

Information Hiding using LSB technique

Jassim Mohammed Ahmed[†] and Zulkarnain Md Ali^{††}

School of Computer Science,
Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia,
43600 Bangi, Selangor Darul Ehsan,
Malaysia

Summary

In the field of Data Communication, security-issues have the top priority. 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 here. An improvement to this technique is suggested by randomly embedding the bits of the message in the image to produce more secured system. The proposed system goal is giving more the complexity to cryptosystems and the execution time does not differ great than the original methods then the hide the messages encrypted inside image in a way that does not allow any Attacker to even detect that there is secret message. As a result, the proposed system is more efficient compared to that introduced earlier.

Keywords:

Steganography, LSB technique .

1. Introduction

Due to advances in ICT, most of the information is kept electronically. Consequently, the security of information has become a fundamental issue to provide confidentiality and protecting the copyright for digital media such as audio, video, and images. Therefore, the steganography is applied to hide some information in digital media. Whereby the message is embedded in digital media. In this paper, we proposed the Secure Information Hiding System (SIHS) that is based on Least Significant Bit (LSB) technique in hiding messages in an image. The system enhanced the LSB technique by randomly dispersing the bits of the message in the image and thus making it harder for unauthorized people to extract the original message. Discrete logarithm calculation technique is used for determining the location of the bit into pixel to embed the message. The proposed algorithm provides a stag-key that will be used during the embedding and extracting of the message.

The growing possibilities of modern communications require the use of secure means of protecting information during transmission against unauthorized access and use.

The Steganography is an approach in information hiding whereby the information is hidden inconspicuously inside a host data set such that its presence is imperceptible [1]

It does not alter the structure of the secret message, but instead hides it inside a cover-image so that it cannot be seen. A message in a ciphertext, for instance, might arouse suspicion on the part of the recipient while an “invisible” message created with steganographic methods will not. In other word, steganography prevents an unintended recipient from suspecting that the secret message exists

In this paper, we present an LSB technique, which randomly select the bit into pixels of the cover-object that is used to hide the secret message. The selection is based on the discrete logarithm.

2. Overview Steganography

The word steganography means covered writing. According to many authors, the definition of the steganographt is the art of science of hiding information to be undetectable and make the communication happening secretly [2] Recently, many ways to send coded messages is developed to hide the message in different application specially media such as imagery, audio, or video. [3], [4].

The history of the stenography is packing to long time. It has been used during times of war. For example, during the American Revolution, General Sir Henry Clinton in the British army had composed a letter to General John Burgoyne and amask or grille was applied to the letter. The second was appeared in (stories of spies and letters) in this type of messages, the apparent message and the actual message is made sense to the reader.

Johnson and Jajodia 1998 have defined the mean goal of the steganography is to communicate secretly in a complete undetectable manner and avoid drawing suspicion to the transmission of the hidden [3]. In another word, it is not to keep others from knowing the hidden information, but in fact it is to keep others from thinking that the information even exists. And according to Proves and Honeyman 2001 that if the steganography

method causes someone to suspect the carrier medium, then the method is absolutely failed. There has been a rapid growth of interest in steganography for two main reasons [5]:

- (i) The publishing and broadcasting industries have become interested in techniques for hiding encrypted copyright marks and serial numbers in digital films, audio recordings, books and multimedia products.
- (ii) Moves by various governments to restrict the availability of encryption services have motivated people to study methods by which private messages can be embedded in seemingly innocuous cover messages.

The methods of hiding data into media cover such as audio, images, and video, were developed about more than a decade ago [6]. And until recently, information hiding techniques received very much less attention from the research community and from industry than cryptography. However, this situation is changing rapidly and the first conference on this topic was organized in 1996. [7] suggested that there were two reasons behind the that rapid growth of interest in steganography, the first reason can be concluded that the publishing and broadcasting industries have become interested in techniques for hiding encrypted copyright marks and serial numbers in digital films, audio recordings, books and multimedia products, and the second reason is Moves by various governments to restrict the availability of encryption services have motivated people to study methods by which private messages can be embedded in seemingly innocuous cover messages. [7]:

The basic model of steganography is consisting of three main objects carrier, message, and password. Figure 1 [8] illustrates the basic model of steganography. The first object which is the carrier is also known as the cover object. The Message can be defined as the data that the sender intends to remain confidential. It can be as plain text, ciphertext, other image or anything that can be embedded in a bit stream such as copyright mark or serial number. The third objects that consist of the steganography are the password, and it's also known as stego-key, which is ensured that the only the official receiver which knows the corresponding decoding key will be able extract the message from the cover-object. The cover-object with the secretly embedded message is then called the stego-object. The cover object and the corresponding key are required in order to recover a message from the stego-object. Sometimes the original image also required in order to extract the message.

Ming et al. 2006 has made a small survey that classifies current steganography tools, according to his survey, there are many information hiding methods. Some of these are substitution methods, transform methods, and other miscellaneous methods.

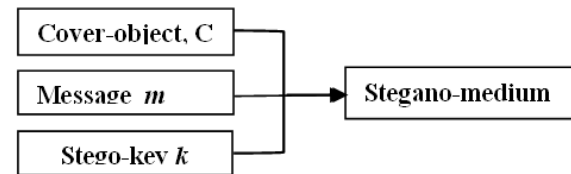


Fig. 1 A Steganography Process (Information Hiding).

There are many information hiding methods. Some of these are substitution methods, transform methods, and other miscellaneous methods. Substitution methods are based on the modification of least significant bits of the cover image using the secret data and some key based random permutations. Transform methods are based on modification or rearrangement of transforming domain (discrete cosine, Fourier, wavelet) coefficients with secret data and some set of rules about the coefficients. Other miscellaneous methods use techniques such as fractals, matrix decomposition and predictive quantitation. Some well-known steganology methodologies are the Least Significant Bit (LSB) Hiding, Regional Hiding with Segmentation (RHS) and SCAN Hiding [9]. is a small survey that classifies current steganography tools.

3. Difference Between Steganography And Cryptography

An important point to note is that both steganography and cryptography provide secure communications and may be used concurrently. Steganography and cryptography differ in execution. In cryptography, the secret message which is the transmitted file itself cannot be recovered without the secret key; however, the encrypted file is identified as being sent. It helps to protect confidentiality but protection vanishes after decryption. In steganography the existence of the stego message is concealed in a cover file in a way that does not allow an enemy to observe that there is a message present [5]. The stego message can be extracted with stego key as long as the stego file is identified by which embedding method is used.

4. Steganography Applications

There are many steganographic applications for digital image, including copyright protection, feature tagging, and secret communication[1]. Copyright notice or watermark can be embedded inside an image to identify it as intellectual property. If someone attempts to use this image without permission, we can prove by extracting the watermark.

In feature tagging, captions, annotations, time stamps, and other descriptive elements can be embedded inside an image. Copying the sfegoimage also copies the embedded features and only parties who possess the decoding stego-

key will be able to extract and view the features. On the other hand, secret communication does not advertise a covert communication by using steganography. Therefore, it can avoid scrutiny of the sender, message and recipient. This is effective only if the hidden communication is not detected by other people.

5. Steganography In Variety Of Covers

As mentioned before, almost all digital file formats can be used for hiding data (steganography). However, the format that is more suitable is those with a high degree of redundancy. Currie 1998 has defined the redundancy as the bits of an object that provide accuracy far greater than necessary for the object's use and display [10]. And Anderson has defined the redundant bits of an object as those bits that can be altered without the alteration being detected easily. And these definitions can work specially with image and audio files, while, other format files that can be used for hiding information are uncovered by researchers yet [11]. Figure 2 illustrates the four main categories of file format that can be used for steganography.

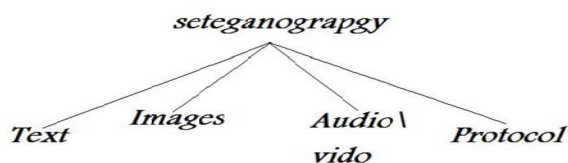


Fig. 2 Type Of The Covers

Nowadays, hiding information inside image is the most popular technique. An image with the secret message can easily be spread over the World Wide Web or in Newsgroups. A German steganographic expert (Niels Provos) had created a scanning cluster which detected the presence of hidden messages inside images that were posted on the net. However, after checking one million images, no hidden information was found. Therefore, Krenn 2004 said that the uses of the steganography are seeming to be limited until now. [12].

There are many different ways to hide information in images. It can be by encoding every bit of information in the image or selectively embed the message in noisy areas in order to draw less attention. The nature color variation considers a great deal to find the noisy areas. Also, the message may be scattered randomly throughout the image. The most common approaches exist to hide information in images are:

- least significant bit insertion
- masking and filtering
- algorithms and transformations

the discussion is only on **Least-Significant Bit (LSB)**. Here offer a short explanation of the LSB Information Hiding technique. The LSB is the technique intended for hiding information in the least significant bits of another information medium for some small n . It is based on the idea that replacing a pixel intensity by modifying the last n the least significant bits for some small n will not create enough intensity change to make a naked eye detect the change. A small value of n is usually 1 or 2. For example, if $n=1$ and a pixel has an intensity of $16 = 00010000_2$, changing the least significant bit of a pixel intensity will only modify the intensity by at most $2^n - 1 = 1$. This means that the given pixel can have the same value of $16 = 00010000_2$ or $17 = 00010001_2$ the change of which is hard to detect by naked eye [13].

As mentioned before, this technique is the most widely used to hide data, Wang H. & Wang S. defined the disadvantage of this method according to the risk of detection, which is happened in the process of hiding data within the image. Therefore, the beginning of series detects storage is much possible to know the rest of the string. The reason behind the popularity of this method according to their explanation is due to the relative easiness to implement the LSB [14].

In order to hide a secret message inside an image, a proper cover is needed. However, this method uses bits of each pixel in the image, when using a 24 bit color image, and in this method a bit of each of the red and green and blue color component can be used. Therefore, a total of three bits can be stored in each pixel,

Thus, a 800×600 pixel image can contain a total amount of 1,440,000 bits (180,000 bytes) of secret data. For example, the following grid can be considered as 3 pixels of a 24 bit color image, using 9 bytes of memory Figure 3 [15].

Anderson 1998 has developed the LSB to be a slightly more secure system by specifying the only certain pixel that has been changed and its known by the sender and the receiver. Therefore, there is no way finding out which pixel is used to hide the data by any intruder [11]. In order to increase the security of the system, either the distribution of the bits can be randomly within the image.

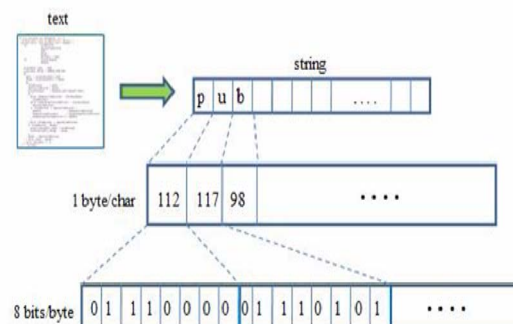


Fig. 3 shows the installation of the message that hidden in the image.

6. Safe Information Hiding System (SIHS)

LSB is the simplest and most straight forward approach to embed or hide a message into a cover-image. The message is embedded with the sequence-mapping techniques 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 embedded in the image.

Therefore, there is a need to enhance the LSB. For this purpose, a system named Safe Information Hiding System (SIHS) is proposed to improve the LSB scheme. SIHS overcame the sequence-mapping problem by embedding 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, not necessarily with the same bit value message. The bits of the secret message are embedded in the pixels of the cover-image that are generated by discrete logarithm calculation.

As the LSB (Least Significant Bit) technique used to hide the messages in images. However, it is decided to enhance the security system by introducing a new technique comprise of randomly dispersing the message bits in images.

It is proposed in this enhancement that the embedding 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 achieved 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, see figure 4 .

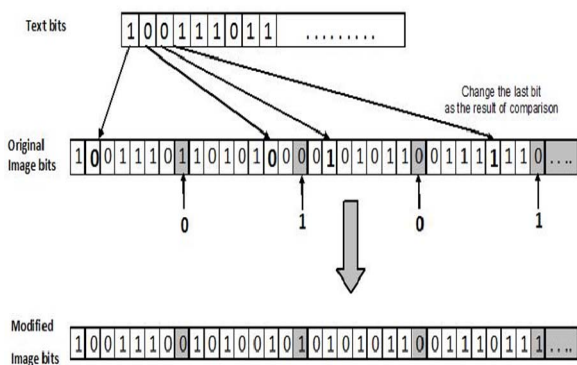


Fig. 4 Embed the bits in image

It is interesting to note that the value of the least significant bit of the image is not always representing the actual value of that from the message, (figure 5) , as expressed in the following table(1):

LSB of modified image (0101) not equal text bits of message (1001)

Fig. 5 Comparison between the message with the image bits

Table 1: Change in the Bits during the process of embed

Random bit (2-8)	Bit of message	Last bit
0	0	1
0	1	0
1	0	0
1	1	1

The attacker can easily extract the message hidden in an image using the LSB technique by identifying the least the significant bit. However, this could not be done with the above described method. Hence, the latter considered an improvement to the common LSB, see figure 6.

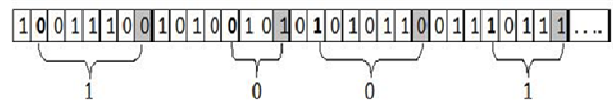


Fig. .6 Extract bits from image

The process of extracting the message from the image includes inverse comparison to that used in embedding. If the least significant bit is 1, then the actual value of the message bit is equal to that compare with image bit value, see figure 3.3, while if the least bit is zero then the message bit is representing the inverse value of the image bits that used in embedding comparison, as expressed in the following table 2.

Table 2 : Change in the Bits during the process of Extract from image

Random bit (2-8)	Last bit	Value extracted
0	0	1
0	1	0
1	0	0
1	1	1

7. Discrete Logarithm

Discrete logarithm calculation can be used to solve the sequence-mapping problem. The main idea here is to generate a series of random numbers of length equal to the message length that ranging from 2 to 8. These series numbers will be use in random-mapping.

We defined discrete logarithm to produce random numbers. These numbers depend on the value of key (k). The values are computed from the following equation, and these numbers will be limited to the length of the message, M:

$$x_t = a * x_{t-1} \bmod p \dots \dots (1)$$

Where,

x_0 = is the sum of K digits.

$$a = 3x_0$$

$$p = k$$

$$t = 1, 2, 3, 4, \dots, m.$$

$$x_t = a * x_{t-1} \bmod p \dots \dots (1)$$

The numbers created from the above equation is then used to generate another numbers ranging from 2 to 8. The latter are used to locate the image bit (in the pixel) that will be used in the comparison with the message bit, as expressed as follow:

$$p_i = (x_i \bmod 7) + 2 \dots \dots (2)$$

The process of stenography by the proposed system SIHS is illustrated in (figure 7).

Recovering a message from a stag-image demands the corresponding decoding key, k that used during the encoding process. Hence, both the sender and receiver have to share the stag-key during the communication. The k key is then used for selecting the positions of the pixel where the secret bits had been embedded. For further clarification for the function of the proposed system, the following is the example is presented:

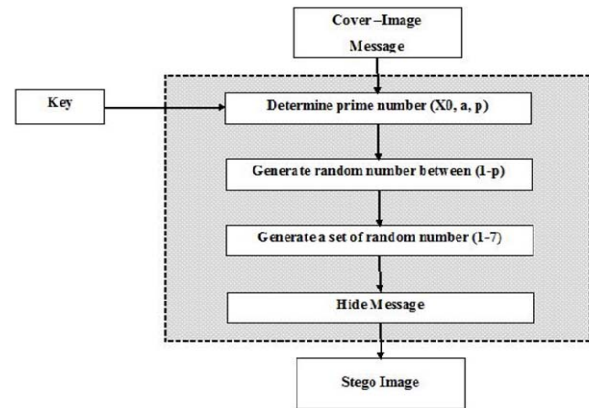


Fig. 7 Process of Stenography

8. Example

To explain the work of the proposed system in the process of generating random numbers, then is the process include the text within the image depending on the figures generated previously, are given this example, suppose the secret key (k) and the message to (m). The sender and the recipient opting key (k), let (1321), let m=(10100011) the length of the message (m) (8), from K, extract value from the x_0 and a,(see Figure 1)

$$x_0 = 1 + 3 + 2 + 1 = 7$$

$$a = 3x_0 = 3 * 7 = 21$$

$$p = k, t = m$$

$$x_t = x_{t-1} * a \bmod p$$

$$x_t = x_{t-1} * a \bmod p$$

$$x_1 = 7 * 21 \bmod 1321 = 147$$

$$x_2 = 7 * 21 \bmod 1321 = 445$$

$$x_3 = 7 * 21 \bmod 1321 = 98$$

$$x_4 = 7 * 21 \bmod 1321 = 737$$

$$x_5 = 7 * 21 \bmod 1321 = 946$$

$$x_6 = 7 * 21 \bmod 1321 = 51$$

$$x_7 = 7 * 21 \bmod 1321 = 1071$$

$$x_8 = 7 * 21 \bmod 1321 = 34$$

$$p_i = (x_i \bmod 7) + 2 \dots \dots (2)$$

$$p_1 = (147 \bmod 7) + 2 = 2$$

$$p_2 = (445 \bmod 7) + 2 = 6$$

$$p_3 = (98 \bmod 7) + 2 = 2$$

$$p_4 = (737 \bmod 7) + 2 = 4$$

$$p_5 = (946 \bmod 7) + 2 = 3$$

$$p_6 = (51 \bmod 7) + 2 = 4$$

$$p_7 = (1071 \bmod 7) + 2 = 2$$

$$p_8 = (34 \bmod 7) + 2 = 8$$

To apply this numbers are in the process of concealment, Section is taken from an image. First eight pixels of the image (Figure 8)

1	0	1	1	0	1	1	0	1	0	1	0	1	0	0
1	1	1	0	0	1	0	0	1	0	0	1	0	0	1
1	0	1	1	0	0	1	1	1	0	1	0	1	1	0
1	1	1	0	0	1	1	0	0	0	1	1	0	1	0

Fig.8 Section of the Image Before the Embed

The extracting of the message from the image is conducted from the following:

1. x_p, a, p are extract from (K):

x_p is the sum of K digits, $a = 3x_p$, $p = k^{\#}$

2. Equation (1) is used to generate random numbers along the hidden message.
3. The numbers are then inserted into equation 2, (see figure 9), these numbers locates the bit within the pixel that compared with the message bit.
4. The process of extracting the message from the image includes inverse comparison to that used in embedding:
 - I. If the least significant bit is 1, then the actual value of the message bit is equal to that compare with image bit value.
 - II. If the least bit is 0 then the message bit is corresponding to the inverse value of the image bits that used in embedding comparison continue this process to extract the entire message..

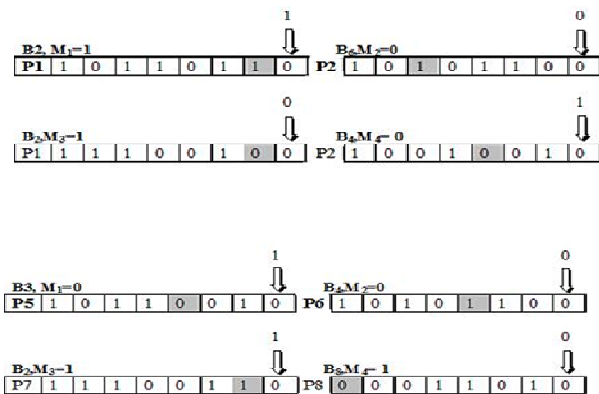


Fig. 9 .A Explain The Embed The Bits In Image

1	0	1	1	0	1	1	1	1	0	1	0	1	1	0
1	1	1	0	0	1	0	0	1	0	0	1	0	0	1
1	0	1	1	0	0	1	1	1	0	1	0	1	1	0
1	1	1	0	0	1	1	1	0	0	0	1	1	0	1

Fig. 9 B Section Of The Image After The Embed

The process of extracting the message from the image includes inverse comparison to that used in embedding. If the least significant bit is 1, then the actual value of the message bit is equal to that compare with image bit value, see figure 3.3, while if the least bit is zero then the message bit is representing the inverse value of the image bits that used in embedding comparison, as expressed in the following table(2):

9. Analysis Of SIHS

The objective of this system, (SIHS), is to solve the LSB's sequence-mapping problem. SIHS process can be summarized in three phases. The evaluation of those phases is carried out using the human eyes. The evaluation process of the proposed system is illustrated below.

In figure 10, a JPG colored image has been embedded with the 1KB message, and secret-key (5372). As shown in (figure 11), the stag-image is identical to that of cover image, and it is impossible for the human eyes to differentiate between the aforesaid pictures. It is believed that the size of the message is very tiny compared to the image size, hence it is difficult to produce considerable changes that may affect the latter's appearance.

For the system analysis, we presented three cases. In all cases, the testing are done through the normal viewing of the human eyes. As mentioned before, this system has been developed to overcome a sequence-mapping problem when using LSB. A JPG image with 400x500 in size, a message of 1 KB and Secret-key as shown in Figure 10 respectively, have been chosen to test the technique.

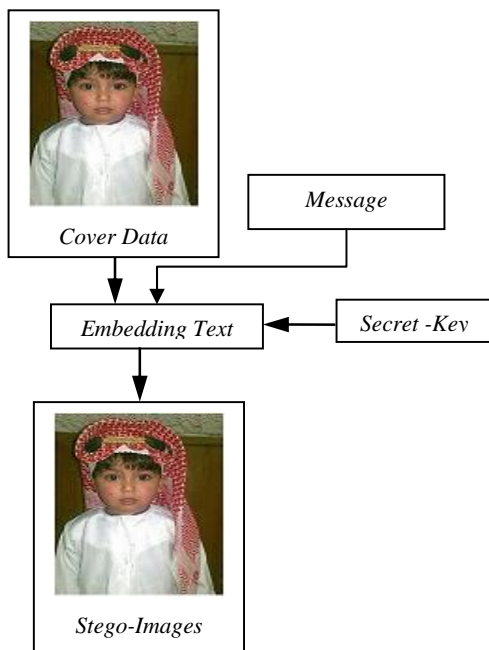


Fig. 10 The process of embedding text within the image

In the first case we used a color image as shown in Figure 11. With a Secret -Key of 5372, we embedded the message into the cover-image and the resulted stego-image is as shown in Figure 12. From normal eyes perception, the result of the stego-image looks identical to the cover-image. This is because there is a little changes of the pixel values and thus no significant difference. The sequence of mapping is subject to risk detection. When the beginning of the series is know, the rest of string can be detected easily. To overcome this weakness, it is proposed to use the discrete logarithm calculation. In this technique, the key, K, is embedded randomly in the message bits, hence, it is difficult for the third party to locate which bit is used to store K. As a result, the message looks to the third party as a nonsense symbols, see figure 12.



Fig. 11 Cover-Image



Fig. 12 Stego-Images.

Process to hide data within the image sequence exposes them to risk of detection So that if the beginning of series detect storage as possible to know the rest of the string. However by using discrete logarithm calculation, the problem of sequence-mapping can be solved. In this technique, the selected pixel for embedding the message bits depends on the random number generated by the SIHS and a key, k. Although the third party could determine where the message bits are embedded, he has a difficulty to recover it because the message bits are embedded in a random order. The recovered message will be a nonsense symbols as shown in Figure 13

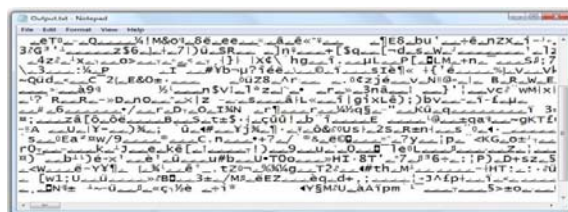


Fig. 13 Recovered Messages with Normal Extraction

Steganography that uses a key has a better security than non-key steganography. This is so because without the knowledge of the valid key, it is difficult for a third party or attackers to recover the embedded the message, bits per pixel inside. Consequently, the values of a least significant bit of the image do not represent the actual value of that from the message. Hence, the SIHS represents the more secure communication way compared to original LSB.

This complexity in the process of embedding, the attacker cannot retrieve the value of the original text without knowing the key and an equation configured random numbers, and Method of embedding.

In this way, the system was strengthened using LSB approach to provide a means of secure communication. A Secret -key has been applied to the system during

embedment of the message into the cover-image. In the proposed approach, the message bits are embedded randomly into the cover-image pixels instead of sequentially.

10. Conclusion

In this paper we have presented an enhancement of the image steganographic system using LSB approach to provide a means of secure communication. A stego-key has been applied to the system during embedment of the message into the cover-image. In our proposed approach, the message bits are embedded randomly into the cover-image pixels instead of sequentially.

It is proposed in this enhancement that the embedding 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 achieved by comparing the message bit to the pixel bit randomly chosen from second to the last bit. It is interesting to note that the value of the least significant bit of the image is not always representing the actual value of that from the message.

References

- [1] M. M. Amin, et al., "Information hiding using steganography," in Telecommunication Technology, 2003. NCTT 2003 Proceedings. 4th National Conference on, 2003, pp. 21-25.
- [2] M. Ramkumar and A. N. Akansu, "Some design issues for robust data hiding systems," in Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on, 1999, pp. 1528-1532 vol.2.
- [3] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," 1998, pp. 273-289.
- [4] R. Popa, "An analysis of steganographic techniques," The Politehnica University of Timisoara, Faculty of Automatics and Computers, Department of Computer Science and Software Engineering, 1998.
- [5] F. A. P. Petitcolas, et al., "Information hiding—a survey," Proceedings of the IEEE, vol. 87, pp. 1062-1078, 1999.
- [6] R. G. van Schyndel, et al., "A digital watermark," in Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference, 1994, pp. 86-90 vol.2.
- [7] F. A. P. Petitcolas, et al., "Information hiding—a survey," Proceedings of the IEEE, vol. 87, pp. 1062-1078, 2002.
- [8] C. Cachin, "An information-theoretic model for steganography," 1998, pp. 306-318.
- [9] C. Ming, et al., "Analysis of Current Steganography Tools: Classifications & Features," in Intelligent Information Hiding and Multimedia Signal Processing, 2006. IIH-MSP '06. International Conference on, 2006, pp. 384-387.
- [10] I. Currie, et al., Surmounting the effects of lossy compression on steganography: 19th On Surmounting Systems Security Conference, 1996.
- [11] R. J. Anderson and F. A. P. Petitcolas, "On the limits of steganography," Selected Areas in Communications, IEEE Journal on, vol. 16, pp. 474-481, 1998.
- [12] R. Krenn, "Steganography and steganalysis," Retrieved September, vol. 8, p. 2007, 2004.
- [13] A. A. Rwabutaza, "A Cryptanalysis Methodology for the Reverse Engineering of Encrypted Information in Images," Wright State University, 2009.
- [14] H. Wang and S. Wang, "Cyber warfare: steganography vs. steganalysis," Communications of the ACM, vol. 47, pp. 76-82, 2004.
- [15] N. F. Johnson and S. Jajodia, "Steganalysis: the investigation of hidden information," in Information Technology Conference, 1998. IEEE, 1998, pp. 113-116.



Jassim Mohammed Ahmed received the B.S. in Computer Science Department at AL MAMON University College in (2002)-Iraq, he is doing M.S in Computer Science at Universiti Kebangsaan Malaysia (UKM). He is interesting in the following Fields (Cryptology, Information Security and Information Hiding)



Zulkarnain MD Ali, He got on B.Sc. from Universiti Teknologi Malaysia, Malaysia, in 1994. M.Sc. from University of Loughborough, UK, in 1997. PhD from Universiti Putra Malaysia (UPM), Malaysia, in 2010. He is lecturer in Universiti Kebangsaan Malaysia (UKM). His research areas are cryptography, programming and parallel computing