Sound based Steganalysis for Waveform Audio File Format (WAV) and Audio File Format (AU)

Muhammad D. Hassan, Murad A. Mohammed Amin
Northren Technique University
Iraq
mdhmm75@yahoo.com
mamin@gazi.edu.tr

Suzan Mahdi
Iraq Northern Oil Company
Iraq
mdhmm75@yahoo.com

ABSTRACT: *Developing of technology today brings to ensure the security of data in the digital media with it. Coding and steganography techniques include the solution techniques which are used to ensure this security. The soundness of those techniques are also tested with the analysis techniques. In this study, steganalysis methods, which are used to reveal the hidden data within sound files, have been examined and the results, which are obtained from those methods, have been compared. The soundness of sound based steganography solutions has been tested via software which was developed in this study. Chi-square attack has been used for the testing of software and the outcomes have been assessed. Very successful outcomes have been obtained in the method which was developed in the sound based steganalysis. However, no good outcomes have been obtained for the sounds which included the prediction in the noisy sounds.*

Keywords: Chi-Square, Double Values, Least Significant Bit Insertion Method

DOI: 10.6025/jes/2018/8/3/103-111

Received: 1 May 2018, Revised 30 May 2018, Accepted 4 June 2018

1. Introduction

Steganography, an important subdiscipline of information hiding, can be described as the concealment of a data within an object [1]. The steganography term roots are derived from the Greek alphabet called “στεγαυοζ” and “γραΦειν”. It is literally meaning “hidden writing”, “covered writing” [2].

Another science which develops in parallel with the development of steganography is steganalysis. The purpose of steganalysis is to detect the existence of a picture, sound, or hidden data in any file, as opposed to steganography. Steganalysis is the whole set of techniques designed to distinguish between the carrier, the original objects and the stego objects.

In the base of the steganalysis studies, it is the idea that the concealed data leaves some fingerprints on the carrier object. That is, the stego object that occurs after concealment carry statistically significant differences, although it cannot be distinguished visually, audibly or functionally from the original [3].

In the field of steganalysis, many studies have been made especially on the steganalysis of picture and sound files.

Type Style and Fonts.

2. Data Hiding Methods in Sound Files

Methods for hiding data in sound files are described in the following subheadings [5].

2.1 Low-Bit Encoding

It is performed in the same way as the LSB insertion method used in image steganography. One bit of information to be hidden in the last bit of each byte of data in the sound file is written. The resulting change causes noise in the sound file. It also has a weak structure. The message can be damaged or destroyed by resampling or noise that may occur in the channel.

2.2 Phase Encoding

The phase encoding method is similar to the JPEG algorithm applied in picture files. In embedding, the sound file is divided into small segments and the phase of each segment is changed with the phase reference of the data to be hidden. The phase encoding procedure is as follows [6]:

- The sound data is divided into *N* short segments.
- Each segment is subjected to Discrete Fourier Transform (DFT) to create phase and magnitude matrixes.
- The phase differences between neighboring segments are calculated.
- A new phase value for each segment is generated by concealing information.
- New phase matrixes and magnitude matrixes are combined to obtain new segments.
- The new segments are combined to produce a coded output.

2.3 Spread Spectrum

Concealment is performed on the frequency spectrum used by the sound signal. Along with not having a strong structure, it causes noise in the sound.

2.4 Echo Data Hiding

The concealment of information is achieved by adding an echo to the carrier sound signal. The information is hidden using values such as the amount of delay, the fading rate, or the magnitude of the information. It is possible to code 0 or 1 at a level that the human ear cannot detect using two different delay values. For every bit encoding, the signal is divided into segments. The echo data hiding method does not cause any noise or use a lossy encoding.

3. Attack Models

The most common of these attack models are following [7,8,9]:

1. **Stego Attack only:** Only stego-object (image file) is known for analysis.

2. **Known Cover Attack:** The state of the image is known before and after the message is hidden.

3. **Known Message Attack:** Hidden message is known.

4. Selected Stego Attack: Steganographic algorithm and stego-object are known.

5. Selected Message Attack: In this method Steganalist selects various messages to analyze the stego-object, uses steganographic tools and tries to find the algorithm.

6. Known Stego Attack: The cover object, stego object, and steganographic tools are known.

4. Sound Based Steganalysis Applications

In steganographic tools, steganalitic approaches for the identification of concealed messages in Wavelet (WAV) and Audio Unit (AU) files were presented using the FE tool. Experimental results show that messages embedded as small as 10% of the steganographic capacity can be reliably determined.

Steganography is becoming more popular with the rapid growth of digital content and the widespread Internet communication system. High definition digital image and the sound closest to natural sound can be obtained in a convenient way using digital camera and digital sound recorder. Many steganographic tools that can store in image, sound and video files are now available on the Internet. People can store their secret data in these digital content using such tools and send it or keep it on their discs to provide privacy. Only target receivers or themselves can receive this hidden data. Third parties will not know hidden data or hidden communications. So in some cases this is much more secure than cryptography.

Although not as much as the picture steganalysis, some studies have been conducted on sound steganalysis [10,11,12].

Ru et al. (2006) have developed a method to detect embedding in WAV files with some existing steganography software [10].

Özer et al. (2006) have achieved 75% to 90% success in their study of identifying the existence of a hidden message in sound files universally, without needing to know the concealment algorithm [11].

Özer et al. have established a method based on the measurement of sound quality [13]. The basic idea is that the distribution of various statistical distance measures calculated on the cover sound signals and the stego-sound signals, and that the one to one sound versions are statistically different. They chose an appropriate sound quality meter as a steganalizer, and made a classifier that comes in two classes.

Micah Johnson et al. [14] recently presented a software for statistical regulation of sound signals and used a non-linear support vector for classification. However, this approach is unlikely to be effective in detecting low-bit-rate embedding. Steghide [15], a tool that can be found free on the internet, can embed messages in JPEG, BMP, WAV, or AU files,. It is primarily designed to counteract the statistical attack.

4.1 Sound Steganalysis Approaches

We can list the approaches to sound steganalysis as follows [16]:

- Auditory detection
- Algorithm specific detections and
- Universal detection.

The auditory detection is similar to that of sound, in that the sound quality of the stego sound file is distinguished from the original sound quality, or the detection of defects that should not be present in a normal sound without the original sound file. However, it will not be easy to distinguish stego sounds from self-loud noises.

Algorithm-specific detections are methods that attempt to estimate the existence of concealment operations performed with specific algorithms.

Universal detection includes methods for detecting differences according to various properties between original sound files and stego sound files, and for separating the stego audio file with the carrier file.

4.2 Techniques used in Steganalysis

Some of the steganalysis approaches are [17]:

1. Visual and Auditory attacks (Visual Detection)
2. Universal Blind Steganalysis(Detection of steganographic Artifacts)
3. Histogram analysis (Steganalysis Based an Image Quality Metrics)
4. First-order statistical Analysis
5. Steganalysis Based on JPEG Compatibility
6. High level statistical analyzes (RS Analysis)
7. Paired Comparison Analysis (Pairs Analysis)
8. Palette Quick Pairs Analysis
9. Raw Quick Pairs Analysis
10. Chi square Attack
11. Other Methods

4.3 Analysis Process and Details

In making these algorithms, a few rules have been taken as follows:

1. The embedding algorithm will have no more than one bit per pixel. Some pixels may not contain embedded bits. This rule has been taken both in terms of simplicity and in order to embed the data into the sound in an inconceivable way.

2. Starting from first line the cover sound, the embedding algorithm bits will embed each message bit to the pixel at the end and to the cover sound. The purpose of this embedding plan is to simplify the large-scale process for debugging, but at the same time, it allows usto test various sections of a sound, i.e. data embedded sections and data not embedded sections.

3. Each message will be generated randomly before embedding. The primary purpose of this design is to minimize the possibility of attributing the results to the specific content of the embedded message. The results are correct if the message changes when the embedding process occurs, and the results match when each new message is embedded,

In constructing these algorithms, several assumptions are made as follows:

1. The number of bits in the message is equal to or less than the number of pixels in the cover sound. This is a necessary condition to satisfy the decision that each pixel has at most one embedded bit.

2. The message is a random distribution of 0s and 1s. Generally for embedding in steganography, the embedded data is encrypted before being embedded, thus forming a random 0's and 1's distribution. So we assume that the embedding data is encrypted.

Pairs of Values 3 Sound (PV3Sound) Algorithm used in Attack.

The PV3Sound detection algorithm is designed to perform a chi-square attack on a sound and output the probability of embedding data.

The details of the chi-square statistical test are given step by step below:

Step 1: Assume that there is a random sampling of k categories and observations. Each observation falls into only one category. The single values of the PSs of suspicious information is considered important.

Step 2: After concealing a uniformly spread out message, the frequency theoretically expected in category *i* is as this:

$$n_i^* = \frac{|\{pixel|(pixel)'\ sorderedindex \in \{2i, 2i+1\}\}|}{2}$$

Step 3: In a random sample, the measured frequency of occurrence is as follows.

$$n_i = |\{pixel \mid (pixel)'\ sordered\ index = 2i\}|$$

Step 4: The chi-square statistic is calculated with the *k*-1 degrees of freedom as follows:

$$X^2_{k\text{-}1} = \sum_{i=1}^{k} \frac{(n_i - n_i^*)^2}{n_i^*}$$

Step 5: If nive n_i^* distributions are equal, *p* is the probability of message embedding. This probability is calculated by taking the integral of the density function (Γ is Euler's gamma function):

$$P = 1 - \frac{1}{2^{\frac{k\text{-}1}{2}}\,\Gamma\left[\frac{k\text{-}1}{2}\right]} \int_0^{X^2_{k\text{-}1}} e^{\frac{-x^2}{2}}\, x^{\frac{k\text{-}1}{2}\text{-}1} dx$$

Chi-square statistical analysis yielded successful results in sequential LSB embedding steganography.

4.4 Experimental Results

I selected five different Wavelet (WAV) files and five Audio Units (AUs) for testing, with varying size and uniformity. The original sounds are labeled as "Wav001", "Wav002" .... "Wav005" and "Au001", "Au002" .... "Au005". I embedded information at the ratio of 10%, 50%, and 100% of the pixels to each one, and saved each of these stego sounds, such as the embedding of information to 10% of pixels in Wav002, the embedding of information to 10% of pixels in "FlipWav00210" and Au002, and in the format of "Flip[sound name][information embedding percentage]" so as to create a stego sound called "FlipAu00210". After creating 15 stego WAV and 15 AU, each of these stego sounds and original sound (total 20 WAV and 20 AU sound) are tested using PV3Sound.

Although the sampling of the sounds I selected is small (only 5WAV 8bit sound file and 5AU 8bit sound file), I got some interesting results. First of all, PV3Sound seems to be analyzing more than just half of the sounds I test, especially large homogeneous regions.

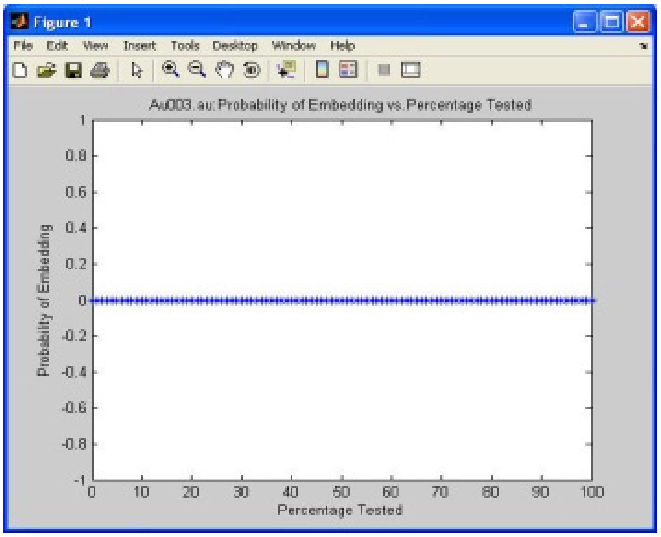

Figure 1. (Au003) Original Sound

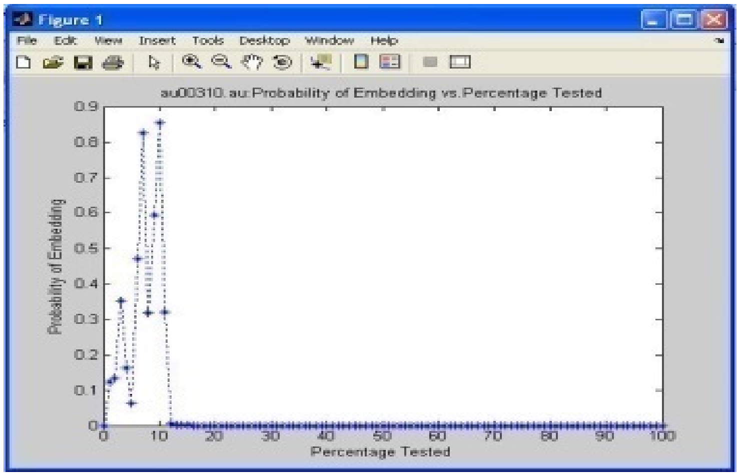

Figure 2. (Au00310) Stego Voice

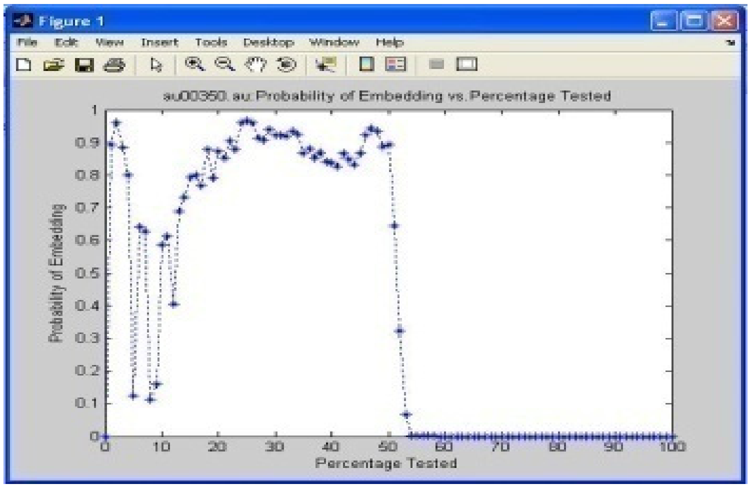

Figure 3. (Au00350) Stego Voice

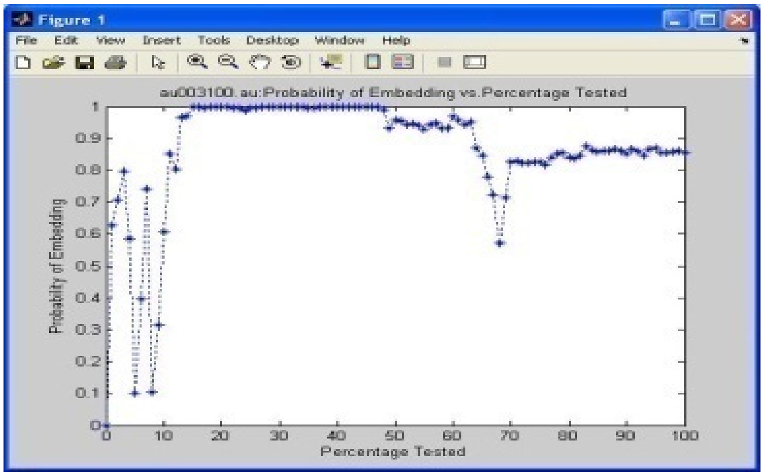

Figure 4. (Au003100) Stego Voice

4.5 Reliability of Chi- square Attack

One of the questions we started to answer is "Can we say that whether an image contains embedded information" and the simple answer is "Sometimes". When 100% of the pixels of each tested sound file are data embedded, the chi-square test returns the probability of embedding very close to one. In other words, when testing any sound file with data embedded in 100% of its pixels (in this study), the chi-square test returns a true positive result for each sound file. In Table 4.1 and Table 4.2, the falls according to the Chi-square attack are given for the data embedded sound files of 100%, 50%, 10%, 0% of their pixels.

Percent of embedding data in audio files	Au1	Au2	Au3	Au4	Au5
% 0	2/0	0	0	0	0
% 10	13	12	11	11	19
% 50	14	65	65	53	90
% 100	14	2/100	100	100	100

Table 1. "Falls" according to the Chi-square attack (Audio Unit)

Percent of embedding data in audio files	Wav1	Wav2	Wav3	Wav4	Wav5
% 0	0	0	1-2/0	2-4/0	2/0
% 10	7/18	10	11	11	2/12
% 50	7/69	89	52	55	2/53
% 100	7/100	100	100	100	2/100

Table 2. "Falls" according to the Chi-square attack"

When there is sufficient data found in a sound file, the chi-square test can detect their existence, but the test may return false positive values for noisy Sound files and regions that are noisy in sound files. The next logical question is, "Can we tell you how much data is embedded in any Sound file?" In order to answer this, we can say with a slight degree of precision where the embedding process stops, paying attention to the fall of r pr. When the pixel after the embedded data is tested, the pr falls drastically. Nevertheless, I have not detected that this is true for most sound files I have tested. Thus, for both 5WAV and 5AU sound files, Tables 1 and 2 show where the falls have taken place, showing both where the pr falls below 0.90 and where it falls below pr.

As noted earlier, the chi-square test does not always accurately show how much data is embedded in the sound files. Although 50% of both Wav00150 and Wav00250 pixels contain data, 69% and 89% of their pixels does not contain data. Similarly, 14% and 90% of the Au00150 and Au00550 pixels are fell until tested. The chi-square test correctly shows how much data is embedded when 50% of the pixels contains data, in ten sound files (Wav, Au), the first and second in Wav and the first in Au, in only three of them in fifth.

Error ratio = % The percentage of existence - % Data embedding ratio

5. Conclusion and Discussion

In recent years, the security of computer systems has become a very important issue. With the spread of the Internet, data exchange and sharing have increased. Text, sound, picture etc. are effectively shared by people in many parts of the world. In this way, it is very easy to hide information that is wanted to be sent into digital media and transfer it to other people. However, the use of this method by malicious people puts the security of society and the environment in danger. For this reason, the examination of whether or not there is confidential information in the data in the digital environment has become a very important subject. Various steganalysis methods have been developed to detect this.

Percent of data embedding	Au1	Au2	Au3	Au4	Au5
% 0	% 2	% 0	% 0	% 0	% 0
% 10	% 3	% 2	% 1	% 1	% 9
% 50	% 36	% 15	% 15	% 7	% 40
% 100	% 86	% 2	% 0	% 0	% 0

Table 3. Error ratio of Wav Sound file

Percent of data embedding	Wav1	Wav2	Wav3	Wav4	Wav5
% 0	% 0	% 0	% 1	% 2	% 2
% 10	% 15	% 0	% 1	% 1	% 4
% 50	% 25	% 39	% 2	% 5	% 5
% 100	% 7	% 0	% 0	% 0	% 2

Table 4. Error ratio of Au Sound file

A sound based software has been developed for the realization of this study and the tests have been carried out successfully and easily.

In the results presented in this study, the following assessments can be made to ensure that steganography studies are more successful:

1. The sound file to be used as the carrier should not be a readily available file of known originals. This may be necessary for attacks of auditory steganalysis.

2. The concealment rate should be kept low enough to increase the soundness to statistical steganalysis although it will reduce the concealment capacity.

3. Steganography methods can be combined with encryption methods to create a more secure system.

4. If it is asked to increase the amount of data to be concealed, this can be achieved by compressing the data to be hidden before the hiding process.

5. On the subject of steganalysis, the chi-square attack, described in the scope of this study and in the framework of experiments presented in this study, is highly effective against data embedding algorithms changing the least significant bit (LSB) when large amounts of data are embedded.

6. As in all over the world, in Turkey as well, steganalysis subject does not draw necessary attention entirely. In general, steganalysis, in particular, sound steganalysis applications, has not yet achieved sufficient progress. As the data transfer techniques become faster, the sound files, which are larger in size than the picture files, are thought to be used in digital communication very intensely. In this study, one of the most common sound file types, PCM-formatted WAV and Audio Unit (AU) files are used, and it is assessed that the realized software will create a basis for the sound based steganalysis applications to be realized in Turkey, and it will contribute. It is predictedthat the picture-based and sound-based software realized in this thesis study may be aimed at meeting the need to provide personal and institutional security in a different and high level in the light of the study results.

Many difficulties have been encountered during this study. First, the lack of a thesis on steganalysis in Turkey, it is observed that creates a misconception among people that there is no need for such a study.People who have hearsay knowledge on this subject, characterize the study an easy field to choose as a thesis subject initially caused a negative motivation. However, when the subject is assessed with all its scope, fitted in the frame that needs to be, and the outlines of the study are revealed, it is realized that the subject is a multidimensional and dynamic subject which is quite comprehensive, important concepts and new technologies are closely related.

It is expected that this study will be helpful to those who will study on this field from now on.

In addition to the English names of the software described in this study, the original Turkish expressions that can be used corresponding to these are identified and used.

Good result is obtained when using sound-based steganalysis in the obtained results. But in noisy sounds, the Chi-square attack is insufficient.

References

[1] Petitcolas, F. A. P., Anderson, R. J., Kuhn, M. G. (1999). Information Hiding–A Survey, Proceedings of the IEEE, *Special Issue on Protection of Multimedia Content*, 87(7)1062-1078.

[2] Murray, A. H., Burchfiled, R. W (eds.). (1933). The Oxford English Dictionary: Being a Corrected Re-issue", Oxford, England: Clarendon Press,

[3] Phan, R. C. W., Ling, H. C. (2003). Steganalysis of Random LSB Insertion Using Discrete Logarithms Proposed At Cita03", *In:* MMU International Symposium on Information and Communication Tecnologies / M2USIC 2003, Petaling Jaya, Malaysia, 2-3.

[4] Provos, N. (2001). Defending Against Statistical Steganalysis, 10th USENIX Securit Symposium, Washington, 323-335.

[5] Internet: Provos, N. (2004). Steganography detection with stegdetect, http://www.outguess. org /detection .php.

[6] Bender, W., Gruhl, D., Morimoto, N., Lu, A. (1996). Techniques for data hiding, *IBM Syst. J.* 35 (3&4), 313-336.

[7] Mallat, S.G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, *IEEE.Transactions on Pattern Analysis and Machine Intelligence*, 674-693.

[8] Cvejic, N. (2004). Algorithms for Audio Watermarking and Steganography, in Department of Electrical and Information Engineering, Information Processing Laboratory, University of Oulu, Finland. http://www.mp3tech.org/programmer/docs/isbn9514273842 .pdf

[9] Pal, S. K., Saxena, P. K., Muttoo, S. K. (2002). The Future of Audio Steganography, *In*: Paper presented at Pacific Rim Workshop on Digital Steganography (STEG'02) 12-14.

[10] Internet: Hetzl, S., Steghide, http://steghide.sourceforge.net/ (2003).

[11] Atýcý, M.A. (2007). Steganografik Yaklaþýmlarýn Ýncelenmesi, Tasarýmý Ve Geliþtirilmesi, Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enistitüsü. Ankara, 23-24, 44-47.

[12] Internet: WetStone Technologies Inc., http://www.wetstonetech.com/faq_stego.html (1997).

[13] Internet: OutGuess - Steganography Detection, http://www.outguess.org/detection (2007).

[14] Fridrich, J., Du, R., Meng, L. (2000). Steganalysis of LSB Encoding in Color Images, Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conferece, New York City, USA, 3. 1279-1282.

[15] Özer, H., Sankur, B., Memon, N., Avcýbaþ, Ý. (2006). Detection Of Audio Covert Channels Using Statistical Footprints Of Hidden Messages, Digital Signal Processing, 16 (4) 389-401.

[16] Westfeld, A., Pfitzmann, A. (2000). Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S-Tools–and Some Lessons Learned, 3rd International Workshop on Information Hiding. http://www.ece.cmu.edu/~adrian/487-s06/westfeld - pfitzmann-ihw99.pdf.

[17] Fridrich, J., Goljan, M. (2002). Practical steganalysis of digilal images state of the art, *In:* Proc. SPIE Photonics West, 4675: 1-13.