GENETIC BASED SUBSTITUTION TECHNIQUES FOR AUDIO STEGANOGRAPHY

MAZDAK ZAMANI

UNIVERSITI TEKNOLOGI MALAYSIA

GENETIC BASED SUBSTITUTION TECHNIQUES FOR AUDIO STEGANOGRAPHY

MAZDAK ZAMANI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

DECEMBER 2010

To my beloved grandmother

To my beloved father and mother

To my beloved sister and brother

ACKNOWLEDGEMENTS

I would like to thank my principal supervisor, Prof. Dr. Azizah Bt Abdul Manaf for her guidance during my research and study. Her perpetual energy and enthusiasm in research had motivated all his advisees, including me. In addition, she was always accessible and willing to help her students with their research. As a result, research life became smooth and rewarding for me. I would also like to thank my co-supervisor, Dr. Rabiah Bt Ahmad, for her support and encouragement. Her contribution by far has dominantly influenced my focus and motivation toward it. I also would like to thank Universiti Teknologi Malaysia (UTM) and the Ministry of Science & Technology and Innovation Malaysia (MOSTI), for funding my research and study. I especially want to thank Prof. Dr. Ghazali bin Sulong, and Prof. Dr. Othman Omran Khalifa as my thesis committee members. They read my thesis rigorously and helped me to correct it. I was also delighted to interact with Dr. Ahamd Reza Naghsh Nilchi by attending his classes and having him as my supervisor when I was doing my Master. His insights to signal processing is second to none.

My deepest gratitude goes to my family for their unflagging love and support throughout my life; this thesis is simply impossible without them. I remember many sleepless nights which my grandmother accompanied me when I was doing my homeworks. I remember, most of all, her delicious dishes. I am indebted to my father for his care and love. He worked, days and nights, to provide our family everything. I cannot ask for more from my mother, as she is simply perfect. I have no suitable word that can fully describe her everlasting love to me. I remember her constant support when I encountered difficulties. I also would like to thank my friends all who helped and supported me during study. I am also proud of my friends in my hometown, Khorramabad. They are all talented and really deserve more than what they are already given.

ABSTRACT

Steganography is a form of security technique through obscurity; the science and art of hiding the existence of a message between sender and intended recipient. Steganography has been used to hide secret messages in various types of files, including digital images, audio and video. The three most important parameters for audio steganography are imperceptibility, payload (bit rate or capacity), and robustness. Any technique which tries to improve the payload or robustness should preserve imperceptibility. The noise which is introduced due to bit modification would limit payload. This research focuses on audio steganography, particularly with respect to Waveform Audio File Format (WAV) files. The Least Significant Bit (LSB) is used as a type of substitution technique. In order to increase the payload and improve the imperceptibility, Genetic Substitution-Based Audio Steganography (GSBAS) was proposed. GSBAS is based on LSB and enhanced with the concept of genetic algorithm. GSBAS is benchmarked against Ordinary Substitution-Based Audio Steganography (OSBAS). The results showed that in comparison with OSBAS, the payload is considerably increased and Peak Signal-Noise Ratio (PSNR) is noticeably raised. The experiments showed that GSBAS gives promising result with high average payload and imperceptibility.

ABSTRAK

Steganografi adalah bentuk teknik keselamatan melalui ketidakjelasan yang melibatkan bidang sains dan seni yang menyembunyikan kewujudan mesej antara penghantar dan penerima. Steganografi digunakan untuk membenam dan menyembunyikan mesej rahsia dalam pelbagai jenis fail seperti gambar digital, audio dan video. Tiga parameter penting dalam pengukuran yang digunakan adalah ketidakjelasan, muatan dan ketahanan. Setiap teknik yang cuba untuk menambahbaik muatan dan ketahanan mesti mempastikan kualiti ketidakjelasan tidak tergugat. Gangguan yang terhasil akibat daripada perubahan bit akan menghadkan muatan. Penyelidikan ini tertumpu pada steganografi audio terutama berkaitan dengan Format Gelombang Fail Audio (WAV) di mana jenis teknik substitusi yang digunakan dalam kajian ini adalah Setidaknya Bit Signifikan (LSB). Untuk meningkatkan muatan dan memperbaiki ketidakjelasan, Steganografi Audio Berasaskan Pengantian Genetik (GSBAS) dicadangkan. GSBAS adalah lebih cekap berbanding Steganografi Audio Berasaskan Pengantian (OSBAS). Keputusan ujian menunjukkan bahawa berbanding dengan teknik LSB substitusi mudah, terdapat peningkatan ketara pada muatan dan peningkatan jelas pada Nisbah Isyarat-Hingar Puncak (PSNR). Hasil kajian menunjukkan bahawa kaedah GSBAS memberikan hasil yang memberangsangkan dengan pengukuran muatan dan ketidakjelasan yang tinggi.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	-	Artificial Intelligence
bps		Bit Per Sample
BPSK	-	Binary Phase Shift Keying
CD	-	Compact Disc
CWT	-	Continuous Wavelet Transform
dB	-	Decibel
DCT	-	Discrete Cosine Transform
DE	-	Differential Evolution
DFT	-	Discrete Fourier Transform
DVD	-	Digital Video Disc
DWT	-	Discrete Wavelet Transform
EOF	-	End-of-File
GA	-	Genetic Algorithm
GSBAS	-	Genetic Substitution Based Audio Steganography
HAS	-	Human Auditory System
HVS	-	Human Visual System
Hz	-	Hertz
Kbits	-	Kilo bits
Kbps	-	Kilo Bit Per Sample
KHz	-	Kilohertz
LSB	-	Least Significant Bit
MDEC	-	Minimizing the Distortion in the Equivalence Class
MPEG	-	Moving Picture Experts Group
MSE	-	Mean-Square-Error
OBS	-	One Bit per Sample
OSBAS	-	Ordinary Substitution Based Audio Steganography
PSNR		Peak Signal-to-Noise Ratio

RIFF	-	Resource Interchange File Format
SNR	-	Signal-to-noise ratio
STFT	-	Short-Time Fourier Transform
WAVE	-	Waveform Audio File Format
WMSE	-	Worst Mean-Square-Error
XPSNR	-	Worst PSNR

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Steganography and watermarking techniques embed information in a media in a transparent manner. Steganography is a technique for covert information, but watermarking may not hide the existence of the message from third persons (Neeta *et al.,* 2006; Cvejic *et al.,* 2004). Watermarking usually requires robustness to withstand against attacks intended to remove or destroy the hidden message from the watermarked media as well as preserving the carrier signal quality (Bhattacharyya *et al.,* 2010; Scagliola *et al.,* 2009). This makes watermarking appropriate for those applications where the knowledge of a hidden message leads to a potential danger of manipulation (Michael *et al.,* 2003; Avcibas *et al.,* 2003; Naji *et al.,* 2009; Yusnita *et al.,* 2007). The most well-known examples of steganography go back to ancient times when Histiaus shaved his slave's head, and then he tattooed a message on his scalp. After that his hair had re-grown the tattooed message was disappeared. He was going to call his men to attack to the Persians (Huayin *et al.,* 2008; Ricardo *et al.,* 1999; Yu *et al.,* 2010; Emelia *et al.,* 2008).

Steganography is the study of methods for hiding the existence of secondary information in the presence of primary information in a way which neither affects on the size nor results in perceptual distortion (Qiao *et al.,* 2009; Ganeshkumar *et al.,* 2009; Francia *et al.,* 2006; Petrovic *et al.,* 2009; Liu *et al.,* 2009). The secondary information is referred to as hidden message, hidden file or hidden information while primary information is referred to as carrier, host or original signal, before embedding and stego signal, file, bit stream or sequence, after embedding (Basu *et*

al., 2010; Khairullah *et al.,* 2009; Alla *et al.,* 2009; Changder *et al.,* 2009; Qi *et al.,* 2009).

Watermarking techniques are principally context-specific, that means, the algorithms must be designed regarding the media type of the data to be watermarked (Lie *et al.,* 2006; Wu *et al.,* 2006). Therefore, watermarking indicates a specific application of steganographic techniques (Baras *et al.,* 2006; Chang *et al.,* 2005a, 2005b, 2007). Specifically, the additional requirement for robustness of digital watermarks against attacks or manipulations during the data processing entails a lower payload of the watermarking methods compared to steganographic algorithms (Lemma *et al.,* 2008; Kejariwal *et al.,* 2006; Wu *et al.,* 2005; Ababneh *et al.,* 2005).

1.1 Fundamental properties

A fundamental tradeoff exists between three key variables: robustness, capacity and imperceptibility which restrict steganography designers. However in some application, computational time in determining the efficiency of a steganography technique is a crucial factor. In some applications, like broadcast monitoring, it is needed real time processing, and thus delays are intolerable under any circumstances (Andres *et al.,* 2002).

1.1.1 Imperceptibility

Imperceptibility is the perceptual similarity between the host and stego audio signal. In audio steganography, imperceptibility is evaluated as an audible distortion caused by signal modifications. There are different methods to measure imperceptibility (listening test, PSNR, and so on). In order to meet fidelity constraint of the embedded information, the perceptual distortion introduced by embedding should not be above the threshold estimated based on the Human Auditory System (Andres *et al.,* 2002).

1.1.2 Robustness

The ability of embedded data or watermark for withstanding against intentional and unintentional attacks is measured as robustness. Unintentional attacks are generally common data manipulations whereas intentional attacks include media degradations such as resizing, and filtering attacks. Robustness is required in some applications that a set of signal processing modifications is predefined, while in some other applications robustness is not desirable and those techniques are so called fragile audio steganography techniques (Nedeljko *et al.,* 2004).

1.1.3 Capacity (Payload)

Payload of an information hiding technique indicates the amount of data that an information hiding technique can successfully embed without introducing perceptual distortion in the changed media. Payload, capacity or bit rate is usually measured in bits per second (Nedeljko *et al.,* 2004).

1.2 Digital data hiding applications

There are many applications for information hiding in today's world. However, information hiding can be used in ethical ways; there are some ways that digital data hiding could be misused. Sometimes a method cannot be easily categorized in either of steganography or watermarking categories. Occasionally, the boundaries between these two disciplines have been blurred. However category can be specified by identifying an application which the method is used for. So, without any classification of the application, most important and common application of data hiding is presented (Chun, 2007).

1.2.1 Secret communication

Some people may use information hiding in order to hide data and be in confidential communication. For example, steganography is a trusted way for those who want to have an undisclosed communication. This is actually the area of steganography rather than watermarking.

1.2.2 Secure storage

Another use of information hiding is in the area of security storage. Obviously many types of sensitive information such as medical records of patients or prescription drug information need security during storage and transmission because in case those are accessible for unauthorized person could lead to illegal activities as identity theft as well as insurance fraud (Boneh *et al.,* 1998).

1.2.3 Covert communication

Some people or organizations need a covert communication for their business operations. For example, military can use information hiding for some technical activities as sending battle plans that if they fall into the wrong hands, then entire tactic would be compromised (Kurak *et al.,* 1992; Kirovski *et al.,* 2002).

1.2.4 Fingerprinting

The recipients or originator of a specific copy of media file could be traced by watermarking. The applied technique has to comprise a high robustness against intentional and unintentional attacks (Wang *et al.,* 2003; Yacobi *et al.,* 2001; Dittmann *et al.,* 2000). For example, before distributing numerous copies of multimedia products to recipients, they can be watermarked by different serial or

identity numbers (Wu *et al.,* 2004; Trappe *et al.,* 2003; Chenyu *et al.,* 2003; Hong *et al.,* 2003).

1.2.5 Copy right protection

While a considerable portion of economic resources is dedicated to the creation of intellectual property, particularly in industrial societies, the cost of reproducing such intellectual creations typically constitutes only a small fraction of the creation (Cox *et al.,* 2003; Bloom *et al.,* 1999; Pan *et al.,* 1995). In the copy right protection, a watermark which contains the information of the owner is embedded into the host media (Noll *et al.,* 1993; Dittmann *et al.,* 2000). The watermark is supposed to be robust and enables the owner to prove his ownership in case is needed. Also with fragile watermarking, watermarks are used to verify if the host signals are tampered (Termont *et al.,* 2000; Termont *et al.,* 1999). Furthermore, in the copy control application, watermarks control access policy or limit to a certain copy (Depovere *et al.,* 1999; Kalker *et al.,* 2000; Craver *et al.,* 2001).

1.3 Rationale of work

Initially steganography techniques were developed for images. But later on researchers became interested in developing techniques for audio. Recently, several algorithms for audio steganography have been presented. Audio steganography is more secure due to the small number of audio steganalysis methods. Since currently most steganography techniques are for still images (Nedeljko *et al.,* 2004). To compare most robust technique and most payload technique, substitution technique and spread spectrum technique are introduced below. More details about data hiding techniques in written in Chapter 2.

1.3.1 Substitution technique

Basic substitution systems try to encode secret information by substituting insignificant parts of the cover by secret message bits; the receiver can extract the information if he has knowledge of the positions where secret information has been embedded (Bender *et al.,* 1996; Möller *et al.,* 1996; Gruhl *et al.,* 1996). Since only minor modifications are made in the embedding process, the sender assumes that they will not be noticed by a passive attacker (Kurak *et al.,* 1992). These approaches are common in steganography and are relatively easy to apply in image and audio. A surprising amount of information can be hidden into carriers imperceptiblly (Johnson *et al.,* 1998; Gerzon *et al.,* 1995; Van *et al.,* 1994).

1.3.2 Spread spectrum technique

Spread spectrum (SS) communication technologies have been developed since the 1950s in an attempt to provide means of low-probability-of-intercept and anti jamming communications. Pickholtz *et al.* (1982) define spread spectrum techniques as "means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by means of a code which is independent of the data, and a synchronized reception with the code at the receiver is used for dispreading and subsequent data recovery. Although the power of the signal which is transmitted can be large, the signal-to-noise ratio in any frequency band would be small. Even if parts of the signal could be removed in several frequency bands, enough information should be present in the other bands to recover the signal. Thus, spread spectrum makes it difficult to detect and/or remove a signal. This situation is very similar to a steganography system which tries to spread a secret message over a cover in order to make it impossible to perceive. Since spreaded signals tend to be difficult to remove, embedding methods based on spread spectrum should provide a considerable level of robustness." According to the landmark paper by Tirkel *et al.* (1993), spread spectrum methods are of increasing importance in the field of information hiding.

Although substitution technique is not as robust as other steganography techniques like spread spectrum technique, the payload of substitution techniques is incomparably higher. Mostly the payload of substitution techniques are more than 40000 bps, while more robust techniques like spread spectrum has a negligible payload that is only about 4 bps (Vinu and Vijayakumar, 2010).

1.4 Background of the problem

The requirements of audio steganography are visualized simply in magic triangle, shown in Figure 1.1. Inaudibility, robustness and the data rate (payload) are in the corners of this triangle. This figure illustrates the required trade-offs between the payload and the robustness, at the same time keeping the quality of steganography algorithm at an acceptable level.

It is not achievable to get a high payload and high robust technique at the same time in steganography. Hence, if it is desired to have a robust steganography algorithm, its payload will be low and vice versa, a steganography algorithm with high payload embedding is usually very fragile. Many applications exist which do not require high robustness in steganography (Nedeljko *et al.,* 2004).

Therefore, a problem remains when an application requires using a high payload technique, while the noise which the technique produces is not tolerable. In other word, there exists a problem to achieve a high payload technique with the noise which is as a result of bit modification.

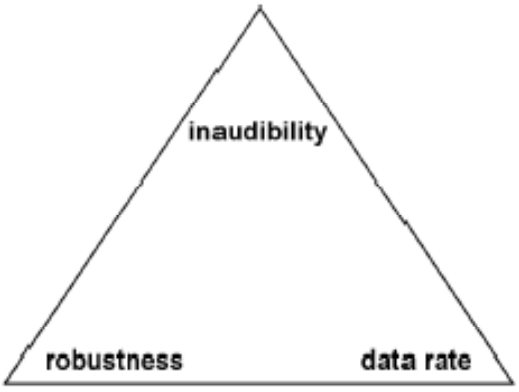

Figure 1.1 Magic triangles of the requirements of information hiding

1.5 Problem statement

Apart from robustness that principally is not desirable for substitution techniques (Nedeljko *et al.,* 2004); the only remaining measure to achieve high payload is imperceptibility. However substitution techniques comparatively are well-known in achieving high capacity, but have to satisfy quality condition yet. The distortion caused by substitution degrades quality.

In a theorical view, quality is measured by PSNR. As such by increasing PSNR, quality would be better. In another sensible view, the quality is measured by imperceptibility. Thus getting noise more imperceptible, quality would be better.

Therefore, to take advantage of a potential high payload, imperceptibility should be retained. In other words, improving imperceptibility, higher payload could be achieved. The general research question this research has aimed to answer is:

I. What is the optimum bit per sample rate (payload) for audio steganography techniques?

This question contains two following sub-questions:

i. What is the threshold of tolerable PSNR for audio steganography techniques?

ii. What is the maximum bit per sample rate (payload) which does not result in exceeding the tolerable level of PSNR for audio steganography techniques?

- Answering this question needs to conduct a listening test and to calculate PSNR for reasonable number of audio files whose level of noise is different. After that the threshold of tolerable PSNR for audio steganography techniques is determined by listening test, the level of PSNR for each bit per sample rate should be calculated. Finally, the bit per sample rate corresponding with the threshold of tolerable PSNR can be determined.

II. How can the quality of audio steganography techniques be improved?

This question contains two following sub-questions:

i. How to maintain high the PSNR when a higher payload is desired?

ii. How to maintain high the payload when a higher PSNR is desired?

- An efficient algorithm could reduce the distortion caused by substitution of samples bits. This reduction of distortion directly improves the PSNR, and indirectly increases the payload. Given the PSNR is already improved; the difference between prior PSNR and improved PSNR could be loaded by more embedding data, which indirectly increases the payload.

1.6 Research objectives

Since this research focuses on two of most important requirements of steganography which are payload and imperceptibility, the following objectives were attempted to be achieved in this study.

- To develop an efficient model in audio steganography in order to reduce the distortion.
- To increase the payload of substitution techniques of audio steganography techniques while maintaining imperceptibility.
- To improve the imperceptibility of substitution techniques of audio steganography techniques preserving the payload.
- To evaluate the performance of a steganography technique before embedding process.

1.7 Significance of the study

Significance of the research is derived from the result of this research that is listed as follow:

- An improved audio steganography method based on a novel.
- The correlation between imperceptibility and payload is determined.

1.8 Assumptions

In this study, it is assumed that, there are some applications for steganography in which payload and imperceptibility are needed but whose robustness is not desirable. Apparently, when the application is to hide the existence of communication, robustness is not important. For example suppose an internet group that is created to connect some students who took the same subject. The lecturer of the subject is registered as a member of the group and as a result can read public messages. The students arrange to embed the links of any website they may find and are related to their assignment. Obviously the secret is not the content of the message but the existence of the message.

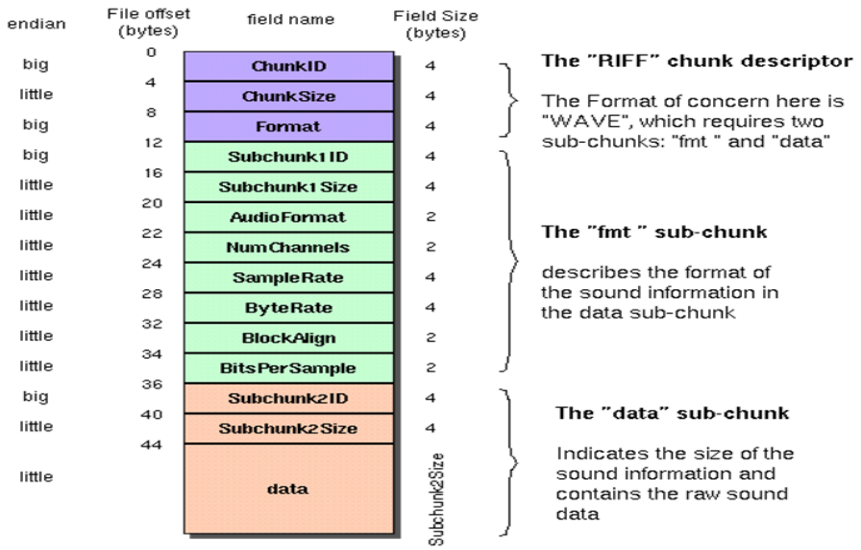

Figure 1.2 The WAVE file format

1.9 Scope of the study

The focus of this research is the audio steganography, particularly with respect to WAV files. This format of audio file is selected because that is the original format of other formats. In other words, the audio files which have other formats, either are converted from WAV format, or can be easily converted to WAV format.

The WAVE file format follows the specification of Microsoft for multimedia storage, as so called RIFF. As Figure 1.2 shows, any file created in RIFF specification consists of a file header and a sequence of data chunks. A WAVE file has only one "WAVE" chunk. This chunk has two sub-chunks: an fmt-chunk to specify the data format and a data-chunk which includes the actual sample data (Scott *et al.*, 2003).

In Figure 1.3, a WAVE file with bytes is assumed that are shown as hexa-decimal numbers.

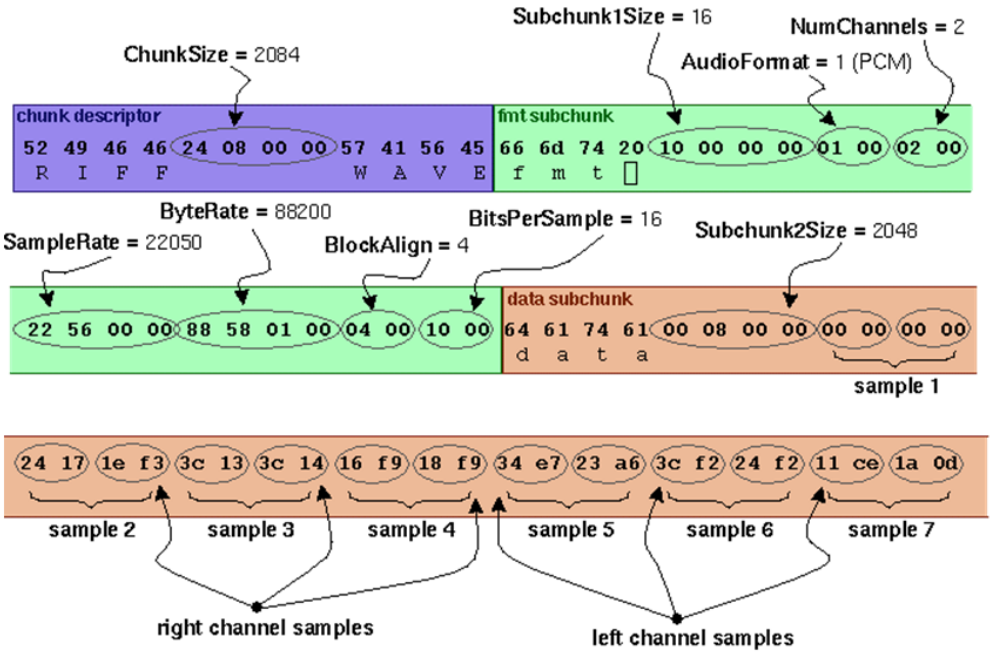

Figure 1.3 An example of WAV file

1.10 Thesis organization

In Chapter 1, an introduction of data hiding, especially for steganography, is given. A brief description of the requirements and applications for steganography is provided. Also steganography definitions are introduced. The objectives of this research are presented; in addition to the problem statement. In Chapter 2, a literature review on data hiding techniques is given. Since there are many types of classification of data hiding techniques, only the most popular techniques have been introduced.

Chapter 3 includes the approaches of the research and the theoretical issues related to the proposed method. Chapter 4 focuses on the implementation of this research and the results of the proposed method after testing a number of samples. Moreover, the results of high capacity steganography in an imperceptible manner are also presented.

In chapter 5, a discussion on the results of the study is given; the analysis of the proposed method and the threshold value are revealed. Besides that, a comparative study between the algorithm proposed in this research and other steganography techniques is also presented and discussed. In chapter 6, the contributions of this study and the proposed future works are given.