Variable Least Significant Bits Stegnography

Sahib Khan[1]
Dept. of Telecommunication Engineering
University of Engineering and Technology, Taxila
Taxila, Pakistan
engrsahib_khn@yahoo.com

Muhammad Haroon Yousaf[2]
Dept. of Computer Engineering
University of Engineering and Technology, Taxila
Taxila, Pakistan
haroon.yousaf@uettaxila.edu.pk

Abstract—**I am presenting a novel Stegnography technique named Variable Least Significant Bits (VLSB) Stegnography. It is an influential and secure data embedding technique, with varying data hiding capacity and distortion. The capacity is increased at the cost of distortion and distortion is decreased by the sacrifice of hiding capacity. A trade-off is made between the two parameters to get desired results. This methodology is much immune to Steganalysis because a large number of combinations (Keys Size) are possible to hide data in a cover image. Key size is proportional to cover image size and the variation in number of bits used for data embedding. The key is kept secret which make it complicated for any third party to retrieve hidden information. The intended party is provided with the key, to extract the information.**

Keywords: VLSB Stegnography; Steganalysis; Key size; Cover Image.

I. INTRODUCTION

The word Stegnography literally means, "covered writing" and encompasses methods of transmitting secret messages through innocuous cover carriers in such a manner that the existence of the embedded messages is undetectable. Stenography hides the secret message within the host data set and its presence is imperceptible [1]. The goal of Stegnography is to hide messages inside other harmless messages in a way that does not allow any enemy to even detect that there is a second secret message present [2]. In the field of Stenography, some terminology has been developed. The term cover is used to describe the original, innocent message, data, audio, still, video or any other digitally represented code or transmission [3] [4]. Presently the technology being mostly used is Digital Images [5]. Care needs to be taken in the selection of the cover image, so that changes to the data will not be visible in the stegno-image [6]. Another very important term is Message, which is the data/information to be concealed in cover file. Stegno-file is the medium having the secret message.

Greeks used it writing message on some material and later covering it with wax, tattooing messages on bald head, later growing hair to cover it up. In World War II invisible inks were used to write messages in between the lines of normal text message [7]. World War II saw the use of microdots by Germans. In microdots technology, photograph of secret message taken was reduced to size of a period. This technology was called “the enemy’s master piece of espionage” by FBI director J. Edgar Hoover [8].

Information hiding is an emerging research area, which encompasses applications such as copyright protection for digital media, watermarking, fingerprinting, and Stegnography. All these applications of information hiding are quite diverse. In watermarking applications, the message contains information such as owner identification and a digital time stamp, which is usually applied for copyright protection. This adds to copyright information and makes it possible to trace any unauthorized use of the data set.

A broad overview of data embedding and watermarking methods is available in [9]. Additional readings, software, and resources used in researching Stegnography and digital watermarking is available at [10]

Many techniques were used to implement Stegnography both in spatial domain or transform domain. Discrete Cosine Transform (DCT) is used and data is hidden by exploiting the coefficient of DCT of the cover image. In this technique Image data is divided into square blocks, for example, 8×8 pixels, which are transformed to DCT coefficients. A DCT coefficient matrix $[dij]$ that corresponds to a block of pixel [bij] is calculated. Secret information is embedded by rewriting DCT coefficients [dij] following a certain rule. If some appropriate rule is used for rewriting of DCT coefficients, the quality of the reconstructed image will be almost the same as the original [11]. Wavelet Transform is also used for data hiding.

The most important technique implemented in spatial domain is 4 Least Significant (4LSB) Stegnography. In 4LSB Stenography four least significant bits are used for substitution [12]. The 4LSB method is implemented for color bitmap images (24 bit and 8 bit i.e. 256 color palette images) and wave files as the carrier media

It would appear that 4 LSB is good method of Stegnography due to its tremendous information capacity, less error and more secured. Using 4-LSB methods we can exchange secret messages over public channel in a safe way [13], [14].

All algorithms employed for any type of format have pros and cons and depend upon the environments used. It also depends upon the information to be embedded. Various techniques developed were compared [15].

II. VLSB STEGNOGRAPHY

The 4LSB Stegnography has a limited and fixed data hiding capacity of 50% i.e. we need a cover file of almost double size

as that of message file. To hide larger data in cover file in more secure way, a new technique called Variable Least Significant Bits (VLSB) Stegnography is devised. In this proposed, VLSB Stegnography, technique variable amount of data is hided in every individual pixel of each sector of the cover file, instead of a fix data as in 4LSB Stegnography.

Variable Least Significant Bits Stegnography is a technique that uses variable number of bits of cover file for data embedding. In this technique cover image is divided in various sections/sectors and "Bi" numbers of bits are used for data hiding in the pixels of that section. The number of bits "Bi" is varied from sector to sector.

The division of cover image in various sectors is a very critical in implementation of VLSB Stegnography. A proper algorithm is needed to implement the suggested Stegnographic technique. The algorithm should be capable of providing larger hiding capacity with least possible distortion. This will open a new research area for the researchers to play with VLSB Stegnography.

A. Hiding Capcity

The selection of pixels for a specific sector and amount of bits to be embedded in each pixel of that sector play vital role in determining hiding capacity of VLSB Stegnography. To a hiding capacity of 50% or more the number bits embedded in each pixel should not be less than four.

If there is an image of rows "R" and Columns "C", then total number of pixels "N" is:

$$N = R \times C \tag{1}$$

The image is divided in a number of sectors "Ns", the number of pixels "Ps" in each sectors are:

$$Ps = \frac{N}{Ns} \tag{2}$$

As in gray scale image each pixels is represented in 8 bits. The total space available in cover image is:

$$B_{total} = N \times 8 \tag{3}$$

Similarly the space available in each sector "Ps" is

$$Bs_{total} = Ps \times 8 \tag{4}$$

The Variable Least Significant Bits Stegnography provides user liberty to embed any number of bits (from 1 to 8 bits) of message in each pixel of a specific sector of cover image. The total data to be hidden in a section is dependent on the number of bits substituted in each pixel.

If a number of bits "Bi" is hided in each pixel of a sector Si, having a number of pixels "Psi" in it. The total amount of data embedded "Di" in sector "Si" is given by:

$$D_i = Psi \times Bi \tag{5}$$

And consequently the total amount of data "D_{total}" embedded in cover image is

$$D_{total} = \sum_{i=1}^{Ns} D_i \tag{6}$$

$$= \sum_{i=1}^{Ns} (Psi \times Bi) \tag{7}$$

The capacity C of VLSB Stegnography is given by

$$C = \frac{D_{total}}{B_{total}} \times 100 \tag{8}$$

$$= \frac{\sum_{i=1}^{Ns} D_i}{N \times 8} \times 100 \tag{9}$$

$$= \frac{\sum_{i=1}^{Ns} (Psi \times Bi)}{N \times 8} \times 100 \tag{10}$$

B. Key Size

The VLSB Stegnography is very immune to Steganalysis; because of its built-in encryption. The aim of Stegnography is to hide data in a cover file in non-perceivable manner but if someone comes to know about the secret he has to try a lot of combination to extract the hidden data.

Consider an image having rows "R" and columns "C", will have a total of "N" pixels in the image. In gray scale image each and every pixel's intensity can be represented with 8 bits. So there is liberty of 1 to 8 bits substitution. The number of bits considered at a time "Bi" vary from 1 to 8 ($1 \le Bi \le 8$), but to get a data hiding capacity of more than 50% the "Bi" should be greater than 3 ($4 \le Bi \le 8$).

If the cover image is divided in "Ns" number of sectors, each with "Psi" number of pixels. The will be a total of "Cpi" possibilities for a pixel to be part of sector "Si'.

$$C_{pi} = C_1^{Ns} \tag{11}$$

Each sector can exposed to a "Bi" number of bits substitution. So the total possible combinations for each sector "Si" are "Csi".

$$C_{Si} = C_{Bi}^8 \tag{12}$$

So the total possible keys/combination "K" for the whole cover image is:

$$K = C_{pi} \times C_{si} \tag{13}$$

$$K = C_1^{Ns} \times C_{Bi}^8 \tag{14}$$

If the number of sectors "Ns" is so large that it becomes equal to the number of pixels "N", the each sector will consists of only one pixel. Then each sector will have "Csi" different

combinations for a single pixel so total possible keys/combinations "K" for a cover image to hide data is given as below.

$$C_{pi} = C_1^{Ns} \tag{15}$$

$$Ns = N \tag{16}$$

$$C_{pi} = C_1^{N} \tag{17}$$

$$C_{pi} = N \tag{18}$$

$$C_{si} = C_{Bi}^{8} \tag{19}$$

Putt value of C_{si} and C_{pi} in $K = C_{pi} \times C_{si}$ we get

$$K = N \times C_{Bi}^{8} \tag{20}$$

Putting the value of N in above equation

$$K = (R \times C) \times C_{Bi}^{8} \tag{21}$$

Number of possible combinations "K" is the number of possible ways to hide data in a cover image. Larger the value of "K" more difficult it would be for an unauthorized person to extract data from the innocent Stego-Image even if he/she came to know that some data is hidden in the image. One must have to try "K" various combinations to reconstruct data by hit and trial method. K is proportional to cover image size, larger the size of the cover media VLSB methodology will be more difficult to break and get the hidden message.

C. SNR and PSNR

The quality of the stego-image is measured quantitatively by calculating signal to noise ratio (SNR) and peak signal to noise ratio (PSNR). For stego image the SNR is calculated in Decibels as

$$SNR = -10Log\left[\frac{sum((Cover - Stego)^2}{sum((Cover)^2)}\right]^{-1} \tag{22}$$

And PSNR in Decibels is calculated as

$$PSNR = -10Log(Mean((Cover - Stego)^2)) \tag{23}$$

III. IMPLIMINTATION OF VLSB STEGNOGRAPHY

To hide data/information in a cover image in more secure manner Variable Least Significant Bits Stegnography is implemented. It's a powerful and influential methodology. It can be implemented in various ways. A proper and well designed algorithm is needed to implement VLSB Stegnography. The algorithm should be capable of providing larger hiding capacity with least possible distortion. I have implemented, in this paper, VLSB Stegnography by selecting a cover image and a message file (image). Message file is converted into a continuous bits stream. Then a pixel of the cover image is selected and "Bi" bits of the message are embedded in that pixel of the cover image by replacing "Bi" number of least significant bits of the pixel with "Bi" bits of the message. The number of bits "Bi" is varied from pixel to pixel. The processed pixels having hidden data are combined, making an innocent Stego Image.

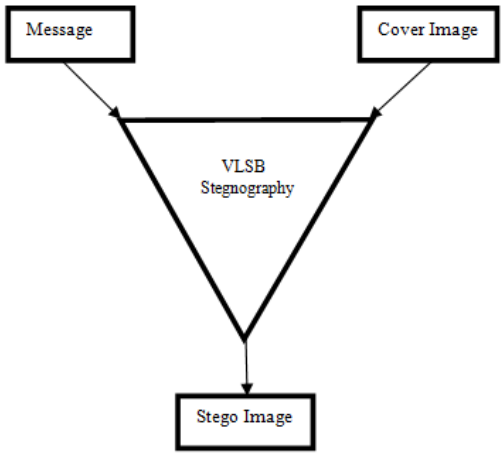

Figure 1: VLSB Stegnography

The VLSB Stegnography can be implemented for any range of values of "Bi" depending application, hiding capacity required and affordable distortion in Stego image.

IV. EXPERIMENTAL RESULTS

The VLSB Stegnography is implemented by embedding "Bi" bits of message in a pixel of cover file. To get more fair results with less distortion all possible values of "Bi" are utilized i.e. "Bi" is varied from 1 to 8 ($1 \le Bi \le 8$) or any other value in this range. The VLSB Stegnography is implemented for different combinations $j \le Bi \le 8$ of "Bi" where $1 \le j \le 8$ the results obtained for j=1 and 2 are significant with capacity of 56% and 63% respectively. The Stego images for $1 \le Bi \le 8$ and $2 \le Bi \le 8$ are shown in Figure 2 (c and d).

But the results obtained for higher values of "j" ($j \ge 3$) are much distorted and the statistics of capacity, SNR and PSNR are listed in Table.1 and Stego image are not shown in the paper. For higher value of "j" the hiding capacity increases but the key size decrease.

(a)

(b)

(c)

(d)

(e)

Figure 2: (a) cover image (b) message (c) Stego Image for (1≤Bi≤8) (d) Stego Image for (2≤Bi≤8) (e) Stego Image of 4LSB Stegnography, of VLSB Stegnography

4 Least Significant Bits (4LSB) Stegnography is applied on the same cover image and message image. The Stego image obtained is shown Figure 2 (e). The hiding capacity, Signal to Noise Ratio and Peak Signal to Noise Ratio are listed in Table.1

Table 1: Hiding Capacity, SNR and PSNR

Sr. No	Bi	Hiding Capacity	SNR	PSNR
1	1 – 8	56.0500	6.2580	-17.8074
2	2 – 8	62.3500	5.8594	-18.2060
3	3 – 8	70.7750	4.8644	-19.2010
4	4 – 8	75.0001	4.2527	-19.8127
5	5 – 8	81.2501	3.3117	-20.7537
6	6 – 8	87.4750	2.3561	-21.7093
7	7 – 8	93.7501	2.9388	-21.1266
8	4LSB	50.0000	11.8667	-12.1987

It is apparent from Table.1 that as the range of "Bi" is decreased and shifted towards most significant bits the hiding capacity increases and consequently the SNR decreases and PSNR decreases. On comparison with 4LSB Stegnography the hiding capacity and SNR of VLSB Stegnography remains higher than 4LSB Stegnography, while the PSNR is smaller than that of 4LSB technique. The VLSB technique implemented by using "Bi" in the range from 1 to 8 is having capacity of 56% and is most secure due to its largest key size among all possible implementation. To decrease distortion in stego images, the capacity is decreased; the range of number of bits "Bi" substituted is shifted towards least significant bits i.e. the minimum value of "Bi" is kept fixed at 1 but the upper limit is changed from 1 to 8 (1≤Bi≤j and 1≤j≤8). The stego image obtained for j=8, 7, 6, 5, 4 and 3 are shown in Figure 3 (a, b, c, d and e) respectively. The stego image for j=2 and 1 are not shown. They are much fine and almost same as the cover image.

(a)

(b)

(c)

(d)

(e)

Figure 3: (a) Stego Image for (1≤Bi≤7) (b) Stego Image for (1≤Bi≤6) (c) Stego Image for (1≤Bi≤5) (d) Stego Image for (1≤Bi≤4) (e) Stego Image for (1≤Bi≤3), of VLSB Stegnography

Table 2: Hiding Capacity, SNR and PSNR

Sr. No	Bi	Hiding Capacity	SNR	PSNR
1	1 – 8	56.0500	6.2580	-17.8074
2	1 – 7	49.8501	7.1191	-16.9463
3	1 – 6	43.6501	8.8419	-15.2235
4	1 – 5	37.5001	11.4318	-12.6336
5	1 – 4	31.2501	16.8266	-7.2388
6	1 – 3	24.9751	20.5118	-3.5536
7	1 – 2	18.7500	25.4019	1.3365
8	1 – 1	12.5000	30.5628	6.4974

Comparing the statistics given in Table.2 with the statistics of 4LSB Stegnography it is quite apparent that the on keeping "Bi" to this range the capacity and SNR gradually increases and remain smaller than that of 4LSB technique for "j" smaller than or equal to 7. The stego images are very close to the original cover image and the key size is also significantly large which make these combinations of "Bi" very suitable for VLSB Stegnography due to innocence of stego image and strong key size. VLSB Stegnography using all possible values

of “Bi” is much efficient, having largest possible key size and good enough hiding capacity although a bit distortion is added to the stego image but that is overwhelmed by the increase in capacity and key size.

V. CONCLUSION

The Stegnographic technique presented is much secure data hiding technique with variable data hiding capacity. The hiding capacity and SNR are inversely related. A trade off is made to get the desire parameters by implementing VLSB Stegnography. If the Capacity is increased the distortion will increases and vice versa. Key size is proportional to number possible values of bits substituted “Bi”. Squeezing the range decreases the key size. VLSB methodology is providing us with full liberty to get desired values of these parameters according our application and needs.

REFERENCES

[1] S. Dumitrescu, W.X.Wu and N. Memon (2002) On steganalysis of random LSB embedding in continuous-tone images, Proc. International Conference on Image Processing, Rochester, NY, pp. 641-644.

[2] S.K. Moon and R.S. Kawitkar,” Data Security using Data Hiding”, International Conference on Computational Intelligence and Multimedia Applications 2007, pp.247-251.

[3] S.K. Moon and R.S. Kawitkar,” Data Security Using Data Hiding” International Conference on Computational Intelligence and Multimedia Applications 2007.

[4] Neil F. Johnson and Sushi1 Jajodia,”Steganalysis: The Investigation of Hidden Information”,

[5] Kafa Rabah. Steganography - The Art of Hiding Data. Information technology Journal 3 (3) – 2004

[6] T. Cedric, R. Adi and I. Mcloughlin (2000), Data concealment in audio using a nonlinear frequency distribution of PRBS coded data and frequency-domain LSB insertion, Proc. IEEE International Conference on Electrical and Electronic Technology, Kuala Lumpur, Malaysia,pp. 275-278.

[7] D. Kahn, the Codebreakers, Macmillan, New York, 1967.

[8] Beenish Mehboob and Rashid Aziz Faruqui,” A Stegnography Implementation”, IEEE, 2008.

[9] M.D. Swanson, M. Kobayashi, A.H. Tewfii, "Multimedia Data Embedding and Watermarking Technologies", *Proc. of the IEEE,* vol. 86, no. 6, June, 1998, pp. 1064-1087.

[10] http://isse.gmu.edu/~njohnson/steganography

[11] Takeshi OGIHARA, Daisuke NAKAMURA and Naokazu YOKOYA, “Data Embedding into Pictorial Images with Less Distortion Using Discrete Cosine Transform”,

[12] J.Fridrich, M.Goljan, and R.Du,”Detecting,” LSB Stegnography in Color and Gray –Scale Images”, Magazine of IEEE Multimedia, Special Issue on Security, October-November issue, 2001, pp.22-28.

[13] S. K. Moon, V. N. Vasnik, “Application of steganography on image file”, National conference on Recent trends in Electronics, pp. 179-185.

[14] Alkhraisat Habes, “Information transmissions in computer network. Information hiding in bmp image Implementation analysis and evaluation” (Jan.2006)

[15] T. Morkel, J. H. P. Eloff, M. S. Olivier, ”An Overview of Image Steganography”, Information and Computer Security Architecture (ICSA) Research Group, Department of Computer Science, University of Pretoria, SA.

[16] http://www.mathworks.com

[17] http://www.math.ucla.edu/

Engr. Sahib Khan is pursuing M.Sc in Telecommunication Engineering at department of Telecommunication Engineering, Faculty of Telecommunication and Information Engineering, University of Engineering and Technology Taxila. He has B.Sc Telecommunication Engineering from N-W.F.P University of Engineering and Technology Peshawar, Pakistan. He is serving as a Lecturer and Course Coordinator at department of Electrical and Computer Engineering, Kohat University of Science and Technology Kohat, Pakistan. His areas of interest are Digital Image Processing and Data Authentication and Information Security.

Muhammad Haroon Yousaf is pursuing Ph. D. in Computer Engineering at University of Engineering & Technology Taxila, Pakistan. He got M.Sc. and B.Sc. in Computer Engineering in the years 2007 and 2005 respectively. He currently holds the position of Assistant Professor in the Faculty of Telecommunication & Information Engineering. His research interests include gesture and activity recognition, human computer interaction, automotive user-interfaces, computer vision and machine learning. He, along with his students, investigating face recognition based solution for automated attendance management system.