Detecting 2-LSB steganography using Extended Pairs of Values analysis

Omed Saleem Khalind
School of Computing
University of Portsmouth

Julio C. Hernandez-Castro
School of Computing
University of Kent

Benjamin Aziz
School of Computing
University of Portsmouth

{Omed.khalind, Benjamin.Aziz}@port.ac.uk
J.C.Hernandez-Castro@kent.ac.uk

Abstract

In this paper we propose an extended pairs of values analysis to detect and estimate the amount of secret messages embedded with 2-LSB replacement in digital images based on chi-square attack and regularity rate in pixel values, as explained in later sections. The method can accurately detect 2-LSB replacement even when the message length is about 10% of the total capacity. However, the method puts no assumptions neither on the image nor the secret message, as it tested with a random set of 3000 images from ASIRRA pet images and embedded with a random message for each case. This method of detection could also be used as an automated tool to analyse a bulk of images for hidden contents, which could be useful for digital forensics analysts in their investigation process.

Keywords: Steganography, Steganalysis, two least-significant bit embedding, Chi-square attack, Pairs of Values, digital forensics.

1. Introduction

There are numerous studies on LSB steganography in images, and lots of methods are proposed to detect the existence of embedded message. The reason behind the interest of LSB steganography is that it is easy to implement and has a high capacity for embedding secret messages.

Nowadays, there are many publicly available steganography tools which allow the use of more than 1-LSB and different colour components (RGB) for embedding that make the development of such a detection method very important. In this paper we focus on detecting 2-LSB replacement, which could be used in both colour and grayscale images. However, 2-LSB replacement is harder to detect than 1-LSB due to the complex changes in pixel values.

The steganalysis methods of 2-LSB detection are quite new and there are only few papers proposing such a method, which we explain them in brief below:

- (Luo, Wang, Yang, & Liu, 2006) Proposed a method of detecting 2-LSB steganography based on the quartic equation, which generates a finite-state machine using the image pixel sample pairs. This method also could detect and estimate the amount of secret message. However, as claimed by (Niu, Sun, Qin, & Xia, 2009), the calculation is too complex and takes a long time for analysis.
- The structural analysis extended by (Ker, 2007) to detect 2-LSB steganography using statistics of many variances to build the equation of estimation, but it involves lots of calculation and is considered a complex method for detection.
- (Zhang, Gao, & Bao, 2009) Proposed a method of detecting 2LSB embedding steganography based on the statistical characteristics in 2LSBs of the image pixel values. As claimed by the author, the detection rate could reach up to 90% when the embedding rate is 20% or more.

- In (Niu, Sun, Qin, & Xia, 2009) a steganalysis algorithm is proposed to detect 2-LSB embedding and estimate the message length. The method constructs the weighted stego image and estimates the message length with a least square equation, which makes the detection faster and more accurate. They also claimed that their results were better than the ones obtained by (Ker, 2007).

Also, there are some methods for detecting multiple LSB embedding steganography, which 2-LSB is a part of it, like;

- Fridrich's concept of weighted stego has been extended by (Yu, Tan, & Wang, ICIP 2005, 2005) to detect n-LSB steganography, which could also estimate the message length. This method, as claimed by (Yu & Babaguchi, Weighted stego-image based steganalysis in multiple least significant bits, 2008), has drawbacks of the low accuracy and having assumptions on the cover image.
- Another estimation method of detecting n-LSB embedding has been proposed based on weighted stego-image by (Yu & Babaguchi, Weighted stego-image based steganalysis in multiple least significant bits, 2008), which puts no assumption on the cover image. As claimed by the author, their method has a very low computation complexity with a clean estimation formula. The method could accurately detect the existence of the secret message and estimate the embedding ratio.
- A method of detecting MLSB (multiple least significant bits) steganography is proposed by (Yang, Liu, Luo, & Liu, 2008) based on the transition relationships among some trace subsets. The method could estimate the amount of embedded secret messages and is also defined as a very accurate method of detection by the author.
- Based on sample pair analysis, (Luo, Liu, Yang, Lian, & Zeng, 2012) proposed a method to estimate the embedding ratios of multiple bit-planes image steganography. It combines suitable trace sets to estimate the modification ratios in Gray code bit-planes. As claimed by the author, the proposed method can estimate the embedding ratios of multiple bit-planes with smaller errors in comparison to previous steganalysis methods.

In this paper we propose an Extended Pairs of Values (EPoV) analysis to detect 2-LSB steganography in colour and grayscale images. We analysed the performance of the classifier over a set of 3000 random images from ASIRRA, after converting them into grayscale with no change to the image dimensions. The embedded message was totally random in each case; this would make it very close to any secret message after encryption. Another feature of the proposed method is that it is very easy to understand and implement and could be used as an automated tool to analyse a bulk of images.

This paper is organized like the following, the conventional pairs of values analysis is explained in section 2. Section 3 clarifies some important basics of this research. Then, in section 4, the proposed method is explained in detail and followed by experimental results. Section 6 is about the estimation of message length. Finally there is a conclusion in section 7.

2. Conventional Pairs of Values Analysis

Most LSB embedding steganography overwrite the LSB of the pixel value with the secret message bits. This will transform pixel values into another value which is different only in their LSB plane. These values are known as pairs of values (PoV), as shown in figure 1.

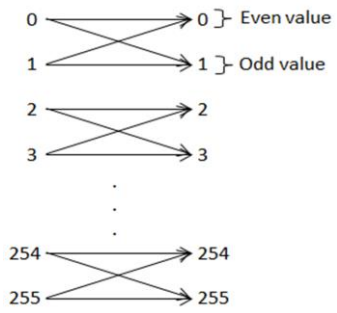

Fig. 1. Possible changes with LSB embedding.

The chi-square attack, proposed by (Westfeld & Pfitzmann, 2000), can detect the sequentially embedded LSB steganography in images. The nearly equal distribution of bits in the secret message, especially in encrypted versions, affect the LSB of the pixel values and will generate a close to equal number of occurrences of values in each PoV after embedding. These close to equal occurrences are usually not found in clean images. As the embedding process transforms values into each other in PoVs, the theoretically expected frequency for stego image will be the arithmetic mean of PoVs. Hence, the probability of having secret messages embedded would be measured by the degree of similarity between the theoretically expected frequency and the observed sample distribution, as explained below.

- The method considers *K* categories (*K*=128 for 8-bit pixel values) of PoVs and each observed pixel value from the image lies in one of them, for example values from (2k and 2k+1) will fall in category k.
- The arithmetic mean of occurrences in each PoV represents the theoretically expected frequencies; any values of theoretically expected frequencies less than 5 will be omitted.

$$n_k^* = \frac{\left|\{color | sortedIndexOf(color) \in \{2k, 2k+1\}\}\right|}{2}$$

- Without restricting the generality, the even values of frequency of occurrences in the observed sample have been taken in each PoV which measured by the following.

$$n_k = |\{color | sortedIndexOf(color) = 2k\}|$$

- Then the chi-square (X^2) is applied with $k-1$ degree of freedom.

$$X_{k-1}^2 = \sum_{i=1}^{K} \frac{(n_i - n_i^*)^2}{n_i^*}$$

- The integration of the density function is used to find the probability of embedding (P), assuming the equal distributions of n_i and n_i^*.

$$P = 1 - \frac{1}{2^{\frac{k-1}{2}}\Gamma\left(\frac{k-1}{2}\right)} \int_0^{X_{k-1}^2} e^{-\frac{x}{2}} x^{\frac{k-1}{2}-1} dx$$

- The probability of embedding P becomes nearly 0 when X_{k-1}^2 approaches to infinity, and it approaches to 1 for small value of X_{k-1}^2.

3. Backgrounds To The Proposed Method

This research considers a number of facts and conclusions based on other studies and experimental findings.

- The correlation between bit planes is different in clean and stego images. As claimed by (Avcıbaş, Kharrazib, Memon, & Sankur, 2005), hiding information in any bit plane lowers down the correlation with other continuous bit planes. Their steganalysis method was based on the binary similarity measure in lower bit planes.
- We define the regularity rate as the average of all similar (xxxxxx00, xxxxxx11) to different (xxxxxx01, xxxxxx10) 2-LSBs in each category (EPoV) for the entire image. As explained in section 6, it will help estimating the message length.
- Our study considers the correlation between 7th and 8th bit planes, as 2LSB steganography overwrites these two. In stego images, the correlation between 7th and 8th bit planes are not like clean images which are expected to be more random based on the fact that LSB steganography is like adding noise to the lower bit planes. To support this consideration we have analysed 44153 images; 24761 of them were random images from Google (Khalind, Hernandez-Castro, & Aziz, 2013), and 19392 images were from ASIRRA[1] (Animal Species Image Recognition for Restricting Access) public corpus pet images. On average, 96.3% of them had a higher rate of same 2LSB (xxxxxx00, xxxxxx11) than different ones (xxxxxx01, xxxxxx10) in their pixel values (RGB) for each EPoV. In other words, their regularity rates were greater or equal to 1, as shown in table I. These rates completely change after 2-LSB steganography has taken place as explained in section 6.

TABLE I
THE PERCENTAGE OF IMAGES WITH OVERALL REGULARITY RATES EQUAL OR GREATER THAN 1

Image Group	No. of Images	R	G	B
Random images from Google	19392	98.0%	96.9%	98.5%
ASIRRA pet images	24761	95.6%	93.7%	96.5%

R, G, and B represent the colour components Red, Green, and Blue respectively.

- The sequential 2LSB embedding is applied starting from the top left pixel to the bottom right and each 2LSB of the pixel value is replaced with 2-bits of the random message.
- The randomness of the sample images used is considered very important in this research, as steganalysis methods should work in real circumstances to be considered as a useful tool. This random set of images was taken with no filtering; they have different size, resolution, noise,

[1] The ASIRRA pet images were downloaded in a compressed folder from the following link (ftp://research.microsoft.com/pub/asirra/petimages.tar) on 11th of June 2012.

texture…etc. However, this approach will negatively affect the performance of the proposed method, but we consider the practical aspect of it especially for digital forensics analysis.

4. Extended Pairs of Values Analysis

Changes to two LSB values are much harder than one LSB to detect, as it leads to complex changes in pixel values. This method uses the chi-square attack with new form of pairs of values for detection; we name it 'Extended Pairs of Values' (EPoV). Each EPoV consists of four values based on the fact that two LSB steganography changes these four values into each other, as shown in figure 2.

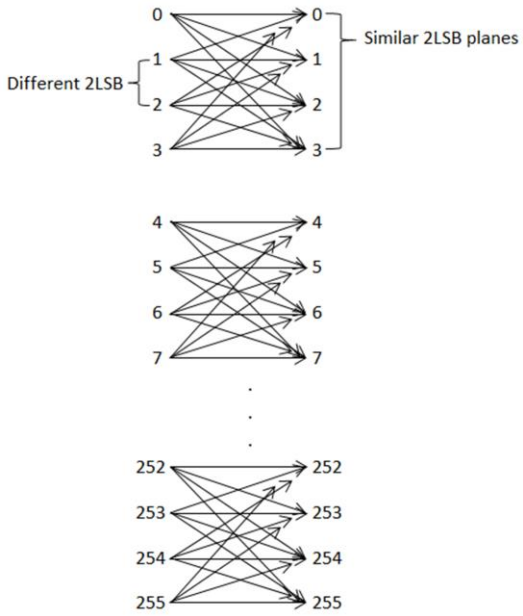

Fig. 2. Possible changes with two LSB embedding.

It could be noticed from figure 2, the sum of frequency of occurrences in each EPoV will stay constant, before and after embedding, as it puts boundaries for each group and values are changing within these scopes. Moreover, because the embedding process inserts noise to the pixel values, it is expected to have more frequencies of different 2LSB values (xxxxxx01, xxxxxx10) than same 2LSB values (xxxxxx00, xxxxxx11) in each EPoV.

The 7^{th} and 8^{th} bit planes of pixel values are not totally random in clean images, which will be the case after embedding process has taken place. Hence, it is uncommon for observed EPoVs (4k+1, 4k+2) and (4k, 4k+3) to be far from their arithmetic means in clean images. Also, for clean images it is more likely to have a higher rate of similar 2LSB values in each EPoV. Thus, the theoretically expected frequency after embedding is expected to be far from the arithmetic mean of the values in each EPoV, because of this, we still stick to the arithmetic mean in each EPoV. The similarity measure between the observed sample and the arithmetic mean would be the base of detection; being close to arithmetic mean indicates that the image is clean. Otherwise, if it was far from the arithmetic mean, indicates the existence of hidden content.

The proposed method of detecting 2LSB steganography uses the chi-squared attack as explained below.

- As shown in figure 2, *K* categories of extended PoVs are considered. Since it groups every four values in one category and pixel values (or colour component values RGB) range from 0 to 255, the value of *K*=64. Each colour value from the image pixels lie in one of those EPoVs, as an example; the values (4k, 4k+1, 4k+2, and 4k+3) all belong to category k.
- Two vectors with K elements are used X^{64x1} and Y^{64x1}, such that;

$$X_k = frequency(4k\ and\ 4k+3); 0 \leq k \leq 63$$

$$Y_k = frequency(4k+1\ and\ 4k+2); 0 \leq k \leq 63$$

  The frequency of values with similar 2LSBs in each category is held by X, and different 2LSBs by Y.
- Without losing the generality, this method considers the similar 2LSB values in the EPoVs in such a way that X_k measures the frequency of occurrences in category k.
- The theoretically expected frequency of occurrences for a stego image should be far from the arithmetic mean in each category. However, this is not the case for clean images which are closer to the arithmetic mean. That is why the arithmetic mean of each category is vital and calculated by;

$$Z_k = \frac{X_k + Y_k}{2}$$

- To measure the degree of similarity between the observed frequency of occurrence and the arithmetic means, the chi-squared (X^2) is applied with $k-1$ degree of freedom.

$$X_{k-1}^2 = \sum_{i=1}^{K} \frac{(X_i - Z_i)^2}{Z_i}$$

- Unlike clean images, the X_{k-1}^2 is expected to be relatively high for stego images as the X_i should be relatively far from Z_i.
- The probability of embedding P is calculated by integration of the density function with an upper limit of X_{k-1}^2, under the condition that the distributions of X_k and Z_k are not equal and relatively far.

$$P = \frac{1}{2^{\frac{k-1}{2}} \Gamma\left(\frac{k-1}{2}\right)} \int_0^{X_{k-1}^2} e^{-\frac{u}{2}} u^{\frac{k-1}{2}-1} du$$

The probability of embedding P converges to 1 as X_{k-1}^2 approaches infinity and for relatively small X_{k-1}^2 becomes much less than 1, which is affected by embedded data or noise insertion.

To analyse the image, the method checks the value of P from (1% - 100%) of the total image pixels. The continuity of P being equal to 1 within the entire image shows the availability of hidden content, otherwise the image is considered as clean, see in figures (3, 4, 5 and 6). However, according to experiments the value of P will not be stable until 5% of the image's total pixels analysed. As a refinement to the result we omitted the first 4% in finding the final value of P, which become the average of all Ps from 5% to 100% of the image's total pixels.

As it is clear from figures 3 and 4, the *P* value in the case of the clean image varies in all colour components (RGB) as coloured accordingly and for the stego image it continues being 1 for all colour components (RGB). It is also the case with grayscale version of Lena image as shown in figures 5 and 6.

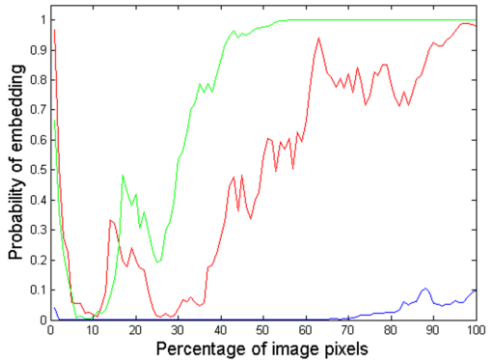

Fig. 3. The probability of embedding for Lena's 512x512 colour clean image.

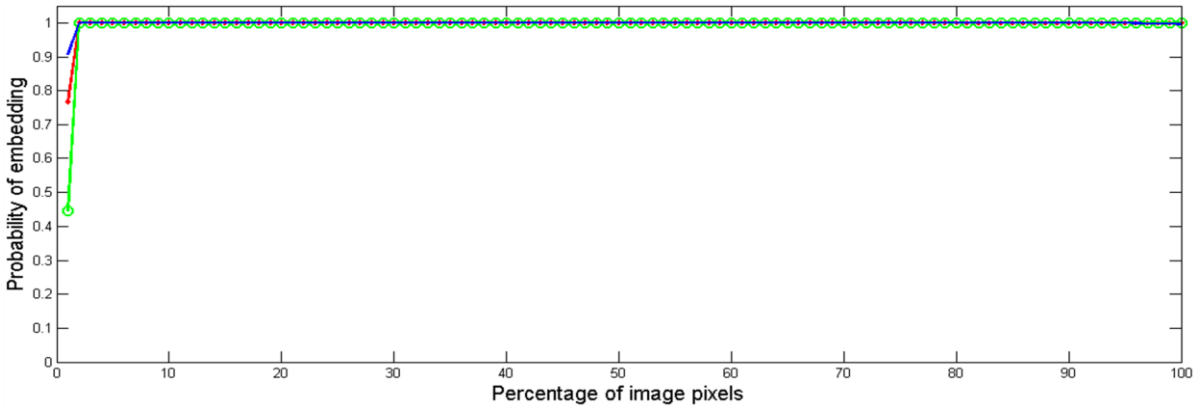

Fig. 4. The probability of embedding for Lena's 512x512 colour stego image.

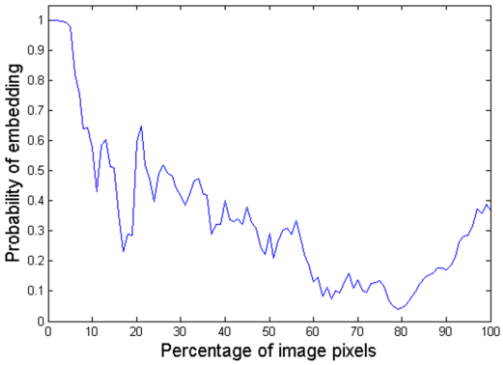

Fig. 5. The probability of embedding for Lena's 512x512 grayscale clean image.

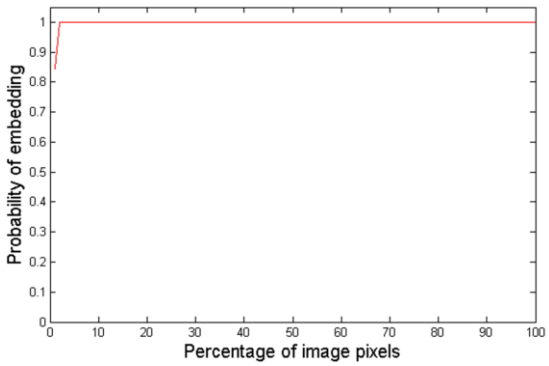

Fig. 6. The probability of embedding for Lena's 512x512 grayscale stego image.

5. Experimental Results

To analyse the performance of the proposed method, a set of 3000 random images from ASIRRA pet images are taken as cover objects after converting them into grayscale. The reason of choosing this set of images is that they are random images, as they originally taken from *petfiner.com*. Then, each image is embedded with a random message with a certain percentage of the total capacity (5%, 10%, 20%, 30%, …, 100%). For each percentage the detection method fed with 6000 images; 3000 stego images with the specified amount of embedded message together with 3000 original ones for classification.

According to the experimental results, shown in table II, the detection method could accurately detect 2LSB replacement, especially when the embedding rate reaches 10% of the image's total capacity. The true positive rate is very high, especially for the message length of 20% - 100% which is 0.997 to 0.999 with the accuracy of higher than 0.96. Also, the false positive rate, 0.074, is very low in comparison to the very well-known steganalysis tool like Stegdetect which was 0.1 for random set of images from Google for the default sensitivity value of 1 (Khalind, Hernandez-Castro, & Aziz, 2013).

TABLE II
THE EXPRIMENTAL RESULTS OF ALERTS, POSITIVE RATES, AND ACCURACY

Amount of embedded data	True Positives	True Negatives	False Positives	False Negatives	True Positive Rate	False Positive Rate	Accuracy
5%	815	2778	222	2185	0.272	0.074	0.599
10%	2499	2778	222	501	0.833	0.074	0.879
20%	2991	2778	222	9	0.997	0.074	0.962
30%	2996	2778	222	4	0.999	0.074	0.962
40%	2995	2778	222	5	0.998	0.074	0.962
50%	2998	2778	222	2	0.999	0.074	0.963
60%	2996	2778	222	4	0.999	0.074	0.962
70%	2998	2778	222	2	0.999	0.074	0.963
80%	2997	2778	222	3	0.999	0.074	0.963
90%	2995	2778	222	5	0.998	0.074	0.962
100%	2997	2778	222	3	0.999	0.074	0.963

Each time 6000 images are used; 3000 original clean images and 3000 stego images with the specified amount of data embedded

There are some very small variances in true positives between different amount of embedded data, especially from 30% to 100%, which are resulted from the randomness of embedded messages for each case.

Figure 7 shows the performance of the classifier in the form of ROC graph, the straight line from (0, 0) to (1, 1) indicates the random guess. Any curve located above this line is considered as better than random guess and the larger the area under the curve indicates the better performance of the classifier. For the proposed method, there are three curves labelled with the specified percentages (5%, 10%, and 20-100%), so the classifier is in its best performance when the amount of data are from 20% to 100%.

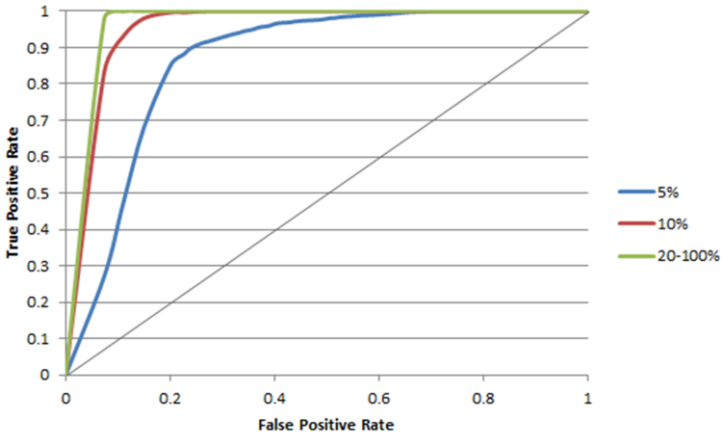

Fig. 7. The ROC curve of the proposed classifier.

6. Estimating The Message Length

The detection method only gives a decision of stego or clean for each image without specifying the embedded message length. To estimate the embedded message length, the regularity rate is considered, which affected directly by the amount of embedded message.

As it groups four values together, there would be 64 categories named as K. For each category it counts two types of frequency of occurrence; same (xxxxxx00, xxxxxx11) and different (xxxxxx01, xxxxxx10) two least significant bits, as the following.

$$Same2LSB(k) = frequencyOfOccurrence(4k \,\&\, 4k+3)$$

$$Different2LSB(k) = frequencyOfOccurrence(4k+1 \,\&\, 4k+2)$$

For each k, if Same2LSB(k) or Different2LSB(K) values were 0, then it set them both to 1 to eliminate the effect of this type occurrence and improving the accuracy.

Now, the regularity rate would be the average of all rates between same to different 2LSBs in each category, as shown in the equation below.

$$RegularityRate = \frac{\sum_{k=0}^{(K-1)} Same2LSB(k)/Different2LSB(k)}{K}$$

Based on experimental results of our set of 3000 images, as shown in table III, the embedding rate can be divided into five ranges; less than 0.65, 0.65-0.8, 0.8-0.95, 0.95-1, and greater than 1.

TABLE III
THE REGULARITY RATE VERSUS THE EMBEDDING RATE

Message amount	Regularity Rate				
	< 0.65	0.65-0.8	0.8-0.95	0.95-1	> 1
0%	0.00	0.00	0.00	0.18	0.82
5%	0.00	0.00	0.05	0.52	0.43
10%	0.00	0.00	0.31	0.45	0.23
20%	0.00	0.01	0.74	0.19	0.07
30%	0.00	0.10	0.83	0.05	0.02
40%	0.00	0.45	0.53	0.01	0.01
50%	0.01	0.78	0.21	0.00	0.00
60%	0.11	0.82	0.06	0.00	0.00
70%	0.46	0.52	0.02	0.00	0.00
80%	0.77	0.23	0.00	0.00	0.00
90%	0.91	0.09	0.00	0.00	0.00
100%	0.97	0.03	0.00	0.00	0.00

The values in each range represent the percentage of the regularity rate.

Each value of regularity rate represents the percentage of values within the specified range. Based on table III, we can derive another table that maps the regularity rate with the embedding rate, as shown in table IV. The proposed method now can accurately estimate the embedded message length. Of course, there is some overlap between certain ranges of the regularity rate and the message size, but still there are very clear boundaries and the ranges are well divided with a very high level of certainty.

TABLE IV
REGULARITY RATE AND THE AMOUNT OF EMBEDDED MESSAGE

Regularity Rate	Estimated amount of embedded data
larger than 1	0%
between (0.95 – 1)	5% - 10%
between (0.8 – 0.95)	20% - 40%
between (0.65 – 0.8)	50% - 70%
Less than 0.65	80% - 100%

This table is derived from the having majority of the specified percentage of embedded data that lies in the specified range of regularity rate.

7. Conclusion

In this study an extended pairs of values, EPoV, analysis is proposed to detect 2-LSB replacement steganography in colour and grayscale images. The experimental results showed that the detection method can accurately detect the existence of the secret message, especially when the embedding rate reaches 10% of the image's total capacity. It also could estimate the amount of embedded message in stego images based on five ranges of regularity rate. Moreover, the method is very simple to understand and implement without any complexity and could actively work on a random set of 3000 images with a completely random messages. As mentioned earlier, the method could also be applied on colour images and indicate which colour components R, G, and/or B are used for embedding.

Finally, it could be used as an automated tool by digital forensics analyst in their investigation process to analyse a bulk of images for hidden contents.

References

Avcıbaş, I., Kharrazib, M., Memon, N., & Sankur, B. (2005). Image Steganalysis with Binary Similarity Measures. EURASIP Journal on Advances in Signal Processing, 17, 2749–2757.

Ker, A. D. (2007, March). Steganalysis of Embedding in Two Least-Significant Bits. IEEE Tansactions on Information Forensics and Security, 2(1), 46-54.

Khalind, O., Hernandez-Castro, J., & Aziz, B. (2013, February). A study on the false positive rate of Stegdetect. Digital Investigation, 9(3-4), 235-245.

Luo, X., Liu, F., Yang, C., Lian, S., & Zeng, Y. (2012, April). Steganalysis of adaptive image steganography in multiple gray code bit-planes. Multimedia Tools and Applications, 57(3), 651-667.

Luo, X., Wang, Q., Yang, C., & Liu, F. (2006). Detection of LTSB steganography based on quartic equation. The 8th International conference of Advanced Communication Technology. 2, pp. 1199-1204. IEEE.

Niu, C., Sun, X., Qin, J., & Xia, Z. (2009). Steganalysis of two least significant bits embedding based on least square method. Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on, 3, 124-127.

Westfeld, A., & Pfitzmann, A. (2000). Attacks on Steganographic Systems. Lecture Notes in Computer Science, 1768, 61-75. Berlin: Springer-Verlag.

Yang, C., Liu, F., Luo, X., & Liu, B. (2008, December). Steganalysis Frameworks of Embedding in Multiple Least-Significant Bits. IEEE Transactions on Information Forensics and Security, 3(4), 662-672.

Yu, X., & Babaguchi, N. (2008). Weighted stego-image based steganalysis in multiple least significant bits. International Conference on Multimedia and Expro (pp. 265-268). IEEE.

Yu, X., Tan, T., & Wang, Y. (2005). Extended optimization method of LSB steganalysis. In Proceedings of IEEE Intetrnational Conference of Image Processing. 2, pp. 1102-1105. IEEE.

Zhang, K., Gao, H.-Y., & Bao, W.-s. (2009). Steganalysis Method of Two Least-Significant Bits Steganography. International Conference on Information Technology and Computer Science. 2, pp. 350-353. IEEE.