Text in Audio Steganography

Salima Raheem Nsayif

Computer Science & Information Technology

Baquba Technical Institute, Diyala, Iraq.

Receiving Date: 02-11-2010 **- Accept Date:** 11-01-2011

Abstract

Because of the great development which occurs in information and the security way of transfer them on world wide web, (Internet) and because of the great interest in keeping on the secret information that have transfer on this net, many methods are used to hide the information such as hiding information inside a sound.

In this paper, a proposed technique is adopted using non serial first bits of the wav format sound file to hide secret message according to a certain equation known by the sender and the receiver.

This paper concludes the possibility of hiding different text in different sound media like music, songs and others. Also, it is possible to make the cipher key variable by using a specific function, this will increases the security of the system, but this requires sending the specific function to the receiver in some way may be by embedding it in stego-image or stego-video clip.

1-Introduction

Recently ,due to the rapid expansion of the Internet, various kinds of digital media (like, picture, video, music, text, and program have increasingly been treated and transferred by many users all over in the world. However, it has been a critical problem to secure multimedia contents against illegal use. This is because such digital content can be easily copied or edited on a computer and delivered through the network by a third party without permission of the copyright owner. In order to solve this problem, data hiding has got great attention as a promising method that plays a complementary role to the conventional cryptographic techniques[1],[3].

2-Information Hiding

The Information hiding system is any method or technique that deals with processing information (data) and puts the output in shapes or forms of media, under the condition that it must not be visible in its new form for human observer [4]. It represents a vast array of techniques for hiding message in a variety of media for example in text , images , sound , signal and some other digitally representative code. The main categories of data hiding can be classified into two classes , the steganography and watermarking [2], as shown in figure (1). The steganography deals with the concealments of message in away that hiding the message objects in another or message [5].

Digital watermarking is to embed a signature within a digital image to signify origin or owner ship for the purpose of copyright protection. Once added watermark must resistant to removal , and such as reliably detected even after typical image transformations , rotation , translation , copying and quantization[4] [6].

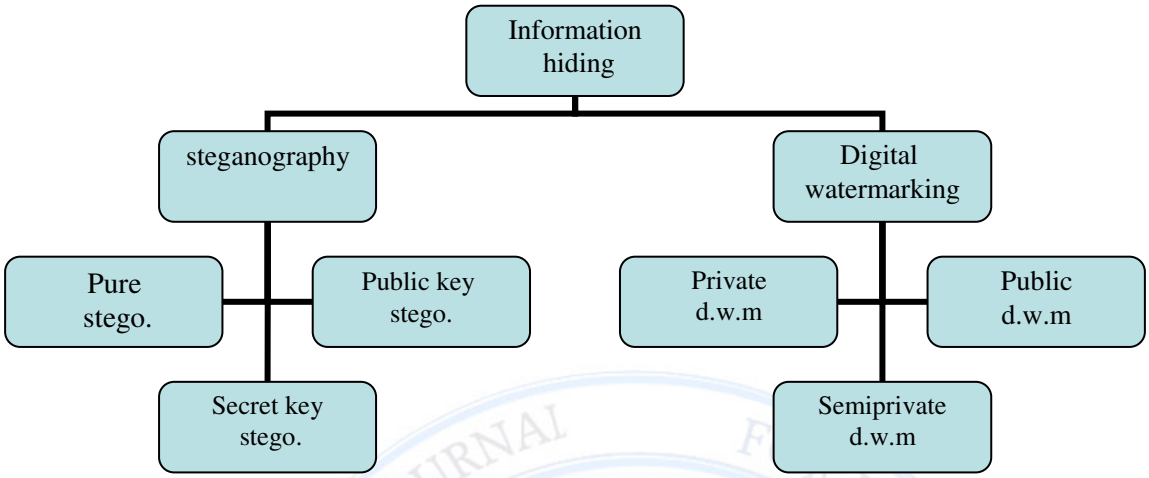


Figure (1) Classification Discipline of Data Hiding

Each data hiding technique consist of :-

1- The embedding algorithm.
2- The detector function.

The embedding algorithm is used to hide a secret message inside a cover (or carrier) document . The embedding process is protected by a key word so that only message which process the secret key word can access the hidden message[1] .

The detector function is applied to a (possibly modified) carrier and returns the hidden secret message as it illustrated in figure (2) [3].

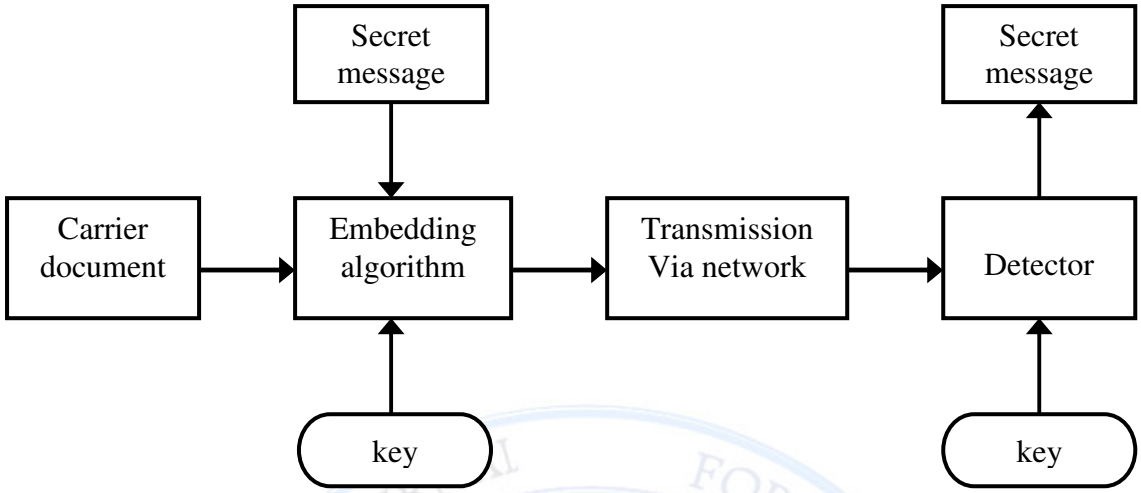


Figure (2) General Scenario for Hiding Information

2.1 Steganography

The most traditional definitions for steganography are the followings : [4] [6]

- Steganography is the art and science of communication in away, which hides the existence of the communication.
- Steganography is the art of passing in formation in a manner that the very existence of the message is unknown.
- Steganography an ideal word users would all be able to openly send encrypted mail or files to each other with no fear of reprisal.
- Steganography encompasses methods of transmitting secret messages through innocuous cover carriers in such a manner that the every existence of the embedded messages is undetectable.
- In data security , the concealment of the existence of message .

Within the past several years, there has been an exponential in the research community and industry focus towards Information hiding techniques as opposed to the traditional cryptography area [2].

Once of the primary drivers of the renewed interest in steganography is for mitigating copyright a buses . As audio , vidio and other works become more readily available in digital forms , the easy with which perfect copies can be made may lead to large – scale unauthorized copying .This Type of copying is naturally of great concern to the music film , book , and software publishing industries . There has significant growth in recent research concerned with digital watermarks or hidden copyright message and digital finger prints or hidden serial numbers [1] [5] .

Steganography includes a vast number of methods of secret communication that conceal the very existence of the message . Among these methods are invisible inks , microdots character arrangement (other than cryptographic methods of permutation on substitution), digital signatures, covert channels and spread spectrum communications [4].

2.2 Steganography methodology

Classical steganography is concerned With the ways of embedding a secret message (which might be a copyright mark , or a covert communication or a serial number) in an innocent cover message (such as video film , an audio recording , or computer code)[1]. A Key typically parameterizes the embedding , without know ledge of this key curare related one) it is difficult for a third part to detect or remove the embedding material . Once the cover object has material embedding in , it is called a stego – object . Thus , for example we might embed a mark in a cover text giving a stego –text , or embed a text in a cover image giving a stego – image , and so on [4].

It is a powerful idea , sine an enemy can not read the hidden information unless they can first detect it . some times steganography methods combine traditional method of cryptography with stganography there for , this process increases the security of the communication process .

2.3 Methods for Hiding Information

There are many different techniques used to hiding information. Some of these techniques are: [1],[5]

2.3.1 Hiding in Text

Documents may be modified to hide the information by manipulating positions of lines and words . HTML files can be use to carry information since adding spaces , tabs , "invisible "characters , and extra line breaks are ignored by web browsers . The extra spaces and line are not perceptible until revealing the source of the web page .

2.3.2 Hiding in Disk space

Other ways to hide information rely on finding unused space that is not readily apparent to an observer. Taking advantage of unused or reserved space to hold cover information provides a means of giding information without perceptually degrading the carrier. The way operating systems store files typically results in unused space that appears to be allocated to files , This "allocated "but available space is known as slack space .

2.3.3 Hiding in Image

Information can be hidden in many different ways in Images. to hide information ,straight message insertion many encode every bit of Information in the Image embed the message in busy areas .Those areas consisting of a great deal of illumination or color variation would be several times through out the image [4].

2.3.4 Hiding in Audio

Because of the range of the Human Auditory system (HAS) data hiding in Audio signals is especially challenging . The HAS perceives over arrange of power greater than one billion to one and range of frequencies greater than one thousand to one . Also, the auditory system is varying sensitive to additive random noise [2],[5].

Any disturbances in a sound file can be detected as a low as one part in ten million (80 dB below ambient level) . However , while the HAS a large dynamic range , it has a fairly small differential range –large sounds tend to drown quiet sounds .when performing data hiding on audio , one must exploit the weaknesses of the HAS , while at the same time being aware of the extreme sensitivity of the human auditory system [1] .

2.4 Audio Environments

When developing a data- hiding method for audio , one of the first considerations is the likely environments the sound signal will travel between encoding and decoding .

There are two main areas of modification which we will consider .First , the storage environment , or digital representation of the signal that will be used , and second the transmission pathway the signal might travel .

2.4.1 Digital Representation.

There are two critical parameters to most digital audio representations: sample quantization method and temporal sampling rate.

The most popular format for representing samples of high - quality digital audio is a 16-bit linear quantization, e.g, Windows Audio-Visual (WAV) and Audio Interchange File Format (AIFF)[2]

Popular temporal sampling rates for audio include 8khz (kilo hertz); 9.6KHZ, 10KHZ, 12KHZ, 16KHz, 22.05KHz, and 44.1KHz. sampling rate impacts data hiding in that it puts an upper bound on the usable portion of the frequency spectrum (if a signal is sampled af 8KHz, you cannot introduce modification that have frequency components above 4KHz) [3].

2.4.2 Transmission Environment

There are many different transmission environments that a signal might experience on its way from encoder to decoder. We consider four general classes for illustrative purposes [2],[4],[5] (see figure (3)).

The first is the digital end-to-end environment (figure 4-A). This is the environment of a sound file that is copied from machine to machine but never modified in any way. As a result ,the sampling is exactly the same at the encoder and decoder.

The next consideration is when a signal is resample to higher or lower sampling rate, but remains digital throughout (figure 4-B). This transform preserves the absolute magnitude and phase of most of the signal, but hangs the temporal characteristics of the signal.

The third case is when a signal is "played" in to an analog state, transmitted on a reasonably clean analog line and resample (figure 4-c). Absolute signal magnitude, sample quantization, and temporal sampling rat are not preserved.

The last case is when the signal is "played in to the air" and " resample with a microphone" (figure 4-D). The signal will be subjected to possibly unknown nonlinear modifications resulting in phase changes, amplitude changes, riff of different frequency components, chooses, etc.

Signal representation and transmission path way must be considered when choosing a data-hiding method. Data rate is very dependent on the sampling rate and the type of sound being encoded.

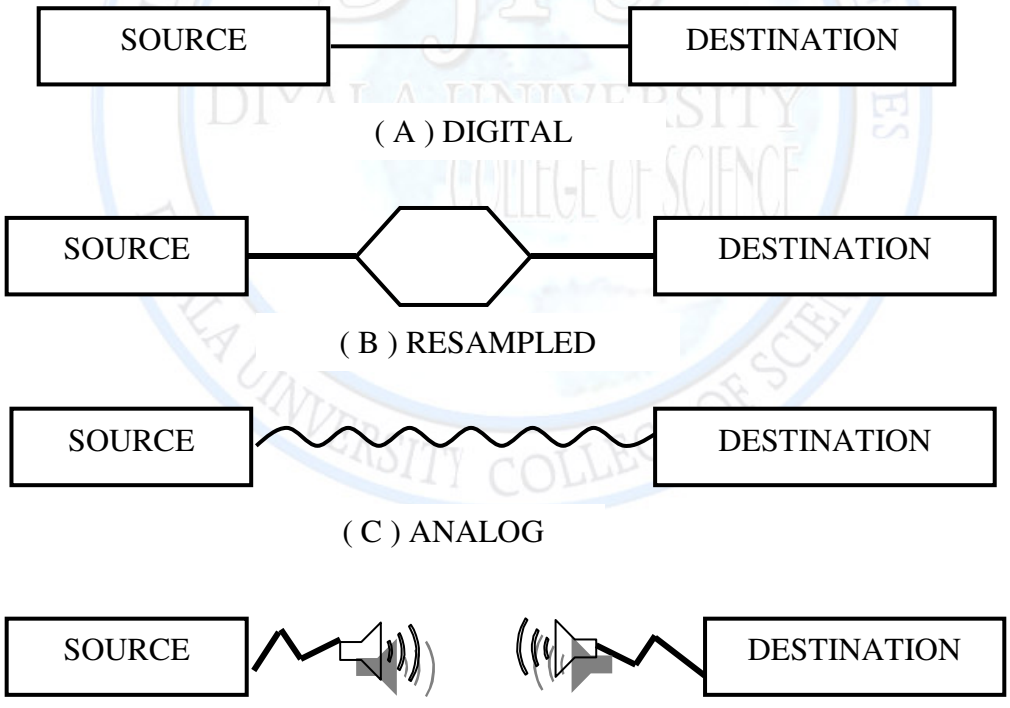


Figure (3) Signal Transmission Environment

3. The Proposed Technique

The suggested technique consists of two main operations, cipher operation and decipher operation. These two operations are programmed in visual basic programming language. (appendix (A) and appendix (B) respectively). These operations can be detailed as follows:

3.1 Hiding Process:

In the hiding operation we used two files. The First file is called (cover file) and the second file called (embed file) as the following .

a- cover file:-

The audio that be extension from the type (WAV) . The cover file must be larger than embed file with four times.

b- Embed File: -

The secret message is text file and must be of size less than cover file size by four times .

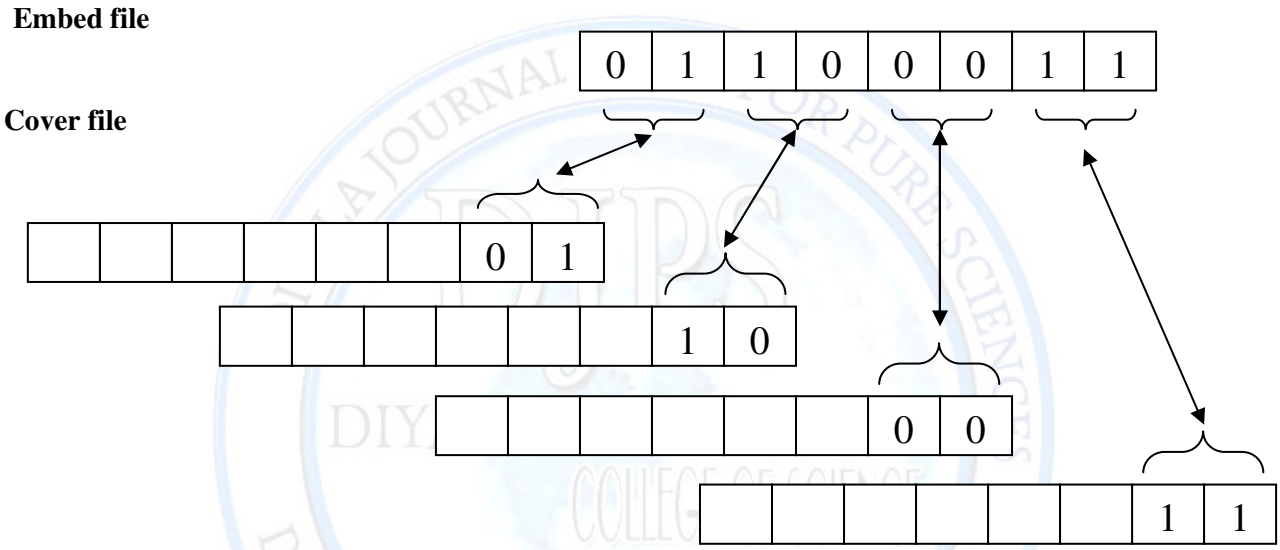
c- embedding method:-

It consists of the following steps, as shown in figure (4):

- Step one : open the first file (cover file) which from the type binary to read byte by byte. Then open the second file (embed file) which from the type binary too .
- Step two : the hidden information be in audio data not be in the header.
  - Step three: take the extension file which is (byte) and store it in the first four byte from the cover file so every byte from the embed file hide in (4-byte) from the cover file.
  - Step four : hide size of the file in the bytes which follow the present bytes.
  - Step five : with the depend on size of the file we work with hiding as he following points:-

1- Take the first byte from the embed file and work on partition it to four part . length every part is (2 bits) , then with work multiple to the byte by number to get on the first part of (2-bit), Then work shift by byte to the right to get the second part of (2- bit) as so on ….

2-we embedded first (2- bits) from the byte in the text file with First byte in the cover file and embedded second (2-bits) from the text file with second byte from the cover file and embedded three (2 bits) and four (2 bits) as explained in figure (4).

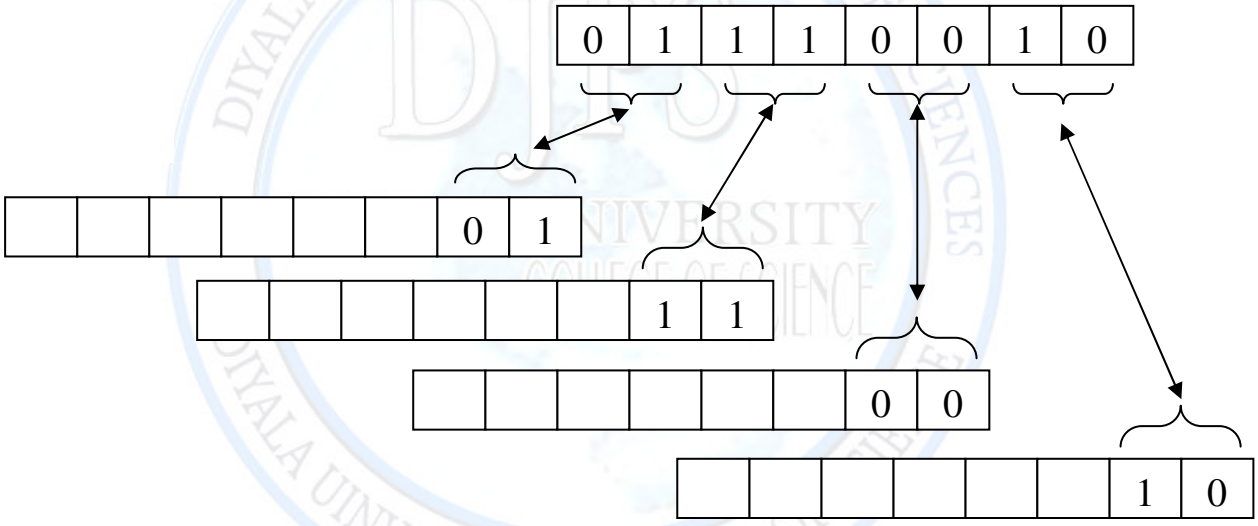


Figure(4) Embedding process

3.2Extracting Method

It consist of the following steps, as shown in figure (5):-

- Step one : - we work with the extension file which be length (3 byte) and which be store in the first (12 bytes) from the cover file because every (4 bytes) is represented by one byte .
- Step two:- we work with the stegocover file according to the size as (Byte , 2 byte , 3 byte …etc).
- Step three:- by using the size we work with the size of file.



Figure(5) Extracting process

4. Conclusions & Suggestions

1- Its possible to make the cipher key as variable value by using a specific function, this will increases the security of the system , but this requires sending the specific function to the receiver in some way may be by embedding it in stego-image in manner .

2- The implementation of the algorithm can be developed to be capable to work with other type of Audio like (Mp3,Mp4).

3- Developing the algorithm to imply the encryption in addition to hiding .

4- The choice of (message and cover) size is an important point therefore the check stage is added in the beginning of project .

References

1- Birgit P. Fitzmann, "Information Hiding Terminology", second edition, Springer-Verlag Berlin Heidelberg, (2003).
2- Corinna John, " Hiding Data in Wave Audio Files", (2005).
3- Moon S. K. , kawitkar r. s. "Data Security Using Data Hiding" , Conference on computational intelligence and multimedia applications, 13-15 Dec. 2007.
4- Shuichi TAKANO, "Data Hiding via Steganographic Image " handbook second edition, AP professional, Academic Press, (2006).
5- Smith M. J., " Information Hiding ", proceeding of first international workshop lecture . Fifth edition , addison-Wesley Longman Inc., (1996).
6- TU Tran, Mills College, " Steganography the art of hiding data ", IEEE Transaction on cipher systems technology, vol.8, no. 2, Iette 2, pp 166-171, March (2005).