Hiding Data in Images Using New Random Technique

Obaida Mohammad Awad Al-Hazaimeh

**Department of Information Technology, AL-BALQA Applied University/
Al-Huson University College, Irbid, Al-Huson, 50, Jordan**

Abstract

Steganography is the art of hiding the fact that communication is taking place by hiding information in other information. In the field of Data Communication, Steganography play a major role. The transmission of information via the Internet may expose it to detect and theft. Some solution to be discussed is how to passing information in a manner that the very existence of the message is unknown in order to repel attention of the potential attacker. We focus on the Least Significant Bit (LSB) technique which is the most common Steganographic technique is employed in this paper. An improvement to this technique is suggested by randomly inserting the bits of the message in the image to produce more secured system. In this paper, the security goals were enhanced via a proposed cryptosystems to maintain the security on the Cover-image. The proposed solution consists of a simple, but strong to hiding the text data and the human eye would be unable to notice the hidden data in the Stego-image.

Keywords: *LSB technique , Steganography, data hiding, Image Quality.*

1. Introduction

Since the rise of the Internet one of the most important factors of information technology and communication has been the security of information. Cryptography was created as a technique for securing the secrecy of communication and many different methods have been developed to encrypt and decrypt data in order to keep the message secret. Unfortunately it is sometimes not enough to keep the contents of a message secret, it may also be necessary to keep the existence of the message secret. The technique used to implement this, is called Steganography [1, 3].

Steganography is the art and science of invisible communication as we mentioned in abstract part. This is accomplished through hiding information in other information, thus hiding the existence of the communicated information. The word Steganography is derived from the Greek words “STEGO” meaning “cover” and “GRAFIA” meaning “writing” [1, 2] defining it as “covered writing”. In image Steganography the information is hidden exclusively in images. Today Steganography is mostly used on computers with digital data being the carriers and networks being the high speed delivery channels.

Steganography differs from Cryptography in the sense that where cryptography focuses on keeping the contents of a message secret, Steganography focuses on keeping the existence of a message secret [4]. Steganography and Cryptography are both ways to protect information from unwanted parties but neither technology alone is perfect and can be compromised. Once the presence of hidden information is revealed or even suspected, the purpose of Steganography is partly defeated [4, 6]. The strength of Steganography can thus be amplified by combining it with cryptography [5].

1.1 Different kinds of Steganography

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object’s use and display [6]. The redundant bits of an object are those bits that can be altered without the alteration being detected easily [7]. Image and audio files especially comply with this requirement, while research has also uncovered other file formats that can be used for information hiding. Figure 1 shows the four main categories of file formats that can be used for steganography.

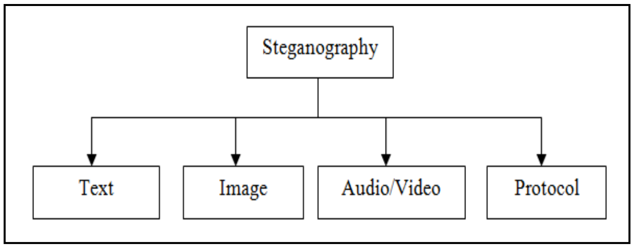

Figure 1: Categories of Steganography

Hiding information in text is historically the most important method of steganography. An obvious method was to hide a secret message in every *n*th letter of every word of a text message. It is only since the beginning of the Internet and all the different digital file formats that is has decreased in importance [2, 5]. Text steganography using digital files is not used very often since text files have a very small amount of redundant data [9].

Given the proliferation of digital images, especially on the Internet, and given the large amount of redundant bits present in the digital representation of an image, images are the most popular cover objects for Steganography [9-10]. This paper will focus on hiding information in images in the next sections.

To hide information in audio files similar techniques are used as for image files. One different technique unique to audio steganography is masking, which exploits the properties of the human ear to hide information unnoticeably. A faint, but audible, sound becomes inaudible in the presence of another louder audible sound [9]. This property creates a channel in which to hide information. Although nearly equal to images in steganographic potential, the larger size of meaningful audio files makes them less popular to use than images [10].

Protocol steganography refers to the technique of embedding information within messages and network control protocols used in network transmission [13]. In the layers of the OSI network model there exist covert channels where steganography can be used [5, 9-10].

1.2 Image Steganography

Images are the most popular cover objects used for Steganography. In the domain of digital images many different image file formats exist, most of them for specific applications. For these different image file formats, different steganographic algorithms exist. Figure 2 shows Categories of image Steganography [15].

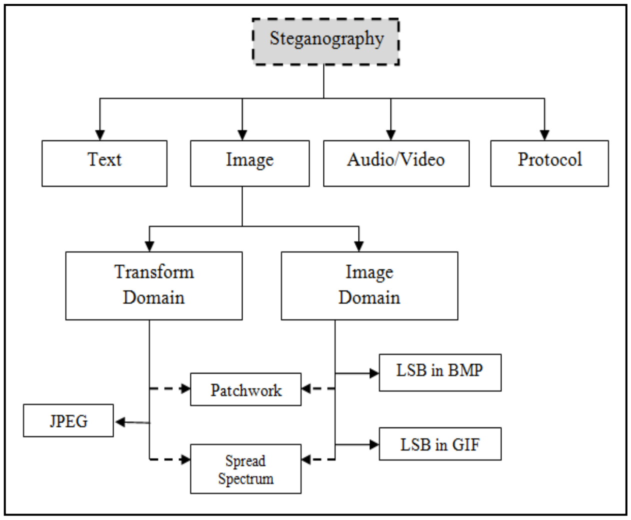

Figure 2: Categories of image Steganography

1.3 LSB Technique

The least significant bit i.e. the eighth bit inside an image is changed to a bit of the secret message. When using a 24-bit image, one can store 3 bits in each pixel by changing a bit of each of the red, green and blue color components, since they are each represented by a byte. An 800×600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data [1, 4-7].

2. Proposed Technique

Least Significant Bits (LSB) is the simplest and most straight forward approach to embed or hide a message into a cover-image. In this paper, the message is inserted in the images in random manner in the pixels of a cover-image. However, LSB hides the message in a way that the humans do not distinguish it, and still possible for the opponent to retrieve the message due to the simplicity of the technique. Malicious people can easily try to extract the message from the beginning of the image if they are doubtful that there exists secret information that was inserted in the image. Therefore, there is a need to enhance the LSB. In this paper, new technique is proposed to improve the LSB scheme by inserting the message bit into a set of random in each pixel within the image, not in the least significant bit, and the least significant bit just a sign to extract data from the image.

It is proposed in this paper that the inserting of message bits into the image is not only in the least bit but also the other bits in the pixel in the random manner. This can be done by comparing the message bit to the pixel bit randomly chosen from second to the last bit. Based on this comparison, 1 is inserted in the least significant bit if the message bit identical to that of the image, whereas, 0 is inserted if the message bit didn't match with the chosen bit from the image as shown in Table 1.

Pixel Bits	Message Bits	Comparison	Result
1	1	Match	1
0	0	Match	1
1	0	Not match	0
0	1	Not match	0

Table 1: Message Bits Comparison

The proposed Technique consists of two algorithms i.e. Steganography and Steganalysis. The following sections will explain each process involved in detail.

2.1 Algorithm for Hiding (Steganography)

The proposed Steganography algorithm consists of the following processes as shown in Figure 3.

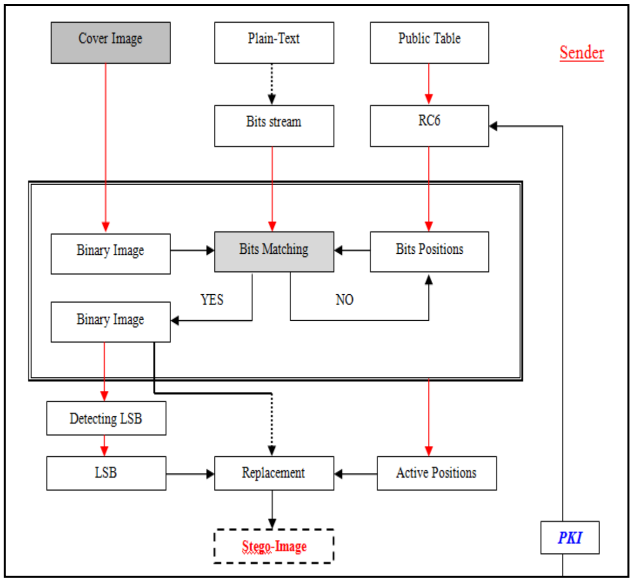

Figure 3: Proposed Steganography architecture

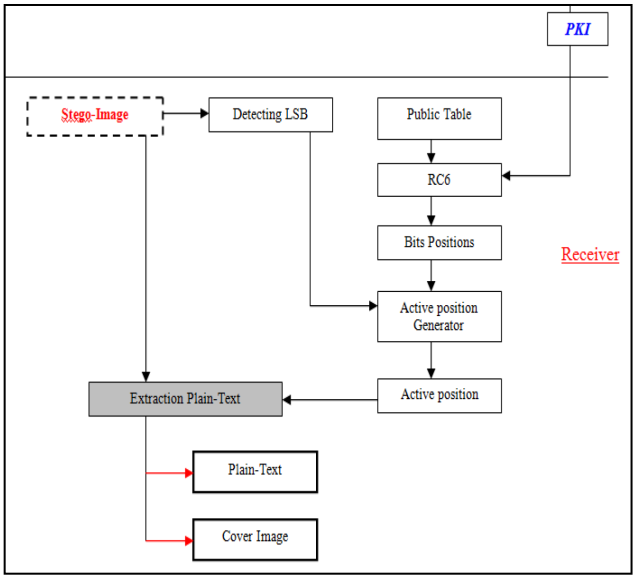

Figure 4: Proposed Steganalysis architecture

A. Public Table: 24 rows by 24 columns in size. Table entry values can be between 0 and F. Since the table is not secret, it can be announced to the public.

B. RC6: Based on Feistel rounds in an attempt to frustrate the attackers to generate private table according to the PKI secret value.

C. Bits Positions: Positions are generated from the values after performed RC6 rounds.

D. Plain-Text: Source data (Information to be hiding in the cover image).

E. Bits Stream: Binary source data streaming.

F. Cover image: Image to be sent (Original image).

G. Binary Image: Convert the original image to the Binary code.

H. Bits Matching: Based on, 1 is inserted in the least significant bit if the message bit identical to that of the image, whereas, 0 is inserted if the message bit didn't match with the chosen bit from the image (Random Manner).

I. Active Position: Sign to the positions are matched between the message bits and the chosen bits from the image.

J. Stego-image: Final look of image (image with the inserted data).

2.2 Algorithm for Steganalysis

The proposed Steganalysis algorithm consists of the following processes as shown in Figure 4.

Steganalysis is a process of reversing all that has happened in the Steganography process. It involves to extracting the hiding data from the Stego-image for the receiver's understanding. The same process is performed at the beginning of the Steganography and Steganalysis process (connection established) as described in the previous part at the sender side to generate the same private table at the receiver side.

As shown in Figure 4, the proposed Steganalysis algorithm consists of the following processes:

1. Detecting LSB: Detect and extract the LSB to generate the active positions according to the bits position table since the bits positions is made available at both the sender and receiver side (connection established).

2. Extraction Process: Extract the original bits of data and the end result of such operation is the plain text data (original text).

For smooth Steganalysis process to be achieved, the accuracy of the Steganalysis bits positions cannot be negotiated. In short, the accuracy of this algorithm is a function of the active positions generator.

3. Experimented and Analytical Results

In this paper, we have used MATLAB for simulating dynamic systems and the analysis and visualization of experimental data. Therefore, JPG image with 400x500 in size was selected (Cover-Image), and a message of 1 KB to be hides as shown in Figure 5.

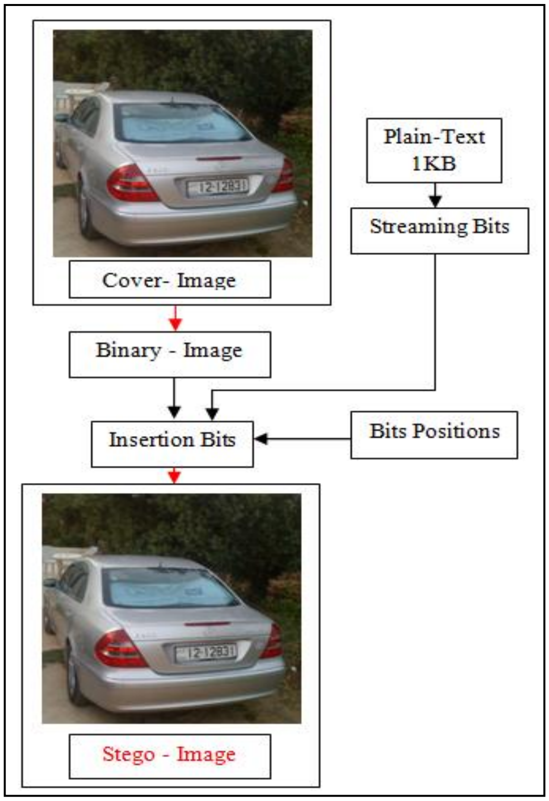

Figure 5: Insertion text process with the image

The complexity in the insertion text process with the image, the attacker cannot retrieve the value of the original text without knowing the bits positions and PKI value to configured random numbers using RC6 algorithm, and Method of insertion processes.

4. 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, 18-19] and Histogram Difference Measure.

1. *Mean-Squared Error (MSE)*

The Mean-Squared Error (MSE) between two images *IMG1(m,n)* and *IMG2*(m,n), the following equation is used to calculate MSE.

$$MSN = \frac{\sum_{M,N}[IMG1(m,n) - IMG2(m,n)]^{2}}{M * N}$$

Where, 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. MSE 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.

2. *Peak Signal-to-Noise Ratio(PSNR)*

Peak Signal-to-Noise Ratio (PSNR), applied to images as a quality metric by scaling the MSE according to the image range, the following equation is used to calculate PSNR [19].

$$PSNR = log_{10} \frac{R^2}{MES}$$

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 Cover-image and Stego-image are shown in Table 2.

Metric	Cover-Image	Stego-Image
MSE	220.944	240.158
PSNR	20.6096	20.2536

Table2: Evaluation of Image Quality

3. *Histogram Difference Measure*

A histogram is an aggregation method that conveys data distribution. To construct a histogram, the data space is partitioned into many small ranges, with each range corresponding to a bin. The height of a histogram bin is determined by the percentage of data points that fall in the corresponding range. It reveals the data density within each sub-range. [15, 19]. Figure 6 and 7 shows the histogram for both Cover-image and Stego-image respectively.

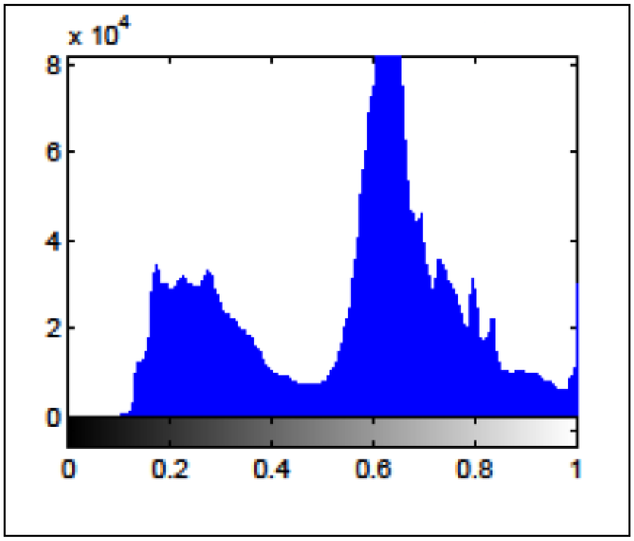

Figure 6: Histogram for Cover-Image

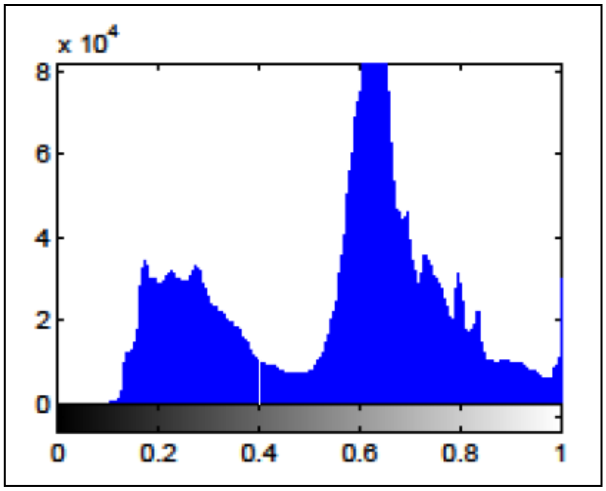

Figure 7: Histogram for Stego-Image

5. Conclusion

In this paper, new Steganographic systems are proposed to enhance/add the security of Steganographic system using LSB approach to provide a means of secure communication. In our proposed approach, the message bits are inserted randomly into the Cover-image pixels instead of sequentially. It is proposed in this enhancement that the inserting of message bits into the image is not only in the least bit but also the other bits in the pixel in the random manner. Thus, we expect that the proposed technique will be efficiently used in Steganographic systems or considered as a good alternative to other technique because of the high level of security.

References

[1] Wang, H & Wang, S, “Cyber warfare: Steganography vs. Steganalysis”, *Communications of the ACM, 47:10*, October 2004.

[2] Anderson, R.J. & Petitcolas, F.A.P., “On the limits of steganography”, *IEEE Journal of selected Areas in Communications*, May 1998.

[3] Dunbar, B., “Steganographic techniques and their use in an Open-Systems environment”, *SANS Institute*, January 2002.

[4] Artz, D., “Digital Steganography: Hiding Data within Data”, *IEEE Internet Computing Journal*, June 2001.

[5] Handel, T. & Sandford, M., “Hiding data in the OSI network model”, *Proceedings of the 1st International Workshop on Information Hiding*, June 1996.

[6] Ahsan, K. & Kundur, D., “Practical Data hiding in TCP/IP”, *Proceedings of the Workshop on Multimedia Security at ACM Multimedia*, 2002.

[7] Johnson, N.F. & Jajodia, S., “Exploring Steganography: Seeing the Unseen”, *Computer Journal*, February 1998.

[8] Owens, M., “A discussion of covert channels and steganography”, *SANS Institute*, 2002.

[9] Johnson, N.F. & Jajodia, S., “Steganalysis of Images Created Using Current Steganography Software”, *Proceedings of the 2nd Information Hiding Workshop*, April 1998.

[10] Venkatraman, S., Abraham, A. & Paprzycki, M., “Significance of Steganography on Data Security”, *Proceedings of the International Conference on Information Technology: Coding and Computing*, 2004.

[11] Lee, Y.K. & Chen, L.H., “High capacity image steganographic model”, *Visual Image Signal Processing*, 147:03, June 2000.

[12] Provos, N. & Honeyman, P., “Hide and Seek: An introduction to steganography”, *IEEE Security and Privacy Journal*, 2003.

[13] Petitcolas, F.A.P., Anderson, R.J. & Kuhn, M.G., “Information Hiding – A survey”, *Proceedings of the IEEE*, 87:07, July 1999.

[14] Bender, W., Gruhl, D., Morimoto, N. & Lu, A., “Techniques for data hiding”, *IBM Systems Journal*, Vol 35, 1996.

[15] N. Jacobsen, K. Solanki, U. Madhow, B. S. Manjunath a, and S. Chandrasekaran, “Image-adaptive high-volume data hiding based on scalar quantization,” *in Proceedings of IEEE Military Communications Conference (MILCOM), Anaheim, CA, USA*, October 2002.

[16] N. Provos and P. Honeyman, “Detecting steganographic content on the internet,” ISOC NDSS’02, San Diego, CA, 2002. [Available at] http://www.outguess.org/.

[17] A. A. Rwabutaza, "A Cryptanalysis Methodology for the Reverse Engineering of Encrypted Information in Images," *Wright State University*, 2009.

[18] R. G. van Schyndel, et al., "A digital watermark," in Image Processing, 1994. *Proceedings. ICIP-94., IEEE International Conference, pp. 86-90 vol.2*. 1994.

[19] F. A. P. Peticolas, et al., "Information hiding–a survey," *Proceedings of the IEEE, vol. 87, pp. 1062-1078*, 1999.

Author

Obaida Mohammad Awad Al-Hazaimeh received the B.S. degree in Computer Science from Applied Science University (ASU), Jordan in 2004, the MSc in Computer Science/ Distributed system from University Science Malaysia (USM), 2005, and PhD in Computer Science/ Network security (Cryptography) for Real-Time Application from University Utara Malaysia (UUM), 2010.