A Novel Technique for Secure, Lossless Steganography with Unlimited Payload

Rahna E. and V. K. Govindan

***Abstract*—Steganography is a way of sending hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message. There exist many methods for digital image steganography. But most of the existing techniques are based on lossy approach. And the major challenges of steganography are security of hidden communication and the size of message that can be embedded in an image. So, this paper is intended to propose an image steganography technique based on matches between cover image and secret data. This proposed method maintains the cover image as such and has unlimited capacity of payload.**

***Index Terms*—Lossless steganography, secure steganography, search partition size, unlimited payload.**

I. Introduction

The difficulties in ensuring individual's privacy become progressively challenging with advancements in digital technologies of communication and the growth of computer power and storage. Different persons will appreciate different degrees of privacy. To protect personal privacy, various methods have been investigated and developed. Encryption is probably the most obvious one, and next comes steganography. Encryption is adaptable to noise and is generally observed whereas steganography is not.

Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity [1]. The word "Steganography" is of Greek origin and means "concealed writing". The main aim of steganography is to hide the existence of the message in the cover medium.

Cryptography and steganography are cousins in the spy craft family [2]. Cryptography scrambles a message with the help of certain cryptographic algorithms for converting the secret data into unintelligible form. On the other hand, steganography hides the message in cover image so that it becomes invisible. Sending a message in the form of cipher text might arouse suspicion on the part of the recipient whereas an "invisible" message created with steganographic algorithms will not. Anyone who needs to perform secret communication can use cryptographic algorithms to scramble the data before performing steganography to achieve additional security. The purpose of steganography is defeated once the presence of secret data is revealed or even suspected, even if the message is not extracted or deciphered.

For a steganography algorithm, a cover image is given or chosen, and the embedding process generates a stego-image using stego-key. The extraction method takes the stego image and applies the inverse algorithm using the shared key to extract the hidden message [3].

A. Challenges

The major challenges of steganography are [4]:

1) Security of Hidden Communication: The hidden contents must be invisible both perceptually and statistically so as to avoid the suspicions of eavesdroppers.
2) Size of Payload: Steganography requires sufficient embedding capacity.

Requirements for higher payload and secure communication are often contradictory. Depending on the specific application scenarios, a tradeoff has to be sought.

B. Applications

Steganography can be used when we need to hide data [5]. The main reason for hiding data is to prevent unauthorized persons from being aware of the existence of a message. Steganography can be used to hide secrets of a company or plans of a new invention. With the help of steganography, we can send out trade secrets without anyone at the company being aware and hence prevents corporate espionage. Steganography can also be used in the non-commercial sector for number of purposes such as secret data hiding and copyright protection.

The rest of this paper is organized as follows: Section II briefly presents the literature survey of the major techniques employed emphasizing the approaches used and the pros and cons. The proposed algorithm of steganography is described in Section III. The Section IV presents the experimental results and analysis, and finally the paper is concluded in Section V highlighting the advantages and issues.

II. Literature Survey

Steganography is an active field of research; many attempts are already been done. Most of them are based on LSB based lossy techniques. This section briefly reviews some of the major work in this topic of research.

Patel and Dave [6] have proposed a new variant of LSB based image steganography. In this, both the parties will have to agree upon a set of carrier images and certain required parameters. Then the sender will select an image, from the set of carrier images which requires least number of bit manipulations on LSB substitution of secret data, and produce stegoimage. Then the receiver on receiving stegoimage will extract LSBs along with the help of the received parameters. The probability of guessing parameters is very less. So extraction without those parameters is very

Manuscript received February 15, 2013; revised April 3, 2013.

The authors are with Department of Computer Science and Engineering, NIT Calicut, Calicut, India (email: rahnaravindran@gmail.com, vkg@nitc.ac.in).

DOI: 10.7763/IJFCC.2013.V2.243

difficult. Here since both the parties agree upon a set of carrier images the visual difference between stegoimage and original image can be reduced.

Johri and Asthana [7] proposed a steganography technique in which data is embedded using alteration component technique. In this, key and secret message will replace each pixel. Then for the security of stegoimage palette based image technique is applied by stretching process. The receiver having the same secret key applies destretching palette process on stegoimage using alteration component extraction process to extract the data. This technique has higher capacity and better imperceptibility.

Swati and Mahajan [8] proposed a secure image steganographic model using RSA algorithm and LSB insertion. In this method, the secret data is first encrypted using recipient's RSA public key. Then each bit of the encrypted message is inserted to the LSBs of image in different images so as to find the best cover image. Best cover image is the one which requires minimum number of LSB changes. The receiver on receiving the stegoimage will extract the message in the encrypted form and will decrypt it using private key.

Lisa M. Marvel *et al.* [9] presented an embedding method, called Spread Spectrum Image Steganography (SSIS). In this method, the data to be hidden is first encoded and a spreading sequence is generated using a wideband pseudorandom noise generator. Then the modulation scheme is used to spread the narrowband spectrum of encoded image with the spreading sequence, thereby composing the embedded signal, which is then input into an interleaver and spatial spreader. The output signal is then combined with the cover image to get the stegoimage.

Hassan Mathkour *et al.* [10] proposed a technique which emphasizes undetectability. It allows for the change of intensity of image planes of (24 bit) colored image to embed secret message in a specific distance between them. It is based on changing the distance of two random selected pixel channels in a specific range that represent hidden data.

Han-ling Zhang *et al.* [11] presented an approach which is based on pixel value differencing. It makes use of the largest difference value between the three pixels nearer to the target pixel to calculate how many secret bits will be embedded into the pixel. In order to enhance the image quality of the stego-image, they applied optimal pixel adjustment process (OPAP).

Ching-Yu Yang [12] proposed a steganography method based on the module substitutions. The secret bits to be embedded in the block are first determined by the base-value (BV) of the block in R-, G-, and B-component of a RGB trichromatic system. Then, the data bits are embedded in each component respectively by Mod u, Mod u-v, and Mod u-v-w module substitutions.

Piyush and Paresh [13] presented a technique that combines the features of cryptography, steganography along with multimedia data hiding. In order to provide higher security levels the algorithm uses a reference database. In this method, they first encrypted the message using DES. And then the cipher is saved in the image using a modified bit encoding technique. For each byte of data one cover pixel will be edited.

Hassan Mathkour *et al.* [14] and Masud Karim et al. [15] proposed variations of LSB substitutions. In the former method key idea was to divide the image into many segments and apply a different processing on each segment. Whereas in the latter one, data is encrypted using a key and is replaced with the LSB of RGB color image. And the length of the hidden message is stored in the 1st row of stego image.

Subba Rao *et al.* [16] presented an image steganography technique that randomizes the sequence of cipher bits. They computed the suitability measure of the various random sequences of the cipher bits against a given image and select the random sequence closest to the image. Then they generated those random sequences by the use of an L.F.S.R. They then embed these random sequences of cipher bits in the image.

Velagalapalli *et al.* [17] proposed a technique known as SteganPEG to hide data in jpeg images. They perform JPEG compression on the data to be hidden. This method uses a new cryptography technique known as 'Rotatocrypt' to encrypt or decrypts data using rotations. A list called 'PassStore' is created from the password used. Then encryption is done by right rotating the bits as guided by the value in PassStore.

Debnath Bhattacharyya *et al.* [18] presented a discrete fourier transformation based Image authentication technique. In this technique they selected 2 x 2 windows for better result of authentication. For achieving more security, insertion and extraction is done in frequency domain rather than on spatial domain. In this they first took 2X2 window of cover image in sliding window manner and applied DFT. Then they replaced the LSB of DFT component by the data bit and applied Inverse DFT.

LIU Tong and QIU Zheng-ding [19] and Vladimir Banoci *et al.* [20] proposed a DWT based color image steganography method. In the former method the secret information is hidden into a publicly accessed color image by a quantization-based strategy. Whereas, the latter case method processes grey scale images as cover object for creating subliminal channel and it utilizes transform coefficients of 2-Dimensional Discrete transform for embedding process.

Mohammad and Adnan proposed [21] an algorithm which uses actual color value of a pixel to determine the number of bits stored in each channel (R, G or B) of that pixel. In this, one of the channels is selected randomly as indicator. Data will be stored in the least significant bits of the channel, having lowest color value among the two channels other than the indicator.

Raja *et al.* [22] proposed a method which is based on a technique that combines Least Significant Bit (LSB), Discrete Cosine Transform (DCT), and quantization. Run length coding algorithms are also used for compressing the stego-image to enhance its security.

Mohammad [23] proposed a technique for steganography in MMS. In this, he hides the data in two media, text and image, so this is more resistant. The approach followed is that, the data is first divided into two parts. The hiding capacity of the text and the image will decide the size of each part. Then hides the first bit in the text. After that the next 5 bits are hidden in the image, and then the 7th bit in the text. After that the next 5 bits are hidden in the image. This hiding process is repeated until the end of data is reached.

The survey carried out reveals that the main issues yet to be further addressed in the field of steganography are payload limitation, quality of stegoimage and the concern of security. We need to develop steganography techniques where we can embed data equal or more than the size of cover image and without any distortion in stego image so that the security of the message is enhanced. In this paper, we propose a method that, instead of substitutions and bit manipulations, makes use of the matches between partitions of the message with the cover image data.

III. Proposed Method

In case of LSB or any other bit substitutions, we have to modify the cover image. Though the change is invisible to human vision system it might be visible to some other visions. So we can go for a system which will not even change a bit of the cover image.

The proposed algorithm tries to find exact matches of message partitions in the image data. These locations of matches and the size of each of the matches form the key to recover the message from the stegoimage. The two procedures required for this purpose, the *embedding* and the *extraction* of messages are given below:

A. Embedding Procedure

Input: Cover image and secret data

Output: Stegoimage

Convert the cover image and secret data into 2 strings, *Str1* and *Str2* respectively.

Fix the number of characters in each partition of *Str2* as 8. (We can also have partitions of size 4, 16, 32 and so on.)

For each partition of *Str2,* search for a match in *Str1*.

If no match is found, divide the partition into partitions of size 6, size 4 and so on until we get a match in *Str1*.

When a match is obtained that location of *Str1* and the size of the match string are saved in a header

Compress the header to obtain a key which is send to the receiver of the message through a secure channel.

Cover image is send to the receiver in plain.

B. Extracting Procedure

Input: Stegoimage

Output: Secret data

Convert the stego image into a string *Str*.

Decode the key to get the uncompressed header consisting of the location index and the size of match.

From the String, *Str*, read the match size from each location index and build the secret message.

IV. Experimental Results and Analysis

By using the proposed algorithm we have hidden secret data into images, Fig. 1 and Fig. 2. Table I and table II give the number of locations for various cases of partition sizes, in the search, used for hiding messages of different sizes. We employed partition sizes of 4, 8 and 16 characters. It is found that the size of the header is less in case of partitions of size 16 and is almost same as in case of partitions of size 8. This implies that the 16 character matches are almost zero. It is higher for 4 sized partitions. We have also seen that the number of locations increases as the size of secret data increases in all of the cases.

From the above two tables, we can see that the number of locations of matches of message partitions are somewhat proportional to the message sizes. Most of the existing steganographic methods' performance is analyzed on the basis of histogram similarity and Peak Signal to Noise Ratio (PSNR) between cover image and the stegoimage. Here, since the stegoimage and cover image is the same; the histograms will be identical and the PSNR value will be infinite, and hence such an analysis is not required in this case.

Fig. 1. Original image & stegoimage (micky.bmp)

TABLE I: Number of Locations Used for Hiding Messages of Various Sizes (in Characters), in 'Micky.Bmp'.

Partition Size	196	1386	4987
4 sized partition	50	348	1227
8 sized partition	42	258	905
16 sized partition	42	257	905

Fig. 2. Original image & stegoimage (nature.bmp)

TABLE II: Number of Locations Used for Hiding, Messages of Various Sizes (in Characters), in 'Nature.Bmp'.

Message Size →	196	1386	4987
4 sized partition	50	355	1250
8 sized partition	39	268	909
16 sized partition	39	266	909

One of the major issues with this approach is that if any of the message symbols is not present in the image, we have to make the partition size less than a character size. This makes the secret key much longer than message size. This can be resolved by choosing the cover image as the one which contains all the message characters.

A second issue is the length of the key which is normally somewhat proportional to the message key. We have to employ some type of lossless compression technique to reduce the size of the key.

V. Conclusion

The ultimate aim of steganography is to hide the very existence of message in the cover medium. There are a number of methods suggested by various researchers attempting to achieve this goal of hiding the messages securely in the cover images. Most of the approaches in the literature surveyed are based on LSB manipulation and their variant. The major issues still unresolved are the payload limitation, quality of stegoimage and the lack of security. In this paper, we have proposed a method that looks for exact matches between partitions of the message with the cover image data. The main advantage of the proposed approach is that the stegoimage will be the cover image itself; we are not even changing a minute portion of the cover image. And also, the payload capacity is very high; it can be higher than the cover image, no limit. A major issue in this technique is that the sizes of secret key which can be even larger than that of the messages in some of the cases of cover images. Otherwise, this technique is a highly secure lossless robust steganography technique unlike the other lossy LSB techniques in the literature. This problem may be eliminated by the choice of cover image appropriate for the message.

References

[1] Wikimedia Foundation. [Online]. Available: http://en.wikipedia.org/wiki/Steganography.

[2] N. F. Johnson and S. Jajodia, “Steganography: Seeing the Unseen,” *IEEE Computer*, 1998, pp. 26-34.

[3] A. Al-Mohammad, “Steganography-Based Secret and Reliable Communications Improving Steganographic Capacity and Imperceptibility,” School of Information Systems, Computing and Mathematics, 2010.

[4] P. Goel., “Data Hiding in Digital Images: A Steganographic Paradigm,” PhD Thesis, Indian Institute of Technology, Kharagpur, 2008.

[5] R. Riasat, I. S. Bajwa, and M. Z. Ali, “A Hash-Based Approach for Colour Image Steganography,” in *Proc. IEEE International Conference on Computer Networks and Information Technology*, 2011, pp. 303-307.

[6] H. J. Patel and P. K. Dave, “Least Significant Bits Based Steganography Technique,” in *Proc. IJECCE 2012*, vol. 3, pp. 97-103.

[7] S. Johri., “An Adaptive Steganography Technique for Gray and Colored Images,” *Journal of Global Research in Computer Science*, vol. 3, pp. 41-45, 2012.

[8] S. Tiwari, R. P. Mahajan, and N. Shrivastava, “Steganography-an Approach for Data Hiding Based on Encryption and Lsb Insertion,” *IJECCE*, vol. 3, pp. 76-83, 2012

[9] L. M. Marvel, C. T. Retter, and C. G. J. Boncelet, “A Methodology for Data Hiding Using Images,” in *Proc. IEEE Military Communications Conference*, 1998, vol. 3, pp. 1044-1047.

[10] H. Mathkour, B. Al-Sadoon, and A. Touir, “A New Image Steganography Technique,” in *Proc. IEEE 4th International Conference on Wireless Communications, Networking and Mobile Computing*, 2008, pp. 1-4.

[11] H. Zhang, G. Geng, and C. Xiong, “Image Steganography Using Pixel-Value Differencing,” *IEEE Second International Symposium on Electronic Commerce and Security*, 2009, pp. 109-112.

[12] C. Y. Yang, “Color Image Steganography Based on Module Substitutions,” in *Proc. IEEE Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2007, vol. 2, pp. 118-121.

[13] P. Marwaha, “Visual Cryptographic Steganography in Images,” in *Proc. IEEE International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2010, pp. 1-6.

[14] H. Mathkour, G. M. R. Assassa, A. A. Muharib, and I. Kiady, “A Novel Approach for Hiding Messages in Images,” in *Proc. IEEE International Conference on Signal Acquisition and Processing*, 2009, pp. 89-93.

[15] S. M. Masud Karim, M. S. Rahman, and M. I. Hossain, “A New Approach for LSB Based Image Steganography Using Secret Key,” in *Proc. 14th International Conference on Computer and Information Technology*, 2011, pp. 286-291.

[16] Y. V. Rao, S. S. Rao, and N. R. Rekha, “Secure Image Steganography Based on Randomized Sequence of Cipher Bits,” in *Proc. IEEE Eighth International Conference on Information Technology: New Generations*, 2011, pp. 332-335.

[17] V. L. Reddy, A. Subramanyam, and P. C. Reddy, “SteganPEG Steganography+ JPEG,” in *Proc. IEEE International Conference on Ubiquitous Computing and Multimedia Applications*, 2011, pp. 42-48.

[18] D. Bhattacharyya, J. Dutta, P. Das, R. Bandyopadhyay, S. K. Bandyopadhyay, and T. Kim, “Discrete Fourier Transformation Based Image Authentication Technique,” in *Proc. 8th IEEE International Conference on Cognitive Informatics*, 2009, pp. 196-200.

[19] T. Liu and Z. Qiu, “A DWT-Based Color Image Steganography Scheme,” in *Proc. IEEE, 6th International Conference on Signal Processing*, 2002, vol. 2, pp. 1568-1571.

[20] V. Banoci, G. Bugar, and D. Levicky, “A Novel Method of Image Steganography in DWT Domain,” in *Proc. IEEE 21st International Conference on Radioelektronika*, 2011, pp. 1-4.

[21] L. M. Marvel, C. T. Retter, C. G. Jr Boncelet, “A Methodology for Data Hiding Using Images,” in *Proc. IEEE Military Communications Conference*, 1998, vol. 3, pp. 1044-1047.

[22] M. T. Parvez and A. A. A. Gutub, “RGB Intensity Based Variable-Bits Image Steganography,” in *Proc. IEEE Asia-Pacific Services Computing Conference*, 2008, pp. 1322- 1327.

[23] H. Zhang, G. Geng, and C. Xiong, “Image Steganography using Pixel-Value Differencing,” in *Proc. IEEE Second International Symposium on Electronic Commerce and Security*, 2009, pp. 109-112.

Rahna E. is currently doing her final semester of M Tech in computer science and engineering in the National Institute of technology Calicut. She has received Bachelor’s degree in computer science and engineering from AWH Engineering College (University of Calicut) in the year 2010. She was born in Calicut, Kerala on 30th October 1988.

V. K. Govindan received Bachelor’s and Master’s degrees in electrical engineering from the National Institute of technology Calicut in the year 1975 and 1978, respectively. He was awarded PhD in Character Recognition from the Indian Institute of Science, Bangalore, in 1989. His research areas include Image processing, pattern recognition, data compression, document imaging and operating systems. He has more than 100 research publications in international journals and conferences, and authored ten books. He has produced six PhDs and reviewed papers for many Journals and conferences. He has more than 34 years of teaching experience at UG and PG levels and he was the Professor and Head of the Department of Computer Science and Engineering, NIT Calicut during years 2000 to 2005. He is currently working as Professor in the Department of Computer Science and Engineering, and Dean Academic at National Institute of Technology Calicut, India. . He was born in Malappuram, Kerala on 30th April 1950.