Algorithm for Double Encryption and Generating Pseudo Random Sequence for Steganography

Rishabh Omar\textsuperscript{1}, Tahseen Anwar Tahir\textsuperscript{2}, Tarun Garg\textsuperscript{3}, Vaibhav Sharma\textsuperscript{4}
\textsuperscript{1, 2, 3, 4}Department of Information Technology, IMS Engineering College, Ghaziabad, U.P.

Abstract: Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. Many different carrier file formats can be used, but digital images are the most popular because of their frequency on the internet. For hiding secret information in images, there exists a large variety of steganography techniques some are more complex than others and all of them have respective strong and weak points. Different applications may require absolute invisibility of the secret information, while others require a large secret message to be hidden. This project intends to give an overview of image steganography, its uses and techniques. It also attempts to identify the requirements of a good steganography algorithm and briefly reflects on which steganographic techniques are more suitable for which applications.

Keywords: Steganography, Cryptography, Encrypt, Decrypt, Pseudo random number.

I. INTRODUCTION

One of the reasons that intruders can be successful is the most of the information they acquire from a system is in a form that they can read and comprehend. Intruders may reveal the information to others, modify it to misrepresent an individual or organization or use it to launch an attack. One solution to this problem is, through the use of steganography. Steganography is a technique of hiding information in digital media. In contrast to cryptography, it is not to keep others from knowing the hidden information but it is to keep others from thinking that the information even exists.

Steganography becomes more important as more people join the cyberspace revolution. Steganography is the art of concealing information in ways that prevent the detection of hidden messages. Besides cryptography, steganography can be employed to secure information. In cryptography, the message or encrypted message is embedded in a digital host before passing it through the network, thus the existence of the message is unknown. Besides hiding data for confidentiality, this approach of information hiding can be extended to copyright protection for digital media: audio, video and images.

Information hiding is an emerging research area, which encompasses applications such as copyright protection for digital media, watermarking, fingerprinting, and steganography.

In watermarking applications, the message contains information such as owner identification and a digital time stamp, which usually applied for copyright protection.

Fingerprint, the owner of the data set embeds a serial number that uniquely identifies the user of the data set. This adds to copyright information to makes it possible to trace any unauthorized use of the data set back to the user.

Steganography hides the secret message within the host data set and presence imperceptible and is to be reliably communicated to a receiver. The host data set is purposely corrupted, but in a covert way, designed to be invisible to information analysis.

II. STEGANOGRAPHY

Steganography is the practice of hiding private or sensitive information within something that appears to be nothing out to the usual. Steganography is often confused with cryptology because the two are similar in the way that they both are used to protect important information.

The difference between two is that steganography involves hiding information so it appears that no information is hidden at all. If a person or persons views the object that the information is hidden inside of, he or she will have no idea that there is any hidden information, therefore the person will not attempt to decrypt the information.

What steganography essentially does is exploit human perception, human senses are not trained to look for files that have information inside of them, although this software is available that can do what is called Steganography. The most common use of steganography is to hide a file inside another file.\textsuperscript{[1]}
III. BLOCK DIAGRAM

A. Encryption

INPUT: Message, Secret Key

Level 1 Encryption

PRN Generation

Embedding in Image

OUTPUT: Encrypted Image

B. Decryption

INPUT: Encrypted Image, Secret Key

PRN Generation

Extraction from Image

Level 1 Decryption

OUTPUT: Message

IV. METHODOLOGY

The Algorithm consists of mainly two independent phases, the Encryption phase and the Decryption phase. The Encryption phase consists of three inputs (message, secret key and an image) and one output (encoded image). The Decryption phase consists of two inputs (encoded image and secret key) and one output (message).\(^3\)

The Encryption phase of algorithm starts with the user entering the secret message he/she wants to send to the desired person, also it takes a secret key as an input. This message and secret key will be passed on to the Level 1 Encryption process, which uses Base64 algorithm for encrypting the message with the secret key, then it converts the level 1 encrypted message into Binary code words which is to be embed into the image. To make it more secure the algorithm is generating the Pseudo Random Sequence to increase randomness of embedded binary encoded level 1 encrypted message into the image. A unique value is generated by the algorithm using the secret key which is used to seed the random function to generate the Pseudo Random Sequence. Afterwards it requires an image input to perform the Steganography process. The input image is then pixelated into the list of tuples containing the RGB of
single pixel. Now with the help of previously generated Pseudo Random Sequence, the algorithm embeds the binary code wards of level 1 encrypted message according to the sequence. To embed the binary code, algorithm picks a particular tuple according to the sequence and starts embedding the particular bit of binary code into RGB of pixel, using the LSB technique, now saving the updated tuple back to list, this process is now repeated for every bit of binary code until whole message isn’t embedded. Finally, the updated list is converted back to the image and algorithm saves that encoded image to the same path. The final image contains the secret message, now which can be send to anyone.

The Decryption phase of algorithm starts with the user entering the encoded image containing the secret message and the secret key which is used to decrypt the secret message. The secret key is used to generate the Pseudo Random Sequence which is to be used for extraction of message from image. The algorithm pixelates the input image to form the list containing the tuples of RGB from where the data has to be extracted. The extraction process starts with the selection of tuple according to the sequence generated previously. With the help of LSB technique, the bit is extracted from each RGB of selected tuple. The extracted binary bits are converted back to cipher text. The cipher text is now deciphered back to original text with help of input secret key, using the Base64 algorithm. If the input secret key is same as the key which is used to encrypt the data, then the deciphered text will be the secret message embedded by the sender.

V. IMPROVEMENTS

It is also important to discuss that though steganography was once undetected, with the various methods currently used, it is not only easy to detect the presence but also retrieving them is easier. By Steganalysis using Machine Learning, the presence of embedded secret message can be easily detected, since LSB technique saves the data in sequential order in the image. This type of Steganalysis visualizes the image pixels and trace the change in the LSB of RGBs of every pixel. After tracing every pixel, the machine learning algorithm checks if the changes are sequential or not. If changes are found sequential, the algorithm confirms that the image contains some embedded secret code. [2]

Now with the help of Pseudo Random Number (PRN), the process of embedding the bits is improved by using the random sequence generated before. The pseudo random sequence provides a random pixel sequence to embed the data in it until whole message is embedded into image, which breaks the sequential order of embedding. Breaking the sequence makes it more difficult for machine learning algorithm to detect the presence, since the change in pixels is distributed throughout the whole image decreasing the chances to trace the change. Thus, the algorithm with pseudo random sequence generator helps in increasing the randomness of data embedding which fails the Steganalysis technique to detect the presence of secret message. [4]

VI. ADVANTAGES

A. The steganography is for protecting the data, such as in the field of media where the copyrighting ensures the authentication. [5]
B. By using the Pseudo Random Sequence, the randomness of data embedding is increased which leads to more secure steganography.
C. The algorithm makes Steganalysis more difficult.
D. The steganography can be used by the intelligence agencies for transmitting their secret information.
E. Nobody apart from receiver and sender can retrieve the message.

VII. CONCLUSION

Steganography is not intended to replace cryptography but rather to supplement it. If a message is encrypted and hidden with a steganographic method it provides an additional layer of protection and reduces the chance of the hidden message being detected. Steganography is a really interesting subject and outside of the mainstream cryptography and system administration that most of us deal with day after day. “You never know if a message is hidden”, this is the dilemma that empowers steganography. As more emphasis is placed on the areas of copyright protection, privacy protection, and surveillance, we believe that steganography will continue to grow in importance as a protection mechanism.

This project deals with Steganography in Image files using Least Significant Bit (LSB) coding. This project can uplift by considering following measures:

A. A more sophisticated approach can be implemented by using a Pseudo-Random Sequence Generator to spread the message over the image file in a random manner.
B. This project can be extended by using other media files like audio, video and other complex formats of audio and image.
REFERENCES

[1] “Wikipedia - The Free Encyclopedia. Steganography.” Available at http://en.wikipedia.org/wiki/Steganography;

[2] “Using steganography for securing data, not concealing it” https://searchsecurity.techtarget.com/tip/Using-steganography-for-securing-data-not-concealing-it

[3] “How to Hide Secret Data Inside an Image or Audio File in Seconds” https://searchsecurity.techtarget.com/definition/steganography

[4] “INTELLIGENT RANDOM IMAGE STEGANOGRAPHY” http://shodhganga.inflibnet.ac.in/bitstream/10603/17471/13/13_chapter_04.pdf

[5] “REVIEW ON STEGANOGRAPHY FOR HIDING DATA” https://www.ijcsmc.com/docs/papers/April2014/V3I4201468.pdf

[6] “Text Steganography Methods and its Tools” http://spublication.com/iist/2014/april14/79.pdf

[7] H. Wang and S. Wang. ”Cyber warfare-Steganography vs. Steganalysis,” Commun. ACM, vol. 47, no. 10, pp. 76-82, 2004.

[8] “Scanning USENET for Steganography” http://niels.xtdnet.nl/stego/usenet.php

[9] “New LSB-based colour image steganography method to enhance the efficiency in payload capacity, security and integrity check” https://www.ias.ac.in/article/fulltext/sadh/043/05/0068

[10] “Image Steganography Techniques: An Overview” https://www.researchgate.net/profile/Osamah_Algersh/publication/292310394_Image_Steganography_Techniques_An_Overview/links/57f642ae08da3ce574080/Image-Steganography-Techniques-An-Overview.pdf