International Journal of Computer and Internet Security.
ISSN 0974-2247 Volume 3, Number 1 (2011), pp. 61-72
© International Research Publication House
http://www.irphouse.com

A New Approach for Hiding Data in Images using Image Domain Methods

[1]P. Mohan Kumar and [2]K.L. Shunmuganathan

[1]Assistant Professor, CSE Dept,
Jeppiaar Engineering College, Chennai, India
(Research Scholar, Sathyabama University)
E-mail: mohankumarmohan@gmail.com
[2]Professor and Head, CSE Dept.,
R.M.K. Engineering College, Chennai, India
E-mail: kls_nathan@yahoo.com

Abstract

Steganography is the approach for hiding any secret message in a variety of multimedia carriers like images, audio or video files. Whenever we are hiding a data, it is very important to make it invisible, so that it could be protected. A number of steganographic algorithms have been proposed based on this property of a steganographic system. This paper introduces a new approach which is based upon spatial domain embedding techniques. And also the cover image we have used is JPEG image. The secret data embedded in images were images, text and audio signals so far. The proposed scheme has also come with the executable file as secret data. Also, the experimentation results show that, the important properties of a steganographic system such as imperceptibility, capacity of the carrier image and also resistance against the various steganalytic tools have also been achieved with this stego-system.

Keywords: Spatial domain, JPEG, executable file, steganalysis.

Introduction

The term steganography means the science of hidden communication. The way in which steganography differs from another secure data communication technique called cryptography is, the visibility of the data exchange. In cryptography, even though the actual data transaction may not be known to a third person, he may get a doubt that some abnormal or suspicious communication is taking place. But, in case of steganography, the hidden communication will never come to the notice of the

eavesdropper. Because, the carrier signal we are using to hide the secret data is going to be innocent. So, we can call the technique as information hiding [1-4].

Another technique which is based on the information hiding strategy is digital watermarking. But, in case of digital watermarking, the important property of information hiding known as resistance to removal is preferred. So, in these applications, we are not worrying about imperceptibility but resistance to removal. This is mainly used in commercial applications like copyright protection of digital forms of media like video or image. Unlike image steganography, digital watermarking techniques mainly concentrate on keeping logos or any other symbols or images in the carrier data. And also it is made sure that those signals embedded are not able to be removed by any other person. There are a number of watermarking techniques have been explained in [5-7].

For a long period of time many researchers have been involved in developing new steganographic systems. Meanwhile, the development of steganalytic tools are also started growing. Steganalysis is a process of finding the existence of a secret data in a cover media [8]. Whenever a suspicious image is received, the main task of a steganalytic tool is to find the algorithm used for hiding secret data in the image. Most of the steganographic algorithm developers are also trying to crack their own algorithm using the existing steganalytic tools, so that the strength and weaknesses of their system may be found.

Related work

Generally, digital image steganography is a way to exchange secret data. So, the important components of a steganographic system include an embedding / extracting algorithm, secret key which is going to be shared by the sender and receiver of the secret data and also a communication channel which is considered to be more secure [9]. The general frame work for a steganographic system is shown as the figure.

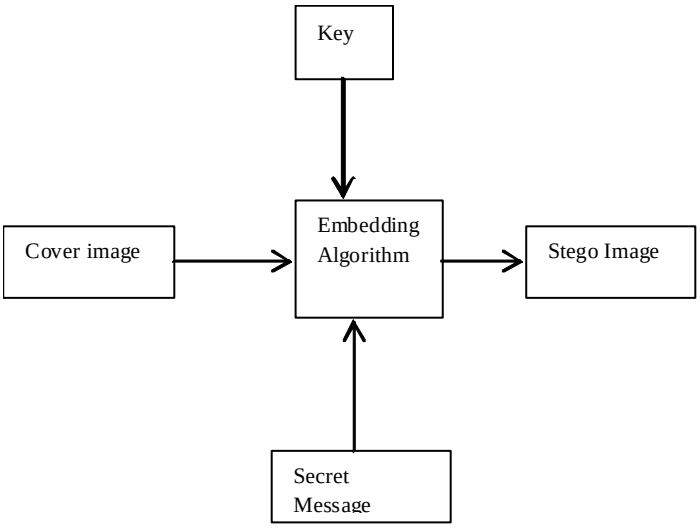

Figure 1: A simple image steganography scheme.

This framework has been derived from the popular idea called prisoner's problem [9]. In this approach Alice and Bob were trying to exchange an escape plan without the knowledge of the warden. Some of the terms used in steganographic system are cover-media (the digital media which is used to hide secret data), secret data (the important data to be hidden) and stego-media (after embedding the secret message in the cover media). The hidden data cannot be detectable when we are performing the embedding phase randomly and also the level of independence between the secret message and cover as well as stego objects [10]. There are many other ways for providing more security includes the usage of encryption-decryption functions for embedding and extraction of secret data [11]. Since JPEG images are widely exchanged through internet, choosing JPEG image for sending secret message to the receiver will never be suspicious. And also, the redundancies that are appearing in JPEG images help us to hide more information securely. Methods for improving the hiding capacity of a JPEG image have been explained in [12].

The Least Significant Bits replacement method or LSB method is the very simple and a commonly used approach for developing steganographic system. Because the amount of space that an image can provide for hiding data will be more comparing with other algorithms. And also the implementation of this technique is also very easy. In this approach, the image pixel's LSB is replaced by one bit of secret data [13]. Spatial domain embedding technique is also known as image domain. The techniques that are following spatial domain embedding are embedding the secret message in the intensity of the cover image pixels. Spatial domain techniques include bit-wise methods that apply bit insertion and noise manipulation techniques [14]. The main disadvantage of LSB replacement is that, while hiding secret data in the image, some of the pixels will never be modified or replaced with the secret bits, since we are using pseudo random generator for placing the secret message bits. As a result, very simple steganalytic tool could trace the existence of the secret message.

But this problem of asymmetry can easily be avoided by an alternate scheme using a LSB matching scheme. In this technique, if the secret bit is not matching with the LSB of cover image, then ±1 will be added randomly. By doing so, we can reduce the probability of increase or decrease in the pixel value modification can be avoided. So, we can eliminate the problem we faced in LSB replacement technique. Also, the steganalytic algorithms which can find the stego-images which were obtained from LSB replacement technique cannot find the stego-images we got from LSB matching.

There are several steganalytic algorithms found for finding stego-images which were got by LSBM (LSB matching) technique. In [15], the image is being taken and its two least significant bit planes are considered. The bit planes are split into 3*3 overlapped sub images. According to the number of gray levels those sub images are classified. In one sub image, the LSBM is applied and found that the alteration rate of cover image is higher than that of stego-image. In [16], the authors have compared the function of LSBM to a low pass filter through the histogram of the image. They found that the no. of high frequency components is very less comparing to the original cover image. But later in [17], this method is found that it will not be working well in case of gray scale images. As a remedy, the author has proposed techniques using down-sampled image and adjacency histogram instead of traditional histogram.

Instead of handling pixel values independently, the other technique proposed by Jarno in [18], is using a pair of pixels for embedding which is known as LSBM revisited (LSBMR). In this technique, the author has proposed an approach for data hiding, in which the LSB of the first pixel carries one bit of information, and a function of the two pixel values carries another bit of information. So, in this approach the changes that are made in the cover image are very few. Also, the modification rates of pixels have been greatly reduced. But all the techniques analyzed above are not taking care of the relationship between the pixel and its neighborhood.

There are many data embedding schemes analyzed which are taking the relationship of a pixel to its neighbor. In [19], a hiding scheme has been proposed by replacing the LSB of a cover according to the difference values between a pixel and its four touching neighbors. This method uses the edges of an image for hiding secret data. Although this method can achieve more visually imperceptible stego-images, the security performance is poor. Since the method just modifies the LSB of image pixels when hiding data, it can be easily detected by existing steganalytic algorithms.

The pixel Value differencing is another type of edge based data hiding scheme, which has been proposed in [20], in which the number of embedded bits is determined by the difference between a pixel and its neighbor. If there is large difference between the pixels, the number of secret bits that can be embedded will also be large. Also based on the experimental results, this approach can provide a larger embedding capacity.

Mostly, all the techniques discussed so far are using a random number generator which will spread the secret data throughout the image which may lead to lower embedding rates. But based on the experimental results, we have found that most of the approaches fail to provide security and stego-images of preferred quality. Also, the usage of JPEG images in internet now a days is very high, we have proposed an approach which is using the JPEG image as cover image. But the thing is, sufficient secret message bits can be embedded into cover image, so that we can achieve better secure payload. To make the JPEG image useful for embedding more data, an encoder scheme has been proposed in [12]. In this research, it is found that, it is necessary to enlarge the hiding capacity of JPEG2000 baseline system because the available redundancy is very limited. In addition, the bit stream truncation makes it difficult to hide information. These problems faced are being eliminated by introducing an encoder which eliminates bit truncation.

Proposed System

The proposed architecture of our steganographic system is shown in the following figure. In the embedding phase, the quantized subband is divided into codeblocks that are units of bit plane encoding. Bit-plane encoding is operated bit-plane by bit-plane, from high to low, to produce independent bitstream for each block. Each bit-plane is encoded in a sequence of three fractional bit-plane coding passes. Then, an adaptive arithmetic coding strategy, known as the MQ coder, is employed to encode the bitstream. The EBCOT algorithm produces a finely embedded bitstream with many

useful truncation points. The bitstream can be truncated at the end of any coding passes to get desired compression ratio. The truncation point of every codeblock is determined by rate-distortion optimization. JPEG2000 uses uniform scalar quantizers with enlarged "deadzones." Truncating the embedded bitstream associated with any given codeblock has the effect of quantizing the wavelet coefficients in that codeblock more coarsely. After that the code stream from the encoder is fed into the stego system.

The system initializes some parameters, which are used for subsequent data preprocessing and region selection, and then estimates the capacity of those selected regions. If the regions are large enough for hiding the given secret message, then data hiding is performed on the selected regions. Finally, it does some postprocessing to obtain the stego image. Otherwise the scheme needs to revise the parameters, and then repeats region selection and capacity estimation until can be embedded completely. Please note that the parameters may be different for different image content and secret message. We need them as side information to guarantee the validity of data extraction. In practice, such side information (7 bits in our work) can be embedded into a predetermined region of the image. In data extraction, the scheme first extracts the side information from the stego image. Based on the side information, it then does some preprocessing and identifies the regions that have been used for data hiding. Finally, it obtains the secret message according to the corresponding extraction algorithm. In this paper, we apply such a region adaptive scheme to the spatial LSB domain. We use the absolute difference between two adjacent pixels as the criterion for region selection, and use LSBMR as the data hiding algorithm. The details of the data embedding and data extraction algorithms are as follows

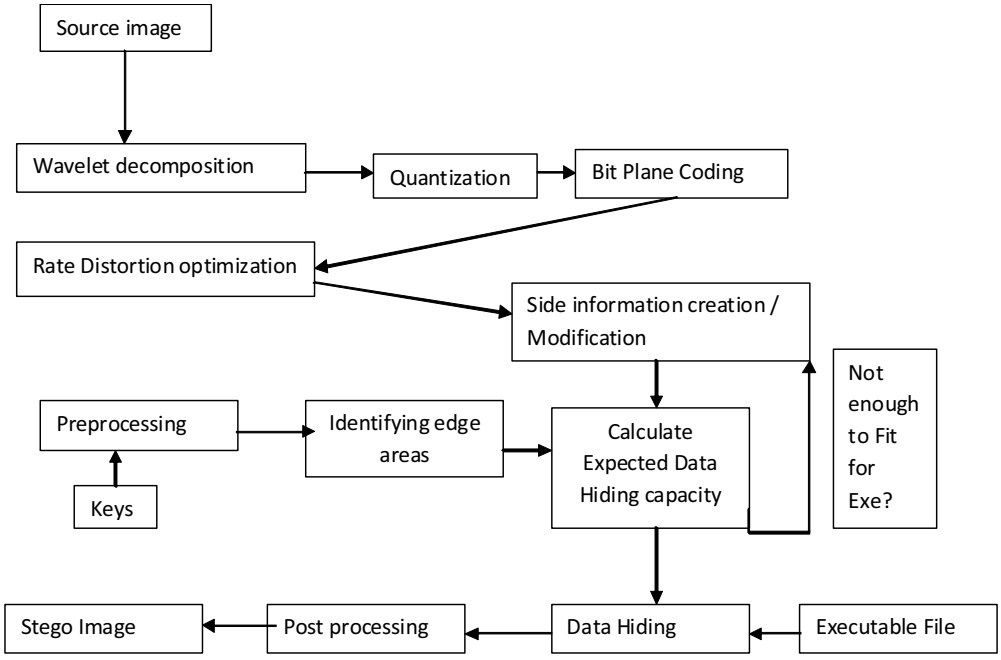

Figure 2a: Proposed Executable file Embedding Architecture.

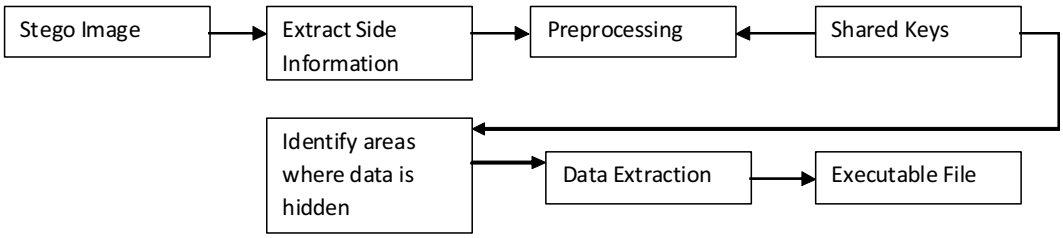

Figure 2b: Proposed Executable file extraction architecture.

Figure 3 a: Sample images used.

Data Embedding

1. The cover image of size of m*n is first divided into non-overlapping blocks of Bz*Bz pixels. For each small block, we rotate it by a random degree in the range of {0,90,180,270} , as determined by a secret key key1. The resulting image is rearranged as a row vector V by raster scanning. And then the vector is divided into non-overlapping embedding units with every two consecutive pixels (xi, xi+1) , where i=1,3,…mn-1 , assuming n is an even number. Two benefits can be obtained by the random rotation. First, it can prevent the detector from getting the correct embedding units without the rotation key key1, and thus security is improved. Furthermore, both horizontal and vertical edges (pixel pairs) within the cover image can be used for data hiding.
2. According to the scheme of LSBMR, 2 secret bits can be embedded into each embedding unit. Therefore, for a given secret message M, the threshold T for region selection can be determined as follows. Let EU(t) be the set of pixel pairs whose absolute differences are greater than or equal to a parameter t

$$EU(t)=\{(xi,xi+1)||xi-xi+1|\geq t\forall(xi,xi+1)\in V\}$$

Then we calculate the threshold T by

$$T=\arg\max t\{2*|EU(t)|\geq|M|$$

Where $t\in\{0,1,..31\}$,|M| is the size of the secret message M, and |EU(t)| denotes the total number of elements in the set of EU(t) .Please note that when T=0 , the proposed method becomes the conventional LSBMR scheme, which means that our

method can achieve the same payload capacity as LSBMR (except for 7 bits).

3. Performing data hiding on the set of

$$EU(T)=\{(x_i,x_{i+1})||x_i-x_{i+1}|\geq T, \forall (x_i,x_{i+1})\in V\}$$

We deal with the above embedding units in a pseudorandom order determined by a secret key key2 . For each unit (xi,xi+1), we perform the data hiding according to the following four cases.

Case 1: $LSB(x_i)=m_i$ & $f(x_i,x_{i+1})=m_{i+1}$

$$(x_i`,x_{i+1}`)=(x_i,x_{i+1})$$

Case 2: $LSB(x_i)=m_i$ & $f(x_i,x_{i+1})\neq m_{i+1}$

$$(x_i`,x_{i+1}`)=(x_i,x_{i+1}+r)$$

Case 3: $LSB(x_i)\neq m_i$ & $f(x_{i-1},x_{i+1})=m_{i+1}$

$$(x_i`,x_{i+1}`)=(x_{i-1},x_{i+1})$$

Case 4: $LSB(x_i)\neq m_i$ & $f(x_{i-1},x_{i+1})\neq m_{i+1}$

$$(x_i`,x_{i+1}`)=(x_{i+1},x_{i+1})$$

Where m_i and m_{i+1}denote two secret bits to be embedded.

4. After data hiding, the resulting image is divided into non-overlapping B_z*B_z blocks. The blocks are then rotated by a random number of degrees based on key_1. The process is very similar to Step 1 except that the random degrees are opposite. Then we embed the two parameters (T,B_z) into a preset region which has not been used for data hiding. Please note that there are two parameters in our approach. The first one is the block size B_z for block dividing in data preprocessing; Another is the threshold T for embedding region selection. In this paper, B_z is randomly selected from the set of {1,4,8,12}, T belongs to {0,1,..31} and can be determined by the image contents and the secret message M (please refer to Step 2). In all, only 7 bits of side information are needed for each image.

Data Extraction

To extract data, we first extract the side information, i.e., the block size B_z and the threshold T from the stego image. We then do exactly the same things as Step 1 in data embedding. The stego image is divided into B_z*B_z blocks and the blocks are then rotated by random degrees based on the secret key key1. The resulting image is rearranged as a row vector V' . Finally, we get the embedding units by dividing V' into non-overlapping blocks with two consecutive pixels. We travel the embedding units whose absolute differences are greater than or equal to the threshold T according

to a pseudorandom order based on the secret key , until all the hidden bits are extracted completely. For each qualified embedding unit, say, (x_i', x_{i+1}') , where $|x_{i+1}' - x_i'| \geq T$, we extract the two secret bits m_i, m_{i+1} as follows:

$$m_i = LSB(x_i''),\ m_{i+1} = LSB(\lfloor x_i'/2 \rfloor + x_{i+1}')$$

Experimental Results and analysis

One of the important properties of our steganographic method is that it can first choose the sharper edge regions for data hiding according to the size of the secret message by adjusting a threshold T. As illustrated, the larger the number of secret bits to be embedded, the smaller the threshold T becomes, which means that more embedding units with lower gradients in the cover image can be released (please refer to the definition of EU(T) in Step 3 in data embedding). When is 0, all the embedding units within the cover become available? In such a case, our method can achieve the maximum embedding capacity of 100% (100% means 1 bpp on average for all the methods in this paper), and therefore, the embedding capacity of our proposed method is almost the same as the LSBM and LSBMR methods except for 7 additional bits.

Table 1: Average Accuracy(%) of RS Features set on FLD with Different embedding Rates.

Embedding Rate Methods	10%	20%	30%	40%	50%
Existing (HBC)	88	91	94	98	99
Proposed	51	52	51	50	53

Table 2: Data for drawing ROC Curves.

False Positive Rate	0	0.1	0.2	0.3	0.5
True Positive Rate	0.3	0.5	0.6	0.65	0.7

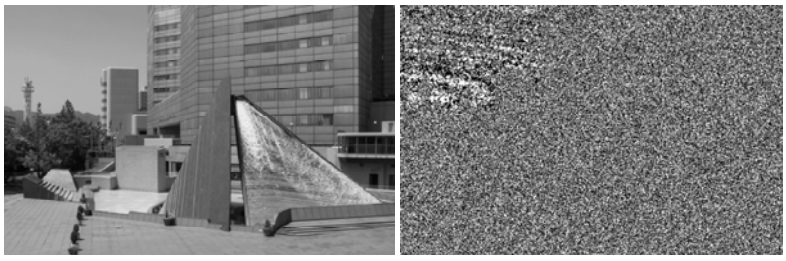

Figure 3a: Cover image and cover image LSB.

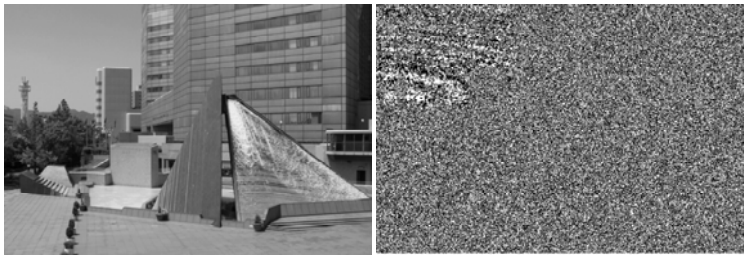

Figure 3b: Stego image and LSB of Stego image.

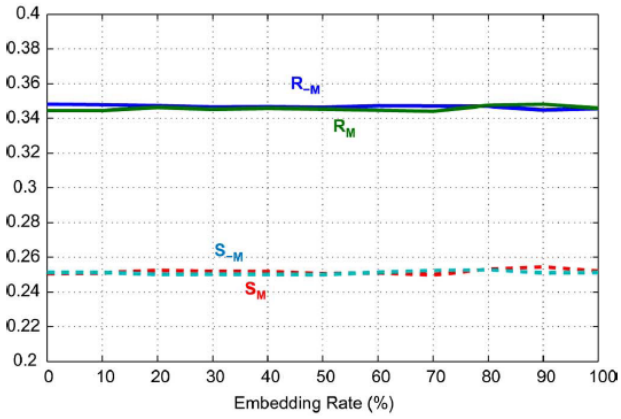

Figure 4a: RS diagram for HBC.

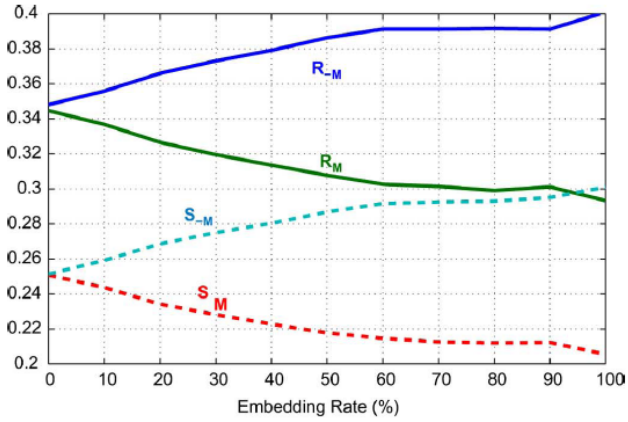

Figure 4b: RS Diagram for proposed method.

Conclusion

In this paper, an image steganographic scheme in the spatial LSB domain is studied in which an edge based scheme also included. Normally, there exist some smooth regions in natural images, which would cause the LSB of cover images not to be completely random or even to contain some texture information just like those in

higher bit planes. If embedding a message in these regions, the LSB of stego images becomes more random, and according to our analysis and extensive experiments, it is easier to detect. In most previous steganographic schemes, however, the pixel/pixel-pair selection is mainly determined by a PRNG without considering the relationship between the characteristics of content regions and the size of the secret message to be embedded, which means that those smooth/flat regions will be also contaminated by such a random selection scheme even if there are many available edge regions with good hiding characteristics. The experimental results evaluated on thousands of natural images using different kinds of steganalytic algorithms show that both visual quality and security of our stego images are improved significantly compared to typical LSB-based approaches and their edge adaptive versions. Furthermore, it is expected that our adaptive idea can be extended to other steganographic methods such as audio/video steganography in the spatial or frequency domains when the embedding rate is less than the maximal amount.

Acknowledgement

We take immense pleasure in thanking our chairman Dr. Jeppiaar M.A, B.L, Ph.D, the Directors of Jeppiaar Engineering College Mr. Marie Wilson, B.Tech, MBA, (Ph.D), Mrs. Regeena Wilson, B.Tech, MBA, (Ph.D) and the principal Dr. Sushil Lal Das M.Sc(Engg.), Ph.D for their continual support and guidance. We would like to extend our thanks to my guide, our friends and family members without whose inspiration and support our efforts would not have come to true. Above all, we would like to thank God for making all our efforts success.

References

[1] N. Provos and P. Honeyman, “Hide and seek: An introduction to steganography,” IEEE security and Privacy Mag., vol. 1, no. 3, pp. 32–44, 2003.

[2] J. Fridrich, “Applications of Data Hiding in Digital Images,” Tutorial for the ISSPA, pp. 22-25, Aug. 1999

[3] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Information hiding—A survey,” Proc. IEEE, vol. 87, no. 7, pp. 1062–1078, Jul.1999.

[4] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, “Techniques for data hiding,” IBM Syst. J., vol. 35, no. 3–4, pp. 313–336, 1996.

[5] M. Swanson, M. Kobayashi, and A. Tewfik, “Multimedia data embedding and watermarking technologies,” IEEE Proceedings, vol. 86, No. 6, pp 1064-1087, June 1998.

[6] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” Proc. First Int. Workshop Information Hiding, R. Anderson, Ed., Cambridge, U. K.: Springer-Verlag, May/June 1996, pp. 183-206.

[7] I.J. Cox, M.L. Miller, and J.A. Bloom, Digital Watermarking. Morgan Kaufmann, 2002.
[8] Tomáˇs Pevný and Jessica Fridrich, Multiclass Detector of Current Steganographic Methods for JPEG Format, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 3, NO. 4, DECEMBER 2008, 635-650.
[9] G. Simmons," The prisoner's problem and the subliminal channel, CRYPTO, pp: 51-67, 1983.
[10] J. Zollner, H. Federrath, "Modelling the security of steganographic systems", 2nd information hiding workshop, pp: 345-355, 1998.
[11] N.J. Hopper, J. Langford, L. Von Ahn, "Provably secure steganography", Advances in cryptology: CRYPTO 2002.
[12] A high capacity Steganography scheme for JPEG 2000 Baseline system, Liang Zhang, Haili wang, Renbiao Wu, IEEE transactions on image processing, 18(8), 2009.
[13] T. Morkel 1, J.H.P. Eloff 2, M.S. Olivier , AN OVERVIEW OF IMAGE STEGANOGRAPHY http://mo.co.za/open/stegoverview.pdf
[14] Johnson, N.F. & Jajodia, S., "Steganalysis of Images Created Using Current Steganography Software",Proceedings of the 2nd Information Hiding Workshop, April 1998
[15] F. Huang, B. Li, and J. Huang, "Attack LSB matching steganography by counting alteration rate of the number of neighbourhood gray levels," in Proc. IEEE Int. Conf. Image Processing, Oct. 16–19, 2007, vol. 1, pp. 401–404.
[16] J. Harmsen and W. Pearlman, "Steganalysis of additive-noise modelable information hiding," Proc. SPIE Electronic Imaging, vol. 5020, pp. 131–142, 2003.
[17] A. D. Ker, "Steganalysis of LSB matching in grayscale images," IEEE Signal Process. Lett., vol. 12, no. 6, pp. 441–444, Jun. 2005.
[18] J. Mielikainen, "LSB matching revisited," IEEE Signal Process. Lett., vol. 13, no. 5, pp. 285–287, May 2006.
[19] K. Hempstalk, "Hiding behind corners: Using edges in images for better steganography," in Proc. Computing Women's Congress, Hamilton, New Zealand, 2006.
[20] D. Wu and W. Tsai, "A steganographic method for images by pixelvalue differencing," Pattern Recognit. Lett., vol. 24, pp. 1613–1626, 2003.

Authors Profile

Dr. K.L. Shanmuganathan B.E, M.E., M.S., Ph.D. works as the Professor & Head of CSE Department of RMK Engineering College, Chennai, Tamil Nadu, India. He has more than 18 years of teaching experience and his areas of specializations are Artificial Intelligence, Computer Networks and DBMS.

P. Mohan Kumar B.E., M.E., (Ph.D.) works as Associate Professor in Jeppiaar Engineering College and he has more than 8 years of teaching experience. His areas of specializations are Network security, Image processing and artificial intelligence.