Volume 2, No. 2, February 2011
Journal of Global Research in Computer Science
ISSN-2229-371X
RESEARCH PAPER
Available Online at www.jgrcs.info

Data Hiding Through Multi Level Steganography and SSCE

Souvik Bhattacharyya*[1], Indradip Banerjee[2] and Gautam Sanyal[3]

*[1] Department of Computer Science and Engineering, University Institute of Technology
The University of Burdwan, Burdwan, India
souvik.bha@gmail.com[1]
[2] Department of Computer Science and Engineering, University Institute of Technology
The University of Burdwan, Burdwan, India
ibanerjee2001@yahoo.com[2]
[3] Department of Computer Science and Engineering, National Institute of Technology
Durgapur, India
nitgsanyal@gmail.com [3]

Abstract: In all over the world maintain the security of the secret data has been a great challenge. One way to do this is to encrypt the message before sending it. Encrypted messages sending frequently through a communication channel like Internet, draws the attention of third parties, hackers and crackers, perhaps causing attempts to break and reveal the original messages. Steganography is an emerging area which is used for secured data transmission over any public media. Steganography is of Greek origin and means "Covered or hidden writing". Considerable amount of work has been carried out by different researchers on steganography. In this paper, a steganographic model combining the features of both text and image based steganography technique for communicating information more securely between two locations is proposed. The authors incorporated the idea of secret key for authentication at both ends in order to achieve high level of security. As a further improvement of security level, the information has been encoded through SSCE values and embedded into the cover text using the proposed text steganography method to form the stego text. This encoding technique has been used at both ends in order to achieve high level of security.. Next the stego text has been embedded through PMM method into the cover image to form the stego image. At the receiver side different reverse operation has been carried out to get back the original information.

Keywords: Steganography, Cover Image, Cover Text, Stego Text, Stego Image, PMM (Pixel Mapping Method), SSCE (Secret Steganography Code for Embedding)).

INTRODUCTION

The term steganography is not new today. In fact several examples from the times of ancient Greece are available in Kahn [5]. In recent years, everything is trending toward digitalization and with the rapid development of the Internet technologies, digital media can be transmitted conveniently over the network. Therefore, messages need to be transmitted secretly through the digital media by using the steganography techniques. Steganography differs from cryptography in the sense that where cryptography focuses on keeping the contents of a message secret, steganography focuses on keeping the existence of a message secret [9, 25]. Another form of information hiding is digital watermarking, which is the process that embeds data called a watermark, tag or label into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object. The object may be an image, audio, video or text only [12]. Although steganography is an ancient subject, the modern formulation of it comes from the prisoner's problem proposed by Simmons [1].An assumption can be made based on this model is that if both the sender and receiver share some common secret information then the corresponding steganography protocol is known as then the secret key steganography where as pure steganography means that there is none prior information shared by sender and receiver. If the public key of the receiver is known to the sender, the steganographic protocol is called public key steganography [4, 8]. For a more thorough knowledge of steganography methodology the reader may see [7, 25]. Although all digital file formats can be used for steganography, but the image and audio files are more suitable because of their high degree of redundancy [25]. Fig. 1 below shows the different categories of file formats that can be used for steganography techniques.

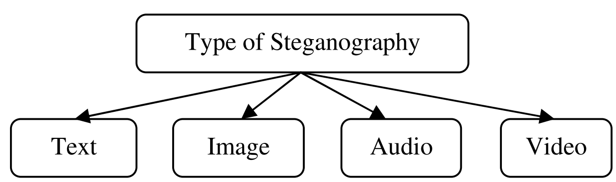

Figure 1: Types of Steganography

Among them image steganography is the most popular of the lot. In this method the secret message is embedded into an image as noise to it, which is nearly impossible to differentiate by human eyes [11, 15, 17]. In video steganography, same method may be used to embed a message [18, 24]. Audio steganography embeds the message into a cover audio file as noise at a frequency out of human hearing range [19]. One major category, perhaps the most difficult kind of steganography is text steganography or linguistic steganography [3]. The text steganography is a method of using written natural language to conceal a secret message as defined by Chapman et al. [16].

A block diagram of a generic form of steganographic system is given in Fig. 2. A message is embedded in a carrier (cover carrier) through an embedding algorithm, with the help of a secret key. The resulting stego carrier is transmitted over a channel to the receiver where it is processed by the extraction algorithm using the same key. During transmission the stego carrier, it can be monitored by unauthenticated viewers who will only notice the transmission of an image without discovering the existence of the hidden message.

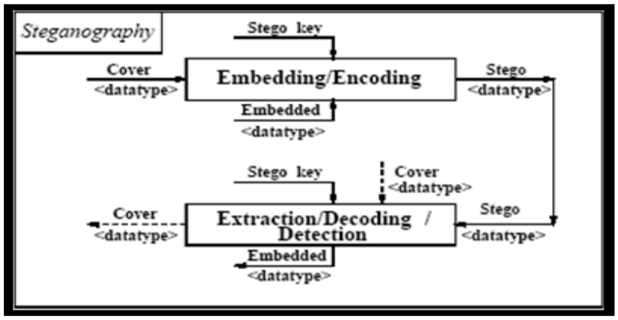

Figure 2: Generic steganographic system

This paper has been organized as following sections:- Section II discusses about some of the related works done based on text steganography and image steganography. Section III describes the SSCE method for text message encryption. Section IV describes proposed text steganography method. Section V describes the PMM method for hiding in image steganography. Section VI and VII deals with proposed data hiding model and the solution methodology, Section VIII describes different algorithms for different processes used at both at sender side and receiver side. Section IX discusses the computer algorithm. Experimental results are shown in Section X. Section XI contains the analysis of the results and Section XII draws the conclusion.

RELATED WORKS ON TEXT & IMAGE STEGANOGRAPHY

Considerable amount of work has been done both on image and text steganography.

Text steganography can be broadly divided into three types. They are format-based, random & statistical generations and Linguistic method shows in Figure 3. Most peoples have suggested various methods for hiding information in text in mentioned three categories. Some of the methods are discussed in this paper. Format-based methods use and change the formatting of the cover-text to hide the data. They don't change any words or sentences, so it does not harm the 'value' of the cover-text. A format-based text steganography method is open space method. In this method extra white spaces are added into the text to hide information. These white spaces can be added after end of each word, sentence or paragraph. A single space is interpreted as "0" and two consecutive spaces are interpreted as "1" [6]. Although a little amount of data can be hidden in a document, this method can be applied to almost all kinds of text without revealing the existence of the hidden data.

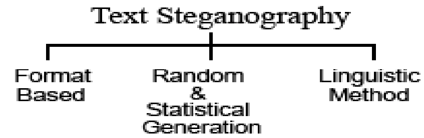

Figure 3: Three broad categories of text steganography

Another two format-based methods are word shifting and line shifting. In word shifting method, the horizontal alignments of some words are shifted by changing distances between words to embed information [21]. These changes are hard to interpret because varying distances between words are very common in documents. Another method of hiding information is, in manipulation of whitespaces between words and paragraph [27]. In line shifting method, vertical alignments of some lines of the text are shifted to create a unique hidden shape to embed a message in it [23]. Random and statistical generation methods are used to generate cover-text automatically according to the statistical properties of language. These methods use example grammars to produce cover-text in a certain natural language. A probabilistic context-free grammar (PCFG) is a commonly used language model where each transformation rule of a context-free grammar has a probability associated with it [2]. A PCFG can be used to generate word sequences by starting with the root node and recursively applying randomly chosen rules. The sentences are constructed according to the secret message to be hidden in it. The quality of the generated stego-message depends directly on the quality of the grammars used. Another approach to this type of method is to generate words having same statistical properties like word length and letter frequency of a word in the original message. The words generated are often without of any lexical value. The last category, the linguistic method considers the linguistic properties of the text to modify it. The method uses linguistic structure of the message as a place to hide information. Syntactic method is a linguistic steganography method where some punctuation signs like comma (,) and full-stop (.) are placed in proper places in the document to embed a data. This method needs proper identification of places where the signs can be inserted. Another linguistic steganography method is semantic method. In this method the synonym of words for some pre-selected are used. The words are replaced by their synonyms to hide information in it [20]. Except the above mentioned methods, there are some other methods proposed for text steganography, such as feature coding, text steganography by specific characters in words, abbreviations etc. [26] or by changing words spelling [28].

For Image Steganography Kevin Curran et al [17] propose an image based steganography methods where he describes a set of steganography methods along with their respective merits and demerits. The most common and simplest image embedding method is the least significant bit (LSB) insertion. The LSB insertion embeds the message in the least significant bit of some selected pixels of the cover image. R.Chadramouli et al. [15] gives an analysis of LSB based steganography techniques. The embedding capacity of LSB method can be increased by using two or more least significant bits. At the same time, not only the risk of making the embedded message statistically detectable increases but also the image fidelity degrades. Hence a variable-sized LSB embedding scheme is presented in [14], in which the number of LSBs used for message embedding /extracting depends on the local characteristics of the pixel. The advantages of LSB-based method are easy to implement. Unfortunately, the hidden message is assailable due to a slight modification from the active warden. Marvel et al. [11] present an image steganographic method, entitled spread spectrum image steganography (SSIS) that hides and recovers the message within digital imagery. The SSIS incorporated the use of error-control codes to correct the large number of bit errors. In recent years many image steganography models have been proposed where the main objective is to protect the transmitted data against any odd. Although increasing the security level of the hidden message of the transmitted data is still an open issue. Silvia Torres Maya et al. [22] presents a steganographic algorithm based on bit plane complexity segmentation, which permits to implement hiding information into images for its sure transmission through a non secure channel. Some Image steganographic algorithm with high security features has been presented in [29-33].

In this paper, a secret key steganographic model combining both text and image based steganography technique for communicating information more securely between two locations has been proposed which first uses a plain text as the cover data and the secret message is embedded in the cover data to form the stego text which in turn embedded into the cover image to form the stego image. The proposed text steganography scheme has been inspired by the author's previous work [35] by inserting indefinite articles '*a*' or '*an*' in conjunction with the non-specific or non-particular nouns in English language based on the mapping information according to the embedding sequence. Here data embedding in an image has been done through Pixel Mapping Method (PMM) [34].The author incorporated the idea of secret key for authentication at both ends in order to achieve high level of security. As a further improvement of security level, the secret message has been compressed and encoded through SSCE values before embedding. This work proposes a new algorithm with higher security features so that the embedded message can not be hacked by unauthorized user.

PROPOSED METHOD FOR DATA ENCODING (SSCE)

The input messages can be in any digital form and are often treated as a bit stream. The input message is first encrypted using a code generation technique SSCE [35]. For the improvement of security level, the SSCE code representation has been used to encrypt the message and then secret message has been embed to the cover text.

Secret Steganography Code for Embedding(SSCE) Table

ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE	ASCII	SSCE
10	1	1	26	2	52	3	78	4	104	5	130	6	156	7	181	8	206	9	231
20	2	11	27	12	53	13	79	14	105	15	131	16	157	17	182	18	207	19	232
30	3	21	28	22	54	23	80	24	106	25	132	26	158	27	183	28	208	29	233
40	4	31	29	32	55	33	81	34	107	35	133	36	159	37	184	38	209	39	234
50	5	41	30	42	56	43	82	44	108	45	134	46	160	47	185	48	210	49	235
60	6	51	31	52	57	53	83	54	109	55	135	56	161	57	186	58	211	59	236
70	7	61	32	62	58	63	84	64	110	65	136	66	162	67	187	68	212	69	237
80	8	71	33	72	59	73	85	74	111	75	137	76	163	77	188	78	213	79	238
90	9	81	34	82	60	83	86	84	112	85	138	86	164	87	189	88	214	89	239
100	10	91	35	92	61	93	87	94	113	95	139	96	165	97	190	98	215	99	240
110	11	101	36	102	62	103	88	104	114	105	140	106	166	107	191	108	216	109	241
120	12	111	37	112	63	113	89	114	115	115	141	116	167	117	192	118	217	119	242
130	13	121	38	122	64	123	90	124	116	125	142	126	168	127	193	128	218	129	243
140	14	131	39	132	65	133	91	134	117	135	143	136	169	137	194	138	219	139	244
150	15	141	40	142	66	143	92	144	118	145	144	146	170	147	195	148	220	149	245
160	16	151	41	152	67	153	93	154	119	155	145	156	171	157	196	158	221	159	246
170	17	161	42	162	68	163	94	164	120	165	146	166	172	167	197	168	222	169	247
180	18	171	43	172	69	173	95	174	121	175	147	176	173	177	198	178	223	179	248
190	19	181	44	182	70	183	96	184	122	185	148	186	174	187	199	188	224	189	249
200	20	191	45	192	71	193	97	194	123	195	149	196	175	197	200	198	225	199	250
210	21	201	46	202	72	203	98	204	124	205	150	206	176	207	201	208	226	209	251
220	22	211	47	212	73	213	99	214	125	215	151	216	177	217	202	218	227	219	252
230	23	221	48	222	74	223	100	224	126	225	152	226	178	227	203	228	228	229	253
240	24	231	49	232	75	233	101	234	127	235	153	236	179	237	204	238	229	239	254
250	25	241	50	242	76	243	102	244	128	245	154	246	180	247	205	248	230	249	255
		251	51	252	77	253	103	254	129	255	155								

Figure 4: SSCE Value Table

PROPOSED METHOD FOR TEXT STEGANOGRAPHY

The proposed secret-key text steganographic model has been discussed in previous work [35]. The input messages can be in any digital form and are often treated as a bit stream. The input message is first encrypted and generates the secret key, (which may be called a message enabled key). Before embedding a checking has been done to find out whether the vowels and consonants are placed in the cover text as per the grammatical order, if not place it in proper order. Secret message has been embed to the cover text by inserting indefinite articles '*a*' or '*an*' in conjunction with the non-specific or non-particular nouns in English language based on the mapping information shown in Fig 5 to form the stego text. At the receiver side other different reverse operation has been carried out to get back the original information.

Words		Bit Sequence
a	consonant	00
an	vowel	11
a	vowel	10
an	consonant	01

Figure 5: Mapping Technique

PROPOSED METHOD FOR IMAGE STEGANOGRAPHY (PMM)

In this section the authors propose a new method for information hiding within the spatial domain of any gray scale image. This method can be considered as the improved version of [34].The input messages can be in any digital form, and are often treated as a bit stream. Embedding pixels are selected based on some mathematical function which depends on the pixel intensity value of the seed pixel and its 8 neighbors are selected in counter clockwise direction. Before embedding a checking has been done to find out whether the selected embedding pixels or its neighbors lies at the boundary of the image or not. Data embedding are done by mapping each four bits of the secret message in each of the neighbor pixel based on some features of that pixel. Fig.6 shows the mapping information for embedding four bits per pixel.

MSG BIT SEQ	2nd SET – RESET BIT	3rd SET – RESET BIT	PIXEL INTENSITY VALUE	NO OF ONES(BIN)
0000	EVEN	EVEN	EVEN	EVEN
0001	EVEN	EVEN	EVEN	ODD
0010	EVEN	EVEN	ODD	EVEN
0011	EVEN	EVEN	ODD	ODD
0100	EVEN	ODD	EVEN	EVEN
0101	EVEN	ODD	EVEN	ODD
0110	EVEN	ODD	ODD	EVEN
0111	EVEN	ODD	ODD	ODD
1000	ODD	EVEN	EVEN	EVEN
1001	ODD	EVEN	EVEN	ODD
1010	ODD	EVEN	ODD	EVEN
1011	ODD	EVEN	ODD	ODD
1100	ODD	ODD	EVEN	EVEN
1101	ODD	ODD	EVEN	ODD
1110	ODD	ODD	ODD	EVEN
1111	ODD	ODD	ODD	ODD

Figure 6: Mapping Technique for data embedding

THE PROPOSED MODEL

Fig. 7 shows the block diagram of the proposed secret-key steganographic model. This input message is first converted into encrypted form using SSCE values. This encrypted message generates the secret key. The encrypted message then embedded in the cover text using the mapping technique method shown in Fig 6 to form the stego text which in turn embedded in to the cover image to form the stego image and transmit to the receiver side. At the receiver side, the stego image will be tested first for a specific feature. If that feature matches, the extraction process starts by extracting the stego text from the stego image. Next the stego text goes through the text extraction and decryption method and finally the receiver may be able to see the embedded message with the help of same secret key generated at the sender side.

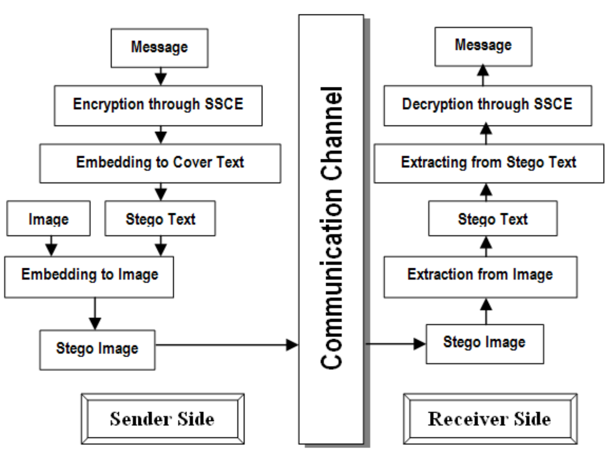

Figure 7: Proposed Steganography Model

SOLUTION METHODOLOGY

The proposed system consists of following two windows, one at the SENDER SIDE and the other at the RECEIVER SIDE.

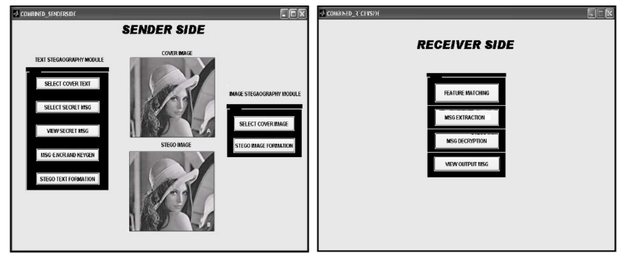

Figure 8: GUI based steganography system

Data encryption and Text steganography method has been included as an option prior to image steganography for generation of the secret key. The user should be able to select secret message as a text file, another text has to be used as the carrier (cover text) and then use the proposed encryption and text steganographic method, which will hide the selected message in the selected carrier text and will form the stego text. This stego text will be embedded in a carrier image (cover image) to form the stego image. The user at the receiver side should be able to extract the secret message from the stego image and stego text respectively with the help of different reverse process in sequential manner.

ALGORITHMS

In this section, algorithms for different processes of text and image based steganography used both in the sender side and receiver side are discussed. Fig.9 shows the algorithm of proposed system.

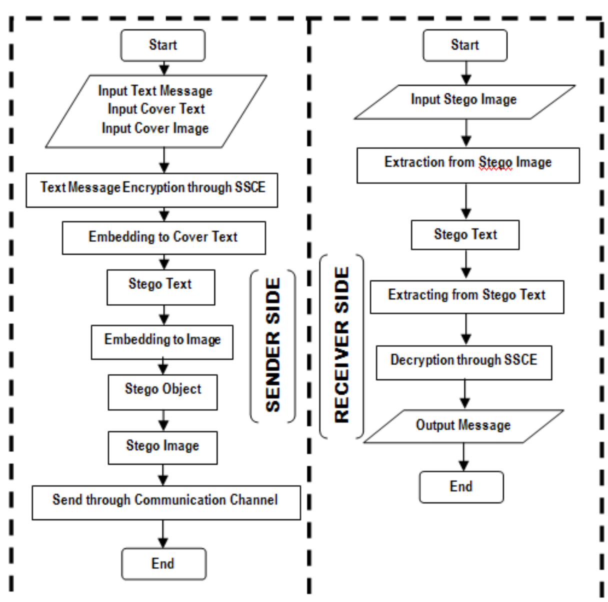

Figure 9: Proposed Algorithm for Steganographic Model

A. Algorithm for Message Encryption / Decryption

- Select the message and pick one by one character.
- Convert to its ASCII equivalent.
- Change ASCII code to our generated code from *SSCE* Table (Figure 4).
- Convert to its character equivalent.

B. Algorithm for Message Embedding for Stego Text formation

- Select the message and encrypt the message with *SSCE value.*
- Select the cover text to embed the message. Check whether the selected text is capable of embedding. If not possible repeat this step otherwise continue.
- Check the message sequence and pick first two bit sequence (MSG).
- Starting from the first word of the cover text (TX)
    - o If MSG='11' then find out the word (an) from the TX and check whether the next word's first character is vowel.
    - o Else If MSG='10' then find out the word (an) from the TX and check whether the next word's first character is vowel. Change (an) to (a).
    - o Else If MSG='01' then find out the (a) from the TX and check whether the next word's first character is consonant. Change (a) to (an).
    - o Else If MSG='00' then find out the word (a) from the TX and check whether the next word's first character is consonant.
- Repeat the above step for the remaining bit sequence of the message (two bit at a time).
- Save the embedding position in a separate file and encode it with *SSCE value* and send it to the receiver separately.

C. Algorithm for Message Extractingfrom the Stego Text

- Select the generated text (stego text) after message embedding and their positions.
- Select the embedding position in TX
    - o If there is word (an) and next word's first character is vowel, then MSG='11'
    - o Else If there is word (a) and next word's first character is vowel, then MSG='10'
    - o Else If there is word (an) and next word's first character is consonant, then MSG='01'
    - o Else If there is word (a) and next word's first character is consonant, then MSG='00'

D. Extraction of cuts of the Cover Image

Segmentation and cut extraction of the cover image is done through combining normalized cut and region growing method.

E. Algorithm of the data embedding in image through PMM

Let C be the original 8 bit gray scale image of size N x N.i.e. C = { P_{ij} I 0≤ i < N; 0 ≤ j < N; Pij є 0,1,….., 255}.Let MSG be the n bit secret message represented as MSG = {m_k I 0 ≤ k < n, m_k є 0, 1}.A seed pixel P_{rc} can be selected with row (r) and column (c). Next step is to find the 8 neighbors $P_{r'c'}$ of the pixel P_{rc} such that r' = r + j, c' = c + j ,-1≤ j ≤ 1. The embedding process will be finished when all the bits of every bytes of secret message are mapped or embedded.

- Input : Cover Image(C), Message (MSG).
- Find the first seed pixel P_{rc}.
- count = 1.
- while (count ≤ n)
- begin (for embedding message in message surrounding a seed pixel).
- m_k=Get next msg bit.
- count = count + 1.
- Mask the 5TH bit from left with the m_k in 'Bincvr'
- m_{k+1}=Get next msg bit.
- count = count + 1.
- Mask the 6TH bit from left with the m_k+1 in 'Bincvr'
- cnt=Count number of ones of one of the $P_{r'c'}$ of intensity (V).
- m_{k+2}=Get next msg bit.
- count = count + 1.
- m_{k+3}=Get next msg bit.
- count = count + 1.
- Bincvr= Binary of V.
- If(m_{k+2} = 0 & m_{k+3} = 1)
- Bincvr (zerothbit) = 0
- If(cnt mod 2 = 0)
- Bincvr(firstbit) = ¬Bincvr(firstbit)
- If(m_{k+2} = 0 & m_{k+3} = 0)
- Bincvr(zerothbit) = 1
- If(cnt / 2 ≠ 0)
- Bincvr(firstbit) = ¬Bincvr(firstbit)
- If(m_{k+2} = 0 & m_{k+3} = 0)
- Bincvr(zerothbit) = 0
- If(cnt / 2 ≠ 0)
- Bincvr(firstbit) = ¬Bincvr(firstbit)
- If(m_{k+2} = 0 & m_{k+3} = 1)
- Bincvr(zerothbit) = 1
- If(cnt mod 2 = 0)
- Bincvr(firstbit) = ¬Bincvr(firstbit)
- End
- Get the next neighbor pixel $P_{r'c'}$ for embedding based on previous $P_{r'c'}$ and repeat.
- End

F. Algorithm of the data extraction method through PMM

The process of extraction proceeds by selecting those same pixels with their neighbors. The extracting processe will be finished when all the bits of every bytes of secret message are extracted. Algorithm of the extraction method is described as:

- Input : Stego image (S) , count.
- count = count / 2.
- BinMsg= " ".
- Find the first seed pixel P_{rc}.
- I=0.
- While (count ≤ N)
- begin (for extract message in message around a seed pixel).
- Get the (First/Next) neighbor pixel $P_{r'c'}$.
- cnt=Count number of ones of one of the $P_{r'c'}$ of intensity (V).
- Bincvr= Binary of V.
- Binmsg(i)=3rd Bit of Bincvr from Right.

- i = i + 1.
- Binmsg(i)=2nd Bit of Bincvr from Right.
- i = i + 1.
- Binmsg(i)=ZerothBit of Bincvr.
- i = i + 1.
- If (cnt mod 2 = 0) (i.e. it is even) Binmsg(i)=0 Else Binmsg(i)=1.
- Binmsg(i)=Enters according to One of ones in the intensity(1 for odd ,0 for even).
- i = i + 1.
- count = count + 1.
- End.
- Get the next neighbor pixel $P_{r'c'}$ for embedding based on previous $P_{r'c'}$ and repeat.
- End loop.
- Binmsg is converted back to Original message.
- Return Original Message.
- End.

COMPUTER ALGORITHM

In this section the two algorithimic approach is discussed one for the function of the Sender Side and another for the Receiver Side.

A. Sender side

- Select the Cover Text from the set of text files.
- Select the Secret message in text form.
- Encrypt the message through SSCE and also generate the Secret key.
- Embed the encrypted form of message in to the Cover text to form the Stego text.
- Select the Cover image from the set of images.
- Check whether the image is in true color (24 bits) or in range of gray, if not error.
- Embed through PMM method the Stego Text in the cuts to generate the Stego Image

B. Receiver side

- Check the Stego Image for a specific feature and if matches continue.
- Extract the Stego text from the Stego Image.
- Extract the encrypted form of secret message from the Stego text.
- Decrypt the message with the help of the previous mentioned SSCE values / Secret key.

EXPERIMENTAL RESULTS

This section presents the obtained results via different processes mentioned in the proposed model. The authors simulated the proposed system and the results are shown in the following figures. Fig 10, 11, 12 and 13 shows the Cover Text, Secret Message to be embedded, Encrypted Message and Stego Text respectively. In this section experimental result of stego image are shown based on two well known images: Lena and Pepper. Fig. 14(A) and Fig. 14(B) shows the original cover image (Lena) and the also the same image after inserting the secret message.Fig.15 shows the same thing considering Pepper as Cover Image.

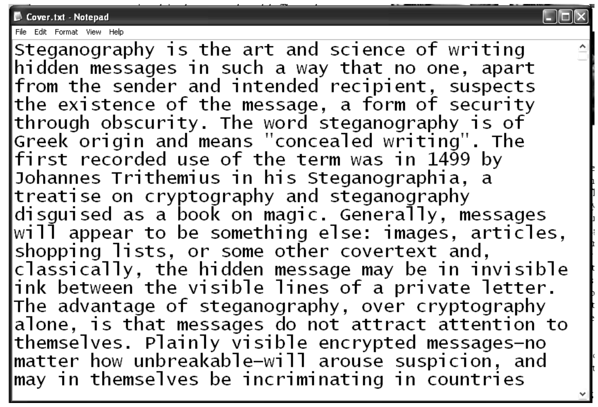

Figure 10: Cover Text

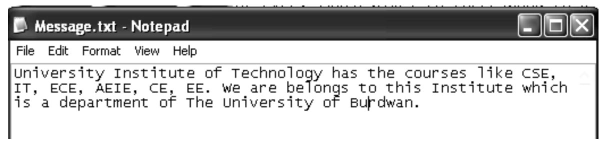

Figure 11: Message to be embedded

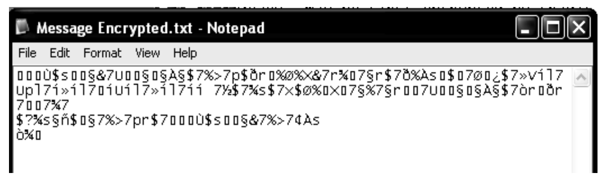

Figure 12: Encrypted Message to be embedded

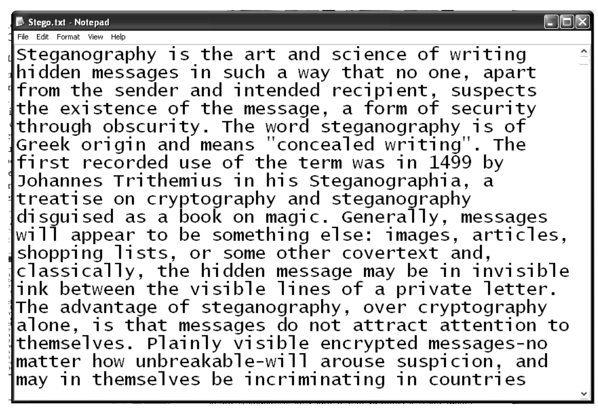

Figure 13: Stego Text

Figure 14. A) Cover Image. B) Stego Image

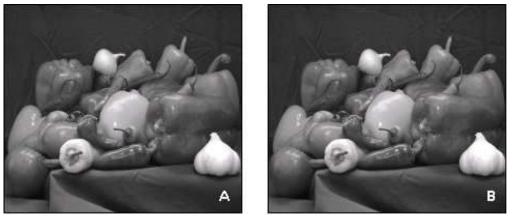

Figure 15. A) Cover Image. B) Stego Image

ANALYSIS OF THE RESULTS

In the previous work made by different researchers it has been seen some of the works has been done on text steganography and some based on image. This work proposes a novel algorithm with higher security features combining both text and image based steganographic methods to prevent the embedded message from unauthorized user. In this work an attempt has been made to increase the level of security of the steganography model by incorporating the idea of secret key along with the use of encoded form of the original message. Besides the data embedding method (PMM) used here for image steganography has been designed in such a way that it can avoid steganalysis also and it will produce a stego image with minimum degradation.

The Levels of Security incorporated in the proposed model:

- Generation of the encrypted form of the secret message.
- Embedding encrypted form of the message in cover text to form the stego text using a new proposed method.
- Embedding of the stego text through PMM in the cover image to form the stego image.
- Use of the secret key.
- Feature matching of the Stego image.
- All the processes both in sender side and receiver side need to be executed in proper sequence.

Similarity Measure of the Cover Text and Stego Text

The most familiar measure of dependence between two quantities is the Pearson product-moment correlation coefficient [37], or "Pearson's correlation." It is obtained by dividing the covariance of the two variables by the product of their standard deviations. Karl Pearson developed the coefficient from a similar but slightly different idea by Francis Galton. The Pearson correlation is +1 in the case of a perfect positive (increasing) linear relationship (correlation), -1 in the case of a perfect decreasing (negative) linear relationship (anti correlation) [37], and some value between -1 and 1 in all other cases, indicating the degree of linear dependence between the variables. As it approaches zero there is less of a relationship (closer to uncorrelated). The closer the coefficient is to either -1 or 1, the stronger the correlation between the variables. If the variables are independent, Pearson's correlation coefficient is 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables.

If we have a series of n measurements of X and Y written as x_i and y_i where i = 1,2,…,n then the sample correlation coefficient can be used in Pearson correlation r between X and Y. The sample correlation coefficient is written

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of X and Y, s_x and s_y are the sample standard deviations of X and Y. The number of matching (but different sequence order) characters divided by two defines the number of transpositions. The Correlation score of comparing cover text and stego text is 5.5460e+003(in case of long message),-611.7406 (in case of too short message), which means this method is not possible in this work.

For comparing the similarity between cover text and the stego text, the Jaro-Winkler distance for measuring similarity between two strings has been computed. The Jaro-Winkler distance [40] is a measure of similarity between two strings. It is a variant of the Jaro distance metric [38], [39] and mainly used in the area of record linkage [5] (duplicate detection). The higher the Jaro-Winkler distance for two strings is, the more similar the strings are. The score is normalized such that 0 equates to no similarity and 1 is an exact match. The Jaro distance metric states that given two strings s_1 and s_2 their distance d_j is $d_j = \frac{1}{3}\left[\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right]$, where *m* is the number of matching characters and *t* is the number of transpositions. Two characters from s_1 and s_2 respectively are considered matching only if they are not farther than $\left\lfloor \frac{\max(|s_1|,|s_2|)}{2} \right\rfloor - 1$. Each character of s_1 is compared with all its matching characters in s_2. The number of matching (but different sequence order) characters divided by two defines the number of transpositions. The Jaro score of comparing cover text and stego text is 0.9022, which means they are closely similar. Besides comparison through histogram technique has been done. It has been observed that the histogram of the cover text and the stego text is almost identical.

Similarity Measure of the Cover Image and Stego Image

For comparing the similarity between cover image and the stego image, the normalized cross correlation coefficient (r) has been computed. In statistics, correlation indicates the strength and direction of a linear relationship between two random variables. The correlation coefficient $\boldsymbol{\rho_{xy}}$ between two random variables X and Y with expected values $\boldsymbol{\mu_x}$ and $\boldsymbol{\mu_y}$ and standard deviations $\boldsymbol{\sigma_x}$ and $\boldsymbol{\sigma_y}$ is defined as:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{E((X-\mu_x)(Y-\mu_y))}{\sigma_x \sigma_y}$$

where E is the expected value operator and cov means covariance. The value of correlation is 1 in the case of an increasing linear relationship, -1 in the case of a decreasing linear relationship, and some value in between in all other cases, indicating the degree of linear dependence between the variables. Cross correlation is a standard method of estimating the degree to which two series are correlated.

Consider two series x(i) and y(i) where i=0,1,2,. . . ,N-1. The cross correlation r at delay d is defined as

$$r = \frac{\sum_i [(x(i) - mx)(y(i-d) - my)]}{\sqrt{\sum_i (x(i) - mx)^2}\sqrt{\sum_i (y(i-d) - my)^2}}$$

where mx and my are the means of the corresponding series. The cross-correlation is used for template matching which is motivated through the following formula

$$r = \sum_{\substack{x \\ y}} f(x,y)t(x-u, y-v)$$

where f is the image and the sum is over x, y under the window containing the feature t positioned at u, v. Similarity measure of two images can be done with the help of normalized cross correlation generated from the above concept using the following formula:

$$r = \frac{\sum_{(C(i,j)-m_1)(S(i,j)-m_2)}}{\sqrt{(\sum_{C(i,j)-m_1})^2}\sqrt{(\sum_{S(i,j)-m_2})^2}}$$

Here C is the cover image, S is the stego image,m_1 is the mean pixel value of the cover image and m_2 is the mean pixel value of stego image. Figure below shows the different parameters i.e. MSE, PSNR and Co-relation of the stego image of the PMM method.

Images		Data Size(in Char)					
		100	500	1000	2000	4000	5000
Lena 512X512	PSNR	63.3	57.5	54.5	51.3	47.8	46.8
	MSE	0.03	0.11	0.22	0.47	1.07	1.34
	Correlation	1.00	1.00	0.99	0.99	0.99	0.99
Lena 256X256	PSNR	57.0	50.5	47.3	44.3	41.5	40.5
	MSE	0.12	0.56	1.20	2.33	4.52	5.71
	Correlation	1.00	0.99	0.99	0.99	0.99	0.99
Lena 128X128	PSNR	51.3	44.5	41.4	38.4	35.4	N.A.
	MSE	0.47	2.29	4.67	9.21	18.4	N.A
	Correlation	0.99	0.99	0.99	0.99	0.99	N.A
Pepper 512X512	PSNR	62.6	56.2	53.2	50.1	47.1	46.3
	MSE	0.03	0.15	0.30	0.63	1.25	1.52
	Correlation	1.00	1.00	0.99	0.99	0.99	0.99
Pepper 256X256	PSNR	57.4	50.5	47.2	44.3	41.4	40.3
	MSE	0.11	0.64	1.22	2.38	4.68	5.94
	Correlation	1.00	0.99	0.99	0.99	0.99	0.99
Pepper 128X128	PSNR	50.9	44.3	41.5	38.4	35.6	34.6
	MSE	0.52	2.40	4.56	9.27	17.7	22.4
	Correlation	0.99	0.99	0.99	0.99	0.99	0.99

Fig 16: Different parameters of PMM Method

CONCLUSION

In this paper authors have used the combination of text steganography and image based steganography to obtain secure stego-image. The SSCE code used for encrypted form of the secret message in order to achieve maximum payload and increase the security level respectively. The encrypted form of the message is embedded into the cover text to form the stego text. The stego image is generated after embedding the stego text through PMM method which is a new and efficient steganographic method for embedding secret messages into images without producing any major changes between cover image and stego image. An exactly reverse procedure is followed at the receiver side to retrieve the embedded message. The integrated approach of SSCE, new method of text steganography and image based steganography using PMM has enabled the secure transfer of the message compared to earlier techniques. However to increase the security level different parameter has been considered for achieving better performance. In our next work steganalysis has also been taken care to build a commercial model.

REFERENCES

[1] Gustavus J. Simmons, "The Prisoners' Problem and the Subliminal Channel", in Proceedings of CRYPTO '83, pp 51-67. Plenum Press (1984).

[2] P. Wayner, “Strong Theoretical Steganography”, Cryptologia, XIX(3), July 1995, pp. 285-299.

[3] J.T. Brassil, S. Low, N.F. Maxemchuk, and L. O’Gorman, "Electronic Marking and Identification Techniques to Discourage Document Copying", IEEE Journal on Selected Areas in Communications, vol. 13, Issue. 8, October 1995, pp. 1495-1504.

[4] “Stretching the Limits of Steganography", RJ Anderson, in Information Hiding, Springer Lecture Notes in Computer Science v 1174 (1996) pp 39-48.

[5] Kahn, The Codebreakers - the comprehensive history of secret communication from ancient times to the Internet, Scribner, New York (1996).

[6] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding", IBM Systems Journal, vol. 35, Issues 3&4, 1996, pp. 313-336.

[7] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," Int. Conf. Computer Vision and Pattern Recognition, San Juan, Puerto Rico, June 1997.

[8] Scott Craver, "On Public-key Steganography in the Presence of an Active Warden," in Proceedings of 2nd International Workshop on Information Hiding, April 1998, Portland, Oregon, USA. pp. 355 - 368.

[9] Ross J. Anderson and Fabien A.P. Petitcolas, "On the limits of steganography," IEEE Journal on Selected Areas in Communications (J-SAC), Special Issue on Copyright & Privacy Protection, vol. 16 no. 4, pp 474-481, May 1998.

[10] N. F. Johnson and S. Jajodia, "Steganography: seeing the unseen," IEEE Computer.,Feb., 26-34 (1998).

[11] L. M. Marvel, C. G. Boncelet, Jr. and C. T. Retter, "Spread spectrum image steganography," IEEE Trans. on Image Processing, 8(8), 1075-1083 (1999).

[12] Digital Watermarking :A Tutorial Review S.P.Mohanty ,1999.

[13] J. Shi and J. Malik, "Normalized cuts and image segmentation.,"IEEE Trans. PAMI, vol. 22, no. 8, pp. 888-905, 2000.

[14] Y. K. Lee and L. H. Chen, "High capacity image steganographic model," IEE Proc.-Vision, Image and Signal Processing, 147(3), 288-294 (2000).

[15] Analysis of LSB Based Image Steganography Techniques ,R. Chandramouli, Nasir Memon, Proc. IEEE ICIP, 2001.

[16] M.Chapman, G. Davida, and M. Rennhard, “A Practical and Effective Approach to Large-Scale Automated Linguistic Steganography”, Proceedings of the Information Security Conference, October 2001, pp. 156-165.

[17] An Evaluation of Image Based Steganography Methods,Kevin Curran, Kran Bailey, International Journal of Digital Evidence,Fall 2003.

[18] G. Doërr and J.L. Dugelay, "A Guide Tour of Video Watermarking", Signal Processing: Image Communication, vol. 18, Issue 4, 2003, pp. 263-282.

[19] K. Gopalan, "Audio steganography using bit modification", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal

Processing, (ICASSP '03), vol. 2, 6-10 April 2003, pp. 421-424.

[20] M. Niimi, S. Minewaki, H. Noda, and E.Kawaguchi, "A Framework of Text-based Steganography Using SD-Form Semantics Model", Pacific Rim Workshop on Digital Steganography 2003, Kyushu Institute of Technology, Kitakyushu, Japan, July 3-4, 2003.

[21] Y. Kim, K. Moon, and I. Oh, "A Text Watermarking Algorithm based on Word Classification and Inter-word Space Statistics", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03), 2003, pp. 775–779.

[22] Silvia Torres Maya,Mariko Nakano and Ruben Vazquez Medina "Robust Steganography using Bit Plane Complexity Segmentation" 1st International Conferenceon Electrical and Electronics Engineering ,2004.

[23] A.M. Alattar and O.M. Alattar, "Watermarking electronic text documents containing justified paragraphs and irregular line spacing ", Proceedings of SPIE - Volume5306, Security, Steganography, and Watermarking of Multimedia Contents VI, June 2004, pp- 685-695.

[24] G. Doërr and J.L. Dugelay, "Security Pitfalls of Frameby-Frame Approaches to Video Watermarking", IEEE Transactions on Signal Processing, Supplement on Secure Media, vol. 52, Issue 10, 2004, pp. 2955-2964.

[25] T Mrkel,JHP Eloff and MS Olivier ."An Overview of Image Steganography,"in proceedings of the fifth annual Information Security South Africa Conference ,2005

[26] M.H. Shirali-Shahreza and M. Shirali-Shahreza, "Text Steganography in Chat", Proceedings of the Third IEEE/IFIP International Conference in Central Asia on Internet the Next Generation of Mobile, Wireless and Optical Communications Networks (ICI 2007), Tashkent, Uzbekistan, September 26-28, 2007.

[27] L.Y. Por and B. Delina, "Information Hiding: A New Approach in Text Steganography", 7th WSEAS International Conference on Applied Computer & Applied Computational Science, April 2008, pp- 689-695.

[28] MohammadShirali-Shahreza: "Text Steganography by Changing Words Spelling" at ICACT 2008.

[29] "Study of Secure Steganography model" by Souvik Bhattacharyya and Gautam Sanyal at the proceedings of "International Conference on Advanced Computing & Communication Technologies (ICACCT-2008),Nov, 2008, Panipat, India"

[30] **"**An Image based Steganography model for promoting Global Cyber Security" by Souvik Bhattacharyya and Gautam Sanyal at the proceedings of "International Conference on Systemics,Cybernetics and Informatics (ICSCI-2009),Jan, 09,Hyderabad,India."

[31] **"**Implementation and Design of an Image based Steganographic model" by Souvik Bhattacharyya and Gautam Sanyal at the proceedings of " IEEE International Advance Computing Conference "(IACC-2009)"

[32] A Novel Approach to Develop a Secure Image based Steganographic Model using Integer Wavelet Transform" at the proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing (ITC 2010)" by Souvik Bhattacharyya, Avinash Prasad Kshitij and Gautam Sanyal. (Indexed by IEEE Computer Society).

[33] A Steganographic Method for Images using Pixel Intensity Value (PIV))" by Souvik Bhattacharyya and Gautam Sanyal at the proceedings of National Conference on Computing & Systems 2010 held at The University of Burdwan in January 2010.

[34] "Hiding Data in Images Using Pixel Mapping Method (PMM) by Souvik Bhattacharyya and Gautam Sanyal accepted as a regular research paper at SAM'10 - 9th annual Conference on Security and Management under The 2010 World Congress in Computer Science, Computer Engineering, and Applied Computing to be held on July 12-15, 2010, USA (The proceedings will be indexed in Inspec / IET / The Institute for Engineering and Technology; DBLP / Computer Science Bibliography, and others.)

[35] "Novel text steganography through special code generation" by Indradip Banerjee, Souvik Bhattacharyya and Gautam Sanyal at the proceedings of International Conference on Systemics, Cybernetics and Informatics (ICSCI-2011), Hyderabad, India. in January 5-8, 2011.

[36] "Design and implementation of a secure text based steganography model" by Souvik Bhattacharyya, Indradip Banerjee and Gautam Sanyal at the Proceedings of 9th annual Conference on Security and Management (SAM) under The 2010 World Congress in Computer Science,Computer Engineering and Applied Computing(WorldComp 2010), LasVegas,USA, July 12-15,2010.

[37] S. Dowdy and S. Wearden. Statistics for research. Wiley. ISBN 0471086029, page 230, 1983.

[38] Kran Bailey Kevin Curran. An evaluation of image based steganography methods. 1999.

[39] Kran Bailey Kevin Curran. An evaluation of image based steganography methods. International Journal of Digital Evidence,Fall 2003, 2003.

[40] G. Doerr and J.L. Dugelay. A guide tour of video watermarking. Signal Processing: Image Communication., 18:263–282, 2003.

ABOUT THE AUTHORS

Souvik Bhattacharyya received his B.E. degree in Computer Science and Technology from B.E. College, Shibpur, India, presently known as Bengal Engineering and Science University (BESU) and M.Tech degree in Computer Science and Engineering from National Institute of Technology, Durgapur, India. Currently he is working as an Assistant Professor in Computer Science and Engineering Department at University Institute of Technology, The University of Burdwan. He has a good no of research publication in his credit. His areas of interest are Natural Language Processing, Network Security and Image Processing.

Indradip Banerjee received his MCA degree from IGNOU in 2009, PGDCA from IGNOU in 2008, MMM from Annamalai University in 2005 and BCA (Hons.) from The University of

Burdwan in 2003. Currently he is working as a Technical Assistant in Computer Science and Engineering Department at University Institute of Technology, The University of Burdwan. His areas of interest are Network Security and Image Processing.

Gautam Sanyal has received his B.E and M.Tech degree National Institute of Technology (NIT), Durgapur, India. He has received Ph.D (Engg.) from Jadavpur University, Kolkata, India, in the area of Robot Vision. He possesses an experience of more than 25 years in the field of teaching and research. He has published nearly 50 papers in International and National Journals / Conferences. Two Ph.Ds (Engg) have already been awarded under his guidance. At present he is guiding six Ph.Ds scholars in the field of Steganography, Cellular Network, High Performance Computing and Computer Vision. He has guided over 10 PG and 100 UG thesis. His research interests include Natural Language Processing, Stochastic modeling of network traffic, High Performance Computing, Computer Vision. He is presently working as a Professor in the department of Computer Science and Engineering and also holding the post of Dean (Students' Welfare) at National Institute of Technology, Durgapur, India.