1. Rajdeep BORGOHAIN

DATA HIDING TECHNIQUES USING NUMBER DECOMPOSITIONS

1. DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, DIBRUGARH UNIVERSITY INSTITUTE OF ENGINEERING AND TECHNOLOGY, DIBRUGARH, INDIA

ABSTRACT: Data hiding is the art of embedding data into digital media in a way such that the existence of data remains concealed from everyone except the intended recipient. In this paper, we discuss the various Least Significant Bit (LSB) data hiding techniques. We first look at the classical LSB data hiding technique and the method to embed secret data into cover media by bit manipulation. We also take a look at the data hiding technique by bit plane decomposition based on Fibonacci numbers. This method generates more bit planes which allows users to embed more data into the cover image without causing significant distortion. We also discuss the data hiding technique based on bit plane decomposition by prime numbers and natural numbers. These methods are based on mapping the sequence of image bit size to the decomposed bit number to hide the intended information. Finally we present a comparative analysis of these data hiding techniques.
KEYWORDS: Data Hiding, Least Significant Bit, Steganography, Steganalysis

INTRODUCTION

Sharing of digital media is one of the main reasons for the boom of the internet. Users can share between them various information in digital format. But with the increasing threat of security [14] in the cyber world, it has become very important that any important information shared through these digital media be concealed in such a way that only the intended recipient will be able to retrieve the information. In this context data hiding in the form of steganography plays a very important role. The word steganography is derived from Greek which means "Concealed Writing" [15]. The intended data is embedded in an image in such a way that no one but the recipient can extract the hidden message.

For data hiding, the requirements [5] are:

- A medium to hold the hidden information which is known as cover media.
- The secret message which the sender intends to embed within the cover media.
- A steganography function and its inverse.
- An optional key (stego-key) that is used to hide or unhide the data.

For hiding the data, the cover media can be any one of the following digital media [6]:

- Plain Text
- Digital Image
- Audio
- Video
- IP Datagram

The process of hiding data, we first take the cover media which can be of any format as mentioned above.

After that we take the secret information which we want to hide in the cover media. The secret message can be in any form, plain text or cipher text or even other image file.

The cover media and secret message is then passed through a steganography function and the message is embedded in the cover media. We can also optionally use a key to hide the message and later decode it using the same key [13].
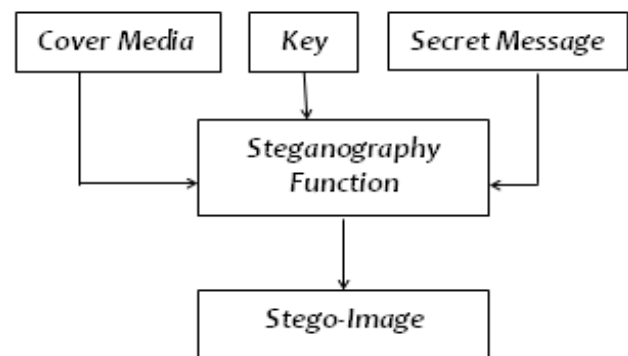

Figure 1. Steganography Process

For embedding data in digital media, two domains are generally considered, spatial domain [16, 21] and the transform domain [17, 22]. In the transform domain, the domains considered are Discrete Cosine Transform (DCT) [7], Discrete Fourier Transform (DFT) [23] and Discrete Wavelet Transform (DWT) [8].

Both the spatial and transform embedding scheme has to fulfill three requirements of visibility, robustness and capacity. Visibility is concerned with the fact that human observers should not be able to detect distortions due to the hidden message in the digital media. On the other hand, it is very important that once a secret message is inserted, it becomes impossible to delete or manipulate that message. This is the requirement for robustness. Moreover, the capacity of the digital media should be kept in consideration while inserting the message. These three factors are dependent on each other and a balance must be maintained between them as

increase in one of the factors leads to the decrease in other [9].

Though, there are many data hiding techniques [2, 3, 4], in this paper, we consider the spatial domain of data hiding. We discuss the various data hiding techniques using bit manipulation of the lowest significant bit (LSB). We take a look at how the bit planes can be increased by various number decomposition methods without compromising on the three requirements of visibility, robustness and capacity.

The rest of the paper is organized as follows. In Section 2 we take a look at the classical LSB data hiding technique. Here we discuss how secret text can be embedded in cover image using LSB manipulation. Section 3 of the paper contains the Fibonacci LSB data hiding technique. Section 4 of the paper contains the LSB data hiding technique using decomposition of prime numbers. The approach for decomposition using prime numbers is discussed here. Section 5 contains LSB data hiding by natural numbers. Here the decomposition of bit planes by natural numbers is discussed. In Section 6 we present a comparative analysis between the various LSB data hiding methods. Finally, in Section 7, we present the conclusion for the paper.

CLASSICAL LSB DATA HIDING TECHNIQUE

One of the simplest implementation of data hiding technique is the classical least significant bit data hiding technique [18]. The technique is based on manipulation of the least significant bits of the carrier image to accommodate the hidden message.

The insertion of LSB varies according to the number of bits in an image. For an 8 bit image, the value of 8th bit, which is the least significant bit of each pixel, would be modified and the secret message would be embedded [19].

Suppose if we wanted to insert the letter **'A'** into an image. The binary equivalent of **'A'** is 10000001. Now if our sample image of 8 bit has the following pixel values:

10101010 10001010 11111111 010110101

10101010 10101010 10001011 10101111

So, embedding the binary equivalent of **'A'** i.e. 10000001would change the pixel values of our sample image to:

1010101**1** 10001010 1111111**0** 010110100**0**

10101010 10101010 1000101**0** 0101111

In our sample image the least significant bit value of pixel 1, 3, 4 and 7 have been changed.

Similarly, for an image of 24 bit, we need to change the RGB (Red Blue Green) color values. Suppose our sample image has the following pixel values,

[10010101 10101010 11110101]
[10001001 10001110 10000010]

[11110101 01001010 10001011]

So to accommodate our desired character, **'A'**, we have to make the following modifications:

[10010101 10101010 1111010**0**]
[1000100**0** 10001110 10000010]
[1111010**0** 01001010 10001011]

So, basically we have changed the values of 3 bit to accommodate our desired character.

Though, classical data hiding is simple to implement, it is vulnerable to image manipulation [15].

LSB DATA HIDING USING FIBONACCI NUMBERS

Battisti et al. [1] proposed a method of embedding data into digital media by decomposition of Fibonacci number sequence which allowed different bit plane decomposition when compared to the classical LSB scheme.

The Fibonacci sequence, named after Leonardo of Pisa, also known as Fibonacci, is a sequence of numbers in the following integer sequence:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89 …

The Fibonacci numbers are defined by a recurrence relation:

$$F_p(0) = F_p(1) = 1$$
$$F_p(n) = F_p(n-1) + F_p(n-p-1), \ \forall n \geq 2, n \in N$$

The bit planes are now decomposed based on the Fibonacci sequence. The main drawback of this approach is that of redundancy and to counter that and obtain a unique representation, Zeckendorf theorem [20] is used.

To embed the intended message in the cover image, it is decomposed into bit planes by using Fibonacci p-decomposition. The Zeckendorf condition is checked for each bit to be modified. If the condition is fulfilled, the bit is inserted otherwise the bit following it is considered.

LSB DATA HIDING USING PRIME NUMBERS

LSB data hiding using prime numbers is a data hiding technique proposed by Dey et.al [10] as an improvement over the Fibonacci numbers data hiding technique proposed by Battisti et. al. The main idea of the work was to use the prime number decomposition and generate new set of bit planes and embed information in these newly generated bit planes with minimal distortion.

In this approach, the researchers took an image of **m** bits and increased the number of bit planes to **n,** where the value of **n** was equal or greater than or equal to the number of bit planes of the image. This was achieved by converting the bit planes of the image to another number system using prime numbers as the weighted function. This resulted in the increase of number of bits and consequently it could be used for hiding data in higher bit planes with minimal distortion.

For decomposition, the weight function was defined as:

$$P(0) = 1, P(i) = p_i \ \forall i \in Z^+, p_i = i^{th} \ Prime$$

In case of any ambiguity, the lexicographically higher number is given preference.

For embedding the data in the image, a number n is chosen in such a way that all possible pixel values in the range of $[0, 2^k - 1]$ could be represented using first **n** prime number, so that **n** virtual are achieved after the decomposition.

The value of **n** can be found out by the formula,

$$\sum_{i+0}^{n-1} p_i \geq 2^k - 1$$

After finding the prime numbers, a k-bit to n-bit sequence is mapped, marking all the valid prime number in the system. Now, for each pixel of the image, a virtual bit plane is chosen and secret data is embedded is by modifying the corresponding virtual bit plane by the desired information bit by bit. If the resulting sequence of data after embedding the data matches the k-bit to n-bit mapping, the bits are kept otherwise it is discarded. After insertion of the hidden data, the resultant prime number system is converted back to the original binary system and the stego-image is achieved.

For extraction, each pixel with the hidden information is converted into prime decomposition and from those bit planes, the secret message is extracted. All the bits are combined together to get the final message.

LSB DATA HIDING USING NATURAL NUMBERS

The approach of LSB data hiding by natural numbers was proposed by Dey et.al [11]. In this approach, the researchers proposed data hiding by decomposition of a pixel value in sum of natural numbers. This resulted in generation of more bit planes than the Classical LSB data hiding, Fibonacci LSB data hiding and the Prime number data hiding [10].

For decomposition, the weight function is defined as:

$$W(i) = N(i) = i + 1, \ \forall I \in Z^+ \cup \{0\}$$

The researchers used the same concept in case of ambiguity which gave higher precedence to lexicographically higher number.

For embedding the data into the k-bit image, a number **n** is chosen in a way such that all pixel values in the range of $[0, 2^k – 1]$ could be represented using first n numbers, which resulted in generation of **n** virtual bit planes .

The value of **n** can be found out by the formula,

$$n \geq \frac{-1 \pm \sqrt{2^{k+3} + 9}}{2}$$

After finding the value of n, a k-bit to n-bit map is created and all valid representations in natural numbers system are marked. Now, for each pixel a virtual bit plane is chosen and the secret data is embedded. If the virtual bit plane matches the

mapping system, the hidden data is kept otherwise it is discarded. After insertion of the secret message, the natural number system is converted back to its original binary form and the stego-image is achieved.

To extract the message from the stego-image, all the pixels with embedded data bit are converted to the natural number decomposition, and the secret message bits are extracted. Finally, all bits are combined together to get the embedded hidden message.

COMPARISON OF LSB DATA HIDING TECHNIQUES

In the previous section, we have looked at the various techniques for data hiding using number decomposition. Here, we present a comparative analysis between the various data hiding techniques using number decomposition.

Table 1. Technique used in LSB data hiding

Classical LSB	Classical LSB data hiding uses the simplest approach. In classical LSB data hiding, the least significant bit of a pixel is manipulated in order to embed the desired image.
Fibonacci Technique	In Fibonacci approach, the bit planes are decomposed so as to generate more bit planes and then the secret message is embedded on following the Zeckendorf theorem.
LSB data hiding by prime numbers	In LSB data hiding using prime numbers, the bit planes are decomposed by using sum of prime numbers. After that the secret message is embedded lexicographically.
LSB data hiding by natural numbers	In LSB data hiding using natural numbers, the bit planes are decomposed by using sum of natural numbers. Here also the secret message is embedded lexicographically.

Table 2. Embedding Techniques used in LSB data hiding

Classical LSB	Data are inserted in the least significant bit of the cover image.
Fibonacci Technique	Technique uses Fibonacci P-sequence for generation of bit planes and data is inserted if it passes the Zeckendorf condition.
LSB data hiding by prime numbers	A k-bit to n-bit map is created where the value of n is, $$\sum_{i+0}^{n-1} p_i \geq 2^k - 1$$ Data is inserted bit by bit matching the k-bit to n-bit mapping sequence.
LSB data hiding by natural numbers	A k-bit to n-bit map is created where the value of n is, $$n \geq \frac{-1 \pm \sqrt{2^{k+3} + 9}}{2}$$ Data is inserted bit by bit matching the k-bit to n-bit mapping sequence.

Table 3. Weight functions

Classical LSB	The least significant bit of the image pixel is manipulated. In case of 8 bit image it is the 8th bit of each byte. For 24 bit image the RGB color code are changed [12].
Fibonacci Technique	$F_p(0) = F_p(1) = 1$ $F_p(n) = F_p(n-1) + F_p(n-p-1),$ $\forall n \geq 2, n \in N$
LSB data hiding by prime numbers	$P(0) = 1, P(i) = p_i \ \forall i \in Z^+, p_i = i^{th} Prime$
LSB data hiding by natural numbers	$W(i) = N(i) = i + 1, \ \forall I \in Z^+ \cup \{0\}$

Table 4. Number of bit planes generated

Classical LSB	8 bit planes using gray scale 8bit Lena image.
Fibonacci Technique	12 bit planes using gray scale 8bit Lena image.
LSB data hiding by prime numbers	15 bit planes using gray scale 8bit Lena image.
LSB data hiding by natural numbers	23 bit planes using gray scale 8bit Lena image.

CONCLUSIONS

In this paper, we looked at the different approaches of the spatial data hiding techniques using Least Significant Bit manipulation.

We discussed about the classical LSB data hiding technique which is based on manipulation of the least significant bit. We took a further look at the different number decomposition techniques by which more bit planes could be generated to conceal more data without causing significant distortion.

By comparative analysis, we could see that among the data hiding techniques, decomposition using natural numbers yielded the highest number of bit planes which was 23. This was followed by 15, 12 and 8 bit of Prime number technique [11], Fibonacci Technique and Classical LSB technique respectively.

Finally we can say that in this era of internet where there is increasing use of digital format to send valuable information and data, steganography will play a very important role in the days to come.

REFERENCES

[1.] F. Battisti, M. Carli, A. Neri, K. Egiaziarian, "A Generalized Fibonacci LSB Data Hiding Technique, 3rd International Conference on Computers and Devices for Communication (CODEC-06), Institute of Radio Physics and Electronics, University of Calcutta, December 18-20, 2006.

[2.] Aruna Ambalavanan and Rajarathnam Chandramouli "A Bayesian Image Steganalysis approach to estimate the embedded secret message," International Multimedia Conference, Proceedings of the 7th Workshop on Multimedia and Security, ACM Press, New York, USA, 2005, pp. 33 –38.

[3.] Liu Shaohui, Yao Hongxun, Gao Wen, "Neural network based steganalysis in still images," Proceedings of the 2003 International Conference on Multimedia and Expo (ICME 2003), vol.2, July 2003, pp. II – 509–512.

[4.] Patricia Lafferty, Farid Ahmed "Texture-based steganalysis: results for color images," Proceedings of the SPIE Mathematics of Data/Image Coding, Compression, and Encryption VII, with Applications, vol. 5561, 2004, pp. 145 – 151.

[5.] Steganographic Techniques and their use in an Open-Systems Environment-Bret Dunbar, The Information Security Reading Room, SANS Institute 2002, http://www.sans.org/readingroom/whitepapers/cover t/677.php

[6.] Steganography Primer - Ruid, Computer Academic underground,2004,http://www.dustintrammell.com/p resentations/Steganography-Primer.pdf

[7.] J. R. Hernandez, J. M. Rodr´ıguez, and F. P´erez-Gonz´alez, "Improving the performance of spatial watermarking of images using channel coding," Signal Process. 80(7), pp. 1261–1279, 2000.

[8.] F. Battisti, K. Egiazarian, M. Carli, and A.Neri, "Data hiding based on Fibonacci-Haar transform," in Mobile Multimedia/Image Processing for Military and Security Applications, SPIE Defense and Security, Vol. 6579, May 2007.

[9.] E. Mammia, F. Battisti, M. Carlia, A. Neria, and K. Egiazarianb, "A novel spatial data hiding scheme based on generalized Fibonacci sequences", Mobile Multimedia/Image Processing, Security, and Applications 2008, Proceedings. of SPIE ,Vol. 6982 69820C, (2008)

[10.] Sandipan Dey, Ajith Abraham and Sugata Sanyal "An LSB Data Hiding Technique Using Prime Numbers", Third International Symposium on Information Assurance and Security, August 29-31, 2007, Manchester, United Kingdom, IEEE Computer Society press, USA, ISBN 0-7695-2876-7, pp. 101-106, 2007.

[11.] Sandipan Dey, Ajith Abraham and Sugata Sanyal "An LSB Data Hiding Technique Using Natural Numbers", IEEE Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2007, Nov 26-28, 2007, Kaohsiung City, Taiwan, IEEE Computer Society press, USA, ISBN 0-7695-2994-1, pp. 473-476, 2007.

[12.] *Nani Koduri, "Information Security through Image Steganography using Least Significant Bit Algorithm", Master's Thesis, Information Security and Computer Forensics, University of East London.*

[13.] *Soumyendu Das, Subhendu Das, Bijoy Bandyopadhyay and Sugata Sanyal, "Steganography and Steganalysis: Different Approaches", International Journal of Computers, Information Technology and Engineering (IJCITAE), Vol. 2, No 1, June, 2008, Serial Publications.*

[14.] *Dhaval Gada, Rajat Gogri, Punit Rathod, Zalak Dedhia, Nirali Mody, Sugata Sanyal and Ajith Abraham, "A Distributed Security Scheme for Ad Hoc Networks", ACM Crossroads, Special Issue on Computer Security. Volume 11, No. 1, September, 2004, pp. 1-17.*

[15.] *Neil F. Johnson and Sushil Jajodia, "Exploring steganography: seeing the unseen," IEEE Computer 31(2), pp. 26-34, 1998.*

[16.] *Jessica Fridrich and Miroslav Goljan, "On estimation of secret message length in LSB steganography in spatial domain," in Security, Steganography, and Watermarking of Multimedia Contents VI, Proceedings of SPIE 5306, pp. 23-34, 2004.*

[17.] *Guorong Xuan, Yun Q. Shi, Zhicheng Ni, "Reversible data hiding using integer wavelet transform and companding technique," IWDW04, Korea, October 2004.*

[18.] *Chi-Kwong Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," Pattern Recognition, pp. 469–474, Mar. 2004.*

[19.] *Walter Bender, Daniel Gruhl, Norishighe Morimoto, and A. Lu, "Techniques for data hiding." I.B.M. Systems Journal, vol. 35, no. 3 & 4, pp. 313{336, 1996.*

[20.] *Verner E. Jr. Hoggatt, "Fibonacci and Lucas Numbers," The Fibonacci Association, Santa Clara, California, USA, 1972.*

[21.] *Raymond B. Wolfgang and Edward J. Delp, "A watermarking technique for digital imagery: further studies." In International Conference on Imaging, Systems, and Technology, pp. 279{287, IEEE, Las Vegas, Nevada, U.S.A., 30 Jun, 1997*

[22.] *Mauro Barni, Franco Bartolini, Vito Cappellini,Alessandro Piva(1998), "A DCT-domain system for robust image watermarking", Signal Processing, Vol. 66 (1998), pp.357–372.*

[23.] *Nabin Ghoshal , Jyostna Kumar Mandal, "A Novel Technique for Image Authentication in Frequency Domain using Discrete Fourier Transformation Technique", Malaysian Journal of Computer Science, ISSN 0127-9094, Vol. 21, No. 1, pp. 24-32, 2008.*