# POSITION IDENTIFICATION USING DIHEDRAL FUNCTION FOR DATA HIDING

**Rajeev K, M. Sethumadhavan, Amritha P. P**

***Abstract* : *Data hiding techniques can be used for protecting data from unauthorized users. Here, data hiding technique called steganography is used to prevent illegal access of data. We introduced a new method of hiding data in images by randomly selecting pixel position using dihedral function. We have used different window size to hide data and applied variable dihedral function based on the size of variable. The security of this approach lies in the knowledge of window size and the function. The performance of our method was analyzed using Chi-square, RS steganalysis attack and histogram attack. The quality of the image was also preserved and found to be 52 dB.***

***Index Terms*: *Steganography, Dihedral function, Steganalysis, PSNR***

## I. INTRODUCTION

Highlight a section that you want to designate with a certain style, Steganography is transporting a message without its existence being detectable by anyone monitoring the communication. Digital steganography is the method of securely embedding secret information in media like text, image, video and audio. Four requisites have to be guarantee while designing a steganography algorithm. They are undetectabilty, robustness, perceptual transparency and security. There is always a tradeoff between these four requisites. All four requirements will not be satisfied at a same time. For example, if we want to hide a large capacity, we cannot obtain more robustness and absolute undetectability at the same time. On the other hand, the message length cannot be too long if robustness to huge distortion is an issue. Suppose a steganography algorithm can achieve these requirements, the final image will likely reach its receiver unnoticed and can say that successful execution of data hiding has been developed.

Steganalysis is a counterpart for steganography, which finds the presence of hidden messages in doubtful media and stop further usage. Steganalysis can be blind or targeted [1, 2]. The targeted steganalysis identifies the secret and then neutralizes it. This necessitates the knowledge of the algorithm used for stego generation and hence its use is limited to specific scenarios.

The blind steganalysis attempts to display the media as stego or not without assuming any knowledge of the algorithm. Here we present an approach to hide data in spatial domain using least significant bit embedding but in random positions of the image. These positions are selected according to the output obtained from the Dihedral function. In this work 4-variable, 6-variable and 8-variable function are applied on different window size in order to select the positions in the image. Embedding rate varies depending on window size and the function chosen.

Depending on the security of application we can select the window size and apply the function. The performance of this method was evaluated using the measure peak signal-to-noise ratio (PSNR) and quantitative methods by using Chi-square analysis and RS steganalysis. We have also compared our method with random least bit significant embedding.

## II. BACKGROUND AND RELATED LITERATURE

### A. Steganographic techniques

A brief survey of the least significant bit steganographic algorithm which has been found useful in our research is given here. The common method in hiding is in the pixel (spatial) domain and in transform domain. The basic technique in pixel domain is least significant bits (LSB) embedding [3] which uses LSB of certain gray levels in the image to embed message. This technique cannot resist visual and histogram attacks and hence random least significant bit embedding was introduced which could resist visual attack. Random LSB is same as LSB embedding but embedding in random locations in image by using a pseudo random generator. This method was attacked using Chi-square attack and RS steganalysis [4]. But our method is resistant to histogram attack, Chi-square attack and RS steganalysis. Authors in [5] have introduced an optimal LSB substitution by hiding data in the rightmost *k* bits of the cover image, to improve the stego image quality over simple LSB method. Even though this method resists visual attack, it could only partially resist histogram attack. LSB matching is a trivial variation of LSB replacement. Instead of substituting the LSB with the desired message bit, the corresponding pixel value is decremented or incremented randomly whenever the least bit needs to be changed.

This method resists against histogram attack [6]. Many other steganographic techniques which are variants of LSB methods are reported in [7, 8].

### B. Steganalysis

Steganalysis is the set of tools and approaches used to find the presence of secret messages. An attacker/adversary is also referred as a warden. Warden

**Revised Manuscript Received on July 05, 2019.**

**Rajeev K**, TIFAC-CORE in Cyber Security, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India, k_rajeev@cb.amrita.edu

**M. Sethmadhavan**, TIFAC-CORE in Cyber Security, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India.

**Amritha P. P**, TIFAC-CORE in Cyber Security, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India,

who only inspects the communication and wishes to know if the communication contains some hidden messages is the passive warden. The warden does not alter the content of the communication. A warden who may introduce alteration to interrupt and destroy the message is the active warden. Active warden cannot identify a stego image if the secret is less than one percent or depending on complexity. On the other hand, normally steganalysis method can detect stego content if it is more than a certain percentage (percentage varies depending on the stego method used). Targeted steganalysis aim at the detection of a message associated with particular steganographic methods, while blind steganalysis aim at the message detection regardless of the embedding algorithm. In these analyses payload content of more than 10 percentages can be detected but less than 10 percent can be detected with some error rate. Active steganalysis methods using filtering approaches are there that can destroy even less than 10 percent of stego content. Types of steganalysis used to measure the system performance:

*Chi-square attack*: Westfeld and Pfitzmann introduced this attack in [9]. This attack was proposed to detect sequential least significant bit embedding, based on the image histogram. This attack concentrates on the histogram count before and after embedding. The observed value and expected value of a histogram count is calculated once we get an image. The Chi-square value calculated for the cover image this value is expected to be more than the stego image. Chi-square value obtained is further used to estimate the percentage of secret embedded using cumulative distribution.

*RS steganalysis:* This attack is used to figure out the extent of the hidden message on an image, for steganographic methods like flipping least significant bits. This attack exploits the image correlation in the spatial domain [4] and each image bit plane is correlated with the remaining bit planes. In this scheme the number of regular (R) and Singular (S) groups of pixels' changes with the increase in payload embedded in LSB plane. To examine an image, groups of pixels are defined depending upon some properties. The percentage of embedding can be predicted with the help of relative frequencies of these groups in the given image from the data got from the original image with LSBs flipped and an image got by randomizing least significant bits of the original image. We are using these two statistical steganalysis for measuring the performance of our method.

### C. Dihedral function

We are considering Boolean functions, a mapping from $\{0, 1\}^n$ to $\{0,1\}$ which plays a vital role in designing cryptographic primitives such as random number generators, hash functions and symmetric ciphers. In this paper we are using a special class of functions naming Dihedral Symmetric Boolean Functions (DSBF) for identifying the embedding positions in the image. These functions are relatively a large class of Boolean functions which is having special properties that are very useful for our current scenario.

DSBFs are invariant under the action of dihedral group [10, 13] and these functions have never been used for data hiding like steganography to our knowledge. The number of DSBFs up to 10 variables is given in Table 1. Any random function from these dihedral classes of DSBFs is a good candidate for identifying the positions in the image. An example for a four variable function could be $f(x_1, x_2, x_3, x_4) = x_2x_4 + x_2x_3 + x_4x_1 + x_3x_4 + x_1x_3 + x_1x_2$. From Table 1: it is clear that for $n$=4 we can choose any function from $2^6$ functions which can be used for identifying the positions for embedding.

Table 1: The number of DSBFs

| *n*-variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| # | $2^2$ | $2^3$ | $2^4$ | $2^6$ | $2^8$ | $2^{13}$ | $2^{18}$ | $2^{30}$ | $2^{46}$ | $2^{78}$ |

## III. PROPOSED SYSTEM

If you are using The primary goal of our paper is to securely transmit data from unauthorized user. Any steganographic method is breakable if given the algorithm and sufficient amount of time. But the purpose of steganography is to pass the message to the end user within the limited time frame. So within this short span of time it is very difficult for an adversary to break the system even though he knows the algorithm. We have designed an algorithm which is two level secure. Even though we make our algorithm public it is very difficult to break our system without knowing function and window size.

We have proposed this method in the pixel domain but can be applied to transform domain also. Figure 1 and 2 shows the block diagram for embedding and extraction of proposed system. The user can select input as any cover media like image or video to hide the secret which can be a text or an image. Given these input to the embedding algorithm user will be asked to select a dihedral function of user choice based on the window size. The proposed algorithm comprises of 4-variable, 6-variable and 8-variable DSBFs. Then we perform least significant bit embedding on the pixel based on positions satisfied by the dihedral function. After all message bits have been embedded a stego image is obtained. This will be send to the receiver by securely sending window size and the function used. The receiver after obtaining the stego image can extract secret by giving the function, window size and stego image to the extraction algorithm (see Figure 2). The steps are summarized below

*Embedding:*

Input: image or video file and secret message

Output stego image

1. Read the input image and secret
2. Select the window size (2x2, 3x3, 4x4, 5x5, 6x6, 7x7 or 8x8) based on the input cover image
3. Select the *n* variable dihedral function (*n*=4, 6, 8)
4. The positions in the window are converted to binary value, for e.g. position (3, 2) in the image is converted to binary as 1110 (*x* coordinate 3 as $(11)_2$ and *y* coordinate 2 as $(10)_2$)
5. This binary value is given as input to the dihedral function
6. If the function gives output 1 we will select that position to embed else ignore.
7. Embedding is done as follows: replace the last bit of the

pixel in selected position by message bit.
8. Thus applying the same procedure all over the image by sliding the window in non-overlapping fashion.
9. Obtained output image after hiding secret is the stego image.

*Extraction:*

Input: Stego image, function and window size
Output: Secret message

1. To extract the secret we need to know the function and window size. Since we know we have embedded in least significant bit by knowing the other two factors we can get the positions and extract the LSB bits from those pixels.
2. All binary inputs sequence is combined to form the secret message.

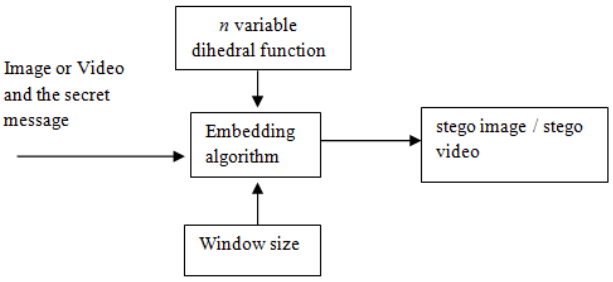


Fig. 1. Block diagram for embedding

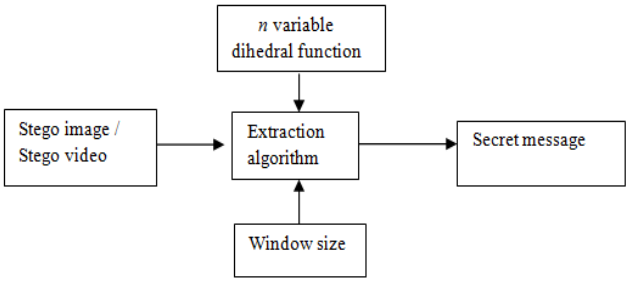


Fig. 2. Block diagram for embedding

### C. Security of proposed system

Security of this system lies in dihedral function chosen from the space $2^{2^n}$ and the window size. The following scenario explains the security based on the computational complexity to brute force the proposed system.

*Case 1: Adversary knows window size*

Computational complexity depends on the probability of finding the number of variable used and to select the function from the space $2^{2^n}$. From Table 1: it is evident that the probability of getting a random function from the class of DSBFs are 0.015, 1.2 x$10^{-4}$, 9.3 x$10^{-10}$ for 4, 6 and 8 variables respectively.

*Case 2: Adversary knows the variable n*

Finding the window size for the known variable n is difficult and identifying the set of functions within the space $2^{2^n}$ increases the complexity to brute force.

## IV. EXPERIMENTAL RESULTS

To conduct experiments, we have used 256 x 256 of image size taken from the database BOSSbase ver.1.01 [11] and also standard data set like Lena and mandrill [12]. All the cover images were subjected to our method and random LSB embedding. We have taken the secret message of varying length 100, 70, 50 and 30 % of image size for analysis. For different window size we have used 4, 6 and 8 variable DSBFs functions and found different embedding location and results are tabulated in Table 2. We can see that by using any of the *n*-variable function and fixed window size the quality is above acceptable level and found to be on average 52dB. We also compared our result with existing random LSB embedding and found that our method could resist Chi-square attack for all embedding rate other than 100%. RS steganalysis failed for detecting payload with 70 % and 30%. When 70% of secret was embedded using our method RS steganalysis could detect only 46 % of it. When 37 % of payload was embedded RS steganalysis was showing as 66 % which is false positive (Table 3). We can see from Table 3 that random LSB failed to resist Chi-square and RS steganalysis for all payload size. Hence by using 8-variable and 6-variable DSBFs function and corresponding window size 6 x 6 and 4 x 4, it is difficult to attack.

We have also seen that both methods could not resist histogram attack completely. The artifact seen in Figure 3 and 4 is because we have used Least Significant Bit embedding. By experiments we can have concluded that if we are using any variant of LSB embedding with our dihedral function and window size, we will be able to resist histogram attack to large extend.

### A. Histogram attack

We have plotted histogram of stego and cover images created by using our method and random LSB method. We were able to see that in both cases histogram artifact was seen. But compared to random LSB our method has less artifact in histogram. Figure 3 and 4 is plotted by calculating the frequency count of gray level values in the cover and stego image created using our method and random LSB embedding with payload 100%. Figure 5 show the image before and after embedding by our method. Left most one is the original image and other two is the stegoed image created by using window sizes of 2x2 and 4x4 with 4-variable DSBF.

Table 2: Performance of our method in terms of image quality metric

| *n*- variable | Window size | Payload (embedding rate in percentage) | PSNR (dB) |
|---|---|---|---|
| 4 | 2x2 | 100 | 51.13 |
| | 3x3 | 88 | 51.6619 |
| 6 | 2x2 | 100 | 51.1532 |
| | 3x3 | 45 | 54.6656 |
| | 4X4 | 37.5 | 55.3768 |
| | 5X5 | 48 | 54.3057 |
| | 6X6 | 47.7 | 54.6656 |
| | 7X7 | 36.7 | 56.6656 |
| 8 | 2X2 | 100 | 51.1532 |
| | 3X3 | 90.2 | 51.6619 |
| | 4X4 | 93.7 | 51.4126 |
| | 5X5 | 84 | 51.8798 |
| | 6X6 | 71.5 | 52.0852 |
| | 7X7 | 57.5 | 53.5541 |
| | 8X8 | 64.1 | 53.0852 |

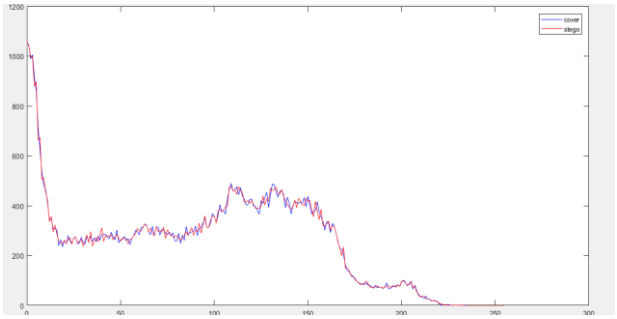

Fig. 3: Histogram of cover and stego for our method

Table 3: Comparison of our method with random LSB method

| Stego methods | Payload (%) | Detection by | | |
|---|---|---|---|---|
| | | RS (%) | Chi-square attack | Histogram attack |
| Random LSB | 100 | 99.1 | Yes | Yes |
| | 70 | 66.7 | Yes | Yes |
| | 50 | 47.6 | Yes | Yes |
| | 30 | 32.9 | Yes | Yes |
| Our method | 100 8 variable (2x2) | 100 | Yes | Yes |
| | 71.5 8 variable (6x6) | 46.2 | No | Yes |
| | 57.5 8 variable (7x7) | 57.6 | No | Yes |
| | 37.5 6 variable (4x4) | 66.3 | No | Yes |

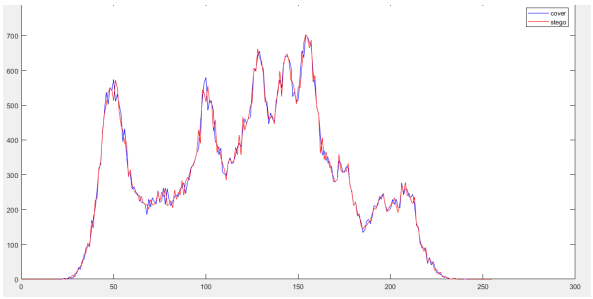

Fig. 4. Histogram of cover and stego for random LSB method

Fig. 5. Left most one is the cover image and other two is the stego image

## V. CONCLUSION

We have introduced new method for protecting message by using steganography. Security of our method lies in the window size and the DSBFs used for identifying positions for embedding. We have also seen that our method could preserve the quality even after embedding 100% and PSNR were found to be 52dB at all embedding rates. This method was able to resist Chi-square attack at all embedding rate other than 100% and could withstand RS steganalysis at 70 and 30% embedding rate.

Hence if 8-variable and 6-variable DSBFs function

and corresponding window size 6 x 6 and 4 x 4 are used, it is more difficult to attack. We concluded that compared to any variant of LSB embedding algorithm our method is computational difficult to brute force given some resource and limited time. Hence any random function from these dihedral classes of DSBFs is a good candidate for identifying the positions in the image to hide data. This method can also be applied to embed data in transform domain.

## REFERENCES

1. R. Chandramouli, M. Kharrazi, N. Memon, “Image steganography and steganalysis: Concepts and practice,” Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 2939, 2003, pp. 35-49,
2. X.-Y. Luo, D.-S. Wang, P. Wang, F.-L. Liu, “A review on blind detection for image steganography,” Signal Processing, vol. 88, 2008, pp. 2138-2157.
3. S. Walton, “Image authentication for a slippery new age,” Dr. Dobb’s Journal, vol. 20, 1995, pp. 18–26.
4. J. Fridrich and M. Goljan, “Practical steganalysis of digital images: state of the art,” Security and Watermarking of Multimedia Contents IV, International Society for Optics and Photonics, vol.4675, 2002, pp. 1–14.
5. Wang, Ran-Zan, Chi-Fang Lin, and Ja-Chen Lin. "Image hiding by optimal LSB substitution and genetic algorithm," Pattern recognition , vol. 34, 2001, pp. 671-683.
6. G. Cancelli, G. Doërr, I. J. Cox, and M. Barni, “Detection of ±1 lsb steganography based on the amplitude of his togram local extrema,” Image Processing, IEEE, 2008, pp. 1288–1291.
7. P. Amritha and T. K. Gireesh, “A survey on digital image steganographic methods,” Cyber Security, Cyber Crime and Cyber Forensics: Applications and Perspectives, IGI Global, 2011, pp. 250–258.
8. Anjana, S., and P. P. Amritha. "A novel method for secure image steganography." Artificial Intelligence and Evolutionary Algorithms in Engineering Systems, Springer, 2015, pp. 151-158.
9. A. Westfeld and A. Pfitzmann, “Attacks on steganographic systems,” in International workshop on information hiding, Springer, 1999, pp. 61–76.
10. Rajeev, K., M. Sethumadhavan, and K. V. Lakshmy. "Results on Dihedral Symmetric Boolean Functions," International Journal of Pure and Applied Mathematics , vol. 119, 2018, pp. 3233-3238.
11. T. Filler, T. Pevny, P. Bas, “BOSS (Break our steganography system)”, Available: http://agents.fel.cvut.cz/boss/, 2010.
12. Weber, g. The USc-SIPI Image Database Version 5, USC-SIPI Report #315, 1997.
13. Cusick, Thomas W., K. V. Lakshmy, and Madathil Sethumadhavan. "Affine equivalence of monomial rotation symmetric Boolean functions: A Polya’s theorem approach," Journal of Mathematical Cryptology, vol. 10, 2016, pp. 145-156.

## AUTHORS PROFILE

**Rajeev K** received his MSc from University of Calicut. He currently serves as Research Associate at TIFAC CORE in Cyber Security, Amrita Vishwa Vidyapeetham, Coimbatore. His research interests include: Pairing based cryptography and cryptographic Boolean functions.

**Dr. M. Sethumadhavan** received his PhD (Number Theory) from Calicut Regional Engineering College. Currently, he is working as a Professor in the Centre for Cyber Security, Amrita Vishwa Vidyapeetham, Coimbatore. His current research interests include: Cryptography and Boolean functions.

**Amritha P. P** received her M. Tech (Cyber Security) from Amrita Vishwa Vidyapeetham, currently pursuing her PhD at Amrita Vishwa Vidyapeetham. Her current research interests include: Steganography and code obfuscation.