UNIVERSIDADE REGIONAL DE BLUMENAU

CENTRO DE CIÊNCIAS EXATAS E NATURAIS

CURSO DE CIÊNCIA DA COMPUTAÇÃO – BACHARELADO

SISTEMA DE OCULTAÇÃO DE DADOS EM ÁUDIO ATRAVÉS DE TÉCNICAS DE COMPACTAÇÃO E ESPALHAMENTO ESPECTRAL

LUIZ DIEGO AQUINO

BLUMENAU
2011

2011/1-26

LUIZ DIEGO AQUINO

SISTEMA DE OCULTAÇÃO DE DADOS EM ÁUDIO ATRAVÉS DE TÉCNICAS DE COMPACTAÇÃO E ESPALHAMENTO ESPECTRAL

Trabalho de Conclusão de Curso submetido à Universidade Regional de Blumenau para a obtenção dos créditos na disciplina Trabalho de Conclusão de Curso II do curso de Ciência da Computação — Bacharelado.

Prof. Aurélio Faustino Hoppe, Mestre - Orientador

BLUMENAU
2011

2011/1-26

SISTEMA DE OCULTAÇÃO DE DADOS EM ÁUDIO ATRAVÉS DE TÉCNICAS DE COMPACTAÇÃO E ESPALHAMENTO ESPECTRAL

Por

LUIZ DIEGO AQUINO

Trabalho aprovado para obtenção dos créditos na disciplina de Trabalho de Conclusão de Curso II, pela banca examinadora formada por:

Presidente:	Prof. Aurélio Faustino Hoppe, Mestre – Orientador, FURB
Membro:	Prof. Paulo César Rodacki Gomes, Doutor – FURB
Membro:	Prof. Fernando dos Santos, Mestre – FURB

Blumenau, 28 de junho de 2011

Dedico este trabalho à minha família e meus amigos, que estiveram sempre presentes fornecendo o apoio e incentivo necessários para a conclusão desta etapa de minha vida.

AGRADECIMENTOS

À minha família, e em especial à minha mãe, Nerci Aparecida Aquino, e minhas avós, Aparecida Ivone Petters e Maria do Espírito Santo Aquino, pelo amor, compreensão e confiança que me incentivaram para o término de mais esta etapa.

Aos colegas e amigos que me acompanharam durante a jornada acadêmica, em especial a Luiz Roberto Leicht e Thyago Schleuss, pelo apoio e companheirismo.

A Raphael Ricardo Moritz Petters, Rodrigo Felipe Moritz Petters e Rubio Luiz Aquino, primos e amigos sempre presentes ao longo de minha vida.

A Paulo Roberto Brandt, pela dedicação ao tema deste estudo e suporte científico prestado para a conclusão do mesmo.

Agradeço, em especial, ao professor Aurélio Faustino Hoppe, pela sabedoria e competência na orientação deste trabalho.

Nunca, nunca, nunca desista.

Winston Churchill

RESUMO

Este trabalho apresenta um sistema robusto de ocultação de dados em áudio baseado na técnica de espalhamento espectral por seqüência direta. O algoritmo apresentado insere marcas d'água em freqüências aleatórias no áudio de entrada, tornando-a imperceptível para o sistema auditivo humano. Os resultados obtidos apontam a eficiência do sistema quanto à capacidade de inserção e transparência dos dados.

Palavras-chave: Esteganografia. Compactação de dados. Espalhamento espectral. Segurança da informação.

ABSTRACT

This work presents a robust audio data hiding system based on the direct sequence spread spectrum technique. The algorithm presented inserts watermarks in random frequencies in the audio input, making it imperceptible to the human auditory system. The results indicated the efficiency of the system regarding the payload and data transparency.

Key-words: Steganography. Data compression. Spread spectrum. Information security.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE SIGLAS

AIFF – *Audio Interchange File Format*

BPSK – *Binary Phase Shift Keying*

CDMA – *Code Division Multiple Access*

DCT – *Discrete Cosine Transform*

DFT – *Discrete Fourier Transform*

DLL – *Dynamic-Link Library*

DSSS – *Direct Sequence Spread Spectrum*

FHSS – *Frequency Hopping Spread Spectrum*

FC – Filtro de Cepstrum

GPS – *Global Positioning System*

LSB – *Last Significant Bit*

MCLT – *Modulated Complex Lapped Transform*

MFSK – *M-ary Frequency Shift Keying*

MOV – *Model Output Variable*

ODG – *Objective Difference Grade*

PCM – *Pulse Code Modulation*

PN – *Pseudorandom Noise*

RF – Requisito Funcional

RNF – Requisito Não Funcional

SNR – *Signal-to-Noise Ratio*

SAH – Sistema Auditivo Humano

SARC – *Steganography Analysis and Research Center*

SS – *Spread Spectrum*

THSS – *Time Hopping Spread Spectrum*

UML – *Unified Modeling Language*

WAV – *Windows Audio-Visual*

WPF – *Windows Presentation Foundation*

XML – *eXtensible Markup Language*

LISTA DE SÍMBOLOS

dB - decibel

SUMÁRIO

1 INTRODUÇÃO

A crescente necessidade de proteção de direitos autorais em materiais multimídia, assim como a possibilidade de agregação de conteúdo, está cada vez mais dependente de avanços científicos que possibilitem a adição de tais características ao objeto alvo sem degradação de seu teor original.

Porém, uma das propostas disponíveis para resolução deste problema encontra-se na aplicação de métodos baseados em uma antiga arte, utilizada para embutir informações secretas em mensagens comuns, cujo formato não desperta atenção. A informação deve ser modelada de tal forma que sua detecção esteja acessível apenas para indivíduos autorizados.

O envio e recebimento de mensagens sigilosas é um ramo vasto de estudo. Empregadas durante séculos, estas práticas vem se beneficiando com a evolução tecnológica.

> É cada vez maior o número de pessoas que tentam a todo custo ludibriar as defesas para ter acesso a um dos bens mais preciosos da sociedade moderna: a informação. Por outro lado, existem outras pessoas que buscam o desenvolvimento e o estudo de técnicas para proteção das comunicações. (ALBUQUERQUE; BRAZIL; JULIO, 2007, p. 55).

Amplamente pesquisadas com propósitos militares, as técnicas de ocultação de informações baseiam-se principalmente em conceitos da criptografia[1] e suas ramificações. Uma delas, a esteganografia, utiliza métodos para inserir mensagens secundárias dentro de mensagens primárias. O termo, originado do alfabeto grego, significa "escrita escondida", pois através desta técnica procura-se camuflar a existência de informações ocultas em mensagens aparentemente inofensivas.

A disseminação da esteganografia em outros meios da sociedade foi auxiliada pelo progresso de equipamentos de gravação e manipulação de dados, possibilitando que a mesma seja utilizada em aplicações diversas.

Entretanto, pode-se dizer que a segurança da informação não gera receita diretamente. Assim, empresas sem interesses bélicos ainda consideram que investimentos neste setor não são prioritários. O resultado desta postura pode ser observado pelo crescente prejuízo gerado oriundo de violação de dados ou em medidas corretivas. Segundo Albuquerque, Brazil e Julio (2007, p. 59), atualmente artistas e gravadoras estão utilizando marcas d'água para proteger suas obras, tendo em vista o crescente aumento da pirataria e de sites na Internet que oferecem

[1] Ramo da ciência matemática que reúne métodos de transformação da informação, com o objetivo de torná-la secreta e ilegível para alguém sem autorização. O processo pode ser revertido para revelação da mensagem original pelo receptor apropriado.

acesso livre e ilícito a filmes, músicas e vídeos.

Diante desta realidade, este trabalho é focado no estudo e aplicação de métodos robustos para esteganografia de texto em áudio digital. Segundo Bender et al. (1996, p. 323), as características do Sistema Auditivo Humano (SAH) tornam o processo desafiador[2], sendo necessário explorar as fragilidades de sua estrutura para obter êxito na ocultação de uma mensagem em um som.

1.1 OBJETIVOS DO TRABALHO

O objetivo deste trabalho foi desenvolver um sistema capaz de ocultar textos em arquivos de áudio de forma imperceptível, com alto nível de segurança.

Os objetivos específicos do trabalho foram:

a) compactar o texto informado pelo usuário, utilizando técnicas de compressão que possuem propósitos distintos de forma integrada;
b) esconder informações compactadas no espectro sonoro do áudio selecionado, sem adicionar ruídos excessivos;
c) gerar um novo arquivo com a mesma informação sonora do áudio original, porém com as modificação necessárias em sua estrutura para ocultar a mensagem secreta;
d) analisar um arquivo de áudio, extraindo e descompactando a mensagem secreta contida nele, quando houver.

1.2 ESTRUTURA DO TRABALHO

Este trabalho está subdividido em capítulos que serão explicitados a seguir.

O primeiro capítulo apresenta a justificativa para o desenvolvimento da proposta do trabalho e alguns dados acerca do tema escolhido.

O segundo capítulo trata de conceitos gerais da esteganografia em áudio, esclarecendo tópicos importantes que devem ser considerados para que seja alcançado êxito nesta tarefa. A

[2] O SAH é capaz de reconhecer uma faixa de freqüência muito grande. Além disso, ele consegue perceber separadamente os sons que compõem um estímulo auditivo e é muito sensível a ruídos.

técnica de espalhamento espectral, utilizada para inserir dados na portadora do sinal do áudio sem adição excessiva de ruído, assim como a compactação de dados, empregada na redução do tamanho de textos pequenos, são apresentadas neste capítulo.

O terceiro capítulo exibe o modelo de dados e a metodologia adotada durante o desenvolvimento do trabalho, assim como os resultados obtidos pelo sistema.

As conclusões são expostas no quarto capítulo, onde sugestões para trabalho futuros também podem ser encontradas.

2 FUNDAMENTAÇÃO TEÓRICA

A seção 2.1 introduz um breve histórico sobre sistemas de esteganografia em áudio digital, suas aplicações e propriedades. Na seção 2.2 são apresentadas considerações sobre o espalhamento espectral, responsável por esconder informações em arquivos sonoros selecionados pelo usuário. Em seguida, na seção 2.3 encontram-se informações sobre compactação de dados e, por fim, na seção 2.4 são descritos trabalhos correlatos ao tema em questão.

2.1 ESTEGANOGRAFIA EM ÁUDIO

A ocultação de dados em conteúdo sonoro exige o estudo e a correta utilização das vulnerabilidades do SAH. Para tanto, pode-se adotar modelos matemáticos que representam a forma como o ouvido humano reconhece e percebe sons.

De acordo com Schütz (2009, p. 21), a abordagem da psicoacústica satisfaz esta necessidade, uma vez que esta área do conhecimento vem sendo estudada há décadas e aplicada em vários codificadores de áudio nos últimos tempos, comprovando sua eficácia.

Albuquerque, Brazil e Julio (2007, p. 79) lembram que, originalmente restrita às atividades militares, a esteganografia em sons foi uma importante evolução nas telecomunicações. A princípio seu objetivo era evitar detecções e tentativas de alteração de mensagens secretas por forças inimigas. Seguindo a evolução dos equipamentos transmissores e receptores de áudio, sua utilização foi sendo expandida para outros propósitos, como "[...] marcas d'água usadas em proteção de propriedade, autenticação e detecção de alterações, rastreamento de cópias etc." (SCHÜTZ, 2009, p. 90).

Contornados os obstáculos técnicos, a tecnologia pôde ser aplicada para outras finalidades. Diversos sistemas relacionados à esteganografia foram desenvolvidos desde a modernização desse tipo de técnica, produto da evolução da computação pessoal iniciada em 1985. O *Steganography Analysis and Research Center* (SARC), por exemplo, possui um catálogo com mais de 800 aplicações comerciais voltadas à esteganografia digital. Localizado em Fairmont, nos Estados Unidos da América, O SARC é um centro de segurança focado exclusivamente em pesquisas voltadas à esteganografia e esteganálise, onde são oferecidos

treinamentos e informações acerca de uma grande quantidade de ferramentas especializadas da área.

Atualmente, a agregação de conteúdo em arquivos de música mostra-se como um mercado potencial para a ocultação de dados em áudio. É possível utilizar as mesmas premissas expostas para acrescentar textos do artista, fotos e demais conteúdos que sejam relevantes ao consumidor. Esta não é uma tarefa trivial, pois conforme Abdulla et al. (2009, p. 54 - 55) e Schütz (2009, p. 17), a capacidade e a transparência da ocultação das informações estão diretamente relacionadas, sendo inútil aumentar a quantidade de informações que pode-se esconder se com isso houver degradação da qualidade do sinal hospedeiro.

Albuquerque, Brazil e Julio (2007, p. 60) enfatizam requisitos que devem ser satisfeitos em qualquer sistema esteganográfico:

a) segurança: a informação oculta deve ser invisível perceptivelmente e estatisticamente, não sendo possível descobrir a presença do conteúdo usando qualquer meio disponível. Este requisito não deve ocasionar complexidade computacional infinitamente grande no algoritmo;
b) carga útil (*payload*): quando direcionada à comunicação escondida, a esteganografia deve firmar um compromisso relativo à quantia de dados que pode ser transmitida. Para esta finalidade, o envio parcial da mensagem não é aceitável;
c) robustez: a resistência à algumas operações comuns realizadas em conteúdos multimídia, como compressão, agrega confiabilidade e utilidade ao sistema.

Estes requisitos são flexíveis e freqüentemente contraditórios. Dependendo dos argumentos da aplicação, um acordo deve ser estabelecido para cada um deles (ALBUQUERQUE; BRAZIL; JULIO, 2007, p. 60).

Segundo Kobuszewski (2004, p. 12), os métodos de ocultamento de dados em áudio mais conhecidos são:

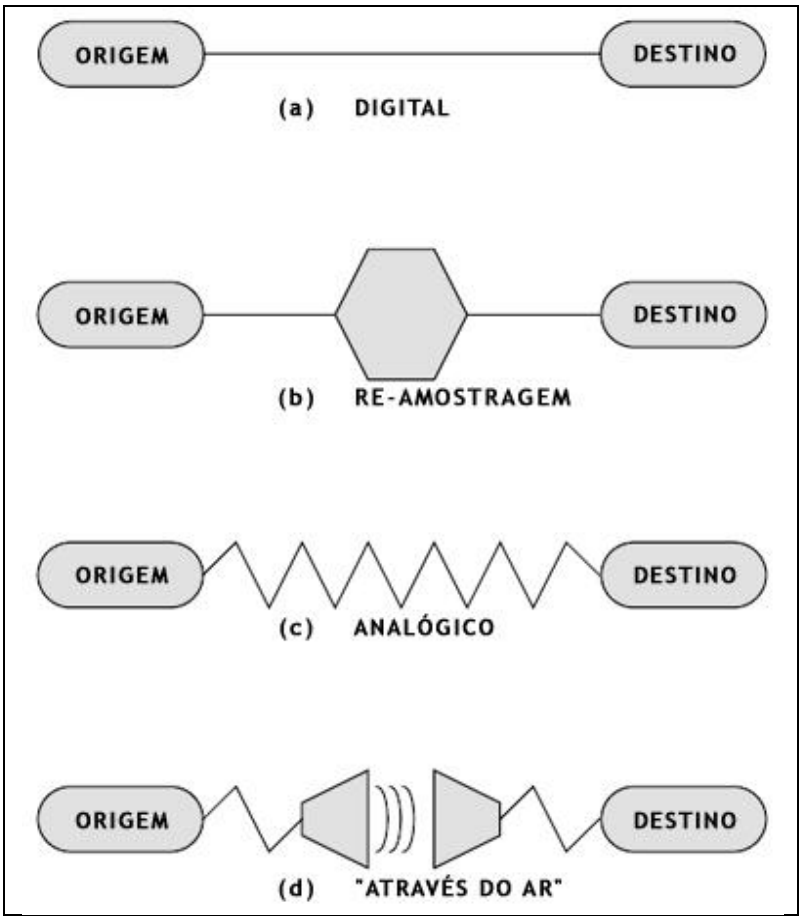
a) *Last Significant Bit* (LSB);
b) codificação de fases;
c) espalhamento espectral;
d) ocultação de dados no eco.

Cheng et al. (2007, p. 3) apontam que a técnica LSB é a forma mais simples e menos robusta de adicionar informação em arquivos de áudio digital, portanto menos utilizada pela indústria. De acordo com Albuquerque, Brazil e Julio (2007, p. 71), na codificação de fase o algoritmo trabalha substituindo a fase de um segmento inicial de áudio por uma fase de referência que representa os dados a serem escondidos. Para uma maior taxa de transmissão

de dados, pode-se utilizar o espalhamento espectral ou a ocultação de dados no eco. Enquanto a primeira técnica difunde informações secretas sobre o espectro de freqüências do som tanto quanto possível, a segunda esconde a informação em um arquivo de som através da introdução de um eco em seu sinal, variando a amplitude, a taxa de deterioração e a variação do sinal original (ALBUQUERQUE; BRAZIL; JULIO, 2007, p. 72 - 73).

O caminho de transmissão do sinal e a representação do mesmo também devem ser considerados no processo de seleção do método de esteganografia. Estes parâmetros definem as propriedades que o áudio deverá possuir para garantir o sucesso do procedimento.

Quanto aos meios de transmissão, eles são representados no esquema da Figura 1.



Fonte: adaptada de Bender et al. (1996, p. 324).

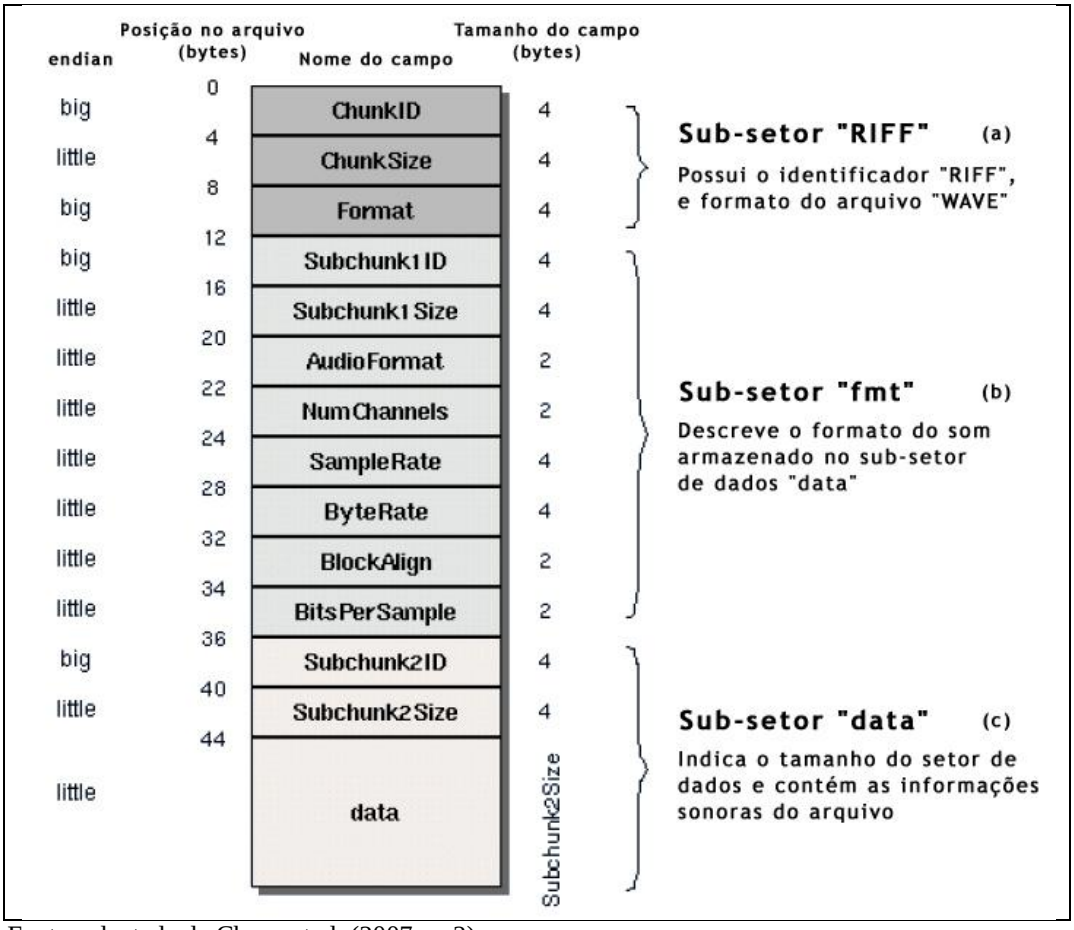
Figura 1 – Ambientes de transmissão de áudio

As classes existentes são ambiente digital fim a fim (Figura 1a), reconfiguração da taxa de amostragem do áudio digital (Figura 1b), reprodução e recepção analógica do som (Figura 1c) e transmissão pelo ar (Figura 1d).

Segundo Albuquerque, Brazil e Julio (2007, p. 70), os formatos mais populares para modular amostras de áudio digital são o *Windows Audio-Visual* (WAV) e o e *Audio Interchange File Format* (AIFF). De acordo com Cheng et al. (2007, p. 2), esse tipo de formato de arquivo torna possível o armazenamento de conteúdo sonoro em sistemas

computacionais. Assim, o som digital é representado através de *samples*, onde a cada fração de segundo uma amostra do som é armazenada no formato digital, ou seja, a informação é representada na forma de *bits* e *bytes* (KOBUSZEWSKI, 2004, p. 27). Atualmente, as formas de representação de áudio digital são agrupadas em formatos descompactados, formatos com compactação depreciativa e formatos com compactação sem perdas.

A Figura 2 ilustra a organização de um arquivo WAV, um dos pioneiros nesta área.



Fonte: adaptada de Cheng et al. (2007, p. 2).

Figura 2 – Formato de um arquivo WAV

Este formato de arquivo é dividido em duas seções principais:

a) cabeçalho do arquivo: possui as informações da configuração do áudio, como a taxa de amostragem, o número de canais e a quantidade *bytes* por segundos. Seu conteúdo agrupa as subseções “RIFF” e “fmt” (Figura 2a e Figura 2b, respectivamente);
b) dados: localização dos *samples* do áudio, em forma binária (subseção “data”, Figura 2c).

A utilização de formatos de áudio que possuem porções de informação bem definidas auxilia a análise dos dados e, portanto, atividades que necessitam alterar sua estrutura, como é o caso de processos esteganográficos.

2.1.1 Sistema Auditivo Humano

O estudo do funcionamento do SAH é um requisito mandatório durante o desenvolvimento de algoritmos de compressão e esteganografia específicos para processamento de áudio. Pesquisas focadas em mapear suas características conseguiram definir os limites de compreensão dos sons e a tolerância a ruídos do sistema auditivo.

Através de um procedimento chamado de modelagem perceptual é possível representar matematicamente a captação dos sons feita por humanos (SCHÜTZ, 2009, p. 19). Esta área de conhecimento pode ser dividida em basicamente três linhas de abordagem: fisiológica, psicoacústica e estatística. A primeira procura modelar o mecanismo neurofisiológico do sistema de audição, enquanto a segunda baseia-se nos efeitos da fisiologia do ouvido na percepção dos sons, através de testes subjetivos. Por fim, a última busca o entendimento das reações a determinados estímulos sonoros (JEHAN, 2005, p. 41).

Por ser mais acessível e menos complexa que as outras técnicas mencionadas, este trabalho abrange apenas a prática psicoacústica.

2.1.1.1 Modelagem Psicoacústica

Os princípios da psicoacústica envolvem vários conceitos relacionados à percepção de sons. De acordo com Painter e Spanias (2000, p. 454), durante as análises realizadas por este modelo, são rastreadas informações irrelevantes na amostra de sinal, ou seja, componentes sonoras imperceptíveis até mesmo para ouvintes treinados, como músicos e maestros. Estes dados são utilizados para configurar limiares de audição e limites para mascaramento de informações no domínio da freqüência, entre outros (SCHÜTZ, 2009, p. 22).

No caso deste trabalho, as informações obtidas são utilizadas para tratar o áudio antes da marca d’água ser inserida. Assim, a fim de evitar a adição excessiva de ruídos no som resultante, o sinal hospedeiro passa por um processo conhecido como conformação de ruído (SCHÜTZ, 2009, p. 35).

Além de ser responsável por tornar o dado escondido no áudio imperceptível aos ouvintes, este procedimento evita que a energia média da portadora seja alterada consideravelmente, fato que ajudaria a evidenciar modificações no áudio original. Os limiares de mascaramento auxiliam de forma decisiva esta etapa da análise de ruídos, indicando quais

as componentes modificadas do áudio podem prejudicar a qualidade do sinal gerado pelo sistema. Quando uma delas é identificada, o algoritmo deve substituí-la pelo sinal original, impedindo que um dos blocos denuncie a violação estrutural que foi realizada pela ferramenta.

2.1.2 Aplicações da Tecnologia

Além de proporcionar uma forma de agregação de conteúdo em um material multimídia, a esteganografia de dados em sons pode ser utilizada para diversos fins, como rastreio de dados e comunicação sigilosa.

A existência de uma marca d'água no conteúdo permite que a comprovação de autenticidade da informação seja realizada sem a utilização de equipamentos complexos de reconhecimento de áudio ou ouvintes treinados. A disponibilização de tal tecnologia torna-se útil também quando vista como mecanismo de detecção da utilização não autorizada de um determinado áudio, seja em vídeos ou em apresentações em geral.

De acordo com Cox e Miller (2001), a gênese da marca d'água eletrônica ocorreu em 1954, quando Emil Hembrooke registrou a patente de um método para adição imperceptível de códigos para identificação em músicas, com o propósito de detectar sua propriedade. Porém, apenas a partir de 1990 esta área obteve atenção significativa como tópico de pesquisa, quando soluções de cunho essencialmente heurístico começaram a ser publicadas.

De uma perspectiva voltada para negócios, o foco se mantém na utilização deste tipo de técnica para o desenvolvimento de soluções econômicas de problemas reais, como adição de interatividade na mídia, monitoramento de propagandas acopladas em conteúdo sonoro, estimativa de audiência etc.

Assim, Cox e Miller (2001) apontam que, de forma geral, as aplicações comerciais da área podem ser categorizadas entre sistemas de segurança e softwares para controle de dispositivos. Em muitas delas, os requisitos de segurança são reduzidos e, em alguns casos, inexistentes, já que não existe motivação justificável para tentativas de remoção da marca d'água. A agregação de informações relevantes ao conteúdo original de uma música para apreciação do usuário final, por exemplo, não é vista como prejudicial ou como um agente de rastreamento, portanto não desperta grande interesse para eventuais ataques.

No âmbito acadêmico, a ocultação de informações em áudio continua oferecendo desafios interessantes para diferentes níveis de pesquisa. Conforme Cox e Miller (2001), a

evolução da tecnologia de esteganografia em áudio trouxe consigo, dentre outros, o avanço dos métodos de ataque com o objetivo de atrapalhar a detecção e extração dos dados mascarados. Desde que modelagens de sistema baseadas em técnicas de espalhamento espectral começaram a surgir, o requinte dos processos aplicados na tentativa de confundi-las também foi aprimorado. Isto ocorreu devido ao fato que a nova metodologia impôs desafios maiores do que aquelas existentes anteriormente, sendo necessário agora que o adversário atacasse diversos pontos no espectro sonoro de maneiras específicas para obter algum sucesso em seu objetivo.

2.1.3 Propriedades de marcas d'água digitais

Conforme Cox, Miller e Bloom (2002, p. 26 - 27), sistemas esteganográficos empregados na criação de marcas d'água digitais podem ser caracterizados por algumas propriedades específicas. Os requisitos da aplicação são responsáveis por definir a importância de cada uma destas características na modelagem da solução adotada.

Os tópicos mais relevantes são apresentados nesta seção. Eles normalmente são agrupados da seguinte forma:

a) inserção de dados: efetividade, fidelidade e taxa de ocultação;
b) detecção de dados: detecção cega e informada, falso positivo e robustez;
c) custos: esforço computacional de inserção e detecção.

Para facilitar a compreensão acerca da relevância e significado de cada um, estes itens são detalhados nas seções 2.1.3.1 (inserção), 2.1.3.2 (detecção) e 2.1.3.3 (custos).

2.1.3.1 Inserção de dados

Cox, Miller e Bloom (2002, p. 27) explicam que as propriedades relacionadas à inserção de dados são relativas à qualidade final da marca d'água adicionada ao material multimídia, indicando o nível de confiabilidade que este indicador oferece. As próximas subseções abordam temas acerca deste fator das marcas d'água.

2.1.3.1.1 Efetividade de inserção

A efetividade de inserção de um sistema de marca d'água digital é definida pela probabilidade de que o resultado gerado pelo módulo de agregação de dados contenha a informação escondida (ALSALAMI; AL-AKAIDI, 2003, p. 3). Assim, pode-se entender que a porcentagem de mídias alteradas pelo algoritmo que resultam em detecção positiva de marcas ocultas é equivalente a este coeficiente.

Entretanto, Cox, Miller e Bloom (2002, p. 27) indicam que o custo para que uma taxa de 100% de efetividade seja atingida é alto, considerando que outras propriedades devem ser preservadas. Em alguns casos, por exemplo, o desempenho do sistema é mais importante do que a efetividade de agregação de dados. Portanto, a manutenção desta propriedade deve ser monitorada para que seus requisitos não prejudiquem o processamento da aplicação como um todo.

2.1.3.1.2 Fidelidade

Esta característica refere-se à similaridade entre a versão original e a versão alterada da mídia que armazena uma marca d'água (COX; MILLER; BLOOM, 2002, p. 27).

Para que a comparação seja realizada de maneira adequada, ela deve ser feita após a transmissão dos dados, quando este for o caso. Esta abordagem é necessária, pois o processo de envio e recepção do sinal pode afetar a qualidade do material que está sendo transmitido, com a adição de ruídos em seu conteúdo. Alsalami e Al-Akaidi (2003, p. 3) recomendam que os dados sejam confrontados no momento em que forem apresentados ao destinatário final, evitando assim resultados parciais e incompletos.

2.1.3.1.3 Taxa de ocultação

Alsalami e Al-Akaidi (2003, p. 3) apontam que em aplicações direcionadas ao processamento de áudio, a taxa de ocultação é delimitada pela quantidade de *bits* agregados por segundo transmitidos no sinal.

Este fator pode ser bastante variável e flexível, dependendo do objetivo da marca

d’água. Sistemas de controle de cópias, por exemplo, possuem requisitos diferentes de soluções projetadas para identificação de anúncios publicitários durante transmissões de radiodifusão; enquanto o primeiro caso deve adicionar um código único no conteúdo da mídia, o segundo é responsável por identificar blocos diferentes da programação nos primeiros segundos de cada um deles, com dados suficientemente relevantes (COX; MILLER; BLOOM, 2002, p. 28).

2.1.3.2 Detecção de dados

Características reunidas na categoria de detecção das informações definem a forma como a mensagem oculta é extraída e sua capacidade de resistir a ataques de agentes maliciosos (COX; MILLER; BLOOM, 2002, p. 27). Neste sentido, propriedades importantes são detalhadas nas próximas subseções.

2.1.3.2.1 Detecção cega e informada

Eventualmente, o armazenamento do sinal original utilizado para mascarar informações pode ser visto como uma prática interessante, facilitando a posterior extração do dado oculto. Porém, esta metodologia limita consideravelmente a utilização de tal sistema, pois pressupõe a estocagem de um grande número de mídias. Conforme Cox, Miller e Bloom (2002, p. 29), módulos de detecção que necessitam de acesso ao material original são apresentados como *detectores informados*.

Por motivos que envolvem tópicos como praticidade e independência do programa, esta solução nem sempre é adotada. Entretanto, algoritmos utilizados para detecção de marcas d’água podem ter seu desempenho comprometido caso não seja disponibilizada a versão sem modificações do áudio que deseja-se analisar. Sistemas desenvolvidos para operar neste tipo de cenário são classificados como *detectores cegos*, pois conseguem extrair dados ocultos sem necessitar de informações adicionais relacionadas ao material original (ALSALAMI; AL-AKAIDI, 2003, p. 3).

Além destas, Cox, Miller e Bloom (2002, p. 29) lembram que outras definições podem ser encontradas na literatura existente acerca de sistemas esteganográficos:

a) sistemas de marca d’água privados: equivalentes aos detectores informados,

podem ser utilizados por apenas um grupo seleto de usuários que detém a mídia original empregada no processo de esteganografia;

b) sistemas de marca d'água públicos: também conhecidos como detectores cegos, permitem que qualquer usuário submeta um sinal à análise, afim de revelar marcas d'água presentes em sua estrutura.

Em geral, o sinal original é disponibilizado apenas em aplicações privadas, portanto detectores informados não podem ser utilizados em aplicações públicas.

Os termos "público" e "privado" se tornam ambíguos quando é abordada a temática de quem deve ser capaz de identificar a marca d'água. Em alguns casos, por exemplo, apenas usuários marcados como confiáveis podem acessá-la, tornando até mesmo um sistema de marca d'água público em uma aplicação privada. Por esta razão, esta termologia é imprópria em certos casos (COX; MILLER; BLOOM, 2002, p. 283).

2.1.3.2.2 Falso positivo

A detecção de marcas d'água em um sinal que não possui efetivamente dados ocultos é chamada de falso positivo (ALSALAMI; AL-AKAIDI, 2003, p. 3).

Novamente, a probabilidade ideal de falsos positivos gerados por um detector é exclusivamente dependente do propósito do software. Em alguns casos, este índice deve ser extremamente baixo para que a confiabilidade no sistema seja assegurada e seu propósito, mantido. Por exemplo, o consenso geral aponta que detectores que analisam vídeos armazenados em DVD devem ter uma probabilidade de falso positivo de 1 *frame* a cada 1000 anos de operação contínua (COX; MILLER; BLOOM, 2002, p. 29 - 30).

2.1.3.2.3 Robustez

Segundo Alsalami e Al-Akaidi (2003, p. 3), esta propriedade está relacionada à resistência da marca d'água em uma mídia que sofreu alterações devido a operações comuns de processamento de sinal, como transmissão, compactação etc.

Apesar desta definição aparentar a idéia de que robustez é um requisito básico de sistemas especializados em marcas d'água, em algumas situações esta característica é completamente irrelevante, quando não indesejada. Neste sentido, marcas d'água frágeis

podem ser empregadas para denunciar alterações, mesmo que mínimas, em um áudio. Qualquer operação aplicada sobre o sinal deve ocasionar a destruição parcial ou total da informação mascarada.

No entanto, Cox, Miller e Bloom (2002, p. 30 - 31) destacam que o conceito de robustez geralmente é aplicado com outro foco. Em grande parte das implementações, nota-se que o objetivo é a prevenção de alterações nas marcas d'água escondidas em um sinal, provocadas por interferências externas. Ataques intencionais são categorizados conforme se segue:

a) remoção não autorizada de marca d'água;
b) adição não autorizada de marca d'água;
c) detecção não autorizada de marca d'água.

A remoção e a adição não autorizada são referenciadas como ataques *ativos*, pois alteram o conteúdo do sinal portador da marca d'água. Como a detecção não autorizada não realiza este tipo de modificação no material, ela é reconhecida como um ataque *passivo* (COX; MILLER; BLOOM, 2002, p. 31).

2.1.3.3 Custos

Tópicos acerca dos custos para agregação e extração de marcas d'água são estudados para avaliar as iniciativas financeiras e operacionais necessárias para a realização de tais processos (COX; MILLER; BLOOM, 2002, p. 27). Informações relacionadas a este tipo de investimento são examinadas na próxima subseção.

2.1.3.3.1 Custo e esforço computacional

Sistemas esteganográficos construídos para atender propósitos comerciais tendem a apresentar custos elevados de desenvolvimento e, conseqüentemente, de venda.

A ocultação de informações em sinais de áudio ainda é um processo pouco difundido, quando comparado com o mascaramento de dados em imagens e vídeos. Cheng et al. (2007, p. 3 - 4) indicam que a dificuldade e complexidade da implementação deste tipo de aplicação são possíveis motivos da escassa existência de referências sobre o tema.

As capacidades auditivas do ouvido humano são extremamente desenvolvidas, assim

como o sistema visual, limitando as opções para resolução do problema (ver seção 2.1.1).

Cox, Miller e Bloom (2002, p. 36) afirmam que a agilidade do algoritmo de inserção e detecção da marca d'água, aliada à quantidade de agregadores e detectores que devem ser distribuídos, são fatores que influenciam diretamente no valor final do sistema. A necessidade de acoplar tal aplicação em hardware incrementa consideravelmente o investimento exigido para a obtenção de uma solução adequada.

Em algumas situações, o sistema deve possuir alto desempenho para atender seu propósito, chegando ao ponto de ter que realizar operações em tempo real. Para que o cronograma de um ambiente real de produção não seja interferido por questões relacionadas a marcas d'água, o esforço computacional exigido é equivalentemente grande (COX; MILLER; BLOOM, 2002, p. 36). Além disso, a utilização de técnicas de espalhamento espectral acarreta *overheads* durante o processamento, pois várias faixas do espectro do áudio precisam ser analisadas durante a inserção e a extração de dados sigilosos.

Concluindo as observações acerca do custo operacional de métodos de esteganografia, é importante ressaltar que, pelo fato da mídia envolvida tratar-se de conteúdo sonoro, é necessário converter os dados recebidos em estruturas que facilitem a execução dos procedimentos listados. Nesta tarefa, são empregadas transformadas matemáticas em diversas iterações, cujas operações são custosas para os microprocessadores em geral.

2.2 ESPALHAMENTO ESPECTRAL

Sistemas de espalhamento espectral (em inglês, *spread spectrum* ou SS) são aqueles onde a banda do espectro utilizada na comunicação é muito maior do que a necessária para transmitir a informação (FERNANDES, 2002, p. 13). De acordo com Schütz (2009, p. 37), o uso deste tipo de proposta oferece maior segurança à comunicação, pois permite a redução ou a eliminação de interferências na transmissão e recepção do sinal, naturais ou intencionais.

Fernandes (2002, p. 12 - 13) indica que dentre as formas disponíveis para realizar o espalhamento espectral, as mais utilizadas são baseadas no emprego de códigos pseudoaleatórios[3] (PN, do inglês *Pseudorandom Noise*), para variar ou modular o sinal

[3] Seqüências pseudoaleatórias são determinísticas e independentes dos dados transmitidos. Na aplicação de técnicas de espalhamento espectral, estas seqüências devem ser de domínio comum do transmissor e do receptor da mensagem (SKLAR, 2001, p. 728 - 729).

portador, tornando a transmissão indistinguível para receptores que não conhecem sua seqüência. Uma etapa importante para que o envio e recebimento do sinal sejam realizados com sucesso é a sincronização entre os envolvidos (GARCIA, 1999, p. 53). O sistema receptor deve executar varreduras na informação captada, orientado pelo código PN utilizado no espalhamento, e assim ajustar-se para recuperar a transmissão da forma correta.

Sklar (2001, p. 719) define alguns requisitos que sistemas SS devem atender:

a) o sinal deve ocupar uma banda muito maior do que a banda mínima necessária para a transmissão dos dados;
b) o espalhamento deve ser realizado com um sinal independente da informação transmitida;
c) no dispositivo receptor, a recuperação do dado original deve ser feita através da correlação do sinal recebido com uma réplica sincronizada do código utilizado para modular a informação.

Entre as aplicações comerciais que utilizam-se deste conhecimento, estão os sistemas de telefonia *Code Division Multiple Access* (CDMA), uma tecnologia de múltiplo acesso por divisão de código onde é possível que vários usuários ocupem a mesma freqüência simultaneamente. Aparelhos de GPS (*Global Positioning System*) também utilizam uma variação de espalhamento espectral para estabelecer comunicação com satélites que capturam a informação disponibilizada por estes equipamentos (FERNANDES, 2002, p. 15 - 16).

Conforme comprovado durante a Segunda Guerra Mundial, as características desta técnica tornam-na apta para realizar processos esteganográficos (SKLAR, 2001, p. 719). Diferentemente de métodos como o LSB, que segundo Kobuszewski (2004, p. 26) distribui a informação que deseja-se ocultar através dos *bits* menos significativos do arquivo de som, as técnicas de espalhamento espectral utilizam o próprio conteúdo sonoro durante o processo. Este aspecto dificulta a interceptação e extração da informação por receptores não autorizados, pois a singularidade de cada som influencia todas as etapas de mascaramento do dado, ou seja, não é utilizada uma regra específica em todos os casos.

2.2.1 Principais técnicas de espalhamento espectral

Basicamente, três técnicas de espalhamento espectral podem ser utilizadas para fins de ocultação de dados (SKLAR, 2001, p. 724): seqüência direta, salto em freqüência e salto no

tempo. Existem também combinações híbridas destas técnicas, normalmente empregadas para obtenção de resultados mais rápidos e transmissões mais seguras. Visto que representam simples extensões das metodologias citadas, as soluções híbridas não são tratadas neste trabalho.

2.2.1.1 Espalhamento espectral por seqüência direta

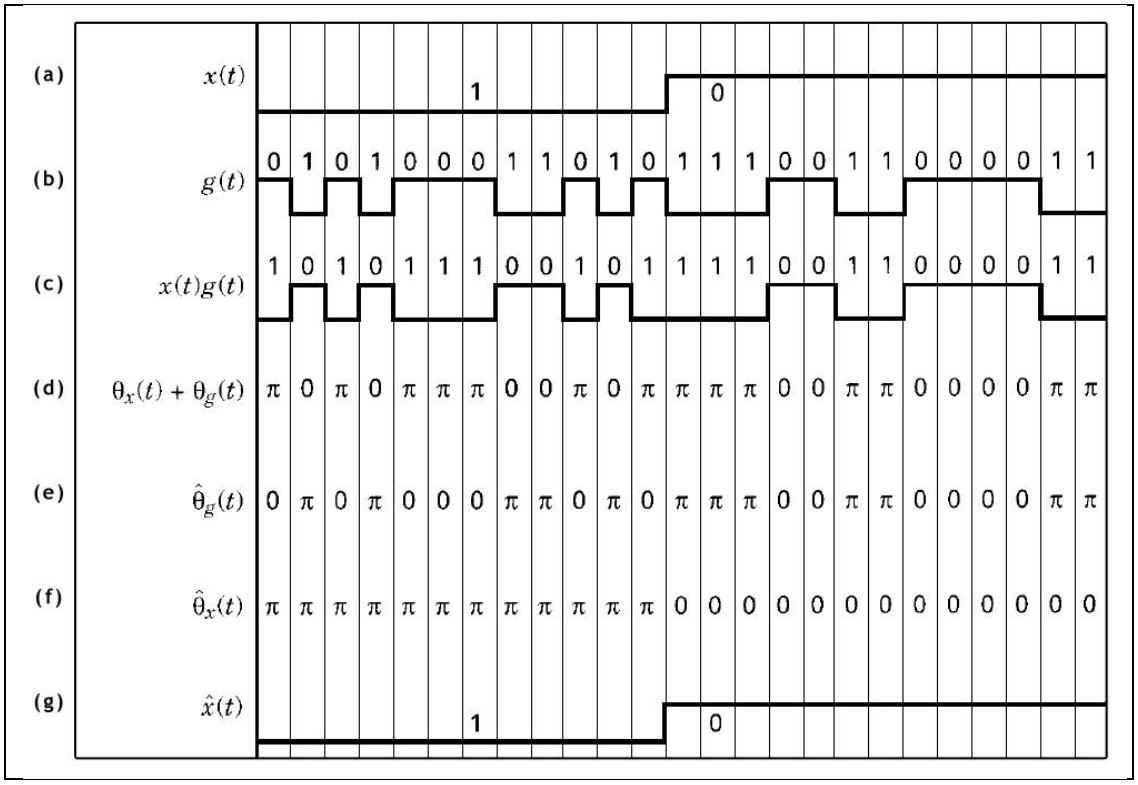
Na metodologia por seqüência direta, mais conhecida como *Direct Sequence Spread Spectrum* (DSSS), uma onda portadora é primeiramente modulada com um sinal de dados, e então esta onda é novamente modulada com um sinal de espalhamento de alta velocidade (SKLAR, 2001, p. 732). Conforme Schütz (2009, p. 45), durante a primeira etapa, o sinal portador binário é modulado em fase com uma portadora senoidal (*Binary Phase Shift Keying* ou BPSK). Já na segunda, a mensagem secreta é modulada por um código pseudoaleatório, que também é utilizado para realizar efetivamente o espalhamento. Desta forma, o dado sigiloso é agregado ao sinal de cobertura.

A seqüência PN, cuja taxa de transmissão é muito maior que a banda do sinal a ser transmitido, deve ser mais rápida do que a seqüência de dados para que ocorra o espalhamento espectral (SCHÜTZ, 2009, p. 51). Sendo assim, o intervalo de um pulso PN, chamado *chip*, é muito menor que a duração de um pulso do sinal a ser modulado.

Após este processo, os *bits* da informação serão representados por uma seqüência modulada pelo código PN. O zero será substituído, por exemplo, por uma cadeia de 16 *bits*, e o um por uma seqüência de outros 16 *bits*.

Durante a detecção, o receptor utiliza uma réplica do código PN para interpretar o sinal recebido, identificando as seqüências de *bits* utilizadas no espalhamento e recuperando a informação original.

A Figura 3 ilustra a aplicação da técnica DSSS durante processo de transmissão, recepção e detecção de um sinal.



Fonte: adaptada de Sklar (2001, p. 734).

Figura 3 – Exemplo de sistema DSSS

Considerando a informação contida na Figura 3a como um sinal binário que deve ser transmitido, na Figura 3b é apresentada a seqüência do código gerado pelo processo de espalhamento. A seguir, pode-se observar a seqüência de dados transmitida (Figura 3c) e sua fase (Figura 3d). As etapas da captura dos dados enviados são listadas a partir da Figura 3e, que representa a fase do código de espalhamento exibida na Figura 3b. A partir deste ponto ocorre o deslocamento de fase (*phase-shift*) produzido pelo receptor. A combinação da fase da portadora, obtida após a decodificação, e do código de espalhamento é ilustrada na Figura 3f. Então, esta informação pode ser convertida novamente em valores binários, resultando no sinal detectado exibido na Figura 3g.

De acordo com Fernandes (2002, p. 20), embora não tenha sido a primeira abordagem de espalhamento espectral, a tecnologia mais usada em sistemas comerciais é a DSSS.

2.2.1.2 Espalhamento espectral por salto em freqüência

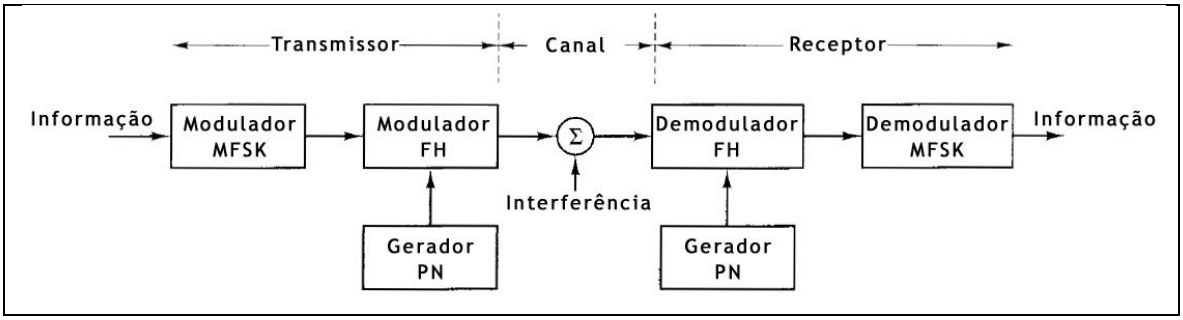
De acordo com Albuquerque, Brazil e Julio (2007, p. 72), a técnica por salto em freqüência, também abordada como *Frequency Hopping Spread Spectrum* (FHSS), codifica o dado secreto segundo um padrão de saltos entre as freqüências do espectro sonoro.

Nesta técnica, o código de espalhamento alimenta um sintetizador de freqüências que gera a portadora da mensagem a ser transmitida, fazendo com que esta varie aleatoriamente dentro da banda de espalhamento. Sendo assim, ao invés da seqüência pseudoaleatória ser responsável por modular um sinal, ela é utilizada para determinar as freqüências geradas pelo sintetizador (SKLAR, 2001, p. 738 - 739).

De posse do mesmo código de espalhamento, o receptor realiza o processo inverso aplicado pelo transmissor. O sinal decodificado então passa por um detector de energia, que indica qual símbolo é equivalente à freqüência recebida nos saltos (SKLAR, 2001, p. 739). O conjunto de símbolos detectados forma a mensagem secreta escondida no áudio.

A modulação mais utilizada com esta técnica é a *M-ary Frequency Shift Keying* (MFSK). Através dela, freqüências do sinal transmitido, escolhidas de forma pseudoaleatória, são alteradas para que representem os símbolos da mensagem.

A Figura 4 representa o processo padrão executado por um sistema FHSS.



Fonte: adaptada de Sklar (2001, p. 739).

Figura 4 – Sistema FH/MFSK

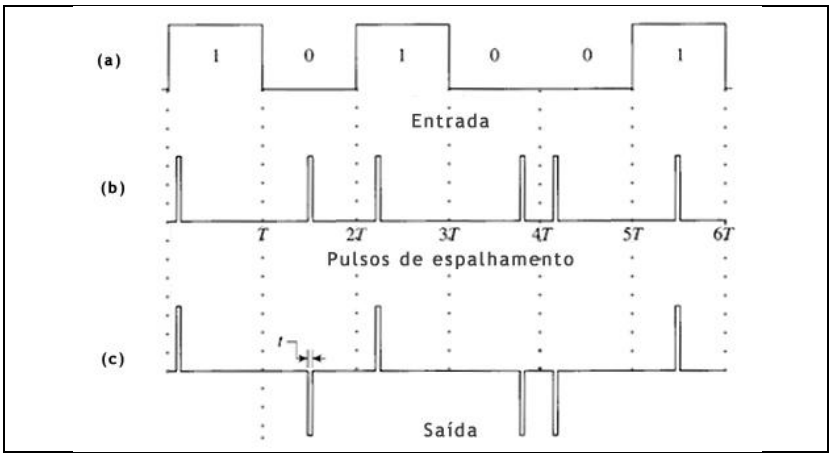
O receptor dos dados reverte o processamento executado durante a transmissão. Novamente, nota-se que a presença de uma cópia do código PN é um requisito crítico para que a mensagem secreta seja compreendida corretamente pelo receptor do sinal, participando ativamente do processo de decodificação do mesmo. Após sua reconstrução, o sinal é analisado pelo demodulador MFSK, que identificará os símbolos recebidos através de um banco de detectores de energia (SKLAR, 2001, p. 738).

2.2.1.3 Espalhamento espectral por salto no tempo

Normalmente referenciada como *Time Hopping Spread Spectrum* (THSS) ou como transmissão por rajada, esta técnica consiste na transmissão da informação através de pequenos blocos de dados do mesmo tamanho. Segundo Cheng et al. (2007, p. 2), a cada

intervalo de tempo é enviada uma rajada de dados para uma das janelas de tempo existentes no período selecionado.

Basicamente, existem duas formas de implementar esta técnica (CHENG et al., 2007, p. 2). Na primeira delas (Figura 5), um gerador de código PN define os intervalos onde são inseridas as informações.

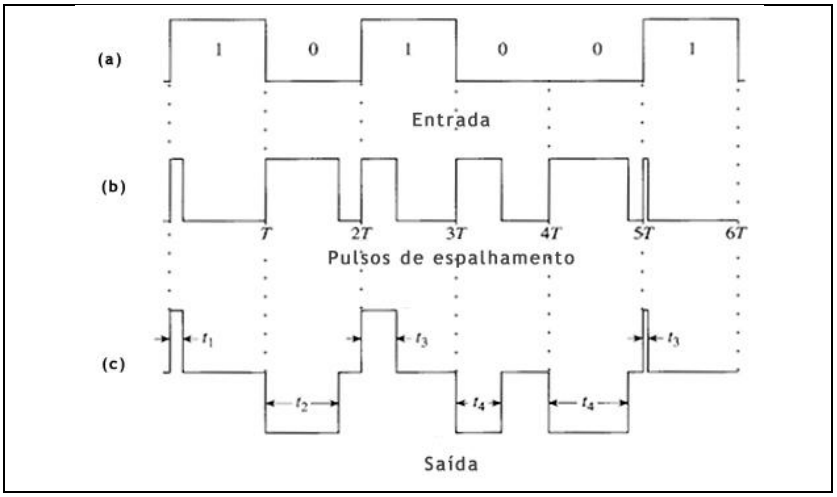


Fonte: adaptada de Cheng et al. (2007, p. 2).

Figura 5 – Sistema THSS utilizando intervalos variáveis

É possível notar na Figura 5b que os pulsos de espalhamento mantêm o mesmo tamanho, porém a freqüência de sua presença no espectro é alterada com o passar do tempo τ. A orientação destes pulsos é mantida quando o bit de entrada vale 1 e invertida se o bit possuir valor 0. Assim, é correto afirmar que a construção da saída do sistema (Figura 5c) é orientada pelo sinal de entrada (Figura 5a) e pelos pulsos gerados.

Opcionalmente, pode-se realizar o THSS de outra forma. Cheng et al. (2007, p. 2) explicam que é possível inserir o *chip* em intervalos de tempo fixos, variando apenas a duração do bloco de informação através de um código PN. A Figura 6 exibe este processo.



Fonte: adaptada de Cheng et al. (2007, p. 2).

Figura 6 – Sistema THSS utilizando blocos de dados de tamanho variável

A variação do tamanho dos blocos de informação espalhados no espectro do áudio é

visível na Figura 6b. Da mesma forma que acontece em sistemas que variam a freqüência dos pulsos, a saída (Figura 6c) é dependente dos *bits* de entrada (Figura 6a) e da seqüência de espalhamento.

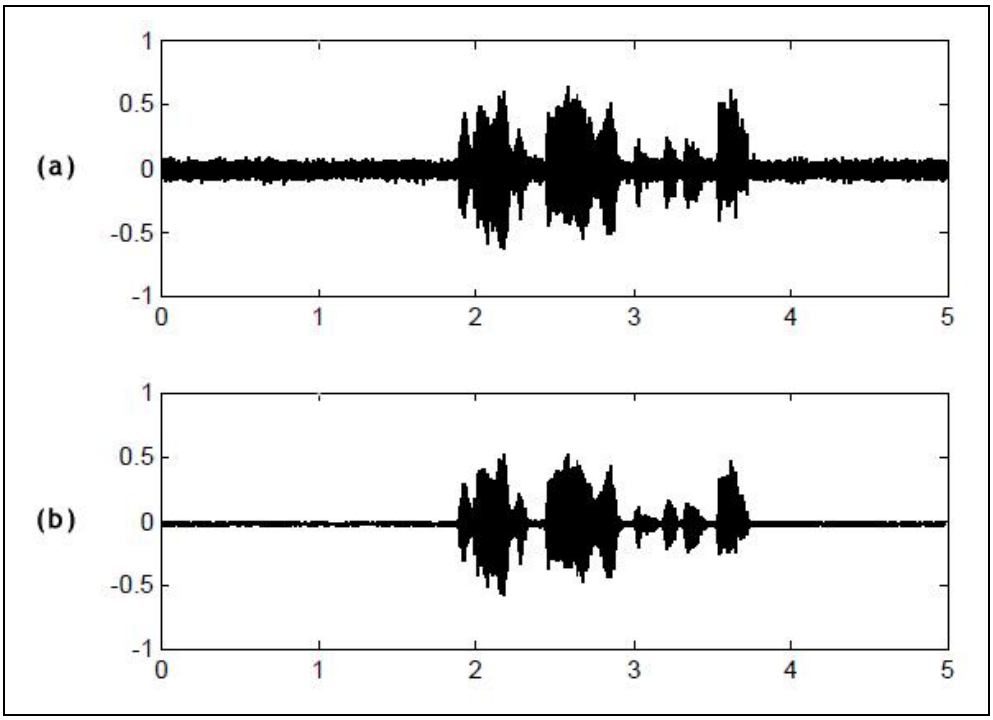
2.2.2 Transformadas

Conforme Schütz (2009, p. 40), é necessário portar o espectro de um sinal do domínio do tempo para o domínio da freqüência antes de analisá-lo. Esta tarefa é realizada através de transformadas, que geram componentes no domínio freqüência a partir de amostras sonoras no domínio tempo. A mais conhecida delas é a Transformada Discreta de Fourier, ou em inglês, D*iscrete Fourier Transform* (DFT), que disponibiliza diversas informações sobre o sinal, como os espectros de magnitude.

Khayam (2003, p. 1) indica que a Transformada Discreta do Cosseno (em inglês, *Discrete Cosine Transform* ou DCT) é outra opção bastante conhecida em processamento digital, com foco na compressão de dados. Ela é uma junção de várias transformadas e permite que o conteúdo presente nas freqüências mais altas do sinal seja descartado, sem com isso gerar degradação significante em sua qualidade. Por este motivo, sua tecnologia vem sendo empregada em padrões como JPEG, MPEG-1, MPEG-2, MPEG-4, H.261 e H.263 (KHAYAM, 2003, p. 28).

Malvar (1999, p. 2) propõe o emprego da *Modulated Complex Lapped Transform* (MCLT), também conhecida como Transformada Superposta Modulada Complexa, em aplicações que envolvem processamento de áudio. Ela é uma ferramenta eficiente para decomposição localizada de sinais, além de permitir a perfeita reconstrução do espectro sonoro. A utilização de suas propriedades em processos de cancelamento de eco e redução de ruído, por exemplo, apresentam bons resultados. Operações que normalmente exigem bancos de filtros complexos podem ser realizadas através de um módulo que aplique esta transformada. Prova disto é a utilização da MCLT em sistemas de codificação de áudio modernos, como o Dolby AC-3.

A Figura 7 exibe o resultado gerado por um redutor de ruído que utiliza a MCLT para realizar a supressão do conteúdo indesejado.



Fonte: adaptada de Malvar (1999, p. 7).

Figura 7 – Redução de ruído com MCLT

A diferença entre o gráfico apresentado na Figura 7a e na Figura 7b evidencia a eficácia da técnica, cuja redução de ruído foi ajustada para cerca de 15 dB. Neste resultado, a relação sinal-ruído (em inglês, *Signal-to-Noise Ratio* ou SNR), responsável por indicar o efeito do barulho de fundo sobre o sinal analisado, foi aumentada de 15 dB para 30 dB. Quanto maior for o valor desta relação, menor é interferência causada pelo ruído (MALVAR, 1999, p. 7).

2.2.3 Jamming

Segundo Garcia (1999, p. 28), a adição intencional de ruídos em um sinal é conhecida como *jamming.* Este tipo de ação é empregado em tentativas de prejudicar a comunicação, degradando a qualidade da transmissão da onda sonora portadora da informação.

Portanto, a arquitetura adotada na construção de receptores de dados suscetíveis a esta problemática deve ser capaz de gerenciar situações desfavoráveis, tornando-se assim mais confiável e segura. Entretanto, conforme Schütz (2009, p. 46), a meta a ser alcançada pelo sistema deve ser de máxima redução das interferências, visto que neste cenário é impossível garantir a ausência de ameaças, pois um agente malicioso pode afetar a transmissão de diversas formas. Dentre elas, as principais são: *jamming* de banda larga, *jamming* de banda

parcial e *jamming* por pulsos. Sklar (2001, p. 759 - 766) explica que, no primeiro caso, o ruído é distribuído por toda a extensão da faixa de freqüências do sinal, ao contrário da interferência de banda parcial, que concentra sua densidade em uma fração limitada do conteúdo que deseja-se adulterar. A terceira opção insere pulsos de ruído com maior potência durante o intervalo finito τ de tempo, embora ele esteja presente de forma moderada ao longo de todo o período.

2.2.4 Sincronização e detecção do sinal

A utilização de técnicas de espalhamento espectral requer a existência de um módulo de sincronização na ponta que fará a recepção do sinal, para que os dados sejam detectados corretamente. Neste extremo do canal de comunicação, uma réplica do código de espalhamento utilizado pelo emissor deve estar disponível, pois é através dela que os dados recebidos serão decodificados (GARCIA, 1999, p. 53). Sklar (2001, p. 745 - 746) indica que este processo normalmente é realizado em duas etapas, chamadas de aquisição e rastreamento.

A aquisição é o estágio onde o alinhamento inicial da réplica do código de espalhamento é feito. Durante esta fase, são realizadas varreduras através de regiões incertas de tempo e freqüência a fim de estabelecer a sincronização entre a seqüência PN local e o sinal recebido. Um procedimento comum neste tipo de método é a realização de cálculos de correlação entre o sinal codificado recebido e o sinal codificado gerado localmente, que resulta em um coeficiente capaz de revelar a similaridade entre estas informações. Quando esta medida atinge determinado limiar, responsável por indicar a existência de sincronismo, o processo de rastreamento é iniciado (SKLAR, 1999, p. 746).

Na etapa de rastreamento, o alinhamento do código é melhorado através de uma estrutura realimentada que regula a fase do sinal de espalhamento no receptor constantemente, mantendo assim a sincronia. Desta forma, a recuperação do sinal original sem espalhamento (*de-spreading*) é possível (SCHÜTZ, 2009, p. 51).

2.3 COMPACTAÇÃO DE DADOS

Essencialmente, compactar dados é colocar a mesma quantidade de conteúdo em um espaço menor, possibilitando que sejam armazenadas grandes quantidades de informações em intervalos limitados (SANCHES, 2001, p. 1). Assim como a esteganografia, a compressão é um estudo proveniente da criptografia, que codifica e decodifica uma informação para torná-la privada aos receptores autorizados.

As técnicas para realizar este tipo de operação classificam-se em dois tipos: *lossy* e *lossless* (do inglês, com perdas e sem perdas, respectivamente). Segundo Sklar (2001, p. 870), alguns tipos de informação (como imagens, áudio e vídeo) permitem que a compressão com perdas seja aplicada sem que haja diminuição perceptível na qualidade do conteúdo. O emprego deste conceito facilita a distribuição de conteúdo multimídia, que pode ser convertido para formatos como JPEG, MPEG e MP3, resultando em arquivos menores do que a fonte original do dado. Entretanto, na compressão de textos deve-se utilizar a compactação sem perdas, pois a alteração do dado original pode comprometer a interpretação da mensagem e ocasionar conseqüências não desejáveis[4].

Dentre os métodos conhecidos, um dos mais referenciados e aperfeiçoados por variações posteriores é o código de Huffman. Desenvolvido por David Huffman, ele é um método estatístico que utiliza a estrutura de árvore binária para codificar textos, sem perda, de forma a obter uma compactação que seja ótima dentro de certos critérios, resultando em um código binário (KOBUSZEWSKI, 2004, p. 30). O alfabeto utilizado pelo algoritmo influencia seus resultados, ou seja, certos domínios podem apresentar taxas de compressão menores que outros (SCHÜTZ, 2009, p. 59).

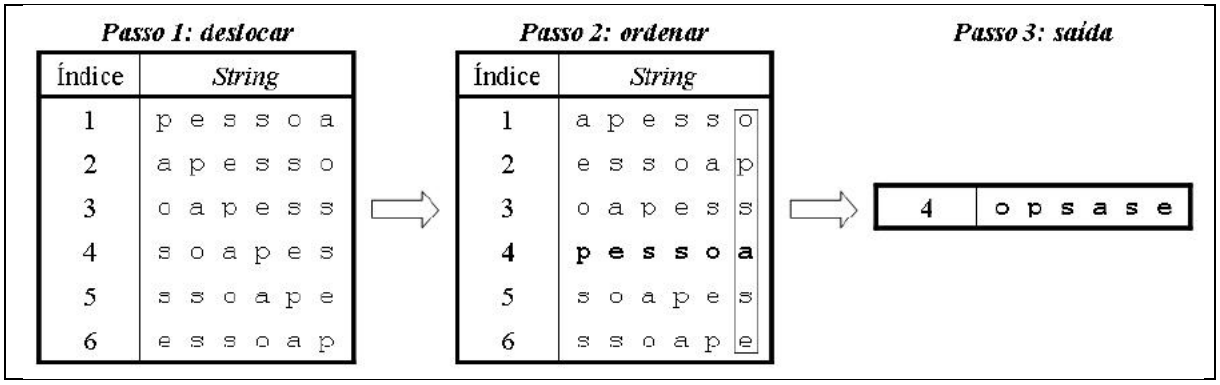
De acordo com Kobuszewski (2004, p. 30) e Schütz (2009, p. 59), a técnica de Huffman não precisa conhecer o caractere que será compactado, porém verifica-se a dependência da informação que indica a probabilidade de ocorrência de todos os caracteres que serão processados. Então, através da listagem dos símbolos em ordem decrescente de probabilidades e agrupamento dos mesmos, este método gera palavras-código em formato binário relativas a cada item do alfabeto de entrada.

Outra opção de implementação é a combinação de algoritmos para a realização da

[4] Largamente utilizada durante a Segunda Guerra Mundial pelas forças do Eixo e dos Aliados, a compressão de dados auxiliava a transmissão de informações cruciais de batalha, como as coordenadas dos pontos que deveriam ser atacados. É possível afirmar que o emprego desta tecnologia foi importante na definição da frente vitoriosa da guerra (SANCHES, 2001, p. 21).

compactação. Quando aliadas, as transformadas de Burrows-Wheeler (em inglês, *Burrows-Wheeler Transform* ou BWT) e *Move-To-Front* (MTF) representam uma alternativa neste sentido.

A BWT é um estudo mais recente que o código de Huffman, que não realiza a compressão dos dados em si, mas por meio do reordenamento dos símbolos que compõe a informação facilita o processo para que outras técnicas, como a MTF, desempenhem este trabalho (BURROWS; WHEELER, 1994, p. 1). Sua definição consiste na construção de uma tabela que contenha todas as rotações possíveis do texto de entrada e, em seguida, a ordenação lexográfica de seus índices. O resultado gerado pelo algoritmo é a combinação de letras da última coluna da tabela e o índice do item que contém a informação em sua ordem original (SCHÜTZ, 2009, p. 68). Este processo é ilustrado na Figura 8.



Fonte: Schütz (2009, p. 69).

Figura 8 – Codificação da palavra “pessoa” utilizando o algoritmo Burrows-Wheeler

A transformada MTF então é aplicada ao resultado fornecido pela BWT, melhorando a eficiência da técnica de compressão utilizada posteriormente, normalmente o código de Huffman (BURROWS; WHEELER, 1994, p. 7). O processo reverso, ou seja, a decodificação da mensagem, emprega o índice fornecido pelo método de compressão para recuperar o texto original. Novamente, uma tabela é construída de acordo com ordenamentos lexográficos, dessa vez guiada pelo bloco de caracteres compactados. No fim do processo, a tabela armazena a mensagem descompactada na posição indicada pelo índice (SCHÜTZ, 2009, p. 69).

Algoritmos heurísticos agrupam opções peculiares nesta área, quando o texto é redigido em um alfabeto limitado de símbolos (COOK, 2005, p. 577). Uma destas opções é a técnica de substituições baseadas em *tokens* (símbolo interno do sistema que deve estar associado a um valor específico). Ela fornece uma solução fundamentada na troca de partes de uma string por marcadores existentes em um dicionário, que ocupam menos espaço. Conseqüentemente, o resultado obtido é uma representação menor do texto original

submetido ao algoritmo (ABEL; TEAHAN, 2005, p. 502).

Neste contexto, Abel e Teahan (2005, p. 500 - 503) indicam que a realização de um procedimento chamado de pré-processamento do texto é uma importante etapa durante a operação de compressão. Esta etapa é responsável por interpretar ou até mesmo gerar o dicionário de *tokens*, antes que a compactação seja efetivamente realizada. Termos representativos do alfabeto e suas respectivas taxas de ocorrências são armazenados de forma correlacionada nesta lista, a fim de evidenciar a relevância de cada elemento dentro da coleção de símbolos.

Ações complementares também podem ser adotadas para auxiliar o processo de compressão heurística. Uma delas é a conversão de letras maiúsculas por suas correspondentes minúsculas, pois desta maneira as palavras do texto analisado se tornam mais equalizadas, facilitando a análise do texto. Em contrapartida, quebras de linhas e pontuações tendem a aparecer próximos às letras neste formato, propriedade que pode ser explorada para melhorar o esquema do algoritmo (ABEL; TEAHAN, 2005, p. 500). Assim, observa-se que regras de manipulação de palavras podem ser adicionadas e modificadas, de acordo com o tipo de texto que será analisado.

2.4 TRABALHOS CORRELATOS

Há um interesse cada vez maior, por diferentes comunidades de pesquisa, no campo da esteganografia. Esta seção destina-se a investigar na literatura alguns dos trabalhos que envolvem algoritmos e/ou técnicas de ocultação de dados.

Os estudos correlatos analisados são: “Protótipo de software para ocultar textos compactados em arquivos de áudio utilizando esteganografia” (KOBUSZEWSKI, 2004), “*Robust audio steganography using direct-sequence spread spectrum technology*” (CHENG et al., 2007), “Sistema de esteganografia em áudio digital que utiliza técnicas eficientes de inserção de dados” (SCHÜTZ, 2009) e “*A genetic-algorithm-based approach for audio steganography*” (ABDULLA et al., 2009).

Kobuszewski (2004) apresentou um protótipo de esteganografia em áudio utilizando a técnica LSB. O software processa áudios digitais no formato WAV e adiciona informações ocultas em alguns segmentos específicos deste tipo de arquivo. Com o objetivo de ocultar uma quantidade maior de texto, foi disponibilizada uma opção ao usuário para compactar a

mensagem utilizando a compressão de Huffman. Segundo o autor, os resultados obtidos foram satisfatórios, pois as alterações feitas no áudio para possibilitar a esteganografia da mensagem são imperceptíveis para o ouvido humano e o processo de extração funcionou de acordo com o esperado durante os testes. Dentre as limitações encontradas neste trabalho, estão a baixa quantidade de informações que podem ser escondidas e o baixo grau de segurança.

Cheng et al. (2007) desenvolveram uma *Dynamic-Link Library* (DLL) capaz de analisar um áudio de entrada, codificar nele uma mensagem escondida e realizar o processo reverso de maneira robusta. A utilização de uma técnica de espalhamento espectral específica (DSSS) foi definida para ocultar informações em sons. Como nenhuma técnica de compressão foi utilizada no projeto, a quantidade de informação que o programa consegue ocultar foi limitada de forma modesta, evidenciando-se como um aspecto negativo deste trabalho.

Schütz (2009) desenvolveu um sistema que permite ocultar, em um arquivo de áudio digital, uma quantidade maior de informações do que sistemas que utilizam métodos convencionais (tais como LSB). De acordo com a pesquisa realizada, a compactação dos dados e a aplicação da técnica DSSS foram fundamentais para o êxito do projeto. Identificou-se neste trabalho que a velocidade de execução e as características do áudio utilizado no processo influenciam diretamente na efetividade do programa.

Abdulla et al. (2009) propõem a utilização de processos inteligentes para selecionar as seções do som que irão mascarar os dados. O estudo indica que ao aplicar algumas características de algoritmos genéticos[5] em sua construção, o método LSB pode ser utilizado como ferramenta segura e poderosa de esteganografia. Apesar de não demonstrar ganhos na quantidade de informações que a técnica consegue ocultar, o estudo prova que a utilização de uma solução convencional também pode ser eficaz, quando modificada apropriadamente.

O quadro 1 mostra de forma resumida as principais características dos trabalhos correlatos relacionados, tendo como base critérios importantes extraídos a partir dos conceitos descritos.

As informações foram dispostas em colunas, representando na vertical os trabalhos analisados e as linhas apresentam as características de cada sistema, indicando semelhanças e/ou diferenças.

[5] De acordo com Pacheco (1999), é um tipo de algoritmo probabilístico, capaz de fornecer mecanismos de busca paralela e adaptativa baseado no princípio de sobrevivência dos mais aptos da espécie e na reprodução entre eles, proposto por Charles Darwin.

características / trabalhos relacionados	Kobuszewski (2004)	Cheng et al. (2007)	Schütz (2009)	Abdulla et al. (2009)
linguagem de alto nível	X	X	-	-
algoritmos de compressão	X	-	X	-
robustez da técnica de esteganografia	-	X	X	X
alta capacidade de ocultação de texto	-	-	X	-
alta velocidade de processamento	-	X	-	X

Quadro 1 - Características dos trabalhos relacionados

A partir do quadro 1 tem-se o estado da arte sobre esteganografia em áudio. É possível observar que a dificuldade e complexidade na implementação limitam o uso de técnicas mais efetivas para a ocultação de dados e o processamento de alto desempenho. Além disso, a resistência da mensagem a ataques e alterações nem sempre é considerada, por exigir a utilização de métodos esteganográficos mais robustos para garantir a confiabilidade das informações escondidas.

Apesar dos relatos indicarem que os sistemas apresentados atingiram os objetivos propostos, nenhum deles combinou técnicas robustas de esteganografia com algoritmos específicos para tratamento de textos curtos, típicos em marcas d'água. Aliado a este fato, a velocidade de processamento do programa não foi considerada como um fator importante nos trabalhos relacionados. A utilização de linguagens de programação com alto nível de abstração foi observada em dois trabalhos, porém estes casos restringiram-se ao emprego de soluções de segurança pouco elaboradas.

Para suprir tais demandas, o sistema apresentado neste trabalho foi desenvolvido guiado por estes requisitos, comumente exigidos em aplicações com finalidades comerciais.

3 DESENVOLVIMENTO

Neste capítulo são apresentadas as etapas do desenvolvimento do sistema de esteganografia, através da especificação de requisitos, modelagem de dados e a implementação do software. Em seguida, são listados e discutidos os resultados obtidos com a aplicação proposta.

3.1 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

A fim de ocultar marcas d'água em arquivos de áudio, o software desenvolvido neste trabalho utiliza conceitos do processamento de sinal em conjunto com uma técnica heurística de compactação de dados.

Neste cenário são identificados dois tipos de usuários, o emissor e o receptor. Enquanto o primeiro é responsável por esconder a informação no espectro sonoro de um áudio, o segundo está apto a visualizar a mensagem que foi oculta neste tipo de mídia.

Os requisitos funcionais (RF) e não funcionais (RNF) do sistema de esteganografia desenvolvido são:

a) ocultar um texto em um arquivo de áudio, ambos definidos pelo usuário, gerando uma cópia do segundo com as alterações realizadas (RF);
b) disponibilizar uma opção de compressão do texto informado, possibilitando assim a ocultação de uma maior quantidade de dados no som escolhido (RF);
c) analisar um arquivo de áudio e verificar a existência de mensagens ocultas pelo programa (RF);
d) extrair a mensagem escondida no áudio, realizando a descompressão dos dados quando necessário (RF);
e) utilizar técnicas de compressão de dados eficientes, cuja metodologia seja otimizada para compactação de textos curtos (RNF);
f) utilizar uma técnica de espalhamento espectral no processo de esteganografia da mensagem no áudio digital (RNF);
g) realizar a ocultação da informação sem adicionar ruído excessivo no arquivo gerado (RNF).

Para atender as condições, optou-se pela utilização da técnica de espalhamento espectral por seqüência direta (item f), conhecida por possuir características favoráveis à construção de ambientes robustos para ocultação de dados, em conjunto com um algoritmo heurístico de compactação direcionado à análise de textos curtos (itens b, d, e). A transformada MCLT foi escolhida para realização da avaliação do áudio (item c).

Buscando garantir que ruídos indesejáveis não fossem adicionados ao arquivo de áudio após a agregação da marca d'água (item g), um modelo psicoacústico foi utilizado para analisar as porções do som que ocultam os dados. Ele é responsável por alertar a rotina de inserção de dados para eventuais problemas deste gênero.

3.2 ESPECIFICAÇÃO

A especificação do sistema foi realizada através da metodologia de orientação a objetos. Ela foi representada em diagramas da *Unified Modeling Language* (UML), utilizando a ferramenta *Enterprise Architect*.

Inicialmente, foi empregado o diagrama de casos de uso, seguido pelo diagrama de classes e pelo diagrama de atividades. Por fim, os diagramas de seqüência detalham a utilização do sistema.

3.2.1 Diagrama de casos de uso

A Figura 9 exibe o diagrama com as ações disponíveis para os usuários no sistema de esteganografia desenvolvido.

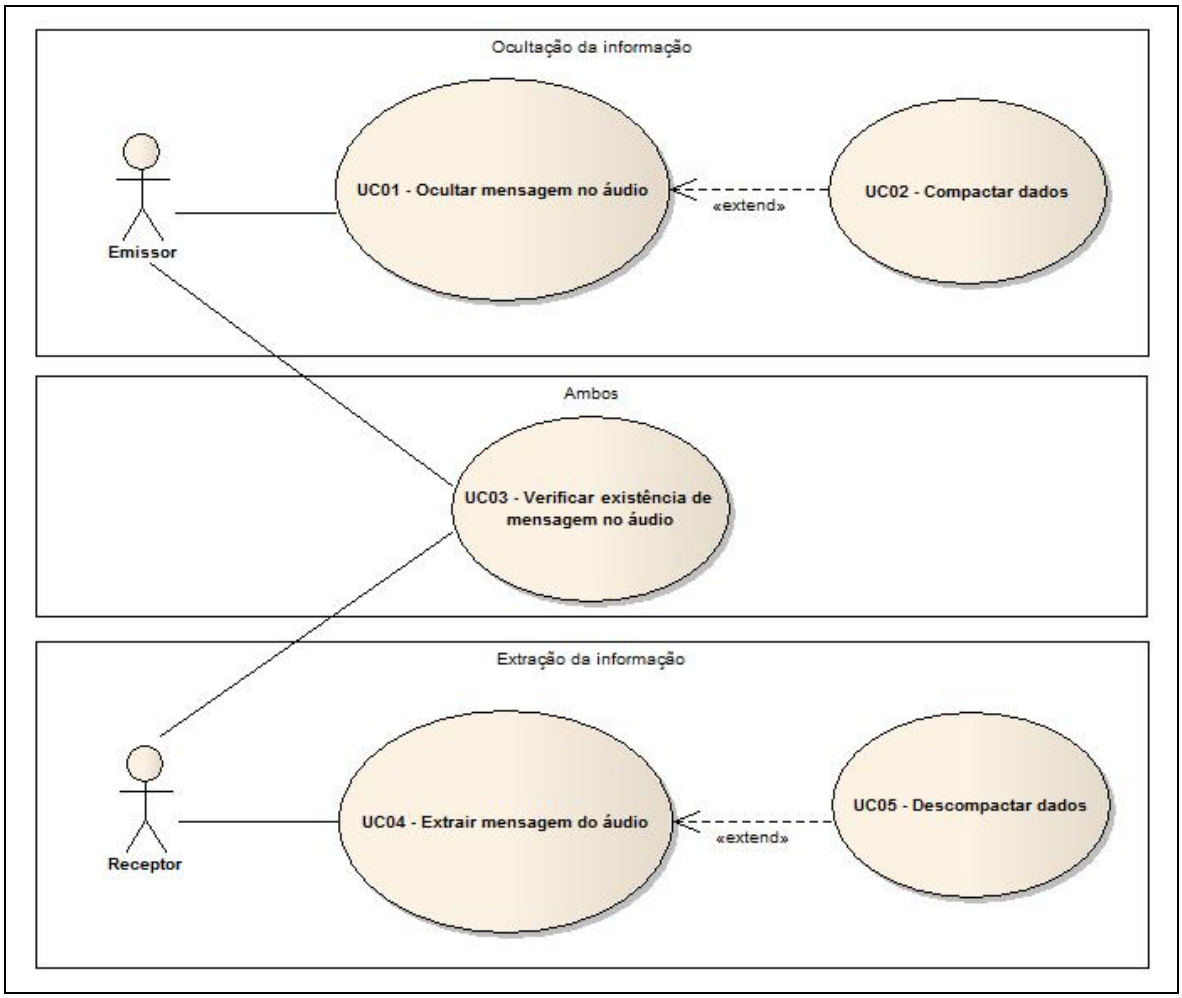


Figura 9 – Diagrama de casos de uso

Os casos de uso da aplicação são descritos a seguir:

a) UC01 – ocultar mensagem no áudio: permite a inserção de uma mensagem secreta no arquivo de áudio;
b) UC02 – compactar dados: permite a compressão da mensagem antes de ocultá-la no arquivo de áudio;
c) UC03 – verificar existência de mensagem no áudio: permite a análise do arquivo de áudio, afim de descobrir informações ocultas em seu espectro sonoro;
d) UC04 – extrair mensagem do áudio: permite a leitura da mensagem escondida em um arquivo de áudio;
e) UC05 – descompactar dados: permite a descompressão da mensagem oculta, quando a mesma tiver sido compactada no processo de inserção de dados.

3.2.2 Diagrama de classes

Responsável por fornecer o panorama geral do relacionamento entre as classes do projeto, o diagrama de classes é apresentado em subseções para facilitar o entendimento da estrutura construída.

3.2.2.1 Inserção e extração da marca d'água

A Figura 10 apresenta o diagrama de classes necessárias para o processo de inserção e extração da marca d'água.

A classe `Append` coordena a inclusão da mensagem, definindo limiares de freqüência, calculando a quantidade de blocos de processamento do arquivo, recuperando possíveis pontos de inserção de *chips*, dentre outras operações.

As operações de varredura do som à procura de dados ocultos, que incluem a recuperação de pontos de decodificação e cálculos de correlação armazenados em objetos do tipo `DataCorrelation`, são comandadas pela classe `Detect`. Ela também é responsável por gerenciar a análise de um áudio durante a verificação que determina a existência ou não de uma marca d'água em seu espectro, além de indicar se a mensagem está compactada.

Durante o processo de extração ou análise de uma marca d'água, os dados são carregados e manipulados em blocos de amostras através de uma instância da classe `DataBuffer`, que também armazena a posição dos pontos inaudíveis do som em sua estrutura.

Os possíveis estados do algoritmo de extração são organizados na entidade `EDetectionState`. Eles são alternados de acordo com os resultados obtidos nas análises do *buffer* de dados. `DetectionResult` é a classe que dispõe destes resultados, além de outros dados periféricos, como a etapa do processamento onde ele foi recuperado (`ProcessTime`).

`MCLT` é a classe que agrupa funções relativas à transformada utilizada durante a análise e alteração do áudio. A classe `SpreadSpectrum` possui operações de seleção e identificação das áreas onde os dados serão ou podem ter sido inseridos. Ambas possuem papel fundamental durante a inclusão e a detecção da marca d'água, pois suas instâncias definem as seções do espectro sonoro onde os dados são escondidos. As estruturas `SpreadSpectrumInsertBand` e `SpreadSpectrumExtractBand` auxiliam estes algoritmos a organizar e armazenar os pontos do som de interesse para a ferramenta.

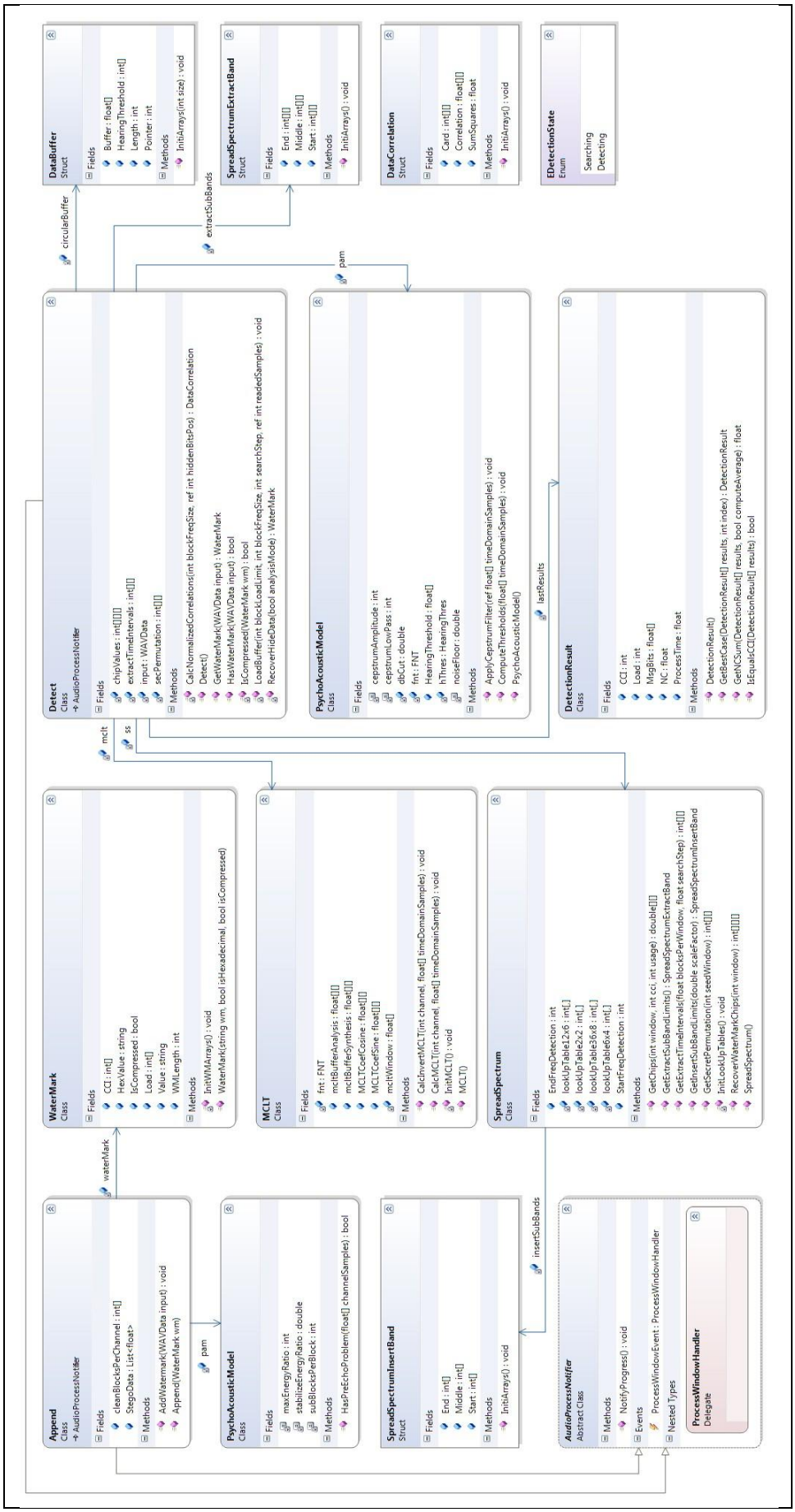


Figura 10 – Diagrama de classes para inserção e extração da marca d'água

A classe abstrata `AudioProcessNotifier` define o comportamento padrão das entidades de inserção e extração para notificação de progresso do processamento.

As informações referentes à marca d'água são armazenadas em instâncias da classe `WaterMark`. Os atributos `CCI` e `Load` desta classe representam as partes hexadecimais dos símbolos da mensagem secreta, pois o algoritmo de espalhamento espectral oculta cada uma das partes de um valor nessa base em uma freqüência diferente, visando aumentar a segurança do sistema. Como esta entidade é alvo de interesse de ambos os procedimentos agrupados nesta seção, ela é compartilhada entre eles.

Duas classes representam o modelo psicoacústico da ferramenta. Organizadas em escopos diferentes, ambas possuem o nome `PsychoAcousticModel`, embora possuam atribuições distintas. A classe pertencente ao domínio da inserção de dados indica a existência de variações bruscas de energia nos blocos de áudio, enquanto a estrutura utilizada durante a extração de dados trata o som e detecta oscilações que denunciam a presença de informações ocultas em seu espectro.

3.2.2.2 Compressão dos dados

O diagrama da Figura 11 apresenta o conjunto de classes que realiza a compressão e descompressão da mensagem oculta no áudio.

A classe `CodeBookLoader` carrega os símbolos utilizados pelos algoritmos heurísticos de compressão e descompressão, pois estes estão gravados em arquivos de texto para uma melhor manutenção do sistema.

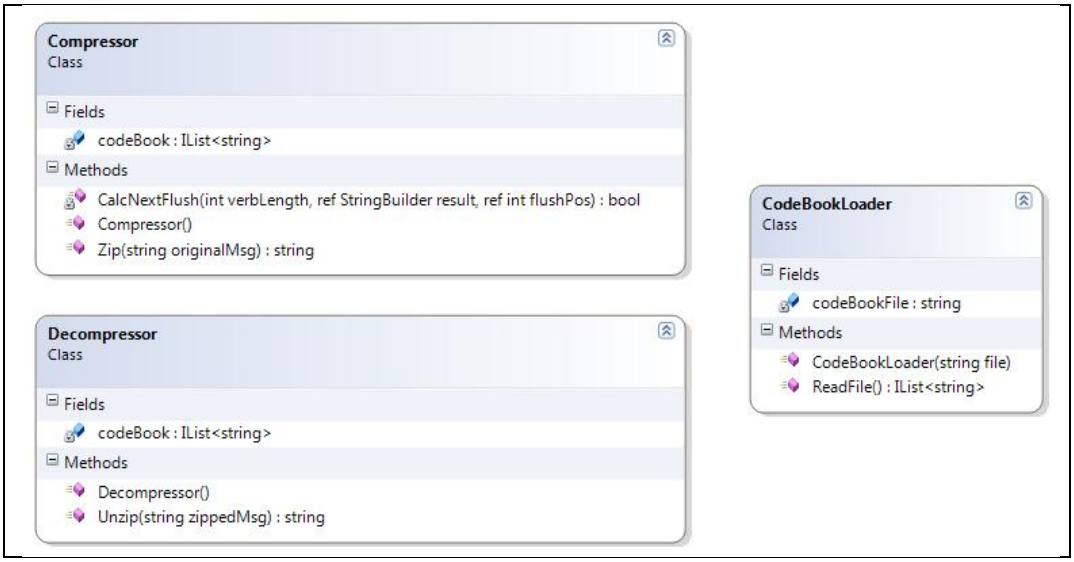


Figura 11 – Diagrama de classes para compactação e descompactação de textos

`Compressor` e `Decompressor` são as classes responsáveis pela compactação (método `Zip`) e descompactação (método `Unzip`) da marca d'água do áudio, respectivamente.

3.2.2.3 Arquivo de áudio

A configuração dos arquivos de entrada e saída é um fator importante para que o sistema cumpra seu objetivo corretamente. A Figura 12 fornece o diagrama que agrupa as classes responsáveis pela manipulação de arquivos WAV do sistema.

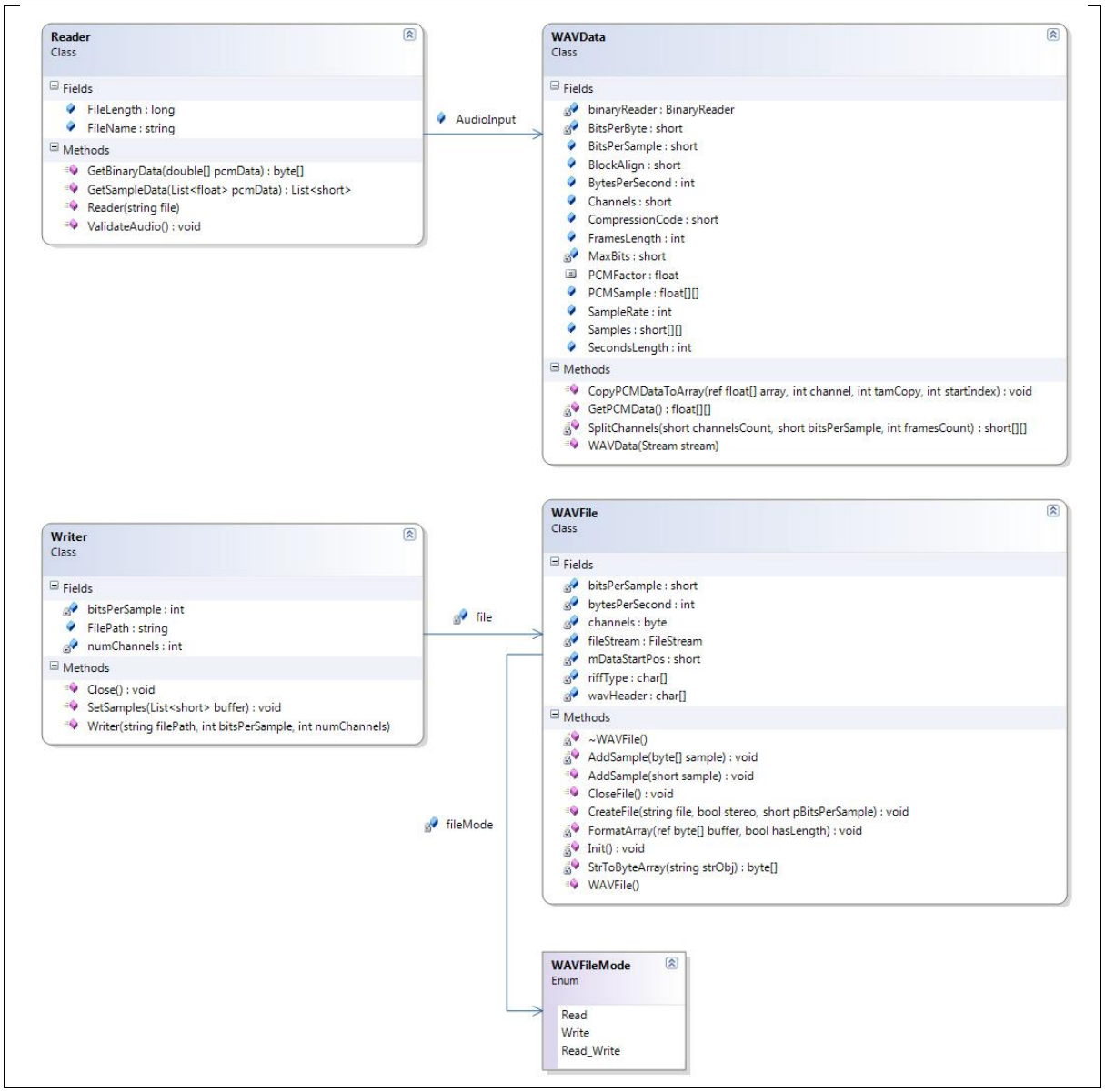


Figura 12 – Diagrama de classes para leitura e escrita do arquivo de áudio

As classes `Reader` e `WAVData` são utilizadas constantemente durante as operações realizadas pela ferramenta. Desde a análise inicial do áudio até o processo de inserção e

extração de dados, valores armazenados em instâncias destas classes possibilitam o estudo do espectro sonoro e disponibilizam a estrutura necessária para a alteração dos mesmos. Dados como quantidades de canais, comprimento em segundos e amostras em formato PCM (*Pulse Code Modulation*) estão acessíveis através delas, além de conversões de amostras para outras bases de codificação.

Após a inserção da marca d'água no áudio, o sistema gera um novo arquivo que contém uma mensagem oculta em sua composição. Esta operação é iniciada através da classe `Writer`, responsável por criar uma instância da entidade `WAVFile` e definir os valores das amostras de som do novo arquivo WAV. O modo de acesso a arquivos manipulados por esta classe é definido pelo enumerador `WAVFileMode`.

3.2.2.4 Estruturas auxiliares

A fim de melhorar a compreensão dos grupos de entidades que formam o sistema, as classes auxiliares são representadas em um diagrama à parte, exibido na Figura 13. Para facilitar o acesso aos métodos e parâmetros mantidos por estas estruturas, todas são classes estáticas.

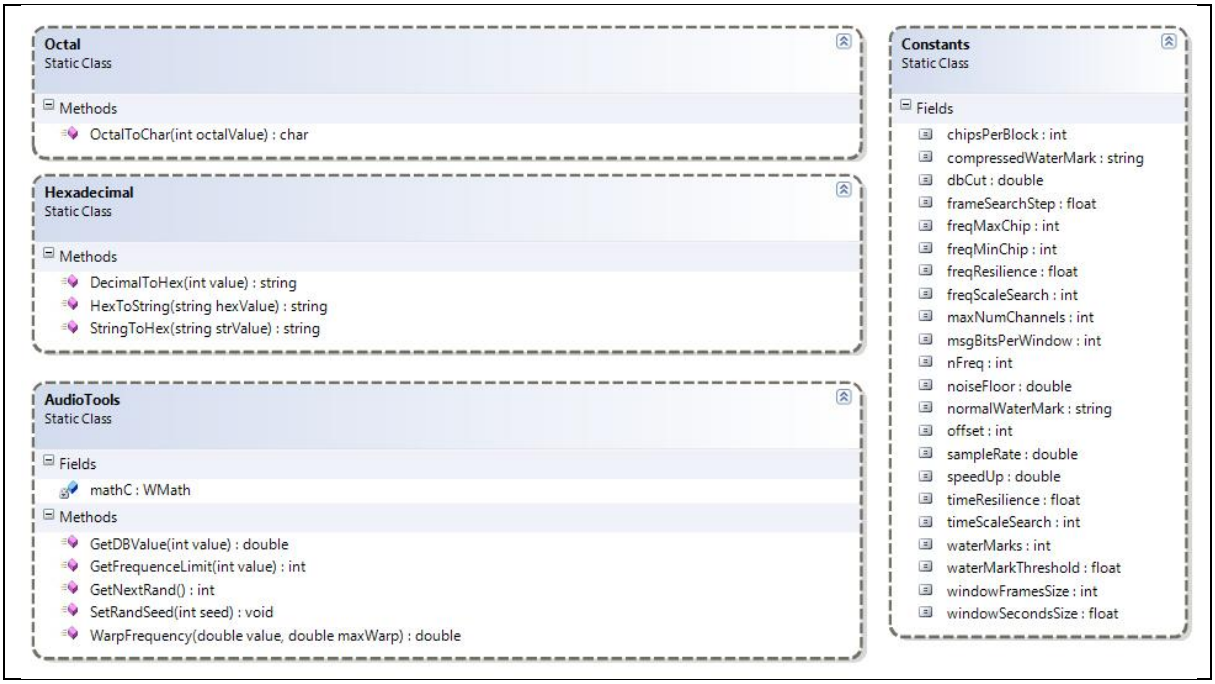


Figura 13 – Diagrama de classes auxiliares do sistema

A principal delas é a classe `Constants`. Nela são definidos valores utilizados em várias entidades do sistema, como a freqüência máxima e mínima onde os *bits* da marca d'água são

inseridos e as configurações que o arquivo WAV deve possuir para ser considerado compatível com o programa.

Na classe `AudioTools` estão agrupados alguns procedimentos matemáticos relativos à análise de áudio, que também são acessados por diversos objetos durante a execução do sistema.

As classes `Hexadecimal` e `Octal` efetuam operações de conversão para as bases indicadas em suas nomenclaturas. Enquanto a primeira é utilizada durante a adição e a recuperação da marca d'água, a segunda auxilia a classe `CodeBookLoader` a carregar a lista de símbolos do algoritmo de compactação.

3.2.3 Diagrama de atividades

O diagrama de atividades é responsável por representar os estados de uma computação. A Figura 14 apresenta os passos seqüenciais disponíveis no sistema para que uma mensagem secreta seja inserida e extraída de um arquivo de áudio.

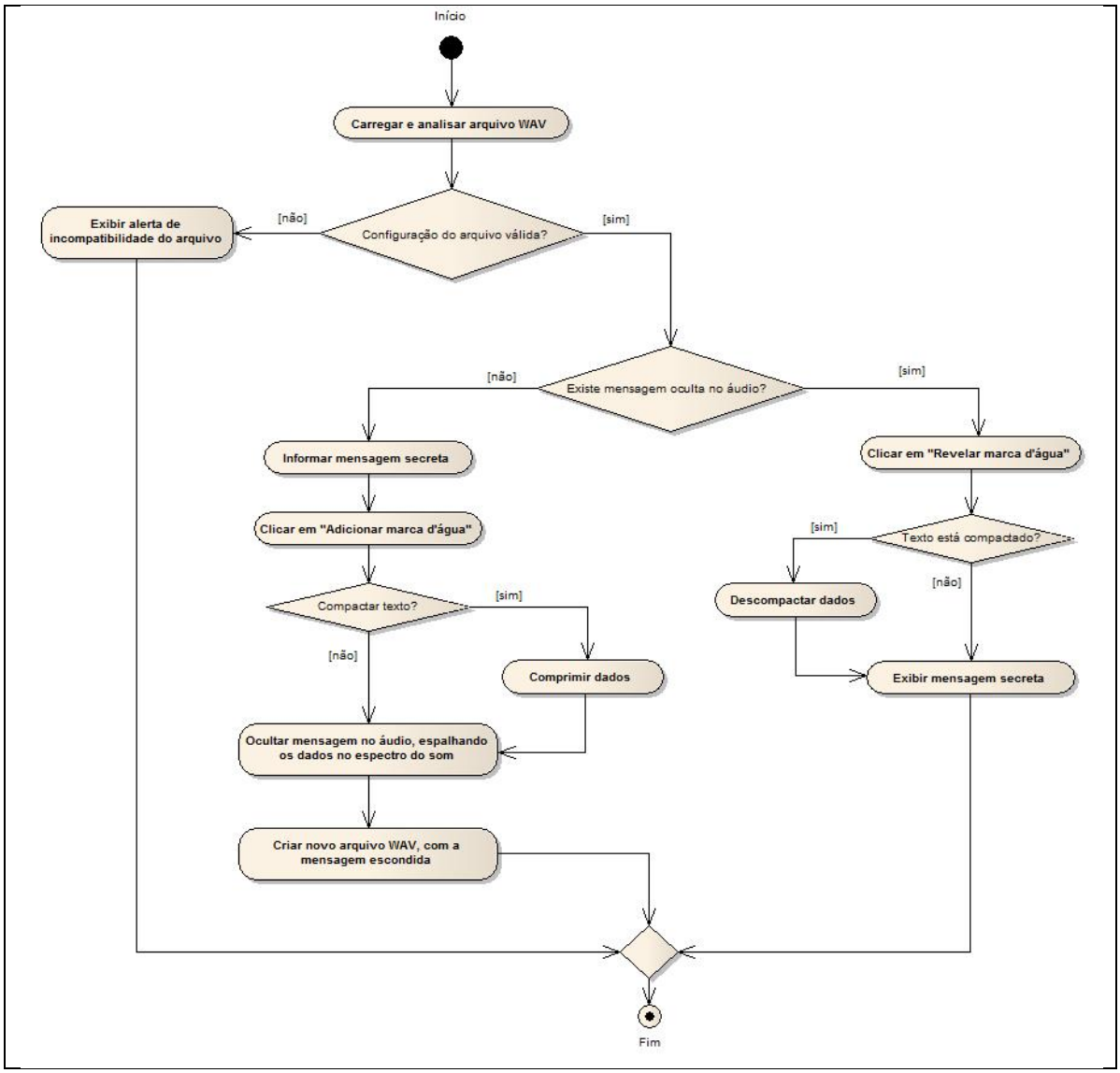


Figura 14 – Diagrama de atividades

A primeira operação realizada é a análise do arquivo, que verifica se o áudio é compatível com o sistema. Em caso afirmativo, uma segunda interpretação do conteúdo é necessária para que o programa defina qual funcionalidade (inclusão ou extração de dados) será disponibilizada.

Na inserção de uma marca d'água, o usuário deve informar um texto, indicar se o mesmo deve ser compactado e acessar o botão "Adicionar marca d'água". Esta ação inicia o processo que esconde a informação no espectro do áudio. O resultado do procedimento é a criação de um novo arquivo WAV com uma informação oculta em sua estrutura.

Caso já exista uma mensagem no áudio, o usuário deve acessar o botão "Revelar marca d'água" para que a marca d'água contida no arquivo seja revelada. Antes de exibir o texto, a aplicação descompacta a informação, quando esta operação é necessária.

3.2.4 Diagrama de seqüência

Diagramas de seqüência demonstram como grupos de objetos colaboram durante a realização de um processo. Neles são registradas as trocas de mensagens entre as classes envolvidas em um caso de uso. Assim, a seguir são mostrados os diagramas de seqüência dos casos especificados para o sistema.

3.2.4.1 Análise do arquivo

O diagrama equivalente ao caso de uso UC03 – verificar existência de mensagem no áudio é apresentado na Figura 15.

O processo de análise do áudio inicia após o usuário ter selecionado o arquivo desejado, através da tela principal do sistema. A partir deste ponto, a classe `Reader` comanda a leitura dos dados presentes no arquivo, instanciando um objeto do tipo `WAVData`. Nesta classe as informações são interpretadas e categorizadas, assim que esta tarefa é distribuída entre o construtor do objeto e os métodos `SplitChannels()` e `GetPCMData()`.

Com a obtenção dos dados do áudio é possível então definir se a configuração do arquivo é suportada pelo sistema. Esta verificação é feita pelo método `ValidateAudio()`.

Caso positivo, uma instância da classe `Detect` é criada para que sejam efetuadas varreduras no espectro sonoro do arquivo. Assim, a execução do método `HasWaterMark()` é realizada. Esta operação tem como objetivo encontrar evidências que denunciem a existência de marcas d'água escondidas na estrutura do som analisado. Neste procedimento, o método `RecoverHideData()` é utilizado parcialmente , já que apenas uma parte relevante do som deve ser estudada em busca destas evidências.

Por fim, o resultado da análise é enviado para a interface do sistema, onde o usuário é notificado quando uma mensagem secreta está presente no áudio.

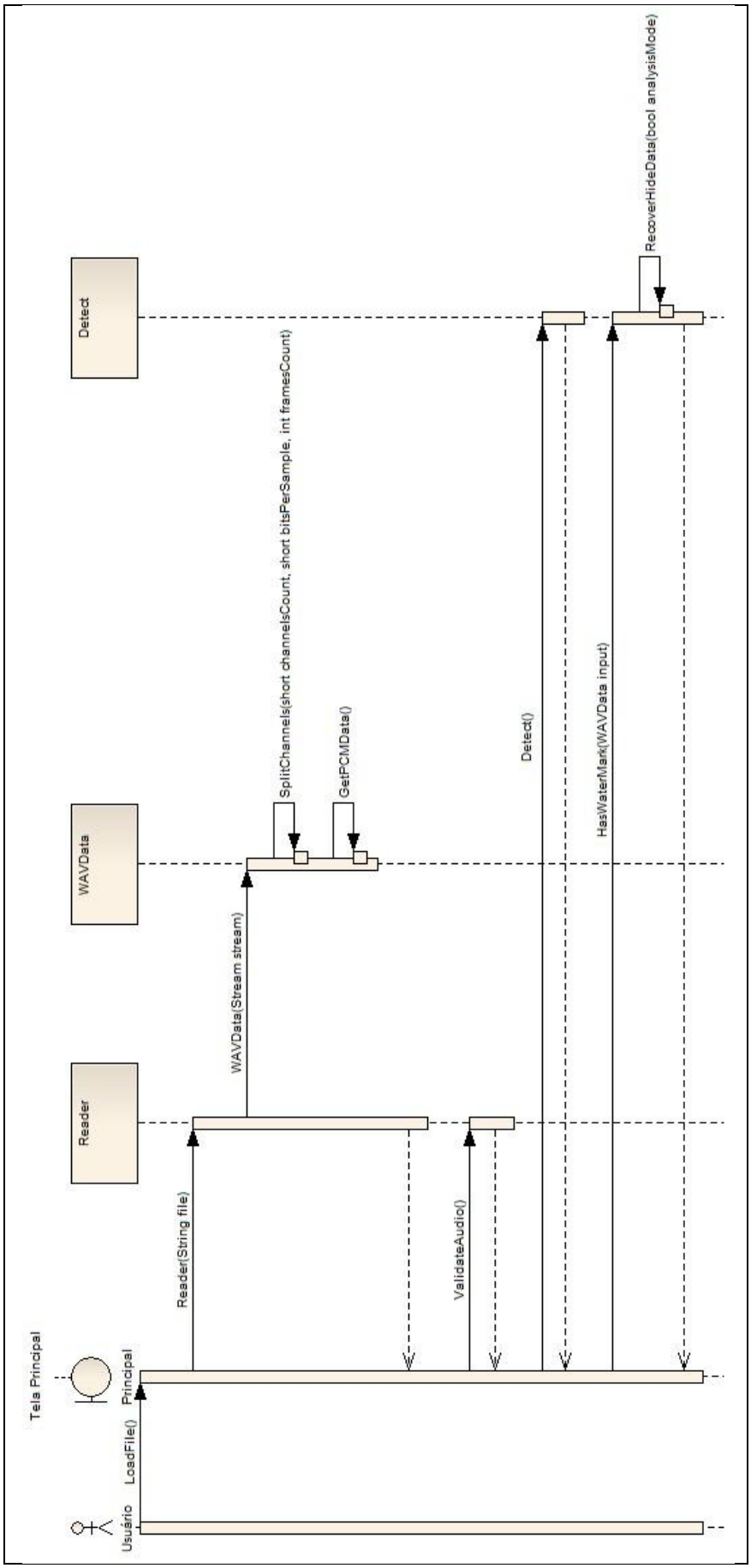


Figura 15 – Diagrama de seqüência “Análise do arquivo”

3.2.4.2 Inserção da marca d'água

O processo de adição de marca d'água no áudio selecionado pelo usuário é composto por dois casos de uso complementares. Para melhor entendimento, os diagramas de seqüência foram divididos entre a Figura 16 e a Figura 17. As rotinas de inserção e de criação do arquivo físico são exibidas na primeira e na segunda imagem, respectivamente.

Durante o caso de uso principal, UC01 – ocultar mensagem no áudio, o emissor aciona o método `HideWaterMark()` através da tela principal do sistema, iniciando assim as etapas para a ocultação da informação. Neste momento é criado o objeto que contém as informações da marca d'água, representado por uma instância da classe `WaterMark`. Esta entidade então é enviada para a classe `Append`, responsável por gerenciar os procedimentos necessários para que o item seja escondido no som, organizado neste ponto como um objeto `WAVData`.

Antes de adicionar o ruído que representará a informação no áudio, é necessário estimar as freqüências que serão empregadas nesta ação, tamanho dos blocos de processamento, entre outros fatores de controle. Além de agrupar estes indicativos como variáveis internas, o método `AddWatermark()` utiliza duas outras classes (`SpreadSpectrum` e `MCLT`) em suas operações mais críticas.

De acordo com o bloco do áudio que está sendo processado, o método `GetChips()` de um objeto `SpreadSpectrum` gera os intervalos onde o sistema realizará pequenas alterações a fim de representar o dado mascarado. De posse desta informação e dos blocos de dados do arquivo (obtidos através da função `CopyPCMDataToArray()` da classe `WAVData`), o programa utiliza transformadas para a inclusão da mensagem. Para desempenhar este papel, os métodos `CalcMCLT()` e `CalcInvertMCLT()` da classe `MCLT` são empregados.

A compressão da mensagem secreta, descrita no caso de uso UC02 – compactar dados, é uma operação opcional. Quando comandada pelo emissor, esta ação é realizada antes da inserção do texto no áudio. O procedimento é efetuado por uma instância da classe `Compressor`, através do método `Zip()`.

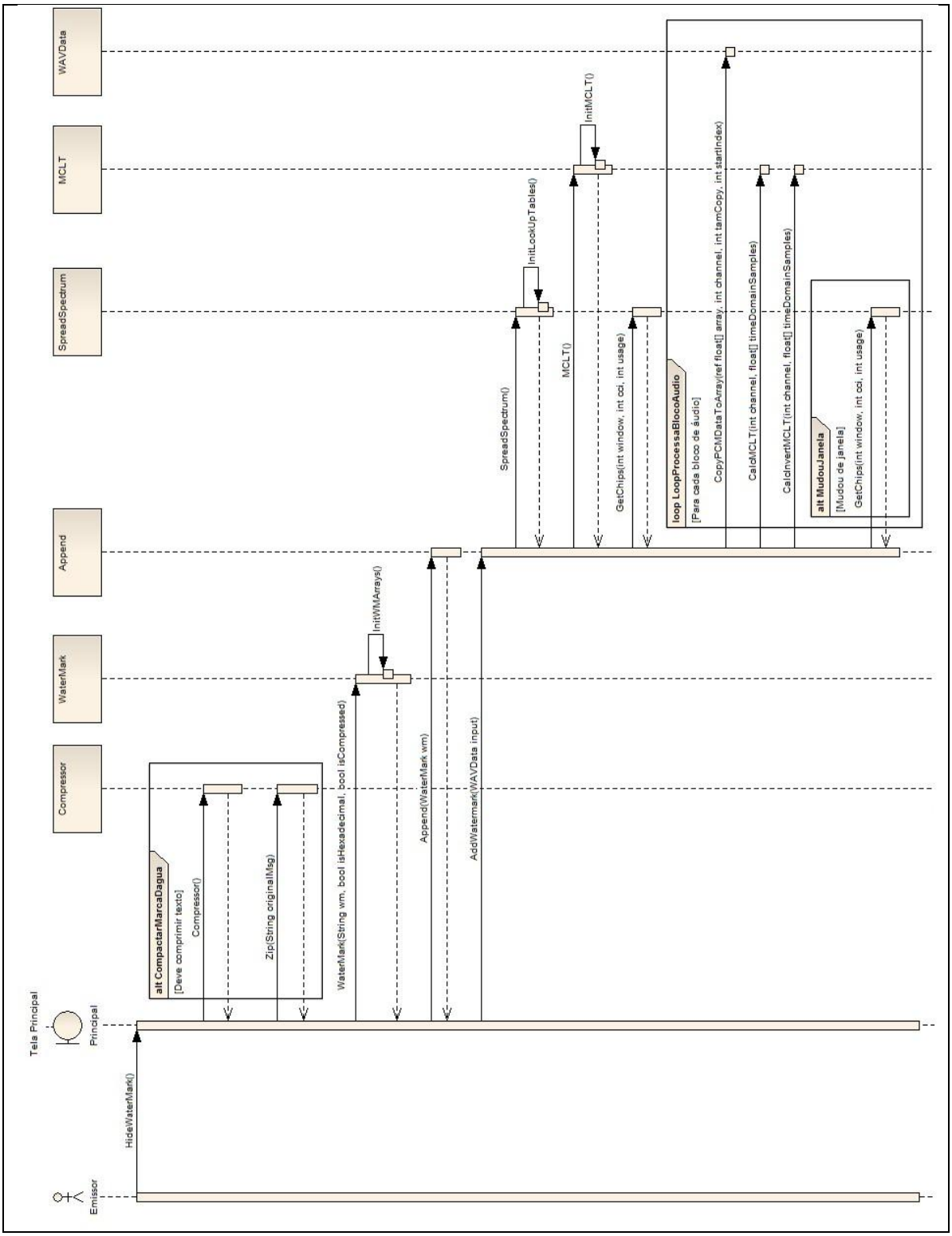


Figura 16 – Diagrama de seqüência “Inserção da marca d’água”

Após a inserção lógica da mensagem no espectro sonoro, o sistema deve criar fisicamente o arquivo. Para tanto, a informação recuperada pelo método `StegoData()` é convertida do formato PCM amostras de áudio WAV, através do método `GetSampleData()` da classe `Reader`.

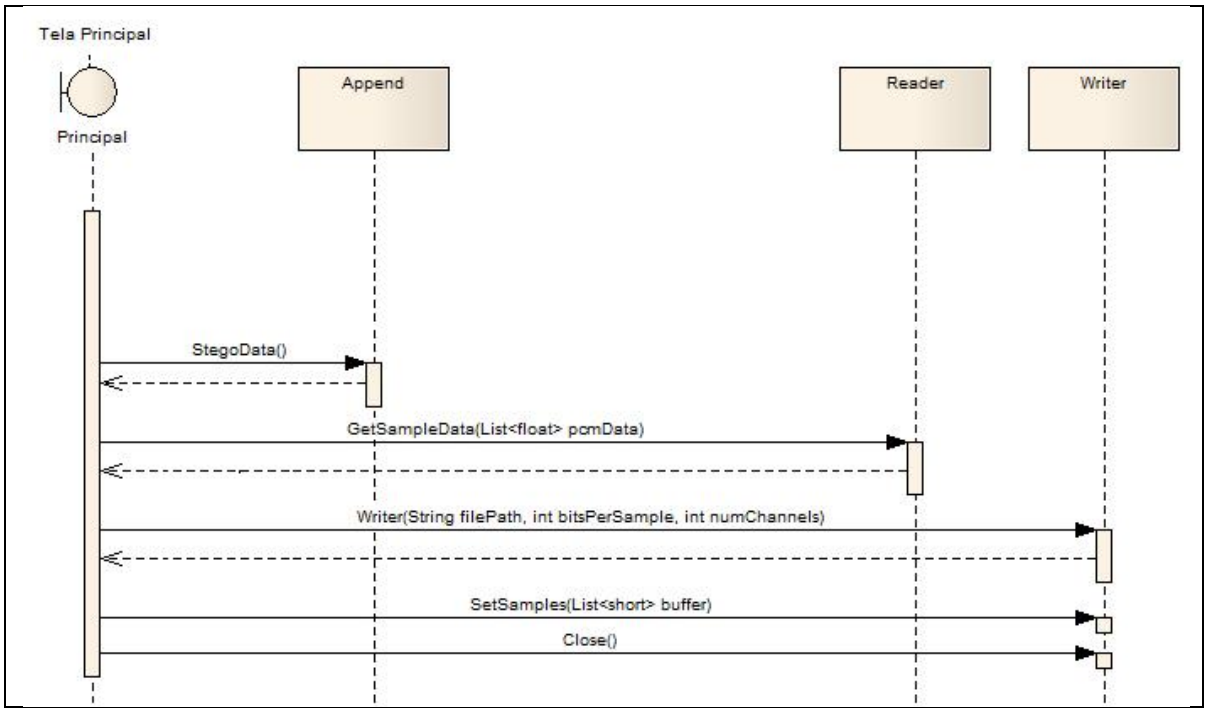


Figura 17 – Diagrama de seqüência “Criação do arquivo”

O arquivo com a mensagem oculta é criado de fato por um objeto da classe `Writer`. O método `SetSamples()` atribui os dados e cria o cabeçalho do novo áudio, enquanto o método `Close()` finaliza sua produção.

3.2.4.3 Extração da marca d’água

O diagrama da Figura 18 contempla o caso de uso executado durante o processo de detecção e extração da marca d’água, UC04 – extrair mensagem do áudio. Nele, o receptor da mensagem inicia o procedimento ao invocar o método `RevealWaterMark()` da tela principal. Uma instância da classe `Detect` gerencia as etapas que extrairão as informações ocultas do áudio, portanto os métodos `GetWaterMark() e RecoverHideData()` são acessados.

A análise do espectro sonoro é realizada em várias escalas de freqüência e tempo. Estes pontos de decodificação, que auxiliam na definição dos limites da procura por dados escondidos, são recuperados pela classe `SpreadSpectrum`, através dos métodos `GetExtractSubBandLimits()` e `GetExtractTimeIntervals()`.

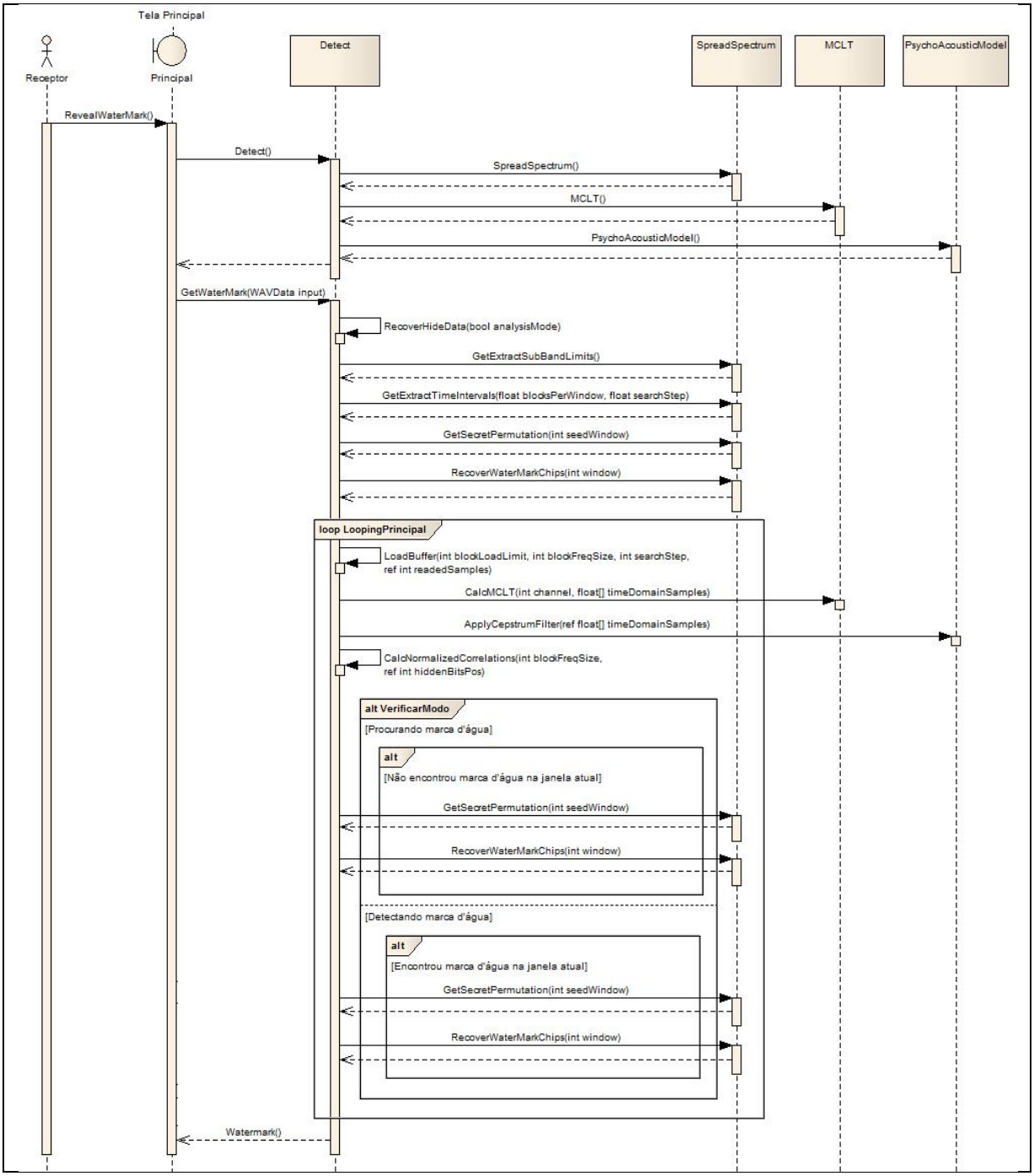


Figura 18 – Diagrama de seqüência “Extração da marca d’água”

Tendo em vista que a inserção da mensagem é realizada para que seja resistente à ataques externos, durante a extração dela é necessário simular dados que possibilitem o rastreamento da informação. Assim, as várias posições onde os *bits* da mensagem podem estar espalhados e os intervalos entre estas marcações são calculados pelas funções `GetSecretPermutation()` e `RecoverWaterMarkChips()`. Estas informações são atualizadas toda vez que o sistema passa a analisar um bloco diferente do áudio.

Durante a iteração principal do caso de uso, a porção sonora que será processada é carregada pelo método `LoadBuffer()`, que utiliza as funções `CalcMCLT()` (classe `MCLT`) e

`ApplyCepstrumFilter()` (classe `PsychoAcousticModel`) para obtenção dos coeficientes utilizados no teste de detecção e para redução de ruídos presentes na portadora do sinal, respectivamente. A verificação que reúne os dados obtidos para indicar a presença de dados ocultos no áudio é realizada no método `CalcNormalizedCorrelations()`.

A Figura 19 exibe a colaboração de algumas classes do sistema para realização do caso de uso UC05 – descompactar dados. Ele é realizado apenas quando o sistema identifica que a mensagem escondida no áudio está compactada, através do método `IsCompressed()` da classe `WaterMark`. Neste caso, é criada uma instância da classe `Decompressor` para posterior utilização do método `Unzip()`, que recupera a mensagem original escondida no áudio.

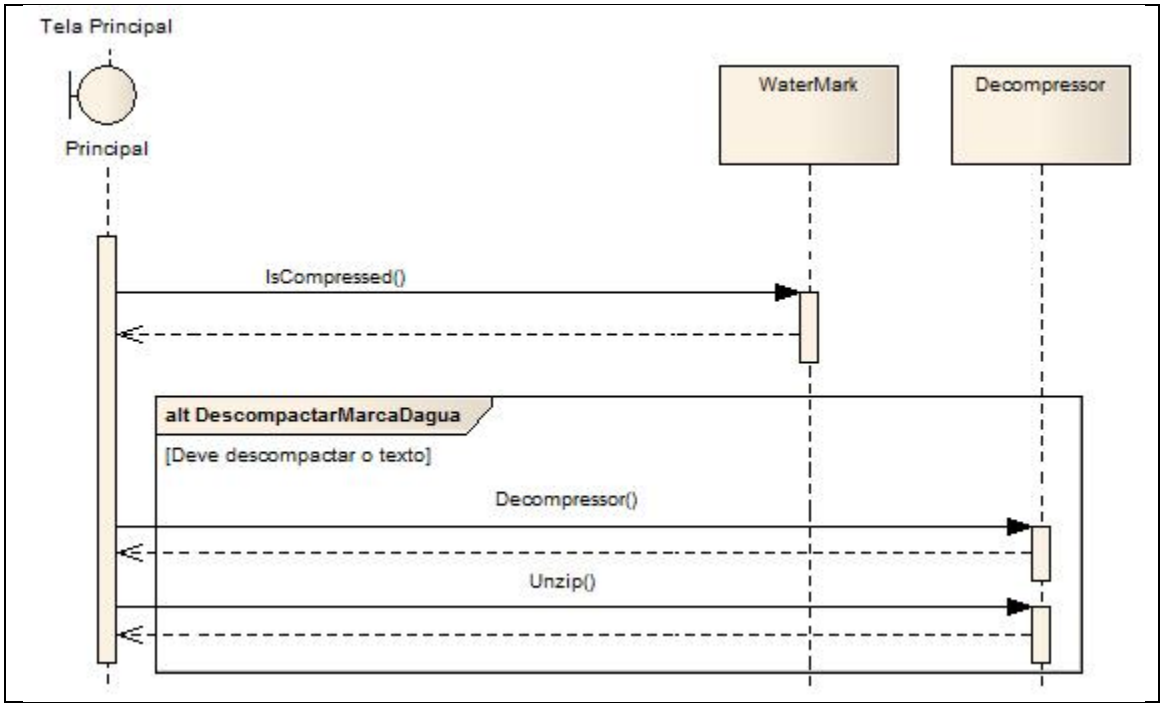


Figura 19 – Diagrama de seqüência “Descompactação da marca d’água”

3.3 IMPLEMENTAÇÃO

Esta seção descreve as técnicas e ferramentas utilizadas durante a implementação do sistema, assim como sua operacionalidade.

3.3.1 Técnicas e ferramentas utilizadas

A aplicação foi desenvolvida utilizando a linguagem de programação C#, com suporte à versão 3.5 do .NET Framework. Nestas condições, a IDE Microsoft Visual Studio 2008 foi selecionada para a realização da implementação.

As técnicas e tecnologias empregadas durante o desenvolvimento são descritas a seguir:

a) compactação de dados: técnica responsável por realizar a compressão do texto que o usuário deseja ocultar em um áudio digital através de métodos heurísticos, aumentando assim a quantidade de informações que pode ser escondida pela ferramenta (seção 2.3). Esta abordagem foi necessária, pois algoritmos consagrados nesta função não realizam-na de forma satisfatória quando são utilizados para compressão de textos de tamanho reduzido;
b) espalhamento espectral: metodologia de dispersão de informações aplicada em conjunto com um modelo matemático do SAH, desenvolvido por terceiros, para ocultar efetivamente a marca d'água em formato de texto no espectro sonoro do áudio selecionado pelo usuário. No caso deste trabalho, optou-se por utilizar o espelhamento espectral por seqüência direta, descrito na seção 2.2.1.1, aliado a uma modelagem psicoacústica (seção 2.1.1.1);
c) *Windows Presentation Foundation* (WPF): framework para desenvolvimento de interfaces gráficas ricas, disponível a partir do .NET Framework 3.0, utilizado na construção da aparência do sistema. Oferece um modelo de desenvolvimento que separa claramente a regra de negócio do design da interface, sendo esta última orientada por uma linguagem de programação semelhante ao *eXtensible Markup Language* (XML).

A inclusão de uma mensagem secreta é permitida apenas quando o arquivo WAV selecionado apresenta formatação compatível com o modelo psicoacústico utilizado pelo

sistema (MALVAR, 1999). Desta forma, a primeira etapa do algoritmo analisa o cabeçalho do arquivo, a fim de validar sua utilização.

Posteriormente, as amostras do espectro sonoro, conhecidas como *samples*, são lidas e agrupadas em vetores que representam os canais do áudio. Como estes dados são utilizados no processo de esteganografia do programa, sua estrutura armazena a mensagem oculta quando esta está presente, ou fornece a localização onde a informação será escondida.

Apesar de o software possuir um módulo de inserção de dados e outro de extração, apenas uma opção torna-se disponível após a análise do arquivo. Durante a tomada desta decisão, a utilização de um processo de verificação para detectar a existência de marcas d'água no sinal torna-se essencial, e por este motivo foi adicionado às funcionalidades da aplicação.

Para facilitar a compreensão da ordem dos eventos realizados pelo sistema, a Figura 20 estabelece os elementos que integram suas funcionalidades principais.

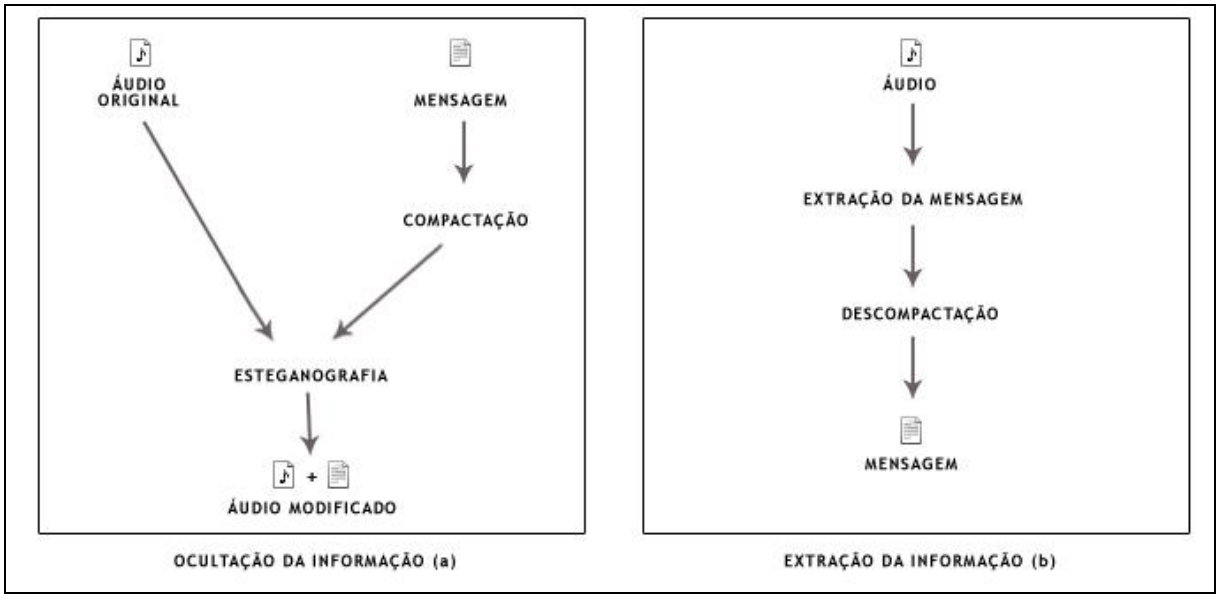


Figura 20 – Etapas realizadas durante o processo de ocultação e extração da informação

Na etapa de ocultação da informação, exibida na Figura 2 (item a), o sistema esconde o dado compactado no áudio selecionado pelo usuário. O resultado deste processo é um arquivo com a mensagem secreta em um áudio modificado.

A extração do dado (item b, Figura 2) é feita através de filtragens e varreduras do som. Após a obtenção da informação, ela é descompactada, gerando assim a mensagem original que foi escondida.

Informações acerca do formato de arquivo de áudio compatível com a ferramenta são encontradas na seção 3.3.1.1. A seguir, as seções 3.3.1.2 e 3.3.1.3 descrevem os passos seguidos pelo software para efetuar a inserção e extração da marca d'água no som,

respectivamente. A verificação responsável por indicar a presença de informações escondidas no áudio é detalhada na seção 3.3.1.4.

3.3.1.1 Arquivo de áudio

No processo de inclusão da mensagem, o usuário deve selecionar um arquivo no formato WAV. Para ser considerado compatível com a ferramenta, o arquivo deve possuir dois canais, freqüência de amostragem de 44100 Hz e no máximo 16 *bits* por amostra.

O tamanho dos blocos de processamento do setor de dados foi fixado em 11 segundos para manter compatibilidade com o modelo psicoacústico utilizado. Assim, a quantidade de dados que pode ser ocultada é diretamente dependente do tamanho em segundos do áudio.

3.3.1.2 Inserção da marca d'água

O processo de inserção de dados inicia com a transformação do arquivo WAV em amostras PCM e o cálculo da quantidade de blocos que serão utilizados para ocultar a informação, chamados de "janelas". A estrutura do algoritmo de inserção dos dados pode ser visualizada no Quadro 2.

A mensagem secreta é compactada através de processos heurísticos que substituem preposições de uso comum e combinações de letras. Por este motivo, a utilização de nomes próprios e siglas no texto podem comprometer a qualidade de compressão do método. Por ser um dos idiomas mais utilizados atualmente e ter proporcionado bons resultados nos testes realizados, o inglês foi escolhido para criação do dicionário de palavras da ferramenta. Entretanto, é possível adaptar este dicionário para outra língua ou para um determinado conjunto de palavras e símbolos. Desta forma, os resultados fornecidos pelo sistema tornam-se mais adequados em situações específicas.

Em seguida, é realizada a análise espectral do sinal, passando-o do domínio tempo para o domínio da freqüência. Para isso, ao invés de utilizar a transformada DFT, popular na área de processamento digital, foi empregada a MCLT, por possuir propriedades de reconstrução do sinal que a tornam atrativa para o processamento de áudio. O uso desta transformada no projeto é dependente da mesma biblioteca externa que disponibiliza o modelo psicoacústico. Este processo é gradativo no algoritmo, sendo aplicado na iteração principal que analisa os

blocos de amostragem do áudio.

```
public void AddWatermark(WAVData input)
{
    //quantidade de blocos
    int blocksLength= (input.FramesLength + Constants.nFreq-1) / Constants.nFreq;

    //limites de freqüências da marca d'água
    int freqMin = AudioTools.GetFrequenceLimit(Constants.freqMinChip),
        freqMax = AudioTools.GetFrequenceLimit(Constants.freqMaxChip);

    ...

    //looping principal para inserir o conteúdo no áudio
    for (int block = 0; block < blocksLength; block++)
    {
        //processa cada canal do áudio
        for (int channel = 0; channel < input.Channels; channel++)
        {
            ...

            //modifica frequências que vão até o valor definido em 'freqMax'
            for (int frequency = freqMin; frequency <= freqMax; frequency++)
            {
               ...
            }

            ...
        }

        //adiciona sample modificado no arquivo de saída
        for (int i = 0; i < Constants.nFreq; i++)
        {
            for (int channel = 0; channel < input.Channels; channel++)
            {
               StegoData.Add(samples[channel][i]);
            }
        }
        ...
    }
}
```

Quadro 2 – Algoritmo de inserção da marca d'água

A marca d'água é definida como uma seqüência de espalhamento espectral *w*, ou seja, um vetor de dados gerado de forma pseudo-randômica, onde $w \in \{\pm 1\}$. Cada elemento w_i da seqüência *w* é chamado de "chip". Na prática, cada caractere da mensagem é modulado em *chips* e escondido em um bloco *x* que está sendo processado, em diversas amplitudes δ. Assim, o sinal marcado gerado *y* pode ser definido como a soma destes fatores, ou seja, $y = x + \delta w$.

Ao término de cada iteração, o vetor *y* e a amostra original do bloco são combinados e transformados em um sinal pertencente ao domínio tempo, através da MCLT inversa.

Por fim, os valores PCM obtidos pelo sistema são reunidos para a criação de um arquivo WAV com a marca d'água.

3.3.1.2.1 Conformação de ruído

A ocultação dos mesmos *chips* em diversas faixas de freqüência aumenta a segurança do sistema e a taxa de detecção do texto oculto, porém pode ocasionar audibilidade de ruídos indesejáveis, provenientes da adição da marca d'água, em faixas mais silenciosas do áudio.

Para evitar este problema, a amostra processada é submetida a um teste de variação de energia que indica se o ruído adicionado será audível ou não. Caso positivo, os *chips* não são esteganografados no bloco em análise.

O Quadro 3 apresenta a etapa do algoritmo de inserção onde ocorre a realização do teste.

```
public void AddWatermark(WAVData input)
{
    ...

    //processa cada canal do áudio
    for (int channel = 0; channel < input.Channels; channel++)
    {
        //copia um bloco de conteúdo PCM para o array 'samples'
        input.CopyPCMDataToArray(ref samples[channel], channel, Constants.nFreq,
indexCount);

        //utiliza a transformada MCLT
        mcltTransform.CalcMCLT(channel, samples[channel]);

        //verifica se existe variação brusca de energia no áudio
        if (!this.pam.HasPreEchoProblem(samples[channel]))
        {

            //esconde a informação no espectro sonoro...
            ...
        }

        //utiliza a transformada MCLT inversa
        mcltTransform.CalcInvertMCLT(channel, samples[channel]);
    }

    ...
}
```

Quadro 3 – Teste de variação de energia

3.3.1.3 Extração da marca d'água

Inicialmente, o arquivo WAV de entrada é convertido em um vetor com valores PCM e suas características são validadas para garantir que o áudio seja compatível com o modelo psicoacústico adotado.

Antes do processamento, o sistema estima a quantidade de blocos com caracteres

ocultos existente no áudio, para que seja possível definir quando a busca pela mensagem deve ser interrompida. As freqüências audíveis do som foram fixadas como o alvo do algoritmo de detecção, pois as porções inaudíveis são mais suscetíveis a ataques de ruído, já que ocupam grande parte do espectro do sinal. É importante mencionar que o módulo de extração dos caracteres ocultos é considerado *Blind Detector* (detector cego), por não necessitar do áudio original durante o processo de extração da marca d'água (seção 2.1.3.2.1).

Para realizar a extração, são gerados coeficientes de decodificação em diversas escalas de tempo e freqüência. Esta operação tornou-se necessária, pois a utilização do MCLT no processo de inserção espalha os *chips* de cada caractere por toda a subbanda da janela que mascara o conteúdo, gerando a redundância de informação que aumenta a segurança da técnica.

A busca pela marca d'água inicia-se pela análise de um bloco do vetor que armazena a informação sonora e em seguida, a geração de seus respectivos coeficientes MCLT. Estes valores são necessários para que sejam realizados testes de audibilidade dos limiares de detecção.

Com o objetivo de reduzir possíveis ruídos da portadora do sinal que atrapalhariam o procedimento, é aplicado um processo chamado de Filtro de Cepstrum (FC) ao resultado gerado pela combinação dos dados PCM com os coeficientes MCLT. Paralelamente são identificadas as freqüências inaudíveis presentes no *buffer* analisado, para que estas sejam ignoradas durante a extração dos caracteres.

O Quadro 4 exibe o esquema geral do algoritmo de extração.

```
private WaterMark RecoverHideData(bool analysisMode)
{
    StringBuilder sbRecover = new StringBuilder();
    //quantidade de blocos
    int blocksLength=(this.input.FramesLength+Constants.nFreq-1)/Constants.nFreq;

    ...

    //pontos de decodificação (domínio FREQUÊNCIA)
    this.extractSubBands = this.ss.GetExtractSubBandLimits();
    //pontos de decodificação (domínio TEMPO)
    this.extractTimeIntervals = this.ss.GetExtractTimeIntervals(blocksPerWindow,
searchStep);

    ...
    //uma verificação dentro do algoritmo realiza a parada do processamento
    while (true)
    {
        this.LoadBuffer(blockLoadLimit, blockFreqSize, searchStep, ref
readedSamples);
        maxCorrelation = this.CalcNormalizedCorrelations(blockFreqSize, ref
hiddenBitsPos);

        lastResults[currentBufferPos] = new DetectionResult();
        lastResults[currentBufferPos].NC = maxCorrelation.SumSquares;

        ...
        //procurando marca d'água
        if (state == EDetectionState.Searching)
        {
            ...

            bestCase = DetectionResult.GetBestCase(this.lastResults,
currentBufferPos);
            bestCase.ProcessTime = currentWindowTime;

            ...
        }
        //detectando marca d'água
        else if(state == EDetectionState.Detecting)
        {
            ...

            sbRecover.Append(Hexadecimal.DecimalToHex(bestCase.CCI));
            sbRecover.Append(Hexadecimal.DecimalToHex(bestCase.Load));
            ...

            if (DetectionResult.GetNCSum(this.lastResults, false) > bestCase.NC)
            {
                bestCase = DetectionResult.GetBestCase(this.lastResults,
currentBufferPos);
                bestCase.ProcessTime = currentWindowTime;
            }
            ...
        }
        currentWindowTime += blockSecLength * blockLoadLimit * searchStep;

        //verifica término da execução
        if (currentWindowTime > ((blocksLength * blockSecLength) -
Constants.windowSecondsSize))
        {
            ...
            break;
        }
    }
    return new WaterMark(sbRecover.ToString(), true, false);
}
```

Quadro 4 – Algoritmo de extração da marca d'água

Neste método, a verificação da existência de algum dado oculto na porção que está sendo processada da janela é realizada através de testes de correlação. Eles são efetuados em todas as escalas de tempo e freqüência utilizadas durante a inserção da informação. Resultados elevados nestes testes denunciam a presença de marcas d'água, conforme pode-se observar na Figura 21.

Quando a soma das últimas três correlações realizadas supera o valor de um limiar de detecção τ, o sistema entende que um caractere da marca d'água foi encontrado.

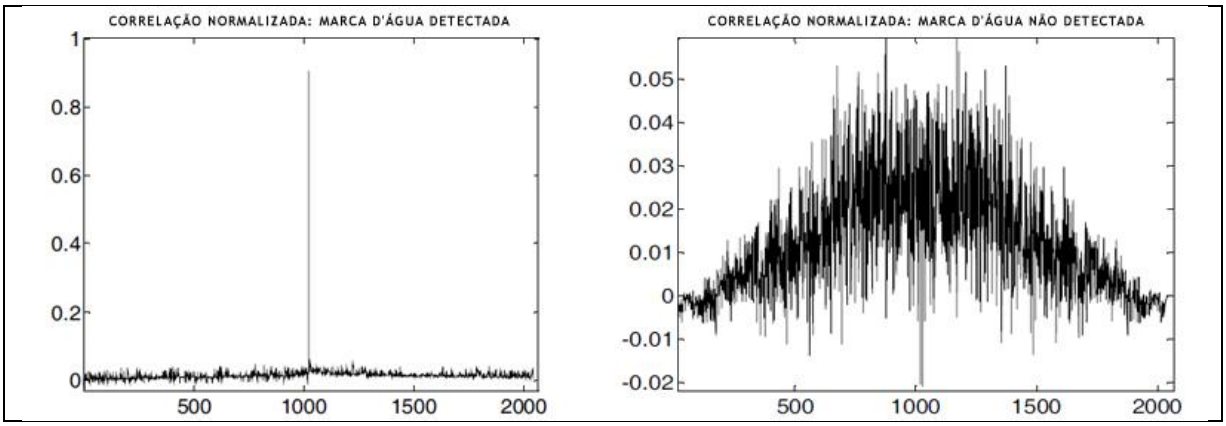


Figura 21 – Resultados da correlação normalizada

As rotinas de aquisição e rastreamento da ferramenta se completam, porém não são executadas simultaneamente. Enquanto a primeira gera um coeficiente baseado nos últimos resultados de correlação obtidos, a segunda analisa este coeficiente e armazena o caractere representado pelo seu valor.

Ao final deste procedimento, os caracteres armazenados são reunidos, formando a mensagem oculta. Antes de apresentá-la, o sistema aplica o algoritmo de descompactação compatível com o método utilizado na inserção, para revelar a marca d'água que foi originalmente adicionada ao áudio.

3.3.1.3.1 Aquisição

Ao encontrar indícios da marca d'água no espectro sonoro, o sistema inicia o monitoramento dos valores das correlações, com o objetivo de encontrar um ponto de estabilização de resultados.

Quando este ponto é alcançado, os últimos dados obtidos são combinados, gerando o coeficiente para detecção do caractere oculto e, consecutivamente iniciando a detecção do texto. A Figura 22 exibe a seqüência de passos executados para encontrar a marca d'água

espalhada pelo áudio.

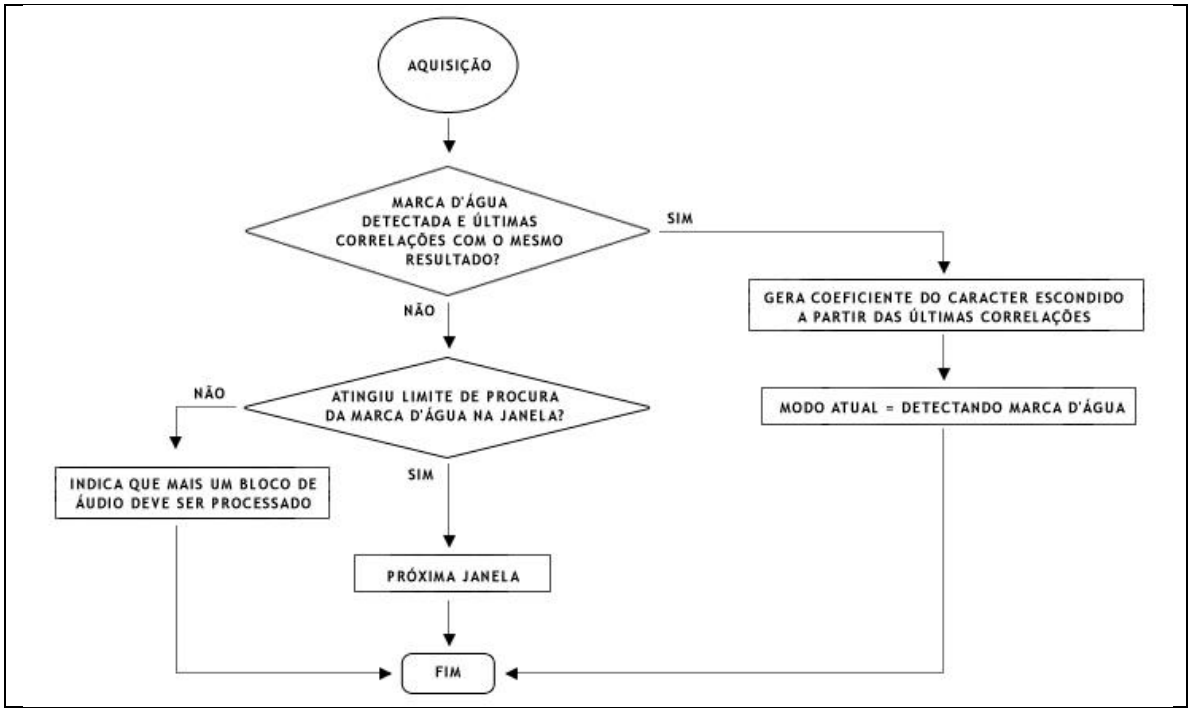


Figura 22 – Esquema de aquisição do sinal

Foi adotado um limite de busca de 5 segundos por janela. Se ele for atingido, a busca é interrompida, pois nenhum texto oculto foi percebido. Quando não são encontrados indícios de mensagens esteganografadas no áudio após a análise de uma parte significativa do bloco, este ponto de parada é importante para que o sistema não permaneça realizando cálculos desnecessários. Neste caso, o sistema passa a realizar varreduras na próxima janela do áudio.

3.3.1.3.2 Rastreamento

A etapa de rastreamento, apresentada na Figura 23, é responsável por verificar quando ocorre dessincronização nos resultados das correlações. Este cenário indica que o caractere deve ser extraído, através do coeficiente armazenado pelo sistema durante a análise da janela atual, e mantido em um *buffer*.

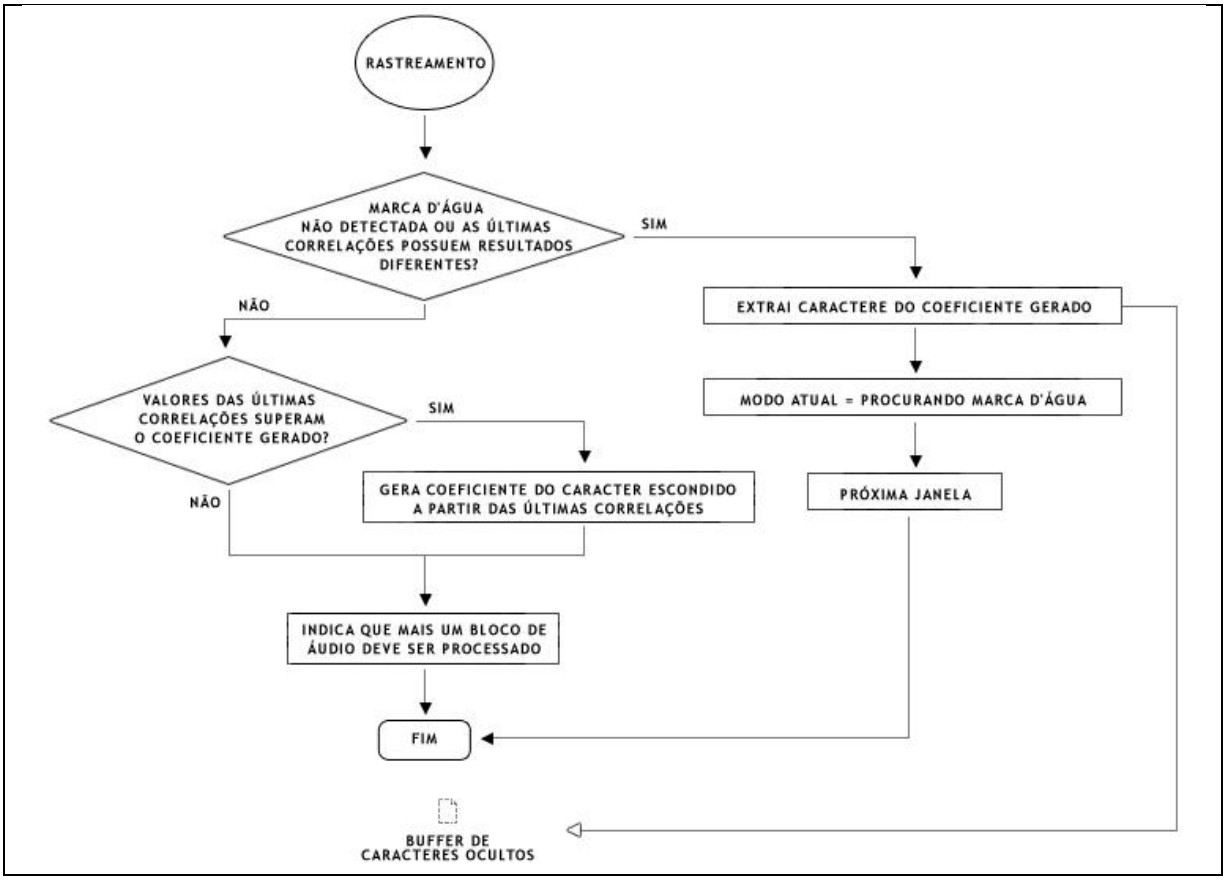


Figura 23 – Esquema de rastreamento do sinal

Enquanto este estado não for atingido (dessincronização), a rotina continua atualizando o coeficiente, caso as últimas correlações apresentem resultados mais favoráveis do que aqueles capturados anteriormente. Estes valores guiam o processo de identificação do caractere oculto, cujo agrupamento formará a mensagem escondida pelo emissor no fim da execução do programa.

3.3.1.4 Presença de marca d'água

Antes de disponibilizar funcionalidades ao usuário, o sistema precisa identificar se existem mensagens ocultas no áudio fornecido.

O método que extrai a marca d'água de um arquivo pode ser utilizado para realização desta verificação. Assim, o procedimento de detecção da mensagem é iniciado, porém em um modo de análise. Neste estado, apenas a primeira janela do áudio passa pela avaliação espectral, evitando processamentos desnecessários. Se as letras "N" (mensagem em formato normal) ou "C" (mensagem em formato compacto) forem encontradas no primeiro bloco de áudio, a operação para recuperação da mensagem oculta é habilitada. Caso contrário, a função

de adição de marca d'água torna-se acessível.

A função responsável por gerenciar o teste é apresentada no Quadro 5.

```
public bool HasWaterMark(WAVData input)
{
    this.input = input;
    WaterMark wm = this.RecoverHideData(true);
    if (wm != null)
    {
        String flagChar = wm.Value[0].ToString();

        switch (flagChar)
        {
            case Constants.normalWaterMark:
            case Constants.compressedWaterMark:
                return true;
            default:
                return false;
        }
    }
    else
    {
        return false;
    }
}
```

Quadro 5 – Algoritmo para verificar existência de dados ocultos

O método realiza a análise apenas do primeiro caractere da mensagem, pois é onde o algoritmo de inserção da marca d'água posiciona o item que indica a presença e o formato do texto.

3.3.2 Operacionalidade da implementação

Para utilizar o sistema, o usuário deve fornecer um arquivo no formato WAV válido e compatível com a aplicação. O software analisa a estrutura do áudio digital e habilita a aba que permite a inserção de uma marca d'água no som ou a aba que possui a opção de extração do dado escondido, de acordo com o estado do espectro sonoro do arquivo. Caso o áudio informado não tenha um formato válido, nenhuma das opções é disponibilizada e o usuário é informado desta situação.

É importante ressaltar que o sistema não permite que o usuário sobrescreva uma marca d'água, inserindo outra em seu lugar. Esta abordagem impede que agentes mal-intencionados corrompam as informações de um arquivo, caso tenham acesso ao programa.

A interface possui três seções principais:

a) informações do arquivo WAV: painel principal do sistema, localizado no parte superior da tela. Nele são encontradas informações do áudio, além do resultado da análise inicial feita pela ferramenta, que identifica se existe alguma marca d'água

oculta no som. O último item exibido nesta seção é um pequeno tocador de áudio, onde é possível realizar a execução do arquivo selecionado;

b) aba de inserção da marca d'água: área destinada à ocultação de uma mensagem no áudio escolhido, situada na parte inferior da tela. Uma opção desta aba permite que o texto seja compactado antes de escondê-lo. Quando esta seção está habilitada, a aba de detecção de mensagens fica indisponível;

c) aba de extração da marca d'água: seção que disponibiliza a ação de detecção do texto oculto no espectro do som. Caso a mensagem esteja compactada, o sistema utiliza o algoritmo apropriado para recuperá-la adequadamente antes de apresentar o resultado ao usuário. Quando esta seção está habilitada, a aba de adição de mensagens fica indisponível.

A Figura 24 apresenta a tela inicial do sistema, com os valores dos campos vazios até que um áudio seja carregado.

Figura 24 – Tela inicial do sistema Cripton

As seções da interface que permitem interatividade com o conteúdo multimídia não são exibidas enquanto nenhum áudio for analisado pela ferramenta. Assim que um arquivo é carregado, estes itens se tornam visíveis e os dados do áudio são exibidos no painel principal do software.

Caso o áudio selecionado seja incompatível, o sistema apresenta uma mensagem informando qual característica de sua estrutura está em desconformidade com o software. A Figura 25 exibe a aparência adotada pelo programa nesta situação.

Figura 25 – Arquivo de áudio incompatível

3.3.2.1 Inserção da marca d'água

Quando o usuário emissor seleciona um áudio que não possui informações ocultas em seu espectro, o sistema disponibiliza a área para ocultação de uma mensagem (Figura 26). Nesta aba existe um campo onde o emissor da mensagem deve informar o nome do novo arquivo de áudio, cujo espectro esconderá uma marca d'água.

A quantidade de informação que pode ser escondida no áudio depende exclusivamente do comprimento do mesmo em segundos. Enquanto o usuário digita o texto secreto no campo apropriado, a ferramenta atualiza um indicador na tela que exibe quantos caracteres ainda poderão ser informados.

A opção "Compactar texto" possibilita que a quantidade de caracteres transmitidos seja maior do que no estado padrão do programa, dependendo do valor informado. O algoritmo heurístico responsável por esta funcionalidade substitui composições comuns de frases por símbolos significativos, diminuindo desta forma o tamanho do texto. Este processo também funciona como uma espécie de criptografia, pois a descompactação da informação depende de uma lista diferente de símbolos, que apenas o sistema conhece. Quando esta opção está marcada, o sistema realiza internamente a compressão do texto cada vez que a frase é alterada pelo usuário, pois este processo pode passar a ser desvantajoso se a frase contiver muitas expressões que não estão no dicionário de símbolos do software. Neste caso, um aviso informando a situação é exibido na interface.

O processo é iniciado quando o usuário clica no botão "Adicionar marca d'água". Sua execução pode ser acompanhada por uma barra de progresso, exibida apenas neste momento,

que indica a etapa de análise que está sendo efetuada pelo sistema.

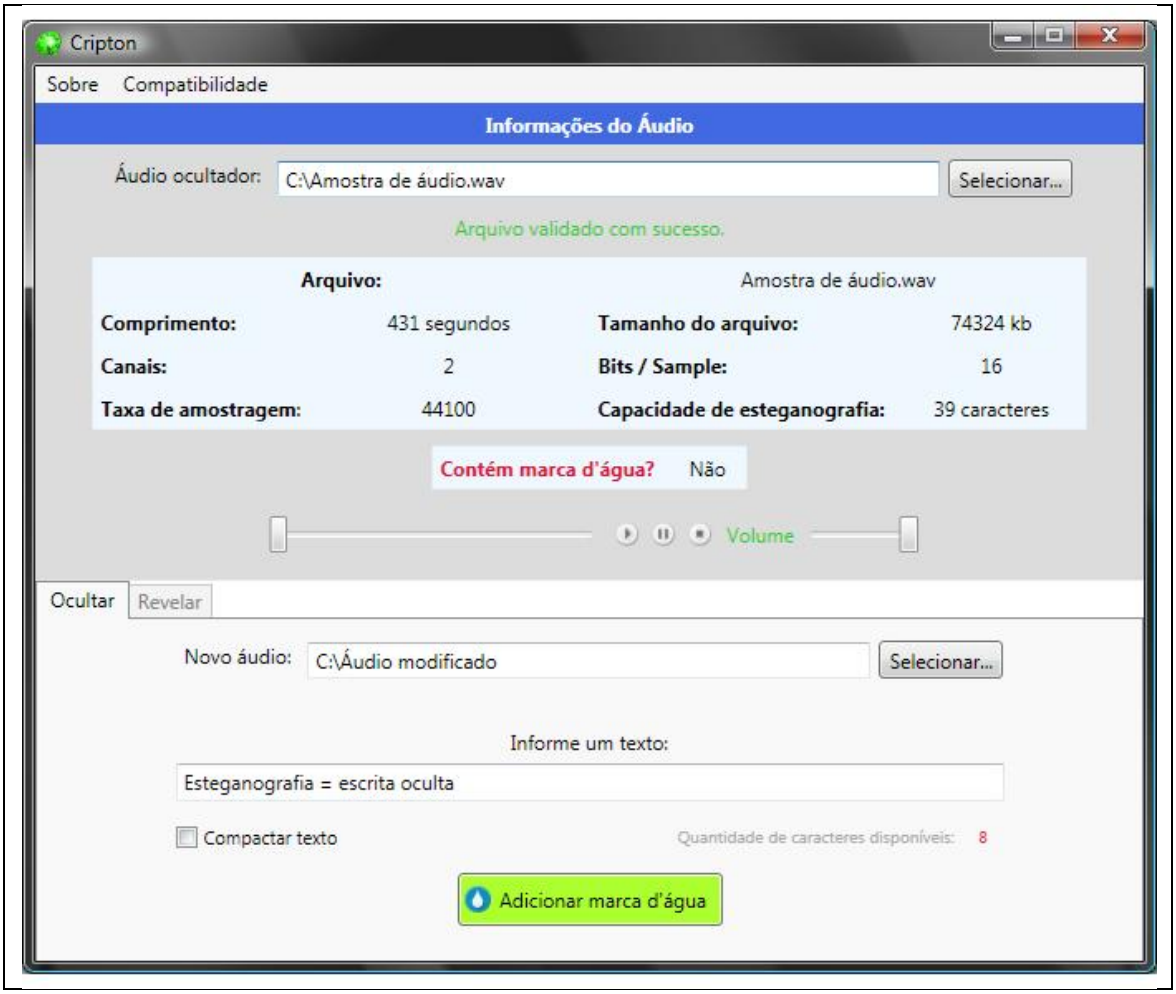


Figura 26 – Adição de marca d'água em um áudio através do sistema Cripton

Um alerta é exibido para o usuário ao término da inserção da marca d'água.

3.3.2.2 Extração da marca d'água

A funcionalidade de detecção e extração de uma marca d'água é habilitada caso a análise inicial realizada pelo sistema identifique a existência de dados ocultos no conteúdo do arquivo selecionado. Durante esta avaliação do sinal, apenas o primeiro bloco do som é processado e testado. Através desta janela é possível identificar a presença de uma mensagem esteganografada e a forma como está representada.

Para visualizar o texto secreto, o usuário deve clicar no botão "Revelar marca d'água". A extração da mensagem exige um esforço computacional maior, pois é necessário que o sistema realize múltiplos testes com os valores do espectro sonoro para obter a sincronia e a detecção das informações. Por este motivo, o tempo de execução deste algoritmo tende a ser

maior do que o período apresentado na adição da marca d'água.

A configuração visual adotada pelo software, quando a função de recuperação da marca d'água está disponível, é ilustrada na Figura 27.

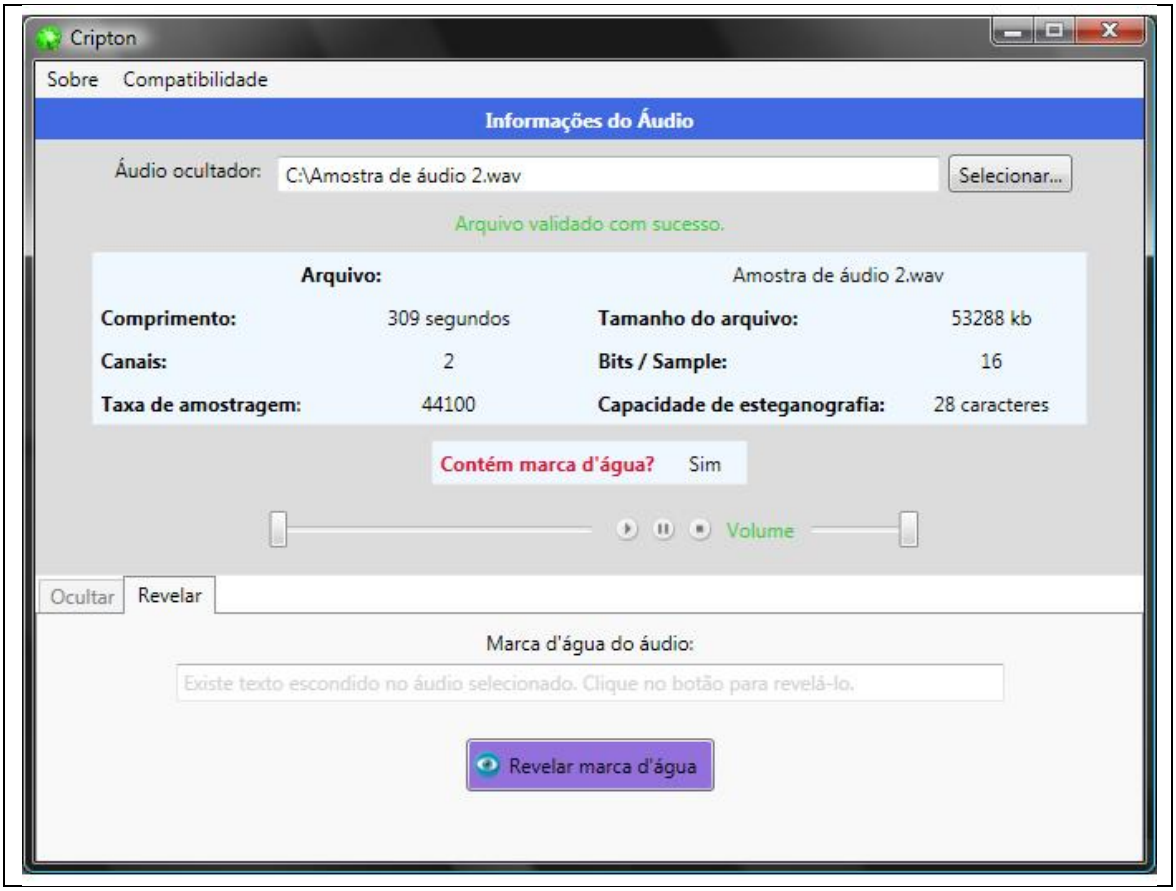


Figura 27 – Extração de marca d'água de um áudio através do sistema Cripton

Quando o programa identifica que a mensagem está compactada, o procedimento de descompressão do texto é efetuado automaticamente. Portanto, não é necessário que o receptor da marca d'água realize nenhuma parametrização ou ação adicional para que este processo seja executado.

Durante a extração, é exibida uma barra de progresso nesta aba para que o usuário consiga monitorar o avanço gradativo do processamento. Assim que a captura da marca é concluída, o sistema exibe um aviso informando o usuário.

3.4 RESULTADOS E DISCUSSÃO

Para validar o sistema, foram realizados testes inserindo a mensagem "Teste de esteganografia" em amostras de áudio. A opção de compactar a mensagem foi utilizada em

todos os casos.

Para mensurar a qualidade perceptual utilizou-se a ferramenta PQevalAudio, descrita em (Kabal, 2002). Nela, variáveis de saída do modelo (em inglês, *Model Output Variables* ou MOVs) são ponderadas e combinadas, gerando uma nota única para o nível de degradação dos sinais de entrada, chamada de *Objective Difference Grade* (ODG). A faixa de valores desta nota varia de –4.0, em casos de distorção muito desagradável, até 0, representando ausência total de distorções. Para serem considerados aceitáveis, os resultados devem apresentar ODGs maiores do que –1,5. Valores maiores que –1,0 são considerados muito bons (SCHÜTZ, 2009, p. 84). Neste software, o sinal de áudio avaliado é comparado com uma referência, que em geral é o áudio original. Os sinais são processados por um modelo psicoacústico e então as MOVs são calculadas em relação ao sinal de referência.

A porcentagem de fidelidade da marca d'água extraída foi mensurada de acordo com o acerto durante a recuperação dos símbolos do texto oculto. A presença de símbolos que não existiam na mensagem original foi considerada durante o cálculo desta taxa, assim como o posicionamento incorreto de letras no texto extraído.

Os resultados obtidos estão listados na Tabela 1.

Tabela 1 – Testes de qualidade perceptual e fidelidade da marca d'água extraída

TESTES DE QUALIDADE E FIDELIDADE			
Áudio	**Autor / Música**	**% de fidelidade da marca extraída**	**ODG**
1	Black Sabbath / Paranoid	95%	0
2	Black Sabbath / Iron Man	100%	0
3	Bob Marley / Buffalo Soldier	100%	-1.195
4	Bob Marley / Could You Be Loved	100%	-1.459
5	Bob Marley / Is This Love	100%	-0.87
6	Bob Marley / No Woman No Cry	100%	-0.032
7	Elvis Presley / Good Luck Charm	90%	-1,645
8	Elvis Presley / It's Now or Never	100%	-1,65
9	Elvis Presley / Love Me Tender	90%	-0,028
10	Elvis Presley - Suspicious Minds	95%	-1,094
11	Elvis Presley / The Wonder of You	90%	-1,085
12	Oasis / Wonderwall	100%	-0.13
13	Oasis / Don't Look Back in Anger	100%	-0.617
14	Scorpions / Still Loving You (live)	50%	-0.296
15	Scorpions / Wind of Change (live)	30%	-0.430
16	The Beatles / If I fell	95%	0
17	The Beatles / Strawberry fields	90%	-0.119
18	The Doors / Alabama Song	100%	0
19	The Doors / Light My Fire	100%	0
20	The Doors / Riders On The Storm	100%	0

Na maioria dos testes, a taxa de acerto na extração da marca d'água foi satisfatória,

com níveis de ruído aceitáveis ou muito bons.

Os itens 7 e 8 obtiverem bons níveis de extração da informação oculta, porém as notas ODG geradas para os áudios produzidos pelo sistema não atingiram o nível mínimo aceitável. A escassa existência de percussão nestas músicas pode ter ocasionado a adição excessiva de ruído no sinal, pois instrumentos desta categoria geram freqüências que auxiliam o processo de ocultação de dados. Uma explicação mais precisa diz que a presença de um ruído ou de um tom com intensidade suficiente para criar uma forte excitação na membrana basilar (estrutura que faz parte do ouvido interno, onde se dá início ao processamento neural dos sons) bloqueia a percepção de um estímulo mais fraco (PAINTER; SPANIAS, 2000, p. 457 - 458)

Nos casos 14 e 15, as amostras de som tratam-se de gravações ao vivo, com volume médio e novamente pouca presença de percussão. Nestes casos, os resultados foram abaixo do esperado. Isso certamente ocorreu devido às características particulares destas faixas. Na primeira, o sinal de áudio possui energia muito baixa e na segunda, possui vários períodos de silêncio ao longo dos trechos utilizados.

Nas amostras 1 e 16, não foi notado nenhum tipo de ruído após o mascaramento do texto, porém a compreensão da mensagem extraída foi comprometida pelo tamanho limitado dos arquivos, que por conseqüência limita a quantidade de caracteres que o sistema consegue ocultar.

A Figura 28 mostra a equalização gráfica de pequenas amostras de som, extraídas através de uma ferramenta de edição de áudio. Os dados exibidos representam variações do item 13 da Tabela 1. A Figura 28a representa a forma do sinal antes do processo de esteganografia. A estrutura da informação é alterada assim que uma mensagem é escondida no espectro do áudio, gerando alterações na amostra selecionada. As diferenças mais significativas são destacadas na Figura 28b. Entretanto, as técnicas de conformação de ruído empregadas pelo sistema garantem que o conteúdo original não seja gravemente comprometido durante o processamento do sinal.

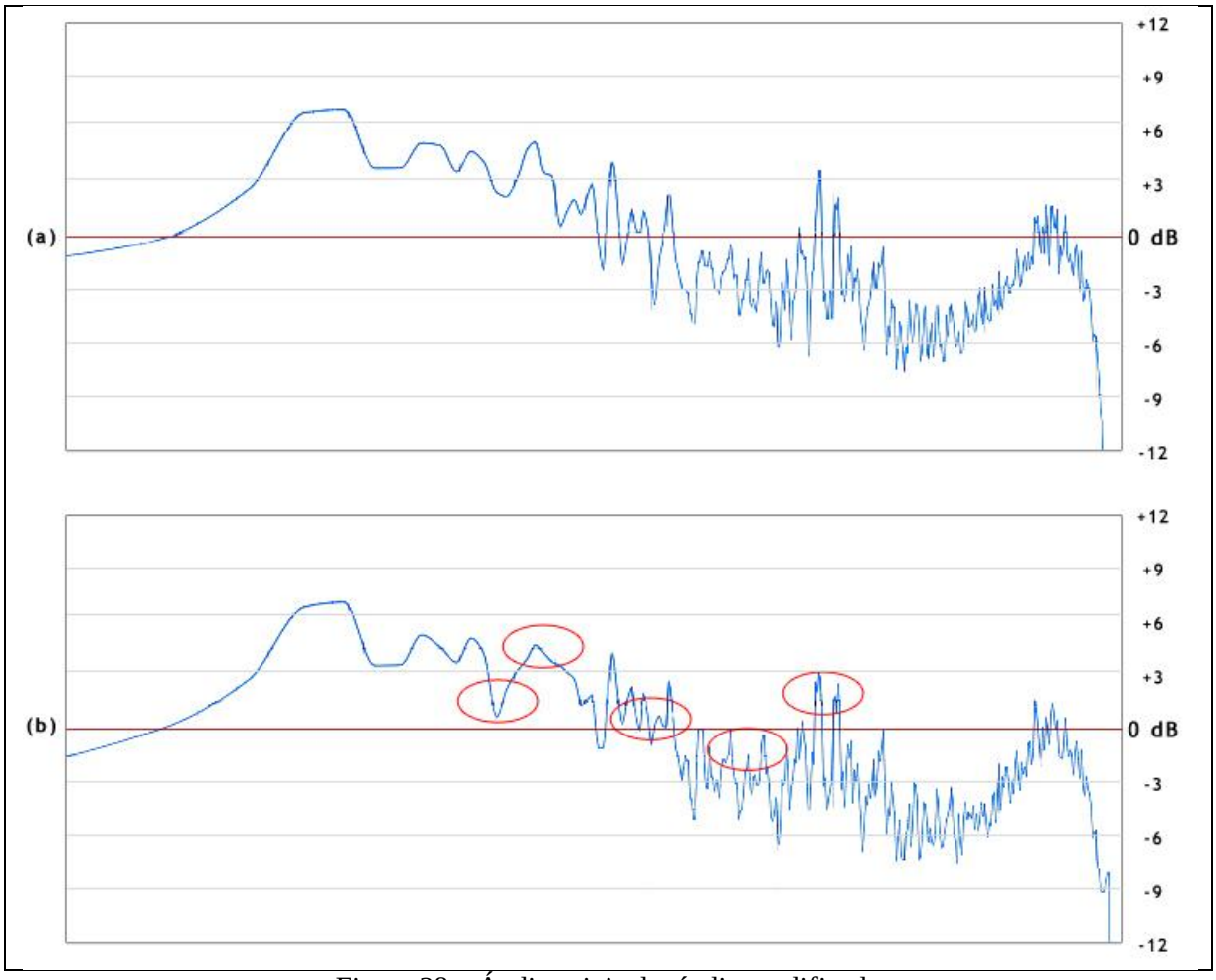


Figura 28 – Áudio original e áudio modificado

As modificações causadas pelo processo tendem a ser imperceptíveis ao ouvido humano, de acordo com o tipo de áudio presente no arquivo. Através da comparação dos gráficos, é possível perceber que a ferramenta não altera consideravelmente o nível de dB. Conseqüentemente, a quantidade de ruído presente após a ocultação da informação é pequena.

A Figura 29 apresenta outra representação gráfica dos mesmos arquivos de áudio, em um panorama mais amplo. A análise de uma quantidade maior de dados torna menos perceptíveis visualmente as alterações presentes no arquivo modificado.

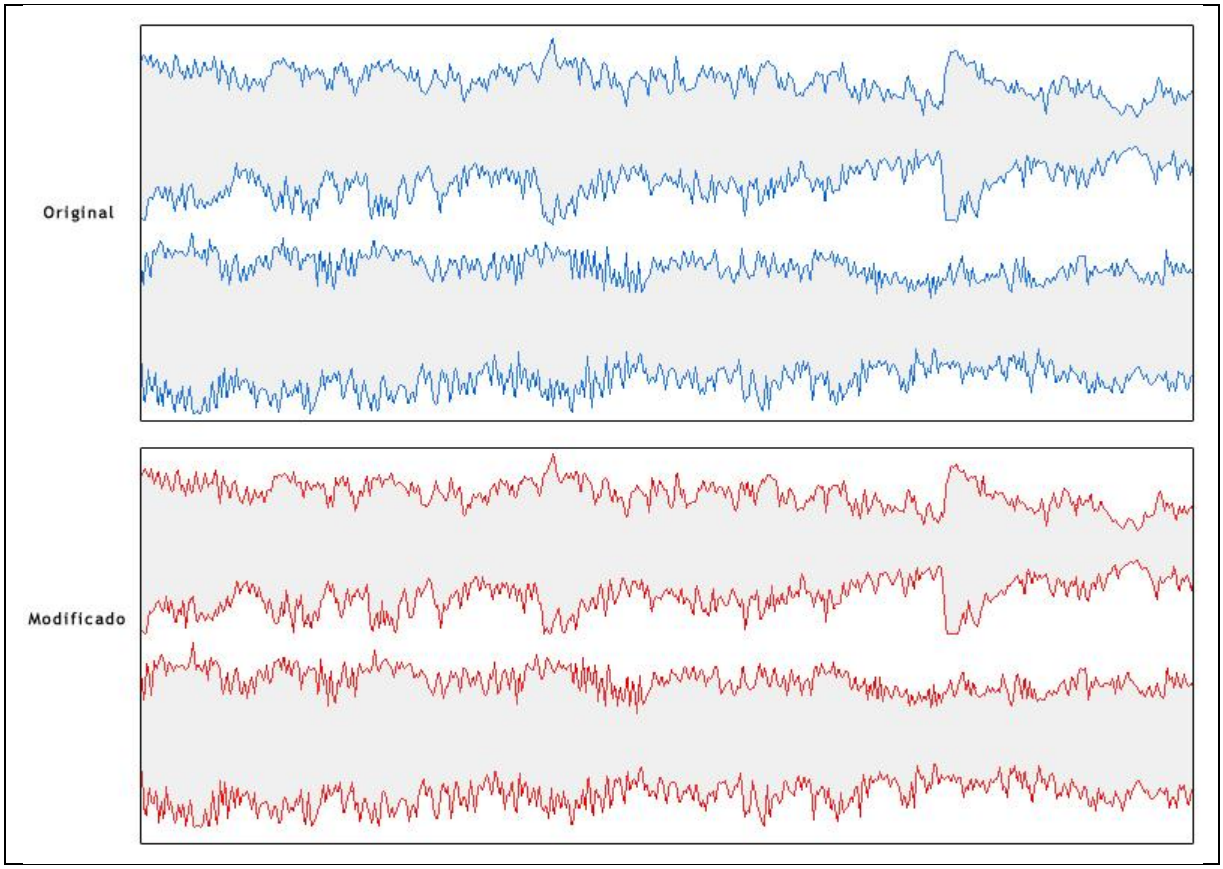


Figura 29 – Visão ampla dos áudios

3.4.1 Análise de transmissão pelo ar

Embora o software tenha sido desenvolvido para manipular apenas arquivos transmitidos em meio digital, também foi realizado um experimento fora do escopo proposto.

Neste ensaio, a mensagem “Informação secreta” foi escondida em formato compactado no conteúdo sonoro original do item 1 (Tabela 1), gerando um novo arquivo de áudio. Em seguida, o som foi reproduzido através de um dispositivo móvel. A recepção do sinal foi realizada por um software, que gera um arquivo digital em formato WAV a partir da informação captada por um microfone conectado ao computador. Este tipo de troca de informações é conhecido como transmissão através do ar (seção 2.1).

A nota ODG obtida pela avaliação do novo arquivo de áudio foi igual a -3.9, evidenciando a existência de muito ruído no espectro do som, oriundo da transmissão do som. Entretanto, quando submetido à análise do sistema desenvolvido neste trabalho, a taxa de fidelidade da marca d’água extraída novamente atingiu 100%, comprovando sua eficácia até mesmo em situações hostis.

3.4.2 Análise dos métodos de compressão

Durante o desenvolvimento da aplicação, as técnicas BWT, MTF e código de Huffman (descritas na seção 2.3) foram implementadas. O resultado da execução destes algoritmos foi confrontado com a compressão obtida pelo método heurístico.

Como strings que representam marcas d'água geralmente possuem tamanho reduzido, nesta avaliação foram utilizados textos com esta característica. Este fator influenciou diretamente a eficácia dos métodos de compressão. A string "Teste de esteganografia" (23 caracteres), por exemplo, é compactada para um texto de 16 caracteres, através do algoritmo heurístico. Por outro lado, um código baseado na metodologia de Huffman gerou uma string com 42 caracteres.

Assim, visando alcançar melhores resultados, optou-se pela abordagem heurística na construção do software.

3.4.3 Publicações

Além dos resultados descritos, um artigo sobre o mesmo tema foi produzido durante o desenvolvimento do presente trabalho, provando a capacidade de pesquisas acadêmicas na produção de soluções robustas para esteganografia em som. Este artigo foi publicado em um congresso nacional voltado especificamente a estudos na área de áudio (AQUINO; HOPPE; BRANDT, 2011).

4 CONCLUSÕES

Este trabalho propôs o desenvolvimento de um sistema de ocultação de dados em áudio através de técnicas de compactação e espalhamento espectral, onde o objetivo foi inserir o maior número de informação no sinal hospedeiro, sem que isso gerasse degradação da qualidade perceptual do áudio. Alguns conceitos de modulação de sinal para transmissão de dados, campo de atuação da engenharia de telecomunicações, foram empregados durante a construção do software.

A aplicação foi implementada com a linguagem de programação C#, utilizando a versão 3.5 do .NET Framework . O modelo psicoacústico e as transformadas para análise do áudio presente no sistema fazem parte de uma biblioteca externa, referenciada no projeto. A interface do sistema foi construída através com o framework WPF, possibilitando a criação de controles visuais com maior usabilidade.

O componente psicoacústico utilizado no sistema para analisar o áudio limita sua capacidade de inserção de dados, pois o tamanho das janelas de processamento fixado em sua estrutura é grande. A utilização de uma técnica de compressão de dados voltada para a compactação de textos pequenos, típicos em marcas d'água, auxilia o programa a contornar esta dificuldade. Por outro lado, esta abordagem aumenta a segurança e a confiabilidade da aplicação, pois cada símbolo da mensagem secreta é escondido em diversas partes da mesma janela. Desta forma, um agente malicioso precisa atacar várias partes do som para conseguir causar dano efetivo à informação oculta, ação que conseqüentemente também degradará a qualidade do áudio e denunciará que o sinal original foi violado.

A compatibilidade do sistema, limitada a arquivos WAV, também foi identificada como uma limitação em sua utilização. A harmonização da funcionalidade com formatos de áudio mais compactos aumentaria a portabilidade do projeto, pois sons em formato WAV tendem a ocupar um espaço grande de armazenagem em disco.

A bibliografia reunida revelou que os estudos acerca deste tema dividem-se entre as soluções focadas na implementação das técnicas de esteganografia em áudio e as que enfatizam estudos mais aprofundados do processo para aumentar sua segurança. Esta realidade demonstra que projetos concentrados em desenvolvimento de software possuem melhores índices de processamento de dados, porém restringem-se ao uso de técnicas mais simples que também permitem a resolução do problema.

A ferramenta desenvolvida aproximou estas duas vertentes de interesse, unindo a

eficiência computacional dos projetos de caráter experimental, com os ganhos proporcionados pelas pesquisas que abrangem o conhecimento teórico envolvido na composição de aplicações esteganográficas. Testes envolvendo transmissão através do ar, não previstos durante a especificação do software, foram realizados e apresentaram bons indicadores. Os resultados obtidos tornam a aplicação interessante em qualquer nicho de mercado em que a venda de conteúdo de áudio possa ser enriquecida por material adicional, sem necessidade de utilização de espaço extra no armazenamento.

4.1 EXTENSÕES

Algumas sugestões de extensões deste trabalho são listadas a seguir:

a) tornar o sistema compatível com o formato de áudio MP3;
b) otimizar o algoritmo de compactação com um dicionário de palavras em português;
c) implementação de técnicas de criptografia para prover maior segurança às informações ocultadas no sinal de áudio;
d) criação de um canal de comunicação seguro, através de um hardware para emissão e recepção de ondas de rádio que possuam mensagens secretas em seu espectro;
e) geração de selos que comprovem a legitimidade de músicas adquiridas via sistemas virtuais, através da adição de marcas d'água com informações relevantes à comprovação de propriedade nos áudios.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLA, Shahidan et al. A genetic-algorithm-based approach for audio steganography. **World Academy of Science, Engineering and Technology**, [S.l.], n. 54, p. 360-363, June. 2009. Disponível em: <http://www.waset.org/journals/waset/v54/v54-63.pdf>. Acesso em: 02 set. 2010.

ABEL, Jürgen; TEAHAN, William. Universal text preprocessing for data compression. **IEEE Transactions on Computers**, [S.l.], v. 54, n. 5, p. 497-507, May 2005.

ALBUQUERQUE, Célio V. N.; BRAZIL, Wagner G.; JULIO, Eduardo P. Esteganografia e suas aplicações. In: _____. **Esteganografia e suas aplicações**. Rio de Janeiro: SBC, 2007. cap. 2, p. 54-102.

ALSALAMI, Mikdam A. T.; AL-AKAIDI, Marwan. M. Digital audio watermarking: survey. In: EUROPEAN SIMULATION MULTICONFERENCE, 17th, 2003, Nottingham, UK. **Proceedings...** Nottingham: [s.n.], 2003. p. 1-14.

AQUINO, Luiz D.; HOPPE, Aurélio F.; BRANDT, Paulo R. Sistema de ocultação de dados em áudio através de técnicas de compactação e espalhamento espectral. In: CONGRESSO DE ENGENHARIA DE ÁUDIO, 9., 2011. **Anais...** São Paulo: AES Brasil, 2011. p. 128-131.

BENDER, Walter et al. Techniques for data hiding. **IBM Systems Journal**, v. 35, n. 3-4, p. 313-336, Feb. 1996. Disponível em: <http://www.research.ibm.com/journal/sj/mit/sectiona/bender.html >. Acesso em: 09 set. 2010.

BURROWS, Michael; WHEELER, David J. **A block-sorting lossless data compression algorithm**. Palo Alto: Digital Systems Research Center, 1994. Disponível em: <http://www.hpl.hp.com/techreports/Compaq-DEC/SRC-RR-124.pdf>. Acesso em: 12 set. 2010.

CHENG, Wei Q. et al. **Robust audio steganography using direct-sequence spread spectrum technology**. [Vancouver], 2007. Disponível em: <http://courses.ece.ubc.ca/412/term_project/reports/2007-fall/Robust_Audio_Steganography_Using_Direct-Sequence_Spread_Spectrum_Technology.pdf>. Acesso em: 13 ago. 2010.

COOK, Robert P. Heuristic compression of an english word list. **Software - Practice and Experience (SPE) Journal**, New York, v. 35, n. 6, p. 577-581, May 2005.

COX, Ingemar J.; MILLER, Matt L. Electronic watermarking: the first 50 years. In: IEEE WORKSHOP ON MULTIMEDIA SIGNAL PROCESSING, 4th, 2001, Cannes, France. **Proceedings**... Cannes: [s.n.], 2001. p. 225-230. Disponível em: <http://www.cs.ucl.ac.uk/staff/ingemar//Content/papers/2001/mmsp01.pdf >. Acesso em: 03 abr. 2011.

COX, Ingemar J.; MILLER, Matt L.; BLOOM, Jeffrey A. **Digital watermarking**. San Francisco: Morgan Kaufmann Publishers, 2002.

FERNANDES, José J. G. **Implementação de espalhamento espectral por seqüência direta**. 2002. 140 f. Dissertação (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre. Disponível em: <http://www.lume.ufrgs.br/bitstream/handle/10183/3582/000390127.pdf?sequence=1>. Acesso em: 11 set. 2010.

GARCIA, Ricardo A. **Digital watermarking of audio signals using a psychoacoustic auditory model and spread spectrum theory**. 1999. 116 f. Thesis (Master of Science in Music Engineering Technology) - School of Music, University of Miami, Coral Gables. Disponível em: <http://www.ragomusic.com/publications/ragothesismiami_nocode.pdf>. Acesso em: 11 set. 2010.

GORDY, James D.; BRUTON, Leonard T. Performance evaluation of digital audio watermarking algorithms. In: MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS, 43rd, 2000, Lansing. **Proceedings...** Lansing: [s.n.], 2000. p. 456-459. Disponível em: <http://www-mddsp.enel.ucalgary.ca/People/gordy/MWSCAS.PDF>. Acesso em: 31 ago. 2010.

JEHAN, Tristan. **Creating music by listening.** Cambridge, 2005. Disponível em: <http://web.media.mit.edu/~tristan/Papers/PhD_Tristan.pdf>. Acesso em: 30 mar. 2011.

KABAL, Peter. **An examination and interpretation of ITU-R BS.1387**: perceptual evaluation of audio quality**.** Montreal, 2002. Disponível em: <http://www.mp3-tech.org/programmer/docs/kabalr2002.pdf>. Acesso em: 14 maio 2011.

KHAYAM, Syed A. **The discrete cosine transform (DCT):** theory and application. East Lansing, 2003. Disponível em: <http://www.egr.msu.edu/waves/people/Ali_files/DCT_TR802.pdf>. Acesso em: 23 maio 2011.

KOBUSZEWSKI, André. **Protótipo de software para ocultar textos compactados em arquivos de áudio utilizando esteganografia**. 2004. 51 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

MALVAR, Henrique S. **A modulated complex lapped transform and its application to audio processing**. Redmond, 1999. Disponível em: <http://research.microsoft.com/pubs/69702/tr-99-27.pdf>. Acesso em: 05 mar. 2011.

PAINTER, Ted; SPANIAS, Andreas. Perceptual coding of digital audio. **Proceedings of the IEEE**, [S.l.], v. 88, n. 4, 2000. Disponível em: <http://ieeexplore.ieee.org/iel5/5/18261/00842996.pdf>. Acesso em: 31 mar. 2011.

PACHECO, Marco A. C. **Algoritmos genéticos:** princípios e aplicações. Rio de Janeiro, jul.1999. Disponível em: <http://www.ica.ele.puc-rio.br/Downloads/38/CE-Apostila-Comp-Evol.pdf>. Acesso em: 02 set. 2010.

SANCHES, Ademir M. **Um estudo de compactação de dados, com a implementação do método de LZW**. 2001. 107 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) – Departamento de Ciência da Computação, Universidade Estadual do Mato Grosso do Sul, Dourado. Disponível em: <http://www.comp.uems.br/trab/pfc/downloads/2001-1.pdf>. Acesso em: 12 set. 2010.

SCHÜTZ, Cristiano A. **Sistema de esteganografia em áudio digital que utiliza técnicas eficientes de inserção de dados**. 2009. 97 f. Dissertação (Mestrado em Engenharia Elétrica) - Curso de Pós-Graduação em Engenharia Elétrica, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre. Disponível em: <http://tede.pucrs.br/tde_busca/arquivo.php?codArquivo=2132>. Acesso em: 12 ago. 2010.

SKLAR, Bernard. **Digital communications:** fundamentals and applications. 2nd ed. Upper Saddle River: Prentice-Hall, 2001.