A Novel Spatial Domain Scheme for Multiple Bitplane Steganography

Arijit Sur, Piyush Goel and Jayanta Mukherjee
Department of Computer Science and Engineering
Indian Institute of Technology,Kharagpur
Kharagpur-721302, India
Email : (arijits,piyush,jay)@cse.iitkgp.ernet.in

***Abstract*—In this paper, a novel spatial domain steganographic scheme is proposed to reduce the amount of noise added during embedding. One of the main goals of steganography is statistical undetectability. As statistical undetectability is explicitly related to the amount of noise added during embedding, reduction of embedding noise is an important factor for secure steganography. In the proposed scheme Single Digit Sum function is used for embedding. We have shown both analytically and experimentally that the proposed scheme adds less noise when multiple bit planes are used for steganography.**

I. Introduction

Least Significant Bits (LSB) embedding in digital images is the most popular steganographic embedding scheme due to its simplicity, large payload and visual imperceptibility. LSB Replacement is vulnerable to targetted attacks which exploit the structural imbalance introduced in the cover due to the embedding procedure [8,9].

In order to overcome the structural weakness of the single plane LSB embedding has been extended to least significant multiple bitplanes [3,4,5]. The most imprtant point to be mentioned here is that even after embedding in multiple bitplanes, no significant visual distortion can be observed in the cover image. The amount of noise added due to steganographic embedding depends upon the number of changes made in the cover signal. Thus a steganographic algorithm should minimize the amount of noise added during embedding. For a given medium, the steganographic algorithm which makes fewer embedding changes or adds less additive noise will be less detectable as compared to an algorithm which makes relatively more changes or adds higher additive noise [6].

In this paper we propose a new scheme in which embedding is done in multiple LSB planes (2-3 planes) using Single Digit Sum Encoding (SDS). The pixels are selected randomly from the image using a shared secret key. Then that pixel is changed such that the Single Digit Sum value of that pixel is changed to the decimal equivalent of the message bits. It must be noted that by the proposed scheme we can embed upto 3 message bits in a single pixel. We show analytically that in terms of noise added while embedding, the proposed scheme outperforms multiple bit plane LSB embedding. The rest of the paper is organized as follows: in section 2, the encoding scheme is described, embedding and extracting algorithms are discussed in section 3. Security of the proposed scheme is analyzed in section 4 using WAM blind steganalysis. In section 5 we compare the performance of the proposed scheme with existing spatial domain steganographic schemes and finally the paper is concluded in section 6.

II. Embedding Algorithm

A. Embedding in Three LSB Planes

Recently some steganographic techniques [3,4] have been proposed where more than one least significant bit planes are used for embedding. The main dwarback of multiple bitplane embedding (specifically more than two bitplanes) is that the additive noise due to embedding is very high. So a blind steganalyzer can be easily designed to detect the statistical dissimilarities between the cover and stego images. But if the structural weakness is considered then embedding in multiple bit planes is harder to detect than single LSB plane embedding [5]. So the goal of this paper is to reduce the additive noise when embedding is done in multiple bit planes. The probability of amount of noise added during 3LSB (we have used the notation nLSB when bit replacement is done in n least significant bit planes) embedding is given in Table 1. If i is the amount of noise and $P(i)$ is the corresponding probability by which i amount of noise is added to the pixel due to embedding, then the total amount of additive noise (ξ_n) during nLSB can be calculated by the following equation

$$\xi_n = \sum_{i=1}^{2^n-1} i \times P(i) \tag{1}$$

So using Table 1 we can calculate the noise for 3LSB embedding with an embedding rate of p,

ξ_3 = $1 \times \frac{7p}{32}$ + $2 \times \frac{3p}{16}$ + $3 \times \frac{5p}{32}$ + $4 \times \frac{p}{8}$ + $5 \times \frac{3p}{32}$ + $6 \times \frac{p}{16}$ + $7 \times \frac{p}{32}$

= $\frac{168p}{64}$

=$2.625p$ *per pixel used for embedding.*

Thus, the total amount of noise added to one pixel of the cover image = $2.625p/3$ since 3 least significant bit planes are used for embedding.

TABLE I
The Probability of Additive Noise when Embedding is Done in Three LSBs with Embedding Rate = p

Amount	0	1	2	3	4	5	6	7
Probability	1- $\frac{7p}{8}$	$\frac{7p}{32}$	$\frac{3p}{16}$	$\frac{5p}{32}$	$\frac{p}{8}$	$\frac{3p}{32}$	$\frac{p}{16}$	$\frac{p}{32}$

TABLE II
The Probability of Additive Noise during SDS Encoding for Pixel values $5 - 251$

Amount	0	±1	±2	±3	±4
Probability	1- $\frac{8}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

B. Single Digit Sum (SDS) Encoding

In this paper, we have used single digit sum encoding as part of the proposed embedding scheme which uses more than two bitplanes. SDS is a *many to one* function which is described by the following recurrence relation

$$T(n) = \begin{cases} n & \text{if } n < 10 \\ \\ T(\sum_{i=0}^{k-1} mod(\frac{n}{10^i}, 10)) & \text{if } n \geq 10 \end{cases} \quad (2)$$

where n is any k digit positive integer.

In SDS encoding method, the message bits are embedded by modifying the pixel value to a nearest value, such that SDS of the new value is same as the SDS of the decimal equivalent of the message bits. It can be observed that for pixel values between $5 - 251$ are altered by a maximum of ±4 during embedding of decimal message symbols $1 - 9$. The probability of amount of noise added during SDS encoding for pixel values $5 - 251$ is given in Table 2.

$$\xi_{SDS_{5-251}} = \sum_{i=1}^{4} i \times P(i) \quad (3)$$

where i is the amount of noise and $P(i)$ is the probability of the noise i to be embedded during SDS encoding as described in Table 2. For the rest of the cover image pixels (i.e. $0 \ldots 4$ and $252 \ldots 255$) the amount of noise added during SDS encoding is given in Table 3.

Using Table 2 and 3, the total additive noise for SDS encoding is calculated using the following equation

$$\xi_{SDS} = \frac{5}{256} + \frac{2}{256} \times (\frac{36}{9} + \frac{29}{9} + \frac{24}{9} + \frac{21}{9}) + \frac{247}{256} \times \xi_{SDS_{5-251}} \quad (4)$$

≈ 2.25

with embedding rate = p, average noise will be $2.25p$ *per pixel used for embedding.*

TABLE III
The Average Additive Noise using Single Digit Sum Encoding for Pixel values $0 - 5$ **and** $252 - 255$

Pixel Values	0	1	2	3	4	252	253	254	255
Noise	5	4	$\frac{29}{9}$	$\frac{24}{9}$	$\frac{21}{9}$	$\frac{21}{9}$	$\frac{24}{9}$	$\frac{29}{9}$	4

Thus, the total amount of noise added to one pixel of the cover image = $2.25p/3$ since 3 least significant bit planes are used for embedding.

We can see that the average noise added in a pixel carrying message bits using SDS encoding method is 2.25 which is relatively less than average noise added in the case of 3LSB embedding.

C. Proposed Steganographic Algorithm

In the proposed scheme, a pixel is selected randomly from the image using a shared secret key, let it be λ. Then we take 3 message bits and find out there decimal equivalent, say δ. The pixel λ is changed to the nearest possible value such that Single Digit Sum value of λ (SDS(λ)) is changed to δ. For the decoding process the pixels numbers can be regenerated using the shared key and their SDS value is found. The binary equivalent of this SDS value represents the message bit sequence embedded in this pixel. It should be noted that SDS based schemes cannot handle long strings of zeros. This limitation can be overcome by making slight modifications to the message bit stream. A Runlength Encoding can be performed on the message stream to remove long string of zeros or we can represent a string of zeros by embedding some other number for example in our experiments we have replaced a string of 3 zeros by 8.

The proposed scheme is able to achieve embedding rates equivalent to 3 LSB scheme , i.e a maximum embedding rate of 3.0, while adding less noise. Next, we make a comparative analysis of the noise added by the two schemes.

We have shown in section II that the noise added per pixel by 3 LSB embedding scheme is approximately $2.625p/3$ i.e $0.875p$ where p is the embedding rate. As we have explained above, the proposed scheme tries to reduce the embedding noise by using SDS Encoding Scheme. This fact can be observed from the scatter plot for the noise added per pixel for 3 LSB vs the proposed scheme. The scatter plot has been drawn using 100 grayscale Tagged Image Format (TIFF) images. All the points lie far below the diagonal line i.e. the noise per pixel is much less for the proposed scheme than 3 LSB scheme.

III. Embedding and Extraction Procedure

A. Embedding Procedure

The embedding algorithm is outlined below:

- Select a pixel from the image in a pseudo-random sequence with a shared secret key. Let the pixel value be ϵ.
- Consider a message strings of 3 bits each and compute its equivalent decimal value δ .
- Calculate the SDS value of ϵ using Eq. (2). Let it be denoted by SDS(ϵ).
- Change this pixel (ϵ) to β such that SDS(β) becomes δ.

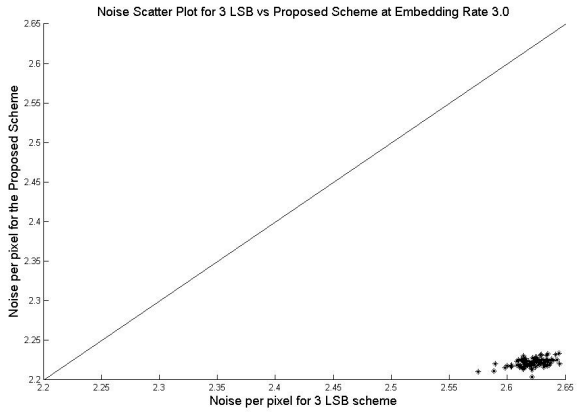

Fig. 1. Noise Scatter Plot for 3LSB scheme vs Proposed Scheme at Embedding Rate 3.0. $x-axis$ Noise per pixel for 3LSB scheme, $y-axis$ Noise per pixel for the Proposed Scheme

B. Extraction Procedure

The extraction algorithm is a simple inverse process of the embedding algorithm and is outlined below:

- Using the pseudo-random sequence with a shared secret key, calculate the pixel to be decoded (ϵ).
- Compute the SDS value of ϵ. The computed value is the decimal equivalent of the message string. Let it be denoted by SDS(ϵ).
- Then calculate the binary equivalent of SDS(ϵ) to get the 3 bit message string.

C. An Illustrative Example

Consider a pixel value 148 (ϵ). Let us assume we have to embed a message string 110. The decimal equivalent of the message string is 6. We change ϵ such the SDS(ϵ) equals 6. So 148 is changed to 150. It should be noted that the value has been changed to 150 instead of 141 to minimize the noise added during embedding. The message bits can be extracted by calculating SDS(150) and then finding its binary equivalent 110.

IV. Steganalysis

A steganographic scheme is considered secure if there are no artifacts in the stego image that could be detected by an attacker with a probability better than random guessing, given the full knowledge of the embedding algorithm, including the statistical properties of the source of the cover images, except the stego key (Kerckhoffs' principle). A formal definition of steganographic security can be found in [1,2].

For evaluating the security of the proposed scheme, we have used the Wavelet Absolute Moment Steganalysis (WAM) proposed in [7]. WAM is a blind steganalyzer which is based on feature extraction from a set of images and then using a Linear Classifier for classifying them as cover or stego image.

A. Feature Extraction

Features for WAM steganalysis are calculated from the noise component of the stego image in the Wavelet domain. Assuming that the stego image is an additive mixture of a non-stationary Gaussian signal (the cover image) and a stationary Gaussian signal with a known variance (the noise), Wiener filter is used to extract noise component in the wavelet domain. All the features (statistical moments) are calculated as higher order moments of the noise residual in the wavelet domain. Moments upto order 9 have been used in [7], thus the total number of features for a grayscale image is $3n_{mom}$ where n_{mom} = 9.

B. Steganalytic Classifier

The 27 dimensional feature space obtained after Feature Extraction is reduced to single dimension using Fisher Linear Discriminant (FLD) Analysis. Then Linear Discriminant Analysis (LDA) classifier is used to classify the projected points on the principal component axis. ROC curves are plotted for evaluating the performance of the classifier on different steganographic algorithms.

V. Experimental Results

Two hundred uncompressed Tiff images of different sizes have been used for our experiment. WAM classifier is trained with 100 cover and 100 stego images and ROC curves have been plotted. The performance comparison of the proposed method with LSB embedding using 3 LSB planes for embedding rates of 0.3 and 0.5 have been given in Fig.2 and Fig.3 respectively. It can be observed that the proposed scheme generates more number of false positives than 3LSB scheme for both the embedding rates and hence is more secure than 3 LSB scheme.

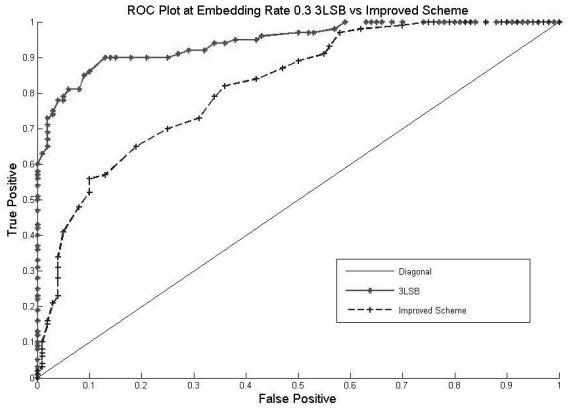

Fig. 2. Performance of Proposed Scheme with 3LSB at Embedding rate 0.30

VI. Conclusion

In this paper, we have proposed a new spatial domain block based steganographic algorithm which can embed same payloads as other spatial domain schemes while adding less noise to the cover signal. The proposed scheme has shown

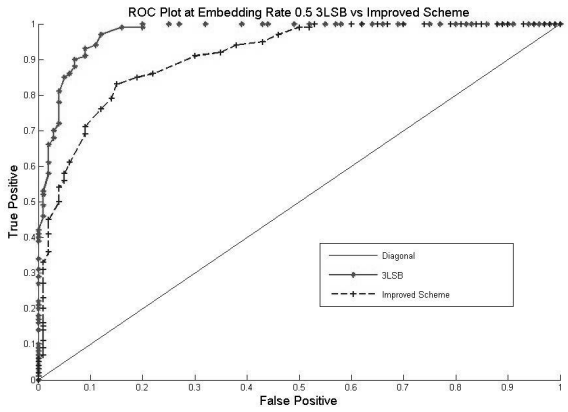

Fig. 3. Performance of Proposed Scheme with 3LSB at Embedding rate 0.50

better performance than normal 3LSB embedding against Wavelet Absolute Moment steganalyzer at same payload. This scheme is especially suitable for the environment when more than two bit planes are used for embedding.

REFERENCES

[1] Zllner, J., Federrath, H., Klimant, H., Pfitzmann, A., Piotraschke, R., Westfeld, A., Wicke, G., Wolf, G., Modeling the Security of Steganographic Systems, In: Aucsmith, D. (ed.): Information Hiding. 2nd International Workshop. Lecture Notes in Computer Science, Vol. 1525. Springer-Verlag, New York, pp. 344-354, 1998.

[2] Katzenbeisser, S., Petitcolas, F. A. P.: Defining Security in Steganographic Systems, SPIE Security and Watermarking of Multimedia Contents IV, Vol. 4675, Electronic Imaging 2000, San Jose, CA, pp. 50-56, 2002.

[3] Zhang, X., Wang, S.: Steganography using multiple-base notational system and human vision sensitivity, Signal Processing Letters, IEEE Volume 12, Issue 1, Jan. 2005 Page(s):67 70.

[4] Wu, D. -C, Tsai W. -H, I.: A Steganographic method for images by pixel-value differencing, Pattern Recognit. Lett., vol. 24, pp. 16131626, 2003.

[5] Ker, A.: Steganalysis of Embedding in Two Least-Significant Bits, IEEE Transaction on Information Forensics and Security, vol. 2, NO. 1, March 2007.

[6] Fridrich, J., Soukal, D.: Matrix Embedding for Large Payloads, Proc. of SPIE Electronic Imaging, Photonics West, January 2006

[7] Goljan, M., Fridrich, J. and Holotyak, T.: New Blind Steganalysis and its Implications, Proc. SPIE Electronic Imaging, Photonics West, January 2006

[8] Dumitrescu, S., Wu, X., Wang, Z.: Detection of LSB steganography via sample pair analysis, In: Proc. 5th Information Hiding Workshop. Volume 2578 of Springer LNCS. (2002) 355-372.

[9] Fridrich, J., Goljan, M., Du, R.: Reliable detection of LSB steganography in color and grayscale images, Proc. ACM Workshop on Multimedia and Security (2001) 27-30.