Review of Various Techniques of Image Steganography

Bhanupriya Katre, Bharti

M.Tech Scholar, Department of Electronics and Communication, SISTech, Bhopal

Assistant Professor, Department of Electronics and Communication, SISTech, Bhopal

Abstract: Steganography – the art and science of hiding information has received much attention in the recent years. It can also be defined as the study of invisible communication that usually involves communication of secret data in an appropriate carrier, e.g., image, audio, video or TCP/IP header file. Many different carrier file formats can be used, but digital images are the most popular because of their frequency on the Internet. The main goal of steganography is to ensure that the transmitted message is completely masked, thereby ensuring that the message is accessible only to the intended receiver and does not attract attention from eavesdroppers and attackers. This paper is an attempt to study the various techniques use in steganography and provide a general overview of different algorithms of image steganography. There exists a large variety of steganography techniques some are more complex than others and all of them have respective strong and weak points.

Keywords: Digital image steganography, cover-image, stego-image, spatial domain; frequency domain.

I. INTRODUCTION

The word steganography is derived from the Greek words "stegos" meaning "cover" and "grafia" meaning "writing" defining it as "covered writing" means it conceals the very existence of the secret message in another medium (audio, video, image, communication)[4]. The media with and without hidden information are called stego media and cover media, respectively. The growing use of Internet need to store, send and receive personal information in a secured manner. Protection of the transmitted data from being intercepted or tampered has led to the development of various steganography techniques [11]. Steganography means is not to alter the structure of the secret message, but hides it inside a cover-object. The characteristics of the cover media depends on the amount of data that can be hidden, the perceptibility of the message and its robustness [3]. Steganography is applied in various fields and applications like intelligence agencies, military agencies , medical imagery, TV broadcasting , Checksum embedding , advanced data structures , radar systems and remote sensing[12].However, it is evident that steganography has numerous useful applications and will remain the point of attraction for researchers.

II. STEGANOGRAPHY IN DIGITAL MEDIUMS

Almost all digital file formats can be used for steganography, however only those with a high degree of redundant bits are preferred. The redundant bits of an object are those bits that can be

altered without the alteration being detected easily. Depending on the type of the cover object there are

many suitable stenographic techniques which are followed in order to obtain security [4].

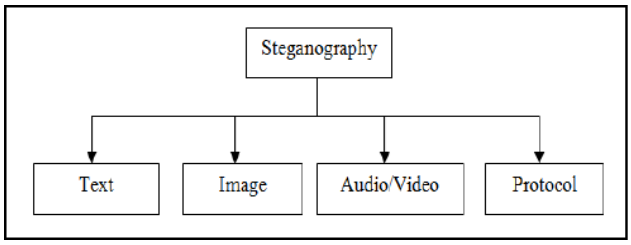

Fig. 1: Different Kinds of Steganography

Text steganography using digital files is not used very often since text files have a very small amount of redundant data [4]. Taking the cover object as image in steganography is known as image steganography. In this technique pixel intensities are used to hide the information [2].When taking audio as a carrier for information hiding it is called audio steganography. One different technique unique to audio steganography is masking, which exploits the properties of the human ear to hide information unnoticeably. Audio steganography uses digital audio formats such as WAVE, MIDI, AVI MPEG or etc. for steganography. Video steganography is a technique to hide any kind of files or information into digital video format. Video steganography uses such as H.264, Mp4, MPEG, AVI or other video formats [2]. The term protocol steganography refers to the technique of embedding information within messages and network control protocols such as TCP, UDP, ICMP, IP etc used in network transmission [4].

III. IMAGE STEGANOGRAPHY

Digital images have high degree of redundancy in representation and pervasive applications in daily life, thus are the most popular cover objects used for steganography. Generally image steganography is method of information hiding into cover-image and generates a stego-image [2].

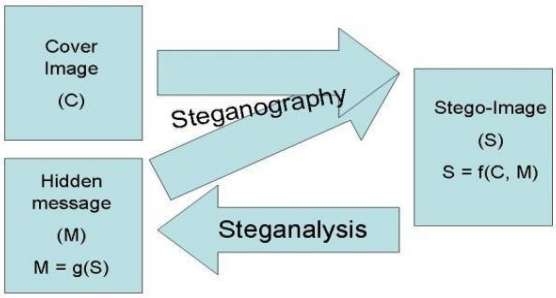

Fig. 2: Image Steganography

IV. CHARACTERIZATION OF STEGANOGRAPHY SYSTEMS

Steganographic techniques embed a message inside cover. Various parameters characterize the strength and weaknesses of the methods. Following are the parameters to measure the performance of the Steganographic system [1]:

A. ***Undetectability****:* The imperceptibility of a steganographic algorithm is the first and foremost requirement; it represents the ability to avoid detection [4]. Undetectability is directly affected by the size of the secret message and the format of the content of the cover image. Truly secure steganographic techniques should be undetectable neither by the human eye nor by the statistical attacks.

B. ***Robustness****:* Robustness refers to the ability of the embedded data to remain intact if the stego-system undergoes transformation, such as linear and nonlinear filtering, addition of random noise; and scaling, rotation, and loose compression [1]. It measures the ability of the steganographic technique to survive the attempts of removing the hidden information.

C. ***Payload capacity****:* It is the Maximum size of information that can be embedded and retrieved successfully by the stego system. Steganography aims at hidden communication and therefore requires sufficient embedding capacity [3].

D. ***Image Steganographic Techniques***

Image steganography techniques can be divided into following classification.

D.1 ***Spatial Domain Techniques***

Spatial domain techniques embed messages in the intensity of the pixels directly [10]. These techniques are also known as substitution techniques because they substitute redundant parts of a cover with a secret message. The receiver can extract the information if he has knowledge of the positions where secret information has been embedded [1]. The advantages of spatial domain methods are high embedding capacity, ease of implementation and imperceptibility of hidden data. The major drawback is its vulnerability to various simple statistical analysis methods [3].

D.1.1 ***Least Significant Bit****:* Least significant bit (LSB)-based steganography is one of the simplest techniques that hides a secret message in the LSBs of pixel values without introducing many perceptible distortions[2].The cover-image is first decomposed into bits planes and then least significant bit of the bits planes are replaced with the secret data bits .In 8 bit image some or all of the bytes inside an image is changed to a bit of the secret message whereas in 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[11]. In other words, one can store 3 bits in each pixel. LSB is most suitable for applications where the focus is on the amount of information to be transmitted and not on the secrecy of that information. There is less chance for degradation of the original image and more information can be stored in an image but it is less robust, the hidden data can be destroyed or lost with image manipulation and by simple attacks [4].

D.1.2 ***Pseudorandom Permutation:*** A more sophisticated approach is the use of pseudorandom number generator to spread the secret message over the cover in a rather random manner [8]. Since in this technique, it is not guaranteed that the subsequent message bits are embedded in the same order, thus increases the complexity for the attacker [1].

D.1.3 ***Palette-Based Image*** *:* Palette based images, for example GIF images cannot have a bit depth greater than 8, thus the maximum number of colors that a GIF can store is 256. GIF images are indexed images where the colors used in the image are stored in a palette, sometimes referred to as a color lookup table [4]. Since pixel values in a palette image are represented by indices into a color look-up table which contains the actual color RGB value, information can be encoded in the way the colors are stored in the palette. Thus even minor modifications to these indices can result in annoying artifacts [1]. In order to minimize the distortion caused by embedding, sort the palette so that the color differences between consecutive colors are minimized. It then embeds the message bits in the LSB of the color indices in the sorted pallet. LSB in GIF is a very efficient algorithm to use when embedding a reasonable amount of data in a greyscale image [6].

D.1.4 ***Cover Regions and Parity Bits:*** Any nonempty subset of $\{c_1 \ldots c_I(c)\}$ is called a Cover region. In this technique by dividing the cover into several disjoint regions, it is possible to store one bit of information in a whole cover region rather than in a single element. A parity bit of a region I can be calculated by [1]:

$$B(I) = \sum_{J \in I} LSB\left(c_j\right) \bmod 2$$

During embedding disjoint cover regions are selected and each region encodes one secret bit in the parity bit. In decoding process the parity bits of all selected regions are calculated and lined up to reconstruct the message. This method is not more robust than simple bit substitution technique but it may assume powerful in many cases [8].

D.1.5 ***Quantization and Dithering****:* Dithering and quantization to digital image can be used for embedding secret information. Some steganographic systems operate on quantized images. The difference ei between adjacent pixels xi and xi +1 is calculated and fed into a quantizer Q which outputs a discrete approximation Δi of the difference signal [1]. Thus in each quantization step a quantization error is introduced. In order to store the i_{th} message bit in the cover-signal, the quantized difference signal Δi is computed. If Δi does not match with the secret bit to be encoded Δi is replaced by nearest Δj where the associated bit equals the secret message bit. At the receiver side the difference signal is dequantized and added to the last signal sample in order to construct an estimate for the sequence [8].

D.1.6 ***Image Downgrading and Covert Channels:***

Image downgrading is a special case of a substitution system in which image acts both as a secret message and a cover. Given a cover-image and a secret image of equal dimensions, the sender exchanges the four least significant bits of the cover gray scale (or color) values with the four most significant bits of the secret image [1]. The receiver extracts the four least significant bits out of the stego-image, thereby gaining access to the most significant bits of the stego-image. Whereas the degradation of the cover is not visually noticeable in many cases, four bits are sufficient to transmit a rough approximation of the secret image.

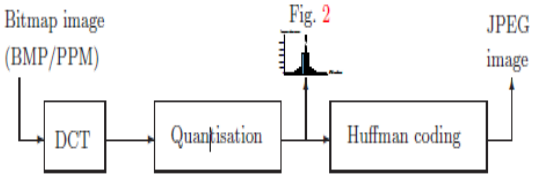

Fig .3. Conversion of Image format

D.2 ***Transform Domain Techniques****:*

Transform domain techniques embed secret information in a transform space of the signal [10] means the process of embedding data in the frequency domain of a signal to make them more robust to attack such as adding noise, compression, cropping, some image processing etc[12]. Many transform domain variations exist. One method is to use the Discrete Cosine Transformation (DCT) as a vehicle to embed information in image. Another method would be the use of wavelet transforms. Advantages of transform domain include higher level of robustness against simple statistical analysis but unfortunately it lacks high embedding means having lower payload [3].

D.2.1 ***JPEG Steganography:*** The JPEG file format is the most popular image file format on the Internet and they use lossy compression. One of the major characteristics of steganography is the fact that information is hidden in the redundant bits of an object and since redundant bits are left out when using JPEG it was feared that the hidden message would be destroyed [4]. Thus it is important to recognize that the JPEG compression algorithm is actually divided into lossy and lossless stages The DCT and the quantization phase form part of the lossy stage, while the Huffman encoding used to further compress the data is lossless. Steganography can take place between these two stages. Using the same principles of LSB insertion the message can be embedded into the least significant bits of the coefficients before applying the Huffman encoding. By embedding the information at this stage, in the transform domain, it is extremely difficult to detect, since it is not in the visual domain means JPEG steganography has high degree of imperceptibility [5]. This technique is is especially suitable for images that have to be communicated over an open systems environment like the Internet.

The major JPEG steganographic methods can be described as follows:

JSteg/JPHide*:* JSteg sequentially replaces the LSB of the non-zero quantized DCT coefficients with secret message bits whereas in JPHide the quantized DCT coefficients are not selected sequentially but are selected randomly by a pseudo-random number generator[5].

OutGuess*:* Outguess was proposed by Provos and Two versions of outguess are available: Outguess 0.13b which is exposed to statistical analysis, and the second is OutGuess-0.2 which has the ability to preserve frequency counts statistics and hence remain undetected [3]. OutGuess goes through the file twice. First it embeds information in LSB of the DCT coefficients by making a random walk, leaving some coefficients unchanged. Second it adjusts the remaining coefficient in order to preserve the original histogram of DCT coefficients so that it will match the cover image [9]. This is important because it will prevent a chi-squared attack.

MB*:* Model-based steganography presented a general framework for performing steganography and steganalysis using a statistical model of the cover media [5]. MB achieves a high message capacity while remaining secure against several first order statistical attacks. MB adapts the division of the carrier into a deterministic random variable and an in-deterministic one. Then a suitable model is employed to describe the distribution of in-deterministic

variable, which reflects the dependencies with deterministic variable [7].

YASS: Yet Another Steganographic Scheme belongs to JPEG steganography but it does not embed data in JPEG DCT coefficients directly. First an input image spatially divided into blocks with a fixed large size called as big blocks (B-blocks). Then within each B-block, an 8×8 subblock is randomly selected known as embedded host block (H-blocks) [5]. Next secret data encoded by error correction codes are embedded in the DCT coefficients of the H-blocks. Then after performing the inverse DCT to the H-blocks, the whole image is compressed and distributed as a JPEG image. For data extraction, image is first JPEG decompressed to spatial domain. Then data are retrieved from the DCT coefficients of the H-blocks [7].

F5: F5 comes after a series of F3 and F4 [9]. F5 steganographic algorithm was introduced by Westfield and it embeds message bits into randomly chosen DCT coefficients. The F5 algorithm employs matrix embedding that minimizes the necessary number of changes to hide a message of certain length [3]. In the embedding process, the message length and the number of non-zero AC coefficients are used to determine the best matrix embedding that minimizes the number of modifications of the cover image. The major strengths of F5 are its high embedding capacity without sacrificing security and its resistance to statistical and visual attacks [7].

D.2.2 Wavelet transforms technique: The discrete wavelet transform (DWT) method is favored over the discrete cosine transform (DCT) method in steganography because DWT provides better image resolution at various levels. DWT converts spatial domain information to the frequency domain information and it clearly partitions the high-frequency and low-frequency information on a pixel by pixel basis [5]. Wavelets are mathematical functions that divide data into frequency components, which makes them ideal for image compression. As compared to JPEG format, they are far better at approximating data with sharp discontinuities. Modifying data by using a wavelet transformation produces good quality with few perceptual artifacts means high capacity and high security steganography can be achieved using wavelet transformation techniques [4].

D.3 Spread Spectrum: In spread spectrum techniques, hidden data is spread throughout the cover-image to provide means of low probability of intercept. Spread spectrum techniques are defined as "Means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by means of a code which is independent of the data, and a synchronized reception with the code at the receiver is used for dispreading and subsequent data recovery"[1]. This band spread of a narrowband signal across a wide band of frequencies makes the energy of the narrowband signal in any one frequency band is low and therefore difficult to detect [4]. In spread spectrum image steganography the message is embedded in noise and then combined with the cover image to produce the stego image. Although the power of the signal to be transmitted can be large, the signal–to–noise ratio in

every frequency band will be small. Even if parts of the signal could be removed in several frequency bands, enough information should be present in the other bands to recover the signal [5]. There are two variants of spread spectrum steganography that are generally used [1]: direct sequence and frequency hopping scheme. In direct-sequence scheme, the secret signal which is to be added to the cover, spread by a constant called chip rate and modulated with a pseudorandom signal. In the frequency hopping schemes the frequency of the carrier signal is altered so that it hops from one frequency to another rapidly [8]. Spread spectrum techniques are generally quite robust against statistical attacks and because of this it is extensively used in military communications.

D.4 Statistical Techniques: Statistical Techniques tends to modulate the statistical properties of an image in addition to preserving them in the embedding process and these modifications are so small that it takes advantage of the human weakness in detecting luminance variation [12]. Statistical Steganography techniques utilize the existence of "1-bit" steganography schemes, which embed one bit of information in digital carrier. This is done by modifying the cover in such a way that some characteristics change significantly if a "1" is transmitted otherwise the cover is left unchanged. So that the receiver can distinguish unmodified covers from the modified ones [5]. To send multiple bits an image is broken into sub-images, each corresponding to a single bit of the message. These sub-images are vulnerable to cropping, rotating, and scaling attacks, along with any attacks that work against the steganography techniques [8]. The payload capacity and invisibility depends on the cover image selected.

D.4.1 ***Patchwork:*** Patchwork is a statistical technique that uses redundant pattern encoding to embed a message in an image. Patchwork algorithm adds redundancy to the secret data and then scatters the hidden information throughout the image. A pseudorandom generator selects two areas of image, patch A and patch B. The intensities of the pixels in the patch a are increased by a constant value means all the pixels are lightened while the pixels of the patch B are decreased with the same constant value means darkened [4]. The contrast changes in this patch subset encodes one bit while not changing the average luminosity and the changes are small and imperceptible. One bit embedding limits the functionality of patchwork. To increase the

performance of the algorithm first divide the image into sub images and then embed data to each of them [8]. In this way more bits are embedded and the secret message is distributed over the entire image, if one patch is destroyed the other may still survive. The patchwork approach is used independent of the host image and proves to be quite robust against malicious or unintentional image manipulation [4]. Patchwork is most suitable for transmitting a small amount of very sensitive information.

D.5 ***Distortion Techniques:*** Distortion techniques require knowledge of the original cover image during the decoding process [2]. The encoder adds a sequence of changes to the cover image and this sequence of modification is chosen in such a way that it corresponds to a specific secret message to be transmitted [1]. Then decoder measures the differences between the original cover image and the distorted cover image in order to restore the secret message. So, information is described as being stored by signal distortion. The message is encoded at pseudo-randomly chosen pixels. If the stego image is different from the cover image at the given message pixel, then the message bit is a "1." Otherwise, the message bit is a "0." The encoder can modify the "1" value pixels in such manner that the statistical properties of the image are not affected [5]. In many applications these techniques are not useful, since the receiver must have access to the original covers. So the need of distributing the original covers through a secure channel limits the benefits of this technique. If an attacker tampers with the stego-image by cropping, scaling or rotating, the receiver can easily detect it[2]. In some cases, if the message is encoded with error correcting information, the change can even be reversed and the original message can be fully recovered. Adopting this technique limits the hidden information capacity, since adding distortion to the cover image is the basis of embedding algorithm [8].

D.6 ***Cover Generation Techniques*** : In above steganographic methods, secret information is added to a specific cover by applying an embedding algorithm whereas some steganographic applications generate a digital object only for the purpose of being a cover for secret communication [1]. Many techniques have been proposed that encrypt messages so that they are unreadable or as secret as possible, same principle can be employed in cover creation, in which a message is converted to picture elements and then collected into a complete stego-image. Parts of the message may be destroyed or lost because of cropping, but it is still possible to recover other parts of the message by encoding the message with error correcting information. Mimic Functions proposed by Wayner is a cover generation technique proposed to hide the identity of a message by changing its statistical profile in such a way that it matches the profile of any innocent looking text[8].

D.7 ***File and Pallette Embedding:*** Different image file formats are known for having different header file structures. In addition to the data values, such as pixels, palette, and DCT coefficients, secret information can also be hidden in either a header structure or at the end of the file. For example, the comment fields in the header of JPEG images usually contain data hidden by the invisible Secrets [5].In some cases, the palette itself can be used to hide secret information. Because the order of the colors in the palette usually does not matter, the ordering of colors can be used to transfer information. In essence, a hidden message can be embedded using the difference between two colors in the palette (i.e., one secret message bit for every two colors in the palette)[4]. Color palettes are used to minimize the amount of information image that are used to represent colors. Since steganographic message within the bits of the palette and/or the indices is embedded in the palette-based steganography, one must be careful not to exceed the maximum number of colors.

D.8 LSB Steganography With AES Encryption: To implement data hiding in the least significant nibble of an image, the source image stored in a Java Buffered Image object and the message are converted to byte arrays. The length of the secret message must also be determined as it is important during the retrieval of the secret message from the image. 128 bit key AES encryption is performed using Java's cryptography libraries. The AES encrypted secret message adds another layer of security to the proposed technique. Both the message length and message data are subjected to AES encryption prior to encoding into image bytes. Nibbles from each byte of the message are extracted. The most significant nibble is extracted first, followed by the least significant nibble. The lower nibble of each image byte are then masked and the message nibble is then stored in the image nibble. This is done using the formula given below:

Image Byte = (Image Byte AND 0xF0) OR Message Nibble

Decoding of the image is done by first extracting the contents of the least significant nibbles of the first 16 image bytes. This gives us the AES encrypted length value for the secret text that is encoded into the image. This encrypted string is decrypted, and this value is then used to obtain the number of image bytes that contain the encrypted secret text.

The extraction is done in accordance to the following simple formula:

Message Nibble = Image Byte AND 0x0F

Once the contents of the image bytes are extracted, the encrypted secret text is obtained. To obtain the original text, AES decryption is performed on the string. This entire process can be demonstrated by the following flowcharts.

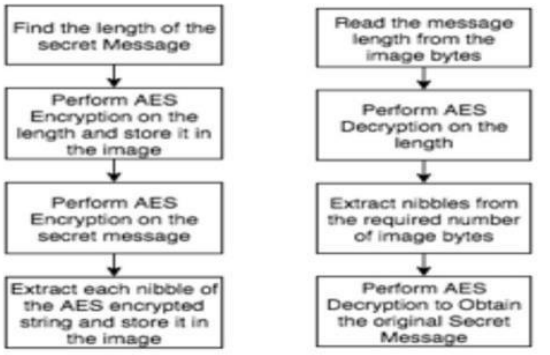

Figure 4 : Flowcharts depicting the encoding of data in images (left) and extraction of data from the images (right)

Fig. 4. Flow Chart Encoding and Decoding of Images

Here, AND and OR signify the logical AND and OR operations respectively.

V. PERFORMANCE MEASURE

As a performance measure for image distortion due to embedding, the well-known peak-signal-to noise ratio (PSNR), which is categorized under difference distortion metrics, can be applied to stego images. It is defined as [5]:

$$PSNR = 10log\left(\frac{C_{max}^2}{MSE}\right)$$

Where MSE denotes the mean square error, which is given as [5]:

$$MSE = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(S_{xy} - C_{xy})^2$$

Here, *x* and *y* are the image coordinates, M and N are the dimensions of the image, S_{xy} is the resultant stego image, and *Cxy* is the cover image. PSNR is often expressed on a logarithmic scale in decibels (dB). PSNR values below 30 dB indicate low quality and PSNR of 40 dB or higher indicates high quality stego image.

VI. CONCLUSION

In past few years we have seen an increasing interest in using images as cover media for steganographic communication. This paper presents different steganographic methods which have been proposed in the literature during the past few years. One can see that there exists a large selection of approaches to hiding information in images, with different strong and weak points respectively. Transform domain techniques make changes in the frequency coefficients instead of manipulating the image pixels directly, thus distortion is kept at minimum level and that's why they are preferred over spatial domain techniques but this advantage is achieved at the expense of payload capacity. Whereas spatial domain approaches are considered not to be robust against lossy compression and filtering but when it comes to embedding capacity, spatial domain techniques give better results. The steganography technique deployed is dependent on the type of application it is designed for means one must have the determination to compromise on some characteristics to ensure the high performance of other characteristics.

REFRENCES

[1] Zaidoon Kh. AL-Ani, A. A. Zaidan, "Overview: Main Fundamentals for Steganography", Journal Of Computing, Volume 2, Issue 3, March 2010, Issn 2151-9617.

[2] Mehdi Hussain and Mureed Hussain, "A Survey of Image Steganography Techniques", International Journal of Advanced Science and Technology Vol. 54, May, 2013.

[3] Babloo Saha and Shuchi Sharma,"Steganographic Techniques of Data Hiding using Digital Images" Defence Science Journal, Vol. 62, No. 1, January 2012, pp. 11-18.

[4] T. Morkel, J.H.P. Eloff, "An Overview Of Image Steganography" Information and Computer Security Architecture (ICSA) Research Group, University of Pretoria, 0002, Pretoria, South Africa.

[5] Nagham Hamid, Abid Yahya, "Image Steganography Techniques: An Overview", International Journal of Computer Science and Security (IJCSS), Volume 6, Issue 3, 2012.

[6] Rajarathnam Chandramouli, Mehdi Kharrazi,"Image Steganography and Steganalysis: Concepts and Practice" T. Kalker et al. (Eds.): IWDW 2003, LNCS 2939, pp. 35–49, 2004.

[7] Bin Li, Junhui He, "A Survey on Image Steganography and Steganalysis", Journal of Information Hiding and Multimedia Signal Processing, ISSN 2073-4212, Volume 2, Number 2, April 2011.

[8] Neil F. Johnson, Stefan C. Katzenbeisser, "A survey of steganographic techniques", Information hiding techniques for steganography and digital watermarking, Chapter 3.

[9] Philip Bateman, Dr. Hans Georg Schaathun, "Image Steganography and Steganalysis",Master's thesis in Security Technologies & Applications, University of Surrey,2008.

[10] C. P. Sumathi, T. Santanam, "A Study of Various Steganographic Techniques Used for Information Hiding", International Journal of Computer Science & Engineering Survey, Vol.4, No.6, December 2013.

[11] Rajesh Kumar Tiwari and Gadadhar Sahoo, "Some New Methodologies for Image Hiding using Steganographic Techniques", Department of Computer Science & Engineering, March 11, 2008.

[12] C. Vanmathi, S. Prabu, "A Survey of State of the Art techniques of Steganography", International Journal of Engineering and Technology, ISSN: 0975-4024, Vol 5 No 1 Feb-Mar 2013.

[13] Utsav Sheth and Shiva Saxena ImageSteganography Using AES Encryption and Least Significant Nibble, nternational

Conference on Communication and Signal Processing, April 6-8, 2016, India.

[14] Amritpal Singh and Harpal Singh an Improved LSB based Image Steganography Technique for RGB Images ,IEEE 2015.

[15] Chi-Yuan Lin, Kai-Ren Chen, and Jyun-Jie Wang, A Steganographic Method for Binary Embedding Using Time-Varying Convolutional Codes, 2014 International Symposium on Computer, Consumer and Control.