International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 2 Issue 7‖ July. *2013* ‖ *PP.*36-42

Steganography Method Based On Data Embedding By Sudoku Solution Matrix

[1,]Suman Chakraborty, [2,]Prof. Samir K. Bandyopadhyay
[1,]B.P. Poddar Institute of Management and Technology, 137 V.I.P. Road,
Kolkata – 700052, India.
[2,]Department of Computer Science and Engineer, University of Calcutta,
92 A.P.C. Road, Kolkata-700009, India

ABSTRACT : In this paper propose a novel staganography method based on data embedding. Embedding is done by Sudoku solution matrix. In this scheme RGB value of Cover-image contains both Secret information and key (Sudoku) .RG component contains Secret information and B component contains key value. Same Sudoku solution matrix is used for embedding and extraction phase. To increase the capacity of embedding, Secret information compressed by DNA sequence compression. Four DNA sequence compression technique has been used to achieve the desire rate of compression.

KEY WORDS: DNA sequence, Compression, Cover-image, Base-9 form, Embedding, Sudoku matrix, Secret information, Extraction, RGB.

I. INTRODUCTION

Using digital images as cover media to conceal secret data is an important issue for secret data delivery applications. From the target of image modification, information hiding techniques can be classified into three domains, namely spatial domain [1], compressed domain [2], and transformed domain [3]. In spatial domain, more redundant spaces are available to secret data embedding so high embedding capacity can be achieved, and less time is needed for embedding and extracting procedures. However, information hiding schemes in spatial domain are vulnerable to common attacks such as statistical stego analysis. So security in steganography can be achieved using different embedding schemes [4-5]. The important factors needed to consider when we are designing a new information hiding scheme are embedding capacity (i.e. the number of secret bits can be embedded into one cover image pixel), visual quality of stego images [6] (i.e. image distortion); amount of data sent (i.e. compression) and secure exchange of data (i.e. Encryption). Desirably, one would want to achieve high embedding capacity, good visual quality, and more data to get embedded and high security. However, embedding capacity and visual quality are inversely proportional to each other. That is, if embedding capacity is increased, then visual quality is decreased and vice versa. Thus, a tradeoff between embedding capacity and visual quality is made by users for different applications. The security and embedding more data can be done using encryption and compression.

DNA sequence, the information in DNA possess some interesting properties which can be utilized to hide data as a code prepared up of four chemical bases (A,C,G,T) : adenine (A), guanine (G), cytosine (C), and thymine (T). Conventionally, data hiding approaches frequently implant a secret message into the congregation images. However, this could deform the congregation image to some degree, and may therefore; medicinal effectiveness and susceptibility for solving complex, highly comparable computational problems have also been demonstrated. The capability to hide, gloss information, and watermarks within this intermediate is clearly meaningful. The order, or sequence, of these bases determines the information accessible for structuring and preserving an organism, comparable to the technique in which correspondence of the alphabet come into vision in a certain categorize to form sentence and words. DNA bases join up with both together, A with T and C with G, to configure units called base pairs that can promote greatly from a data hiding scheme, and a bit of surroundings in retroviral DNA sequences is required to understand this method. Each base is also closed to a sugar molecule and a phosphate molecule [7].

A massage representing a DNA sequence, with the combination of a, c, t, g is equivalent to DNA sequence of the original data. Compressing the DNA sequence by-1) 2-bits encoding method, 2) Exact matching method, 3) Approximate matching method, 4)For the approximate matching method. These compression techniques would be produce equivalent digital form of the DNA sequence and one of procedure will produce minimum number of bits [8-9].

Sudoku, English pronunciation soo-DOH-koo is a logic-based, combinatorial number-placement puzzle. The proposed method derives the secret data and generates the meaningful image steganography using DNA sequence and sudoku .The objective of this game is to filled a 9×9 grid with the digits so that each of column, each of row, and all of the nine 3×3 sub-grids that compose of the grid contains all of the digits from 1 to 9. A partially completed grid produced by the puzzle setter which typically has a unique solution. For example, the same single integer may not appear twice in the same 9×9 playing board row or column or in any of the nine 3×3 sub regions of the 9×9 playing board [10]. Figure 1 shows a puzzle and Figure 2 indicates its solution.

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Fig-1: A Puzzle [10]

5	3	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9

Fig-2: Solve Puzzle [10]

In this paper proposed a staganography process which changed minimum in RGB value of pixels of Cover-image after embedding the secret data. In Embedding procedure Sudoku matrix act as Key. Specialty of this method is secret information and key both are embedded into Cover-image .So, key(Sudoku matrix) isn't required to send separately. Another important aspect is secret information embedding after DNA compression which increases security and less no. of bits required to embed. The remainder of the paper is organized as follows. In Section 2 described related work, In Section 3, the description of the proposed steganography in flow diagram is presented. In Section 4, the description of the proposed embedding procedure is presented. In Section 5, , the description of the proposed steganography algorithm.In Section 6,conclution.

II. RELATED WORK

2.1data Embedding

Steganography using Sudoku is used to hide data or secret information onto an image using Sudoku solution. The image used in our proposed method is a 24 bit colored image. Initially the data to be hidden is chosen which can be any digital media file such as text, image, audio, video etc. Sudoku solution is taken and every value in it is subtracted by '1'. This is done so as to ensure all values lie between 0 and 8 in Sudoku so as to maintain compatibility between Sudoku and secret data which is a three bit value from the input data. 9x9 Sudoku is used as reference matrix M for both data embedding and extraction. Before embedding, one or more media files are compressed and encrypted to increase the efficiency and security of the method. DES technique is used for encrypting.

Any image onto which secret data has to be embedded is chosen. Two pixels of this image are chosen and RGB values of both the pixels are paired as C1(R1,G1), C2(B1, R2), C3(G2, B2) we can generalize each pair as Ci(x, y). For each pair Pi.x =Ci.x % 9, pi.y = Ci.y % 9 Then Pi.x and Pi.y are chosen as X-axis and Y-axis components of reference matrix M. Then three candidate elements Horizontal (CEH), Vertical (CEV) and Boxed (CEB) are chosen. Here CEH is shown by green line, CEV sis shown by purple line and CEB is shown by black square in Figure 5. CEH and CEB are chosen so that M (Pi.x, Pi.y) from M is put in middle position of the candidate element array. The remaining positions are filled by respective left and right elements from position of M (Pi.x, Pi.y) in reference matrix. The difference in index positions of secret digit and M (Pi.x, Pi.y) is always less than or equal to 4 reducing the distortion in cover image.

Initialization of CE_H:

```
For (i: 0 to 8)
    pos = (i+4) %9
    CE_H [pos] = M (Pi.x, Pi.y)
    Pi.x = (Pi.x+1) %9
End For
```

Initialization of CE_V:

For (i: 0 to 8)
pos = (i+4) %9
CE_V [pos] = M (Pi.x, Pi.y)
Pi.y = (Pi.y+1) %9
End for

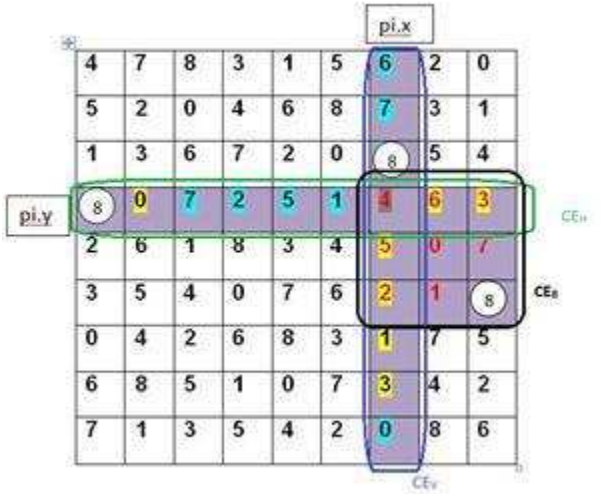


Figure 5: Illustration of CEH (green line), CEV (purple line), CEB (black square)

Pi.x = Ci.x%9
Pi.y = Ci.y%9
Initialization of CE_B:
For (i: 0 to 2)
For (j: 0 to 2)
posx = |_ Pi.x/3_|*3
posy = |_ Pi.y/3_|*3
CE_B [posx] [posy] = M (Pi.x, Pi.y)
posx++
End for
posy++
End for

Selecting optimum modified pixel value:
Positional Difference between Si and M (Pi.x, Pi.y) in CE_H
Find the position of Si in CE_H say it as pos.
DH=pos-4
Positional Difference between Si and M (Pi.x, Pi.y) in CE_V
Find the position of Si in CE_V say it as pos.
DV=pos-4
Find the position of Si in CE_B say it as posx and posy.
SQX=posx-(Pi.x%3)
SQY=posy-(Pi.y%3)
SQD=|SQX|+|SQY|
The minimum distance is calculated by,
Min=minimum (|DH|, |DV|, SQD)
If min =|SQD| Ci.x = Ci.x + SQX
Ci.y = Ci.y + SQY
Else If min=|DH| Ci.x= Ci.x+ DH
Else min=|DV| Ci.y =Ci.y + DV

If Ci.x<0 Ci.x=9+Ci.x or Ci.x>255 Ci.x=Ci.x-9.
If Ci.y<0 Ci.y=9+Ci.y or Ci.y>255 Ci.y=Ci.x-9.

As in above Fig 5 Pi.x=6, Pi.y=3, so M (Pi.x, Pi.y) = 4 and Si=8.
The candidates elements are selected as CE_H= {7, 2, 5, 1, 4, 6, 3, 8, 0},

CE_V= {0, 6, 7, 8, 4, 5, 2, 1, 3}, CE_B={{4,6,3},{5,0,7},{2,1,8}}.
Here DH=7-4=3, DV=3-4= -1, SQX=8-6=2 & SQY=5-3=2.
SQD=SQX+SQY=2+2=4. Min=minimum{|DH|,|DV|,|SQD|}=minimum{3,1,4}=1,
So Ci.x=Ci.x Ci.y=Ci.x+DV.

As a result 9 bits are embedded in two pixels. Similarly apply above method for C2 & C3. The above method ensures the each component of pixel is modified maximum by 4 when its value is greater than 3 and less than 252. Repeat the above procedure until the data gets embedded in cover image, if cover image is not big enough to hold all the data then new cover image should be used until all the data is embedded. Figure 6 shows images before embedding and after embedding. The first 10 pixels of cover image are reserved to embed the size of input data file which is a zip file consisting of variable input media data files. If a cover image can't embed all input data file then remaining data is embedded in new images until all the data is embedded. Sudoku solution is encrypted and then it is embedded using the LSB method in which 3 bits of encrypted Sudoku is embedded so that each bit is at least significant bit of R, G, and B component of cover image pixel. [11]

2.2Data extraction

Two pixels of this image are chosen and RGB values of both the pixels are paired as C1 (R1, G1), C2 (B1, R2), C3 (G2, B2) we can generalize each pair as Ci(x, y).
For each pair

Pi.x =Ci.x % 9, Pi.y = Ci.y % 9

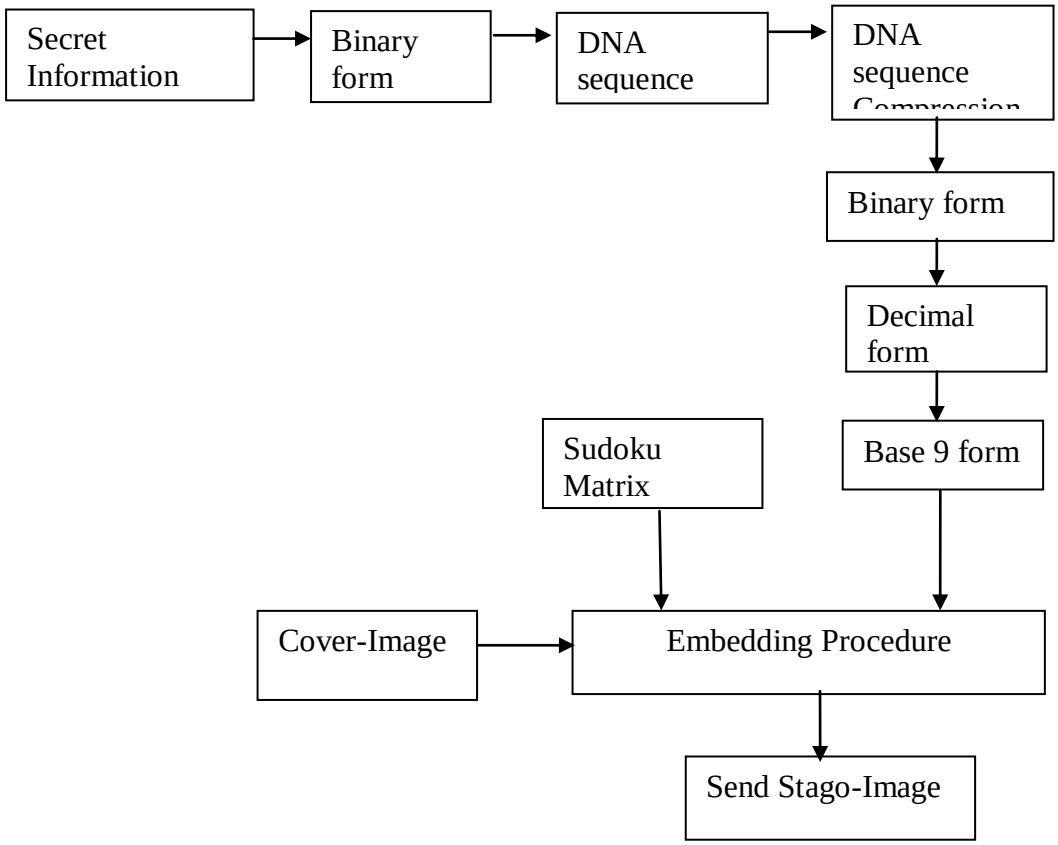
Then Pi.x and Pi.y are chosen as X-axis and Y-axis of reference matrix M. M (Pi.x, Pi.y) is secret data extracted. This process is repeated for C1 and C2 pairs .Then again the whole process is repeated till the required size of secret data is retrieved which is obtained from extracting ten pixels of first cover image.Encrypted Sudoku solution from the received cover image is extracted from LSB of R, G and B components of cover image. The pixels from cover image are used for extraction until size of encrypted Sudoku is obtained which is embedded in cover image. Sudoku solution is retrieved by decrypting Encrypted Sudoku.[11]

Binary	Nucleotide
00	A
01	C
10	G
11	T

Table-1.Binary to Nucleotide mapping

III. FLOW DIAGRAM OF STAGANOGRAPHY PROCESS

In this proposed method Secret-information converted into Base-9 form, after compression.DNA sequence compression technique has been used for compression. Following flow diagram showing the pictorial representation of proposed method.



IV. PROPOSED METHOD

In this Steganography process Sudoku is used as key to hide data or secret information onto an image. RGB color image used as Cover media. Secret information can be any form of digital media, such as text, image, audio, video etc.Sudoku solution matrix converted into Base-9 form by subtracted 1 from all values. 9x9 Sudoku solution matrix (M) is used as key for both data embedding and extraction. Here M is called reference matrix. Before embedding, one or more media files are compressed and encrypted to increase the efficiency and security of the method. DNA sequence compression is used for encrypting.

4.1Secret Information embedding

Any image onto which secret data has to be embedded is chosen. Two pixels of this image are chosen and RG values of both the pixels are paired as C1(R1,G1), C2(R2, G2), C3(R3, G3) we can generalize each pair as Ci(x, y). For each pair Pi.x =Ci.x % 9, pi.y = Ci.y % 9 Then Pi.x and Pi.y are chosen as X-axis and Y-axis components of reference matrix M. Then three candidate elements Horizontal (CEH), Vertical (CEV) and Boxed (CEB) are chosen. Here CEH is shown by green line, CEV sis shown by purple line and CEB is shown by black square in Figure 5. CEH and CEB are chosen so that M (Pi.x, Pi.y) from M is put in middle position of the candidate element array. The remaining positions are filled by respective left and right elements from position of M (Pi.x, Pi.y) in reference matrix. The difference in index positions of secret digit and M (Pi.x, Pi.y) is always less than or equal to 4 reducing the distortion in cover image.

Initialization of CE_H:
For (i: 0 to 8)
pos = (i+4) %9
CE_H [pos] = M (Pi.x, Pi.y)
Pi.x = (Pi.x+1) %9
End For
Initialization of CE_V:
For (i: 0 to 8)
pos = (i+4) %9
CE_V [pos] = M (Pi.x, Pi.y)
Pi.y = (Pi.y+1) %9

End for

Pi.x = Ci.x%9
Pi.y = Ci.y%9
Initialization of CE_B:
For (i: 0 to 2)
For (j: 0 to 2)
posx = |_ Pi.x/3_|*3
posy = |_ Pi.y/3_|*3
CE_B [posx] [posy] = M (Pi.x, Pi.y)
posx++
End for
posy++
End for

Selecting optimum modified pixel value:
Positional Difference between Si and M (Pi.x, Pi.y) in CE_H
Find the position of Si in CE_H say it as pos.
DH=pos-4
Positional Difference between Si and M (Pi.x, Pi.y) in CE_V
Find the position of Si in CE_V say it as pos.
DV=pos-4
Find the position of Si in CE_B say it as posx and posy.
SQX=posx-(Pi.x%3)
SQY=posy-(Pi.y%3)
SQD=|SQX|+|SQY|
The minimum distance is calculated by,
Min=minimum (|DH|, |DV|, SQD)
If min =|SQD| Ci.x = Ci.x + SQX
Ci.y = Ci.y + SQY
Else If min=|DH| Ci.x= Ci.x+ DH
Else min=|DV| Ci.y =Ci.y + DV

If Ci.x<0 Ci.x=9+Ci.x or Ci.x>255 Ci.x=Ci.x-9.
If Ci.y<0 Ci.y=9+Ci.y or Ci.y>255 Ci.y=Ci.x-9.

As in above Fig-5 Pi.x=6, Pi.y=3, so M (Pi.x, Pi.y) = 4 and Si=8.
The candidates elements are selected as CE_H= {7, 2, 5, 1, 4, 6, 3, 8, 0},

CE_V= {0, 6, 7, 8, 4, 5, 2, 1, 3}, CE_B={{4,6,3},{5,0,7},{2,1,8}}.
Here DH=7-4=3, DV=3-4= -1, SQX=8-6=2 & SQY=5-3=2.
SQD=SQX+SQY=2+2=4. Min=minimum {|DH|,|DV|,|SQD|}=minimum{3,1,4}=1,
So Ci.x=Ci.x Ci.y=Ci.x+DV.

As a result 9 bits are embedded in two pixels. Similarly apply above method for C2 & C3. The above method ensures the each component of pixel is modified maximum by 4 when its value is greater than 3 and less than 252. Repeat the above procedure until the data gets embedded in cover image.Sudoku solution is encrypted and then it is embedded using the LSB method in which 2 bits of encrypted Sudoku is embedded so that 2 bits is at least significant bits of B component of cover image pixel.

4.2Data extraction

Two pixels of this image are chosen and RG values of both the pixels are paired as C1 (R1, G1), C2 (R2,G2), C3 (R3,G3) we can generalize each pair as Ci(x, y).
For each pair

Pi.x =Ci.x % 9, Pi.y = Ci.y % 9

Then Pi.x and Pi.y are chosen as X-axis and Y-axis of reference matrix M. M (Pi.x, Pi.y) is secret data extracted. This process is repeated for C1 and C2 pairs .Then again the whole process is repeated till the required size of secret data is retrieved which is obtained from extracting ten pixels of first cover image.

Encrypted Sudoku solution from the received cover image is extracted from LSB of B component of cover image. The pixels from cover image are used for extraction until size of encrypted Sudoku is obtained which is embedded in cover image. Sudoku solution is retrieved by decrypting Encrypted Sudoku.

V. STAGANOGRAPHY ALGORITHM

[1] Change the Secret information into DNA sequence by Table-1.
[2] Compress DNA sequence by one of four compression techniques which will provide minimum no. of bits.
[3] Convert this binary form into decimal form and from decimal form to Base-9 form.
[4] Embedding Secret information and Sudoku matrix.
[5] Send Stago-image.

Output of this algorithm:


Before Embedding


After Embedding

VI. CONCLUSION

Secret information can be any form of digital media, such as text, image, audio, video etc. Sudoku solution matrix converted into Base-9 form by subtracted 1 from all values. Same Sudoku solution matrix is used for embedding and extraction phase. To increase the capacity of embedding, Secret information is compressed by DNA sequence compression. Four DNA sequence compression technique has been used to achieve the desire rate of compression. The output is encouraging.

VII. REFERENCES

[1] C.-C. Chang, T. D. Kieu, and Y.-C. Chou. High capacity data hiding for gray scale images. In Proceedings of the
[2] First International Conference on Ubiquitous Information Management and Communication, pages 139–148. Seoul,
[3] Korea, February 2007.
[4] C.-C. Chang and C.-Y. Lin. Reversible steganography for vq- compressed images using side matching and
[5] relocation. IEEE Transactions on Information Forensics and Security, 1(4):493–501, 2006.
[6] Y.-T. Wu and F. Y. Shih. Digital watermarking based on chaotic map and reference register. Pattern Recognition,
[7] 40(12):3754–3763, December 2007.
[8] Yung-Chen Chou, Chih-Hung Lin, Pao-Ching Li, Yu-Chiang Li A (2, 3) Threshold Secret Sharing Scheme Using
[9] Sudoku 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing
[10] 978-0-7695-4222-5/10, 2010
[11] C.C. Chang, Y.C. Chou and T.D. Kieu, An Information Hiding Scheme Using Sudoku, Proceedings of the Third
[12] International Conference on Innovative Computing, Information and Control (ICICIC2008), June 2008.
[13] Wien Hong, Tung-Shou Chen, Chih-Wei Shiu, Steganography Using Sudoku Revisited Second International
[14] Symposium on Intelligent Information Technology Application 978-0-7695-3497-8/08, 2008
[15] Prof. Samir Kumar Bandyopadhyay and S Chakraborty, Image Hiding in DNA Sequence Using Arithmetic
[16] Encoding, Journal of Global Research in Computer Science, Volume 2, No.4, April 2011.
[17] M. Crochemore, W. Rytter, Jewels of Stringology, World Scientific, 2002.
[18] European Bioinformatics Institute, <http://www.ebi.ac.uk/>.
[19] http://en.wikipedia.org/wiki/Sudoku.
[20] Sanmitra Ijeri, Shivananda Pujeri, Shrikant B, Usha B A. Image Steganography using Sudoku Puzzle for Secured
[21] Data Transmission. International Journal of Computer Applications (0975 – 888) Volume 48– No.17, June 2012.