Implementation of LSB Steganography and its Evaluation for Various File Formats (LSB, JSTEG)

Vivek Kumar, Sandesh Kumar, Lavalee Singh, Prateek Yadav

MANGALAYATAN UNIVERSITY[1, 2,3,4]
ALIGARH

ABSTRACT:

Steganography is derived from the Greek word steganos which literally means "Covered" and graphy means "Writing", i.e. covered writing. Steganography refers to the science of "invisible" communication. For hiding secret information in various file formats, there exists a large variety of steganographic techniques some are more complex than others and all of them have respective strong and weak points. The Least Significant Bit (LSB) embedding technique suggests that data can be hidden in the least significant bits of the cover image and the human eye would be unable to notice the hidden image in the cover file. This technique can be used for hiding images in 24-Bit, 8- Bit, Gray scale format. This paper explains the LSB Embedding technique and Presents the evaluation for various file formats.

Keywords*:* Steganography, Least Significant Bit (LSB), GIF, PNG, BMP.

1. INTRODUCTION

In the current trends of the world, the technologies have advanced so much that Most of the individuals prefer using the internet as the primary medium to transfer data from one end to another across the world. There are many possible ways to transmit data using the internet: via e-mails, chats, etc. The data transition is made very simple, fast and accurate using the internet. However, one of the main problems with sending data over the internet is the security threat it poses i.e. the personal or confidential data can be stolen or hacked in many ways. Therefore it becomes very important to take data security into consideration, as it is one of the most essential factors that need attention during the process of data transferring.

Data security basically means protection of data from unauthorized users or hackers and providing high security to prevent data modification. This area of data security has gained more attention over the recent period of time due to the massive increase in data transfer rate over the internet.

In order to improve the security features in data transfers over the internet, many techniques have been developed like: Cryptography, Steganography and digital watermarking. While Cryptography is a method to conceal information by encrypting it to cipher texts and transmitting it to the intended receiver using an unknown key, Steganography provides further security by hiding the cipher text into a seemingly invisible image or other formats.

2.1. IMAGE STEGANOGRAPHY

Image compression techniques are extensively used in steganography. Among the two types of image compressions, lossy compression and loss less compression; lossless compression formats offer more promises. Lossy compression may not maintain the original image's integrity. Lossless compression maintains the original image data exactly, hence it is prefered. Example of Lossy compression format is JPEG format files. Examples of Lossless compression formats are GIF[3] and BMP formats.

We have used an 8-bit image size for implementation of our steganography. Improvement in stegnographic techniques is make it possible to apply the Detecting LSB Steganography in Colour and Gray- Scale Images which were confined to gray scale images in the initial stages The difficulty in colour images control is solved later on in many techniques such as the analysis of the variation of the gradient energy. The secret message embedded in the target image is detected in both gray and colour images, and the length of the embedded message is estimated [5, 6].

2.3 HIDING METHODS IN IMAGE STEGANOGRAPHY

In Image Steganography, There are a variety of methods using which information can be hidden in images.

Least Significant Bit Replacement Technique: In image steganography almost all data hiding techniques try to alter insignificant information in the cover image. Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image. For instance, a simple scheme proposed, is to place the embedding data at the least significant bit (LSB) of each pixel in the cover

image[7,8,9] . The altered image is called stego-image. Altering LSB doesn't change the quality of image to human perception but this scheme is sensitive a variety of image processing attacks like compression, cropping etc. We will be emphasizing more on this technique for the various image formats.

2.4 MODERATE SIGNIFICANT BIT REPLACEMENT TECHNIQUE:

The moderate significant bits of each pixel in the cover image can be used to embed the secret message. This method improves sensitivity to modification, but it degrades the quality of stego-image.

Experiments have shown that the length of hidden messages embedded in the least significant bits of signal samples can be estimated with relatively high precision.

2.5 LEAST SIGNIFICANT BIT SUBSTITUTION

LSB (Least Significant Bit) substitution is the process of adjusting the least significant bit pixels of the carrier image. It is a simple approach for embedding message into the image. The Least Significant Bit insertion varies according to number of bits in an image. For an 8 bit image, the least significant bit i.e., the 8th bit of each byte of the image is changed to the bit of secret message. For 24 bit image, the colours of each component like RGB (red, green and blue) are changed. LSB is effective in using BMP images since the compression in BMP is lossless. But for hiding the secret message inside an image of BMP file using LSB algorithm it requires a large image which is used as a cover.

Consider an 8-bit grayscale bitmap image where each pixel is stored as a byte representing a grayscale value. Suppose the first eight pixels of the original image have the following grayscale values

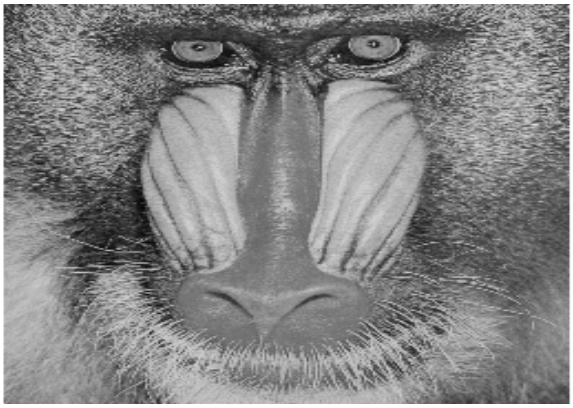
11010010
01001010
10010111
10001100
00010101
01010111
00100110
01000011

To hide the letter C whose binary value is 10000011, we would replace the LSBs of these pixels to have the following new grayscale values:

11010011
01001010
10010110
10001100
00010100

01010110
00100111
01000011

Note that, on average, only half the LSBs need to change. The difference between the cover (i.e. original) image and the stego image will be hardly noticeable to the human eye. Figure 2.5 that show a cover image and a stego image (with data is embedded); there is no visible difference between the two images.



Above, replaces the LSBs of data values to match bits of the message. It can equally alter the data value by a small amount ensuring the a legal range of data values is preserved. The difference being that the choice of whether to add or subtract one from the cover image pixel is random**.** This will have the same effect as LSB replacement in terms of not being able to perceive the existence of the hidden message. This steganographic technique is called LSB matching. Both LSB replacement and LSB matching leave the LSB unchanged if the message bit matches the LSB. When the message bit does not match the LSB, LSB replacement replaces the LSB with the message bit; LSB matching randomly increments or decrements the data value by one. LSB matching is also known as ±1 embedding.

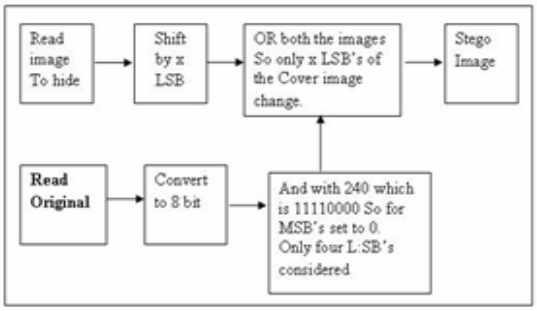


Fig.1 Block Diagram for implemented Logic of LSB embedding.

2.6 DESIGN DETAILS

This section focuses on algorithms of LSB Steganography and Steganalysis[10]

A. ALGORITHM FOR HIDING (STEGANOGRAPHY)

1. Read the original image and the image which is to be hidden in the original image
2. Shift the image to hide in the cover image by X bits.
3. And the original image or cover image.
4. The shifted hidden image and the result of step 3 are bitored. This makes changes only in the X LSB bits so that the image is hidden in the original image.

In MATLAB we convert it to unit8 format. This image can be called as the stego image

B. ALGORITHM FOR STEGANALYSIS

1. The stego image is bit shifted by 4 bits since it was shifted by 4 bits to insert it into the original image.
2. The image is the ANDED with 255 i.e., 11111111, which gives the original image. It is ANDED with 255 because initially all the LSB's were made 0. Now it is recovered back.
3. To get it to Unit8 format we, convert it back to unit8 which is the extracted.

C. LSB in BMP

The BMP file format also called bitmap or DIB file format (for *device-independent bitmap*), is an image file format used to store bitmap digital images. Since BMP is not widely used the suspicion might arise, if it is transmitted with an LSB stego. When image are used as the carrier in Steganography they are generally manipulated by changing one or more of the bits of the byte or bytes that make up the pixels of an image. The message can be stored in the LSB of one colour of the RGB value or in the parity bit of the entire RGB value. A BMP is capable of hiding quite a large message. LSB in BMP is most suitable for applications, where the focus is on the amount of information to be transmitted and not on the secrecy of that information. If more number of bits is altered, it may result in a larger possibility that the altered bits can be seen with the human eye. But with the LSB the main objective of Steganography is to pass a message to a receiver without an intruder even knowing that a message is being passed is being achieved.

D. LSB in PNG

Portable Network Graphics (PNG) is a bitmapped image format that employs lossless data compression. PNG was created to improve upon and replace GIF. Since PNG is widely used the suspicion might not arise if it is transmitted with an LSB stego. When images are used as the carrier in

Steganography they are generally manipulated by changing one or more of the bits of the byte or bytes that make up the pixels of an image. The message can be stored in the LSB of one colour of the RGB value or in the parity bit of the entire RGB value .A PNG is capable of hiding quite a large message. LSB in PNG is most suitable for applications where the focus is on the amount of information to be transmitted and not on the secrecy of that information. If more number of bits is altered it may result in a larger possibility that the altered bits can be seen with the human eye.

But with the LSB the main objective of steganography Is to pass a messege to a receiver without an intruder even knowing that a message is being passed is being achieved.

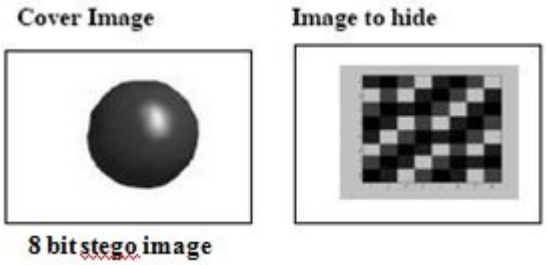
C. LSB in GIF

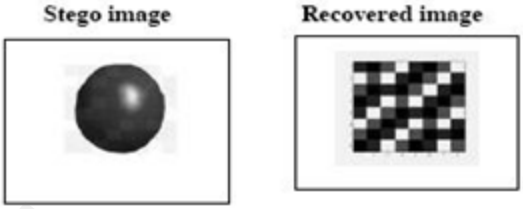
Graphics interchange format also known as GIF is one of the machine independent compressed formats for storing images. Since GIF images only have a bit depth of 8, amount of information that can be hidden is less than with BMP. Embedding information in GIF images using LSB results in almost same results as those of using LSB with BMP.LSB in GIF is very efficient algorithm to use when a reasonable amount of data in a Grayscale.

2.7 EXPERIMENTED RESULTS

Following experimental results highlights on 8 bit LSB Steganography.

A. Results for Png image





B. Results for .bmp file

2.8 EVALUATION OF IMAGE QUALITY

For comparing stego image with cover results requires a measure of image quality, commonly used measures are Mean-Squared Error, Peak Signal-to-Noise Ratio[3] and histogram.

A. MEAN-SQUARED ERROR

The mean-squared error (MSE) between

two images I1(m,n) and I2(m,n) is:

$$MSE = \frac{\sum_{M,N} [I_1(m,n) - I_2(m,n)]^2}{M * N}$$

M and *N* are the number of rows and columns in the input images, respectively. Mean-squared error depends strongly on the image intensity scaling. A mean-squared error of 100.0 for an 8-bit image (with pixel values in the range 0-255) looks dreadful; but a MSE of 100.0 for a 10- bit image (pixel values in [0,1023]) is barely noticeable.

B. PEAK SIGNAL-TO-NOISE RATIO

Peak Signal-to-Noise Ratio (PSNR) avoids this problem by scaling the MSE according to the image range

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right)$$

PSNR is measured in decibels (dB). PSNR is a good measure for comparing restoration results for the same image, but between-image comparisons of PSNR are meaningless. MSE and PSNR values for each file format is shown in table 1

	Cover image	Stego image	Cover-Stego image
MSE	224.948	244.162	69.826
PSNR	24.6100	24.2540	29.690

2.9 EVULATION OF DIFFERENT TECHNIQUES

There are different algorithms available.

One should select best available algorithm

For given applications. Following characteristics are to be evaluated while selecting a particular file format steganography. Steganography says that secret message is to be hidden and result in distortion less image.

The distortion must not be visible to the human eye. The amount of data embedded in the image also plays an important role. The algorithm decides how much amount of data could be embedded in the image resulting in a distortion less image. Steganalysis is the technique of detecting the hidden information in the image. The algorithm for Steganography must be such that the steganalysis algorithms should fail. i.e the Steganography algorithms must not be prune to attacks on steganalysis. During communication the intruder could check the original image to remove the hidden information.. He/she may manipulate the image. This manipulation may include cropping or rotation etc of the images. The manipulations done may cause the image distortion. steganographic algorithms chosen must be such that it overcomes such manipulation and the steganographic data reaches the destination in the required format.

Table 2: Comparison of LSB technique for various file formats

	LSB In BMP	LSB in GIF	LSB In PNG
Percentage Distortion less resultant image	High	Medium	High
Invisibility	High	Medium	Medium
Steganalysis detection	Low	Low	Low
Image manipulation	Low	Low	Low
Amount of embedded data	High	Medium	Medium
Payload capacity	High	Medium	Medium
Independent of file format	Low	Low	High

2.10 CONCLUSION

Since BMP uses lossless compression, LSB secret message inside a BMP file, one would require a very large cover image. BMP images of 800×600 pixels found to have this accepted as valid.

References

[1] Vaishali S. Jabade, Dr. Sachin R.Gengaje, ‘‘Literature Review of Wavelet Based Digital Image Watermarking Techniques,’’ *Volume 31– No.1, October 2011*.

[2] Mr. Manjunatha Prasad. R, Dr. Shivaprakash Koliwad “A Comprehensive Survey of Contemporary Researches in Watermarking for Copyright Protection of Digital Images”, *International Journal of Computer Science and Network Security (IJCSNS), Vol.9 No.4, April 2009*, pp.91-102.

[3] Vidyasagar M. Potdar, Song Han, Elizabeth Chang, “A Survey of Digital Image Watermarking Techniques”, *3rd IEEE InternationalConference on Industrial Informatics (INDIN)*, 2005, pp.709-713.

[4] B.N. Chatterji, Manesh Kokare, A. Adhipathi Reddy, Rajib Kumar Jha., “Wavelets for Content Based Image Retrieval and Digital Watermarking for MultimediaApplications”,*IEEETransactions ,Vol.2*, 2003, pp.812-816.

[5] F. Kumgollu, A Bouridane, M A Roula and S Boussaktd, “Comparison of Different Wavelet Transforms for Fusion Based Watermarking Applications”, *IEEE Transactions, Vol.3,2003*, pp. 1188-1191.

[6] F.Mintzer, et al., “Effective and Ineffective Digital Watermarks”, IEEE International Conference on Image Processing, ICIP-97, 1997, Vol.3, pp. 9-12.

[7] R. G. Van Schyndel,“A Digital Watermark”, Proc. IEEE International Conference onImage Processing, ICIP-94, 1994, Vol.2, pp.86-90.

[8] Feng Wen-ge, Liu Lei, “SVD and DWT Zero-bit Watermarking Algorithm”, *2nd IEEE InternationaAsia Conference on Informatics in Control, Automation and Robotics*, 2010, pp. 361-364.

[9] Wei Cao, Yixin Yan, Shengming Li, “Robust Image Watermarking Based on Singula Value Decomposition in DT-CWT Domain”, *IEEE International Workshop on Imaging Systems and Techniques, IST*, 2009.

[10] Ben Wang, Jinkou Ding, Qiaoyan Wen, Xin Liao, Cuixiang Liu, “An Image Watermarking Algorithm Based on DWT, DCT and SVD”, *Proceedings of IC-NIDC*, 2009,1034-1038.

[11] Quan Liu, Xuemei Jiang, “Design and Realization of a Meaningful Digital

Watermarking Algorithm Based on RBF Neural Network" *IEEE, 2005* pp. 214-218.

[12] Chuan-Yu Chang, Hung-Jen Wang, Sheng-Jyun Su, "Copyright Authentication For Images With A Full Counter-Propagation Neural Network" Science Direct, Expert Systems with Applications 37,2010,pp. 7639–7647

[13] Yuanhai Shao, Wei Chen, Chan Liu, "Multiwavelet based Digital Watermarking with Support Vector Machine Technique", *IEEE, 2008*, pp. 4557-4561.

[14] Chin-Shiuh Shieha, Hsiang-Cheh Huangb, Feng-Hsing Wangc, Jeng-Shyang Pana, "Genetic watermarking based on transform-domain techniques", *Science Direct Pattern Recognition 37, 2004*,pp. 555–565.

ABOUT AUTHORS

Sandesh Kumar received the B. Tech degree in Electronics and Communication Engineering from U.P.T.U in 2009 and M. Tech degree in Electronics and Communication Engineering from Mangalayatan University in 2013. Currently Working as Assistant Professor in R.B Group of Institutions India. Research interests include Image Processing and Wireless Communication.

Lovely Singh received the B. Tech degree in C.S from U.P.T.U in 2007 and M. Tech U.P.T.U in 2011. Currently Working as Assistant Professor in, R.B Group of Institutions, India. Research interests include Image Processing and Web Mining.

Vivek Kumar received the B. Tech degree in Electronics and Communication Engineering from Mangalayatan University in 2011 and M. Tech degree in Electronics and Communication Engineering from Mangalayatan University in 2013. Research interests include Image

Processing and Wireless Communication.