EasyChair Preprint

№ 3764

Information Hiding in Images Using Steganography Techniques

Anoop Kumar

July 6, 2020

Information Hiding in Images Using Steganography Techniques

ANOOP KUMAR

ABSTRACT

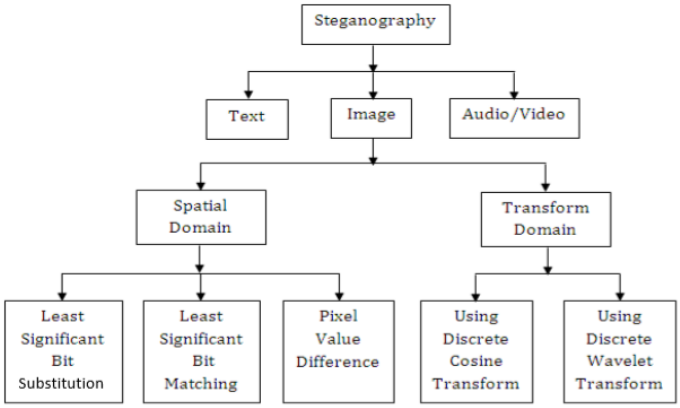
Steganography is the process of encrypting users messages within an image that others can't read the content of the encode message. The main purpose of Image Steganography is to help secret communication. Image Steganography is mostly used in securing high tech information and user's privacy. The paper will explain how Image Steganography method is utilise in today's world and it will provide practical understanding of Image Steganography and its uses. It is difficult to find secret message and methods used to hide data. It allows for copyright protection on media files using the secret message as a digital watermark. The other main uses for Image Steganography is for the transportation of top secret or high level documents and files between international governments. Image Steganography can be used to send viruses and Trojans by hackers or terrorists to compromise machines and other organizations.

Keywords: Image Steganography, Security and Privacy.

INTRODUCTION

The rapid development of modern technologies that's help high-speed communication and along with the innovation of new technologies security threats are also increasing rapidly. Secured communication is the primary requirement of all international communication to personal communication. One of the techniques which only secures the content of messages like cryptography increases the chance of compromising the security of communication, but Steganography is the process to hide the fact that communication is taking place. It is the simplest way to encode the communicating messages because if attacker is unaware about the communication and the chances of attacks are automatically decreased. The term used in Steganography are: stego image, cover image, secure message and steganalysis. Encrypted message is the message which we want to keep secure. Cover image is the carrier image which contains encrypted message. So, the stego image is that cover image which is going to be transferred with an encrypted message. There are some web-based Image Steganography application that's enables secret communication to send message or share data anywhere at any time. It's a platform that independent and more dynamic compare to others. Sender encrypts the message with private key before sending. Receiver decrypts the message with sender's private key. The private which is used for encryption is shared between sender and receiver only. Steganography usually deals with the ways of hiding the communicated data in such a way that it remains confidential. It maintains security between two communicating parties. In image steganography, secrecy is achieved by embedding data into cover image and generating a stego-image (Cover image). There are different types of steganography techniques, each have their strengths and weaknesses. We review the different security and data hiding techniques that are used to implement a steganography such as LSB, ISB and MLSB etc.

TYPES OF STEGANOGRAPHY



1. **Text Steganography**

It consists of hiding information inside the text files. In this method, the secret data is hidden behind every nth letter of every words of text message. Numbers of methods are available for hiding data in text file. These methods are:

i) Format Based Method
ii) Random and Statistical Method
iii) Linguistics Method

Text steganography can be achieved by updating the text formatting or by updating certain characteristics of textual elements (e.g., characters). The goal in the design of coding methods, is to develop alterations that are reliably decodable (even in the presence of noise) yet largely indiscernible to the reader. These criteria, reliable decoding and minimum visible change, are somewhat conflicting. Here by lies the main challenge in designing document marking techniques. The three coding techniques that are proposed, illustrate different approaches. The techniques can be used either separately or jointly. These are following:

a. Line-Shift Coding: This is a method of altering a document by vertically shifting the locations of text lines to encode the document uniquely.

b. Word-Shift Coding: This is a method of altering a document by horizontally shifting the locations of words within text lines to encode the document uniquely.

c. Feature Coding: This method is applied either to a format file or to a bitmap image of a document.

2. Image Steganography

Hiding the data by taking the cover object as image is referred as image steganography. In image steganography pixel intensities are used to hide the data. In digital steganography, images are widely used cover source because there are number of bits presents in digital representation of an image. Some data hiding techniques that are used to implement a steganography such as LSB, ISB, MLSB etc.

3. Audio Steganography

Audio steganography is the method of hiding secret message into audio signal which result slightly alteration of binary sequence of the resulting audio file. There are many methods that are available for audio steganography. We are going to have a brief introduction on some of them. 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. These methods are

a) Low Bit Encoding
b) Phase Coding
c) Spread Spectrum.

4. Video Steganography

This is a technique of hiding any kind of files or data into digital video format. In this type video (combination of pictures) is used as carrier for transmitting the hidden data form sender to receiver. Generally discrete cosine transforms (DCT) alter the values (e.g., 8.667 to 9) which is used to hide the data in each of the images in the video, which is not noticeable by the human eye. H.264, Mp4, MPEG, AVI are the formats used by video steganography.

In all of these methods, the basic principle of steganography is that a secret message is to be embedded in another cover object which may not be of any significance in such a way that the encrypted data would finally display only the cover data. So it cannot be detected easily to be containing hidden information unless proper decryption is used.

LEAST SIGNIFICANT BIT TECHNIQUE

One of the earliest stego-systems to surface were those referred to as Least Significant Bit Substitution techniques, so called because of how the message data m is embedded within cover image c. The term Least Significant Bit (LSB) refers to the smallest bit of a binary sequence. The structure of binary is such that each integer may only be either a 0 or a 1, often thought of as off and on respectively. Starting from the right, the value (if on) denotes a 1. The value to its left (if on) denotes a 2, and so on where the values double each time. Now let us consider the following 8-bit binary sequence:

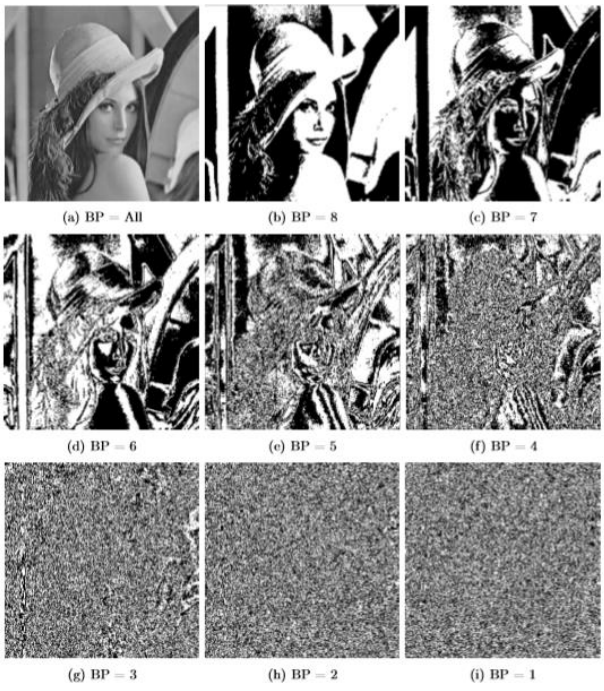
1 0 1 1 0 0 1 1

Adding all the values equal to 1 yields a result of 179. The right-most value is the LSB of this sequence. This value essentially determines whether the total sum is odd or even. If the LSB is a 1, then the total will be an odd number, and if 0, it will be an even number. However, changing the LSB value from a 0 to a 1 does not have a huge impact on the final figure; it will only ever change by +1 at most. If we now think of each 8-bit binary sequence as a means of expressing the colour of a pixel for an image, it should be clear to see that changing the LSB value from a 0 to a 1 will only change the colour by +1 - a change that is unlikely to be noticed with the naked eye. In fact, the LSBs of each pixel value could potentially be modified, and the changes would still not be visible. This highlights a huge amount of redundancy in the image data, and that means we can effectively substitute the LSBs of the image data, with each bit of the message data until the entire message has been embedded. This is what is meant by Least Significant Bit Substitution. Finally, when we talk of Least Significant Bit Substitution algorithms. We should mention that this encompasses two different embedding schemes: sequential and randomised. Sequential embedding often means that the algorithm starts at the first pixel of the cover image c(0,0) and embeds the bits of the message data in order until there is nothing left to embed. Randomised embedding however, scatters the locations of the values that will be modified to contain the bits of the message data. The main reason for randomising the approach is to make things a little trickier for the steganalysts that are looking to determine whether the image is a stegogramme or not.

Overview

In the early stages of image steganography development, many steganographers believed that the least significant bits of an image were an ideal place to embed the secrete data, not only because their modification yields no perceptible loss of quality, but also because they believed the LSBs were completely random in terms of their overall significance to the complete image. In other words, it was common belief that if the LSBs of an image were viewed in isolation as a binary image (where 0 = black, and 1 = white) then the distribution will appear so scattered that modifying the values will make no difference to its appearance - it would still look very random. Figure illustrates why this assumption was made. The Figure shows a grayscale image (a) and allows for a comparison of each of its bit planes (b) → (i), where "BP = 8" corresponds to the most significant bit plane, and "BP = 1" corresponds to the LSB plane.

An image and each of its bit planes (BP) in descending order



We look at each bit plane in this manner, it does appear as though the LSB plane in (i) is more random than that of a bit plane higher up in the scale such as (d), thus it is understandable why such an assumption was made. However, Andreas Westfield and Andreas Pfitzmann found that this

hypothesis was incorrect. Their works suggests that the LSBs - whilst perhaps random in terms of appearance - are no more so than any other bit plane in terms of design. If we consider an image of natural life, the image is almost guarantee to include objects that contain gradual colour changes due to natural filters such as light, shadows, and the texture of the object itself. Typically, a shadow that is cast on an object for example, produces a pattern in the pixel values such that the values in that area of the image decrease by very small amounts as the shadow gets stronger. It is also suggested by Wayner that some cameras pad the data by adding extra detail to produce 24-bit picture, and it is also true that JPEG compression incorporates averaging that may result in large areas of the image having the same LSBs. It therefore seems to be the case that the LSBs of an image are more structured than the steganographers originally believed, and this poses a huge weakness in the authors' systems that can be exploited through visual attacks.

Hide & Seek:

The Sequential Approach is the most common form of image steganography method which is known as Hide & Seek which replaces the LSBs of pixel values with the bits from the message bit stream. This algorithm is simple and very straight forward because it does not require any key to be implemented which makes the things a lot simpler to compute and exchange the secret, it simply means that the security resides solely in the algorithm. If a key were used, then it might become impossible for the user to decode the hidden information, because the key would usually index the manipulated regions of the image. But in the Hide & Seek algorithm, the user needs to understand how the algorithm works, and they will be able to decrypt the message.

Encoding

Algorithm 1 The encoding process of the *Hide & Seek* algorithm in *sequential* mode.

1: **for** $i = 1, ..., l(m)$ **do**
2: $p \leftarrow \text{LSB}(c_i)$
3: **if** $p \neq m_i$ **then**
4: $c_i \leftarrow m_i$
5: **end if**
6: **end for**

The encoding process (as shown by the pseudocode in Algorithm 1) represents that the complete algorithm can be shown by just writing a few lines of code. This algorithm works by considering the first pixel of the image c_i and generating its LSB value (as shown by line 2 of the Algorithm). This is basically obtained by calculating the modulus 2 of the pixel value. This will return 0 if the number is even, and 1 if the number is odd, which tells us the LSB value. We further compare this value with the message bit m_i that we are trying to embed. If they are already the same, then we have nothing to change, but if they are different then we replace c_i with m_i. This process continues till the values in m that needs to be encoded.

Decoding

The decoding process is much simpler than encoding. As the encoder replaced the LSBs of the pixel values in c in sequence, we already know the sequence that should be followed to retrieve the data. Therefore all we need to do is to calculate the modulus 2 of all the pixel values in the stegogramme s, and we are able to reconstruct m as m'_i. Algorithm 2 shows the pseudocode of the decoding process.

Algorithm 2 The decoding process of the *Hide & Seek* algorithm in *sequential* mode.

1: **for** $i = 1, ..., l(s)$ **do**
2: $m'_i \leftarrow \text{LSB}(s_i)$
3: **end for**

Note that this time we run the loop for l(s) instead of l(m). This is because the decoding process is completely separate from the encoding process and so it has no logic of knowing the l(m). If a key were used in this method, it would probably tell this information, but we simply retrieve the LSB value of every pixel. When we convert this to ASCII, the message only be readable up to the certain point that the message was encoded, and will then appear as rubbish when we are reading the LSBs of the image data.

SCOPE OF THE WORK

- The scope of this project is to develop a web application which would help to create a secure communication of data between the sender and receiver with the help of a web domain.
- Steganography is the process of hiding private or sensitive data within the image.
- Steganography involves hiding text so it appears to be a normal image.
- If the user views that object which has hidden information inside, he or she will have no idea that there is any secret information stored inside it.
- Steganography essentially does is to exploit human perception, because human senses are not trained to search for files that have information inside of them.
- Actually this system performs to let the user send information as secret message inside an image file, then user uploads that image and enters the text to send secretly, and also gives a key to lock the text, what this key does is, it encrypts the data, so that even if it is hacked by hacker then also he will not be able to read the text.
- You also need the key to decrypt the text which is hidden by the user.
- The user transfers the image and key to the receiver and the receiver first opens the image, and then he/she enters the key for decryption of text, he/she then press decrypt key to get secret text which is send by the sender.
- With the help of this technique you can ensure that your secret information is sent secretly without any outside interference of hackers or crackers.
- If sender transmits this image in public others will have no idea about what is it, and it will be received by receiver.

LITERATURE REVIEW AND RELATED WORK

Steganography is basically the technique for hiding data within a carrier file such that it is imperceptible for unauthorized users. From this study, it is intended to merge many techniques to produce a totally new technique for colour image steganography to obtain

enhanced efficiency, increased payload capacity, possess integrity check and security with cryptography at the same time. The proposed work supports various different formats as payload. In this proposed method, the codeword is formed with secret data and its CRC-32 checksum, then the codeword is then compressed by Gzip just before encrypting it by AES, and then it's finally added to the encrypted header information for the further process and then implanted into the cover image. Embedding the encrypted data and header information process make use of Fisher-Yates Shuffle algorithm for choosing the next pixel location. To hide one byte of data, different LSB (least significant bits) of all colour channels of the selected pixel is abused. To examine the proposed method, comparative performance tests are performed against the different dimension of image Steganographic techniques by using some of the well-known image quality metrics. For security purpose, histogram, enhanced LSB and Chi-square analyses are performed. The outcome suggest that with the proposed method that it has an improved payload capacity, security and integrity check for common issues of simple LSB method. Moreover, it has been resulted that the proposed method improves the visual quality of the stego image as compared to other studied methods, and makes the secret data hard to be discovered.

Today, with the expansion of information and communication technology, the world, through the digital data, is transforming to the digital world and communications. In meanwhile, the role of internet as a public communication channel is enhancing more and more important in the world of communication every day. In addition to that, maintaining the security and creating the confidential communications are also important regarding the general structure of this communication channel. So, cryptography and information steganography are two important problems in security systems. Both encryption and steganography techniques are not much effective for high security information alone, but when combining these two techniques can greatly enhance the confidentiality and security of confidential information. Recently, the new hybrid algorithm have been developed using cryptography and steganography. However, in these techniques, various attempts have been made to increase the security of censorship by using the random factors and hidden keys, most of these methods are broken by determining the statistical features of the images. In this paper, a high-security hybrid approach is proposed to the digital images steganography based on the Imperialist Competitive Algorithm and Symmetric Cryptography Algorithm. In this proposed method, by selecting the Imperialist Competitive Algorithm, generates a high quality, high-security image. Before the data insertion, symmetric encryption of information takes place, and then encrypted information is placed in the cover image. The outcome of the proposed method shows that in addition to improving the image quality of the steganography, it is much more secured as compared to other methods.

The image files are one of the most commonly used file types today. This paper describes the use of JPEG image files in Steganography. Steganography is the method of hiding a message or information in an image file (cover image) such that it should not be known by users who do not have rights to access. This insertion uses the smallest bit of pixel units in an image file (Least Significant Bit). In this paper, steganography will be combined with vigenere cipher. Steganography make use of the weakness of the human eye in viewing the image file, steganography also utilizes the mathematical calculations in inserting messages into the image file. This type of insertion basically uses the binary of the ASCII code of a character.

EXISTING SYSTEMS

Today, there are numerous Steganographic software tools are available on the internet. The basic concept behind these different tools is the same: to develop a steganographic software that can able to hide image or text in another medium. The various existing softwares are S-Tools, VSL, OpenPuff, CryptaPix and Quick Crypto.

- S-Tools software uses images or audio files to hide data. In this software it has an Action Window that displays the user that what steps are being carried out by the software.
- VSL software uses LSB technique to perform steganography and it also uses much advanced encoding techniques such as Karhunen-Loeve Transform technique.
- OpenPuff software make use of carrier chains by splitting the data into carrier chains and then hiding data. Image, Audio, Video and PDF files can be hidden using this software.
- CryptaPix software performs image editing such as rotating, resizing, cropping and removing red eye from images.
- Quick Crypto software includes encryption of files, emails and password. It basically uses AES, Triple DES and Blowfish software.

DRAWBACKS OF EXISTING SYSTEMS

- The main limitation of the above systems are its portability, they are created to serve two users at different system with separate software for encryption and decryption.
- Both end users must have an application program to encrypt or decrypt.
- The system requirements must be fulfilled in order to perform the operation.
- These softwares are basically platform dependent and are capable for encrypting only the few data formats.

PROPOSED SYSTEM

- In this web application, user can encrypt and decrypt the message or content anywhere at any moment.
- For better security, DES cryptography technique has also been utilized in this proposed method. Before applying the Steganography technique, DES cryptography will convert the secret message into secret text to ensure that the two layer security of the message is achieved.
- In this proposed technique, a new Steganography technique is developed to hide large amount of data in image. This method is basically an improvement of LSB method for hiding information in images.

BENEFITS OF STEGANOGRAPHY

1. It provides the security to the contents or messages without even knowing to third party.
2. The number of bits have swapped according to the sender, therefore the third party cannot even guess password.
3. The normal network user cannot guess the image.
4. In steganography anyone can't suspect by seeing the image.
5. Reliable.
6. Easy to use.
7. Easy Maintenance.
8. System have been protected by password authentication.

VARIOUS APPLICATIONS

1. Storing authenticated and secured data.
2. Protection of Data Modifications.
3. It is the access control system for the Digital Content Distribution.
4. E-Commerce.
5. Media.
6. Database Systems.
7. Digital Watermarking

CONCLUSIONS

It is observed that with LSB Substitution Steganographic method, the results generated in data hiding are quite impressive as it make use of the fact that any image could be broken up to individual bit-planes each holding different levels of information. It is to be observed that as discussed earlier, this method is only used for bitmap images because these involves lossless compression techniques. But this technique can also be extended to be utilized for colour images where, bitplane slicing is to be performed individually for the top four bit-planes for each of the RGB of the message image.

It's also important to note that though steganography was once unrecognised, with the various techniques currently used, it is not only easy to detect the presence but also retrieving them is easier. For instance, without the need to use of a software or complex tools for detection, simple techniques to observe if an image file has been manipulated are: 1. Image size: A steganographic image has a huge storage size as compared to a regular image of the same dimensions i.e. if the storage size of the original image would be a few KBs, the steganographic image could be several MBs in size. This changes with the resolution and type of image used. 2. Noise in image: A steganographic image has a noise as compared to a regular image. So, this is the reason why initially little noise is been added to the cover image, such that the steganographic image does not appear too noisy as compared to the original cover image.

REFERENCES

[1] Ramadhan J. Mstafa, Christian Bach, 2013 https://www.researchgate.net/publication/259893801_Information_Hiding_in_Images_Using_Steganography_Techniques

[2] Dr. V. R. Sadasivam, R. Abishake, R. Tamilarasan, 2019 https://www.eleyon.org/journals/index.php/sajet/article/view/264/204

[3] Dr. Ekta Walia, Navdeep and Payal Jain. "An analysis of LSB & DCT based Steganography." Global Journal of Computer Science and Technology, April 2010. https://globaljournals.org/GJCST_Volume10/gjcst_vol10_issue_1_paper8.pdf

[4] Nick Nabavian. "Image steganography" Nov. 28, 2007. http://www1.chapman.edu/~nabav100/ImgStegano/download/ImageSteganography.pdf