Breaking the Steganographic Utility MP3Stego

Yawai Tint,
University of Computer Studies, Yangon
yawai.ywt@gmail.com

Khin Than Mya
University of Computer Studies, Yangon
khinthanmya@gmail.com

Abstract

Steganography is the art and science of hiding information by embedding messages within other, seemingly harmless messages. In this paper, techniques are presented which aim at the breaking of steganography usage in digital audio data. The propose system enhances the audio steganalysis by breaking the steganographic utility. In this system, white noise is added to the audio signal and lowest bit of part2_3_length is changed by other bit. The propose scheme analyzes this algorithm by using MP3Stego. The performance of this system is calculated by Perceptual Evaluation of Speech Quality (PESQ) and Bit Error Rate (BER).

***Keywords**: MP3Stego, Perceptual Evaluation of Speech Quality (PESQ), Bit Error Rate (BER)*

1. Introduction

Steganography is the science which deals with hidden information exchange. While cryptography protects the content of secret data, the goal of steganography is to conceal the mere existence of secret information. In technical steganography, digital data like images, audio and video files are used as cover media. In this paper, the focus has been set to audio steganography. The characteristics of the human auditory system (HAS), the application of steganography in this area is a challenging task. This is due to the fact that the HAS notices slight changes in the audio data more quickly than e.g. the human visual system (HVS) reacts on small distortions in pictures [13, 14]. Due to the sensitivity of the HAS, there are three important requirements for steganography when applied to audio data: undetectability and inaudibility, robustness and capacity.

Steganalysis is the detection of embedded data present in a given cover medium. This process can be done through statistical analysis as well as more sophisticated techniques, dealing with specific parameters of the audio signal. In contrast, the presented attack techniques shall be applicable on all kinds of audio material, having the goal of interfering the steganographic receiver if there exists a steganographic communication. On the other hand, if no steganographic communication is present, the techniques can be applied as well. There have been two main research approaches to the problem of steganalysis, namely, technique-specific steganalysis and universal steganalysis. The former group of techniques performs very accurately when used against the steganographic technique it is targeted for. The latter group of technique, on the other hand, are effective over a wide range of techniques, while performing less accurately overall. However, since universal steganalysis is better suited to the practical setting, it attracted more interest and many effective steganalyzers are proposed.

The major contribution of this paper is a novel combination of various basic signal processing operations for the generic prevention of steganographic communications in audio cover media. The system in [15] has done detection the existence of hidden messages and separation signal from mixture source. The rest of this paper is organized as follows; in section 2 related approaches, dealing with the prevention of steganographic communication. Section 3 discuss the MP3Stego technique and Section 4describe the propose system. Section 5 present the implementation of the propose system and Section 6 shows the outcome of evaluating the quality of audio signals. The conclusions are given in Section 5.

2. Related Work

The first version of the StirMark benchmark [1] was built for image watermarking and steganography algorithms, hence [2] mainly focuses on digital images. However some notes on the introduction of jitter into audio data are given, a method which took up and implemented it as “variable time delay”. After StirMark was released the initial researchers as well as other scientists improved the first version and in 2001 a StirMark benchmark for audio cover media was described in [7]. This work has been enhanced by the addition of lossy compression techniques like MP3 and Ogg.

In 2000 a competition has been started on digital watermarking robustness, announced by the Secure Digital Music Initiative (SDMI). The goal was to break various audio watermarking techniques, but no restriction was given about the resulting quality. In [10] attack approaches on SDMI watermarks have been published. While some of these attacks require the study of the embedding algorithm (non-blind attacks), they also mention pitch shifting and “time axis warping”. But although it is said that their approaches successfully attacked the SDMI watermarks, no actual evaluation results of the strategies are given.

Therefore, performed attacks may as well degrade the audio quality by an unacceptable amount while our techniques try to keep up the original quality as much as possible. Related to this approach is the work from [11] which deals with attacks on digital audio watermarking systems. In this work watermarking attacks are grouped into four categories: Removal, desynchronization, embedding and detection. For each group, several attacks are described but no evaluation is done to show the actual efficiency of the attack categories as well as their processing time. The propose system enhances this work by giving an overview of the performance as well as the audibility of the prevention techniques.

In [15], that proposed steganalysis of audio signal by using Independent Component Analysis. A detection method is used for detecting hidden message in compressed audio files produced by MP3Stego. Steganography can be successfully detected during the Principle Component Analysis (PCA) whitening stage. A nonlinear robust batch ICA algorithm, which is able to efficiently extract various temporally correlated sources from their observed linear mixture are used for blind steganography extraction.

In this paper enhances this approach by breaking the steganography usage in digital audio data and minimizing the audio quality degradation, making them usable for different fields of application. And then this method show to recover embedded data from media which has been damaged by some kind of attack.

3. MP3Stego

MP3 was the research result of the Fraunhofer-IIS Institute. MP3Stego embeds compressed and encrypted data in an MP3 bit stream during the compression process. The MP3 audio encoding process is shown in Figure 1.

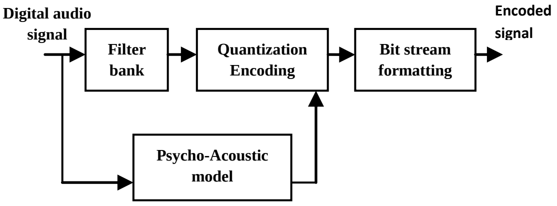


Figure 1. The MP3 Encoder

The important parts that carry out the MP3 encoding are the two loops [14]. The inner loop is a quantization and encoding loop. The outer loop is a noise control loop. It controls the quantization noise according to the threshold value and adjusts the scale factor. MP3Stego was carried out on the base of the 8 Hz-mp3 encoder. The tool uses the power of parity principle to embed data in part2_3_length of a granule in a MP3 file. The Part2_3_length variable indicates the total number of bits required to encode the scalefactors and the Huffman coded data. The hiding process of MP3Stego could be explained in Figure 2.

```
Static int inner_loop (...)
{    ……
do {
quantizerStepSize += 1.0;
bits=quantize ( );
switch (hiddenBit)  {
case 2:
embedRule = 0; break;
case 0 : case 1:
embedRule = ((bits + part2length)%
2)!=hiddenBit;
break ;
default ;
ERROR (“inner loop: unexpected hidden
bit”);}
} while ((bits>max_bits)|embedRule);
Return bits;
}
```

Figure 2. The simplified inner iteration loop of MP3Stego

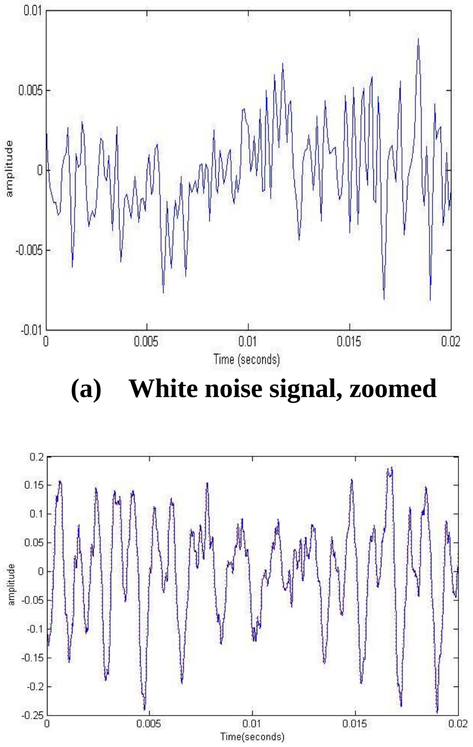
4. Proposed system for “Natural” Modifications

Techniques from this category model existing behaviors which have real-world causes. As a consequence these techniques are not suspicious, even if they are detectable e.g. by analyzing the waveform of a signal.

4.1. White noise

An obvious technique for the disruption of steganography in any media is the addition of white noise, where individual samples are changed by a random value hence modeling electrical noise.

Noise addition or interference happens to transmitted signals on various communication channels, e.g. due to thermal noise or crosstalk. As a consequence audio data is affected as well when transmitted over such channels. Due to an increased noise level and changed sample values the steganographic receiver may not be able to extract all the embedded information correctly, depending on the actual embedding algorithm used. This operation directly influences the quality of the audio signal, which degrades with increasing noise amplitude. In general, a signal-to-noise ratio (SNR) above 20dB guarantees for a reasonable audio quality. As a consequence this approach uses very little noise amplitudes, therefore keeping the SNR above 20dB and minimizing the influence on the audio quality while still interfering with the steganographic receiver. Figure 3 (a) shows an example of a noise signal which has been added to a block of the audio signal. The result shows Figure 3 (b), where the original signal is shown as solid curve and the audio signal including the noise is shown as dotted curve. There are not different clearly because the noise included signal is nearly the same with original signal.



(a) White noise signal, zoomed

(b) Original and modified signals

Figure 3. White noise addition

4.2. Bit Changing

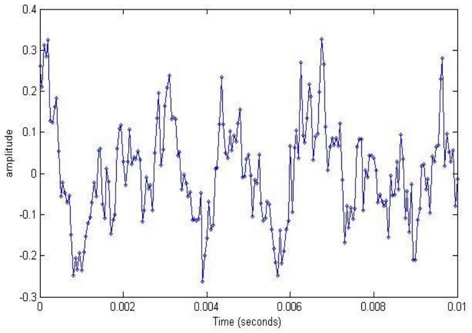
MP3Stego is attacked by changing the lowest bit at part2_3_length because the secret information is hidden in the part2_3_length offset. MP3Stego can be detected by analyzing the statistics of part2_3_length. The inner iteration loop of the MP3 encoding process can be ended when the part2_3_length is less than the specified max_length, but the loop will continue to the hidden bit. The final part2_3_length becomes smaller and the next frame's part2_3_length becomes larger. Directly proportional the block length variance becomes larger. For MP3Stego algorithms such changing the lowest bit points may be a problem for the extraction process. This is expecially true if no synchronization is done at the receiver. Therefore this technique can disrupt the steganographic communication while having no influence on the quality of the underlying audio signal.

Below is a snip of MP3Stego calculating part2_3_length: Part2_3_length is represented by 12 bit for mono mode and 24 bit for other mode. Part2_3_length exist in the side information of audio frame. MP3's frame consists of five parts; header, CRC, side information, main data and ancillary data. The side information part of the frame consists of information needed to decode the main data. The size depends on the encoded channel mode. If it is a single channel bit-stream the size will be 17 bytes, if not, 32 bytes are allocated. The different parts of side information are main_data_begin, private_bits, scfsi, Side_info gr.0 and Side_info gr.1. Part2_3_length can be extracted from the Side_info gr.0. Figure 4 express the calculating part2_3_length from this side information.

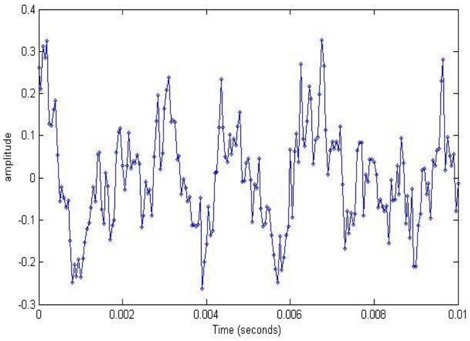
```
    void ResvFrameEnd(L3_side_info_t
*l3_side, int
    mean_bits )
    {  …
      if(stuffingBits) {  …
    #ifdef MP3STEGO //to satisfy the request of
odevity
          if (stuffingBits % 2) {
            gi->part2_3_length += stuffingBits -
1;
            stuffingBits = 1;
            l3_side->resvDrain=stuffingBits;
          }
          else gi->part2_3_length +=
stuffingBits;
    #else
        gi->part2_3_length += stuffingBits;
    //in normal case this request is not needed
    #endif
      }  …
    }
```

Figure 4. Calculating part2_3_length

MP3Stego's variance distribution is dependent on the encoder being used, a difference in encoders can make it appear that the MP3 either does or does not have any hidden data. The wide variety of open source encoders makes detection of block length variances marking steganography very difficult. Figure 5 (a) shows an original audio signal; Whereas Figure 5 (b) shows the same signal with bit changing the lowest bit at part2_3_length.



(a) Original signal



(b) Modified signal

Figure 5. Bit Changing of part2_3_length

5. Implementation

The propose technique is implemented and tested on a set of 400 MP3 audio files. The audio samples include music types (piano, symphony, violin and rock), songs (pop ,blue, classical, country and folk) and nature noise etc. Each audio has duration of 20 seconds. This system produced the same amount of stego audio by hiding random messages in these audios. Then these MP3 files have been detected and separated by PCA and ICA [15]. These audio files are used for steganography prevention. White noise is generated with a random number generator by matlab function. Bit changing is done by replacing the lowest bit of part2_3_length with another bit.

6. Evaluation

This section shows the performance of attack techniques. First the quality of audio files before and after modifications have been applied is tested in order to make a statement about the resulting quality difference due to this techniques. Afterwards the impact of the individual attack methods on the steganographic receiver is demonstrated, showing bit error rates (BER) for MP3Stego. Further the relationship between attack technique and embedding algorithm, based on the test results, is investigated and the possibility of error-correction codes (ECC) as a countermeasure against attack techniques is considered.

6.1. Audio Quality Evaluation

In Perceptual Evaluation of Speech Quality (PESQ) the original and degraded signals are mapped onto an internal representation using a perceptual model. The difference in this representation is used by a cognitive model to predict the perceived speech quality of the degraded signal. This perceived listening quality is expressed in terms of Mean Opinion Score. Table 1 describes the ACR (Absolute Category Rating) listening quality opinion scale used in the development of PESQ. R-value is below the 60 that is unacceptable.

Table 1. Opinion scale used in the development of PESQ.

Quality of the speech	Score	R-value Range
Excellent	5	100-90
Good	4	90-80
Fair	3	80-70
Poor	2	70-60
Bad	1	60-0

R-value is calculated from the below equation.

$$R = 94.2 - I_d - I_e \quad (1)$$

- •I_d is impairment from mounth-to-ear delay
  - – Encoding (packetization)
  - – Network (transmission, propagation and queuing)
  - – Playout (buffering)
- •I_e is impairment from distortion (loss)

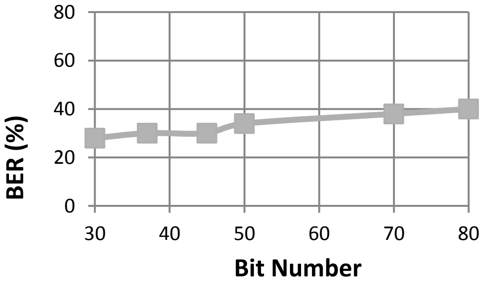
For this test audio signals have been composed, each contain 20s of MP3 file. On this signals all attack techniques were performed in a row and tested via the PESQ reference implementation. Table 2 shows the results and the actual configuration per attack technique, comparing the average MOS grade of the original signal to those of the signal after manipulations.

Table 2. MOS grades for original and modified audio signals

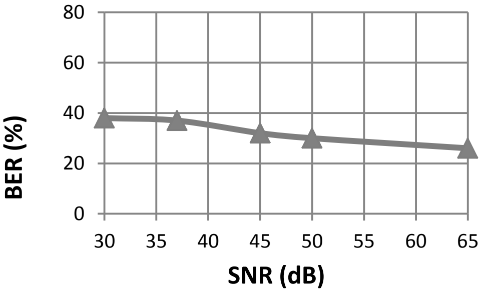
Manipulation	MOS grade
Original signal	4.500
Noise addition	3.112
Bit changing	3.319

6.2. Prevention Efficiency

To grade the efficiency of the techniques the BER from the steganographic extraction process have been used. That is, the higher the BER the more effective the attack technique has been on the particular steganographic algorithm. For all tests 80 bits have been embedded into MP3 files. For these tests all attack techniques have been integrated into a system which originally implemented for steganographic communications over audio channels. The following figures show the BER percentage of white noise and bit changing of part2_3_length.



(a) Bit Changing



(b) White Noise

Figure 6. BER for MP3Stego

7. Conclusion

In this paper describes the prevention of steganographic algorithm and shows to recover embedded data from media which has been damaged by some kind of attack. The attacks on MP3Stego with audio data as cover medium are given. The proposed techniques can be divided into two categories. The first category detects the embedded data in audio file. The second category performs noise addition and changing the lowest bit value at part2_3_length. Such modifications can happen at any time. Methods from second category perform “malicious” modifications for steganography algorithms.

Reference

[1] F. A. P. Petitcolas, StirMark Benchmark 4.0. http://www.petitcolas.net/fabien/watermarking/stirmark/.

[2] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, Attacks on copyright marking systems, Proc.

[3] FFTW, Fastest fourier transform in the west, http://www.fftw.org.

[4] G. Fisk, M. Fisk, C. Papadopoulos and J. Neil, Eliminating steganography in internet traffic with

[5] H. Wang and S. Wang, Cyber warfare: steganography vs. steganalysis, Communications of the ACM,vol. 47, no. 10, pp. 76- 82, 2004.

[6] J. Dittmann, M. Steinebach, A. Lang and S. Zmudizinski, Advanced audio watermarking bench-marking, Proc. of SPIE, pp. 224-235, 2004.of the 2nd International Workshop Information Hiding, pp. 219-239, 1998.

[7] J. Dittmann, A. Lang and M. Steinebach, Stirmark benchmark: audio watermarking attacks based on lossy compression, Proc. of SPIE, pp. 79-90, 2002.

[8] L. Shelley and J. Picarelli, Methods not motives: Implications of the convergence of international organized crime and terrorism, International Journal of Police Practice and Research, vol. 3, pp.305-318, 2002.

[9] M. Arnold, Attacks on digital audio watermarks and countermeasures, Proc. of International Conference on Web Deliv. Music, pp. 55-62, 2003. International Conference on Acoustics, Speech, Signal Processing, vol. 3, pp. 1369-1372, 2001.

[10] M. Wu, S. Craver, E. Felten, and B. Liu, Analysis of attacks on sdmi audio watermarks, Proc. Of International Conference on Web Deliv. Music, pp. 55-62, 2003.

[11]Lee, T.W Independent Component Analysis: Theory and Applications. Kluwer Academic Publishers (1998)

[12]P. N. Basu and T. Bhowmik, On embedding of text in audio-a case of steganography, Proc. Of International Conference on Recent Trends in Information, Telecommunication and Computing, pp.203-206, 2010.

[13]W. Bender, D. Gruhl, N. Morimoto and A. Lu, Techniques for data hiding, IBM Systems Journal,vol. 35, no. 3, pp. 313-336, 1996.

[14] Y.Huang, H.Song “Detecting MP3Stego and Estimating the Hidden Size” In Proceedings of the 20th International Joint Conference in Artificial Intelligence (IJCAI). 2808–2813.

[15] Y.Tint,“Audio Steganalysis Based on Independent Component Analysis” In proceeding of the 10th International Conference on Computer Applications .