

Embedding Change Estimation for Universal Steganalysis using 3-way Tensor Model

Arunvinodh C^{1*} and S.Poonkuntran²

^{1*}Research Scholar, Department of Computer Science and Engineering, JIT University, Rajasthan, India

²Professor, Department of Computer Science and Engineering, Velammal College of Engineering and Technology, Madurai, Tamilnadu, India

e-mail: ¹arunvinodh@gmail.com, ²s.poonkuntran@gmail.com

Available online at: www.ijcseonline.org

Received: Oct/26/2016

Revised: Nov/05/2016

Accepted: Nov/24/2016

Published: Nov/30/2016

Abstract— This paper proposes a novel Universal Steganalysis framework which can be applied for spatial domain and JPEG Domain Steganography algorithms. The objective is to develop a steganalysis algorithm which has to identify any distribution (uniform or non-uniform) of stego-payloads. The framework proposed uses a 3-way tensor model to extract the image features which is important for estimating the embedded change irrespective of domain. To obtain the accurate results and to analyze the error, 360 degree bit change estimation is done. The experimental results evaluated on 3000 images which shows a good detection rate in both domains and a reasonable false acceptance rate and false rejection rate based on the pay load when tested with most of the steganography algorithms.

Keywords- Spatial Domain, Jpeg Domain, Steganalysis, Tensor, SVM

I. INTRODUCTION

Steganalysis is the art of discovering hidden data in images/video. The Steganalysis techniques are classified into two categories. 1. Specific Steganalysis 2. Universal Steganalysis. Specific Steganalysis are designed for a targeted embedding technique [13],[5]. Thus they yield very accurate decisions when they are used against the particular steganographic technique. In universal technique, the dependency on the behavior of the individual embedding techniques is removed by determining and collecting a set of distinguishing statistics that are sensitive to a wide variety of embedding operations. It has been identified that most of the research objectives focused on identifying the embedded data instead of extracting the data [1],[2],[3],[5],[6]. Quantitative Steganalysers are important forensic tools because the number of embedding changes correlates with the message length.

Steganography can be done in two major domains. 1. Spatial domain. 2. JPEG Domain. It has been identified that design of steganalysis algorithm is focused mainly on JPEG Domain[2],[4]. The algorithm which has been designed for JPEG domain is partially working for Spatial Domain [2],[3],[14]. Therefore the spotlight of this paper is mainly on Spatial Domain Steganalysis as well as Jpeg Domain which can be a Universal Steganalysis frame work.

Tomas Pevny etal [4] focused their assumptions more on JPEG domain. Since this research proposes to concentrate more on spatial domain steganography algorithm as well as JPEG Domain, it is difficult to predict how well the results will compare to Tomas Pevny results. Jan Kodovsky and

Jessica Fridrich [1] assumed that both training and testing images were generated based on uniformly distributed payloads. This assumption leads to a development of a steganalysis algorithm which has to identify any distribution (uniform or non-uniform) of stego-payloads. The validation is done by the literature survey based on non-uniform distribution of stego-payloads as follows.

T.Pevny etal.[15] highlighted about the challenge of steganalysis researchers for advanced content adaptive steganographic methods. Even though J.fridrich etal. [2] proposes a universal steganography detector which successfully attacked LSB matching revisited algorithm (LSBMR) proposed by LUO etal.[16]. Edge adaptive Image Steganography Based on LSB Matching Revisited (EALSBMR)[17] algorithm was not successfully attacked by [2] because of the adaptive methodology. Shunquan Tan etal. [18] proposes Targeted steganalysis of EALSBMR and it was successful only for the proposed steganography algorithm.

There are many adaptive steganography algorithm where an embedding redundancy in LSB matching to select modification direction and takes the dependency of

neighbouring pixels into consideration. Since the neighboring pixel dependency is considered the universal steganalysis may be a challenging part in my research [18]. A combined spatial domain embedding and transform domain embedding makes difficulty in the attack [19].

Based on the above validation and discussion, This paper proposes an unique frame work based on 3 way tensor model which will accompany adaptive steganalysis as well as steganalysis of uniform payload distribution. Also the frame work satisfies the requirement of steganalysis in the spatial domain as well as the JPEG Domain

This paper is organized as follows. In section II , a Frame work has been proposed using 3-way tensor model and we discussed the details of SCI, Forward cycling and backward cycling of matrices and bit change rate estimation. The Results and Discussion claims the successful working of the framework by analyzing in various test bed created based on the steganography algorithms which is mentioned in the section III. Finally the conclusion is summarized in section V.

II. FRAME WORK

In our system, the frame work proposed in figure 1 clearly shows the importance of tensor representation. The mathematical model of this frame work proposed is adopted from [20], [21] and [22] which give the foundation of tensor representation and manipulation

A. SCI Generation

The first step in SCI Generation is to partition the image into N slices, where N is denoted as the number of bit slices. If the image is composed of $N-1$ bit slices, ranging from slice 0 for least significant bit to slice $N-1$ for the most significant bit. In terms of N bit slices, slice 0 contains all the lowest order bits in the bytes comprising the pixels in the image and slice $N-1$ contains all the high order bits. Therefore by separating the image into bit slices, we immediately have a method of identifying more important and less important information which is suitable for extracting the image features.

The image can be divided into bit slices by the following steps.

- Let I be an image where every pixel value is n -bit long
- Express every pixel in binary using n bits
- Form out of I n binary matrices

where the i -th matrix consists of the i -th bits of the pixels of I .

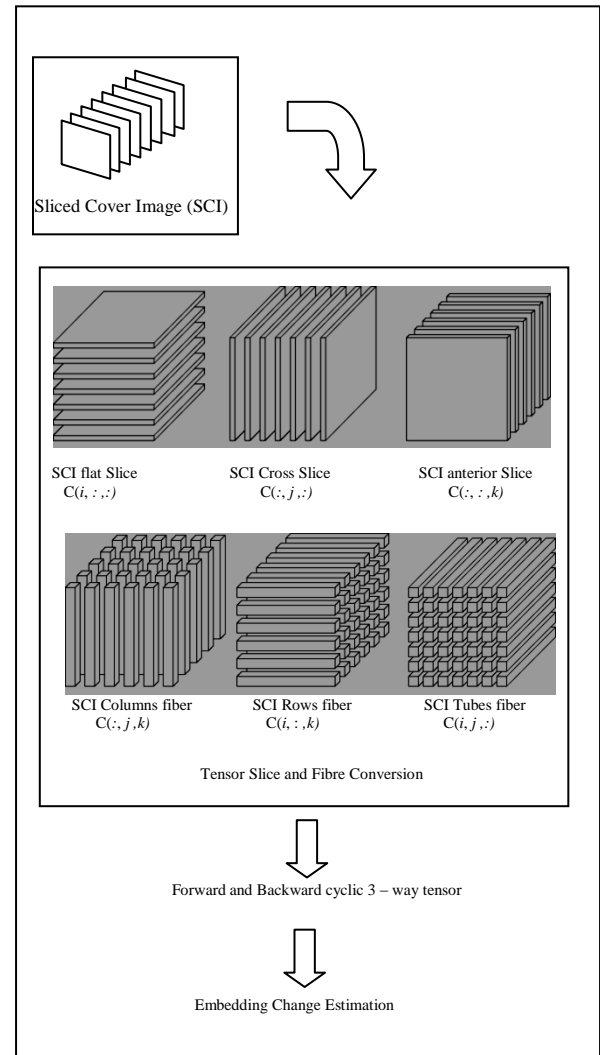


Figure 1: Proposed Frame work

B. Tensor Slice and Fibre Conversion

Let C be a Bit Sliced Tensor of dimension

$$C_1 \times C_2 \times \dots \times C_N.$$

The order of C is N . The n th dimension (or mode or way) of C is of size C_n .

Let $C(i, :, :)$ acquiesce the i th SCI flat slice, $C(:, j, :)$ the j th SCI Cross slice, and $C(:, :, k)$ the k th SCI anterior slice $C(:, j, k)$ yields a column fibers, $C(i, :, k)$ yields a row fibers, and $C(i, j, :)$ yields a so-called tube fibers as shown in Figure 1..

Typically, a tensor is matricized such that all of the fibers associated with a particular *single* dimension are aligned as columns of the resulting matrix. In other words, we align the fibers of dimension n of tensor C to be the columns of the matrix. The resulting matrix is typically denoted by $C_{(n)}$. The columns can be ordered in many ways. As discussed in [21], the ordering can be given as

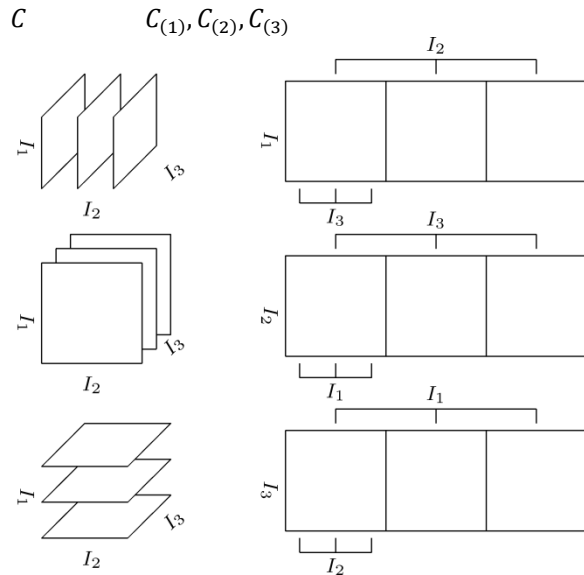


Figure 2: Backward cyclic matricizing a three-way tensor.

$$\{c_1, \dots, c_L\} = \{n - 1, n - 2, \dots, 1, N, N - 1, \dots, n + 1\},$$

and this ordering is named as *backward cyclic*. As per [23], the ordering is specified as follows

$$\{c_1, \dots, c_L\} = \{n + 1, n + 2, \dots, N, 1, 2, \dots, n - 1\},$$

and this ordering is mentioned as *forward cyclic* or “fc” for short. This framework uses both backward and forward cyclic which is helpful for identifying the bit change rate.

Based on the matricizing process an Nth-order tensor is represented as follows

$$C \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$$

The matrix Unfolding is represented as follows.

$$C_{(n)} \in \mathbb{R}^{I_n \times (I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N)}$$

$$C \quad C_{(4)}, C_{(5)}, C_{(6)}$$

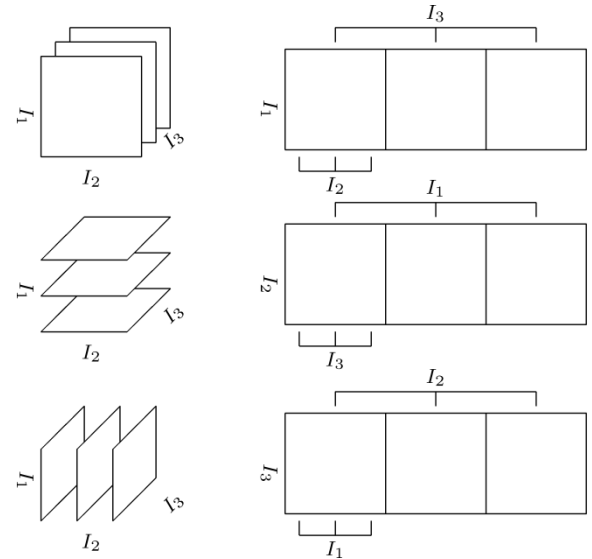


Figure 3: Forward cyclic matricizing a three-way tensor.

Which contains the element $a_{i_1 i_2 \dots i_N}$ at the position with row number i_n and column number equal to

$$(i_{n+1} - 1) I_{n+2} I_{n+3} \dots I_N I_1 I_2 \dots I_{n-1} + (i_{n+2} - 1) I_{n+3} I_{n+4} \dots I_N I_1 I_2 \dots I_{n-1} + (i_N - 1) I_1 I_2 \dots I_{n-1} (i_1 - 1) I_2 I_3 \dots I_{n-1}$$

C. Embedding Change Estimation

Forward cyclic and backward cyclic matricizing creates 6 matrices from 3 way tensor. Rate of intensity change in a particular region of an image always have slight variation. Therefore by analyzing the change in the bit rate, embedding change rate can be estimated. To estimate the embedding change rate we use a methodology called as hamming distance.

Let us take the same elements which are shown in the previous section.

The element $a_{i_1 i_2 \dots i_N}$ at the position with row number i_n and column number equal to

$$(i_{n+1} - 1) I_{n+2} I_{n+3} \dots I_N I_1 I_2 \dots I_{n-1} + (i_{n+2} - 1) I_{n+3} I_{n+4} \dots I_N I_1 I_2 \dots I_{n-1} + (i_N - 1) I_1 I_2 \dots I_{n-1} (i_1 - 1) I_2 I_3 \dots I_{n-1}$$

which is XORed with row number i_{n+1} and the corresponding column number is

$$(i_{n+2} - 1) I_{n+3} I_{n+4} \dots I_N I_2 I_3 \dots I_{n-2} + (i_{n+3} -$$

$$1) I_{n+4} I_{n+5} \dots I_N I_2 I_3 \dots \dots I_{n-2} + (i_N - 1) I_2 I_3 \dots I_{n-2} (i_2 - 1) I_2 I_3 \dots I_{n-1}$$

The number of change in bits can be estimated by counting the number of ones. This process can be repeated for all unfolded matrices.

Once the process is over, a graphical analysis can be done by plotting the rows of the unfolded matrix in the x axis and bit count change in the y axis as shown in the figure 4 and Table 1 and Table 2. Table 1 show 6 unfolded matrix bit change where $C_{(j)}$ and $C_{(k)}$ shows the bit change in original and embedded image in spatial domain using LSB matching steganography algorithm [12]. Table 2 show 6 unfolded matrix bit change where $C_{(j)}$ and $C_{(k)}$ shows the bit change in original and embedded image in Transform domain using LSB matching steganography algorithm [12].

The graph shown in the figure clearly shows that the input image given is the stego image since the bit change is not linear and it scatters.

III. RESULTS AND DISCUSSIONS

The framework proposed is successfully tested in the data base provided by BOWS-2[34]. All images are tested with spatial domain steganography algorithms and Transform Domain Steganography algorithms. False Identification Rate (FIR) represents that the framework incorrectly identified that some bits are modified which concludes as a

stego image. False Rejection Rate represents that the framework incorrectly identified that not bits are modified which concludes as a normal image. Table 3 clearly shows that the average detection rate for spatial domain is good with the proposed framework on the steganography algorithms [12], [16], [29], [30] is having less detection rate since the framework is applied to the full image. Detection rate can be increased by extracting the regions of an image and for each region the frame work can be applied. The average detection rate is based on the pay loads. The detection rate is less for low pay load where as good for high pay load.

IV. CONCLUSION

In this paper, a novel framework has been proposed based on 3 way tensor model. As a start of this invention, excellent results obtained for algorithms irrespective of domain. Also the framework was successful in both uniform and non uniform distribution of Stego-payloads. The disadvantage was the low average detection rate for [29] & [30] for spatial domain steganography algorithms and [27], [32] & [33] for transform domain steganography algorithms. To increase the average detection rate and to decrease the Average false acceptance rate and average false rejection rate, the proposed framework can be applied to regions in an image. If such implementation is done, then the steganalysis can be done for less payload and hence the average detection rate increases.

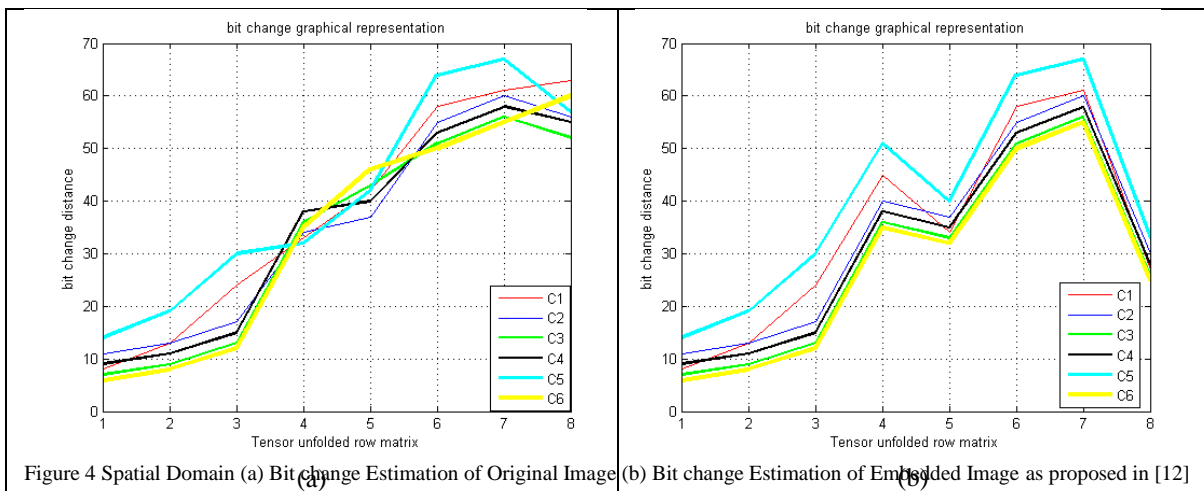


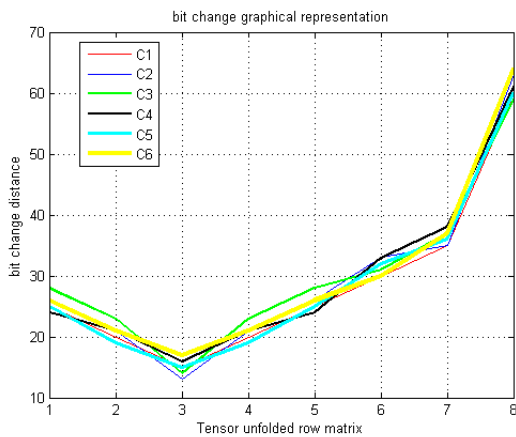
Figure 4 Spatial Domain (a) Bit change Estimation of Original Image (b) Bit change Estimation of Embedded Image as proposed in [12]

Table 1: Change Estimation in the Forward and backward cyclic Matrix in spatial domain as per [12]

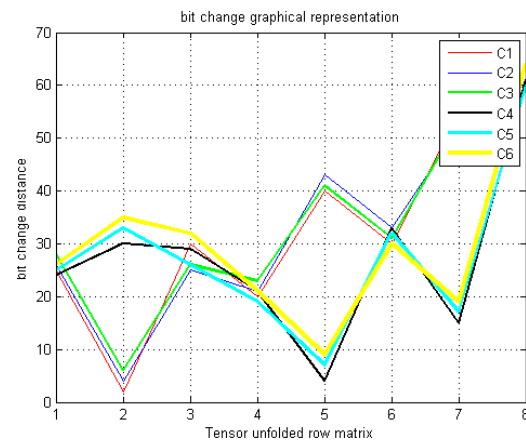
Unfolded Row Matrix	C ₍₁₎		C ₍₂₎		C ₍₃₎		C ₍₄₎		C ₍₅₎		C ₍₆₎	
	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)
i _n	8	8	11	11	7	7	9	9	14	14	6	6
i _{n+1}	13	13	13	13	9	9	11	11	19	19	8	8
i _{n+2}	24	24	17	17	13	13	15	15	30	30	12	12
i _{n+3}	33	45	34	40	36	36	38	38	32	51	35	35
i _{n+4}	42	34	37	37	43	33	40	35	42	40	46	32
i _{n+5}	58	58	55	55	51	51	53	53	64	64	50	50
i _{n+6}	61	61	60	60	56	56	58	58	67	67	55	55
i _{n+7}	63	27	56	30	52	26	55	28	57	33	60	25

Table 2: Change Estimation in the Forward and backward cyclic Matrix in transform domain as per [27]

Unfolded Row Matrix	C ₍₁₎		C ₍₂₎		C ₍₃₎		C ₍₄₎		C ₍₅₎		C ₍₆₎	
	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)	C _(j)	C _(k)
i _n	2	2	3	3	2	2	0	0	1	1	1	1
i _{n+1}	5	5	6	6	4	4	7	7	5	5	6	6
i _{n+2}	15	3	13	4	14	3	16	24	15	25	17	26
i _{n+3}	20	4	21	5	23	3	21	33	19	32	21	30
i _{n+4}	25	25	26	26	28	28	24	24	25	25	26	26
i _{n+5}	30	30	33	33	31	31	33	33	32	32	30	30
i _{n+6}	35	35	35	35	37	37	38	38	36	36	37	37
i _{n+7}	40	40	39	39	38	38	39	39	41	41	40	40



(a)



(b)

Figure 5 Spatial Domain (a) Bit change Estimation of Original Image (b) Bit change Estimation of Embedded Image as proposed in [16]

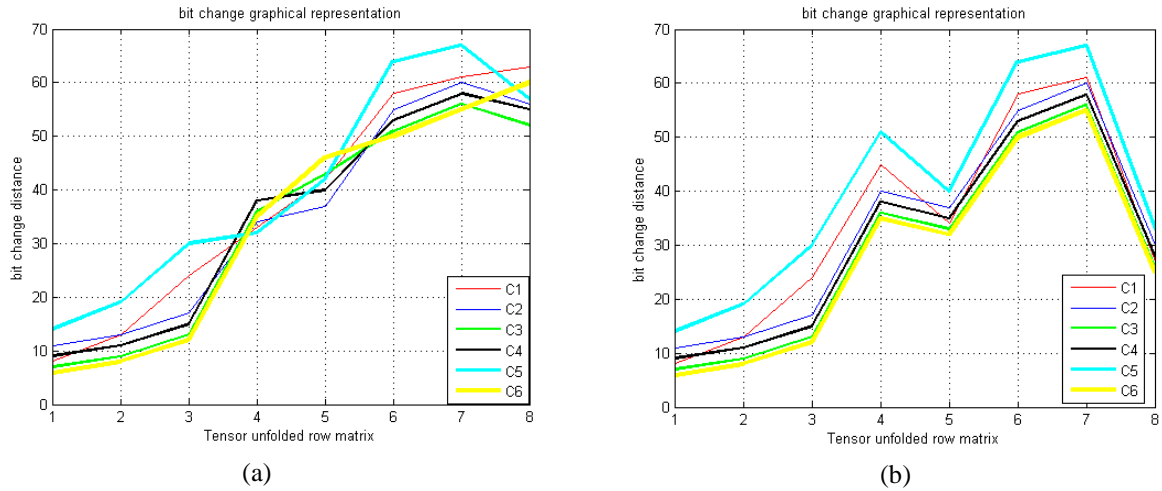


Figure 6 Spatial Domain (a) Bit change Estimation of Original Image (b) Bit change Estimation of Embedded Image as proposed in [29]

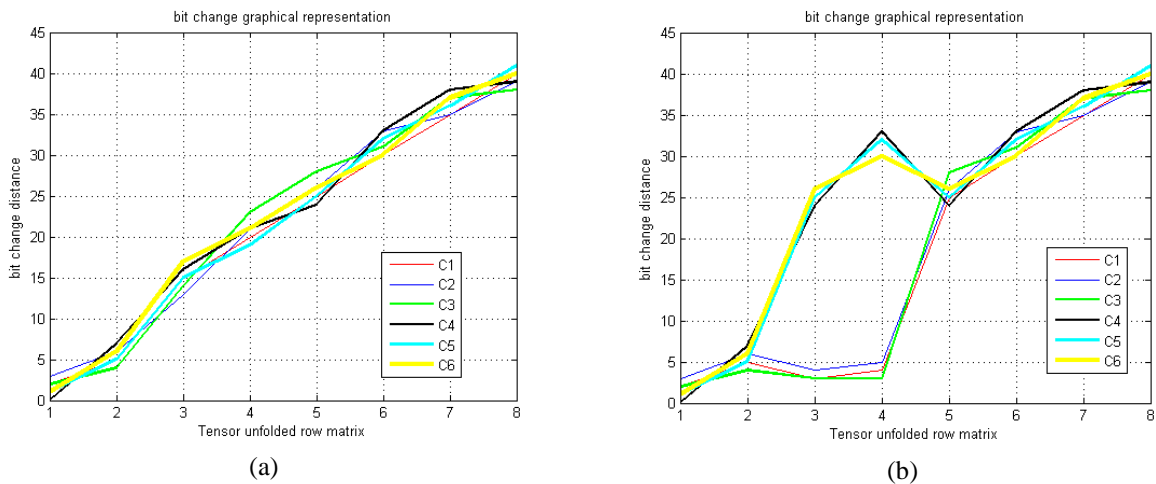


Figure 7 Transform Domain (a) Bit change Estimation of Original Image (b) Bit change Estimation of Embedded Image as proposed in [27]

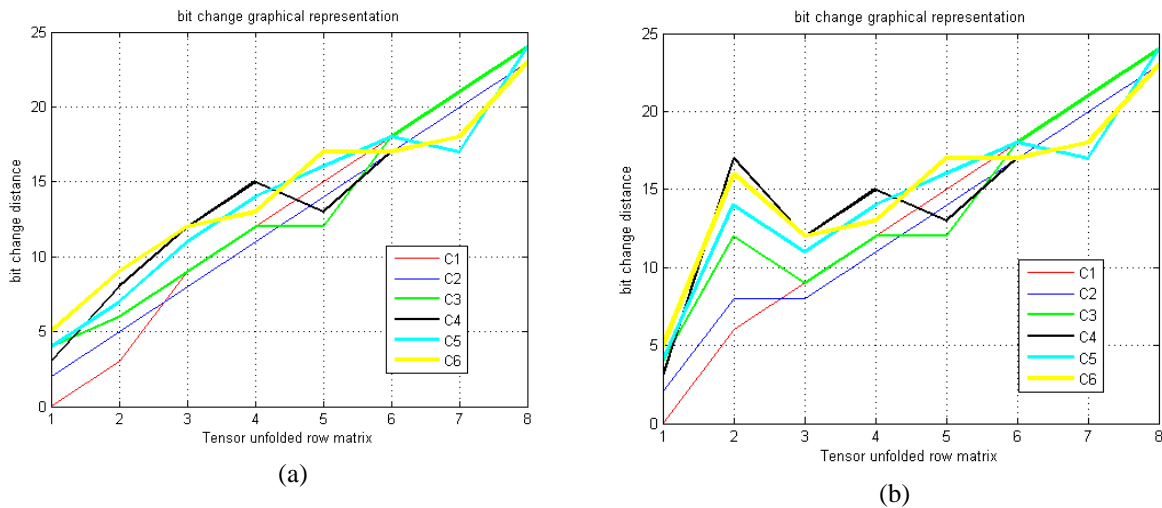


Figure 8 Transform Domain (a) Bit change Estimation of Original Image (b) Bit change Estimation of Embedded Image as proposed in [28]

Table 3: Detection Analysis of Spatial Domain Algorithms

Algorithm	Average Detection Rate	Spatial Domain	
		Average False Acceptance Rate	Average False Rejection Rate
[12]	76.23%	3.6%	.25%
[16]	82.1%	4.12%	1.67%
[29]	65.76%	9.55%	8%
[30]	72.03%	5.21%	5.34%

Table 4: Detection Analysis of Transform Domain Algorithms

Algorithm	Average Detection Rate	Spatial Domain	
		Average False Acceptance Rate	Average False Rejection Rate
JP Hide&Seek (JPHS) [31]	86.23%	0.32%	0
Jsteg [28]	89.51%	0.76%	0.47%
MBS1 [27]	76.39%	1.43%	0.238%
MMx [32]	74.12%	1.58%	1.3%
nsF5 [33]	70.28%	3.2%	1.72%

REFERENCES

- [1] Jan Kodovský and Jessica Fridrich, "Quantitative Steganalysis Using Rich Models", Proc. SPIE 8665, Media Watermarking, Security, and Forensics, March 22, 2013.
- [2] J. Fridrich and J. Kodovský. Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 7(3):868–882, June 2012.
- [3] J. Kodovský, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media", IEEE Transactions on Information Forensics and Security, 7(2):432–444, April 2012
- [4] T. Pevný, J. Fridrich, and A. Ker, "From blind to quantitative steganalysis," IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, pp. 445–454, Apr. 2012.
- [5] Chunfang Yang, Fenlin Liu, Xiangyang Luo, and Ying Zeng, "Pixel Group Trace Model-Based Quantitative Steganalysis for Multiple Least-Significant Bits Steganography", IEEE Transactions On Information Forensics And Security, vol. 8, no. 1, January 2013
- [6] Chunfang Yang, Fenlin Liu, Xiangyang Luo, Ying Zeng, "Fusion of Two Typical Quantitative Steganalysis Based on SVR", Journal Of Software, Vol. 8, No. 3, March 2013
- [7] Tomas Pevny and Andrew D. Ker, "The Challenges of Rich Features in Universal Steganalysis", Proc. SPIE 8665, Media Watermarking, Security, and Forensics, 86650M, March 22, 2013.
- [8] Ziwen Sun, Hui Li, "Quantitative Steganalysis Based on Wavelet Domain HMT and PLSR", 10th IEEE International Symposium on Distributed Computing and Applications to Business, Engineering and Science, 2011
- [9] Gikhan Gul and Faith Kurugollu, "SVD-Based Universal Spatial Domain Image Steganalysis", IEEE Transactions on Information Forensics and Security, Vol.5, No.2, June 2010
- [10] Changxin Liu, Chunjuan Oujang, "Image Steganalysis Based on Spatial Domain and DWT Domain Features", IEEE Second International Conference on Network Security, Wireless Communications and Trusted Computing, 2010.
- [11] Souvik Bhattacharyya and Gautam Sanyal, "Steganalysis of LSB Image Steganography using Multiple Regression and Auto Regressive (AR) Model", Int. J. Comp.Tech. Appl., Vol 2 (4), 1069-1077
- [12] Zhenhao Zhu, Tao Zhang, Baoji Wan, "A special detector for the edge adaptive image steganography based on LSB matching revisited", 2013 10th IEEE International Conference on Control and Automation (ICCA)
- [13] Weiqi Luo, Yuangen Wang, Jiwu Huang, "Security analysis on spatial +_1Steganography for jpeg decompressed images", IEEE Signal Processing Letters, (Volume:18, Issue: 1)
- [14] Tomas Pevny, Tomas Filler, Patrick Bas, "Using high-dimensional image models to perform undetectable steganography", Lecture Notes in Computer Science Volume 6387, 2010, pp 161-177
- [15] Mielikainen, " LSB Matching revisited", Signal Processing Letters, IEEE (Volume:13, Issue: 5), May 2006
- [16] Weiqi Luo, Fangjun Huang, and Jiwu Huang, "Edge Adaptive Image Steganography Based on LSB Matching Revisited", IEEE Transactions On Information Forensics And Security, vol. 5, no. 2, June 2010
- [17] Shunquan Tan, "Targeted Steganalysis of Edge Adaptive Image Steganography Based on LSB Matching Revisited Using B-Spline Fitting", Signal Processing Letters, IEEE (Volume:19, Issue: 6), June 2012
- [18] Chao Wang, Xiaolong Li, Bin Yang, Xiaoqing Lu, Chengcheng Liu, "Content-adaptive approach for reducing embedding impact insteganography", IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 14-19 March 2010
- [19] Yi-zhen Chen, Zhi Han, Shu-ping Li, Chun-hui Lu, Xiao-Hui Yao, "An Adaptive Steganography algorithm based on block sensitivity vectors using HVS features" 3rd International Congress on Image and Signal Processing (CISP) 16-18 Oct. 2010
- [20] Brett w. Bader and Tamara g. Kolda, "MATLAB Tensor Classes for Fast Algorithm Prototyping" ACM Transactions on Mathematical Software, Volume 32 Issue 4, December 2006 Pages 635-653

- [21] De Lathauwer, L., De Moor, B., and Vandewalle, J. 2000a, "A multilinear singular value decomposition" *SIAM J. Matrix Anal. Appl.* 21, 4, 1253–1278, 2000.
- [22] De Lathauwer, L., De Moor, B., and Vandewalle, J. 2000b, "On the best rank-1 and rank- (R_1, R_2, \dots, R_N) approximation of higher-order tensors". *SIAM J. Matrix Anal. Appl.* 21, 4, 1324–1342, 2000.
- [23] Kiers, H. A. L., "Towards a standardized notation and terminology in multiway analysis" *J. Chemometrics* 14, 105–122., 2000
- [24] The MathWorks, Inc. 2004a. Documentation: MATLAB: Programming: Classes and objects
http://www.mathworks.com/access/helpdesk/help/techdoc/matlabprog/h11_mat.html.
- [25] The MathWorks, Inc. 2004b. Documentation: MATLAB: Programming: Multidimensional arrays.
http://www.mathworks.com/access/helpdesk/help/techdoc/matlabprog/ch_dat32.html#39663.
- [26] www.sandia.gov/~tgkolda/TensorToolbox/
- [27] P. Sallee, "Model-based steganography," in *Digital Watermarking, 2nd International Workshop*, ser. Lecture Notes in Computer Science, T. Kalker, I. J. Cox, and Y. M. Ro, Eds., vol. 2939. Seoul, Korea: Springer-Verlag, New York, October 20–22, 2003, pp. 154–167.
- [28] D. Upham, <http://zooid.org/~paul/crypto/jsteg/>.
- [29] Guangjie Liu, Zhan Zhang and Yuewei Dai, "Improved LSB-matching Steganography for Preserving Second-order Statistics", *Journal of Multimedia*, vol. 5, no. 5, October 2010
- [30] J. Fridrich and D. Soukal, "Matrix embedding for large payloads," in *Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII*, E. J. Delp and P. W. Wong, Eds., vol. 6072, San Jose, CA, January 16–19, 2006, pp. W1–W10.
- [31] A. Latham, <http://linux01.gwdg.de/~alatham/stego.html>.
- [32] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *Information Hiding, 8th International Workshop*, ser. Lecture Notes in Computer Science, J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, Eds., vol. 4437. Alexandria, VA: Springer-Verlag, New York, July 10–12, 2006, pp. 314–327.
- [33] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities," in *Proceedings of the 9th ACM Multimedia & Security Workshop*, J. Dittmann and J. Fridrich, Eds., Dallas, TX, September 20–21, 2007, pp. 3–14.

Author's Profile



C. Arunvinodh received his B.E degree in Electrical and Electronics Engineering from Annamalai University, India in 2002 and the M.Tech. degree in Computer and Information Technology from Manonmaniam Sundaranar University, India in 2005. He is currently doing his P.hD in JJT University, Rajasthan. He is having 9 years of experience in teaching. He has published papers in 1 international journal, 4 International conference and 2 National Conferences. His research interests include Information Security, Digital image processing and Biometrics.



Dr. S. Poonkuntran received B.E in Information Technology from Bharathidasan University, Tiruchirapalli, India in 2003, M.Tech in Computer and Information Technology from Manonmaniam Sundaranar University, Tirunelveli, India in 2005 and Ph.D in the department of computer science and engineering, Manonmaniam Sundaranar University, Tirunelveli, India in 2011. He is having 10 years of experience in teaching and research. He is a life time member of IACSIT, Singapore, CSI, India and ISTE, India. He has published papers in 4 national conference, 22 international conferences, 1 national journal and 10 international journals on image processing, information security and soft computing. Presently he is working on Computer Vision for under water autonomous vehicles and Information Security for Healthcare Information Systems. His areas of research interest include digital image processing, soft computing and energy aware computing in computer vision.