Steganography: Hiding encrypted information

Course project of Análisis numérico

Julieth Stefanny Escobar Ramírez
Ingenieria Matemática
Universidad EAFIT
Medellín, Colombia
jsescobarr@eafit.edu.co

Sara Gallego Villada
Ingenieria Matemática
Universidad EAFIT
Medellín, Colombia
sgallegov@eafit.edu.co

Hernán Moreno Mora
Universidad EAFIT
Medellín, Colombia
homorenom@eafit.edu.co

Jacobo Rodriguez Ochoa
Ingenieria de Sistemas
Universidad EAFIT
Medellín, Colombia
jrodrigueo@eafit.edu.co

***Abstract*—Steganography, a technique used to hide a message inside a file (or in general, any media), is broadly used today in the digital media. In this work we study different techniques and make emphasis in the encryption of messages before its hiding in a file using some Steganography method.**

***Index Terms*—Hide, message, cryptography, spy, tampering**

I. Introduction

The security and prevention of your information to be hacked is one of the main concerns we have in the digital world. The main subject of this project is to show why steganography and encryption together make the transfer of information hidden in images or any other file more secure. We will show some algorithms which help to hide the information and in some cases, to encrypt it in an easy but safe way.

II. Objectives and justification

1) General objective:

- To hide encrypted messages in digital files by algorithms in order to transmit some information in a secure way.

2) Specific objectives:

- To define in a clear way the different concepts used in the process.
- To apply the concepts presented in a code
- To explain the importance of numeric methods in the process.

A. Justification

The sharing of sensitive information might be dangerous, as we don't want unauthorized persons to look into it, so we could want to hide it or encrypt it. One way to do it,is to hide it in an unrelated file; it is possible to hide the original file we want to hide, or its encrypted version. The hiding process is the same but in the second case we must encrypt the file before hiding it. In this way, if somebody find the hidden message, he will not be able to read it due to the encryption.

III. Concepts

A. What is steganography?

Steganography is a technique to hide information in a file of different characteristics to the hidden message, for example an image or an audio file. It is different to encryption, as the hidden message is not encoded, but hidden in the file. But it is required to know where the message is in order to be able to read it.

The word Steganography comes from the greek word *Steganographia*, which is a combination of two words: *steganos* meaning "covered or concealed" and *graphia* meaning "writing".

By contrast, Cryptography encodes the message, but does not conceal the fact that there is an encoded message. When using steganography, nobody suspects that there is a message hidden in the file.

Actually, the word is used for the concealment of messages in digital files, but the practice has been used since many years ago, going back in time to 1499 when the practice was first recorded in the writing of magic tricks.

And if we go to the ancient Greece, Herodotus mentions that Histiaeus sent a message to his friend Aristagoras, by writing the message in the shaved scalp of one of his servants, and sending him to his friend once his hair had grown, with the instruction to tell Histiaeus to "shave your head and look on".

During the Second World War it was common the use of invisible ink to hide messages. Vinegar, milk, fruit juices and urine were some of the substances used for it.

The **Steganalyis** is the practice of discovering or detecting hidden messages in some file (Pasive steganalysis) and/or extracting the message (Active steganalysis).

IV. Description

The process has different steps as follows:

1) Message File: this contains the message or the data to be concealed
2) Cover File: The file in which the message will be concealed, also called the Carrier File.
3) Secret Key: required to retrieve the message
4) Steganography Tool: the program or other media used to conceal the Message File into the Carrier File.
5) Stego-file: this is the file with the hidden message in it: Cover File + Message File
6) Communication channel: the media used to transmit the Stego-file to the receiver.

V. Techniques

A. Physical

It is the use of any physical media to conceal the message. The examples given of the shaved head and invisible ink are physical techniques. Other creative method used during the war, was to encode a Morse written message using knots in a yarn, and then using it to knit some piece of clothing used by the messenger.

B. Digital

With the advent of the personal computers around 1985, digital steganography has been steadily growing, and today it is possible to choose from among 800 different applications using different techniques. Some of them are:

- Use the lowest bits of image or sound files
- Use executable files, concealing the message in redundant parts of the program
- Video files: sometimes must be played at slower than normal speed to see the hidden message.
- Blog-steganography: using some social media to hide a partitioned message in different blogs on social media in some predetermined order and sequence. In this case, the carrier is the blogosphere.

VI. Digital Techniques

There are many digital techniques for steganography, here we will talk about some of them and focus on the one we will use.

A. LSB:

This is a well-known technique of steganography. It consists of modifying the least significant bits of an image or audio file, for the bits that constitute the hidden message you want to transmit. As the bits changed are a small percentage of the file, they will go unnoticed by the human eye [3].

The main technique for detecting LSB steganography is known as value pair analysis. Specifically, a byte can represent 256 values. If these 256 values are grouped by contiguous pairs, 128 possibilities are obtained. The key is to note that, despite changing the last bit of each byte, the value is still within the same pair, regardless of the new value of the last bit.

One of the disadvantages of the method is its limited capacity: only up to 12.5 percent of the image can be dedicated to the message.

B. Multiple substitution:

In this technique, it is proposed to replace more than one bit of the message in each byte of the carrier, multiple in fact, in general N, where N=1...8. / A 24-bit RGB image can be understood as composed of 3 layers or color channels (RGB), and composite would be those made up of the three simultaneous channels. A layer is one of its component binary images. When filtering by layers is applied, what is obtained is the view of one of them, or more than one overlapped [4].

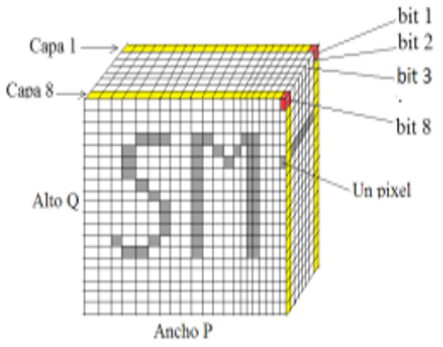


Figure 1 Example

C. Histogram Modification:

The histogram of an image contains the number of pixels that have the same gray level, which provides information about the brightness and contrast of the image, and can be used to adjust these parameters, remove certain annoying tones, etc [5].

VII. Methodology

A. State of the art

In the document we define our objectives so during the investigation we pretended to explain clearly the concept of Steganography, some of the applications and techniques used so in the development of the code we show the discoveries.

B. Implementation

Our implementation consist of an algorithm to hide a previously encrypted message. We decided to hide the message on images, however it can be done on any archive. First, the code reads the image where the message is going to be hidden, and prepares the image such that the less significant bits of each pixel are 0. Then receives the message wich is encrypted with a given password, this message is then converted to binary and subsequently the binary encrypted message is hidden in the carrier file using the least significant bits of each pixel and lastly saves the file with the hidden message.

The second part of the code, is used to retrieve the message back. Initially the code reads the file and stores the values of the less significant bits of each pixel in a list. Then this list wich contains the message in binary is converted to an alphanumeric message, afterwards this message is decrypted with the given password and the hidden message is returned. Although the objective of this work is steganography, we decided to include the message encryption to make it more secure. So, if the hidden message is detected, its contents cannot be known without the proper key. The algorithm we use, Steganography code .This part was done in python

1) Cryptography: The algorithm we use is Vigenere Cipher Program "is a form of polyalphabetic substitution. It is an algorithm of encrypting alphabetic text (e.g. "a" , "b", "c" etc) by using a series of interwoven Caesar ciphers based on the letters of a keyword. "[1] **Formula**

$$C \equiv T_i + K_i(mod \quad m)$$

Where Ci — ith char of the ciphertext
Ti — ith character of the open text
Ki — ith character of the key phrase (if the key phrase is shorter than the open text, which is usual, than the keyphrase is reapeated to match the length of the open text)
m — length of the alphabet [1] This part was done in Javascript Encryption code

C. User Manual

When you run the program, it will show a menu with 3 options. The user will have to write 1, for hiding a message; 2, to retrieve the message of a file; and 3, to exit the program. If you chose 1, the program will ask you to write the name of the file (with the extension) in which you want to hide the message. Then it will ask you to write the message you want to hide, consider that if your carrier file has a size of 1920x1080, the maximum number of characters your message can have is 388800 (this number changes depending on the size of the file). Afterwards the program will ask you for a password which will be used to encrypt the message and lastly it will save the image named "Encrypted.tif" on the folder you ran the program. If you chose 2, the program will ask you to write the name of the file that contains the hidden message. Then it will ask you to write the password in order to decrypt the message. Finally the program shows you the message that was hidden in the file.

D. Timeline

Activity	Tasked	Sources	Start date	Final date
Create repository	Stefanny E	Source	July 25	July 31
Tematic election	Everybody	Internet	August 5th	August 18th
Concepts investigation	Hernan H y Sara G	Source	August 19th	August 31st
Objectives and methodology	Stefanny E	Source	August 19th	August 31st
Implementation	Jacobo R	Results	September 1st	September 30th
Conclusions	Everybody	Source	October 25th	November 6th

VIII. Conclusions and Results

Even though Steganography has been used since thousands of years ago by humans, the digital era offers new easy alternatives to use. In this paper, we present only one of them using encryption to hide a message on an image file. The big power of modern computers, and the use of quantum computers, will make obsolete the encryption methods, as most of them could be broken in a short time. Is in this context where Steganography recovers its importance, as the message will need to be found before its decryption is attempted. Any digital file of the modern era could be the carrier of hidden messages. But you will need to know where the message could be, before attempting to read it.

IX. Steganography and Numerical Methods

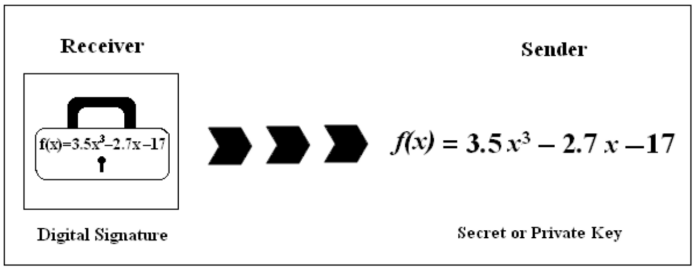
Numerical methods are not directly used in the steganography process, but they are broadly used during the encryption process. One of the most popular is the Newton-Raphson method, for finding successively better approximations to the roots of a function of real-values. A short description of the method follows:

Given a function $f(x)$ and its derivative $f'(x)$ the method begins with a first guess at x_0. Then we find $x_1 = x_0 - \frac{f(x_0}{f'(x_0}$.

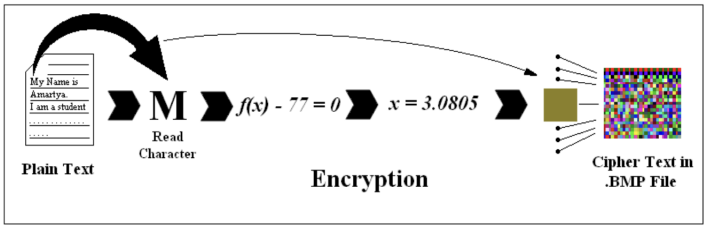
The process is repeated until a value with the desired accuracy is found:

$$x_{n+1} = \frac{f(x_n)}{f'(x_n)}$$

Some encryption methods are based on a polynomial function [6] obtained from the intended receiver of the message, which is used as the key for the encryption.



The receiver then uses the same polynomial function to find its roots, and use those roots in the decryption of the message. Newton-Raphson method is normally used during this process.



References

[1] Randerson. 'Programming Encryption Algorithms' 2017. [Online] Reference
[2] Ginni 'What are the application of Steganography?' 2022. [Online] Reference
[3] Jesús Díaz. 'Esteganografía y estegoanálisis básicos' 2015. [Online] Reference
[4] Mg. Ing. Guillermo Sergio Navas, Mg. Ing. Carlos Gustavo Rodríguez Medina. 'Esteganografía por sustitución múltiple, Implementación en Matlab'. [Online] Reference
[5] Carlos L, Velasco-Bautista Julio, López-Hernández Mariko Nakano-Miyatake Héctor M, Pérez-Meana. 'Esteganografía en una imagen digital en el dominio DCT' 2007. [Online] Reference
[6] Amartya Gohsh and Anirban Saha. 'A numerical method based encryption algorithm with steganography'. [Online] Reference