JPEG versus GIF Images in forms of LSB Steganography

[1] ELTYEB E. ABED ELGABAR, [2] FAKHRELDEEN A. MOHAMMED

[1, 2] Information Technology, College of Computer Science and Information Technology - Khulais,
King Abdul Aziz University, Jeddah, Khulais, Saudi Arabia

Abstract - Steganography (from Greek steganos, or "covered," and graphie, or "writing") is the hiding of undisclosed message (such as text, image, audio and video) within an ordinary message (such as text, image, audio and video) and the extraction of it at its target (receiver). Steganography takes cryptography a step farther by hiding an encrypted message so that no one suspects it exists. This paper compares and analyses Least Significant Bit (LSB) algorithm using the cover object as an image with a focus on two types: JPEG and GIF. The comparison and analysis are done with deference number of criteria (Robustness against statistical attacks, Invisibility, Steganalysis detection, Robustness against image manipulation, Efficient when amount of data reasonable, Payload capacity, Unsuspicious files and Amount of embedded data) to understand their strengths and weaknesses.

Keywords - *Steganography, Steganographic, Least significant bit (LSB), Lossless, lossy*

1. Introduction

When we want to send data safely in a communication channel or media, the very first idea that automatically and spontaneously jumps to mind is that we have to encrypt them when we use a secured or unsecured communication channel. However well known, Encryption science is one of the ancient and effective sciences whose codes can be resolved via surveillance and within the course of time if we put into consideration the high speeds of the modern apparatuses used in the realm of decoding. Therefore, there has been a dire necessity and need for the use of modern and updated technologies to protect these data such as the science of ‘Data Hiding’, (steganography) which is an ancient, and at the same time modern science, which proved has its effectiveness, efficiency and accuracy in securing data. Steganography is an ancient Greek word composed of two syllables meaning covered or concealed writing. Steganography also known as art and science of hiding information by embedding messages within other, seemingly harmless message. Steganography means “covered writing” in Greek [3].The main goal of steganography is to hide the message within another message called cover message such as text, image video and audio, so steganography can be seen as the complement of cryptography whose goal is to hide the content of a message.

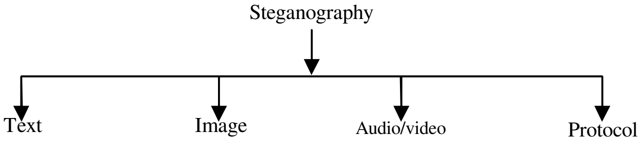

Fig1. Categories of Steganography

1.1 Types of Steganography

Steganography can be classified into various types (General types) [30]:

1. Pure Steganography.
2. Secret key Steganography.
3. Public key Steganography.

1.2 Depending upon the cover medium used [5]:

1. Text Steganography.
2. Image Steganography.
3. Audio Steganography.
4. Video Steganography.

1.3 Steganography Carrier Files [30]

1. Jpeg.
2. Gif.
3. Wav.
4. Mp3.

1.4 Steganography Tools [30]

1. Steganos.
2. S-Tools (GIF, JPEG).
3. StegHide (WAV, BMP).

4. Invisible Secrets (JPEG).
5. JPHide.
6. Camouflage

1.5 Methods of detecting the use of Steganography

1. Visual Detection (JPEG, GIF , BMP, etc.)
2. Audible Detection (WAV, MPEG, etc.)
3. Statistical Detection (changes in patterns of the pixels or LSB – Least Significant Bit) or Histogram Analysis
4. Structural Detection - View file properties/contents
    a. Size difference.
    b. Date/time difference.
    c. Contents – modifications.
    d. Checksum

1.6 Basic Terms

- *Cover-object*, *c*: the original object where the message has to be embedded. Cover-text, cover-image,
- *Message*, *m*: the message that has to be embedded in the cover-object. It is also called stego-message or in the watermarking context mark or watermark.
- *Stego-object*, *s*: The cover object, once the message has been embedded.
- *Stego-key, k:* The secret shared between A and B to embed and retrieve the message [7].

1.7 The steganographic process

- *Embedding function, E*: is a function that maps the tripled cover-object c, message m and stego-key k to a stego-object s. $E(c,m,k) = s$
- *Retrieving function, D:* is a mapping from s to m using the stego-key k. $D(s,k) = m$
- *A secret key* steganographic system [12] can be defined as the quintuple $\delta =< C,M,K,E,D >$ where C is the set of possible cover-objects, M is the set of messages with $|C| \geq |M|$, K the set of secret keys, $E: C \times M \times K \rightarrow C \text{ and } D: C \times K \rightarrow M$ with the property that $D(E(c,m,k),k) = m$ for all $m \in M, c \in C \text{ and } k \in K$. [13].

2. Image Definition

To a computer, an image is a collection of numbers that constitute different light intensities in different areas of the image [6]. This numeric representation forms a grid and the individual points are referred to as pixels. Most images on the Internet consists of a rectangular map of the image's pixels (represented as bits) where each pixel is located and its color [10]. These pixels are displayed horizontally row by row. The number of bits in a color scheme, called the bit depth, refers to the number of bits used for each pixel [12]. The smallest bit depth in current color schemes is 8, meaning that there are 8 bits used to describe the color of each pixel [12]. Monochrome and grayscale images use 8 bits for each pixel and are able to display 256 different colors or shades of grey. Digital color images are typically stored in 24-bit files and use the RGB color model, also known as true color [12]. All color variations for the pixels of a 24-bit image are derived from three primary colors: red, green and blue, and each primary color is represented by 8 bits [5]. Thus in one given pixel, there can be 256 different quantities of red, green and blue, adding up to more than 16-million combinations, resulting in more than 16-million colors [12]. Not surprisingly the larger amount of colors that can be displayed, the larger the file size [10].

2.1 Image Format

There are several types of image file formats that can be used for steganography such as, JPEG, GIF, TIFF, BMP and PNG; each has certain advantages and disadvantages for hiding messages.

2.1.1 Joint Photographic Experts Group (JPEG)

The term actually stands for "Joint Photographic Experts Group," because that is the name of the committee that developed the format. But you don't have to remember that because even computer nerds will think you're weird if you mention what JPEG stands for. Instead, remember that a JPEG is a compressed image file format. JPEG images are not limited to a certain amount of color, like GIF images are. Therefore, the JPEG format is best for compressing photographic images. So if you see a large, colorful image on the Web, it is most likely a JPEG file. While JPEG images can contain colorful, high-resolution image data, it is a lossy format, which means some quality is lost when the image is compressed. If the image is compressed too much, the graphics become noticeably "blocky" and some of the detail is lost. Like GIFs, JPEGs are cross platform, meaning the same file will look the same on both a Mac and PC [6].

2.1.1.1 General JPEG format properties

- Are commonly used for photo.

- Can be compressed to a smaller size.
- JPEG files allow only 8 - 24-bit indexed color.
- JPEG files use lossy compression.

2.1.2 Graphics Interchange Format (GIF)

GIF is used for the purpose of storing multiple bitmap images in a single file for exchange between platforms and images. It is often used for storing multi-bit graphics and image data. GIF is not associated with a particular software application but was designed “to allow the easy interchange and viewing of image data stored on local or remote computer systems”. GIF is stream based and is made up of a series of data packets called blocks (which can be found anywhere in the file) and protocol information. GIF files are read as a continuous stream of data and the screen is read pixel by pixel.GIF is used also because it applies lossless file compression method.

2.1.2.1 General GIF format properties

- Can be compressed to a small size**.**
- Are commonly used for images presented on the web.
- GIF files allow only 8-bit indexed color.
- GIF files use lossless LZW compression.
- GIF files support transparency.
- Animated GIF files can be created by sequences of single images.
- GIF files can be saved in an interlaced format that allows progressive download of web images (low-resolution version of an image first then gradually comes into focus the rest of the data is downloaded**.**

GIF images uses indexed color, which contain a color palette with up to 256 different colors out of 16,777,216 possible colors [14], and the Lempel- Ziv-Welch (LZW) compressed matrix of palette indices. Thus, LSB method in GIF is efficient when used for embedding a reasonable amount of data in an image [15].

Table 1: Comparison of JPEG & GIF Images

	JPEG	*GIF*
File types	Joint Photographic Experts Group	Graphics interchange format
File extensions	.jpg, .jpeg, .jpe	.gif, .gfa
File Size	Small	Large
Resolution	High	Low
Support Color	16 Million Colors	256 Colors
support transparency	No	Yes
Ideal for	Photo	Animation, icons or symbols
Color Depth	8-24 bit color	8-bit color
Compression algorithms	lossy	Lossless(LZW)

3. 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 preferred. Example of Lossy compression format is JPEG format files. Examples of Lossless compression formats are GIF [29].

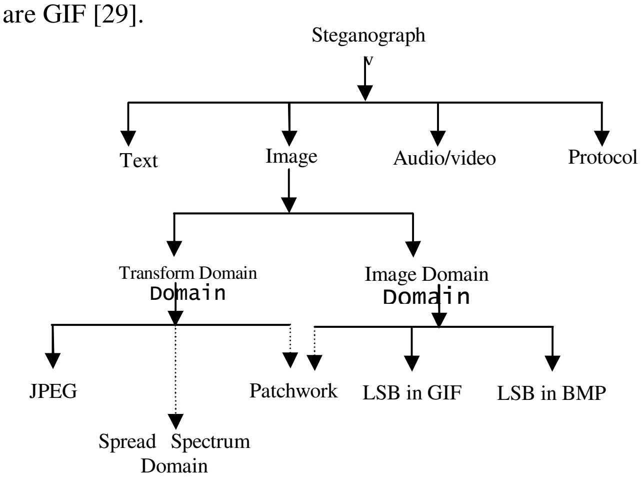

Fig.2 Categories of image steganography

3.1 JPEG Steganography

There are two broad categories of image-based steganography that exist today: frequency domain and spatial domain steganography. The first digital image steganography was done in the spatial domain using LSB coding (replacing the least significant bit or bits with embedded data bits) [30]. Since JPEG transforms spatial data into the frequency domain where it then employs lossy compression, embedding data in the spatial domain before JPEG compression is likely to introduce too much noise and result in too many errors during decoding of the embedded data when it is returned to the spatial domain. These would be hard to correct using error correction coding. Hence, it was thought that steganography would not be possible with JPEG images because of its lossy characteristics. However, JPEG encoding is divided into lossy and lossless stages [23]. DCT transformations to the frequency domain and quantization stages are lossy, whereas entropy encoding of the quantized DCT coefficients (which we will call the JPEG coefficients to distinguish them from the raw frequency domain coefficients) is lossless compression. Taking advantage of this, researchers have embedded data bits inside the JPEG coefficients before the entropy coding stage [17].

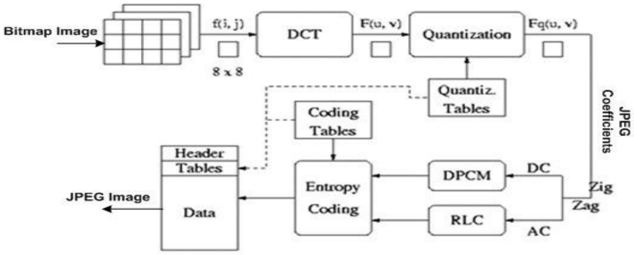

Fig.3 Most popular image format used [27].

4. Overview of LSB Algorithm

A digital image consists of a matrix of color and intensity values. In a typical gray scale image, 8 bits/pixel are used. In a typical full-color image, there are 24 bits/pixel, 8 bits assigned to each color components.

Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image [5]. The least significant bit in other words, the 8th bit of some or all of the bytes inside an image is changed to a bit of the secret message. When using a 24-bit image, a bit of each of the red, green and blue color components can be used, since they are each represented by a byte. In other words, one can store 3 bits in each pixel. An 800 × 600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data [9]. For example a grid for 3 pixels of a 24-bit image can be as follows:

(00101101 00011100 11011100)
(10100110 11000100 00001100)
(11010010 10101101 01100011)

When the number 200, which binary representation is 11001000, is embedded into the least significant bits of this part of the image, the resulting grid is as follows:

(0010110**1** 0001110**1** 1101110**0**)
(1010011**0** 1100010**1** 0000110**0**)
(1101001**0** 1010110**0** 0110001**1**)

Although the number was embedded into the first 8 bytes of the grid, only the 3 underlined bits needed to be changed according to the embedded message. On average, only half of the bits in an image will need to be modified to hide a secret message using the maximum cover size [10]. Since there are 256 possible intensities of each primary color, changing the LSB of a pixel results in small changes in the intensity of the colors. These changes cannot be perceived by the human eye - thus the message is successfully hidden. With a well-chosen image, one can even hide the message in the least as well as second to least significant bit and still not see the difference [5].

In the above example, consecutive bytes of the image data from the first byte to the end of the message are used to embed the information. This approach is very easy to detect. A slightly more secure system is for the sender and receiver to share a secret key that specifies only certain pixels to be changed. Should an adversary suspect that LSB steganography has been used, he has no way of knowing which pixels to target without the secret key [6]. In its simplest form, LSB makes use of BMP images, since they use lossless compression. Unfortunately to be able to hide a secret message inside a BMP file, one would require a very large cover image.

The advantage of LSB embedding is its simplicity and many techniques use these methods [5]. LSB embedding also allows high perceptual transparency. However, there are many weaknesses when robustness, tamper resistance, and other security issues are considered. LSB encoding is extremely sensitive to any kind of filtering or manipulation of the stego-image. Scaling, rotation, cropping, addition of noise, or lossy compression to the stego-image is very likely to destroy the message. Furthermore an attacker can easily remove the message by removing (zeroing) the entire LSB plane with very little change in the perceptual quality of the modified stego-image.

4.1 Advantages of LSB [5][6]

1. Major advantage of the LSB algorithm is it is quick and easy.
2. There has also been steganography software developed which work around LSB color alterations via palette manipulation.
3. LSB insertion also works well with gray-scale images

4.2 The LSB Algorithm [5][6]

1. Select *cover-object* (BMP/JPEG) *c* as an input.
2. Encode the *c* in binary [16].
3. The Secret Message, *m*.
4. Encode the *m* in binary [16].
5. Choose one pixel of the *c* randomly.
6. Use a pixel selection to hide information in the *c* .
7. Save the new image (*Stego-object*) *s*.

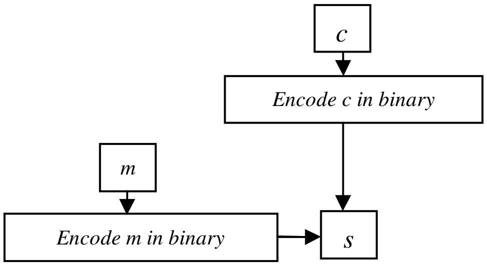

Fig. 4 the LSB Algorithms

4.3 LSB in JPEG

The most commonly used method to embed a bit is LSB embedding, where the least significant bit of a JPEG coefficient is modified in order to embed one bit of message. Once the required message bits have been embedded, the modified coefficients are compressed using entropy encoding to finally produce the JPEG stego image. By embedding information in JPEG coefficients, it is difficult to detect the presence of any hidden data since the changes are usually not visible to the human eye in the spatial domain. During the extraction process, the JPEG file is entropy decoded to obtain the JPEG coefficients, from which the message bits are extracted from the LSB of each coefficient.

LSB embedding [24], [25], [26] is the most common technique to embed message bits DCT coefficients. This method has also been used in the spatial domain where the least significant bit value of a pixel is changed to insert a zero or a one. A simple example would be to associate an even coefficient with a zero bit and an odd one with a one bit value. In order to embed a message bit in a pixel or a DCT coefficient, the sender increases or decreases the value of the coefficient/pixel to embed a zero or a one. The receiver then extracts the hidden message bits by reading the coefficients in the same sequence. And decoding them in accordance with the encoding technique performed on it. The advantage of LSB embedding is that it has good embedding capacity and the change is usually visually undetectable to the human eye. If all the coefficients are used, it can provide a capacity of almost one bit per coefficients using the frequency domain technique.

4.4 LSB in GIF [17]

We can use GIF images for LSB steganography [17], although extra care should be taken. The main issue with the palette based approach is that if one changes the least significant bit of a pixel, it could result in an entirely different color since the index to the color palette gets modified. One possible solution to this problem is to sort the palette so that the color differences between consecutive colors are minimized. The strong and weak points regarding embedding information in GIF images using LSB is that since GIF images only had a bit depth of 8, the total amount of information that could be embedded will be less. GIF images are vulnerable to statistical as well as visual attacks, since the palette processing which has to be done on the GIF image leaves a clear signature on the image. This approach was dependent on the file format as well as the image itself, since a wrong choice of image could results in the message being visible.

5. The Applying and Evaluation

5.1 The original image (before hiding)

Fig.5a JPEG Image

Fig.5b GIF Image

Fig.6a JPEG Image

Fig.6b GIF Image

Name	***JPEG(a)***			***GIF (b)***		
	Size ***MB***	*Dimension* ***X*Y***	*Depth* ***bpp***	*Size* ***MB***	*Dimension* ***X*Y***	*Depth* ***bpp***
Fig.5	0.08	800*600	24	0.24	800*600	8
Fig.6	1.38	1920*2560	24	2.38	1920*2560	8

Table 2: Properties of JPEG & GIF Images

5.2 The Images After Hiding

Fig7a. JPEG Image

Fig7b.GIF Image

Fig8a. JPEG Image

Fig 8 b. GIF Image

Table 3: Comparison of LSB for JPEG GIF Images

	JPEG	***GIF***
Robustness against statistical attacks	Medium	Low
Invisibility	High	Medium
Steganalysis detection	Medium	Low
Percentage Distortion less resultant image	Medium	Medium
Robustness against image manipulation	Medium	Low
Efficient when amount of data reasonable	Medium	Medium
Independent of file format	Low	Low
Payload capacity	Medium	Medium
Unsuspicious files	High	Low
Amount of embedded data	Low	Low

Table 4: Comparison of LSB for JPEG & GIF Images

	JPEG	***GIF***
Robustness against statistical attacks	1	0
Invisibility	2	1
Steganalysis detection	1	0
Percentage Distortion less resultant image	1	1
Robustness against image manipulation	1	0
Efficient when amount of data reasonable	1	1
Independent of file format	0	0
Payload capacity	1	1
Unsuspicious files	High	0
Amount of embedded data	0	0

(high = 2, medium =1 and low =0)

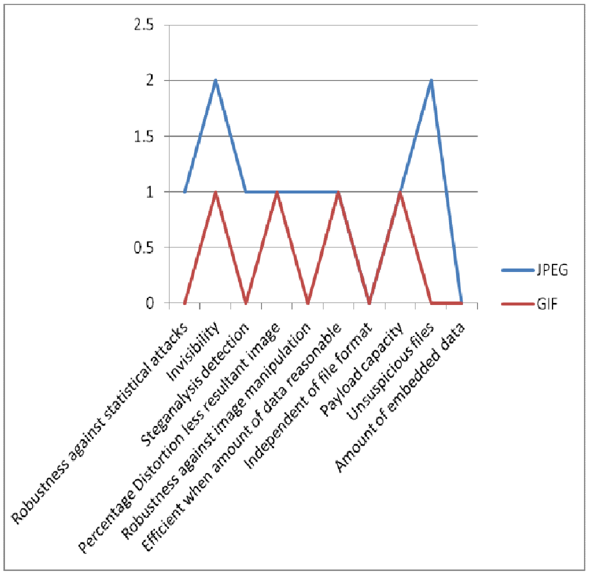

Fig 9.Comparison of LSB for JPEG & GIF Images

6. Conclusion

In the image of kind JPEG we find medium data embedded, high unsuspicious, medium robustness against statistical attacks, high invisibility and low Independent of file format. For the image of kind GIF we find very little data embedded, low unsuspicious, medium invisibility and low robustness against statistical attacks.

.

References

[1] Eltyeb E.Abed Elgabar, Haysam A. Ali Alamin, “Comparison of LSB Steganography in GIF and BMP Images ”, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4, September 2013

[2] Eltyeb E.Abed Elgabar “Comparison of LSB Steganography in BMP and JPEG Images ”, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-5, November 2013.

[3] "Watermarking Application Scenaros and Related Attacks ", IEEE international Conference on Image Processing, Vol. 3, pp. 991 – 993, Oct. 2001.

[4] Moerland, T., “Steganography and Steganalysis”, Leiden Institute of Advanced Computing Science, www.liacs.nl/home/ tmoerl/privtech.pdf.

[5] Henk C. A. van Tilborg (Ed.), "Encyclopedia of cryptography and security", pp.159. Springer (2005).

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

[7] Krenn, R., “Steganography and Steganalysis”, http://www.krenn.nl/univ/cry/steg/article.pdf

[8] Pallavi Hemant Dixit, Uttam L. Bombale, " Arm Implementation of LSB Algorithm of Steganography", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-3, February 2013.

[9] “Reference guide:Graphics Technical Options and Decisions”, http://www.devx.com/ /Article/1997.

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

[11] NXP & Security Innovation Encryption for ARM MCUs ppt.

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

[13] “MSDN:About Bitmaps” <http://msdn.microsoft. com/library/default.asp?url=/library/enus/gdi/bitmaps_99ir.asp?frame=tru>, 2007,M Corporation.

[14] V. Lokeswara Reddy, Dr. A. Subramanyam and Dr.P. Chenna Reddy, " Implementation of LSB Steganography and its Evaluation for Various File Formats", Int. J. Advanced Networking and Applications, Volume: 02, Issue: 05, Pages: 868-872 (2011).

[15] Neeta Deshpande, Snehal Kamalapur and Jacobs Daisy, “Implementation of LSB steganography and Its Evaluation for Various Bits”, 1st International Conference on Digital Information Management, 6 Dec. 2006 pp. 173-178.

[16] J. E. Boggess III, P. B. Nation, M. E. Harmon, “Compression of Colour Information In Digitized Images Using an Artificial Neural Network”, Proceedings of the IEEE 1994 National Aerospace and Electronics Conference, Issue 23-27 May 1994 Page(s):772 - 778 vol.2.

[17] Atallah M. Al-Shatnawi, "A New Method in Image Steganography with Improved Image Quality", Applied Mathematical Sciences, Vol. 6, 2012, no. 79, 3907 - 391.

[18] Ze-Nian Li and Marks S.Drew, "Fundamentals of Multimedia, School of computing Science Simon Faster University, Pearsoll Education, Inc, 2004.

[19] J. Fridrich, M. Goljan, and R. Du, “Detecting LSB steganography in color, and gray-scale images,” IEEE Multimedia, vol. 8, no. 4, pp. 22–28, Oct. 2001.

[20] Priya Thomas," Literature Survey On Modern Image Steganographic Techniques", International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 5, May - 2013 ISSN: 2278-0181.

[21] V. Lokeswara Reddy, Dr. A. Subramanyam, Dr.P. Chenna Reddy ,"Implementation of LSB Steganography and its Evaluation for Various File Formats", Int. J. Advanced Networking and Applications Volume: 02, Issue: 05, Pages: 868-872 (2011).

[22] Roshidi Din and Hanizan Shaker Hussain, “The Capability of Image In Hiding A Secret Message”, Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing, September 2006.

[23] D. Llamas, C. Allison, and A. Miller, \Covert channels in internet protocols: A survey," in Proceedings of the 6th Annual Postgraduate Symposium about the Convergence of Telecommunications, Networking and Broadcasting, 2005.

[24] Neil R. Bennett, JPEG STEGANALYSIS & TCP/IP STEGANOGRAPHY, University of Rhode Island , 2009.

[25] H. Wu, N. Wu, C. Tsai, and M. Hwang, “Image steganographic scheme based on pixel-value differencing and LSB replacement methods,” IEE Proceedings-Vision, Image and Signal Processing, vol. 152, no. 5, pp. 611–615, 2005.

[26] R. Chandramouli and N. Memon, “Analysis of LSB based image steganography techniques,” in Image Processing, 2001. Proceedings. 2001 International Conference on, vol. 3, 2001.

[27] Y. Lee and L. Chen, “High capacity image steganographic model,” IEE Proceedings-Vision, Image and Signal Processing, vol. 147, no. 3, pp. 288–294, 2000.

[28] W. Pennebaker and J. Mitchell, JPEG still image data compression standard. Kluwer Academic Publishers, 1993.

[29] Moerland, T., “Steganography and Steganalysis”, Leiden Institute of Advanced Computing Science, Silman, J., “Steganography and Steganalysis: An Overview”, SANS Institute, 2001 Jamil, T.,“Steganography: The art of hiding information is plain sight”, IEEE Potentials, 18:01, 1999.

[30] Hamid.A.Jalab, A.A Zaidan, B.B Zaidan, “New Design for Information Hiding with in Steganography Using Distortion Techniques”, International Journal of

Engineering and Technology (IJET)), Vol 2, No. 1, ISSN: 1793-8236, Feb (2010), Singapore.

DR.ELTYEB ELSAMANI ABD ELGABAR ELSAMANI, Assistant Professor(2009) in the Computer Science at Faculty of Computer Science and Information Technology, Information Technology Department - Khulais - King Abdul Aziz University- Jeddah - Saudi Arabia. Assistant Professor in the Computer Science at the Department of Computer Science, Faculty of Computer Science and Information Technology - Alneelain University - Khartoum - Sudan. . Main specialization is Information Security in particular and Encryption in specific. A member of the committee of Standard specifications for Computers Hardware and Peripherals in the National Information Center (NIC) - Khartoum -Sudan , member of Standard specifications for Network Hardware in the National Information Center (NIC) - Khartoum -Sudan, and member of Curriculum of information technology department - Faculty of Kamleen Ahlia- Gazera Sudan.

Dr.Fakhreldeen is an Assistant Professor in the Computer Science at the Department of Information Technology, Faculty of Computer and Information Technology in khlais, King AbdulAziz University, Saudi Arabia. He is an Assistant Professor in the Computer Science at the Department of Computer Science, Faculty of Computer Science and Information Technology at Alneelain University, Sudan. His main specialization in particular is Performance Evaluation of Computer System. The researches interests include Network Technology & Application, Internet Security and Performance Evaluation of Internet Application. He is a member of the committee of the software standards in the public sector, NIC, Sudan. He is a member of the academic committee, faculty of CSIT, Alneelain University, Sudan.