A Hybrid Stegano- Cryptographic Approach to Data Obfuscation Using LSB Technique

Dipta Mukherjee[1], Anandarup Mukherjee[2], Somen Nayak[3]
[1,3]Department of Computer Science Engineering
[2]Department of Electronics & Communications Engineering
University of Engineering & Management, Jaipur
India
[1]amidipta@gmail.com, [2]anandarupmukherjee@ieee.org, [3]somen.live@gmail.com,

***Abstract*—This paper proposes a hybrid approach to enhanced data security using special matrices in conjunction with steganography. The special matrices are used for mapping the actual characters to numbers which are actually pointers to the location of the characters in the matrices. The generated numbers corresponding to each character of the message are embedded in the image pixels at the least significant bit position prior to transmission. The change of LSB values of the image pixels do not distort the image significantly, allowing for undetected transmission of any message to its intended receiver.**

Keywords-Crypto-systems, Steganography, Random number, LSB Technique, Data Obfuscation

I. INTRODUCTION

Data security, in the modern world is an inherent part of our lives. It is achieved by various techniques such as passwords, cryptography, bio-metrics, steganography and other such techniques [1]. In this work a hybrid cryptographic approach is presented along with its parameters. This work discusses the importance of cryptography along with the need for our hybrid approach in enhancing cryptographic based security. Cryptography is an ancient art that has passed through many paradigms starting from simple letter substitutions, poly-alphabetic substitutions, and digital encryption to public-key cryptosystems. Cryptography is a method of saving or protecting data or information from the unauthorized access. Password based security is not foolproof because they are often easily guessed by automated programs [2]. The main use of cryptography is to prevent potential adversaries from gaining accesses to encryption technologies that might reveal important characteristics of information security products. In cryptography we cannot achieve cent percent security. The encrypted data is momentarily safe but in the long run its safety cannot be guaranteed. If we use the same technique every time, it is prone to repeated attacks if the security of the encryption is breached or hacked even once [3]. This is the main reason for the emergence of bi-fold or hybrid systems in recent times. In this paper, the technique proposed by us, randomly chooses the encryption method to encrypt data. Whenever, a hacker tries to access the encryption system, the entire process changes, preventing the hacker from guessing or breaking into the system or accessing the information. Using a random technique to encrypt our data will make the system even harder for the hacker to decrypt. If a random technique is not used, then security can be increased by increasing the key bit length resulting in increase in processing overheads. An efficient cryptographic technique must have minimum number of overheads and have a simple encryption algorithm [4]. Our aim is to develop an efficient algorithm which can be used for encrypting data with random approach although keeping the key bit length and its associated overheads manageable. Cryptography can be broadly classified into Symmetric cryptography, Asymmetric cryptography and Steganography.

A. Symmetric Cryptography:-

In this approach we use the key to encrypt the message at the sending end; the same key is used to decrypt the message at receiving end. This approach is quite fast and efficient approach in order to transmit large amount of data. Some commonly used algorithms are Data encryption standard, triple data encryption standard, etc. The major drawback of this approach is the key itself, the key has to be present within the sender and receiver. The security of the key may be compromised during transmission which results the whole method useless [5].

B. Asymmetric Cryptography:-

This cryptographic approach uses one key to encrypt data and a matching key to decrypt the data. The two different keys used in encrypting and decrypting the messages are referred to as key pairs. One of these key pairs is called the secret key or the private key, which is kept secure; the other key, called the public key is distributed in the public domain, to its intended users. For example, Public key Infra Structure, Pretty Good Services, etc. are famous Asymmetric key cryptography techniques [5].

C. Steganography:-

It is a technique by which written messages can be hidden in such a way, so as to, nobody except the sender and the corresponding receiver know the existence of message or information. Steganography, derived from Greek can be loosely termed as concealed writing or secret writing. This technique is used for security purposes where images are incorporated as the secret carrier medium to transmit the messages. The unauthorized user has no prior way to guess the image being tampered for transmitting the message within the image due to the presence of a mind boggling number of

 ISBN: 978-81-923777-9-7

similar images available online, in the public domain, rendering this method virtually hack free [6].

II. Image Information Assessment

Image Quality assessment plays a crucial role in providing metrics about an image and its derivatives. In the field of image processing one of the crucial and primary tools for providing these metrics is the Signal to Noise Ratio (SNR) which is the ratio of signal power to the noise power within that signal. Similarly, Peak Signal to Noise Ratio (PSNR) is the ratio of the maximum power of the signal to the power of the noise present in the signal. Previous studies and statistics show that if the resultant image has a PSNR value between 35dB to 50dB, it is less prone to attacks by unauthorized users. An image with PSNR in the above mentioned range appears to have more noise content than the signal itself rendering the image useless for automated decryption attacks, although this is not a rigid and only condition for considering a stego-image as un-decryptable [7].

A. SNR (signal-to-noise ratio):-

It is the ratio of signal power to the noise power. If the ratio is greater than 0dB, it indicates the presence of more signal than background noise. SNR can be represented by the relation shown in equation 1 below [2-6].

$$SNR = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \quad (1)$$

Where

$P_{signal} = Average\ power\ of\ signal$

$P_{noise} = Average\ power\ of\ noise$

Both signal and noise power are measured using the same system parameters and conditions, within the same points of the system to generate a credible SNR figure.

B. PSNR (peak signal-to-noise ratio):-

It is the ratio between the maximum possible power of the signal and the power of corrupting background noise. Peak signal to noise ratio is represented in logarithmic decibel scale. Statistically it is proved that an image with a secret message encoded into it, having a PSNR value between 35dB to 50dB is less prone to unauthorized decrypting attacks as discussed above. So, we have used this knowledge to restrict the stego-images formed, to lie within the above mentioned PSNR range. PSNR for a given signal can be calculated as given in equations 2 and 3 below. [1-7]

$$PSNR = 10.\log_{10}\left(\frac{MAX^{2}_{I}}{MSE}\right) \quad (2)$$

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j)\right]^{2} \quad (3)$$

Where,

$MSE = Mean\ Squared\ Error$

$MAX_{I} = Maximum\ possible\ pixel\ value\ of\ image$

III. Methodology

In this paper, an information hiding system has been developed using cryptography and Least Significant Bit (LSB) based steganography. Here, we have taken a random image and are using it as a carrier to hide secret messages. We name the unaltered carrier image as cover-image, while we call the transformed stego-object as the transformed-image. The implementation of the system focuses on the method of Least Significant Bit (LSB) as one of the steganographic techniques. The process of embedding data into an image requires two important things; the original image so called the cover-image which will be used for hiding the data and second is the message itself, which is the information to be hidden in the image. Here, to generate the information, a cryptographic approach is used in which there exist six different matrices

	0	1	2	3	4	5	6	7	8
0	A	B	C	D	E	F	G	H	I
1	J	K	L	M	N	O	P	Q	R
2	S	T	U	V	W	X	Y	Z	a
3	b	c	d	e	f	g	h	i	j
4	k	l	m	n	o	p	q	r	s
5	t	u	v	w	x	y	z	!	@
6	#	$	%	^	&	*	(	)	[
7	]	{	}	;	:	“”	.	?	/
8	<	>	,	+	-		‖	_	~

Figure 1: Figure denoting a sample of the first Special Matrix which is used for encrypting the secret message.

	0	1	2	3	4	5	6	7	8
0	G	O	z	t	;	)	@	J	Q
1	[	h	“”	p	l	y	X	e	d
2	%	n	B	*	F	>	m	&	A
3	<	c	{	i	g	Y	-	,	M
4	S	/	$	Z	L	.	b	C	s
5	V	(	j	W	+	_	E	‖	R
6	~	w	f	D	#	H	@	}	r
7	I	^	T	v	:	]		!	P
8	K	o	?	k	u	U	x	N	q

Figure 2: Figure denoting a sample of the fifth Special Matrix which is used for encrypting the secret message.

ISBN: 978-81-923777-9-7

inspired from Polybius matrix. But unlike the Polybius matrix, these matrices consist of 81 cells i.e., these are 9 X 9 matrices. These matrices are made up of all alphabets both upper case and lower case, along with some special characters which are used very frequently like @, #, $, %, etc. These 9 X 9 matrices are significant because they contain alpha- numeric characters as well as special characters.

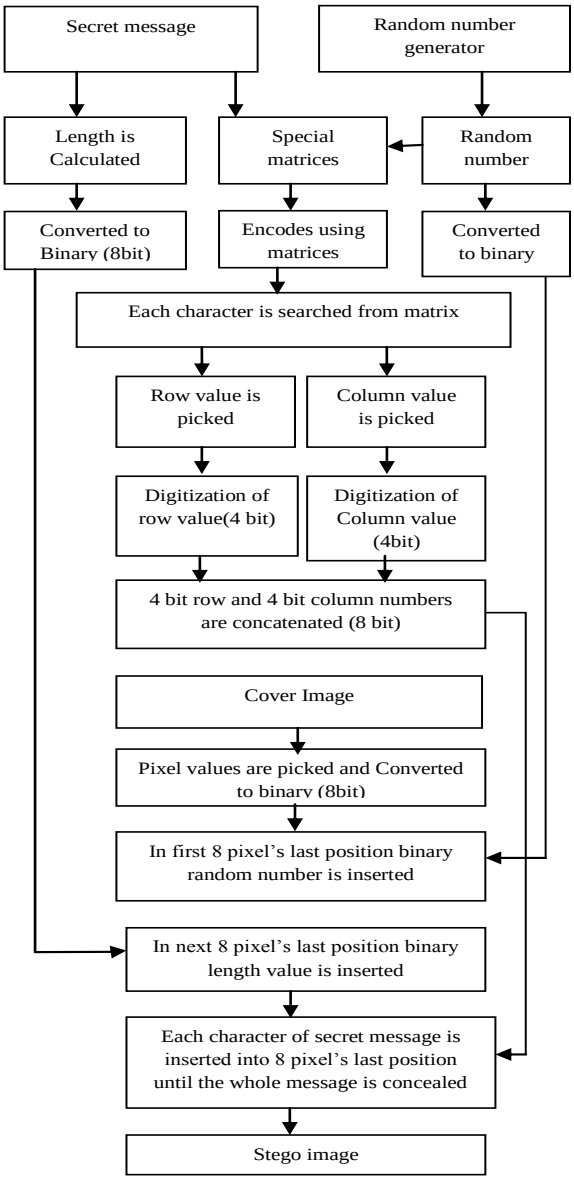

Figure 3: A flowchart depicting the encoding (encrypting) algorithm used in this paper.

A. Matrix Based Character Encryption System:-

The matrices for encrypting the data are shown in figs. 1 and 2. Whenever the system encounters a message in plain text, it encrypts the same using these matrices. When a plain text is input into the system, the algorithm counts the message length and stores the value in an array. A random number is generated using a random number generator function; depending upon that very random number any one of the above mentioned six matrices is chosen. A searching process comes into play which locates every character of the plaintext in the chosen matrix. The row and column value of each selected character is determined and converted into a four bit binary sequence corresponding to the value of the row and the column so determined. The binary value for the determined row number and the binary value for the determined column number are concatenated in the same order to generate an 8-bit binary

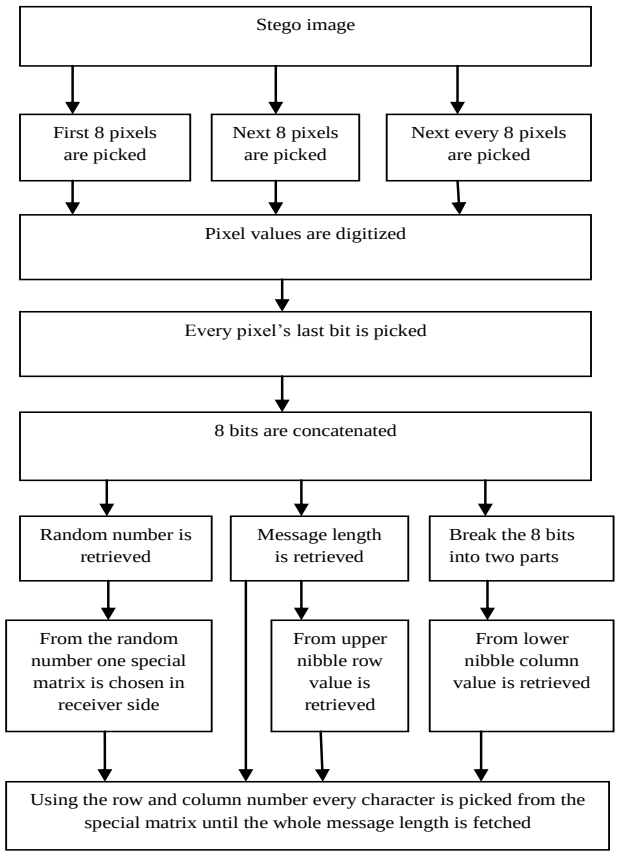

Figure 4: A flowchart depicting the decoding (decrypting) algorithm used in this paper.

sequence for each character. The sequence generated has now been successfully camouflaged as it cannot be determined without knowing the matrix being used to encode it. The matrix itself is being chosen randomly, making it virtually impossible to guess. So, each character of the message generates a unique 8-bit code.

B. Steganography Based Image Obfuscation

Prior to the embedding process, the size of image and the message must is defined by the system. This ensures that the image is able to support the message to be embedded. For example, we can loosely define a condition that the image dimension should be eight times the data we want to hide without distorting the image considerably. The final image is named and saved as "transformed.pgm". The cover-image is combined with the message to produce the above mentioned output stego image. Figure 3 illustrates the process. At first glance, the stego-image seems identical to the cover-image, albeit there are hidden messages in the image that are

ISBN: 978-81-923777-9-7

imperceptible. The simplicity of the LSB technique makes it a lucrative choice to provide an additional layer of security to this algorithm. This technique embeds the bits of the messages directly into the LSB plane of cover-image. Modulating the LSB does not result in a humanely perceptible difference because the amplitude of the change is small enough to be easily overlooked. Therefore, to the human eye, the resulting stego-image (Fig. 7) will look identical to the cover-image (Fig. 6). This allows high perceptual transparency of the LSB.

The primary stage involves picking the pixel values of the

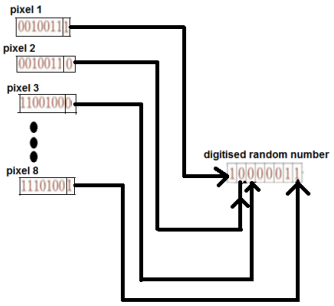

Figure 5: Figure depicting the reconstruction of the message data from the LSB values of the image pixels.

cover- image. The pixel values in any image lie in a range of 0 to 255; hence, each pixel value can be represented as an 8-bit binary number as per the following condition shown below.

$$2^n = k \tag{4}$$

Where, $n = number\ of\ bits = 8$

The next step involves digitizing the random number generated as it is the key to the whole encryption system. The key is indispensible for decoding the encrypted message at the receiver end. The consecutive step involves inserting every single bit of the digitized random number at the least significant bit position of the first 8 pixel values of the cover-image. The second stage involves inserting the length of the

Figure 6: The undistorted carrier image, known as "cover image", containing no embedded messages.

secret message in the next eight bits. This stage is followed by hiding the cipher in the same manner within the cover-image. Each character of the cipher is located from the chosen matrix selected by the random number generator. The position of each located character from the matrix is denoted by an 8-bit binary number, the first 4-bits of which denote the row value and the last 4-bits denoting the column value of the located character in the chosen matrix. Similarly, the rest of the characters are converted to binary data. Each of these bits of the secret message is embedded into the LSB position of the digitized pixels of the image. Once the data obfuscation is complete, the new stego-image named "transformed.pgm" is ready to be digitally transmitted.

Figure 7: The distorted image at the receiver's end, known as "transformed image", containing embedded messages within the image.

The reverse process is used for decoding the encrypted message from the received image. The image is read and its pixel values are gathered. The LSB values from each pixel converted to binary are stripped off (Fig. 5). Combining the first 8 stripped LSB values generates the key which was the random number used for selecting the special matrices. Similarly, the next 8 stripped LSB values give the length of the secret message. Since, the message embedded in the image is actually a reference to the position of the character in one of the special matrices; the length of the message obtained in the previous step is multiplied by 8 to indicate the last pixel of the image which contains legible information about the secret message. The bits starting from the sixteenth stripped LSB values and ending at the position indicated by the length of the message multiplied by 8 contain the reference to the position of the characters comprising of our secret message. The bits are grouped 8 at a time forming a byte. Each byte can be further divided into 2 nibbles consisting of 4 bits each; the upper nibble denoting the row value and the lower nibble denoting the column value of each character of the secret message in the chosen matrix. This process is repeated for the rest of the information.

Figure 8: A screenshot of the image pixels. The highlighted portions contain the label, key, length and the message along with the original image data.

ISBN: 978-81-923777-9-7

IV. RESULT

Figure 6 shows the cover- image without any distortions to the pixels. Figure 7 shows the transformed image with the secret message embedded into it. Figure 8 shows the fields of the pixel values of the transformed image containing the key, message length and secret message embedded along with the pixel values of the carrier image. It can be seen that there is only a slight change in the values of the pixels of the two images if we open both images in text format and sometimes there is no change at all, as the last bit of the pixels correspond to the bit values of the embedded data. The visual difference between the two images is virtually non- existent.

V. CONCLUSION

This method of obfuscation provides three layers of security to the transmission of data as compared to other steganography based approaches [7] [8]. The first layer being the random number generator, followed by the second layer of special matrices and a final layer of LSB based steganographic approach. The existence of the message is virtually untraceable to an unauthorized user. The application of brute force approaches in decoding the message is also futile against this method as it will turn up nonsensical data and garbage values. Even if the random number used for selecting the matrix is compromised, the encrypted data is still safe as the matrices are unique for each set of users hence, preventing further penetration of this system. The main handicap of this system is its point to point nature of operation. Each user pair will have their own individual matrices, which are completely different from other user pairs. This feature can be reliably employed in military communications as they mostly require a one to one approach to data communication. Ad-Hoc networks can also reliably use this mode of communication. Data loss during transmission can be checked by employing a redundancy based data check algorithm.

REFERENCES

[1] Stallings W, “Cryptography and Network Security: Principals and Practice”. Prentice-Hall, Upper Saddle River, NJ, 1999.

[2] James Goodman “An Energy-Efficient Reconfigurable Public-Key Cryptography Processor”, Member, IEEE, and Anantha P. Chandrakasan, Member IEEE, IEEE.

[3] Dr. Ekta Walia, Payal Jain, Navdeep, "An Analysis of LSB & DCT based Steganography", Global Journal of Computer Science and Technology, Page 4, Vol 10, Issue 1(Ver 1.0), April 2010.

[4] Vimal,Mahendra Kumar Patil, "Review of Various Data Hiding Algorithms in Encrypted Images", International Journal of Engineering Research and Development, e-ISSN: 2278- 067X, p-ISSN: 2278-800X, www.ijerd.com, Volume 7, Issue 6 (June 2013), PP. 42-49.

[5] Menezes AJ, van Oorschot PC and Vanstone SA, Handbook of Applied Cryptography; CRC Press, Boca Raton, FL, 1996.

[6] M.M. Amin, M. Salleh, S. Ibrahim, et al., “Information Hiding Using Steganography”, 4th National Conference On Telecommunication Technology Proceedings (NCTT2003), Shah Alam, Malaysia, pp. 21-25, January 14-15, 2003.

[7] OzturkS, Soukp_nar B. “Analysis and comparison of image encryption algorithms” Int J Inf, Technol 2004; 1(2): 64-7.

[8] B.Veera Jyothi, S.M. Verma, C. Uma Shanker, “Implementation and analysis of Email Messages Encryption and Image Steganography Schemes for Image Authentication and Verification”, International Journal of Computer Applications, Vol.5, No. 5, August 2010, ISSN 0975-8887, p:22-27.

ISBN: 978-81-923777-9-7