Steganographic Algorithm for hiding Text in Digital Speech Signal

Dr. Hayder Mahmood Salman

Abstract

In this paper, a method of a steganography is introduced for hiding secret data in digital speech signal. Hiding in speech becomes a challenging discipline, since the Human Auditory System is extremely sensitive.

The proposed method is used to embed binary sequence with high data rate by modulating the amplitude of WAVE file. The embedding process utilizes the amplitude modulation of the cover signal; the manipulation of the sample depends on its two previous samples and next two samples. By using this hiding method, good hiding rate is achieved.

All of the fidelity measures (Mean Absolute Error, Mean Square Error, Signal to Noise ratio and Peak Signal to Noise ratio) obtained in the tests have indicated good results for PSNR. In this method the reconstructed data is exactly the same as secret data.

الخلاصة

في هذا البحث جرى عرض طريقة لأخفاء البيانات السرية في ملفات الاشارة الصوتية الرقمية. أن الأخفاء في الصوت هو بالغ الدقة ,لأن النظام السمعي للأنسان حساس جدا. الطريقة المقترحة استخدمت لأخفاء البيانات السرية عن طريق تعديل سعة عينات الملف الصوتي , ويعتمد التعديل لأية عينة على العينتين السابقة و العينتين اللاحقة , وقد حققت هذة الطريقة نسبة جيدة من الاخفاء .

النظام المقترح تم اختباره باستخدام أربعة مقاييس معولية هي متوسط الخطأ المطلق ومتوسط الخطأ المربع ونسبة الاشارة الى الضوضاء ونسبة الاشارة العليا الى الضوضاء. كل المقاييس المعولية التي استخدمت اظهرت قيم جيدة لنسبة الاشارة العليا الى الضوضاء . أما البيانات المسترجعة فكانت بالضبط هي نفسها البيانات السرية التي تم اخفاءها بهذه الطريقة.

1. Introduction

Embedding secret data in digital sound is generally more difficult than embedding information in digital images. Human auditory system is extremely sensitive; perturbations in a sound file can be detected as low as one part in 10 million.

Trade off exists between the quantity of embedded data and the degree of immunity to host signal modification. By constraining the degree of host signal degradation, a data hiding method can operate with either high embedded data rate, or high resistance to modification, but not both [2].

Steganography is the art of hiding and transmitting data through apparently innocuous carries in an effort to conceal the existence of the data. Though steganography is an ancient craft, the onset of computer technology has given it a new life. Computer-based steganographic techniques introduce changes to digital covers to embed information foreign to the native covers. Such information may be communicated in the form of text, binary files [1].

There are three types of steganographic protocols: Pure Steganography, Secret key Steganography and Public key Steganography; The suggested algorithm use Secret key Steganography, in which the sender embed the message in to

selective cover and the receiver extract the secret message from the cover by using the same secret key.

Steganography can be classified according to media type, such as Steganography in Text, Video, Audio, Image, Disk space, ...etc. The work in this paper is considered with the Steganography in Audio.

In this paper, we built a system for hiding a text data in digital speech signal to embed binary sequence with high data rate by modulating the amplitude of a wave file working in time domain. The proposed system uses the concept of blind steganography which does not need original audio file in the extraction stage.

2. Digital Speech Signal Steganographic Techniques

Digital Speech Signal Steganographic Techniques may be grouped into five categories as follow [3]:

1. **Substitution Techniques:** Substitute redundant parts of a cover with a secret message. 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. The Substitution Technique was used in the suggested work.
2. **Transform Domain Techniques:** Embed secret information in a transform space of the signal.
3. **Spread Spectrum Techniques:** (SS) is a mean of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information.

4. **Statistical Techniques:** Encode information by changing several statistical properties of cover.
5. **Distortion Techniques:** Store information by signal distortion and measure the deviation from the original cover in the decoding step.

3. Digital Speech Signal Environments

Speech is the result of a very complex and not completely understood process. A "*concept*" that is formed in the brain is somehow converted to neural signals that travel to the muscles of the speech production mechanism. These components then produce an acoustic waveform that is radiated out from the head as the speech signal. Speech can be though of as being composed of elements of discrete information [4].

When air pulsates, sound waves are created; these waves are then registered as sound. The amplitude and frequency are the attributes of these waves [7].

The amplitude of a wave refers to half the distance between the wave's highest point and it is lowest. The larger the amplitude of a wave, the louder its volume, which is typically measured in decibels (dB). The frequency of a wave refers to how many times per second the wave transitions from its highest point to its lowest point and back again. This is typically measured in hertz (Hz), or number of cycles per second. Figure 1 shows a typical example of sound that oscillates several times. The human ear is sensitive to a wide range of sound frequency, normally about 20 Hz to about 20,000 Hz, depending on the person's age and health. This is called range of audible frequencies [6].

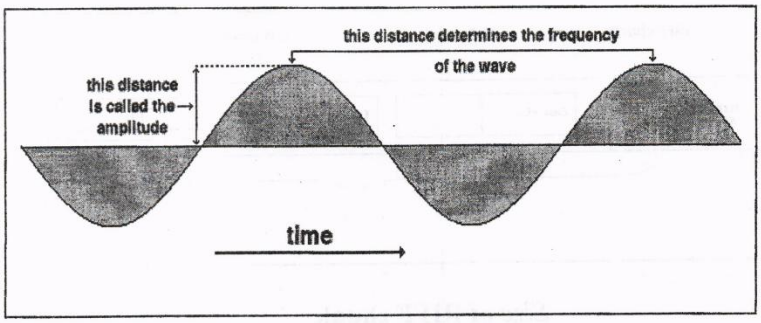


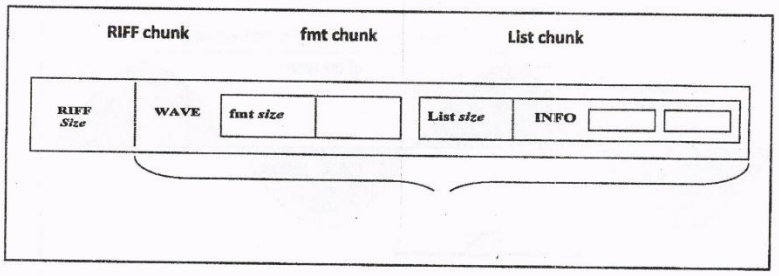
Figure 1. Sound wave, with amplitude and frequency [6].

Sounds from the real world can be recorded and digitized using an Analog-to-Digital (A/D) Converter. Digital sound storage involves two processes, the first process is sampling and the second process is quantization [5].

3.1 Wave Audio File Format

Wave is the native sound format used by Microsoft Windows. The over all structure is based on the general file format RIFF (Resource Interchange File Format) [4].

RIFF file consists of a collection of nested chunks. Each chunk contains a four-character code (such as: RIFF, fmt_, or list; shorter codes are padded with spaces). This code indicates the type of chunk, for example: an "fmt" chunk contains information about the format of the sound. Figure 2 illustrates the nesting of chunks in a RIFF file [4].



***Size* of RIFF chunk**

Figure 2. The nesting of chunks in a RIFF file [4].

A wave file consists of a number of chunks, each of these chunks includes:

- An identifier,
- The size of chunk in bytes and
- Any data associated with the chunk.

There are two chunks, a format chunk and the sample data chunk. Table 1 illustrate the Native Wave File Format [7].

Table 1. The Native Wave File Format [7].

Size in bytes	Description
4	Chunk type: RIFF
4	Total file size minus 8 bytes
4	RIFF contains type: WAVE
4	Chunk type: fmt_
4	Format chunk data length: usually 16
16	Format chunk data
4	Chunk type: data
4	Length of sound data (NO. of samples in file)
N	Actual sound samples

1. RIFF Header

The header consists of :

1-The characters "RIFF".

2-The size (in bytes) of the rest file.

3-The characters "WAVE".

After this, the file contains chunks, format chunk and data chunk [6].

2. Format Chunk

The format chunk has a chunk identifier whose length is 4 characters "fmt_", followed by the following fields [6]:

A. Chunk size (4 bytes).

B. Format tag (2 bytes).

C. Number of channels (2 bytes).

D. Samples per second (4 bytes).

E. Total number of bytes per second (4 bytes).

F. Block alignment (2 bytes).
G. Bits per sample (2 bytes).

3. Data Chunk

The data chunk contains the actual samples of the waveform [6];

- It starts with a chunk identifier of data (4 bytes),
- Followed by the chunk size in bytes (4 bytes),
- And finally the samples themselves (n bytes).

4. Proposed Audio Data Hiding System

This hiding method used for hiding high data rate operating in time domain. The proposed hiding algorithm consists of two modules:

1. Embedding module.
2. Extracting module.

Embedding module

This module is used to embed secret message within audio file, and it consists of the following stages:

1. Ciphering the secret message.
2. Extracting the header of the original audio file (cover).
3. The embedding process, in this stage, the manipulation of current sample depends on its two previous samples and the next two samples. The overall scheme of embedding is shown in the Figure 3. The suggested embedding process can be summarized as follow:

For each five successive host samples (S1, S2, S3, S4 and S5), determine:

- $\delta = S1 + S2 + S4 + S5$
- $\beta = \delta / 4$

- S3= { Round (β) if embedded bit=0 ; Round (β+1) if embedded bit=1 }

The steps of embedding process are demonstrated by algorithm (1).

Algorithm (1) Embedding process

Input : (Secret_bit) a vector represent secret message, (Len_msg) the length of secret bits, (Len_cover) the length of the cover, (Wav) a vector contains samples of the original file(cover).

Output : stego_ file.

Step 1 : *Write header of original audio data (cover) into stego_file*

Step 2 : *Set j=0, i=0*

Step 3 : *While j<= Len_msg and i <= Len_cover-2*

S1= Wav (i): S2= Wav (i+ 1): S3= Wav (i +2)

S4= Wav (i+ 3): S5= Wav (i +4)

$\delta = S1 + S2 + S4 + S5$

$\beta = \delta/4$

If secret_bit (j) = 1 then

S3= Round (β+1) ' Embedding 1

Else

S3= Round (β) ' Embedding 0

End of if

j=j +1: i = i+1

End of While

Step 4 : *Write Wav vector into stego_file.*

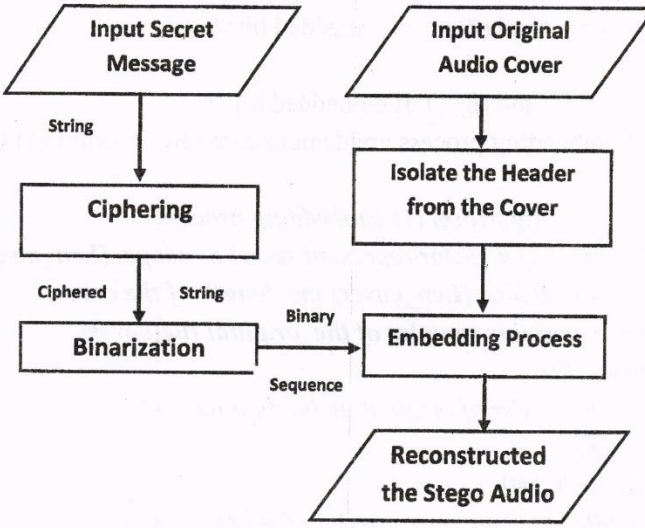


Figure 3. **The steps of the embedding module of the suggested Audio steganography system.**

Table 2, illustrates the embedded bits (110).

Table 2. Embedding process.

Position	*Embedded Bit*	*S1*	*S2*	*S3*	*S4*	*S5*	*S3*
15670	1	129	128	128	128	127	129
15673	1	128	127	125	121	119	125
15676	0	121	119	121	126	128	124

4.2 Extracting module

This module consists of the following stages:

1. Parse header.
2. Read the audio stego file.
3. The extracting process. The overall scheme of extracting is shown in the Figure 4. The extracting process can be summarized as follows:

For each five successive stego samples (S1, S2, S3, S4 and S5), determine:

- $\delta = S1+S2+S4+S5$
- $\beta = \text{Round}(\delta/4)$
- Secret bit $= S3 - \beta$

4. Deciphering the secret bits

The steps of extracting process are demonstrated by algorithm (2).

Table 3, illustrates the extraction of bits (110).

Algorithm (2) Extraction process

Input **:** ***Audio file (stego_file).***
Output **:** ***Secret message (txt_file).***
Step 1 : Parse the header of the stego_ file.
Step 2 : Write the audio data into Rec_Wav, Len_stego = length of stego file.
Step 3 : Set i=0, j=0
Step 4 : While i < = Len_stego - 2 and flag <> End_of_Msg
S1 = Rec_Wav (i): S2= Rec_Wav (i+1): S3= Rec_Wav (i +2)
S4 = Rec_Wav (i+3): S5= Rec_Wav (i+4)

δ=S1+S2+S4+S5
β= Round (δ/4)
Secret bit (j) =S3- β: j = j + 1
i = i+ 3
End of While

Table 3. Extracting process.

Position	*S1*	*S2*	*S3*	*S4*	*S5*	*Secret Bit*
15670	**129**	**128**	**129**	**128**	**127**	**1**
15673	**128**	**127**	**125**	**121**	**119**	**1**
15676	**121**	**119**	**124**	**126**	**128**	**0**

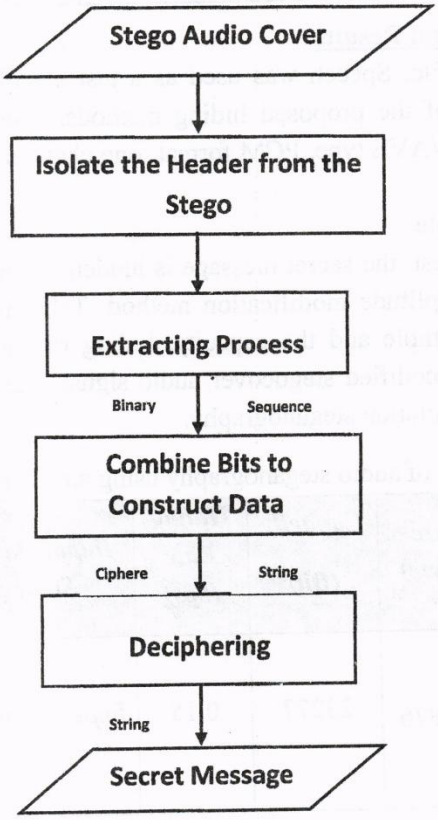


Figure 4. **The steps of the extracting module of the suggested Audio steganography system.**

5. Experimental Results

Audio file, Speech was used as a test sample to assess the performance of the proposed hiding methods. The test sample is audio file of WAVE type, PCM format, one channel (mono), and 8-bit sample size.

5.1 Test Sample

In this test, the secret message is hidden in the WAVE audio file by the amplitude modification method. Table (4) presents the size of test sample and the capacity, hiding rate and impairment grade for the modified stegocover audio signal after implementing amplitude modulation steganography.

Table 4. Result of audio steganography using amplitude modulation.

File Name	*Size (Byte)*	*Capacity (Bit)*	*Hiding Rate (bps)*	*Impairment Scale*	*Quality*
Speech_1	69,876	23277	0.15	*Imperceptible*	*Excellent*

Table (5) illustrates the objective and subjective results after hiding by using amplitude modulation method. Figure 5 and Figure 6, show the audio signal before hiding and after hiding using amplitude modulation method.

Table 5. Objective and subjective test results.

Sample	*Length of secret message*	*Hiding Rate (bps)*	*MAE*	*MSE*	*PSNR*	*SNR*	*Impairment Scale*	*Quality*
Speech_1	10632	0.152	0.39	0.83	48.95	43.09	*Imperceptible*	*Excellent*

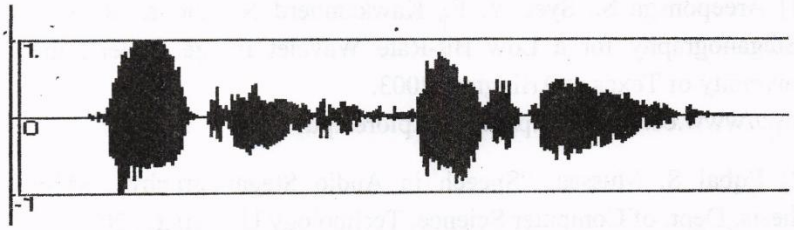


Figure 5. The test sample Speech_1 before hiding by using amplitude modulation method.

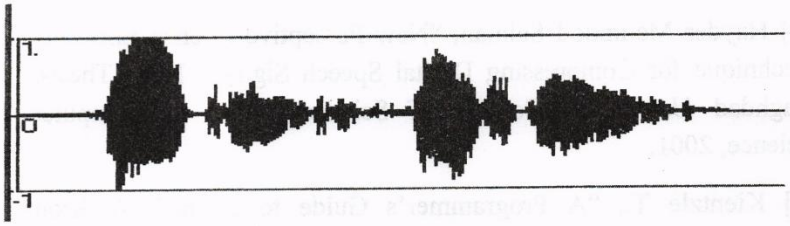


Figure 6. The test sample Speech_1 after hiding by using amplitude modulation method.

6. Conclusions

A summary of some important conclusions will be given:

1. This hiding method is successful in hiding high data rate of bits. The hiding rate is about 0.15 bps.
2. The determined SNR values are high, because only one sample out of five samples of the cover is slightly modified.
3. Speech audio files are good covers for hiding secret data by using amplitude modulation method.

References:

[1] Areeponsga S., Syed Y. F., Kawkamnerd N., and K. R. Rao, "Steganography for a Low Bit-Rate Wavelet Image Coder", the university of Texas at Arlington, 2003.
http://www.ee.uta.edu/dip/paper/explore2.pdf

[2] Eqbal S. Mussaa, "Speech in Audio Steganography", M.Sc Thesis, Dept. of Computer Science, Technology University, 2006.

[3] Haider Fethi Abd Al-Wahhab, "Text Steganography in Audio Media", M.Sc Thesis, Al-Nahrain University, College of Science, Dept. of Computer Science, 2007.

[4] Hayder Mahmood Salman, "New Perceptive Silence Encoding Technique for Compressing Digital Speech Signal", M.Sc Thesis, Baghdad University, College of Science, Dept. of Computer Science, 2001.

[5] Kientzle T., "A Programmer's Guide to Sound", Addison Wesley Longman Inc., 1998.

[6] Nidaa F. Hassan AL-Alousi, "Robust Method for Hiding Text inWave File Against MP3 Compression", Ph.D. Thesis, Dept. of Computer Science, Technology University, 2005.

[7] Raed M. Salah, "Information Hiding in Wave Media file by Using Low Bit Encoding", M.Sc Thesis, Dept. of Computer Science, Technology University, 2001.