TEXT, AUDIO AND VIDEO STEGANOGRAPHY

Information Technology Department, National Institute of Technology Karnataka, Surathkal

Atharv Belagali (181IT208)
Information technology
National Institute of Technology Karnataka
Surathkal, India

Gagandeep KN(181IT215)
Information Technology
National Institute of Technology Karnataka
Surathkal, India

Prithvi Raj (181IT234)
Information Technology
National Institute of Technology Karnataka
Surathkal, India

***Abstract*— Steganography is basically the art of hiding a message inside another message. In the digital form (computers), a text, audio or an image maybe even a video is concealed within another file.**

It is better when compared to cryptography because the third party doesn't really know that there exists an hidden message inside a file. Whereas in case of cryptography the content is kind of blended in and is plainly visible but in a puzzling format but nevertheless arises interest to whoever is looking at the file

Looking into digital steganography, e-communications consists of steganographic coding within a layer usually the transport one, or instance a document/image file. To accommodate lager file for better steganographic transmission media files are used. An instance for it is: Let's say a sender might start with an innocuous image file and adjust the color of every hundredth pixel to correspond to a letter in the alphabet. The change is so subtle that someone who is not specifically looking for it is unlikely to notice the change.

In our paper we try to implement the three major parts of Steganography namely:

1. Text

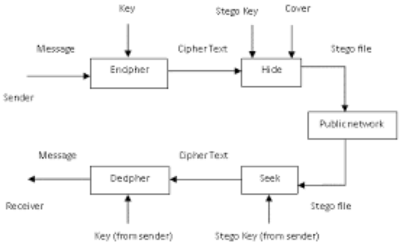
2. Image

3. Audio

The Ultimate goal is to be able to conceal data in the above three given formats and be able to reveal them by the recipient.

INTRODUCTION

In the modern era, the data transmission has reached a paramount pinnacle in terms of speed but what about the security of it? The result has made it taht much easier for the data to be interrupted by a third party whose intentions maybe malicious. The data can be altered destroyed ,etc.. , by anonymous users or for unauthorized access by eavesdropper, attacker, and etc. Therefore, secure and safety of data transmission has gained so much relevance in the modern era. Information security has 2 major techniques: Data encryption and data hiding to avoid the above mentioned circumstances.



Encryption basically converts the data into a cypher text during the data transmission phase, which is tough to comprehend unless presented with a decryption key. Data hiding basically consists of hiding a secret data into some random data . The random data are usually digital media like text, audio, image, video, and multimedia and called the cover for secret data. Data hiding has two main branches, steganography and watermarking. We will be looking specifically at steganography in this paper.

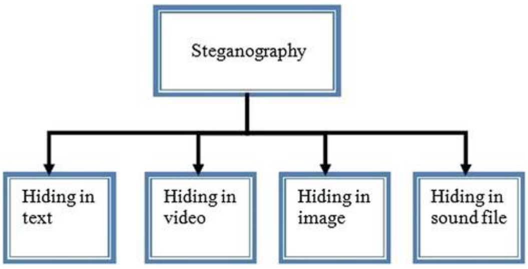


Image Steganography: .

This technique is used for hiding data inside the image. Hiding the data in an image file causes a slight alteration to the image such as: the bit quality will be changed (e.g., resolution or color),pixels also change when compared to the original image . On the other hand, the limitations on the binary string length, size, and number of pixels in the image, which each pixel hides one bit in it using one of the image steganography algorithms. The simplest method is to append the message to the end of the image. So the image is viewed by some image viewer application and the text at the EOF(end of file) is generally ignored. The image steganography can be classified as spatial domain steganography and transform domain steganography.

Text Steganography:

Text steganography is a mechanism of hiding secret text message inside another text as a covering message or generating a cover message related with the original secret message.

There are three main categories used to hide text-in-text messages:

1.Format based
2.Random and statistical generations
3.linguistic method.

Audio Steganography: .

Here a secret message is embedded into a digitized audio signal which results in slight altering of binary sequence of the corresponding audio file. There are several methods available for audio steganography. It involves hiding data in audio files. This method hides the data in WAV, AU and MP3 sound files. There are different methods of audio steganography which are listed below:

i) Low Bit Encoding
ii) Phase Coding
iii) Spread Spectrum.

OBJECTIVES:

1. Differentiate and understand what Steganography and Cryptography are.
2. Jot down the different types and their sub types in Steganography
3. Implement each of their sub types and be able to conceal data in any of the formats I.e. audio, text, image.

SYSTEM REQUIREMENTS:

1. Software requirements:
   Windows/Linux /macOS

2. Hardware Requirements:
   i3 processor
   4GB RAM

METHODOLOGY:

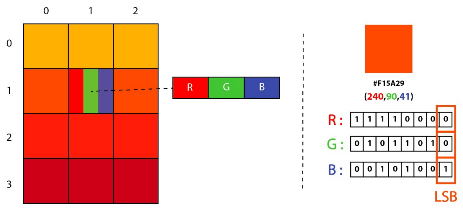
IMAGE STEGANOGRAPHY:

1. As we know the image is made up of set of pixels and each pixel consists of 3 values(red, blue, green).

2. We take the secret message and each byte of it is converted into 8 bit binary code using its ASCII values.

3. For every byte we take up first three pixels which are not encoded and alter it to even number if it has a corresponding zero value in the binary code and odd number if it has a corresponding value of 1 in the binary code.

4. The first 8 values of 3 pixels are used to signify the message as shown above and the last value is used to signify weather the message is completed or not. If its more data then the last value is even.



TEXT STEGANOGRAPHY :

1. It uses a program called SNOW(Steganographic Nature of White Spaces)

2. As the name suggests, the main idea behind this type of steganography is to hide text

inside a text by appending tabs and spaces at the end of the line.

3. The secret message is written with hidden characters among the white spaces available

between the normal text file.

4. The text message is not visible to the attacker and hence this kind of technique becomes

quite effective.

5. To hide a message “xyz” using the password “nitk_student” using the file “infile” and

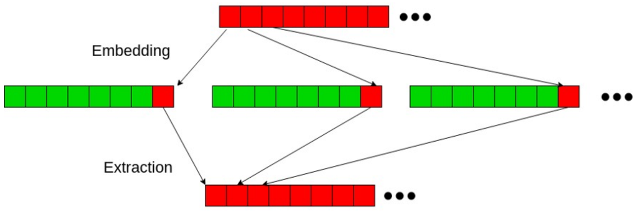
generating a new file called “outfile” following command is used in snow:

```
./snow -C -m "xyz" -p "nitk_student" infile outfile
```

AUDIO STEGANOGRPAHY:

LSB (Least Significant Bit) Algorithm:

Here we will look at a popular algorithm that is used to hide secret audio, text , or image in an audio file. The main idea behind this algorithm is to exploit the least significant bit of every or some bytes of data to hide the secret message into it. As it is the least significant bit that is being manipulated in the audio data it very much resembles the original audio file making it possible to maintain the secrecy of the hidden message.



Each byte of original audio data's least significant bit is replaced with the bit from hidden data. Hence the number of bits in a hidden message is equal to the number of bytes required to store the hidden message. For decoding the hidden message, all the least significant bits are placed in an order which gives us the hidden message as shown in the above figure.
Innovation done by us Instead of manipulating every byte of the original data continuously to store the hidden message, we use alternate bytes of data. This modification helps to improve the secrecy of the message by resembling it very closely to the original audio file.

LITERATURE SURVEY:

1. ***An overview of steganography : a data hiding technique -Priya Pareek1, N. Monica2***

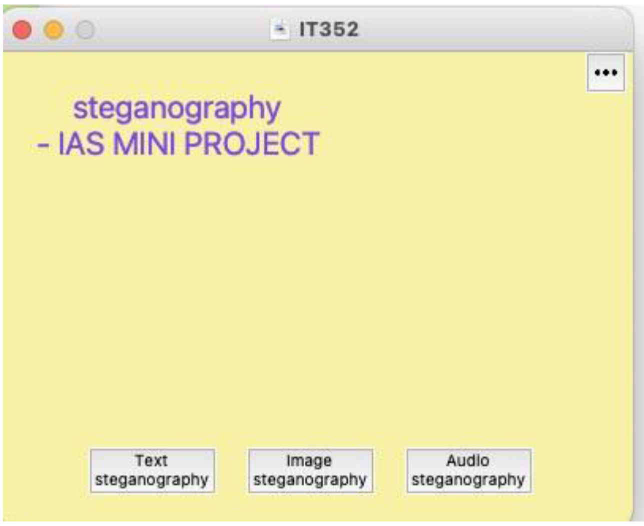
From this paper , we extracted information regarding the three formats of Steganography i.e., via text, image, audio. It has a pictorial and theoretical representation of the data hiding techniques mentioned above including video format too.

In the text steganography secret message is hidden using word or line shifting and the techniques employed are format based, statistical and linguistic. Coming to image part the classification has been done into two categories namely spatial domain and transform domain. Spatial domain consists of hiding secret data into the bits and LSB(Least Significant Bit) also plays a role here. Whereas in the transform domain steganography hides the information in the frequency domain by altering magnitudes of Discrete Cosine Transform (DCT).Next up we have the audio steganography which has 5 categories to it: echo hiding, parity coding, phase coding, spread spectrum , tone insertion.

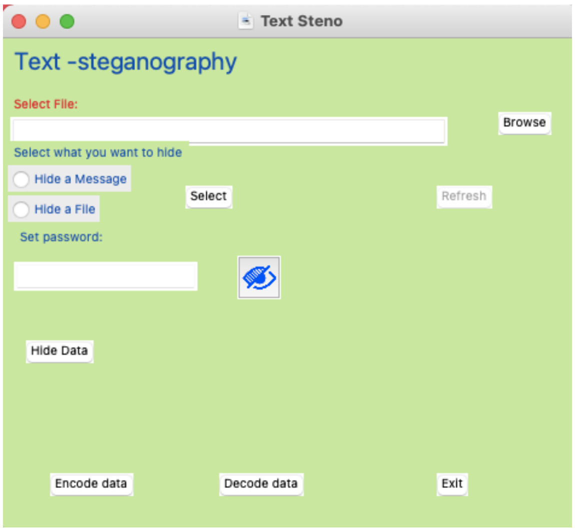
2. An efficient steganographic technique for hiding data-Dalia Nashat &Loay Mamdouh

This paper describes and gives a profound understanding on LSB which falls under the category of Image Steganography. The goal of the proposed method is to enhance the capacity taking high visual quality into consideration. To achieve this goal, some LSB of the cover image are inverted depending on the secret data for embedding instead of replacing LSB with the secret data. First, the maximum and minimum values in the secret data are determined then subtract all values of the secret data from this maximum value. Finally, make a division for the results and embed the new results into the cover image to obtain the steno image. The results show that the proposed method gives high capacity and good imperceptibility in comparison with the previous methods.

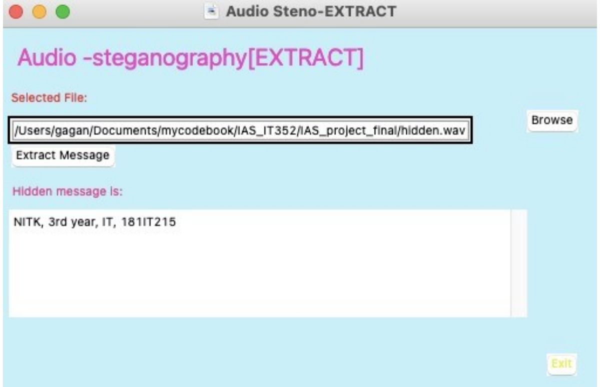
Results:



Landing page:



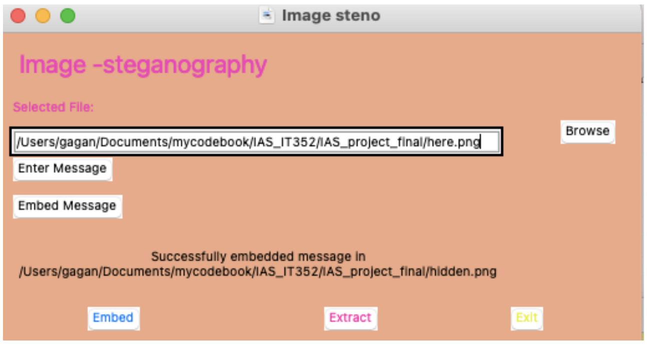
Text Steganography :



/Users/gagan/Documents/mycodebook/IAS_IT352/IAS_project_final/hidden.wav

NITK, 3rd year, IT, 181IT215

Audio steganography:



/Users/gagan/Documents/mycodebook/IAS_IT352/IAS_project_final/here.png

Successfully embedded message in
/Users/gagan/Documents/mycodebook/IAS_IT352/IAS_project_final/hidden.png

Image steganography:

Conclusions

As we know that cryptography gives privacy. Steganography provides secrecy, it gives a method to communicate covertly. We pick Steganography over cryptography because it provides us with the following advantages: the potential capability to conceal the existence of confidential data, the degree of hardness increase to detect concealed data, enhances the secrecy of the hidden data. In this paper we have successfully implemented 3 major types of Steganography namely : text, audio, image. We have taken reference from various methods described as given in paper 1 and paper 2 in our literature survey and have implemented our model based around them.

REFERENCES

1. *https://searchsecurity.techtarget.com/definition/steganography*

2. *https://en.wikipedia.org/wiki/Steganography*

3. *https://www.ijcaonline.org/volume9/number7/pxc3871887.pdf*

4. *https://link.springer.com/article/10.1186/s42787-019-0061-6 (paper 2)*

5. https://www.academia.edu