

## IMPROVING THE RELIABILITY OF DETECTION OF LSB REPLACEMENT STEGANOGRAPHY

Shreelekshmi R<sup>1</sup>, Wilscy M<sup>2</sup> and C E Veni Madhavan<sup>3</sup>

<sup>1</sup>Department of Computer Science & Engineering, College of Engineering, Trivandrum, India

[shreelekshmir@cet.ac.in](mailto:shreelekshmir@cet.ac.in)

<sup>2</sup>Department of Computer Science, University of Kerala, Trivandrum, India

[wilsyphilipose@gmail.com](mailto:wilsyphilipose@gmail.com)

<sup>3</sup>Department of Computer Science & Automation, Indian Institute of Science, Bangalore, India

[cevm@csa.iisc.ernet.in](mailto:cevm@csa.iisc.ernet.in)

### ABSTRACT

*We propose a simple method for increasing the reliability of detection of LSB replacement steganography. RS Steganalysis Method (RSM), Sample Pair Method (SPM), and Least Square Method (LSM) are three of the most reliable steganalysis methods in the literature for LSB replacement steganography on digital images in spatial domain. However all these methods are prone to error. We propose a simple method to identify images which give very accurate results and images which give highly inaccurate results by these methods. The novelty of our method is that it does not require any knowledge about the cover images as the image classification is done based on certain statistical properties of the image, which are invariant with embedding. Thus it helps the steganalyst in attaching a level of confidence to the estimation he makes and thereby increasing the reliability of steganalysis of digital images. It also helps to choose between images for more secure LSB replacement steganography.*

### KEYWORDS

*LSB replacement steganography, steganalysis, confidence level, reliability*

## 1. INTRODUCTION

In this paper we focus our studies to steganalysis of digital images in spatial domain employing LSB replacement steganography. Many methods (see [1], [2], [3], [4], [5], [6] and [7]) have been proposed in the literature for steganalysis of digital images. These methods give very accurate results on most of the images. However they give highly inaccurate or moderately inaccurate results on some images.

With the development of very accurate steganalysis techniques, methods have been proposed for increasing reliability of LSB steganography also. Fridrich et. al developed a general coding method called matrix embedding[8] that can be applied to most steganographic schemes for improving their steganographic security.

RSM[1], SPM[2], and LSM[3] are three of the most reliable quantitative steganalysis methods based on the probability of transitions due to LSB embedding. In general performance of SPM is better than RSM and that of LSM is better than SPM. However all these methods show moderately inaccurate results on some images and highly inaccurate results on some other images.

Attempts were made in estimating the error in various steganalysis methods. Ker derived error distribution in Least Square steganalysis[8]. He showed that error distribution of LSM has heavy tails indicating extreme outliers due to image specific properties.

Therefore when the steganalyst makes an estimation on the hidden message length, he is not able to attach a confidence level to the results. Our attempt is to improve the reliability of steganalysis by classifying images into highly accurate and highly inaccurate categories.

In this paper we present a simple image classification to identify a class of images that give accurate results and another class of images that give highly inaccurate steganalysis results. The method we propose is based on the statistical properties of the images which are invariant with embedding.

The rest of this paper is organized as follows: Section 2 explains the notations we use in this paper. Section 3 introduces the new method for image classification. Section 4 shows the experimental results we obtained. Section 5 is the conclusion.

## 2. NOTATIONS

- $P$  : Multiset of sample pairs  $(u, v)$  drawn from digital image
- $X_n$  : Sub multi set of  $P$  that consists of sample pairs drawn from cover signal and whose values differ by  $n$  and in which even value is larger
- $Y_n$  : Sub multi set of  $P$  that consists of sample pairs drawn from cover signal and whose values differ by  $n$  and in which odd value is larger
- $C_m$  : Sub multi set of  $P$  that consists of sample pairs drawn from cover signal and whose values differ by  $m$  in the first  $(b-1)$  bits (i.e., by right shifting one bit and then measuring the difference of  $m$ )
- $D_0$  : Sub multi set of  $P$  that consists of sample pairs drawn from cover signal and whose values differ by 0
- $p$  : Estimated length as percentage of number of pixels in the image
- $b$  : Number of bits to represent a sample value
- $k$  : 5 for LSM, 30 for SPM and 127 for RSM
- $$U = \sum_{i=1}^{2^b-1} X_i$$
- $$V = \sum_{i=2}^{2^b-1} Y_i$$
- $$\alpha_k = \frac{\sum_{m=1}^k (X_{2^{m-1}} - Y_{2^{m-1}})}{\sum_{m=1}^{2^{b-1}-1} C_m}$$
- $$\beta_k = \frac{\sum_{m=0}^k (X_{2^{m+1}} - Y_{2^{m+1}})}{\sum_{m=1}^{2^{b-1}-1} C_m}$$

## 3. IMAGE CLASSIFICATION

RSM, SPM, and LSM are based on probabilities of transitions between sample pairs due to LSB embedding operations. RSM is based on the finite state machine shown in (Figure 1)[2]. LSM

and SPM are based on the finite state machines shown in (Figure 2) and (Figure 3). The arrows are labelled with the probability of transition.

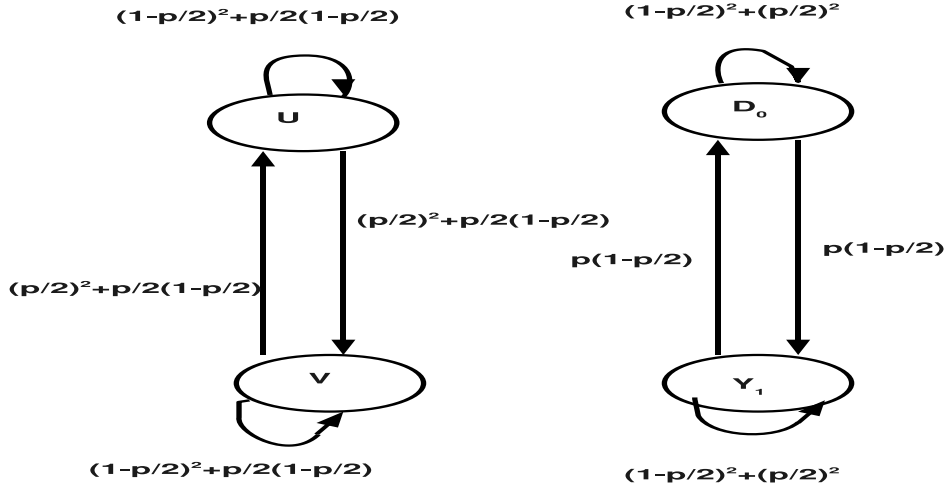


Figure 1. Finite state machine associated with RSM

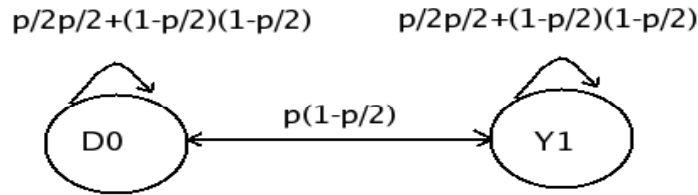


Figure 2. Finite state machine associated with C0

The precision of RSM is based on the assumption that [1]

$$E\left\{\sum_{i=1}^{2^b-1} X_i\right\} = E\left\{\sum_{i=1}^{2^b-1} Y_i\right\} \quad (1)$$

The precision of SPM depends on following hypotheses[2]

$$E\{|X_{2m+1}|\} = E\{|Y_{2m+1}|\} \quad (2)$$

or a more relaxed condition

$$E\left\{\left|\bigcup_{i=1}^j X_{2m+1}\right|\right\} = E\left\{\left|\bigcup_{i=1}^j Y_{2m+1}\right|\right\} \quad 1 \leq i \leq j \leq 2^{b-1} - 1 \quad (3)$$

LSM makes the hypothesis[3] that a small parity difference occur in natural signals for each m. Estimation errors occur when the hypotheses do not hold.

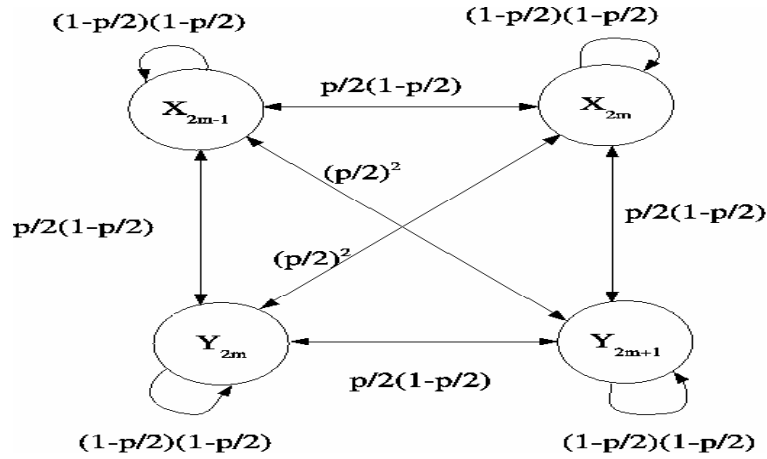


Figure 3. Finite state machine associated with  $C_m$ ,  $m > 0$

### 3.1 Identification of images which give highly inaccurate results

As suggested in SPM, we analyse the transitions due to embedding, as follows. Image is considered as the set  $P$ .  $\{C_0, C_1, \dots, C_{2^{b-1}-1}\}$  defines a partition on  $P$ . Further  $\{X_{2m-1}, X_{2m}, Y_{2m}, Y_{2m+1}\}$  is a partition on  $C_m$ ,  $1 \leq m \leq 2^{b-1} - 1$  and  $\{D_0, Y_1\}$  is a partition on  $C_0$ . When data is embedded into the image, sample pair in one sub multi set change to another sub multi set and thus the cardinalities of these sub multi sets change with embedding.

In most of the images,  $|X_{2m-1}| > |X_{2m}| \approx |Y_{2m}| > |Y_{2m+1}|$  and  $|D_0| > |Y_1|$ . Hence due to embedding  $|X_{2m-1}|$  decreases and that of  $|Y_{2m+1}|$  increases.  $|X_{2m}|$  and  $|Y_{2m}|$  increase or decrease depending on their initial values. At 100% embedding, all these cardinalities become equal. Typical change in cardinalities of sub multi sets of a  $C_m$ ,  $0 \leq m \leq 2^{b-1} - 1$  are shown in Figure 4.

In cover images,  $|X_{2m+1}| \approx |Y_{2m+1}|$ ,  $0 \leq m \leq 2^{b-1} - 1$  and with embedding the difference between  $|X_{2m+1}|$  and  $|Y_{2m+1}|$  increases and the difference is maximum at 100% embedding. In a cover image parity difference is negligible. Therefore a cover image, where the cardinalities of sub multi sets of a  $C_m$  are equal, meet the following criteria.

$$|X_{2m-1}| \approx |Y_{2m-1}| \approx |X_{2m+1}| \approx |Y_{2m+1}| \approx |X_{2m}| \approx |Y_{2m}|, m > 0 \quad (4)$$

$$|D_0| \approx |X_1| \approx |Y_1|, m=0 \quad (5)$$

The probability of transition from a sub multi set  $X$  to a sub multi set  $Y$  and that from  $Y$  to  $X$  are same. Therefore when a  $C_m$  meet the criteria given in (4) or (5), the cardinalities of sub multi sets of that  $C_m$  do not change with any amount of embedding. Such a  $C_m$  does not help in estimating the length of embedding. If the image contains only such  $C_m$ s, steganalysis is unreliable using RSM, LSM and SPM.

More precisely

$$RSM, LSM \text{ and } SPM \text{ are unreliable if the conditions (4) or (5) hold good for all } C_m, 0 \leq m \leq k \quad (6)$$

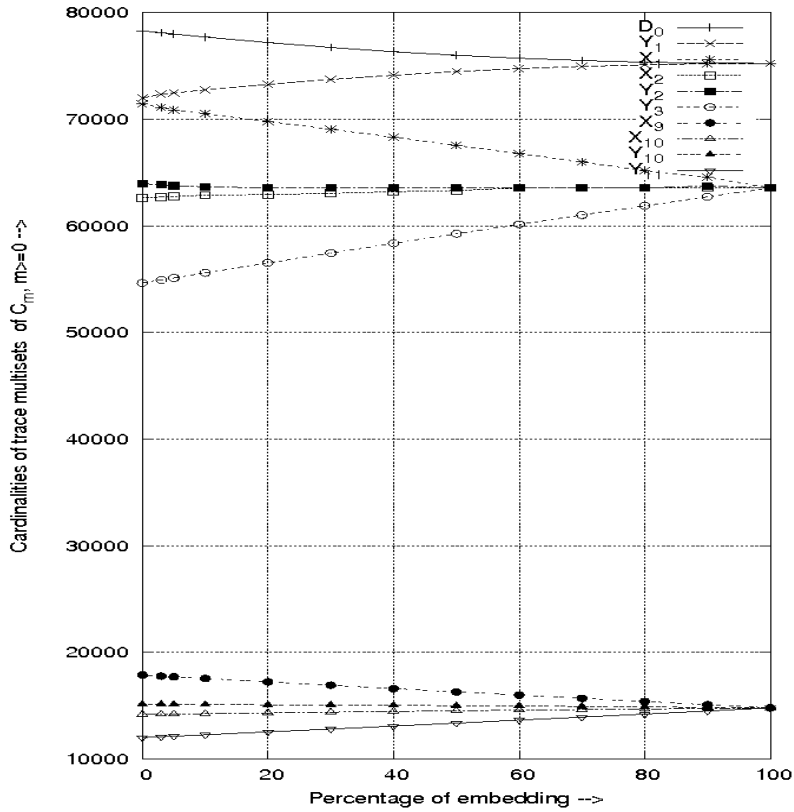


Figure 4. Change in cardinalities of sub multi sets of  $C_m$  due to embedding in general

In a cover image where the cardinalities of sub multi sets of a  $C_m$  are not equal, at 100% embedding, these cardinalities become equal and the difference between  $|X_{2m+1}|$  and  $|Y_{2m+1}|$ ,  $0 \leq m \leq 2^{b-1} - 1$ , becomes maximum. Hence the criteria given in (6) can not be true for such images with any amount of embedding.

Therefore by testing the criteria given in (6), we can identify images which yield inaccurate results for LSM, SPM and RSM. Thus the steganalyst is able to attach a confidence level to his results.

### 3.2 Identification of images which give accurate results

The cardinalities of  $C_m$ s in the image are important as the estimation is done based on the law of large numbers. As pointed out in [2] parity play a role in the accuracy of results.  $|C_0|$  is involved in estimation using RSM, SPM and LSM. When  $|C_0|$  is very large, the negative effect of parity is minimum.  $|C_0|$  is an invariant with embedding. So an image, if  $|C_0|$  is sufficiently large, gives reliable results. However image might give mild estimation error due to the effect of parity.

The images which meet the criteria in (6) will give highly inaccurate results by LSM, SPM and RSM. Hence an image which does not meet the criteria in (6) and for which  $|C_0|$  is large, gives

highly accurate estimated results. Thus the image classification we suggest involves finding  $|C_0|$  and checking the criteria in (6).

More precisely

*RSM, SPM and LSM are highly accurate if  $|C_0|$  is large and does not meet condition in (6)* (7)

The criteria given in (7) is an invariant with embedding, hence by testing the same on the stego image, we can predict results are going to be accurate or not. The prediction does not require knowledge about cover image.

Using the classification proposed we can detect images which give accurate results and images which give highly inaccurate results from the properties of stego images. Thus the novelty of the proposed classification is that it is done based on the statistics from the stego images alone. Since the classification is done using the statistics of the stego image, the steganalyst is able to add a confidence level also along with the results he gives.

#### 4 EXPERIMENTAL RESULTS

We downloaded few hundred 24-bit color images which were originally stored as high-quality JPEG images. For our test purposes, we resized them to 800X600 pixels. Then we selected one hundred 24-bit color images meeting the condition (6).

Table 1. Estimated length of hidden message when the cardinalities of sub multi sets of  $C_m$  are almost equal in cover image ( $\alpha_{30} = 0.00005$ ,  $\beta_{30} = 0.0115$ )

| Embedded message length (%) | RSM   | SPM    | LSM |
|-----------------------------|-------|--------|-----|
| 0                           | 80.10 | 72.20  | Nan |
| 3                           | 82.12 | 74.70  | Nan |
| 5                           | 84.40 | 74.50  | Nan |
| 10                          | 82.50 | 78.70  | Nan |
| 20                          | 88.30 | 86.11  | Nan |
| 30                          | 87.40 | 87.11  | Nan |
| 40                          | 82.10 | 90.80  | Nan |
| 50                          | 88.50 | 95.20  | Nan |
| 60                          | 94.30 | 96.50  | Nan |
| 70                          | 92.40 | 98.20  | Nan |
| 80                          | 98.20 | 99.01  | Nan |
| 90                          | 99.80 | 100.10 | Nan |
| 100                         | nan   | 99.60  | Nan |

We created a series of stego images by embedding messages of length 0%, 3%, 5%, 10%, ...,100% into these images using random LSB replacement method. Then we estimated the hidden message length from these stego images using RSM, SPM and LSM, the three most accurate steganalysis methods in the literature.

The estimated message lengths using a sample cover image with  $\alpha_{30} = 0.00005$  and  $\beta_{30} = 0.0115$  are given in table 1. In some cases RSM and SPM are not able to estimate message length as the

determinant turns out to be negative. In all other cases all three methods show very high estimation error. The results are similar for all other images satisfying the criteria given in (6).

We selected one hundred images meeting the classification criteria ( $|C_0| > 50\%$ ). We created a series of stego images by embedding messages of length 0%, 3%, 5%, 10%, ..., 100% into these images using random LSB replacement method. We estimated the hidden message length from these stego images using RSM, SPM and LSM. The test results are given in table 2.

Thus it is very evident that by testing the above criteria which is an invariant with embedding, we can predict whether the results given by the three steganalysis methods are highly accurate or highly inaccurate.

Table 2. Average estimated message length (in percent) for 100 images with  $|C_0| > 50\%$ .

| Embedded message length (%) | RSM   | SPM   | LSM   |
|-----------------------------|-------|-------|-------|
| 0                           | 0.78  | 0.38  | 0.17  |
| 3                           | 3.21  | 3.41  | 2.91  |
| 5                           | 5.41  | 5.21  | 4.91  |
| 10                          | 10.78 | 10.60 | 10.21 |
| 20                          | 21.98 | 21.20 | 20.35 |
| 30                          | 30.98 | 31.22 | 30.45 |
| 40                          | 41.20 | 40.65 | 40.24 |
| 50                          | 50.97 | 50.42 | 49.98 |
| 60                          | 60.78 | 60.35 | 60.14 |
| 70                          | 70.90 | 70.20 | 70.17 |
| 80                          | 79.10 | 78.17 | 79.35 |
| 90                          | 90.90 | 89.57 | 89.97 |
| 100                         | 96.50 | 97.02 | 98.85 |

## 5. CONCLUSIONS

In this paper we discussed a method for classifying images for increasing the reliability of steganalysis methods on digital images in spatial domain employing LSB replacement steganography. The proposed classification is done based on certain properties of images which are invariant with embedding. Using the image classification proposed, the steganalyst can identify images that give accurate results and images which give highly inaccurate results without having any knowledge on cover images. Thus steganalyst can attach a confidence level to his results. Thus the image classification we proposed increases the reliability of staganalysis of digital images in spatial domain employing LSB replacement steganography. It also helps to choose images for more secure steganography. Influence of other factors like parity on accuracy of steganalysis results of images in spatial domain are to be investigated further.

## REFERENCES

1. Fridrich J, Goljan M and Du R, (2001) "Detecting lsb steganography in colour and greyscale images", Magazine of IEEE multimedia, Special Issue on Security, October- November issue, 2001
2. Dumitrescu S Wu X and Wang Z, (2003) "Detection of lsb steganography via sample pair analysis", IEEE Transactions on Signal Processing, vol. 51, No.7, 2003, pp1995-2007

3. Lu P, Luo X, Tang Q and Shen L, (2005) "An improved sample pairs method for detection of lsb embedding", 6th Information hiding Workshop, Springer LNCS, vol.3200, pp116–127
4. Fridrich J, Du R and Meng L, (2000) "Steganalysis of lsb encoding in colour images", Proceedings of IEEE International conference on Multimedia and Expo New York City, NY, Jul 30 - Aug2, 2000
5. Tao Z and Xijian P, (2003) "Reliable detection of lsb steganography based on the difference image histogram", Proc. IEEE ICAAP, Part III, 2003, pp545–548
6. Ker A.D, (2005) "Improved detection of lsb steganography in greyscale images", In: Proc. The 6th Information Hiding Workshop, Springer LNCS, vol. 3200, pp97-115
7. Liu B Luo X and Liu F, (2005) "Improved RS method for detection of lsb steganography", In: Proc. Information Security & Hiding workshop , Springer LNCS, vol. 3481, pp508-516
8. Fridrich J and Soukal D, (2008) "Matrix embedding for large payloads", IEEE Transactions on Information Security and Forensics, vol. 7, pp12–17
9. Ker A, (2007) "Derivation of error distribution in least squares steganalysis", IEEE Transactions on Information Security and Forensics, vol. 2, pp140–148

### Authors

Shreelekshmi R received her B Tech in Computer Science & Engineering from University of Kerala, India and M E degree in Computer Science & Engineering from Indian Institute of Science, Bangalore, India. She is an Assistant Professor in College of Engineering, Trivandrum, India and is pursuing her Ph.D. Her research interests include pattern recognition, steganography and steganalysis.



M Wilscy received her B.Tech degree in Electrical Engineering from University of Kerala, ME in Computer Science from Indian Institute of Science, Bangalore and Ph.D from Indian Institute of Technology, Madras, India. She is the professor and head of the Dept. of Computer Science, University of Kerala, India. Her research interests include Image Processing, Pattern Recognition, Network Security and soft computing Techniques.

C E Veni Madhavan received his B.E. in Electrical Engineering from College of Engineering, Guindy, Madras, M.E. in Control Systems from BITS, Pilani and Ph.D in Control Theory from Indian Institute of Science, Bangalore. He is a Professor in the Department of Computer Science and Automation, Indian Institute of Science, Bangalore. His research, teaching, and development interests include cryptography, steganography, and arithmetic, algebraic and combinatorial algorithms.

