Message Segmentation to Enhance the Security of LSB Image Steganography

Dr. Mohammed Abbas Fadhil Al-Husainy

Department of Multimedia Systems,
Faculty of Sciences and Information Technology,
Al-Zaytoonah University of Jordan.
Amman, Jordan

***Abstract*—Classic Least Significant Bit (LSB) steganography technique is the most used technique to hide secret information in the least significant bit of the pixels in the stego-image. This paper proposed a technique by splitting the secret message into set of segments, that have same length (number of characters), and find the best LSBs of pixels in the stego-image that are matched to each segment. The main goal of this technique is to minimize the number of LSBs that are changed when substituting them with the bits of characters in the secret message. This will lead to decrease the distortion (noise) that is occurred in the pixels of the stego-image and as result increase the immunity of the stego-image against the visual attack. The experiment shows that the proposed technique gives good enhancement to the Classic Least Significant Bit (LSB) technique.**

Keywords—Security; Distortion; Embedding; Substitution.

I. INTRODUCTION

Steganography is one of many techniques that are used to hide secret information to prevent any attackers to make damage in this information or use it in illegal form. Steganography can be defined as the technique used to embed data or other secret information inside some other object commonly referred to as cover, by changing its properties. The purpose of steganography is to set up a secret communication path between two parties such that any person in the middle cannot detect its existence; the attacker should not gain any information about the embedded data by simply looking at cover file or stego file. Steganography is the art of hiding information in ways that prevent the detection of hidden messages. Steganography, derived from Greek, literally means "covered writing." It includes a vast array of secret communications methods that conceal the message's very existence. These methods include invisible inks, microdots, character arrangement, digital signatures, covert channels, and spread spectrum [1, 2].

The basic model of steganography uses a cover object (any object that can be used to hold secret information inside), the secret message (the secret information that is to be sent to some remote place secretly), a stego key that is used to encode the secret message to make its detection difficult and a steganography algorithm/technique (the procedure to hide secret message inside cover object). The outcome of the process is the stego object which is the object that has the secret message hidden inside. This stego object is sent to the receiver where receiver will get the secret data out from the stego image by applying decoding algorithm/technique [1].

Recently, steganography is implemented by using digital media. Secret message is embedded inside digital cover media like text, images, audio, video or protocols depending upon the requirement and choice of the sender. Compared with the other types of steganography, the image steganography is most widely used. The reason behind the popularity of image steganography is the large amount of redundant information present in the images that can be easily altered to hide secret messages inside them, and because it can take advantage of the limited power of the human visual system (HVS). With the continued growth of strong graphics power in computer and the research being put into image based steganography, this field will continue to grow at a very rapid pace [1, 3, 4, 5].

Steganography has a wide range of applications. The major application of steganography is for secret data communication. Covert channels in TCP/IP involve masking identification information in the TCP/IP headers to hide the true identity of one or more systems. Cryptography is also used for the same purpose but steganography is more widely used technique as it hides the existence of secret data. Another application of steganography is feature tagging. Captions, annotations, time stamps, and other descriptive elements can be embedded inside an image, such as the names of individuals in a photo or locations in a map [1, 2, 5, 6].

Steganography can be also used to combine explanatory information with an image (like doctor's notes accompanying an X-ray). Steganography is used by some modern printers, including HP and Xerox brand color laser printers. Tiny yellow dots are added to each page. The dots are barely visible and contain encoded printer serial numbers, as well as date and time stamps. The application list of image steganography is very long [1, 6].

The Steganography technique is the perfect supplement for encryption that allows a user to hide large amounts of information within an image. Thus, it is often used in conjunction with cryptography so that the information is doubly protected; first it is encrypted and then hidden so that an adversary has to first find the hidden information before decryption take place [7, 8, 9, 10, 26(11)]. The problem with cryptography is that the encrypted message is obvious. This means that anyone who observes an encrypted message in

transit can reasonably assume that the sender of the message does not want it to be read by casual observers. This makes it possible to deduce the valuable information. Thus, if the sensitive information will be transmitted over unsecured channel such as the internet, steganography technique can be used to provide an additional protection on a secret message [2].

A good technique of image steganography aims at three aspects. First one is capacity (the maximum data that can be stored inside cover image). Second one is the imperceptibility (the visual quality of stego-image after data hiding) and the last is robustness [7].

The idea in this paper is that when substitute the LSB of the pixels (in the stego-image) with the bits of all characters (in the secret message) as one segment, this will result a large number of changes that are happen in LSB of pixels. It is normally come from the truth that it is rarely find a best match between very long sequence of bits of all characters in the secret message and the LSB of the pixels in the stego-image. A message segmentation LSB was proposed in this paper to overcome this problem by splitting the secret message into set of segments of same length (same number of characters). And try to find the best match between the bits of the characters in each segment and the LSB of different sequences of pixels in the stego-image. When the proposed technique split the long secret message into number of small segments, this will lead to increase the probability of finding best matching between the bits of the characters in the secret message and the LSB of the pixels in the stego-image. The best match between bits will decrease the number of LSB of the pixels that are changed when replace the bits of characters in the secret message in it. As a result of that, the distortion/noise that will appear in the pixels of the stego-image will be decrease and the immunity of the stego-image against the attack by human visual system (HVS) becomes strong.

II. Related Works

When hiding information inside images usually Least Significant Bit (LSB) method is used. In the LSB method the 8th bit of every byte of the carrier file is substituted by one bit of every bit of the secret information [12]. The LSB method usually does not increase the file size, but depending on the size of the information that is to be hidden inside the file, the file can become noticeably distorted.

Ross J. Anderson and Fabien A.P. Petitcolas argued that every steganographic approach will have its limitations; they proposed an information theoretic approach using Shannon's theory for perfect secrecy [13]. In the methods that are proposed by H. Motameni and his colleague's one can embed at the dark corners of an image [14]. One can also embed the secret information in frequency domain by using Discrete Wavelet Transform method [15]. In this method the embedding should be done at high frequency coefficients. P. Mohan Kumar and D. Roopa suggested that one can apply block matching procedure to search the highest similarity block for each block of the secret image and embed in LSBs of the cover image [16]. Mohammed A.F. AlHusainy employed different strategy in image steganography art by mapping the pixels of image to English letters and special characters [17]. Lisa M Marvel and Charles G Boncelet proposed to hide at the inherent noise places [18]. Ran-Zan Wang and Yeh-shun Chen also did the two way block matching for image in image steganography [19]. But this approach is suspicious to the hackers. Xinpeng Zhang and his colleagues proposed an approach called "multibit assignment steganography for palette images", in which each gregarious color that possesses close neighboring color in the palette is exploited to represent several secret bits [20]. In reference [21] authors have discussed a double substitution algorithm for encrypting at sender and decrypting at receiver and the embedding process was at 7th and 8th bit positions alternatively. In [22] an image steganography with palette based images is suggested. The method is based on a palette modification scheme, which can iteratively embed one message bit into each pixel in a palette based image. In each iteration, both the cost of removing an entry color in a palette and the benefit of generating a new one to replace it are calculated. If the maximal benefit exceeds the minimal cost, an entry color is replaced. It is found that the fundamental statistics of natural images are altered by the hidden non-natural information [23]. But if we do not touch the bytes those carry the image features and embed in the other bytes then the problem can be solved. As LSB embedding is very common, many steganalysis tools are available for it [24]. So LSB embedding is no more secured now-a-days. So, new embedding techniques are to be welcomed to the steganographic world. Due to the large number of steganographic tools available over the internet, a particular threat exists when criminals use steganography to conceal their activities with in digital images in cyber space. Reference [25] presents two JPEG steganographic methods using Quantization Index Modulation (QIM) in the Discrete Cosine Transform (DCT) domain. The two methods approximately preserve the histogram of quantized DCT coefficients, aiming at secure steganography against histogram-based attacks. Sukhpreet Kaur and Sumeet Kaur in [26] developed a technique for hiding text using image steganography that use 7 bits per pixel as a full capacity of the cover image to hide data and still no visual changes in the stego image.

III. Classic-Lsb Image Steganography Technique

The Least Significant Bit (LSB) steganography technique works by representing each character (byte) of the secret message as a set of 8-bits (where 1 byte ≡ 8 bits). And then hide/replace the bits of the characters in the least significant bit of the pixels in the stego-image. If the secret message has n characters, then LSB technique need at least (n*8) pixels in the stego-image to hid the bits of the n characters.

By substitute the LSB of each pixel in the stego-image with one bit (from the 8-bits) of each character in the secret message, the substitution operation will cause some distortion/noise in the stego-image. By using Human Visual System (HVS), the attackers may doubt that the stego-image contain a secret information in it. In general, whenever the length of the secret message (number of characters) is long, then the noise in the stego-image probably will increase as a result. This will make restriction to hide a very long message in a small stego-image. Therefore, we will tend to choose a short message to hide it in a large stego-image to minimize the noise that is happen in the pixels of the stego-image and to put aside

the doubt about containing the stego-image any secret information.

Also, when an attacker success know that the stego-image contains a secret message, it is easy to get this message by recompose the secret message from the LSB of the pixels in the stego-image.

IV. The Proposed LSB Image Steganography Technique

The message segmentation LSB technique is suggested here to enhance the performance of the Classic-LSB technique by supporting it through three strong points:

- Decrease the distortion/noise that will be appearing in the pixels of the stego-image.
- Increase the capability of hiding very long secret message in a small stego-image.
- Increase the immunity of the stego-image against the attacks of Human Visual System (HVS).

In the following paragraphs, the detail explanation of the operations that are doing in the proposed technique will be given. Two definitions used in this technique for secret message and stego-image are listed below:

A secret message is an English message might be contains alphabetic letters (‘a’…’z’) or numbers (‘0’…’9’) or any special symbols like: ('space character', '.', ',', '(', ')').

A stego-image, for the purpose of testing, a candidate image to be used in this work is a bitmap images (.bmp) type. In general, each file of type (.bmp) is consisting of a header part which is containing much information like (Width and Height of the image, number Palette, number of bits for each pixel) followed by the data of the bitmap image pixels. The pixels of each image represent as a two dimensional list, but the proposed technique treat the pixels of the image as a one dimensional list of bytes, (where each byte has a value between (0…255)), by reading the bytes of the two dimensional image row by row and stores them as a one dimensional list.

Before listing the steps of the algorithm that describe the operations of the proposed technique, some data structures used in the algorithm are defined below:

1) MessageB: is a list that contains a binary representation (bits) of all characters in the secret message. The number of elements (size) of this list is (n*8), where n is the number of characters in the secret message.

2) ImageB: is a list of the Least Significant Bit (LSB) of all pixels in the stego-image. The number of elements (size) of this list is (m), where m is the size of the image and its equal (Width × Height × Palette).

3) SegmentLength: is a positive integer number between (2 … (n*8)/2) which represents the length of each segment (number of bits) in the SegmentList.

4) SegmentsList: is a list of segments that is created from the MessageB by splitting it to k segments, where k = (n*8) / SegmentLength. And each segment has number of bits equal SegmentLength.

5) *SegmentIndex:* is a list of indices, each index represents the first index of a sequence of bits in ImageB which have a best match with the bits for one of the segments in SegmentsList. We must note that there is no overlapping between the matched bits sequences in this technique.

Algorithm:

// Hiding Operation

Step1: Calculate the ***TotalSize*** (in byte) that is required to store:

(1) Length of secret message (number of character)

(2) ***SegmentLength***

(3) Size of ***SegmentList***

Step2: Store the bits representation of the above three information in the Least Significant Bit (LSB) at the start of the ***ImageB*** list (from **bit #1** to **bit #(*TotalSize**8)**).

Step3: For ***i* = 1** To **((*n**8) / *SegmentLength*)**

{

For ***j* = ((*TotalSize**8)+1)** To ***m***

{

***x* = 1**

***BestMatch* = 0**

***BestIndex* = -1**

For ***w* = *j*** To **(*j*+ *SegmentLength*)**

{

Find the number of matched bits ***MBits*** in ***Segment*[*i*][*x*]** with the bits of ***ImageB*[*w*]**

***x* = *x*+1**

}

If (***MBits*** > ***BestMatch***)

{

BestMatch* = *MBits

BestIndex* = *j

}

}

SegmentIndex*[*i*] = *BestIndex

Substitute the bits of ***Segment*[*i*]** instead of the bits in ***ImageB*** starting at ***BestIndex***

}

// Extracting Operation

Step1: Read from the stego-image the information that is stored in the first **(*TotalSize**8)** LSB of the pixels.

Step2: Reconstruct the segments of the secret message by using the extracted information in Step1.

Step3: Reassembling the all the segments that are constructed in Step2 to regenerate the characters of the secret message.

V. Experimental Results And Discussion

The performance of the proposed Segmented-LSB image steganography technique has been tested by using both the Classic-LSB and the proposed LSB to hide some messages in different (.bmp) images and record the results to enable the reader to makes a good comparison, in the performance, between these two techniques. Table I shows the stego-images, of different sizes, that are used in the experiments. Table II summarizes the recorded results from the experiments using SegmentLength = 10.

To clarify the effect of SegmentLength on the performance of the proposed Segmented-LSB image steganography technique, different selected values of the SegmentLength used on the above experiments. Fig. 1 shows the effect of the SegmentLength on (a) Number of LSB changed, (b) Signal to Noise Ratio (SNR) of the stego-image, (c) Time of hiding operation.

The required programs to implement the Classic-LSB and the proposed LSB techniques written by using C++ programming language and executing them on a computer system are of 2.53GHz processor with 4.0 GB memory and Microsoft Windows 7 operating system.

From Table I, we note that the proposed LSB decrease the number of LSB that are changed in the stego-image when it compares with the Classic-LSB, this certainly enhance the SNR of the stego-image. The main challenge of the proposed LSB is the time of the hiding operation, but this comes from performing the exhaustive search to find the best matching between the bits in each segment with all non-overlapped bits sequences in the ImageB list.

When we see the three parts of Fig. 1, we can note that the proposed LSB produce a stable performance when the SegmentLength change:

- When increase the SegmentLength the number of LSB that are changed will increase and vice versa. This is because when the SegmentLength be large the possibility of finding best match between bits becomes less.
- When increase the SegmentLength the SNR will decrease and vice versa. This is because the value of SNR of the stego-image is proportional with the number of LSB that are changed in the pixels of the stego-image.
- The time of the hiding operation of each image was increase/decrease with few changes. It stays suitable when the size of the stego-image is small, but it will be long when the size of the stego-image becomes large. This is because the search time for best matching becomes huge when we using a stego-image of large size.

TABLE I. Stego-Images (.Bmp) Of Size (Width × Height × Palette)

Butterfly (128×128×3)

Garden (160×121×3)

Girls (300×225×3)

TABLE II. Recorded Results of Performance Experiments

	Classic-LSB Technique			Proposed LSB Technique		
Stego-Image	***Butterfly***	***Garden***	***Girls***	***Butterfly***	***Garden***	***Girls***
Length of Secret Message (Characters)	500	1000	2000	500	1000	2000
Number of LSB Changed	1972	4001	8081	1508	2987	5760
Signal to Noise Ratio (SNR) of the Stego-Image	51.950	50.107	51.890	53.115	51.377	53.360
Time of Hiding Operation (Second)	0.047	0.140	0.421	3.76	7.005	60.372
Time of Extracting Operation (Second)	0.110	0.125	0.421	0.109	0.124	0.421

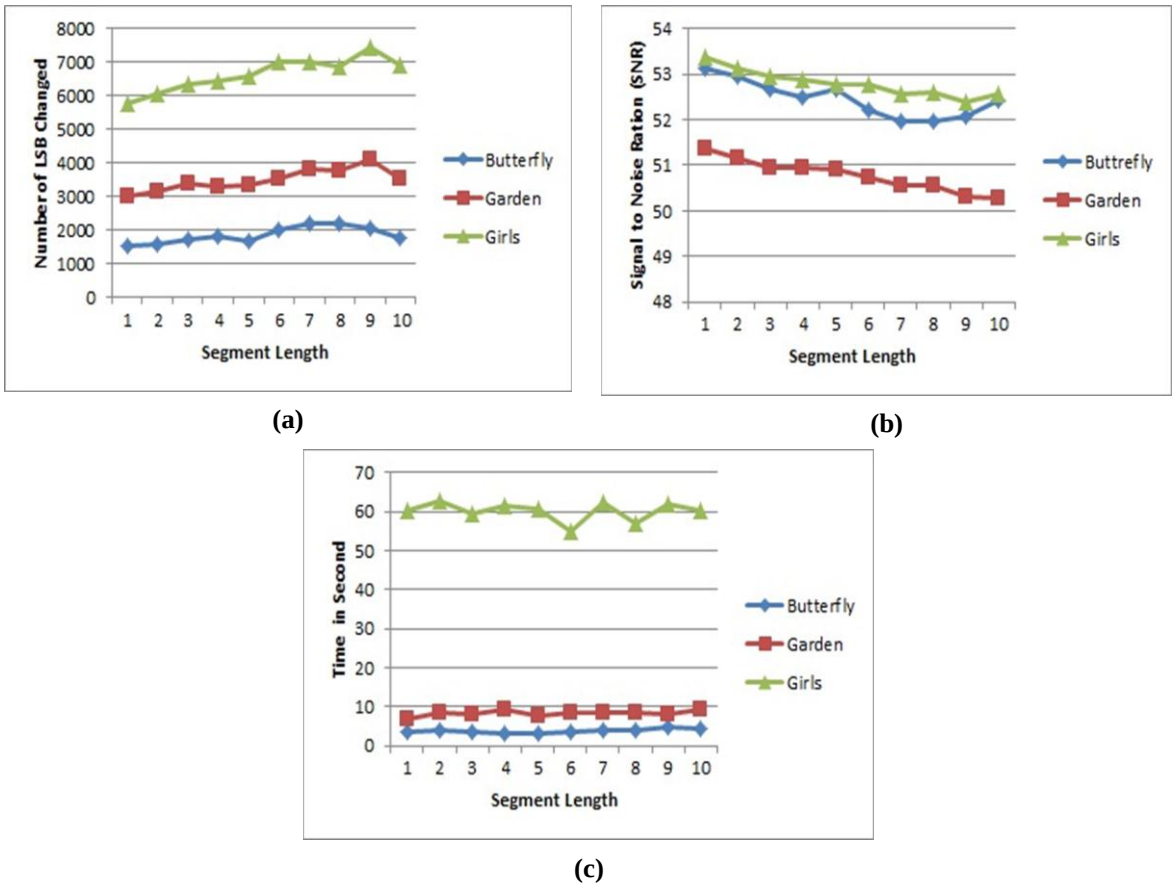

Figure 1. The effect of the SegmentLength on (a) Number of LSB changed, (b) Signal to Noise Ratio (SNR) of the stego-image, (c) Time of hiding operation.

VI. Conclusion

The idea to enhance the performance of the Classic-LSB image steganography technique was present in this paper. The message segmentation LSB image steganography technique was suggested here by splitting the long secret message into number of short segments. Then hide these short segments in different parts of the best matched LSB in the pixels of the stego-image. The main goal behind this suggested technique is to decrease the number of LSB that are changed of the pixels in the stego-image and as a result increase the immunity of the stego-image against the attack by human visual system (HVS). The recorded results from the experiments showed that the proposed LSB image steganography technique success in increase the security of the secret message that is hid in the stego-image by decreasing the number of LSB that are changed in the pixels of the stego-image.

The challenge point of the proposed LSB is in the long time of the hiding operation that is spend during the exhaustive search to find the best matching when using a large size stego-image. I

n the next work, we will try to minimize the effect of this weak point on the performance of the proposed LSB. But in spite of this point, the Segmented-LSB still can be used instead of the Classic-LSB to satisfy more security for the secret message.

References

[1] Cheddad, J. Condell, K. Curran, & P. Kevitt, (2010). Digital image steganography- survey and analysis of current methods. Signal Processing, 90, 727–752. doi: http://10.1016/j.sigpro.2009.08.010

[2] Adnan Gutub, Ayed Al-Qahtani, & Abdulaziz Tabakh. (2009). Triple-A: secure RGB image steganography based on randomization. AICCSA, IEEE/ACS International Conference on Computer Systems and Applications, Rabat, Morocco, 400-403, doi: http://doi.ieeecomputersociety.org/10.1109/AICCSA.2009.5069356

[3] Kaur, R. Dhir, & G. Sikka. (2009). A new image steganography based on first component alteration technique. International Journal of Computer Science and Information Security (IJCSIS), 6, 53-56. http://arxiv.org/ftp/arxiv/papers/1001/1001.1972.pdf

[4] Alvaro Martin, Guillermo Sapiro, & Gadiel Seroussi. (2005). Is Steganography Natural. IEEE Transactions on Image Processing, 14(12), 2040-2050. doi: 10.1109/TIP.2005.859370

[5] Bhattacharyya, A. Roy, P. Roy, & T. Kim. (2009). Receiver compatible data hiding in color image. International Journal of Advanced Science and Technology, 6, 15-24. http://www.sersc.org/journals/IJAST/vol6/2.pdf

[6] EE. Kisik Chang, J. Changho, & L. Sangjin. (2004). High Quality Perceptual Steganographic Techniques. Springer. 2939, 518-531. doi: 10.1007/978-3-540-24624-4_42, http://www.springerlink.com/content/c6guuj5xnyy4wj3c/

[7] C. Kessler. (2001). Steganography: Hiding Data within Data. An edited version of this paper with the title "Hiding Data in Data".

Windows & .NET Magazine. [Online] Available: http://www.garykessler.net/library/steganography.html (October 4,2011)

[8] Gandharba Swain, & S.K.lenka. (2010). Steganography-Using a Double Substitution Cipher. International Journal of Wireless Communications and Networking. 2(1), 35-39. ISSN: 0975-7163. http://www.serialspublications.com/journals1.asp?jid=436&jtype

[9] Hideki Noda, Michiharu Nimi, & Eiji Kawaguchi. (2006). High-performance JPEG steganography using Quantization index modulation in DCT domain. Pattern Recognition Letters, 27, 455-46. http://ds.lib.kyutech.ac.jp/dspace/bitstream/10228/450/1/repository6.pdf

[10] Kathryn (2005). A Java Steganography Tool. http://diit.sourceforge.net/files/Proposal.pdf

[11] Gandharba Swain, Dodda Ravi Kumar, Anita Pradhan, Saroj Kumar Lenka, (2010). A Technique for Secure Communication Using Message Dependent Steganography. Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010 for International Conference [ICCT-2010], 3rd - 5th December. http://interscience.in/SpIss_ijcct_icct2010vol2_no234/32_EC31.pdf

[12] Motameni, M.Norouzi, M.Jahandar, & A. Hatami. (2007). Labeling method in Steganography. Proceedings of world academy of science, engineering and technology, 24, 349-354. ISSN 1307-6884. http://www.waset.org/journals/waset/v30/v30-66.pdf

[13] Zhang, & H. Tang. (2007). A novel image steganography algorithm against statistical analysis. Proceeding of the IEEE, 19, 3884-3888. doi: 10.1109/ICMLC.2007.4370824

[14] Lisa M. Marvel, & Charles G. Boncelet. (1999). Spread Spectrum Image Steganography. IEEE Transactions on Image Processing, 8(8), 1075-1083. doi: 10.1109/83.777088

[15] Mei-Yi Wu, Yu-Kun Ho, & Jia-Hong Lee. (2004). An iterative method of palette-based image steganography. Pattern Recognition Letters, 25, 301-309. doi: 10.1016/j.patrec.2003.10.013

[16] Mohammed A.F Al Husainy. (2009). Image Steganography by mapping Pixels to letters. Journal of Computer Science, 5(1), 33-38. ISSN 1549-3636. doi: 10.3844/jcssp.2009.33.38, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.165.7818&rep=rep1&type=pdf

[17] Mohammad Ali Bani Younes, & Aman Jantan. (2008). A New Steganography Approach for Image Encryption Exchange by using the LSB insertion. IJCSNS International Journal of Computer Science and Network Security, 8(6), 247-254. http://paper.ijcsns.org/07_book/200806/20080634.pdf

[18] M.T. Parvez , & A. Gutub. (2008). RGB intensity based variable-bits image steganography. APSCC 2008 –Proceedings of 3rd IEEE Asia-Pacific Services Computing Conference, Yilan, Taiwan, doi: 10.1109/APSCC.2008.105

[19] N.F. Johnson, & J. Suhil. (2006). Exploring Steganography:Seeing the Unseen. Computing Practices. http://www.jjtc.com/pub/r2026.pdf

[20] P.Mohan Kumar, & D.Roopa (2007). An Image Steganography Framework with Improved Tamper Proofing. Asian Journal of Information Technology, 6(10), 1023-1029. ISSN: 1682-3915. http://medwelljournals.com/abstract/?doi=ajit.2007.1023.1029

[21] Po Yuch Chen, & Hung Ju Lin. (2006). A DWT Based Approach for Image Steganography. International journal of Applied Science and Engineering, 4(3), 275-290. http://www.cyut.edu.tw/~ijase/2006/4-3(Microsoft%20Word%20-%20010-009-6).pdf

[22] Ran-Zan Wang, & Yeh-Shun Chen. (2006). High Payload Image Steganography Using Two-Way Block Matching. IEEE Signal Processing letters, 13(3), 161-164. doi: 10.1109/LSP.2005.862603

[23] Ross J. Anderson, & Fabian A.P. Petitcolas. (1998). On The Limits of steganography. IEEE Journal of selected Areas in communication, 16(4), 474-481. Special Issue on Copyright and Privacy protection. ISSN 0733-8716. http://www.cl.cam.ac.uk/~rja14/Papers/jsac98-limsteg.pdf

[24] Sorina Dumitrescu, & Xiaolin (2005). A New Framework of LSB Steganalysis of Digital Media. IEEE Transactions on Signal Processing, 53(10), 3936-3947. doi: 10.1109/TSP.2005.855078

[25] Xinpeng Zhang, Shuozhong Wang, & Zhenyu Zhou. (2008). Multibit Assignment Steganography in Palette Images. IEEE Signal Processing Transactions, 15, 553-556. doi: 10.1109/LSP.2008.2001117

[26] Sukhpreet Kaur, Sumeet Kaur (2010). A Novel Approach for Hiding Text Using Image Steganography. (IJCSIS) International Journal of Computer Science and Information Security, 8(7), October. http://www.scribd.com/doc/40763180/A-Novel-Approach-for-Hiding-Text-Using-Image-Steganography

AUTHORS PROFILE

Mohammed Abbas Fadhil Al-Husainy was born in Mosul, Iraq, in January 1973. He received the M.Sc. and Ph.D. degrees in 1996 and 2002, respectively. From 1997 to 2002, he was a lecturer in the Department of Computer Science, Al-Hadba University of Mosul. Since 2002 he has been an assistant professor in the Departments: Computer Science and Multimedia Systems, Faculty of Science and Information Technology, Al-Zaytoonah University of Jordan. He lectures in the areas of microprocessors, data structures, algorithm design and analysis, digital design systems, operating systems, cryptography, computer organization, programming languages. His research interests are in the broad field of algorithm design, including multi-media data processing, scheduling algorithms, Information Security and cryptography algorithms.