Stenography: the science behind the hidden message

Manuela Tolosa Garcia
Ingenieria de Sistemas
Universidad EAFIT
Medellín, Colombia
mtolosag@eafit.edu.co

Jorge Alfredo Villarreal Marquez
Ingenieria de Sistemas
Universidad EAFIT
Medellín, Colombia
javillarrm@eafit.edu.co

Jaime Rodrigo Uribe
Ingenieria de Sistemas
Universidad EAFIT
Medellín, Colombia
jruribem@eafit.edu.co

Julian Rojas Gallego
Ingenieria de Sistemas
Universidad EAFIT
Medellín, Colombia
jrojasg1@eafit.edu.co

May 31, 2023

Abstract

Audio steganography is the process of hiding secret information within an audio signal without detection. It has been widely used for covert communication and information security, but attempting audio steganography can be challenging due to various factors such as the limited capacity of audio signals, the human auditory system, and detection algorithms.

In this paper, we explore the challenges and limitations of attempting audio steganography. We begin by discussing the properties of audio signals and the steganographic capacity of different audio carriers. We then describe the state-of-the-art techniques used for audio steganography, including frequency domain and time domain methods. Finally, we propose an LSB-like technique to hide information on different bits whose order is repeated and pre-configured by the user.

1 Objective

General objective:

• Hide messages in audio files through algorithms that transmit desired information.

Specific objectives:

• Associate and implement numerical methods and steganography techniques to hide information in audio files.

• Transmitting information in a hidden way, trying to maintain the integrity of the message.

2 Introduction

In today's digital world, information security and privacy protection have become critical concerns. In this context, steganography, the art of hiding information within the media, has acquired increasing relevance. A specific branch of this discipline, audio steganography, focuses on hiding sensitive data within sound files without arousing suspicion.

Audio steganography techniques take advantage of sound's intrinsic characteristics and psychoacoustic properties to imperceptibly hide information within audio files. This offers a discreet way to transmit sensitive data, since the modified audio files do not necessarily show obvious signs of alteration.

This research article aims to explore in detail the techniques used in audio steganography, as well as its applicability and effectiveness in the protection of confidential information. Through the analysis of numerical and psychoacoustic methods, it will be examined how these concepts can be associated and implemented to hide data in audio files without significantly affecting their perceptual quality.

In addition, the different audio steganography techniques used today will be investigated in depth, such as LSB modification, frequency domain keying, wavelet transform keying, phase coding and spread spectrum. The theoretical foundations and algorithms associated with each of these techniques will be analyzed, evaluating their concealment capacity, resistance to detection and their impact on the quality of the resulting audio.

3 Advantages and Disadvantages of Audio Steganography

Audio steganography is the technique of hiding information in audio files without it being detected by people outside the communication. But like everything, it has its advantages and disadvantages, which are:

Advantages:

1. Effective Concealment: Audio steganography allows messages to be effectively hidden within audio files without making them easily detectable by unauthorized persons. This provides an additional level of security for the transmission of sensitive information.

2. Stash Capability: Audio files have a significant capacity to store additional data without compromising perceptible sound quality. This means that relatively long messages can be hidden in audio files without raising suspicion.

3. Message integrity: Audio steganography focuses on maintaining the integrity of the hidden message. Unlike cryptography, where the primary goal is confidentiality, steganography focuses on ensuring that the hidden message is not altered or corrupted during transmission.

4. Diversity of Methods: There are several techniques and algorithms available to implement audio steganography, providing flexibility and options when it comes to hiding messages. This allows adapting to different scenarios and specific requirements.

Disadvantages:

5. Advanced detection: As steganography detection and analysis techniques advance, audio masking methods must also evolve to remain effective. If robust enough techniques are not used, there is a risk that the hidden message may be discovered.

6. Audio Quality: When hiding information in audio files, there is a chance that the sound quality may be affected. If a hidden message is introduced, there may be a perceptible degradation in audio quality, which could arouse suspicion in a detailed analysis.

7. Limited capacity: Although audio files have a reasonable capacity to hide additional information, this capacity is still limited compared to other storage media, such as text files or images. This means that the amount of information that can be hidden in an audio file can be restricted.

8. Implementation complexity: The implementation of audio steganography can be more complex than other information hiding methods. It requires specialized knowledge in signal processing techniques and steganography algorithms, which can make it difficult for people without experience in the field to apply.

4 Overview of Techniques in Audio Steganography

The following techniques were identified and evaluated for implementation:

1. Least significant bit (LSB) embedding: It is the simplest technique that works by replacing the least significant bits of the audio signal with the bits of the hidden message (Kamble & Chaurasia, 2021). It is the most efficient, but the least secure technique. The presence of a hidden message can be relatively easy to detect in an audio file encoded using LSB embedding. It finds applications in audio watermarking, copyright protection, and content distribution.

2. Frequency domain embedding: This technique is more secure than LSB embedding, and work by modifying the frequency components of the signal. Then, he has hidden message is embedded, making it more difficult to detect compared to LSB embedding. However, it is also less efficient in terms of the amount of data that can be embedded. It is commonly used in content distribution, copyright protection, and authentication.

3. Wavelet transform embedding: It is the most secure technique among the three discussed. It hides the message in the wavelet transform of the audio signal. The wavelet transform provides a more complex representation of the audio signal, making it highly resistant to detection. However, this enhancement in security comes at the cost of efficiency. Wavelet transform embedding requires more bandwidth compared to the other techniques. It is used in content distribution, copyright protection, and authentication.

4. Phase coding: It is a relatively simple technique that works by modulating the phase of the audio signal to encode data (Kamble & Chaurasia, 2021). It is relatively efficient but also relatively insecure. The phase of the signal can be easily manipulated, which can make it difficult to recover the hidden message. It finds applications in audio watermarking.

5. Spread spectrum: Spread spectrum is a more complex technique that spreads the signal in the frequency domain using a spreading code. The spreading code is a pseudo-random sequence of bits that expands the signal over a wider frequency range. This technique makes it difficult to identify the original signal, thus providing increased security (Kamble & Chaurasia, 2021). In terms of security, spread spectrum is comparable to Wavelet transform embedding, but is also more complex and requires more bandwidth. It is commonly used in military communications, secure audio transmission, and content distribution. It can require more bandwidth compared to Wavelet transform embedding, while being comparable in terms of security.

For ease of implementation, the LSB technique will be used throughout this work.

5 Implementation

LSB is one of the most popular methods used to hide information. Generally, it uses the principle of embedding each bit of the message in the least significant bit of the audio (see Figure 1). It is important to clarify that, in computing, the least significant bit is the rightmost bit. As we can see in Figure 1, using this technique, it allows hiding a single bit for every byte; Also, note that there is a 50/50 chance that the bit you are replacing is the same as its replacement, which means that half the time it takes, the bit does not change, helping to minimize wear on bits. the audio quality. This method stands out for its ease of implementation and the great opportunity to combine it with other techniques to hide information. In addition, it has the advantage over other methods, the large amount of data that it allows to encode. However, LSB is characterized by low robustness in terms of security, since the data is stored in a very deterministic way, allowing attackers to easily discover the message. However, to address its security limitations, a modification based on changing which bits are modified per sample.

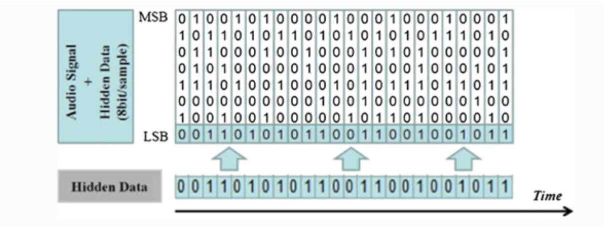

Figure 1

As an alternative proposal regarding the security and resistance of the hidden data, a minimum 4-bit replacement method was implemented, where the order can be exchanged with values from zero to 7, being zero the least significant bit. For example, if the values chosen for the order are [1,4,2,3], the bit that will be hidden in each byte will correspond to the position of the byte vector. In this case, security is significantly increased, since in order to extract a secret message from an LSB-encoded sound file, the receiver needs access to the sequence of sample indices used in the embedding process. This also makes the signal more resistant to some statistical analyses used for detection.

With this we were able to obtain a greater ability to hide information within the audio, allowing us to reduce the 80%-20% gap, where the audio had to be 80% larger than the message to hide. On the other hand, greater resistance to loss of information and audio compression is added by adding four bits instead of one; This is mainly because the last few bits tend to be more susceptible to modification or removal during audio compression. By using more bits, the impact of those operations is reduced. On the subject of security, by using more bits to hide information, we are practically adding another layer of security. For example, stego-analysis algorithms look for patterns and statistical features in your data to detect hidden information. By using more bits, hidden data becomes less distinguishable from noise and natural audio features, making it more difficult to detect by statistical analysis. Also, it's important to note that using more bits allows hidden information to blend better with noise and audio characteristics, making it even more difficult to detect. This can make hidden information visually indistinguishable from normal noise present in the last few bits of an audio file.

On the other hand, the implemented algorithm was optimized by almost 90. We achieve this by entering a special character at the end of the message to be encrypted; once this character is found, the process ends, and it is hidden in the audio. This prevented the algorithm from filling missing bits with zero spaces, which, while they may be insignificant, were taking up space in memory and slowing down the process.

5.1 Tests

For testing the impact of this technique, which should increase because of the modification of bits of higher significance, the following metrics are utilized for evaluation of the results. Ultimately, the results show that these metrics get increasingly worse whenever the order includes bits of higher significance, more frecuently.

MSE (Mean Square Error): The MSE is a metric commonly used to assess the quality of the reconstruction of hidden information. By comparing the original audio to the stenographic audio, the MSE provides a measure of how different the two are. A lower MSE indicates that the hidden information has been kept more faithful to the original audio. On the other hand, the MSE also provides information about the level of distortion introduced by hiding information in the audio. A higher MSE indicates a greater difference between the original audio and the stenographic audio, which may indicate greater perceived distortion.

PSNR (Peak Signal-Noise Ratio): is another metric commonly used in the compression and processing of images and audio. Provides a relationship between the power of the original signal and the level of noise or distortion introduced. A higher PSNR indicates less perceived distortion and better hidden audio quality. A high level indicates better signal quality and less noticeable distortion. PSNR is expressed in decibels (dB) and is calculated using the following formula:

$$PSNR = 20 * log10(MAX) – 10 * log10(MSE)$$

Where MAX is the maximum possible value of the audio samples and MSE is the root-mean-square error, which measures the root-mean-square difference between the original signal and the processed or distorted signal.

Algorithm structure

- Audio File: Original MP3 Audio
- Message: Message to hide
- Order: Order in which you want to hide the information in the audio bits.
- Conversion: translation to binary
- Encode: hide information according to the order established in the last bits
- New Sample: Audio with the hidden message

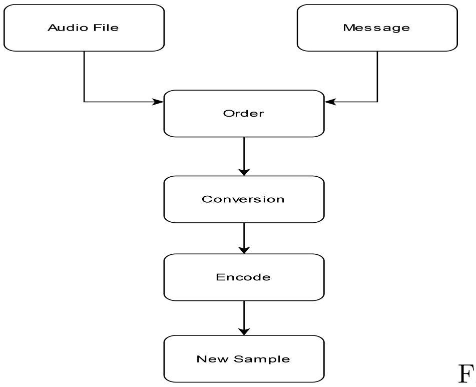

Figure2

6 References

1. Milind y Shivam Chaurasia, K. (21d. C.). ESTEGANOGRAFÍA DE AUDIO.

2. Kamble, Mr. K., & Chaurasia, S. (2021). AUDIO STEGANOGRAPHY.

3. Baneen Q Abd Muayad Kod Haider Ismael Shahadi Hameed R. Farhan. (2021, january). A Review and Comparison for Audio Steganography Techniques Based on Voice over Internet Protocol. https://www.researchgate.net/publication/360244669

4. Figure 1: Djebbar, F., Ayad, B., Meraim, K. A., & Hamam, H. (2012). Comparative study of digital audio steganography techniques. EURASIP Journal on Audio, Speech, and Music Processing, 2012(1). https://doi.org/10.1186/1687-4722-2012-25

5.Darsana, R., Vijayan, A. (2011). Audio Steganography Using Modified LSB and PVD. En Trends in Network and Communications (Vol. 2, pp. 11–20). Springer Berlin Heidelberg.

6. AUGUSTO y SALAZAR MAYO SILVIA PANDORA, C. N. M. (2010). "Investigación, análisis y pruebas de los Procesos de Esteganografía". UNIVERSIDAD TECNICA DE COTOPAXI.