VOL. 11, NO. 13, JULY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

SECURE DATA HIDING IN IMAGE OVER ENCRYPTED DOMAIN

Priyanka Kumari, Preeti Reddy M. and B. Bharathi
Department of Computer Science and Engineering, Sathyabama University, Chennai, Tamil Nadu, India
E-Mail: priyankakumari1895@gmail.com

ABSTRACT

Steganography is a concept of hiding or concealing data in a cover object while the sender and receiver are able to communicate within themselves. The security of data has been prime concern for the people from past and many research works are still going to find out convenient methods to safeguard the communication between people. Intruder can easily detect the information in the links as internet does not have secure links. The transmission of data is required to be protected by limiting the chances for its detection while transmission. In this paper we have proposed a technique for image steganography that embeds or hides important or private messages or files into files like images, audio and video without affecting the quality of actual files. It is achieved by using the Least Significant Bits (LSB) of these files for embedding data in the portion of image, which are not used by the image viewing tools. It allows to embed the messages or files in encrypted form using steganography and uniform embedding algorithm which means that once encrypted, the message or file could be retrieved back (or decrypted) from the image only after specifying the correct password which was given by the sender at the time of embedding data into the image(or encryption).

Keywords: cryptography, steganography, least significant bits (LSB), key, encryption, decryption.

1. INTRODUCTION

Cryptography is an art of hiding and passing any information in text. The concept for the paper is somewhat derived from cryptography but a similar concept named as steganography. “Steganography” is used for transmission of data secretly. Steganography is basically derived from two Greek word steganos that literally means “covered” and graphic means “writing” in Greek, therefore Steganography means “covered writing” [6][5].

“Steganography” can be called as art and science of writing secret messages in a way such that only the sender and intended receiver is able to suspect the existence of the message. It can rather be called as a form of security through obscurity. The word Steganography is derived from Greek origin meaning “concealed writing.” The first ever use of the term steganography was in 1499 by Johannes Trithemius in his papers namely: Steganography, a treatise on cryptography and Steganography disguised as a book on magic [6]. Basically cryptography is all about encryption and decryption and the most important element i.e., key. Based on Key, cryptography is of two types symmetric and non-symmetric where symmetric uses same key for both sender and receiver side and non-symmetric uses both public and private key. The earlier used concept in cryptography was DES (DATA ENCRYPTION STANDARD) which failed due to brute force attacks [4]. Due to the brute force attack and its effect, the concept of steganography came into the frame. In steganography, the messages do not attract attention to themselves, giving it as advantage over cryptography. The sender uses an image to embed the secret message using different embedding algorithms and techniques and while encryption of data; a secret key is provided [Figure-1]. Now the message is send to receiver and the receiver is able to retrieve the message only after he provides the correct secret key given at the time of encryption [Figure-2]. The main goal is intended to provide the transfer of secret message m embedded in the image data d, to obtain new image data d', practically indistinguishable from the data d, by people, in such a way that a normal person cannot detect the presence of m in d’ with the objective of hiding information in an undetectable way to provide extra security by means of cryptography, to prevent extraction of the hidden information.

The basic pictorial representation of steganography using the cryptography concept is shown in Figure-1 and Figure-2.

At sender side:

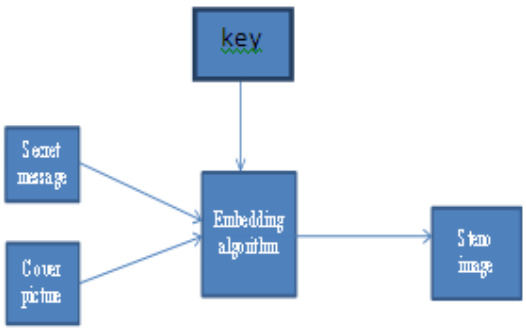


Figure-1. The sender side representation of steganography.

At receiver side:

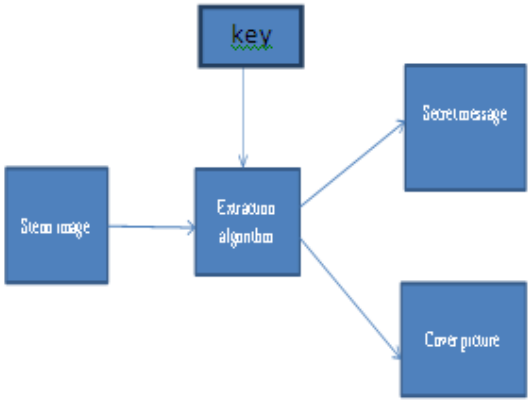


Figure-2. Receiver side representation of steganography.

2. LSB ENCODING

Different types of data like word file, images, pdf file, win word file or even multimedia file can be embedded in other images or multimedia files such that the original data is not disturbed or distorted, it appears to be normal [7]. In a normal image, the rough space that does not contain any portion of the image is used for embedding data in any multimedia form (images, pdf file, word file). AES (Advanced Encrypted Standard) algorithm is used for this purpose. In AES algorithm key of 128, 196, 256 bit are used which is difficult to crack via brute force method as it will lead to billions of combinations [1]. Whereas in DES algorithm just 56 bit key is used that is 256 combination which is prone to brute force attack.

Least Significant Bit (LSB) technique is used where the file to be embedded is converted into binary code [5] [6]. The number of bits in binary code will decide the fragment of clipart of image. It is the most common technique used to encrypt data in image by just fixing the least significant bit of a pixel of the selected image. The variation in insertion of bits takes place according to the size of image. One pixel of an image contain RGB (Red, Blue, Green) component. In LSB technique, for an 8 bit image the 8th bit of each byte will be changed by any other bit. In this LSB substitution technique, the 8th bit or the least significant bit of an image is replaced by the secret message bit and thus a new image is formed in which the secret message is embedded which is known as steno image. The image on which LSB operation is to be done is broken into RGB components and side by side the message to be embedded is also converted into binary format [4]. Now the secret message taken 8 bits at once will be embedded to the RGB component of the pixel of the image in the order of 3, 3 and 2 respectively[6]. This order of insertion is followed because the chromatic influence of red and green colour is less than the blue colour to normal human eyes. So, only two bits are inserted to the blue colour component [5] [6]. The formula used to find the position to hide the bits of secret message in image is:

$$P = a \% b \quad (a)$$

Where,

P = position of the LSB where the secret message bit is to be inserted

A = position of secret message bit which is to be embedded in image bit

B = number of bits considered for LSB

While encrypting data this formula is used for embedding and then in decryption side after combination of bits and the information will lead to retrieval of the secret message. Suppose we take a pixel of image whose RGB component is (150, 55, 77) and the message to be sent is 143(i.e., 10001111) [equation (a)].

The RGB pixel component of cover image will be:

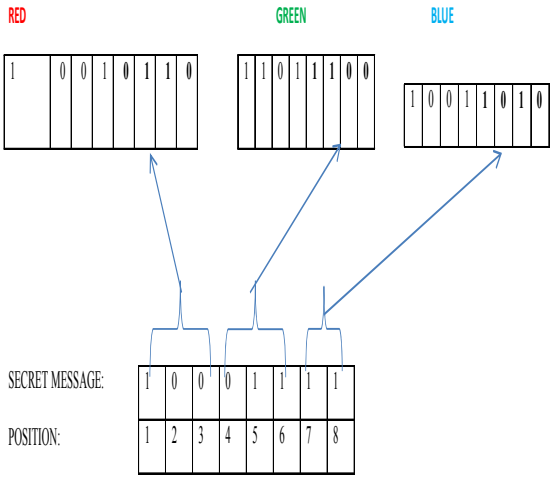
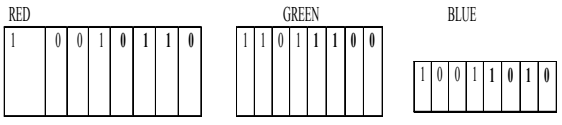


Figure-3. The RGB value of pixel and secret data.

Now we have to find the position in the last significant bit of the image component to insert the bits of secret message. The position can be easily calculated by using the formula (a). Here we have taken number of LSB bits to be four. The position value for red will be P=1, 2, 3, for green P= 4, 1, 2, for blue P= 3, 4. Now after the positions are calculated we need to replace the least significant bit of image with secret message.

Before embedding secret message



After embedding secret message

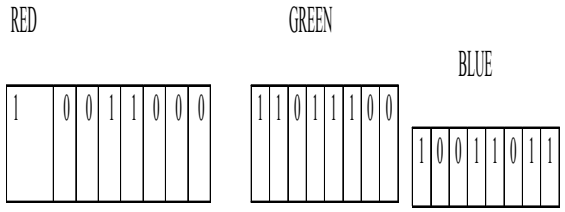


Figure-4. Image bits after change in least significant bit.

3. ARCHITECTURE AND MODULES

The architecture gives a detailed description where the image and the object are uploaded and the secret message or file to be embedded is also taken. Thereafter their binary conversion and the least significant bit formulation is done and the data is embedded in the image and a secure key is provided so as to send the data with high security. The image which is ready to be sent is known as steno image. At receiver side, they will decode and retrieve the message using the key which was used while sending the image. The architecture is illustrated in Figure-5.

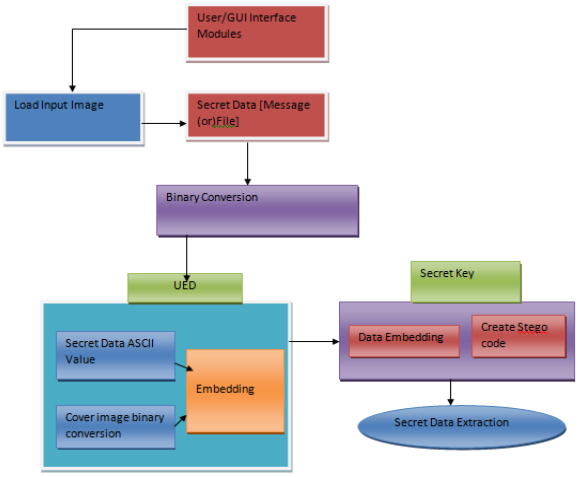


Figure-5. Architecture of the data embedding and retrieval.

The process of data embedding consists four phases:

a. Embedding a message
b. Embedding a file and compression
c. Retrieving message from a master file
d. Retrieving embedded file from a master file

Phase 1: embedding a message

In this we embed a message within an image, audio and video files. Here we have to specify the master file (image) and the output file. The process of embedding information during JPEG compression results in a steno image with a high level of invisibility, as embedding occurs in transform domain. On internet, JPEG file format is considered as the most popular and the sizes of the image are also small because of the compression, which makes it to be used as least suspicious algorithm. However, it is difficult to implement compression process as it is a very mathematical process. Over an open system environment for communication, JPEG file format serves most suitable for images and it is also used by almost every applications.

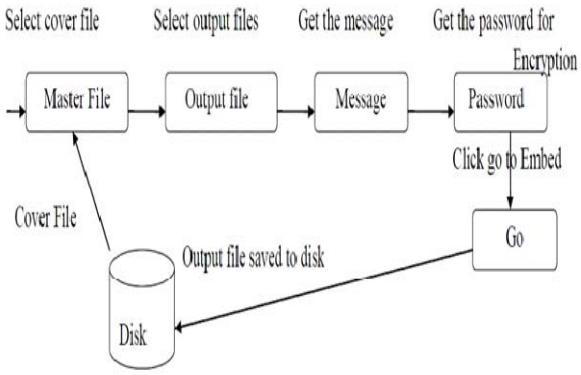


Figure-6. Data flow diagram for embedding a message.

Phase 2: embedding a file and compression

Here we have to specify the master file and the output file. At first the RGB color representation is converted to a XYZ representation, for compression of image into JPEG format. In this representation the X component represents the luminance (or brightness) and the Y and Z components stand for the chrominance (or color). According to research, the eye of a human being is more perceptive towards the change in the brightness (luminance) of a pixel than to the change in its color. This fact is imposed by the JPEG compression by down sampling the color data to diminish the size of the file. The color components (Y and Z) are divided between horizontal and vertical directions, which will decrease the size of file by a unit known as "factor". In JPEG quantization coefficient is used to divide all the values in a block. The coefficients and the rounded integer values of the results are encoded using Huffman coding to further minimize the size.

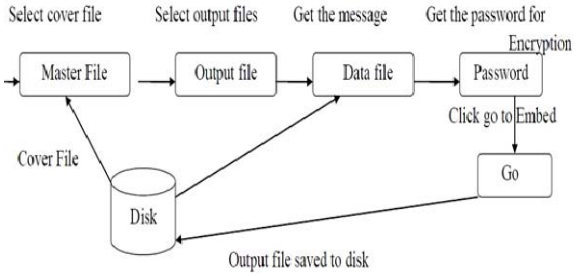


Figure-7. Data flow diagram for embedding file.

Phase 3: retrieve embedded message from image

We will retrieve a message which is embedded on an image, audio and video files. Here we have to specify

password and the steno file. The retrieving of the embedded message includes extraction of secret message from the file which is independent of their respective file format. Once the message has been received by the receiver, the original message needs to be retrieved. This can be done by reading the embedded data from the retrieved file. The data is in the bytes format when we read it from file. This message has to be converted into the suitable output file format. The process of embedding data that is to be hidden into an image needs two files. The first one is the innocent looking image that will allow embedding of the hidden information, called the cover image. The second file is the message- the content to be hidden. A message could be any plain text, cipher text, other images, or anything that can be embedded in a bit stream. The combination of cover image and the embedded message make a steno- image.

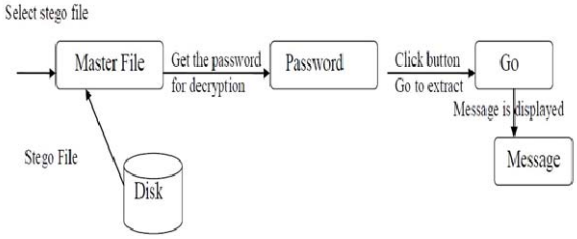


Figure-8. Data flow diagram for retrieving message from image.

Phase 4: retrieve embedded file from image

We will retrieve a file which is embedded in an image, audio and video files. Here we have to specify password and the steno file. Most software used for steganography does not supports the usage of JPEG images. Instead the use of lossless 24-bit images such as BMP is recommended. 256 -color or gray scale images is next best recommended alternative for 24-bit image. In 8-bit color images like GIF files, each pixel is represented by a single byte, and each pixel nearly refers to a color index Table (a palette) having 256 possible color [8]. The pixels value is between 0 and 255. At the selected position of pixel, software paints it with the color on which it is currently indicating. Many steganography experts recommend 256 shades of gray to be used in images. There is gradual change in shades from byte to byte, so mostly Gray scale images are preferred. And while considering an image in which we have to hide information, it can hide information better if the value change is less between palette entries.

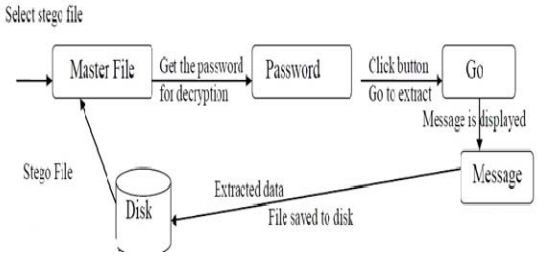


Figure-9. Data flow diagram for retrieving file from image.

4. CONCLUSIONS

This technique is used to induce a strong and effective ways for securing the data from outsiders and hackers and to safely send the information to their respective destination. It is suitable for different types of audio, video and text file formats without affecting its quality and size post the encryption process. The Sign encryption and encryption techniques contributes to the robustness of the security system. For an instance, the approach for patchwork provides effective robustness against a variety of security attacks but is capable of hiding very less amount of data and secure information. Least Significant Bit (LSB) in both BMP and GIF can be done, but both approaches result in suspicious files that increase the probability of detection when in the presence of a hacker.

REFERENCES

[1] 2001. Federal Information Processing Standards, Advanced Encryption Standard (AES). Federal Information Processing Standards Publications (FIPS PUBS).

[2] J. Fridrich J. Kodovský and T. Pevný. 2007. Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities. In: Proc. 9thACM Workshop Multimedia Security, Dallas, TX, USA.

[3] J. Judas, J. Fridrich and T. Filler. 2011. Minimizing additive distortion in steganography using 00000000000000000000 syndrome-trellis codes. IEEE Trans. Inf. Forensics Security. 6(3): 920-935.

[4] Kousik Dasgupta, J. K. Mandal, Paramartha Dutta. 2012. Hash Based Least Significant Bit Technique for Video Steganography (HLSB). International Journal of Security, Privacy and Trust Management (IJSPTM). 1(2).

[5] Ankit Chaudhary, J. Vasavada, J. L. Raheja, S. Kumar, M. Sharma. 2012. A Hash based Approach for Secure Keyless Steganography in Lossless RGB Images. 22nd International Conference on Computer Graphics and Vision.

[6] Anil Kumar *, Rohini Sharma. 2013. A Secure Image Steganography Based on RSA Algorithm and Hash-LSB Technique. ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering. 3(7).

[7] Miss. Priyanka V Deshmukh, Prof. Aniket K Shahade, Prof. Gajendra Y Patil. 2015. Higher LSB Optimize Data Hiding Mechanism on Encrypted Image. 2015 International Conference on Pervasive Computing (ICPC).

[8] S. Doyle. 2001. Using short message service as a marketing tool. Journal of Database Marketing. 8(3): 273-277.