso da segunda senha e lugar de inserção do texto para o nível 4.

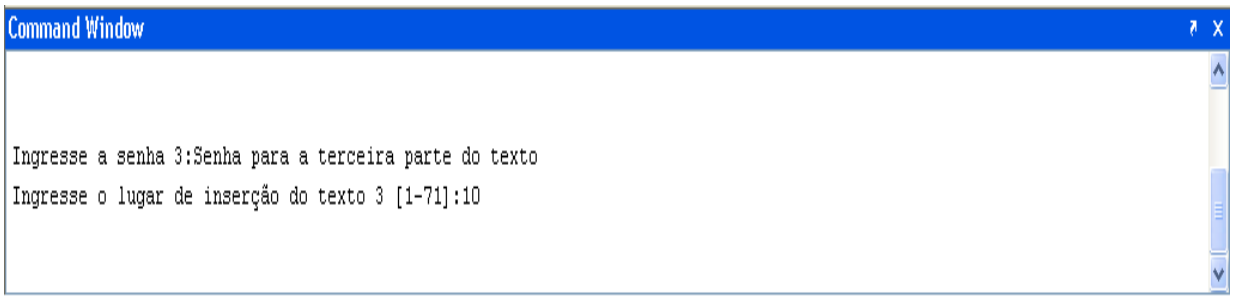


Figura 57. Janela de ingresso da terceira senha e lugar de inserção do texto para o nível 7.

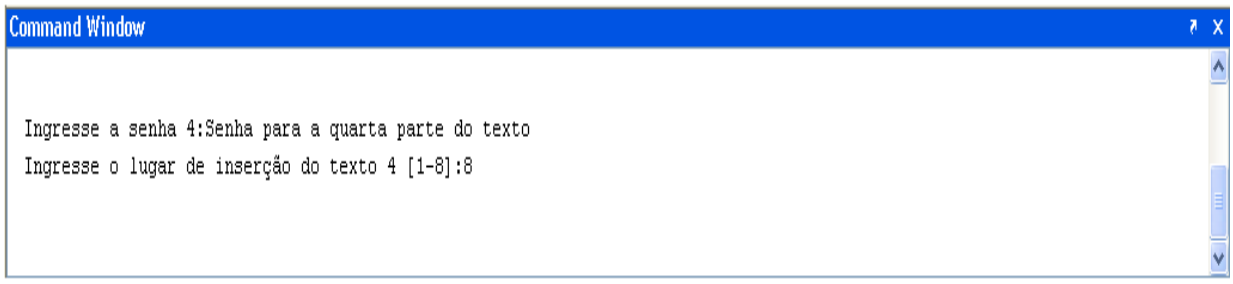


Figura 58. Janela de ingresso da terceira senha e lugar de inserção do texto para o nível 10.

Por último se realiza a reconstrução do arquivo de áudio e o som é gravado contendo a informação oculta com o nome especificado: **TESTE**, Figura 59.

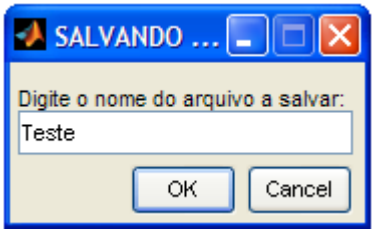


Figura 59. Janela de gravação do áudio contendo a informação oculta.

O arquivo de áudio se grava por *default* na pasta de trabalho do Matlab®.

5.4.2. Esquema Gráfico da Recuperação de Dados num Arquivo de Áudio

O primeiro passo para a recuperação da mensagem é ingressar o nome do som que contem os dados, Figura 60.

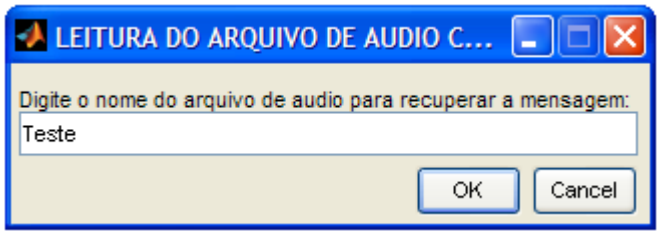


Figura 60. Janela para ingressar o nome do arquivo de áudio contendo dados.

Imediatamente são requeridas as informações ou chave-estego.

Os três primeiros dados a solicitar-se são:

- Tipo de *wavelet*.
- Tamanho da mensagem.
- E número total de níveis onde foram inseridas as mensagens.

Esta informação é verificada na Figura 61.

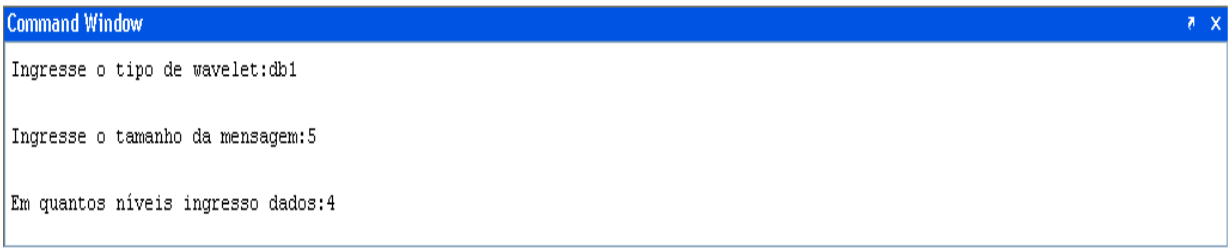


Figura 61. Janela de ingresso dos três primeiros dados da chave-estego.

As seguintes informações também formam o grupo do que se conhece como chave-estego:

- Senha(s).
- Especificação de cada nível onde foi inserida uma porção da mensagem.
- Tamanho de cada porção de mensagem.

- E lugar de inserção no texto.

Esses dados são especificados para cada nível onde foi inserida uma porção da mensagem, assim as Figuras 62, 63, 64, e 65 detalham a inserção dessas informações:

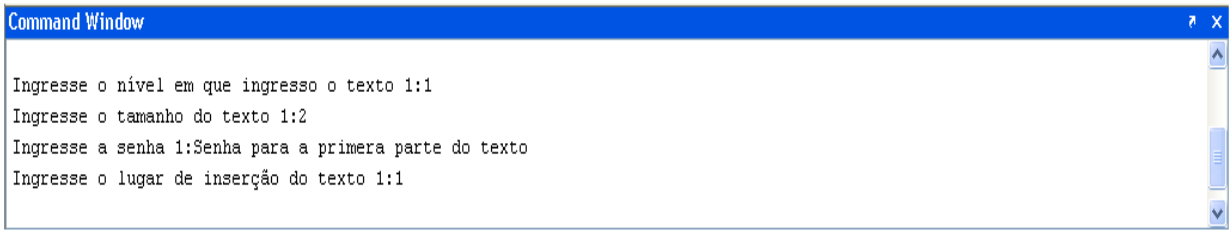


Figura 62. Janela de inserção das informações restantes que constituem a chave-estego – nível 1.

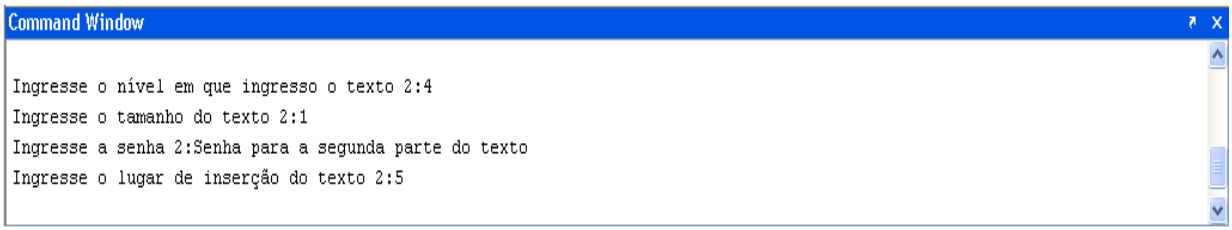


Figura 63. Janela de inserção das informações restantes que constituem a chave-estego – nível 4.

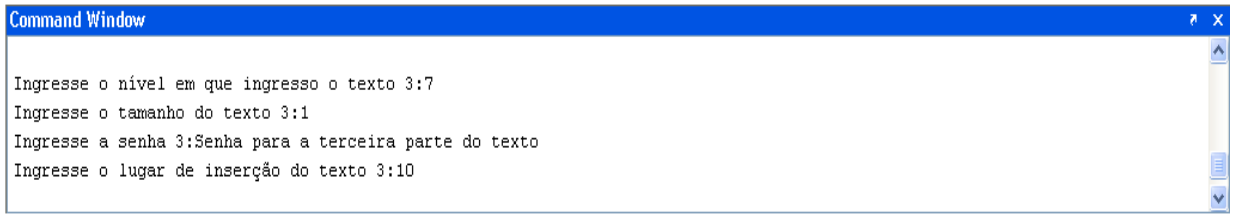


Figura 64. Janela de inserção das informações restantes que constituem a chave-estego – nível 7.

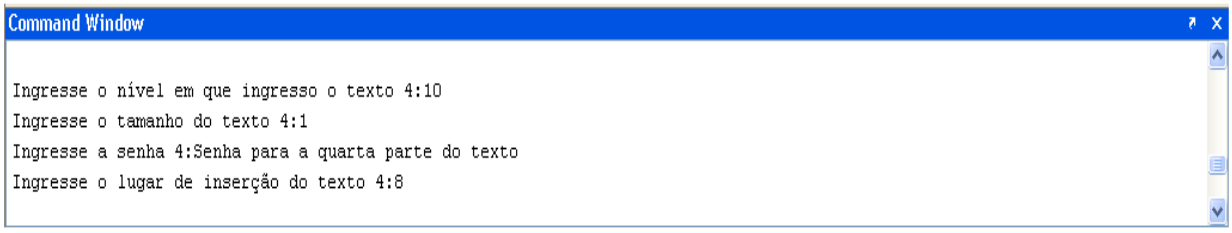


Figura 65. Janela de inserção das informações restantes que constituem a chave-estego – nível 10.

Uma vez ingressados todos os dados corretos a mensagem é recuperada, obtendo-se o texto inserido: **PROVA,** Figura 66.

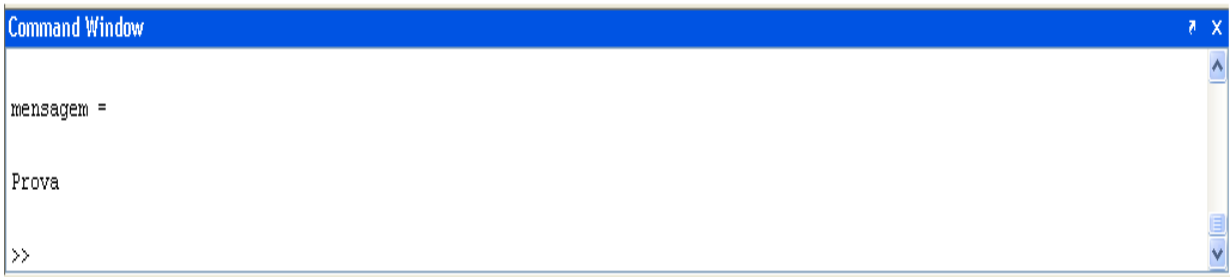


Figura 66. Janela de recuperação da mensagem.

Capítulo 6

CONCLUSÕES E SUGESTÕES

6.1. Conclusões

A esteganografia permite ocultar a presença de uma mensagem dentro de outros objetos, e à simples vista deste objeto, a informação introduzida é imperceptível. E é justamente esta característica que foi explorada durante o passar dos anos, até a época presente, onde se pode ocultar facilmente informação em arquivos digitais, mediante *softwares* que se encontram disponíveis na rede. Cada vez são em maior número os métodos criados e de livre circulação na Internet, o que constituiu um desafio à implementação do programa, de forma que constitua uma aplicação diferenciada do restante dos algoritmos que inserem dados.

Um dos maiores objetivos na implementação foi adequar convenientemente o método, tanto para a inserção como para a recuperação da informação. Foram realizados testes com áudios inserindo texto nas diferentes posições decimais dos coeficientes da matriz de decomposição *wavelet* e o melhor resultado foi obtido quando foram substituídos os dados nas três primeiras casas decimais, recuperando quase na sua totalidade a informação. Ao contrário, quando foram introduzidos dados nos bits menos significativos (LSB), como usual, a informação acabava-se deteriorando; isto aconteceu porque ao se utilizar os comandos próprios do Matlab® para a reconstrução do sinal, realiza-se uma aproximação nos valores dos coeficientes, fazendo com que as quantidades das últimas casas decimais se alterem (erros numéricos de arredondamentos) e a informação inserida se perde. Também foi verificado mediante ensaios que a inserção de texto nessas posições não modifica a essência do sinal desde que sejam tomadas as devidas precauções com a quantidade de dados que vai inserir e o respectivo tamanho do arquivo de som.

O fato de que a inserção de dados deva ser cautelosa, a recuperação da informação aceitável, os testes realizados com áudios e resultados obtidos satisfatórios, pode se concluir que a aplicação atende às expectativas, e torna o programa viável, podendo ser disponibilizado e empregado para “esconder” dados em arquivos de áudio com formato wav. Todo este esquema se conseguiu graças às propriedades que apresentam as *wavelets* na análise de sinais.

O Matlab®, plataforma empregada como ferramenta junto com o conjunto de *wavelets* padrão, permitiu a execução do programa, de forma que o áudio decomposto em níveis possibilita uma distribuição (espalhamento) dos dados a ser inseridos. A informação fica disseminada e de forma intuitiva, pode-se afirmar que, tanto as modificações marcantes no arquivo, como a detecção da informação, serão menos prováveis do que se os dados secretos estivessem concentrados.

Uma forte aliada da Esteganografia é a Criptografia, reforçando a segurança do sistema. A proposta de montar um sistema estego-cifrado em dois passos é bastante atrativa, e não é descrita de maneira direta na literatura. Os dados inseridos podem ser previamente cifrados mediante um criptosistema conhecido, o atacante deverá possuir a chave-estego para a obtenção dos dados e, se conseguisse obter a informação, possuiria outro desafio que é a decifragem da mensagem.

A implementação satisfaz duas propriedades importantes que caracterizam sistemas esteganográficos: Robustez e Imperceptibilidade. Pode-se dizer que o programa é suficientemente *robusto*, já que é capaz de recuperar a informação de forma quase inalterável, além de que a inserção dos dados é realizada nas casas decimais significativas do arquivo o que também proporciona *robustez*. Com as respectivas restrições tomadas o estego-objeto gerado não delata a presença de informação oculta, tornando-a *imperceptível*. Com estas propriedades satisfeitas (e outras que não foram mencionadas neste parágrafo), poderiam realizar-se breves alterações ao algoritmo e originar um sistema de inserção de marcas de água.

Não poderia deixar de se mencionar uma aplicação prática da esteganografia como é a Marca de Água. Pesquisas, trabalhos e aplicações são realizados abordando este tema, devido ao fato de que atualmente as marcas são largamente aproveitadas na indústria, áreas tecnológicas, e em diferentes áreas como um auxílio, para controle de cópias indevidas, proteção de direitos autorais, modificações de arquivos, pirataria, entre outras.

O estudo de estego-análise não é com o propósito de incentivar a remoção de informações valiosas que são inseridas em arquivos digitais com finalidade de proteção e segurança, mas, com a intenção de detectar fraquezas dos sistemas para futuras melhorias. Aplicando estego-

análise pode-se detectar arquivos - que se encontram livremente circulando pela rede, e que com um simples olhar parecem "inocentes". Através da observação e estudo detalhado de arquivos, a capacidade de descobrir e revelar informações ocultas aumenta.

Presume-se que na Internet encontram-se circulando arquivos que contêm informação maliciosa (sobretudo em imagens). Para detectar este tipo de arquivos – sobretudo as autoridades, devem estudar as técnicas e ferramentas de inserção e empregar táticas de estego-análise para a detecção destas informações indesejadas.

Neste trabalho foram abordados temas complementares, mas não menos importantes, que incluem: um breve estudo fisiológico do Sistema Auditivo Humano e sua influência na escuta; o que é um sinal de áudio e como é afetado por um ruído; tudo com o objetivo de obter melhores resultados no desenvolvimento da dissertação.

6.2. Sugestões para Trabalhos Futuros

- A estego-análise é uma proposta oportuna para ser realizada. Mediante a estego-análise não apenas se é capaz de verificar as fraquezas que possui o sistema, mas ao mesmo tempo se pode avaliar a robustez do método. A estego-análise é a arte de tentar detectar a existência (ou ausência) de informação oculta. Baseia-se em quebrar o princípio da esteganografia, que ninguém desconfie que se está transmitindo informação. Além da suspeita e detecção, a modificação da mensagem é considerada como um ataque bem sucedido. Assim, pode-se sugerir realizar um estego-análise em arquivos de áudio onde foram inseridos dados empregando o algoritmo desenvolvido, para detectar possíveis fraquezas que possa apresentar o sistema.

- As Marcas de Água e a Esteganografia baseiam-se no mesmo princípio (ocultar informação), e a principal diferença consiste na função que cumpre cada uma delas; enquanto a marca é empregada como um meio de proteção para arquivos digitais, o foco da esteganografia consiste em tornar "invisível" uma mensagem. Obviamente o sistema desenvolvido poderia ser adaptado para a inserção de marcas de água em sinais de áudio, tomando em consideração as propriedades e características exigidas para as marcas.

- Para o estudo e processamento de sinais existem diversos tipos de transformadas e poderia citar-se algumas delas como exemplo: a transformada discreta de Fourier, a transformada discreta do Cosseno, etc. Outros métodos que empregam a inserção de dados em arquivos de áudio são: transformação do espectro, geração de ecos, inserção de ruído, mascaramento de tons, etc. Qualquer destas técnicas, sem limitar a escolha, pode ser adotada para conceber um novo algoritmo que encubra informação. Além da implementação, seria interessante realizar uma comparação entre os processos (se possível), determinando vantagens e desvantagens que apresenta cada procedimento.

- Considerando que existem diversos formatos de áudio digital, o algoritmo desenvolvido poderia ser ajustado para trabalhar com outros tipos de áudios, por exemplo, áudios com formato AU (Unix Audio), etc.

ANEXOS

ANEXO A – O SISTEMA AUDITIVO HUMANO

A1.1. Características do Sistema Auditivo Humano

A intensidade sonora que um ser humano pode perceber é uma magnitude subjetiva, muitos fatores podem influenciar na escuta de cada pessoa, tanto fatores internos (estado de ânimo) ou externos (ambiente) que afetam ao indivíduo, e é dessa forma, que cada pessoa interpreta de maneira distinta o que considera uma melodia, um ruído [51] ou uma deturpação em um som.

O ouvido humano se divide em três partes, que serão representadas na Figura 67:

- O ouvido externo.
- O ouvido médio.
- O ouvido interno.

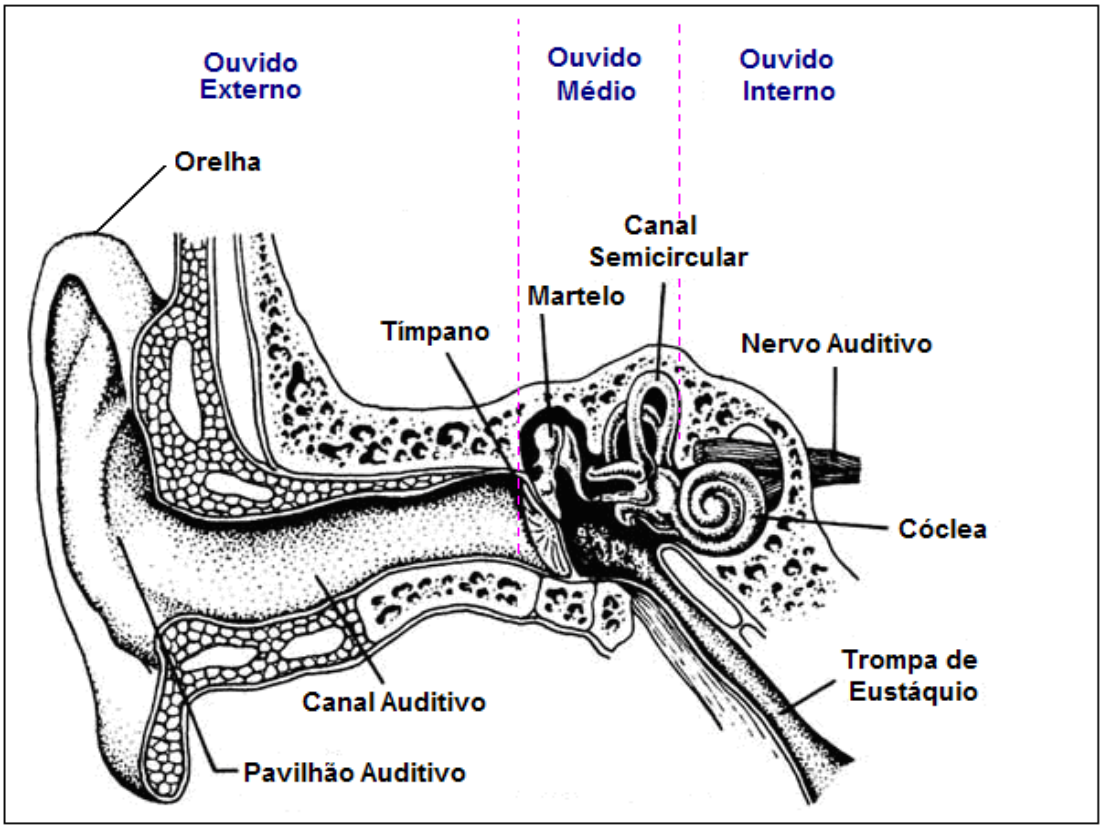


Figura 67. Esquema do ouvido humano.

- **O ouvido externo.**

Encontra-se constituído pelo *pavilhão auditivo*, que permite localizar as fontes sonoras, o *canal auditivo* (de aproximadamente 3 centímetros de comprimento) e o *tímpano* (membrana fina e elástica) que divide o ouvido externo do ouvido médio. A principal função do *ouvido externo* é concentrar as ondas sonoras na cavidade do ouvido e direcioná-las ao interior onde são transformadas em vibrações mecânicas (no ouvido médio). O *ouvido externo* apresenta algumas funções acústicas (cf. Figura 68) e funciona de forma similar a alguns instrumentos musicais (como o órgão com tubo fechado), as ondas sonoras que percorrem o canal auditivo sofrem um aumento de 5 a 20 dB, isso, devido à ressonância existente nessa região (3 - 5 kHz) o que permite um aumento de sensibilidade nos sons nessa zona de frequência. O ouvido também atua como um amplificador direcional, o que permite detectar as fontes sonoras em três dimensões: de frente para trás, de cima para baixo e de um lado para outro.

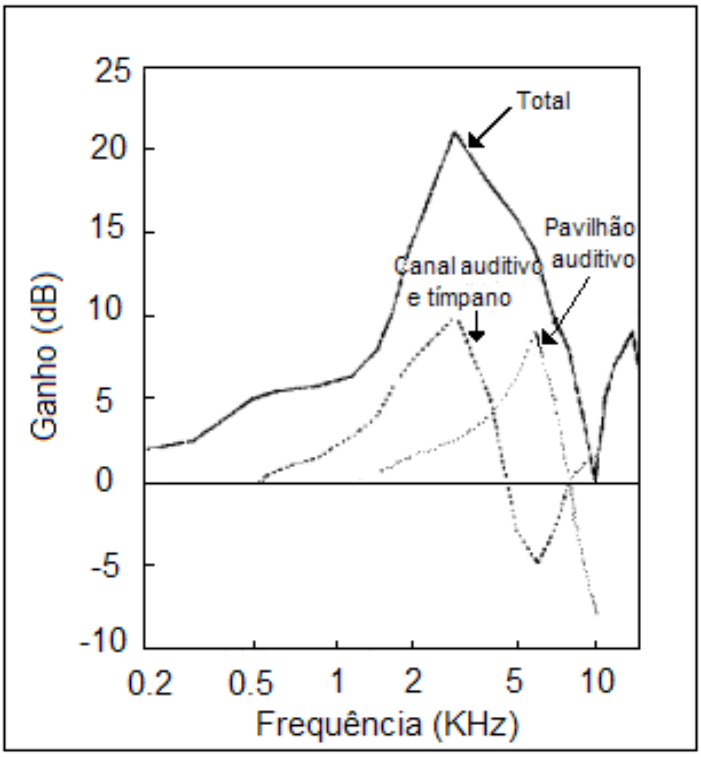


Figura 68. Efeitos acústicos do ouvido externo.

- **Ouvido médio.**

Encontra-se após o tímpano, e está constituído por três ossículos: *martelo, bigorna* e o *estribo* e pela *trompa de Eustáquio* que se conecta com a garganta e mantem a pressão atmosférica (ver Figura 69). A função do ouvido médio é amplificar as ondas sonoras recebidas para serem transmitidas ao ouvido interno, devido a uma perda que acontece quando as ondas sonoras se chocam com a janela oval. Os ossículos, além de transmitir os movimentos do tímpano sem perda de energia, resguardam o sistema auditivo de sons muito altos, ativando um mecanismo de defesa (stapedius reflex) que reduz a transferência do som protegendo o ouvido interno. Sons com níveis de pressão sonora acima de 75 dB contraem os músculos que se juntam ao tímpano e ao estribo, diminuindo a transmissão de energia de 12 a 14 dB de atenuação (para sons abaixo de 1 kHz). A proteção leva um tempo para ativar-se fazendo com que não se tenha resposta a sons impulsivos.

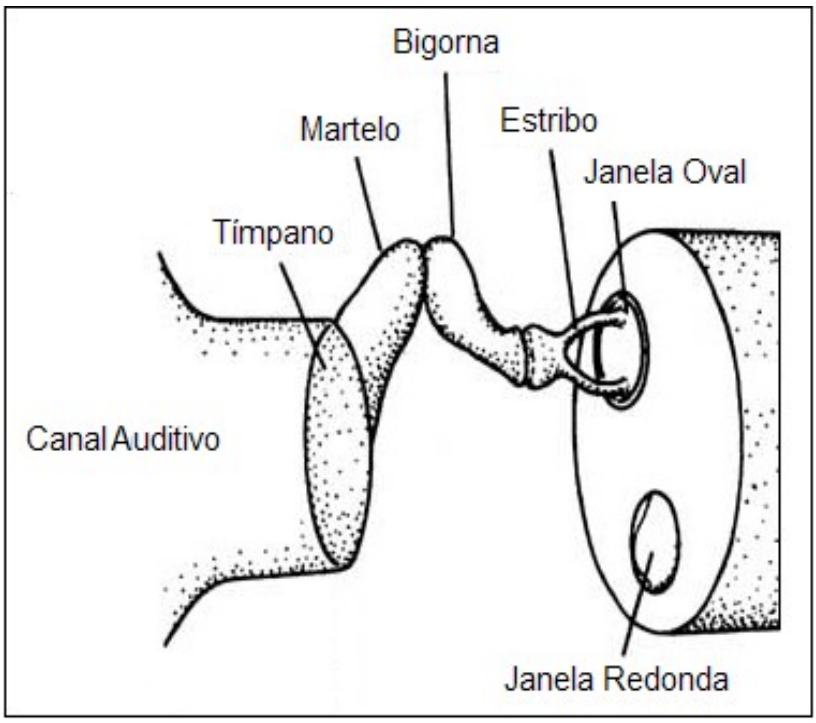


Figura 69. Conjunto de ossículos (martelo, bigorna e estribo).

- **Ouvido interno.**

Localizado dentro do crânio, é considerado a parte mais complexa do ouvido. Se encontra conformado pela *cóclea* (forma de caracol de aproximadamente três voltas), os *canais semicirculares* (regulam e mantêm o equilíbrio do corpo) e o *nervo auditivo*. Através da janela oval são recebidas as vibrações emitidas pelo ouvido médio. No *ouvido interno* se realiza a conversão das ondas sonoras em sinais elétricos para serem transmitidas ao cérebro. A membrana basilar permite a percepção do som, assim, a parte mais fina e estreita da membrana possui melhor resposta a frequências altas (agudas), enquanto a parte grossa e mais larga possui melhor resposta a frequências baixas (graves), como ilustrado na Figura 70.

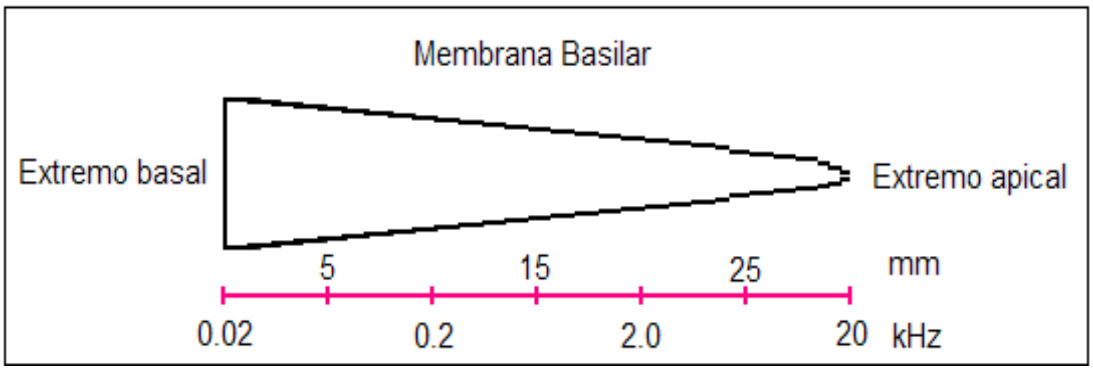


Figura 70. Membrana basilar estendida.

Um fenômeno que acontece na membrana basilar como resposta dos diferentes tipos de frequência é o deslocamento da onda sonora, como se ilustra na Figura 71. Sons de alta frequência percorrem uma distância pequena, sons com frequência média viajam uma distância maior, enquanto sons de frequência baixa alcançam o final da membrana antes de seu desvanecimento.

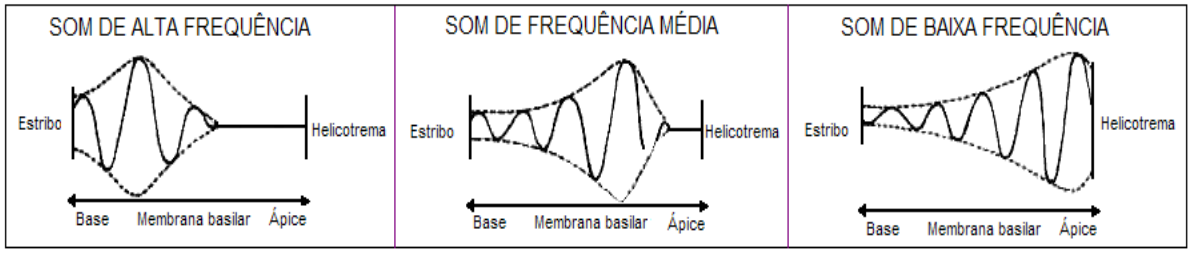


Figura 71. Deslocamento da onda sonora na membrana basilar.

Quando acontece o fenômeno de ressonância na membrana basilar, os cílios, que constituem um conjunto de pelos sensíveis, se acionam gerando sinais elétricos. Os cílios são polarizados e, dependendo da quantidade de impulsos elétricos, gera-se a hiperpolarização (envio mínimo de impulsos) ou despolarização (envio máximo de impulsos) das células auditivas, como mostra a Figura 72.

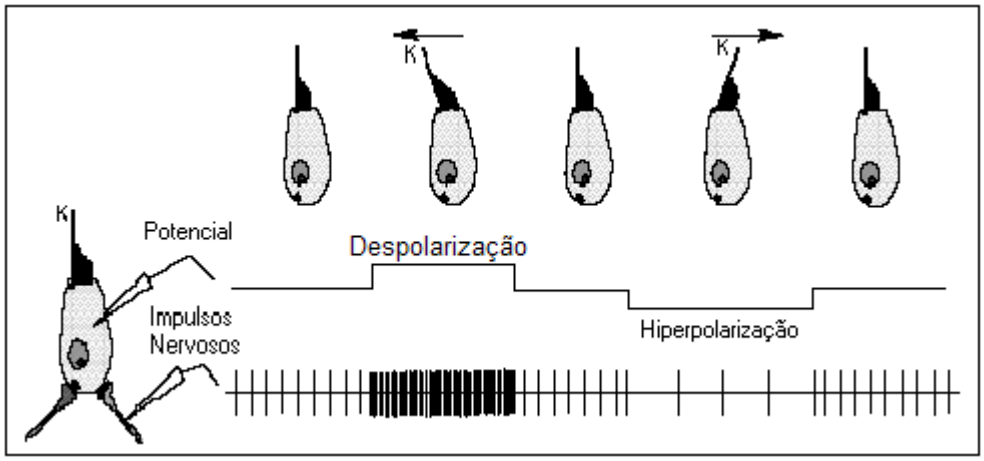


Figura 72. Polarização das células auditivas.

O ouvido interno funciona como um analisador de sons como se mostra na Figura 73. Na escuta, o ouvido capta sons com diferentes frequências que excitam a membrana basilar transformando-os em sinais nervosos, e o cérebro é capaz de reconhecer altura, intensidade, frequência e qualidade do som, pelo qual é possível reconhecer uma pessoa, instrumento ou qualquer ruído emitido. O sistema auditivo também é modelado como uma banda de filtros [52].

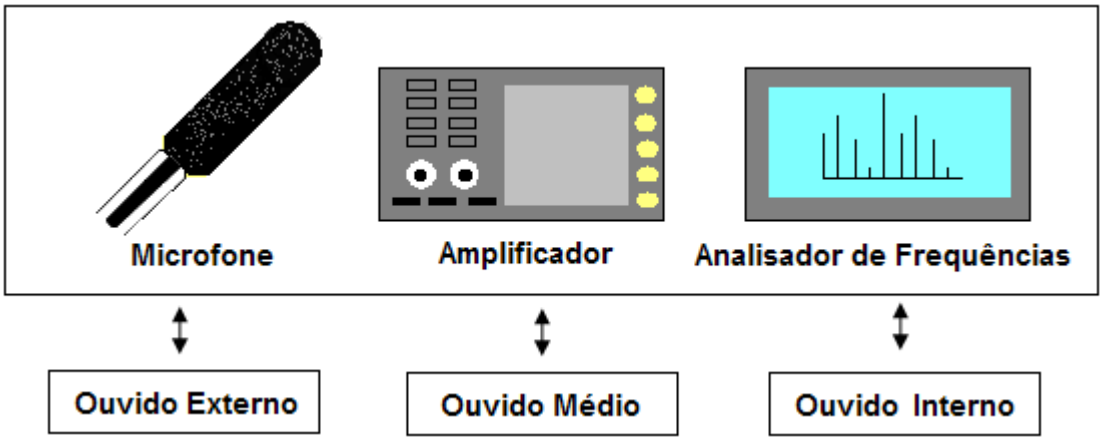


Figura 73. Relação entre o ouvido humano e um analisador de frequência.

A faixa de frequência audível do ouvido humano varia de 20 Hz a 20 kHz, onde duas bandas críticas limitam a faixa da frequência, gerando um limite superior e inferior. Intensidades sonoras menores e maiores que os limiares não podem ser captados pelo ouvido humano. Tons muito baixos (infra-som) não podem ser escutados e tons muito altos (ultra-som) causam dor no ouvido, ou seja, uma faixa dinâmica entre 0 dB e 120 dB é tolerável na escuta. Uma escala logarítmica é apresentada na Figura 74, mostrando a variação de pressão de um som na qual o ouvido tem reação.

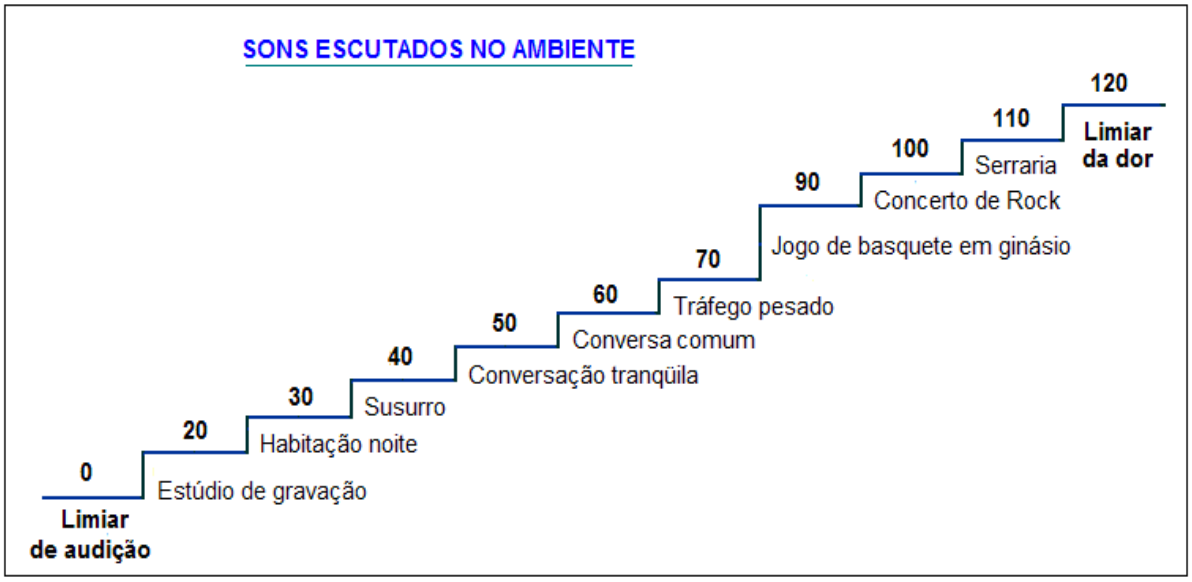


Figura 74. Escala logarítmica de pressão de um som em Decibel (dB) – pessoa jovem.

A membrana basilar se comporta como um analisador de frequências, em determinados pontos são atribuídos zonas de resposta. Por exemplo, frequências baixas excitam a parte interna da membrana e frequências altas estimulam a parte externa, isto pode ser apreciado na Figura 75.

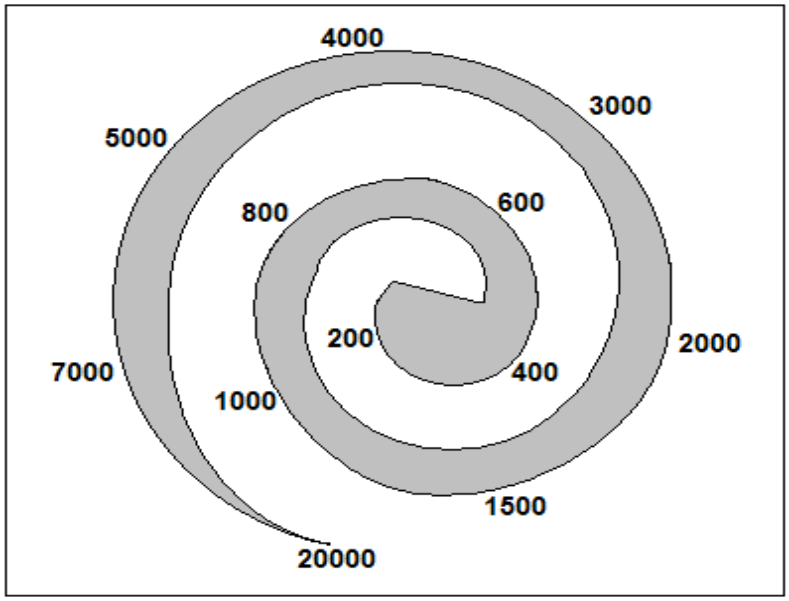


Figura 75. Zona de resposta de frequência na membrana basilar.

A sensibilidade auditiva determinada por Fletcher e Munson em 1933 através de medições da sensibilidade do ouvido humano, é definida mediante curvas características de igual nível de sensação para tons puros que determinam a variação da sensibilidade do ouvido na audição (cf. Figura 76). Devido a estrutura do ouvido ser semelhante a um tubo (com um lado aberto e outro fechado) se geram picos de sensibilidade acima de 1 kHz. O primeiro pico pode ser observado em torno de 3.4 kHz e o outro um pouco menor em torno de 13 kHz, estes efeitos de ressonância provocam o aumento de sensibilidade na escuta. O ouvido se mostra mais sensível para frequências que se encontram entre 3500 e 4000 Hz (próximo da primeira zona de ressonância no ouvido externo). Frequências baixas (sons graves) possuem menor sensibilidade e frequências altas (sons agudos) apresentam maior sensibilidade.

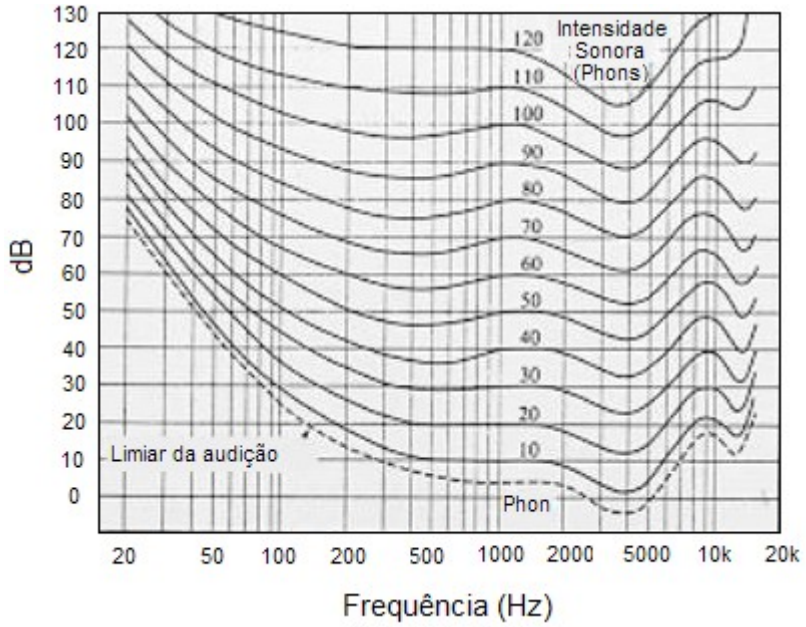


Figura 76. Curvas de Fletcher e Munson.

O ouvido humano possui um comportamento não linear, frequências entre 1000 e 5000 Hz possuem melhor resposta, enquanto que frequências de 20 a 1000 Hz e de 5000 Hz adiante são atenuadas. Sons agudos têm maior excitabilidade no ouvido que sons graves [53]. A sensibilidade do ouvido humano varia de indivíduo para indivíduo dependendo da frequência, do ambiente e do tipo de som, pelo qual testes e conclusões não poderiam basear-se apenas em experiências de escuta para determinar mudanças em arquivos de áudio. Por isso, é necessário basear-se em critérios quantitativos, algoritmos que permitam medir a variação do som depois de ter sofrido algum tipo de processamento ou inserida alguma informação no seu interior.

Um fenômeno acontece quando se juntam dois tons puros: um "mascara" o outro [54], ou seja, um som impede que outro seja escutado. Por exemplo, quando se escutam dois sons com as seguintes características: um som forte e grave com um som débil e agudo se tem a impressão que se estivesse escutando somente aquele de maior intensidade e de menor frequência, ou seja, o som forte e grave, (cf. Figura 77). Esse fenômeno é mais perceptível quando os dois sons são de frequências próximas. O aproveitamento desse fenômeno peculiar pode ser de grande interesse em ocultação de dados, onde o sinal cobertor resguarda o sinal que se deseja ocultar, ficando imperceptível ao ouvido.

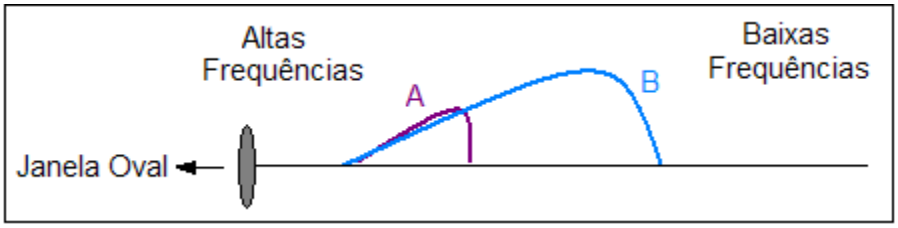


Figura 77. Resposta da membrana basilar, mascaramento de dois sons puros.

A1.2. A Função do Cérebro na Escuta

O cérebro exerce a função de distinção das características de um som. O cérebro é o órgão que armazena a informação de cada tom e permite comparar um ruído ou som escutado com padrões previamente gravados em nossa memória. Quando uma informação é enviada pelo ar, todo um processo é gerado: o som é captado pelo ouvido e percorre o conduto auditivo, a amplificação é realizada e a membrana basilar excitada, transmitindo impulsos elétricos ao cérebro, onde, os dados recebidos são comparados com informação armazenada na memória e interpretados. Se os dados fornecidos formam parte de um conjunto de informação da memória, será associado a tal grupo e reconhecido pelo ouvinte [52], se sons ou padrões forem desconhecidos e não tivesse nenhuma interligação, o cérebro pode guardá-lo como um novo padrão ou simplesmente esquecê-lo.

No que se tem conhecimento, existem três etapas realizadas pelo cérebro no processamento da informação recebida mediante a escuta:

1. Determinação do local de onde provém o som.
2. Análise das características do sinal de áudio.
3. Relação do som escutado com outros sons.

A1.3. As Ondas Sonoras

Depois de haver realizado uma breve introdução pelo Sistema Auditivo Humano, é importante especificar os tipos de ondas ou vibrações que interatuam na escuta. Conforme ao efeito que geram sobre o ouvido (forma subjetiva), as ondas sonoras podem ser classificadas em:

- Som.
- Ruído.

Mas, essa classificação é muito pessoal pelo fato de ser subjetiva, dependera de cada ouvinte, assim, por exemplo, certo tipo de música (rock) para uma pessoa pode ser considerado um som, enquanto para outra poderia ser considerado um ruído.

A1.3.1. O Som

É uma vibração que se propaga num meio elástico geralmente pelo ar e que o ouvido humano pode detectar.

Características do som.

A característica do som é representada por três variáveis físicas:

- **Intensidade.**

  Depende da quantidade de energia do som, amplitude das vibrações. A intensidade permite distinguir entre um som *forte* ou *fraco*. É representada mediante uma escala logarítmica, já que a percepção é baseada na razão entre valores de amplitude e pode ser expressa em decibéis (veja equação 36).

- **Altura.**

  Depende da sua frequência, assim, sons com maior frequência são considerados agudos, e sons com baixa frequência serão graves, ou seja, sons altos e baixos respectivamente. Esta característica permite distinguir entre um som *grave* ou *agudo.* A frequência se expressa em Hertz (ciclos / segundo).

- **Timbre.**

  Permite a distinção da fonte sonora, por exemplo: vozes, instrumentos musicais ou sons com a mesma intensidade e altura. Se dois sons possuírem a mesma intensidade, frequência e duração eles poderiam

ser identificados através do timbre. O timbre encontra-se relacionado com as componentes senoidais; além da frequência fundamental, existem frequências secundárias (ou harmônicos) que dão origem a diferentes timbres.

A resposta de audição de nosso mecanismo é proporcional ao logaritmo da razão entre a frequência e também entre amplitude. Na escuta se empregam escalas logarítmicas devido a que o intervalo entre o valor mais baixo que se pode escutar e o limiar da dor, existe uma grande distância, desse modo, são definidos os níveis sonoros em decibel (dB) das três grandezas acústicas: intensidade, pressão e potência (*I*, *P*, *W*) do som/ruído:

- **Intensidade:** A intensidade é o fluxo de energia por unidade de área, e se mede em *watts* por *metro quadrado* ($watts/m^2$). O nível de intensidade sonora pode ser deduzido de acordo a seguinte equação:

$$L_I = 10\log\frac{I}{I_0} \tag{36}$$

em que:

L_I é o nível de intensidade

I é a intensidade do som, e

I_0 é o valor de referência da intensidade sonora: $I_0 = 10^{-12}\, watts/m^2$

- **Potência:** A potência é a energia acústica total emitida por uma fonte por unidade de tempo, medida em *watts* (1 *watt* = 1 Joule/segundo). O nível de potência sonora é deduzido de acordo a seguinte expressão:

$$L_W = 10\log\frac{W}{W_0} \tag{37}$$

em que:

L_W é o nível de potência

W é a potência, e

W_0 é o valor de referência da potência sonora: $W_0 = 10^{-12}\, watts$

- **Pressão:** A pressão sonora é medida pela pressão que as ondas sonoras exercem sobre uma superfície, e é medida em *Newton* por *metro quadrado* ($Newton/m^2$), também é denominada *Pascal* (*Pa*).

$$L_P = 20\log\frac{P}{P_0} \tag{38}$$

em que:

L_P é o nível de pressão

P é a amplitude da pressão, e

P_0 é o valor de referência da pressão acústica: $P_0 = 2x10^{-5}\,Newtons/m^2$

I_0, W_0, P_0, são níveis de referência escolhidos como limiar da audibilidade do ouvido humano para dar origem a uma escala logarítmica.

Para que se produza o som se requer que intervenham os seguintes elementos:

- O emissor ou fonte sonora.
- O meio propagador.
- Receptor.

O *emissor ou fonte sonora* é qualquer aparelho capaz de criar ondas sonoras como, por exemplo, um motor, um violino, uma flauta, etc.

O *meio propagador* permite a propagação das ondas, o meio pode ser sólido, líquido ou gasoso; neste caso o meio é o ar.

O *receptor* é o meio que recebe a onda.

A1.3.2. O Ruído

Definir o que se considera um ruído envolve vários aspectos tanto fisiológicos, pessoais e externos e é por esses motivos que o ruído é considerado como uma apreciação subjetiva, então o que se poderia considerar um ruído?

Sob um enfoque *fisiológico*, o ruído será meramente considerado como um som não desejado, desagradável ao ouvido.

Desde o ponto de vista *físico*, o ruído pode ser definido como um sinal não suave. Comporta-se como um fenômeno acústico não periódico e seus harmônicos não são claramente definidos.

Mediante uma *Análise Esteganográfica*, um ruído é mais que uma sensação desagradável pode ser uma pequena perturbação que danifica a qualidade do som e pode delatar a presença de alguma alteração em um arquivo de áudio.

A1.4. Como Influência o Ruído na Perda da Audição?

O ruído auditivo pode ser definido como um sinal que perturba o ouvido, causando danos que podem ser irreversíveis, podendo ocasionar até perda da audição dependendo do tipo de ruído, intensidade e o tempo a que foi submetido o ouvido a dita perturbação. Ruídos muito fortes tais como aqueles de uma explosão, afetam consideravelmente na escuta, por tratarem-se de sons de grande intensidade e que se produz a uma velocidade maior que a reação do sistema de resguardo do ouvido contra ruídos fortes [51], pelo que o ouvido não é capaz de reagir e ativar o sistema de proteção que ocorre depois de 150 milésimos de segundos. A Tabela 14 apresenta diferentes ruídos que expostos ao ouvido humano influenciam a perda de audição [55].

Tabela 14: Ruídos que afetam ao ouvido humano em Decibel (dB).

RUIDO	dB
▪ Exposição a um ruído por um tempo prolongado de 8 horas.	85
▪ Ferramentas que geram ruído.	100
▪ Escuta de auriculares estereofônicos.	110
▪ Concerto de *rock*.	120
▪ Disparo de uma arma de fogo.	140-170

ANEXO B – O ÁUDIO DIGITAL

B1.1. Arquivo de Áudio Digital

No anexo A abordaram-se temas de sinais que se propagam no meio (o ar) tais como o som e o ruído, os quais atingem o ouvido humano, ocasionando o que se conhece como escuta; como parte complementar deste trabalho será realizada uma breve introdução sobre arquivos de áudio digital com formato wav, já que ele é considerado o estego-objeto no qual são ocultados os dados.

B1.2. O Áudio Digital

O comportamento de uma melodia, áudio, voz, ruído ou qualquer tipo de som escutado é analógico; para poder representar digitalmente tais sons faz-se necessário realizar uma transformação do sinal: de análogo a digital (conversão A/D). O processo de transformação é realizado mediante a amostragem – o sinal passa de contínuo a descontínuo ou discreto no tempo. Do arquivo de áudio são armazenadas amostras, onde cada amostra representa um número que será aproximado (quantizado) de acordo com a escala elegida, e posteriormente transformada em código binário [56], este processo é recíproco um sinal análogo pode ser convertido em digital e vice-versa. *Então um arquivo de áudio digital é a representação binária de um som*.

B1.3. Vantagens e Desvantagens que apresenta um Sinal de Áudio Digital

- **Vantagens.**

  Algumas das vantagens que apresentam os sinais de áudio digital são:

  - Menor tamanho.
  - Menor custo.
  - Eliminação de ruídos.
  - Facilidade para armazenar, transportar e reproduzir.
  - Facilidade para editar.

- **Desvantagens.**

    - Dependendo do formato de áudio pode-se perder a qualidade do som.
    - Precisa-se maior proteção e segurança contra pirataria.
    - Controle de cópias.

B1.4. Tipos de Formato de Áudio Digital

Existem diversos tipos de formato de áudio digital, alguns deles são:

- **WAV (*Windows Audio Visual*)**. Arquivo de áudio do Windows.

- **AIFF (*Audio Interchange File Format*)**. Formato de áudio empregado nos computadores Apple.

- **AU (*Unix Audio*)**. Empregado para criar arquivos de som em computadores Unix.

- **MP3 (*MPEG Audio Layer 3*)**. Arquivo de áudio compactado usando modelos perceptuais.

- **WMA (*Windows Media Audio*)**. Desenvolvido pela Microsoft, empregado para distribuir música gravada pela Internet.

B1.5. Arquivos de Áudio com Formato Wav

A palavra Wav se deriva de *Wave* (onda), e é um formato de arquivo de áudio da Microsoft e IBM para armazenamento de áudio em computadores, compatível com Windows e Macintoch. O formato wav é um método de armazenamento de áudio não comprimido, isto é, sem perdas. Para a reprodução dos arquivos com formato wav não é necessário um programa adicional.

O emprego deste tipo de áudio na prática e nos testes realizados foi considerado viável devido ao fato de que ele mantem a qualidade máxima do áudio e é de fácil manipulação, entretanto, existe a desvantagem de requerer grande espaço de armazenagem.

Um arquivo de áudio wav pode conter áudio compactado, apesar de que os formatos mais comuns contêm áudio em formato de modulação de pulsos PCM (*Pulse Code Modulation*).

Existem algumas limitações neste tipo de arquivo e uma delas se refere ao tamanho de gravação, podem ser armazenados arquivos com tamanho máximo de até 4 Gigabytes, e com certos programas, esse tamanho pode ser ainda menor.

B1.6. Formato de Modulação de Pulsos: PCM (*Pulse Code Modulation*)

PCM é um sistema binário; que é largamente empregado em sistemas de transmissão digital. A longitude dos pulsos e a amplitude são valores fixos; a presença ou ausência de pulsos representa uma condição lógica de um ou zero (1 ou 0) respectivamente, tal como ilustrado na Figura 78.

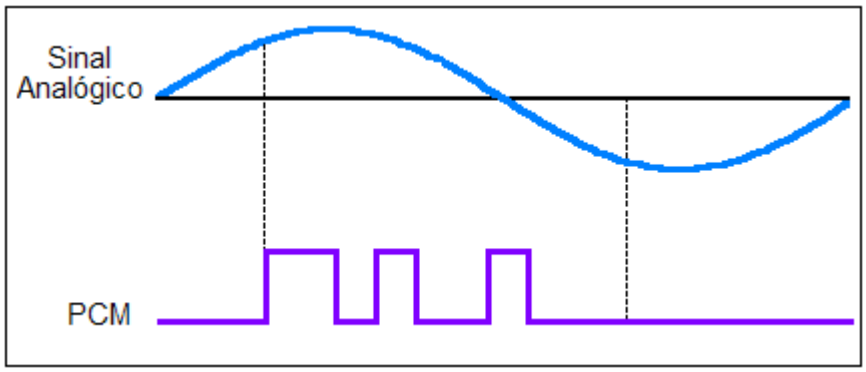


Figura 78. Representação da modulação de pulsos por código.

B1.6.1. Taxa de Amostragem

A taxa de amostragem é a quantidade de amostras colhidas ou reproduzidas por unidade de tempo, ou seja, o número de vezes por segundo que um sinal é medido, e essa medida se expressa em Hertz (Hz). Quando se emprega uma maior taxa de amostragem, esperar-se-ia que a representação do sinal seria melhor.

O teorema de Nyquist, entretanto, estabelece a mínima quantidade de amostras f_s que pode ser usada para um sistema de conversão analógico-digital, em particular, para PCM. A taxa de amostragem deve ser pelo menos duas vezes a maior frequência do sinal analógico f_a [57], então a mínima razão de amostragem de Nyquist corresponde ao lado direito da desigualdade:

$$f_s \geq 2f_a, \tag{39}$$

em que:

f_s = é a mínima taxa de amostras colhidas (Hertz).

f_a = frequência mais alta presente no sinal (Hertz).

B1.7. Avaliação da Qualidade do Áudio

Junto com a diminuição de tamanho dos arquivos de áudio e a expectativa de ganhar cada vez maior espaço no armazenamento, pode-se perder qualidade no áudio. Com base nessa e várias outras perspectivas, a União Internacional de Telecomunicações - ITU (*International Telecommunication Union*) estabeleceu um padrão para a medição da qualidade do áudio.
Nos últimos anos, vários testes para medição da qualidade do áudio foram desenvolvidos [58], mas nenhum deles foi padronizado. Com a finalidade de efetuar medições objetivas da percepção da qualidade do áudio, estabeleceu-se a recomendação ITU-R BS1387-1 [59], sendo hoje em dia considerada um padrão.

Em 1994 a ITU realizou um chamado para apresentação de propostas que permitam realizar a avaliação e percepção da qualidade do áudio; vários métodos foram apresentados, e seis propostas foram recebidas, mas nenhuma abarcava todos os requerimentos. Assim, foram incorporadas as melhores idéias numa única proposta que abrange todas as necessidades.

Diferentes tipos de medições podem ser realizados no áudio: medição objetiva seja em tempo real (demanda maior poder computacional e maior precisão) ou tempo não real; e medição subjetiva.
Depois de ensaios e estudos realizados, as conclusões indicaram que a avaliação independente (seja objetiva ou subjetiva) não pode ser realizada individualmente. Deve existir uma relação mútua entre ambas, isto é, resultados objetivos devem ser validados através de testes subjetivos, e a análise subjetiva - a escuta, deve ser reforçada mediante testes ou provas objetivas.
Para testes de escuta – baseados no Sistema Auditivo Humano (HAS) – foram geradas bases de dados de áudios e avaliados através da recomendação ITU-R BS 1116 a qual estabelece uma escala de 5 pontos. Nela é especificado se certo áudio possui alguma distorção ou perturbação, sendo atribuído o valor máximo (5) para mudanças imperceptíveis, até chegar ao valor mínimo da escala (1) que corresponde a alterações muito incômodas.

ANEXO C – TABELAS DE TESTES COM FATOR DE 600

C1.1. Tabelas com Fator de 600

Além do fator condicional de 1000, foram realizadas provas de áudios com um fator condicional de 600, no qual se constatou que com esse fator (600), se pode ingressar uma maior quantidade de dados, mas a porcentagem de variação também aumenta dependendo da disposição dos dados nos diferentes níveis. Por exemplo, na Tabela 15 o número máximo de dados a introduzir em um som de 10 kB são 15 caracteres, enquanto na Tabela 6, com um som do mesmo tamanho, o número máximo de caracteres a inserir foram 9.

Tabela 15: Variação percentual para um arquivo de som de 10 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5	5										-	-	-	0.1089799
5			5								-	-	-	0.4359198
5					5						-	-	-	1.743679
5							5				-	-	-	5.579773
10		10									-	-	-	0.4359198
10						10					-	-	-	5.579773
10	5			5							-	-	-	0.9808195
10					6			4			-	-	-	11.33391
10							9			1	-	-	-	7.672188
15	15										-	-	-	0.3269398
15				15							-	-	-	1.743679
15		5			5			5			-	-	-	15.21360
15						9	3		2	1	-	-	-	12.55449
15	3	2	2	2	1	1	1	1	1	1	-	-	-	4.642546
TAMANHO DO ARQUIVO DE AUDIO: 10 kB														

Tabela 16: Variação percentual para um arquivo de som de 25 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5	5											-	-	0.08108985
5			5									-	-	0.3243594
5					5							-	-	1.297438
5							5					-	-	5.189750
10		10										-	-	0.3243594
10						10						-	-	5.189750
10	5			5								-	-	0.7298086
10					6			4				-	-	9.860525
10							9			1		-	-	17.64515
15	15											-	-	0.2432695
15				15								-	-	1.816413
15		5			5			5				-	-	11.67694
15						9	3		2	1		-	-	24.39183
20	20											-	-	0.3243594
20						20						-	-	10.37950
20							13	3	2	1	1	-	-	36.32825
20	3	3	2	2	2	2	2	1	1	1	1	-	-	20.01297
TAMANHO DO ARQUIVO DE AUDIO: 25 kB														

Tabela 17: Variação percentual para um arquivo de som de 50 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5		5												0.07951970
5					5									0.6361576
5								3			1		1	16.28563
10			10											0.3180788
10						7				3				13.99547
10	2		2				3				1	1	1	40.79361
15	15													0.1192796
15				15										0.9542364
15							14						1	17.81241
20		20												0.3180788
20					10			10						11.45084
20			6			7			7					16.03117
20			5		5		5		5					13.51835
30				30										1.908473
30						28				2				15.26778
30		15						14					1	23.64916
41	20						21							10.84649
41	2	5	5	5	4	4	4	3	3	3	1	1	1	45.48527
TAMANHO DO ARQUIVO DE AUDIO: 50 kB														

Tabela 18: Variação percentual para um arquivo de som de 105 kB.

QUANTIDADE DE DADOS (Número de Caracteres)	NIVEIS (Níveis onde se insere dados)													VARIAÇÃO PORCENTAGEM
	1	2	3	4	5	6	7	8	9	10	11	12	13	
5		5												0.03706930
5					5									0.2965544
5								3			1		1	9.015254
10			10											0.1482772
10						7				3				6.524197
10	2		2				3				1	1	1	12.13649
15	15													0.05560395
15				15										0.4448316
15							14						1	10.91320
20		20												0.1482772
20					10			10						5.337979
20			6			7			7					7.562137
20			5		5		5		5					6.301781
30				30										0.8896632
30						28				2				7.117306
30		15						14					1	14.23461
40	20						20							4.581766
40	2	5	5	4	4	4	4	3	3	3	1	1	1	25.13299
60	30	15	15											0.3706930
60				30		30								4.300039
89					89									5.278668
89							50	30	9					19.92846
89					10	10	10	30	15	7	3	2	2	44.36454
TAMANHO DO ARQUIVO DE AUDIO: 105 Kb														

ANEXO D – ARTIGO SUBMETIDO A COMUNICAÇÕES CIENTÍFICAS

A Low-throughput Wavelet-based Steganography Audio Scheme

P. Carrión, H.M. de Oliveira, R.M. Campello de Souza

Abstract — **A novel method of embedding covert text in an audio signal via the wavelet transform is presented. The two-step steganography approach combines plaintext encryption by a standard cryptosystem followed by the embedding of the encrypted data in a .wav file. The audio signal is decomposed in twelve levels through a selected discrete wavelet. The stego-key indicates the decomposition level, the wavelet coefficients, and the decimal positions of coefficients where the secret text will be embedded. This novel stego-tool was fully implemented using Matlab™, a freeware version of which is available in the Internet.**

Keywords —- **steganography, audio signals, wavelet analysis.**

INTRODUCTION

Steganography literally means "covered writing", which is derived from the Greek στεγοσ (roof) and γραπηια (writing). It is defined as "the art and science of communicating in a way that hides the existence of the communication" [1-2]. Indeed, cryptography, watermarking and stenography are closely related [3-7]. Cryptography scrambles a message so it cannot be understood. Steganography hides the message. Humans always had the desire to conceal their messages from the curious eyes of others. Steganography uses range from the trivial to the illegal [1, 8].

This paper introduces the basis of a wavelet-based low-throughput secret key steganography system that requires the exchange of a secret key (stego-key) prior to communication. We try to take apart the fact that the message can only be read with a secret key by using a standard cryptosystem [9] before hiding a text in an audio file. Steganography makes possible a secret communication so as to hide a trade secret, a blueprint, or extra sensitive information without alerting potential attackers or eavesdroppers. The user should first select a suitable cryptography scheme and key length (e.g. 128 or 256 bits) prior to the implementation of the proposed stego-technique.

Certainly, trade-offs do exist between the amount of embedded data and the degree of immunity of the host signal. The amount of actual text hidden within the audio file must be chosen in such a way that a normal cover should hardly be distinguishable from a stego-object, neither by the human sensory system nor by a computer looking for statistical patterns. An interesting recent survey lists about 32 different audio stego-tools (.wav 50%, .mp3 28%, midi 6%, others 16%) [10]. When analysing the file in which the steganography was applied, the hidden information must be disguised in such a way that no clue is given about the possible existence of a stored message. Applications of stego such as digital watermarks and digital fingerprinting are possible, but they were not considered in this paper.

WAITER, PLEASE: THERE IS A TEXT IN MY AUDIO

Developing a data-hiding technique for audio is particularly tricky, because the human auditory system works over a wide dynamic range [11-12]. The sensitivity to additive noise is incredibly acute and perturbations in a sound file can be detected as low as 60 dB below ambient level. Nevertheless, there are some environmental distortions so common as to be ignored by the listener in most cases. Also, loud sounds tend to mask out quiet sounds. When embedding secret messages in an audio source, two features should be taken into account, namely, the digital format of the audio and the likely environments the signal will travel between encoding and decoding. The most popular format for representing samples of high-quality digital audio is a 16-bit linear quantization [11], e.g., Windows Audio-Visual (WAV) and Audio Interchange File Format (AIFF). These digital audio formats provide a proper cover for low-throughput steganography. The audio sampling rate rules data hiding by putting an upper bound on the working portion of the frequency spectrum (if a signal is sampled at ~8 kHz, no modifications can have frequency components beyond ~4 kHz). At 16 bits per sample and sampled at a rate of 44.1 kHz, digital audio has the bit-rate to support large messages. Possible strategies for embedding data inside the host audio include the use of frequencies inaudible to humans [13-15], low-bit encoding, phase-coding [16] (the ear is unable to perceive absolute phase, only relative phase), echo hiding [4], spread spectrum [17], embedding data using the LSB [3, 18-19], transform embedding techniques [20] (a variant of this method is precisely the approach

The authors are with the Federal University of Pernambuco - UFPE, Signal Processing Group, C.P. 7.800, 50.711-970, Recife - PE, Brazil (e-mail: jpcc1@hotmail.com, hmo@ufpe.br, ricardo@ufpe.br). This work was partially supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under research grant #301996 (HMdO). The first author also thanks Brazilian National Council for Scientific and Technological Development (CNPq) for a scholarship grant.

described here), and encoding musical tones [16]. This paper presents the preliminary of a novel scheme of steganography, and introduces the idea of combining two secret keys in the operation (Fig. 1). The first secret key encrypts the text using a standard cryptographic scheme (e.g. IDEA, SAFER+, etc.) prior to the wavelet audio decomposition [21]. The way in which the ciphertext is embedded in the file requires another key, namely a stego-key, which is associated with features of the audio wavelet analysis.

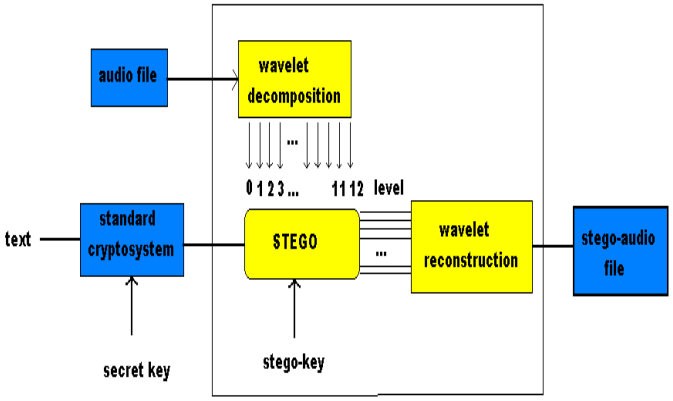


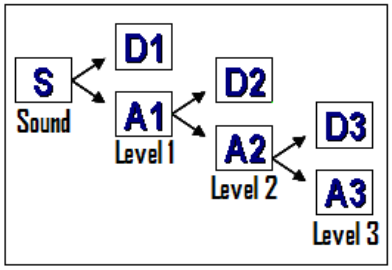
Fig. 1 –A novel method of embedding cover text in an audio signal by wavelet transform technique.

Audio File Specification

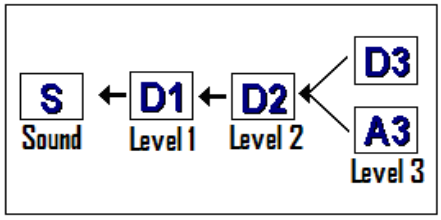
An extension .wav audio file can be read by the command *wavread* [22]. It supports *multichannel data*, with more than 32 bits per sample, and reading 24 and 32 bits of files .wav, which returns the sample data in a variable. The amplitude range is normalized (by Matlab™) in the range $-1.0 \le y \le 1.0$. Two Matlab commands are used:
Sound: converts a vector to a sound format to be played.
wavedec: performs 1D multilevel wavelet decomposition.

Description of the Tool

Wavelets are nowadays a powerful and well-established tool for signal analysis [21]. Approximation and details are derived from the wavelet decomposition of the audio, in an iterative procedure so that the approximations are split in each level. Fig. 2(a) illustrates the standard signal decomposition.



(a)



(b)

Fig. 2 – (a) Wavelet signal Decomposition, (b) Signal reconstruction.

In a fourth-level decomposition, the signal is reconstructed from the approximation ($A3$) and details components ($D3$, $D2$, $D1$) according to

$$S = A3 + D3 + D2 + D1, \text{ as shown in Fig.2(b).}$$

There are many applications of the wavelet decomposition as far as audio is concerned. Thanks to the advantages and features of wavelets, the data hiding can be done at different levels, so as to spread the text throughout the file in a non-suspicious manner.
A program was written using the Matlab™ platform [22] for embedding secret information into a sound, which is carried out by wavelet analysis. Two matrices are generated in the audio decomposition– the first one ($n \times 1$) contains wavelet coefficients, the column specifies the number of channels (mono, in this prospective study), and the second furnishes the length of decomposed signals in each level. The text to be hidden is first converted in ASCII code. Differently from all previous schemes [10, 23-24], data are used to **replace** the three most significant decimal positions of the wavelet coefficients of the audio decomposition matrix (unfortunately computational errors prevent using less significant positions).
Caution must be taken about the text length and audio file extent. Constraints were introduced in the program to ensure that the message is at least 1,000 times less than the original sound file, avoiding thus a large changing in the audio file.
Three parameters are required in the audio decomposition process: The wavelet matrix from the audio file in which the text is to be hidden, the number of levels where text will be inserted and the type of wavelet used in. Standard (Matlab tool-box) and non-standard families of wavelets are available built-in in the software. The user can select among 55 possibilities; the number of decomposition levels was preset as 12, with the option of performing the data embedding or not at each of the levels, thereby fragmenting the message. There are between 2^{10} and 2^{12} choices, depending on the length of the audio stretch, i.e., from 1,024 to 4,096 different possibilities to input the stego-key. This adds at least 10 bits to the secret key.
Alphanumeric passwords are used at each level with the aim of encrypting and scattering the hidden data through the audio signal. The password length is proportional to the amount of hidden data in each sub-block and it is previously converted into a binary stream. The 1's positions indicate where to insert encrypted data. In order to specify the initial position where the text starts within the sub-block, some constraints are made according to the length of each level. After all these operations, the audio signal is reconstructed storing the text message. For the converse procedure, text retrieving, the knowledge of an ensemble of information (the stego-key) is required, which allows extracting the hidden message in the correct way:

- Type of wavelet.
- Text message length.
- Number of levels containing text info.
- Level specification.
- Length of the subblocks of the recorded text at each selected level.
- Alphanumeric password, one at each level.
- Place(s) of text insertion.

ON IMPLEMENTING THE SCHEME

Embedding a text in an audio file

1. Enter an alphanumeric text.
2. Load a sound vector from an audio file.
3. Select a discrete wavelet for audio decomposition.
4. Audio decomposition, preset number of levels equals to 12.
5. Parsing the text into subblock for allocating characters in different wavelet scales (Fig.3b).
6. Enter the key (password).
7. Local (file position) for text insertion.
8. Data replace decimal position in the matrix of wavelet coefficients of the decomposition wavelet.
9. Save the stego-file containing the hidden text.

The (possibly ciphered) text is converted into a decimal string using the ASCII code as shown in Fig. 3a. It is then parsed (step 5) and each subblock allows allocating information at different wavelet scales (Fig.3b).

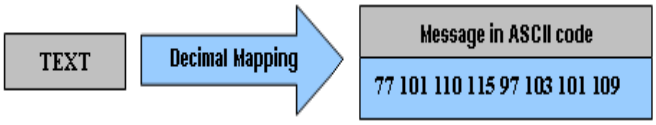


Fig. 3 – (a) alphanumeric-to-decimal conversion.

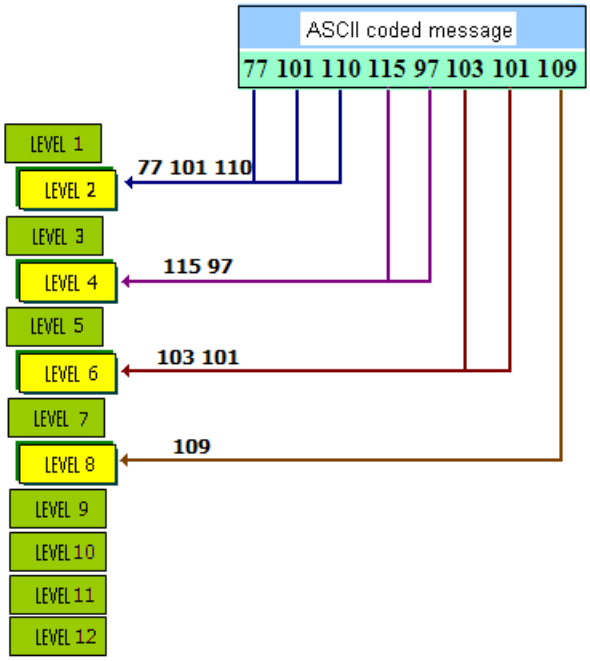


Fig. 3 – (b) In this figure, the stego-key selects decomposition levels 2, 4, 6 and 8. The number of data assigned to each level is also furnished by the stego-key.

Once step 5 is implemented, different alphanumeric passwords are provided to each decomposition level (they are also part of the stego-key) so as to spread the hidden information. Longer passwords are required at wavelet decomposition levels where much information is selected to be embedded. This is precisely step 6 (Fig. 4).

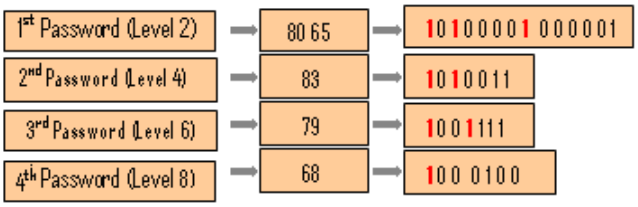


Fig. 4 – Passwords at selected levels (2, 4, 6 and 8) and their corresponding binary sequences.

Fig. 3 shows the four levels assigned by user through his stego-key (illustrated as 2, 4, 6 and 8 in this case). Another part of the stego-key gives the "start position" where information will be placed at each selected decomposition level. Fig. 5 shows only the data embedding at level 2, supposing that the start position indicated in the key was 3. Replacing the first three decimal positions in the floating-point representation of the wavelet coefficients carries out data insertion. A similar procedure is performed at the other levels.

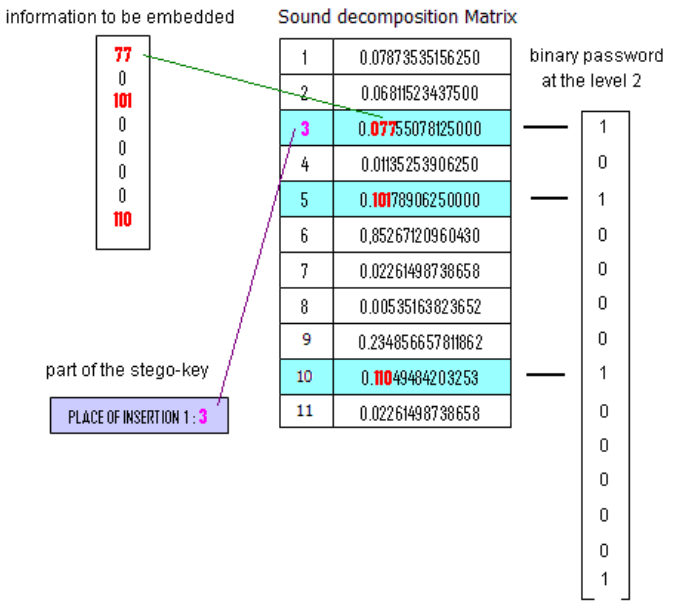


Fig. 5 – Passwords at selected levels (2, 4, 6 and 8) and their corresponding binary sequences.

Text retrieving from stego-audio file

1. Read the audio file containing the hidden message
2. Input of parameters:
   - Type of wavelet.
   - Text message length.
   - Number of levels containing text information.
   - Level specification.
   - Length of the recorded text.
   - Stego-password(s).
   - Place(s) of text insertion.
3. Audio decomposition using the selected wavelet.
4. Text extraction in each specified level.
5. Recover the original text.

FIRST-ROUND TEST AND VALIDATION

The text length to be dealt with is constrained, that is to say, the number of elements in the matrix derived from the sound file should be at least 1,000 times the number of message symbols, so as to maintain a small and imperceptible audio modification. An analysis of the audio changes shows that three factors play a major role the when embedding data:

- Quantity of levels and place where portions of the text are inserted
  Besides the number of levels where data are inserted – the amount of hidden text must take into account the insertion level, since modifications are less important at higher levels than at lower levels, where the changes are more remarkable.
- Message length and Sound length
  The greater the message, the greater the percentage of variation in the file. The sound length should be proportional to the inverse of the percentage of change – less variation is obtained, the greater the audio file is.

The percentage of variation was computed by comparing the two wavelet decomposition matrices (original and stego). If data are inserted solely at one of the first five levels (Table 1), the variation is smaller; if it is inserted exclusively at one of further levels, more alteration is perceived; broad data distribution engenders greater modification.

There is a natural trade-off between the distribution of data within the levels, and the possibility of the hidden information be unveiled by a third part. If the data is more scattered, there is a reduction on the probability of the information being unveiled. On the other side, if data are inserted at only the first levels, the variation is smaller, but the probability of message recovering increases.

The condition for the data to be hidden into the audio was established according to the following inequality:

$$x.f \leq M\,, \tag{1}$$

where x denotes the length of the message, M denotes the length of the original file and f is a conditional factor. The interpretation of the inequality is straightforward, the greater the factor f is, the smaller is the message that can be inserted into a given audio file. The length of the message (x) can be expressed as the quantity of characters of the message (n) times the average length of the inserted characters ($\overline{c}$):

$$x = n.\overline{c} \tag{2}$$

The inequality (1) yields

$$n.\overline{c}.f \leq M\,. \tag{3}$$

When the maximum tolerated amount of characters is to be introduced in the audio, the equality is met:

$$n_{max} \approx \frac{M}{\overline{c}.f}\,. \tag{4}$$

A very preliminary analysis is shown in Table I (appendix). Both the source code (extension .m) developed to implement this steganography technique, and a straightforward example are freeware available at the URL http://www2.ee.ufpe.br/codec/stego.html. In this version, only the stego-key is provided and the secret key should be supplied by external available software. The input text is expected to be previously enciphered. Indeed, the insertion of a very long text may sporadically introduce a few audible clicks or noticeable distortion.

CLOSING REMARKS

This paper is focused on the software implementation of a novel steganography approach, which dissociated robustness and transparency issues. Of course, a small amount of stegoanalysis results is presented. No comparisons with other stego-audio systems are made nor deeper stegoanalysis presented [12, 25-26]. Nevertheless, working software is the first step in such a direction. Despite steganography could provide also copyright protection, digital fingerprinting, feature tagging, and assurance of content integrity, we are only concerned with the ability of hiding a text message into a host audio file. Appending a cryptographic hash into the audio file so as to determine whether or not the file has been tampered with, could provide an efficient and undemanding protection [9].

ACKNOWLEDGEMENTS- The authors are grateful to Mr. Jorge Rehn for a number of valuable comments.

REFERENCES

[1] J.C. Judge, "Steganography: Past, Present, Future," Lawrence Livermore National Laboratory, U.S. Department of Energy UCRL-ID-151879 University of California, 2001.

[2] B. Dunbar, "A detailed look at Steganographic Techniques and their use in an Open-Systems Environment", Sans InfoSec Reading Room, 2002.

[3] N.F. Johnson, S. Jajodia, "Exploring steganography: Seeing the unseen," *IEEE Computer* **31** (2), pp.26–34, 1998.

[4] F.A.P. Petitcolas, R.J. Anderson, M.G. Kuhn, "Information hiding—a survey," *Proc. of the IEEE,* **87**, pp.1062–1078, July, 1999.

[5] M. Swanson, M. Kobayashi, A. Tewfik, "Multimedia data embedding and watermarking technologies," *Proc. of the IEEE*, **86** (6), pp.1064-1087, June, 1998.

[6] M.A.T. Alsalami, M.M. Al-Akaidi, "Digital Audio Watermarking: Survey," *17th European Simulation Multiconferece*, ESM, 2003.

[7] R. Anderson (Ed.), Information hiding, *Lecture Notes in Computer Science*, #1174, special issue: proc 1st Int. Workshop, Cambridge, UK, June, 1996), Springer-Verlag, 1996.

[8] USA Today, San Francisco, Reuters "Researchers: No secret Bin Laden messages on sites" URL: **www.usatoday.com/life/cyber/tech/2001/10/17/bin-laden-site.**

[9] B. Schneier, *Applied Cryptography - Protocols, Algorithms and Source Code* in C, John Wiley, 1996.

[10] J. Dittmann, C. Kraetzer, ECRYPT-European Network of Excellence in Cryptology, "Audio Benchmarking Tools and Steganalysis," Jan., 2006.

[11] M.K. Johnson, S. Lyu, H. Farid, "Steganalysis of Recorded Speech", *Proc. SPIE*, **5681**, pp. 664-672, 2005.

[12] H. Özer, İ. Avcýbas, B. Sankur, N. Memon, "Steganalysis of audio based on audio quality metrics," *Proc. of SPIE* **5020**, pp.55–66, June, 2003.

[13] K. Gopalan, S. Wenndt, "Audio steganography for covert data transmission by imperceptible tone insertion," *Wireless and Optical Communication MultiConference*, WOC, A.O. Fapojuwo Ed., 2004.

[14] K. Gopalan, S. Wenndt, A. Noga, D. Haddad, S. Adams, "Covert Speech Communication via Cover Speech by Tone Insertion," *Proc. of the 2003 IEEE Aerospace Conference*, MT, **4** (3), pp.1647-1653, 2003.

[15] I. Cox, M. Miller, "A review of watermarking and the importance of perceptual modeling," *Proc. of the SPIE/IST&T Conference on Human Vision and Electronic Imaging* II, SPIE 3016, San Jose, CA, pp. 92-99, February, 1997.

[16] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for data hiding." *IBM Systems Journal*, **35** (3&4), 1996, pp.313-336.

[17] D. Kirovski, H. Malvar, "Robust spread-spectrum audio watermarking," *IEEE Proc. of the Acoustics, Speech, and Signal Processing*, ICASSP, **3**, pp.1345-1348, 2001.

[18] S. Dumitrescu, X. Wu, Z. Wang, "Detection of LSB steganography via sample pair analysis,"*IEEE Transactions on Signal Processing,* **51**, pp.1995–2007, July, 2003.

[19] R. Chandramouli, N. Memon, "Analysis of LSB image steganography techniques," *Proc. ICIP* **3**, pp. 1019–1022, 2001.

[20] E.T. Lin, E.J. Delp, "A Review of Data Hiding in Digital Images," *Proc. of the Image Processing, Image Quality, Image Capture Systems Conference*, Savannah, Georgia, pp. 274-278, 1999.

[21] H.M. de Oliveira, *Análise de Sinais para Engenheiros*: Uma abordagem via WAVELETS. 1ª edição, Rio de Janeiro: Brasport, 2007 Série da Soc. Bras. de Telecomunicações ISBN 978-85-7452-283-8 (244p.)

[22] E.W. Kamen, B.S. Heck, *Fundamentals of Signals and Systems Using Matlab*. Englewood Cliffs, NJ: Prentice Hall, 1997.

[23] M. Noto, "MP3 Stego: Hiding Text in MP3 Files" available at http://rr.sans.org/covertchannels/mp3stego.php

[24] R. Petrovic, J.M. Winograd, K. Jemili, E. Metois, "Data Hiding Within Audio Signals," Series: Electronic and Energetics, **12** (2), pp.103-122, 1999.

[25] A. Westfeld, A. Pfitzmann, "Attacks on steganographic systems," in Information Hiding, LNCS 1768, pp.61-66, Springer-Verlag, Heidelberg, 1999.

[26] N.F. Johnson, S. Jajodia, "Steganalysis: The investigation of hidden information," *Proc. of the IEEE Information Technology Conference*, pp.113–116, 1998.

[27] R. Machado, "Stego," Available at http://www.nitv.net/~mech/Romana/stego.html (1994).

Table 1 – Preliminary results on changes in the audio file (930 kB) of as a function of the number of characters of the text and the levels where text is embedded. The variation is stego-key dependent.

Number of characters	Level (level where data is embedded) 1	2	3	4	5	6	7	8	9	10	11	12	13	Variation (%)
20					10			10						0.212
20			6			7			7					0.260
20			5		5		5		5					0.159
60	30	15	15											0.024
60				30		30								0.172
100								100						1.719
100	50								50					1.193
150					75			75						1.622
150		40		66					30		14			1.089
200	200													0.041
200					100			100						1.753
300		300												0.122
300						100		160		40				3.983
617	300			317										0.541
617			174				400				43			5.720
952							786			133		33		11.064
952		300		352		150		150						3.417

ESTA DISSERTAÇÃO FOI OBJETO DAS SEGUINTES COMUNICAÇÕES CIENTÍFICAS

1) CARRIÓN P., DE OLIVEIRA H.M., SOUZA, R.M.C.,
Transformada de Wavelets em Esteganografia: Ocultando textos simples em arquivos de áudio
XXXI Congresso Nacional de Matemática Aplicada e Computacional,
CNMAC 2008, 8 a 11 de Setembro de 2008.
(Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC)

2) CARRION, P., DE OLIVEIRA H.M., CAMPELLO DE SOUZA R.M.,
A Low-throughput Wavelet-based Steganography Audio Scheme
8th Brazilian Symposium on Information and Computer System Security,
September 1st to 5th in Gramado, Brazil, 2008.
(Sociedade Brasileira de Computação – SBC)

ANEXO E – QUALIDADE DO ÁUDIO

E1.1. Avaliação Subjetiva da Qualidade do Áudio

A avaliação de um arquivo de áudio pode ser realizada tanto de uma forma *objetiva*, mediante um algoritmo que permita detectar distorções, ou de uma forma *subjetiva* mediante testes de escuta para determinar a qualidade do som.

Um dos objetivos deste trabalho consiste em realizar testes de qualidade dos arquivos de áudio que foram submetidos ao processo esteganográfico. Para a realização desta prova é necessário enfatizar algumas recomendações adotadas pela ITU (International Telecommunication Union), padrões que são empregados em medições subjetivas. Estas indicações podem ser acolhidas por qualquer sistema de escuta [60] (por exemplo: sistema de telefonia, sistema de alto-falante, etc.), assim também, podem ser tomadas em conta características individuais relevantes de cada sistema.

A seguir detalham-se pontos específicos que devem considerar-se na avaliação conforme a norma ITU-T P.800 [60]:

- **Experimentador.** Deve decidir qual será o melhor método a ser utilizado conforme os resultados que deseja obter, levando em conta várias condições: número de testes, número de pessoas a submeterem-se à avaliação, recursos disponíveis, precisão dos resultados e a capacidade de extrair juízos e conclusões.

- **Ouvinte.** Os participantes podem ser eleitos aleatoriamente; para obter uma precisão adequada o número de ouvintes deve ser de no mínimo 12 pessoas e as respostas devem ser individuais. Cada ouvinte deve satisfazer aos seguintes critérios:

    - Não deve ter nenhuma relação direta com o trabalho.
    - Não pode haver escutado os áudios antes da avaliação.
    - Não tenha participado de provas subjetivas, há no mínimo seis meses.
    - Não tenha participado em um ensaio de conversação, há no mínimo um ano.

- **Ambiente.** O local a efetuar-se a avaliação deve ser uma sala silenciosa. Devem-se considerar fatores externos, tais como o ruído ambiente (50 dBA), o ar condicionado, a presença de pessoas na sala, ou qualquer ruído que possa comprometer os resultados.

- **Gravação.** Os áudios podem ser gravados através de um sistema de armazenamento digital ou mediante um computador. Para a gravação o microfone deve estar posicionado entre 140 e 200 milímetros de distância dos lábios. O falante deve encontrar-se confortável para conseguir um nível constante de voz, mantendo uma pronúncia fluente das palavras ou frases (sem ser de forma drástica nem com deficiências).

- **Experiência.** O teste não deve durar mais de 20 minutos. Se a sessão for muito demorada poderia ser dividida em várias sessões.

- **Alto-falante.** A distância mínima entre cada alto-falante deve ser de 1.5 metros, eles devem ser capazes de reproduzir fielmente os sons.

- **Coleta de dados.** As respostas dos ouvintes podem ser recolhidas em meios adequados como: papel, teclado, botões eletrônicos, computador, painéis (touch-screen), etc.

- **Escala.** Uma alternativa para medir a qualidade do áudio, baseia-se na comparação mediante uma escala (recomendada pela ITU) de 5 pontos, onde 5 é Excelente e 1 é Ruim. O método permite ao ouvinte avaliar a qualidade elegendo uma entre cinco opções de degradação, como está apresentado na Tabela 19:

Tabela 19: Escala de qualidade adotada pela ITU (Recomendação ITU-T P.800).

PARECER	ESCALA
▪ Excelente.	5
▪ Boa.	4
▪ Regular.	3
▪ Pobre.	2
▪ Ruim.	1

E1.2. TESTE DE AVALIAÇÃO SUBJETIVA DA QUALIDADE DO ÁUDIO

Nesta experiência você escutará um conjunto de seis áudios (monofônicos) três dos quais contém informação oculta, sem que o ouvinte tenha conhecimento de quais são os arquivos que possuem dados ocultos.

Os arquivos de áudio serão escutados e avaliados mediante a percepção subjetiva de cada ouvinte que informara o grau de qualidade do áudio escutado mediante a seguinte escala:

5. Qualidade Excelente.
4. Qualidade Boa.
3. Qualidade Regular.
2. Qualidade Pobre.
1. Qualidade Ruim.

TESTE.

Áudio 1: SOUND50.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Áudio 2: Happy.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Áudio 3: Blue.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Áudio 4: Memória.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Áudio 5: Números.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Áudio 6: Ingredientes.

Qualidade:

☐ Excelente ☐ Boa ☐ Regular ☐ Pobre ☐ Ruim

Obrigada pela sua ajuda neste teste.

E1.3. Conclusões do Teste

O método de avaliação consistiu em que cada ouvinte deveria escutar o conjunto dos áudios e julgar a qualidade de cada um mediante uma escala de cinco (5) alternativas recomendada pela ITU para esse tipo de teste subjetivo.

O conjunto de sons a serem analisados foi constituído por áudios selecionados do *Windows* como também gravação de voz mediante um microfone. Para arquivos de áudio que foram gravados se tomaram as precauções necessárias como ter uma pronuncia fluente, ou a distância que deve existir entre o microfone e os lábios, etc., para não introduzir algum tipo de distorção que comprometa o arquivo.

A coleta dos dados foi realizada em papel e os testes foram aplicados a doze (12) pessoas, elegidas de forma aleatória numa faixa etária de 18 a 35 anos de idade. As pessoas que foram submetidas cumpriam com os requerimentos básicos mencionados anteriormente, como por exemplo, nenhum deles tinha escutado os áudios antes da avaliação, além de que nenhum deles possuía relação direta com o trabalho. As respostas obtidas foram individuais de cada participante.

Para executar o teste tratou-se de procurar uma sala com as condições apropriadas, ou seja, deve ser silenciosa, para o qual, procurou-se um horário no qual não se tenha fatores externos ruidosos que prejudicassem a avaliação. Assim como a reprodução do som tinha que ser clara.

Tratou-se, na medida do possível, de procurar todos os recursos para obter os melhores resultados do teste. As respostas obtidas foram consideradas satisfatórias. Considerando o maior comprimento de mensagem de texto permitido pelo aplicativo em todos os casos nos quais foi inserido texto. Em nenhum caso foi possível identificar auditivamente a presença de alguma alteração que sugerisse a presença de informação lateral.

ANEXO F – MARCAS DE ÁGUA

F1.1. As Marcas de Água

Com a tecnologia o resguardo de mensagens em arquivos digitais se direcionou, na prática, para aplicações como:

- Proteção de direitos autorais.
- Número de série.
- Monitoramento e localização de arquivos.
- Impressões digitais.
- Prova de titularidade.
- Integridade de documentos.
- Etiquetagem.
- Não repúdio, entre outros.

A *marca de água* é uma das aplicações com maior interesse – no que diz respeito a dados ocultos – para a tecnologia, a indústria, empresas, pesquisa e outras áreas, devido a representar um meio de proteção para arquivos digitais. As marcas surgiram como uma prova de propriedade, e apesar de não constituírem uma prova concisa em tribunais [3], sua utilização continua crescente.

F1.2. Breve História das Marcas de Água

As *marcas de água* já eram empregadas em papel para diferenciar um fabricante de outro desde a idade média. O documento mais antigo que se conhece contendo uma *marca de água* registra-se na Itália e data de 1292. Naquela época existiam diversas fábricas que produziam e distribuíam papel de diferente qualidade, formato e preço. Foi devido à necessidade de identificar cada fabricante e dar o seguimento apropriado ao papel que surgiram as marcas; com o passar do tempo elas também foram empregadas para indicar datação e autenticação do papel.

Atualmente a aplicação prática da esteganografia se reflete nas *marcas de água* que pode encontrar-se em notas, como ilustrado na Figura 79, traveler's check (cf. Figura 80), etc., que são inseridas através de raio ultravioleta ou mediante tintas especiais.

Figura 79. Nota de 50 Francos Suíços.

Descrição dos símbolos que aparecem na nota de 50 Francos Suíços:

A: As cifras com a tinta Iriodin®: número mágico.

B: As cifras em *Marca de água.*

C: As cifras em talhe doce: O número que tinge.

D: O número perfurado (microperf ®).

E: A tinta com efeito óptico variável: O número camaleão.

F: As cifras com ultravioleta.

G: As cifras metalizadas: O número cintilante.

H: O efeito basculante.

1: Frente e verso.

2: *Marca de água* do rosto.

3: Guillochis.

4: Kinegram®: A cifra dançante.

5: Microtexto.

6: Símbolo para deficientes visuais.

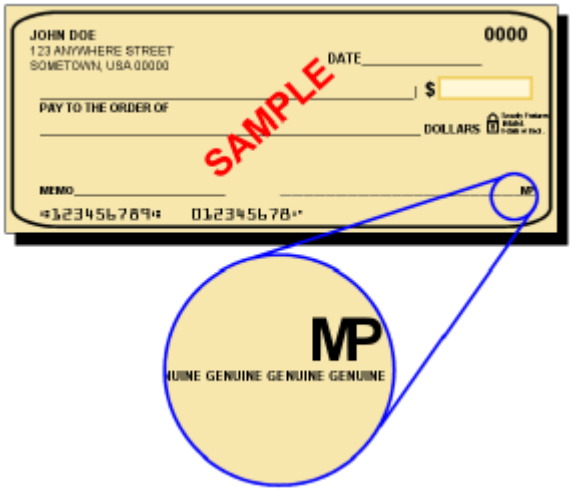


Figura 80. Traveler's check.

A evolução da tecnologia permite a inserção de *marcas de água* não só impressas, mas também de forma digital, podendo constituir uma porção de bits adicionados a um arquivo.

F1.3. Técnica de Inserção de Marcas de Água

A técnica para inserir *marcas de água* é similar à esteganográfica. Para realizar a inserção de uma *marca de água* precisa-se do sinal cobertor, da marca (que pode ser um logotipo, texto, números, etc.) e da chave secreta, estes elementos vão gerar o arquivo com a *marca de água*. Se o arquivo assinado é transmitido por um detector de marcas se obtém como resultado a *marca de água* inserida, tal como ilustra a Figura 81.

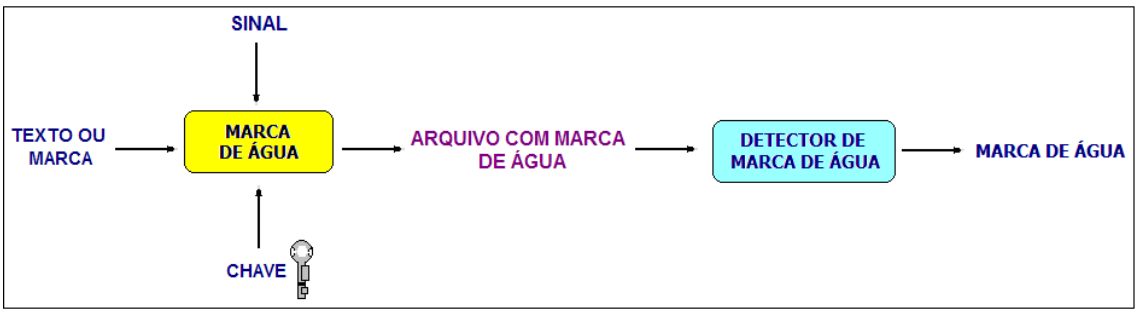


Figura 81. Diagrama de inserção de *Marca de água*.

Mas nem todas as técnicas que são aplicadas para ocultar dados são aplicáveis quando se trata da inserção de *marcas de água*, por exemplo, a tática do Bit menos significativo (LSB), não é recomendada para a inserção de marcas por ser muito vulnerável contra ataques maliciosos.

Geralmente o propósito de um ataque é inutilizar a *marca de água*, seja mediante a distorção ou eliminação da marca. Podem existir ataques do tipo malicioso e não malicioso (com o propósito de testar a marca).

Uma contramedida para evitar que a *marca de água* se perca quando o arquivo esteja sendo submetido a algum tipo de processamento (compressão, filtragem, adição de ruído, etc.), ou ataques que pretendem danificá-la ou retirá-la, pode ser à redundância, a qual consiste na inserção de mais de uma *marca de água* dentro do arquivo, considerando o tamanho (geralmente este método é válido para marcas pequenas), já que se a marca for muito extensa poderá ser inserida apenas uma vez a fim de não ocasionar distorção no arquivo cobertor. Essa estratégia aumenta a probabilidade de recuperação da informação.

F1.4. Tipos de Marcas de Água

As marcas poderiam classificar-se segundo a utilização (cf. Figura 82).

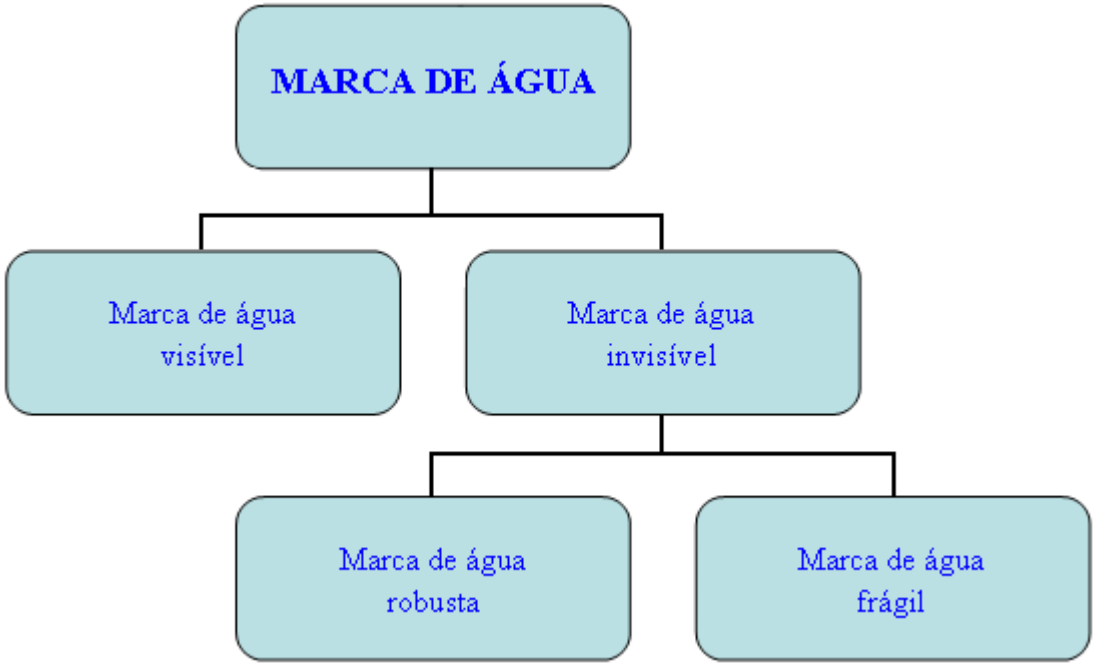


Figura 82. Classificação das *Marcas de água*.

- **Marca de água visível.** É um tipo de logotipo, ou assinatura visível que indica o autor(s), entidade ou proprietário. Este tipo de marca apenas deve ser inserida em fotos, vídeos, documentos, notas (ver Figura 79 - 80), etc., com o propósito de realizar o seguimento, identificar ou simplesmente marcar como propriedade privada podendo evitar a comercialização ilícita. Apesar que a *marca de água* seja conhecida, não poderá ser modificada nem destruída.

- **Marca de água invisível.** A marca invisível é um método de segurança e vigilância para arquivos digitais que circulam livremente na mídia. Permite a identificação de arquivos, seguimento pela rede, controle de modificações, autenticação, etc. Poderia considerar-se como uma estratégia para lutar contra a pirataria, plagio e outros tipos de adulterações. Uma marca invisível poderia subdividir-se de acordo a exigência requerida e à dificuldade de remoção:

    - **Marca de água Frágil.** A marca de água frágil permite verificar se um arquivo foi ou não modificado mediante a alteração da marca, isto é, permite verificar a integridade do arquivo. Para alcançar este objetivo a marca não deve ser muito robusta, mas deve ser imperceptível. Um ataque específico para este tipo de marca é a falsificação, intrusos poderiam modificar o conteúdo do arquivo e logo inserir a marca falsificada. Marcas frágeis são empregadas em gravações de voz, fotografias, em medicina, etc.

- **Marca de água Robusta.** Ao contrário da marca frágil, esta marca possui alto grau de robustez, resistindo a processamentos do sinal e ataques maliciosos. O principal objetivo é o de proteger os direitos de propriedade. A eliminação da marca de água robusta pode levar a comprometer a qualidade do sinal.

F1.5. Diferença entre Marca de Água e Esteganografia

A diferença entre ambas consiste na intenção, enquanto à esteganografia se propõe a ocultar uma mensagem para ser enviada a uma entidade específica, à *marca de água* oculta dados como meio de proteção e não possui um alvo específico. Algumas diferenças entre estas aplicações são abordadas na Tabela 20.

Tabela 20. Diferença entre Esteganografia e *Marca de água*.

ESTEGANOGRAFIA	MARCA DE ÁGUA
▪ A transmissão é ponto a ponto.	▪ Geralmente não existe um destino específico.
▪ Os dados inseridos são sigilosos.	▪ Nem sempre os dados inseridos são sigilosos.
▪ Deve possuir robustez.	▪ O tipo de marca (frágil ou robusta) determina a robustez.
▪ Geralmente a inserção de dados é maior.	▪ A quantidade de dados a inserir é menor.

F1.6. Características das Marcas de Água

- **Robustez.** A robustez vai depender da exigência da aplicação que se estiver realizando. Consiste na resistência que possua a marca depois de ter sido submetida a processamento ou ataques.

- **Capacidade.** É a quantidade máxima de dados que podem ser inseridos num determinado arquivo.

- **Imperceptibilidade.** Consiste na transparência da marca, a quantidade de dados a inserir e a técnica empregada para a ocultação definem a imperceptibilidade.

- **Redundância.** Garante a recuperação dos dados. Se modificações ou ataques são realizados sobre o arquivo que possui a marca, a recolocação dos dados proporciona solidez, permitindo recuperar a marca

em diversas partes do sinal. Este método é aplicável quando o tamanho da marca não é muito grande, já que não ocasiona modificações no arquivo cobertor.

- **Segurança.** A segurança para a recuperação da informação baseia-se na senha e pode ser reforçada mediante a encriptação dos dados.

- **Possibilidade de verificação.** A marca deve ser clara, de modo que se consiga facilmente a detecção e reconhecimento. A verificação da marca não deve levar a ambiguidade.

Considerando as três primeiras características, robustez, capacidade e imperceptibilidade, pode verificar-se mediante a Figura 83 que existe um *trade-off* entre elas.

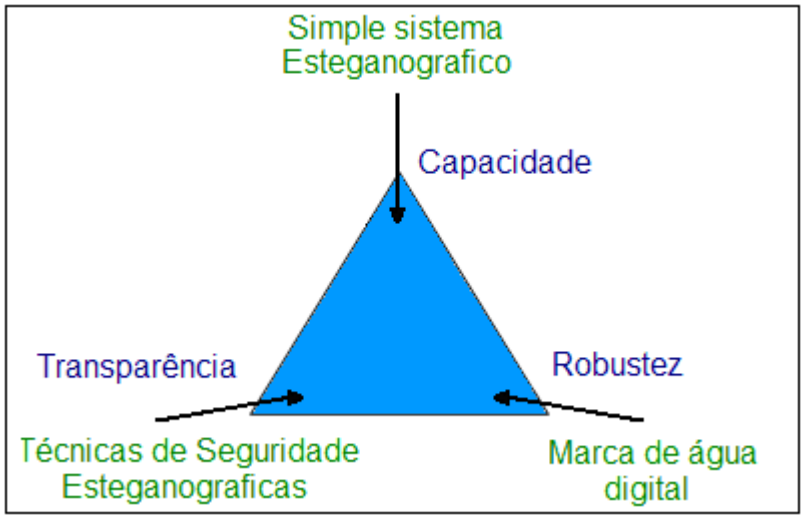


Figura 83. *Trade-off* entre Capacidade, Imperceptibilidade e Robustez.

É difícil que um algoritmo possa oferecer a capacidade máxima e máxima transparência ao mesmo tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

Principais Referências Bibliográficas

[1] I. S. Moskowitz; E. Garth Longdon; LiWu Chang, *A New Paradigm Hidden in Steganography*, Center for High Assurance Computer Systems, Naval Research Laboratory, 2000.

[2] Sociedade das Ciências Antigas, *Jean Trithemius*, http://www.sca.org.br/biografias/Trithemius.pdf, acessada em novembro de 2008.

[3] Stefan Katzenbeisser; Fabien A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, editors Artech House computing library, Inc., Norwood, MA, 2000.

[4] J.C. Judge, *Steganography: Past, Present, Future*, Lawrence Livermore National Laboratory, U.S. Department of Energy UCRL-ID-151879 University of California, 2001.

[5] B. Dunbar, *A detailed look at Steganographic Techniques and their use in an Open-Systems Environment*, Sans InfoSec Reading Room, 2002.

[6] E. Cole, *Hiding in Plain Sight: Steganography and the Art of Covert Communication*, Published by Wiley Publishing, Inc., Indianapolis, Indiana, 2003.

[7] R. Machado, *EzStego, Stego Online, Stego*, <http://www.stego.com>, acessada em abril de 2009.

[8] N. F. Johnson, *Steganography Tools and Software*, http://www.jjtc.com/Security/stegtools.htm, acessada em abril de 2009.

[9] D. Upham, *Jpeg-Jsteg, modification of the independent JPEG group's JPEG software (release 4) for 1-bit steganography in JFIF output files*, <ftp://ftp.funet.fi/pub/crypt/steganography/>, acessada em abril de 2009.

[10] H. Hastur, *Mandelsteg*, <ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/steg.tar.Z>, acessada em abril de 2009.

[11] F. Hansmann, *Steganos, Deus Ex Machina Communications*, <http://www.steganography.com/>, acessada em abril de 2009.

[12] N. F. Johnson; S. Jajodia, *Steganalysis of Images Created Using Current Steganography Software*, http://www.jjtc.com/ihws98/jjgmu.html, acessada em abril de 2009.

[13] N. F. Johnson; S. Jajodia, *Exploring Steganography: Seeing the Unseen*, http://www.jjtc.com/pub/r2026.pdf, acessada em abril de 2009.

[14] Bruno da Rocha Braga, *Análise de Frequências de Línguas*, Ravel; COPPE; Universidade Federal Rio de Janeiro, 2003.

[15] *Privacidade Online versus Combate ao Terrorismo*, http://www.ime.usp.br/~is/ddt/mac339/projetos/2001/alessandro/5_falhas.html, acessada em outubro de 2008.

[16] W. Stallings, *Cryptography and Network Security Principles and Practices*, Fourth Edition, Prentice Hall, November 16, 2005.

[17] M. Misiti; Y. Misiti; G. Oppenheim; J-M. Poggi, *Wavelet Toolbox For Use with* MATLAB®, User's Guide, Version 2.1, 2000.

[18] Y. Meyer, *Wavelets: algorithms and applications*, Philadelphia, SIAM, 1993, 133p.

[19] K. Randy Young, *Wavelet Theory and its Applications*, Kluwer Academic Publishers.

[20] I. Daubechies, *Ten lectures on wavelets*, Philadelphia, SIAM, 1992. (CBMS-NSF Regional Conference Series on Applied Mathematics) 357p.

[21] J. Morlet, *Sampling theory and wave propagation*, Springer, 1983.

[22] A. Bultheel, *Learning to Swim in a Sea of Wavelets*, Bull. Belg. Math. Soc. 1995, vol. 2, pp. 1-46.

[23] W. Sweldens, *The lifting scheme: a new philosophy in biorthogonal wavelet constructions*. Proceedings of SPIE, v.2569, p.68-79, 1995.

[24] A. Bruce; D. Donoho; M-Y. Gao, *Wavelet Analysis*, IEEE Spectrum, October, 1996, pp. 26-35.

[25] P. Kumar e E. Foufoula-Georgiou, *Wavelet Analysis in Geophysics:* An Introduction, Reviews of Geophysics, 1997, vol. 35, pp. 385-412.

[26] A. Akay, *Wavelet Applications in Medicine*, IEEE Spectrum, May, 1997, pp. 50-56.

[27] H.M. de Oliveira, *Análise de Sinais para Engenheiros*: Uma abordagem via WAVELETS. 1ª edição, Rio de Janeiro: Brasport, 2007 Série da Soc. Bras. de Telecomunicações ISBN 978-85-7452-283-8 (244p.).

[28] W. Bender; D. Gruhl; N. Morimoto; A. Lu, *Techniques for data hiding*, IBM Systems Journal, v.35 n.3-4, p.313-336, 1996.

[29] N.F. Johnson; S. Jajodia, *Exploring steganography: Seeing the unseen*, *IEEE Computer* **31** (2), pp.26–34, 1998.

[30] F.A.P. Petitcolas; R.J. Anderson; M.G. Kuhn, *Information hiding—a survey*, *Proc. of the IEEE,* **87**, pp.1062–1078, July, 1999.

[31] M. Swanson; M. Kobayashi; A. Tewfik, *Multimedia data embedding and watermarking technologies*, *Proc. of the IEEE*, **86** (6), pp.1064-1087, June, 1998.

[32] M.A.T. Alsalami; M.M. Al-Akaidi, *Digital Audio Watermarking: Survey*, *17th European Simulation Multiconferece*, ESM, 2003.

[33] R. Anderson (Ed.), *Information hiding*, *Lecture Notes in Computer Science*, #1174, special issue: proc 1st Int. Workshop, Cambridge, UK, June, 1996), Springer-Verlag, 1996.

[34] B. Schneier, *Applied Cryptography - Protocols, Algorithms and Source Code* in C, John Wiley, 1996.

[35] J. Dittmann; C. Kraetzer, ECRYPT-European Network of Excellence in Cryptology, *Audio Benchmarking Tools and Steganalysis*, Jan., 2006.

[36] M.K. Johnson; S. Lyu; H. Farid, *Steganalysis of Recorded Speech*, *Proc. SPIE*, **5681**, pp. 664-672, 2005.

[37] H. Özer; İ. Avcýbas; B. Sankur; N. Memon, *Steganalysis of audio based on audio quality metrics*, *Proc. of SPIE* **5020**, pp.55–66, June, 2003.

[38] M.K. Johnson; S. Lyu; H. Farid, *Steganalysis of Recorded Speech*, *Proc. SPIE*, **5681**, pp. 664-672, 2005].

[39] K. Gopalan; S. Wenndt, *Audio steganography for covert data transmission by imperceptible tone insertion*, *Wireless and Optical Communication MultiConference*, WOC, A.O. Fapojuwo Ed., 2004.

[40] K. Gopalan; S. Wenndt; A. Noga; D. Haddad; S. Adams, *Covert Speech Communication via Cover Speech by Tone Insertion*, *Proc. of the 2003 IEEE Aerospace Conference*, MT, **4** (3), pp.1647-1653, 2003.

[41] I. Cox; M. Miller, *A review of watermarking and the importance of perceptual modeling*, *Proc. of the SPIE/IST&T Conference on Human Vision and Electronic Imaging* II, SPIE 3016, San Jose, CA, pp. 92-99, February, 1997.

[42] W. Bender; D. Gruhl; N. Morimoto; A. Lu, *Techniques for data hiding*. *IBM Systems Journal*, **35** (3&4), 1996, pp.313-336.

[43] D. Kirovski; H. Malvar, *Robust spread-spectrum audio watermarking*, *IEEE Proc. of the Acoustics, Speech, and Signal Processing*, ICASSP, **3**, pp.1345-1348, 2001.

[44] S. Dumitrescu; X. Wu; Z. Wang, *Detection of LSB steganography via sample pair analysis*, *IEEE Transactions on Signal Processing,* **51**, pp.1995–2007, July, 2003.

[45] R. Chandramouli; N. Memon, *Analysis of LSB image steganography techniques*, *Proc. ICIP* **3**, pp. 1019–1022, 2001.

[46] E.T. Lin; E.J. Delp, *A Review of Data Hiding in Digital Images*, *Proc. of the Image Processing, Image Quality, Image Capture Systems Conference*, Savannah, Georgia, pp. 274-278, 1999.

[47] E.W. Kamen; B.S. Heck, *Fundamentals of Signals and Systems Using Matlab*. Englewood Cliffs, NJ: Prentice Hall, 1997.

[48] J. Dittmann; C. Kraetzer, ECRYPT-European Network of Excellence in Cryptology, *Audio Benchmarking Tools and Steganalysis*, Jan., 2006.

[49] M. Noto, *MP3 Stego: Hiding Text in MP3 Files*, available at https://www.sans.org/reading_room/whitepapers/stenganography/550.php, acessada em novembro de 2008.

[50] R. Petrovic; J.M. Winograd; K. Jemili; E. Metois, *Data Hiding Within Audio Signals Series: Electronic and Energetics*, **12** (2), pp.103-122, 1999.

[51] F. Miyara, *Control de ruído*, 1999, http://www.ingenieroambiental.com/4023/control%20de%20ruido,federico%20miyara.pdf, acessada em abril de 2009.

[52] H. Lins de Barros, *Música, Pintura, Física e as Leis Universais*, Centro Brasileiro de Pesquisas Físicas, 2007.

[53] V. E P Lazzarini, *Elementos de Acústica*, Music Department of National University of Ireland, Maynooth, 1998, http://www.fisica.net/ondulatoria/elementos_de_acustica.pdf, acessada em abril de 2009.

[54] R. A. E. Rüncos, *Equalização Adaptativa para Ambientes*, Curitiba 2007, http://www.eletrica.ufpr.br/marcelo/TE072/012007/Rudolfo-Equaliza.pdf, acessada em abril de 2009.

[55] *Audición: Pérdida de la audición inducida por el ruído*, http://familydoctor.org/online/famdoces/home/healthy/safety/work/226.html, acessada em novembro de 2008.

[56] J. Watkinson, *An Introduction to Digital Audio*, Second Edition, Elsevier Publishers, 2001.

[57] W. Tomasi, Sistemas de Comunicaciones Electrónicas. 2ª edición, Prentice Hall, México 2003.

[58] J. G. A. Barbedo e A. Lopes, Member, IEEE, *Uma Nova Estratégia para a Estimação Objetiva da Qualidade de Sinais de Áudio*, IEEE Latin America Transactions, Vol. 2, No. 3, September 2004.

[59] IEEE, *RECOMMENDATION ITU-R BS.1387-1: Method for objective measurements of perceived audio quality*, 1998-2001.

[60] International Telecommunication Union – ITU, *Methods for subjective determination of transmission quality*, ITU-T Recommendation P.800, Telecommunication Standardization Sector of ITU, 1996.