ENHANCING LSB STEGANOGRAPHY AGAINST STEGANALYSIS ATTACKS USING COMBINATIONAL LSBS

Yahya Belghuzooz and Ali Al-Qayedi
Etisalat University College, U.A.E

Keywords: LSB Steganography, Steganalysis.

Abstract: This paper describes an enhanced approach for hiding secret messages in the spatial domain of digital cover images such that the resulting stego-images are robust to steganalysis attacks. Firstly, different methods of hiding in the Least Significant Bits (LSBs) are comparatively discussed including the Sequential and the Random algorithms. Then our approach is illustrated which uses a combination of LSBs to store large amounts of secret information while maintaining robustness against detection by steganalysis attacks. The results achieved are commensurate to those obtained using widely available stego tools.

1 INTRODUCTION

Steganography, or the art of covert communication, has exhibited a considerable focus over the past few years following the claim that it could be heavily utilised for secret hidden communication between criminals. Consequently, *steganalysis*, a field that is concerned with how to detect the presence of secret messages and possibly reveal its content has become an important topic on its own.

A common approach to steganography involves hiding secret information within the Least Significant Bits (LSBs) of a cover image. However, LSB steganography is a spatial domain hiding technique which is known to be relatively weak compared to other transform domain hiding techniques (Westfeld and Pfitzmann, 2000), (Cole and Krutz, 2003), (Katzenbeisser and Petitcolas, 2000), (Provos and Honeyman, 2003) .

Using a single or a few LSBs in the hiding process can be invisible in visual terms; however, it provides a limited space for hiding secret information and is also easily guessable. On the other hand, using more LSBs can accommodate larger information but can result in clear visual and statistical discrepancies between the cover image and the stego-image, hence revealing the presence of a hidden secret for steganalysis algorithms.

In this paper we present a novel approach that enhances the robustness of LSB steganography against steganalysis attacks by using a combination of sequentially or randomly selected LSBs. Our approach enables storing large amounts of secret information while maintaining secrecy of its presence in the stego-image.

The remaining sections of this paper are organised as follows: Sections 2 and 3 describe the commonly used spatial domain LSB stegonagraphy and steganalysis methods. Section 4 gives an overview of the proposed approach. Results with discussion are shown in section 5. Finally concluding remarks with recommendations for future work are given in section 6.

2 LSB STEGANOGRAPHY IN THE SPATIAL DOMAIN OF IMAGES

There are many techniques for embedding secrets in the spatial domain of cover images most of which are weak methods to the extent that the secret message can be fully retrieved rather than just detecting its presence using steganalysis attacks.

One of the methods is hiding a signature in the header of the image using an application like FortKnox 3.55 (FortKnox, 2007) which may lead to destroying the resulting image.

Another method is hiding the secret message at the end of the image (or fusion within the image) which can be easily broken. Many systems are using

Belghuzooz Y. and Al-Qayedi A. (2007).
ENHANCING LSB STEGANOGRAPHY AGAINST STEGANALYSIS ATTACKS USING COMBINATIONAL LSBS.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 237-242
DOI: 10.5220/0002133602370242

this technique, a few examples are: Camouflage (Camouflage, 2007) , JpegX (JpegX, 2007) , Safe & Quick Hide Files (Safe, 2007) and Data Stash (DataStash, 2007).

The other main steganography embedding methods are discussed next according to their order of robustness against statistical steganalysis attacks.

2.1 LSB Sequential Embedding

In this process every byte of the image represents a color value which can be changed by 1 without leaving a trace in the output image. Thus the LSB of the image at position 2^0 is used (Cole and Krutz, 2003) (Katzenbeisser and Petitcolas, 2000), (Provos and Honeyman, 2003). The secret message is distributed sequentially on the LSB of each byte of the image.

The main limitation of this approach is that the secret message may change the LSB by a probability of 0.5, thus in the steganalysis process this can be utilized by checking the pairs which has 0.5 distribution of 1's and 0's. Where a pair is considered as any two bytes in which the 7 MSB are the same, for example 0101 1010 and 0101 1011 represent a pair.

2.2 LSB Random Embedding

This approach is somewhat similar to the previous one except that the LSB embedding is done randomly instead of being sequential. This technique is more robust to steganalysis attacks compared to the sequential technique because of distributing the secret message across the image without affecting the statistical property of the contiguous color values.

However; if the secret message uses all the available LSBs in the cover image then the embedding process can be easily detected by the steganalysis operations since all the LSBs were modified which is similar to the sequential method (Katzenbeisser and Petitcolas, 2000), (Lenti, 2000), (Johnson and Jajodia, 1998).

Any randomly selected LSB should not be reused again in the embedding process otherwise it will be overwritten. In our proposed system in section 4 we describe a swapping process that is developed to obtain an unrepeated sequence of LSBs.

2.3 Changing Pairs

Changing pairs can be considered as one of the hardest LSB embedding methods to detect by steganalysis. It is relatively better than the random method because it randomizes the embedding process without modifying the bit that will be used in the embedding process.

This process increments or decrements the color value by an odd value (mainly 1 or 3) which may result in changing the whole 8 bits used to represent a color value but the actual colour value (intensity) is only incremented or decremented by a small amount such that the difference is not easily noticeable by the human eye (Soukal and Goljan, 2005) .

For example, if the color value is 127 (0111 1111 in binary) and the secret bit is 0, then in the normal LSB embedding the result will be 0111 1110 which is 126 but the pair 0111 111 was not changed, however when using the changing pair method if the selection was to increment then 127 will become 128 which is 1000 0000, so the LSB still contains the secret bit which is 0 but at the same time the pair has been changed from 0111 111 to 1000 000 which clearly makes a difference in statistical terms for the steganalysis.

The *Changing Pairs* algorithm can be summarized as follows:

1. Read all Bytes of the Image.
2. Select one byte randomly using a key.
3. Decide whether the LSB of that byte is to be changed by comparing the LSB with the secret bit.
4. If yes, flip a coin to decide whether to increment or to decrement by an odd value.
5. Repeat processes 1, 2, 3 and 4 until the end of the secret message.
6. Write all bytes to the output Image.

3 SPATIAL-DOMAIN STEGANALYSIS

Even though stego-images can rarely be spotted by the naked eye, they usually leave behind some traces or statistical hints that they have been modified. It is that discrepancy which an analysis tool may be able to detect. Since some techniques and their effects are commonly known, a statistical analysis of an image can be performed to check for the presence of a hidden message.

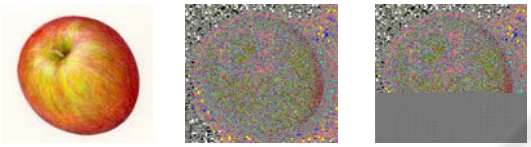
There are two main types of steganalysis methods: the visual steganalysis and the statistical steganalysis. These tests can be applied on a given image to check if a secret message is embedded in it or not.

3.1 Detection using Visual Steganalysis

Some images have the same pattern in the LSB and if that pattern is changed by embedding a secret message, which is simply a random noise, then the original pattern will become random. To detect that kind of stego-images; an enhancement of the LSB and a visualization of it can be done. If there is a secret message that was embedded sequentially then it can be seen as a strip of random LSBs (Westfeld and Pfitzmann, 2000).

Figures 1 (a), (b), and (c) illustrate this effect by showing the visual difference in steganalysis terms between the original cover image and the resulting stego-image after embedding the secret message sequentially.

This process of steganalysis is not applicable for randomly embedded stego-images or for images that already have a randomly patterned LSBs like an image with a colorful background instead of the one shown in Figure 1 (a).



(a) **(b)** **(c)**

Figure 1: (a) Cover Image (b) Steganalysis of Cover Image (c) Steganalysis of Stego-Image.

3.2 Statistical Steganalysis

This type of steganalysis uses statistical properties of the stego-image. One type of steganalysis is the Chi-Square test which checks the number of occurrences of pairs in the secret message as shown in (1) and (2) (Westfeld and Pfitzmann, 2000). The chi-square attack is a steganalysis method developed to recognize some types of steganographic embedding in the LSBs of an image's pixel values. When the chi-square attack is applied to an image, it produces a graph of the probability of steganographic embedding vs. the sample size of the image tested. By examining this graph an analyst can determine whether or not an image contains steganographic embedding.

$$x_{k-1}^{2} = \sum_{i=1}^{k} \frac{(n_i - n'_i)^2}{n'_i} \quad (1)$$

$$p = 1 - \frac{1}{2^{\frac{k-1}{2}}\Gamma\left(\frac{k-1}{2}\right)} \int_{0}^{x_{k-1}^{2}} e^{\frac{-x}{2}} x^{\frac{k-1}{2}-1} dx \quad (2)$$

Where:

n_i is the observed population in the i_{th} bit and n'_i is the expected population in the i_{th} bit.

p is the probability of the chi-square statistics when the distributions of n_i and n'_i are equal.

4 SYSTEM OVERVIEW

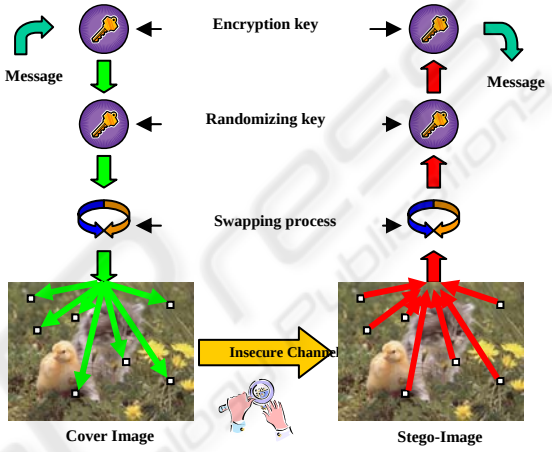


Figure 2: The behavior model of the system.

In our system, as shown in Figure 2, we use LSB randomization and swapping to avoid overwriting a previously selected LSB. This is achieved as follows:

1. Predefine a sequence of unrepeated random numbers.
2. Generate another random sequence that is equal in size to the sequence generated in 1.
3. Swap between the numbers of the unrepeated sequence using the random number generated in step 2 as an index to the selected candidate from the unrepeated sequence.
4. Repeat steps 1 to 3 until all secret message bits are filled into the cover image LSBs.

The system enables embedding secret messages in 24-bit BMP images. It allows the user to control the use of a combined number of LSBs from the Red, Green and Blue color values to highlight the difference of using combinational LSBs compared to the previously discussed steganalysis methods.

Our system is not intended for secret communication over low bandwidth channels since it is not practical to use a 24-bit cover image as a

carrier to transmit it over a limited bandwidth channel. Also the use of a lossy compression such as JPEG can destroy the LSBs containing the secret message. Thus, the system can be used for secure communication of high quality imaging applications such as medical imaging.

Also the system could be modified to handle a set of small GIF images that can be used as carriers for the secret message. For example, in a Web gallery where a single secret message could be distributed over a set of GIF images. This would result into an extra storage capacity for the secret message and at the same time an increase in the systems robustness against steganalysis attacks.

5 RESULTS AND DISCUSION

The chi-square steganalysis test is applied on the stego-images generated by all the steganographic methods described before using 1 LSB from each color value. Figure 3 represents the stego-image generated using 1 LSB sequential embedding of a 60 KB secret message into a 24-bit BMP cover image of (896 by 674) pixels.

It can be seen from Figure 4 that the probability of embedded secret is high, thus the chi-square successfully detected the presence of a secret message, whereas Figure 5 shows the chi-square of the original cover image with a zero probability due to the absence of a hidden secret. Figures 6, 7 and 8 respectively show the chi-square of the previous experiment but now using the Random method, then using a public domain steganography tool called S-Tools (S-Tools, 2007), and finally using the changing pair method.

As can be seen in Figures 6 and 7, the chi-square is not effective in detecting the size of the secret message but it can detect a small portion of the embedding that has taken place at the beginning of the image. A similar result was observed with the S-Tools chi-square graph in Figure 7, but with a lower probability.

Figure 3: Stego-image by 1 LSB sequential method.

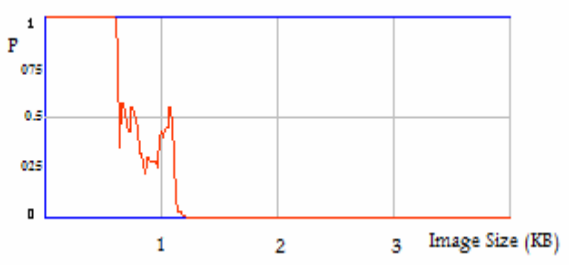


Figure 4: Chi-square of stego-image by the 1 LSB sequential.

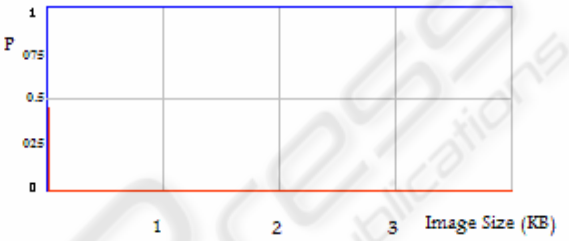


Figure 5: Chi-square of the cover image.

In Figure 8 the chi-square could not detect the presence of the secret hidden with the changing pair method since it changes the pair to another pair that results in a totally deferent statistics. In this case, the result of the chi-square is similar to that of the original cover image.

The system was then used to generate a combination of 3 LSBs from Red, 3 LSBs from Green and 2 LSBs from Blue. The results of Figure 12 and 13 show an improvement in the stego-images against steganalysis detection because some pairs were changed in the embedding process.

Figures 4 and 9 clearly show an improvement in the robustness of the stego-image generated by the combination LSBs method over the single LSB sequential method. A similar conclusion can also be deduced from Figures 6 and 10 for the random methods.

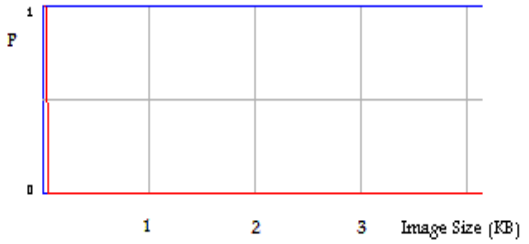


Figure 6: Chi-square of stego-image by 1 LSB random method.

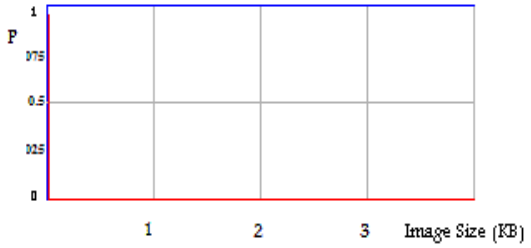


Figure 7: Chi-square of the S-Tools stego-image.

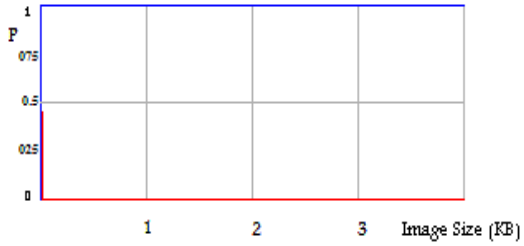


Figure 8: Chi-square of stego-image by changing pair method.

The changing pair method is still better in terms of steganalysis but it is limited to 1 LSB only which reduces the capacity of the carrier for the secret message. On the other hand, the combination method, especially the random one, provides 3 times the capacity of the changing pair method, and is not easily detectable by steganalysis.

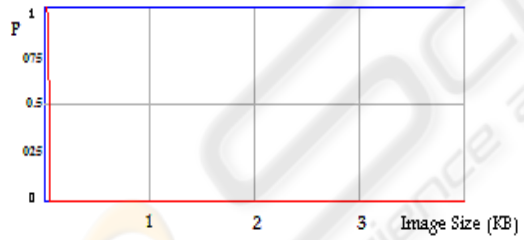


Figure 9: Chi-square generated using 3, 3, 2 LSBs combinational sequential method.

Figure 11 shows the average value of the LSBs of the cover image starting at 1 and then dropping to 0.5 which indicates that the image at the start contains no random distribution of 1s and 0s in the LSBs but then the distribution becomes random which is common in some patterned images.

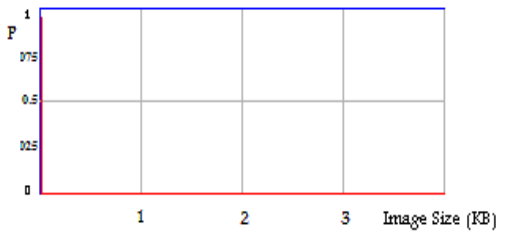


Figure 10: Chi-squa re of stego-image by 3, 3, 2 LSBs random method.

In Figures 11 to 15 the sequential method forces the average value to become close to 0.5. The random and the S-Tools average numbers are less than 1 but still not close to 0.5. In the pair method however, the average LSB is still at 1, which is indicating an embedded secret.

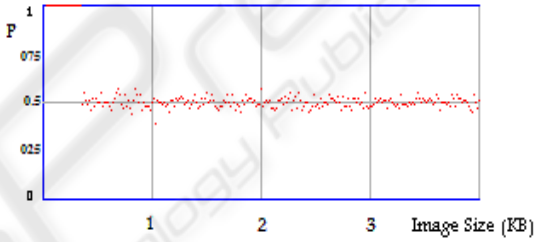


Figure 11: Average LSB value of the cover-image.

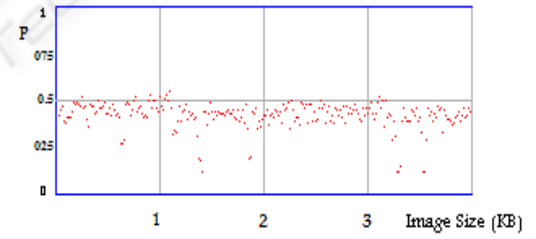


Figure 12: Average LSB value of the stego-image generated by the 1 LSB sequential method.

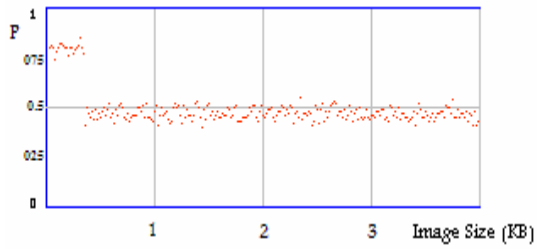


Figure 13: Average values of stego-image generated by the 1 LSB random method.

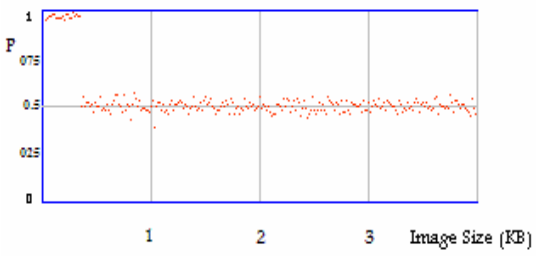


Figure 14: Average values of stego-image generated by S-Tools.

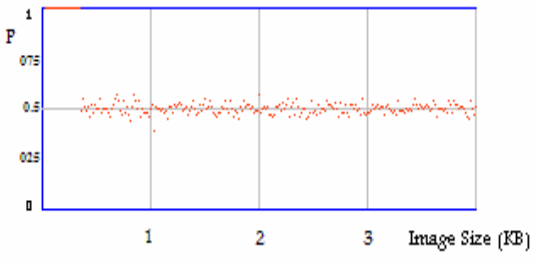


Figure 15: Average values of stego-image generated by the pair method.

Figure 16 shows that the average values generated using 3, 3, 2 LSBs combinational sequential method drop below 0.5, this is an advantage over the pair method as it improves robustness against steganalysis.

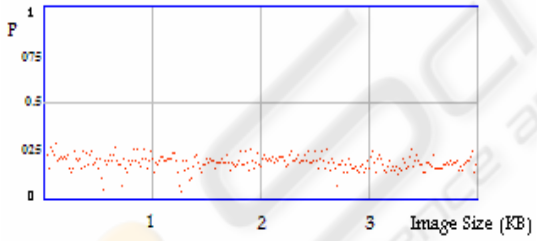


Figure 16: Average values generated by the 3, 3, 2 LSBs combinational sequential method.

6 CONCLUSION

In this paper we have presented a steganographic system that uses a combination of LSBs to improve the storage capacity of the stego-image and to increase its robustness against steganalysis attacks. The chi-square and the average value LSB results obtained from our combinational algorithm are significantly better than those achieved with the sequential and random 1 LSB, S-Tools and changing pair methods.

REFERENCES

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

Cole, E. and Krutz, R. (2003) Hiding in Plain Sight: Steganography and the Art of Covert Communication, Wiley, John & Sons.

Katzenbeisser S. and Petitcolas, F. (2000) Information Hiding Techniques for Steganography and Digital Watermarking, Artech House.

Provos, N. and Honeyman, P. (2003) Hide and Seek: An Introduction to Steganography, IEEE security & privacy 1:33, 32-44, IEEE Computer Society .

FortKnox (2007), Online, Available: http://www.softpedia.com/get/Security/Firewall/FortKnox-Personal-Firewall.shtml , 31 May 2007

Camouflage (2007), Online, Available: http://camouflage.unfiction.com/, 31 May 2007.

JpegX (2007), Online, Available: http://nerdlogic.org/jpegx/ , 31 May 2007

Safe (2007), Safe & Quick Hide, Online, Available: http://www.freedownloadscenter.com/Utilities/File_Encryption_Utilities/Safe_Quick_Hide_File_and_Folder.html [31 May 2007]

DataStash (2007), Online, Available: http://www.skyjuicesoftware.com/software/ds_info.html, 31 May 2007

Lenti, J. (2000) Stegangraphic Methods, PERIODICA POLYTECHNICA,VOL. 44, NO. 3–4, PP. 249–258.

Johnson, N. and Jajodia, S. (1998) Steganography: Seeing the Unseen, IEEE Computer, vol. 31, no. 2, pp. 26-34.

Soukal, D. and Goljan, M (2005) Maximum Likelihood Estimation of Secret Message Length Embedded Using +-K Steganography in Spatial Domain, Proc. SPIE Electronic Imaging San Jose, CA, January 16-20, pp. 595-606.

S-Tools (2007), Online, Available: http://www.snapfiles.com/get/stools.html, 31 May 07