International Journal of Computers, Communications & Control
Vol. II (2007), No. 2, pp. 149-158

Blind Steganalysis: Estimation of Hidden Message Length

Sanjay Kumar Jena, G.V.V. Krishna

Abstract: Steganography is used to hide the occurrence of communication. Discovering and rendering useless such covert message is an art of steganalysis. The importance of techniques that can reliably detect the presence of secret messages in images is increasing as images can hide a large amount of malicious code that could be activated by a small Trojan horse type of virus and also for tracking criminal activities over Internet. This paper presents an improved blind steganalysis technique. The proposed algorithm reduces the initial-bias, and estimates the LSB embedding message ratios by constructing equations with the statistics of difference image histogram. Experimental results show that this algorithm is more accurate and reliable than the conventional difference image histogram method. It outperforms other powerful steganalysis approaches for embedded ratio greater than 40% and comparable with RS steganalysis technique for shorter hidden message length.

Keywords: Steganography, steganalysis, hidden message extraction

1 Introduction

Steganography is the art of passing information through apparently innocent files in a manner that the very existence of the message is unknown. The term steganography in Greek literally means, "Covered Writing". The innocent files can be referred to as cover text, cover image, or cover audio as appropriate. After embedding the secret message it is referred to as stego-medium. A stego-key is used to control the hiding process so as to restrict detection and/or recovery of the embedded data. While cryptography is about protecting the content of messages (their meaning), steganography is about hiding the message so that intermediate persons cannot see the message.

Historically, steganography has been a form of security through obscurity where the security lies in that only sender and receiver know the method in which the message is hidden. This is in violation of Kirchoff's principle, which states that the security should lie in key alone. Steganography can be either "linguistic steganography"or "technical steganography"[1]. The ancient techniques that hide messages physically are called as technical steganographic systems. They include microdots, tattoos, invisible inks and semagrams. Recent techniques belong to the linguistic steganography. These techniques hide message in the cover images, which are of digital form.

Steganalysis is the process of detecting the existence of the steganography in a cover medium and rendering it useless. Current trend in steganalysis [4] seems to suggest two extreme approaches (a) little or no statistical assumptions about the image under investigation where statistics are learnt using a large database and (b) a parametric model is assumed for the image and its statistics are computed for steganalysis detection. The messages embedded into an image are often imperceptible to human eyes. But there exists some detectable artifacts in the images depending on the steganographic algorithm used [2,5]. The steganalyst uses these artifacts for the detection of the steganography.

By far the most popular and frequently used steganographic method is the Least Significant Bit embedding (LSB). It works by embedding message bits as the LSB's of randomly selected pixels. Several techniques for the steganalysis of the images for LSB embedding are present. Fridrich and Goljan [6,7] proposed LSB Steganography dual detection method, named RS method, based on probability statistics in the color or grayscale images. The basic idea is that LSB plane seems random in the typical cover images, but to some extent the other 7 bit planes could predict it. This method is suitable for detection of the non-sequential steganography reliably.

Pfitzmann and Westfeld [8] introduced a method based on statistical analysis of Pairs of Values (PoVs) that are exchanged during message embedding This method, which became known as the χ^2 attack, is quite general and can be applied to many embedding paradigms besides the LSB embedding. It provides very reliable results when the message placement for sequential embedding. Fridrich et al. [9] developed a steganographic method for detecting LSB embedding in 24-bit color images-the Raw Quick Pairs (RQP) method. The new method is based on analyzing close pairs of colors created by LSB embedding. On the condition that the number of unique colors in the cover image will be less than 30 percent that of the total pixels, it works reasonably well. When the number of unique colors exceeds about 50 percent that of total pixels, the results gradually become unreliable. This frequently happens for high resolution raw scans and images taken with digital cameras stored in an uncompressed format. Another disadvantage of the RQP method is that it can't be applied to grayscale images.

There are few papers in the field of the detecting pixels that contain the hidden message. Ian Davidson and Goutam Paul [10] proposed the hidden message location problem as outlier detection using probability/energy measures of images. Pixels contributing the most to the energy calculations of an image are deemed outliers. Though results for grayscale images are quite accurate; they are not as good as for color images. The algorithm can be defeated if the steganography algorithm has knowledge of probability/energy function or if the message is carefully embedded in the high-energy regions of an image.

The difference image histogram proposed by T.Zhang and X.Ping [10] consists of an initial-bias. The proposed algorithm constructs the embedding ratio-estimate equations using the difference image histogram and reduces the initial bias. Experimental results show that the novel algorithm is more accurate than the conventional difference image histogram method and other steganalysis techniques.

In the following section, we review principle of the difference image histogram method, and then in section 3, describe the improved difference image histogram method (IDIH) algorithm. Sections 4 show the experimental results and conclude this paper in section 5.

2 Principles Of Difference Image Histogram

Tao Zhang and Xijian Ping introduced the difference image histogram method, which uses the measure of weak correlation between successive bit planes to construct a classifier for discrimination between stego-images and cover images. Considering the property of LSB steganography, the difference image histogram is used as a statistical analysis tool. The difference image is defined as

$$D(i,j) = I(i+1,j) - I(i,j). \tag{1}$$

Where $I(i,j)$ denotes the value of the image I at the position (i,j). T.Zhang and X.Ping found that there exists difference between the difference image histograms for normal image and the image obtained after flipping operation on the LSB plane. This fact is utilized to realize the steganalysis technique. To explain the details of difference image histogram method (DIH), we need to define some notions, Let I be the test image, which has $M \times N$ pixels. The embedding ratio p is defined as the percentage of the embedded message length to the maximum capacity.

If the difference image histogram of an image is represented by h_i, that of the image after flipping all bits in the LSB plane by f_i, and that of the image after setting all bits in the LSB plane to zero by g_i, there exist the following relationships between hi, f_i and g_i.

$$\begin{aligned} h_{2i} &= f_{2i} = a_{2i,2i} g_{2i}, \\ h_{2i+1} &= a_{2i,2i+1} g_{2i} + a_{2i+2,2i+1} g_{2i+2}, \\ f_{2i+1} &= a_{2i,2i-1} g_{2i} + a_{2i+2,2i+3} g_{2i+2} \end{aligned} \tag{2}$$

in which $a_{2i,2i+j}$ is defined as the transition coefficient from the histogram g_i to h_i. When $j = 0, 1, -1$ then $0 < a_{2i,2i+j} < 1$, otherwise $a_{2i,2i+j} = 0$, and they satisfy

$$a_{2i,2i-1} + a_{2i,2i} + a_{2i,2i+1} = 1 \tag{3}$$

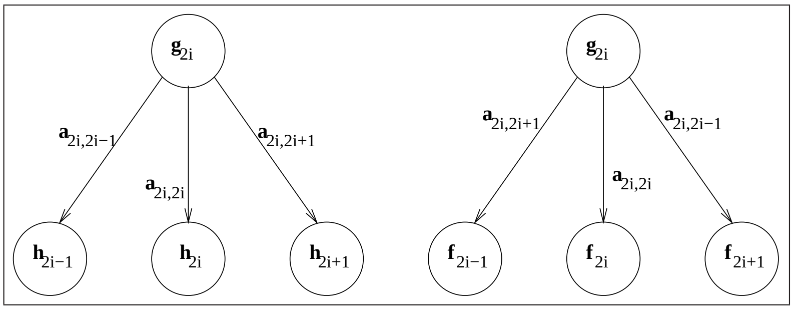

Figure 1: The transition diagram from g_i to h_i,f_i

Starting from the approximate symmetry of the difference histogram about $i = 0$, we first get $a_{0,1} \sim a_{0,-1}$ From the above Equations (2) we obtain the following iterative formula for calculating transition coefficients for all positive integers i:

$$\begin{aligned} a_{0,1} = a_{0,-1} &= \frac{g_0 - h_0}{2g_0}, \\ a_{2i,2i} &= \frac{h_{2i}}{g_{2i}}, \\ a_{2i,2i-1} &= \frac{h_{2i-1} - a_{2i-2,2i-1}g_{2i-2}}{g_{2i}}, \\ a_{2i,2i+1} &= 1 - a_{2i,2i} - a_{2i,2i-1}. \end{aligned} \tag{4}$$

Assuming the embedded hidden message forms a random bit sequence, for the stego-image with the LSB plane fully embedded (i.e.$p = 100\%$) the LSB plane is independent of the neighboring bit planes. Therefore, for such stego-images we have $a_{2i,2i+1} \approx 0.25$, $a_{2i,2i+1} \approx 0.5$, $a_{2i,2i+1} \approx 0.25$.

The h_{2i+1} consists of two parts: $a_{2i,2i+1}g_{2i}$ and $a_{2i+2,2i+1}g_{2i+2}$. Statistical tests show that for natural images these two parts make an approximately equal contribution to h_{2i+1}, that is

$$a_{2i,2i+1}g_{2i} \approx a_{2i+2,2i+1}g_{2i+2} \tag{5}$$

If $\alpha_i = (a_{2i+2,2i+1})/(a_{2i,2i+1})$, $\beta_i = (a_{2i+2,2i+3})/(a_{2i,2i-1})$ and $\gamma_i = g_{2i}/g_{2i+2}$ then the statistical hypothesis of the steganalytic method is that for a natural image the following equation should be satisfied:

$$\alpha_i \approx \gamma_i \tag{6}$$

while for stego-images with the LSB plane fully embedded

$$\alpha_i \approx 1. \tag{7}$$

The physical quantity α_i, can be viewed as the measure of the weak correlation between the LSB plane and its neighboring bit planes. From further experiments they got that for a given i the value of α_i, decreases monotonically with the increasing length of embedded secret messages (p) and when the embedding ratio p increases to 100%, α_i decreases to 1 approximately.Figure-2 shows the functional relation between α_i and the embedding ration p when $i = 0$ for the “Lena”image.

The relationship between α, and the embedding ratio p will be modeled using a quadratic equation $y = ax^2 + bx + c$. The following four critical points $P_1 = (0, \gamma_i)$, $P_2 = (p, \alpha_i)$, $P_3 = (1, 1)$, and $P_4 =$

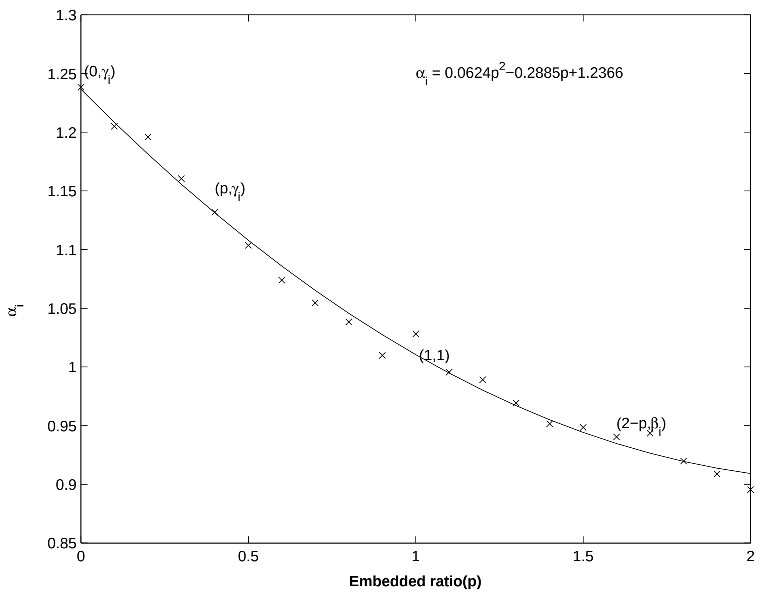

Figure 2: The functional relation between α_i and p$(i = 0)$

$(2-p, \beta i)$. Now the following equation set is obtained:

$$\begin{aligned} c &= \gamma_i, \\ ap^2 + bp + c &= \alpha_i, \\ a + b + c &= 1, \\ a(2-p)^2 + b(2-p) + c &= \beta_i. \end{aligned} \tag{8}$$

Assume $d_1 = 1 - \gamma_i$, $d_2 = \alpha_i - \gamma_i$, $d_3 = \beta_i - \gamma_i$, then above equation set (8) can be simplified to

$$2d_1p^2 + (d_3 - 4d_1 - d_2)p + 2d_2 = 0 \tag{9}$$

The embedding ratio p can be obtained from the root of above whose absolute value is smaller. If the discriminant is smaller than zero, then $p \approx 1$.

3 Principles of Improved Difference Image Histogram steganalysis

The Difference image histogram algorithm was primarily based on the statistical hypothesis that for the natural images

$$\alpha_i \approx \gamma_i \tag{10}$$

and for a stego-images with the LSB plane fully embedded

$$\alpha_i \approx 1. \tag{11}$$

Obviously, the hypotheses given in Equations(10) and (11) will affect the precision of the Difference image histogram method. Once in these hypotheses there exists some initial bias, the estimate value via the Equation(9) will not be reliable. When the embedding ratio is low, the bias of these hypotheses will lead the incorrect decision, and if there are no embedding messages in images, the false alarm rate is high. Table 1 will show the mean and variance of the γ_i to αi value. With the increase in i the variance increases and the mean begins to deviate from 1. In some cases the detection lead to an incorrect decision

	i=0	i=1	i=2	i=3
mean	1.0013	1.0034	1.0079	1.0443
variance	7.6E-04	1.4E-03	2.8 e-03	4.9E-03

Table 1: Statistical data on the ratio of γ_i to αi for natural images

of estimating more than 1% embedding, for the normal images.

The Figure "Proposed" shows the initial value of difference between α_i and γ_i for a lena image and Figure "Proposed (b)" shows a close look at the α_i and γ_i at $p = 0$ values. This initial deviation may lead a serious estimate error. The initial-bias in detection of the message in the normal image is affecting the detection of stego-images, as the error present in the normal image will effect the estimation of the hidden message length for stego-image.

If the stego-image created with embedded embedding ratio p is denoted as S_p, and the image created by flipping all bits in the LSB plane of S_p as R_p, the value of α_i can be calculated for the images S_p and R_p(note that the value of α_i for the images R_p is equal to the value of β_i for the image S_p). Moreover, we note that in S_p only $p/2$ of the pixels are flipped by message embedding, while in R_P about $1-(p/2)$ of pixels are flipped. Therefore, R_p is equivalent to a “stego-image”with “embedding ratio”$2-p$. So given a stego-image we can calculate the values of α_i at p and $2-p/2$, as the value of β_i at p is equal to the value of α_i at value $2-(p/2)$.

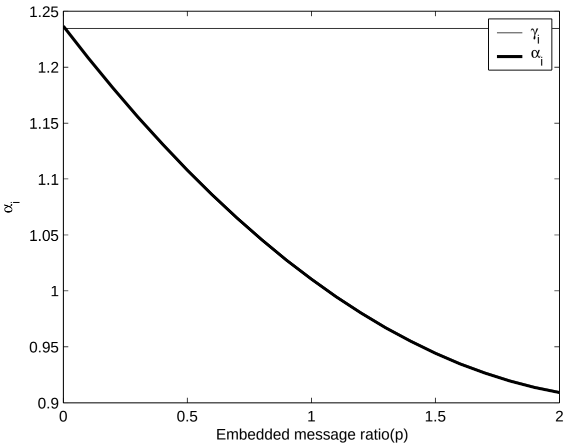

Figure 3: Proposed

Let $\alpha_i(0)$ be the initial value of α_i and $\gamma_i(0)$ be the initial value of γ_i (i.e. when embedding ratio is zero). The error ‘ε’be the initial-bias between the γ_i and α_i. So we have

$$\varepsilon = \gamma_i(0) - \alpha_i(0). \tag{12}$$

From the difference image histogram method, $\gamma_i = g_{2i}/g_{2i+1}$ where g is the difference image histogram after setting all bits in LSB plane to zero. The grayscale value of pixels in the image will be even as the LSB are set to zero. When the image is embedded with hidden message using the LSB insertion and then performing the operation of setting all LSB plane bits to zero for the stego-image will result in the values of g_{2i} and g_{2i+1} unmodified. Hence

$$\gamma_i(0) = \gamma_i \quad \forall\ p \tag{13}$$

so the value of error ε will become

$$\varepsilon = \gamma_i - \alpha_i(0). \tag{14}$$

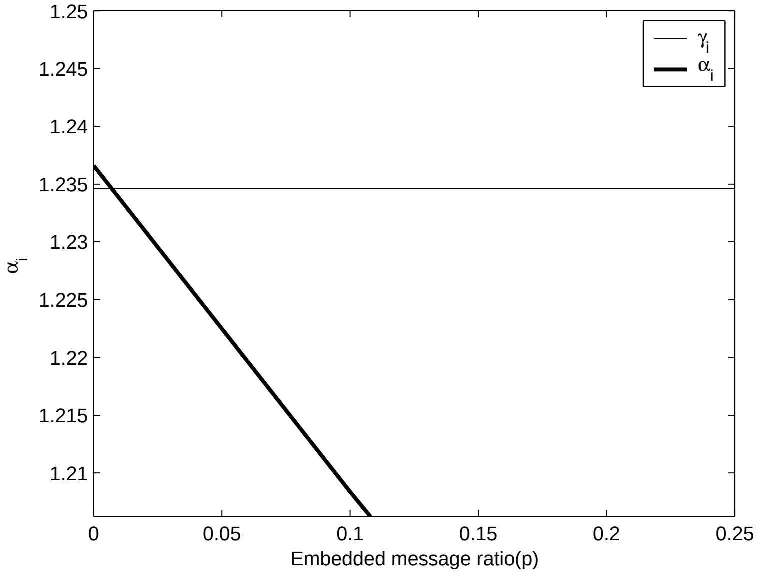

Figure 4: Proposed-(b)

The Difference image histogram models the relationship between α_i and the embedded ratio (p) using a quadratic equation $y = ax^2 + bx + c$. Considering the statistical hypotheses given in Equation (15) to be correct initially, we can find (p,α_i), (1,1), $(2-p,\beta_i)$ are the three points on the curve $y = ax^2 + bx + c$. Now we obtain the following equation set:

$$\begin{aligned} ap^2 + bp + c &= \alpha_i, \\ a + b + c &= 1, \\ a(2-p)^2 + b(2-p) + c &= \beta_i. \end{aligned} \tag{15}$$

Assume $e_1 = 1 - p$, $e_2 = 1 - \alpha_i$, $e_3 = 1 - \beta_i$ and then constant value “c”in equation 4.21 can be simplified to

$$c = \frac{2e_1^2 - 2e_1e_2 - (e_2 + e_3)(1 - e_1)}{2e_1^2}. \tag{16}$$

The value of ‘c”in Equation (15) will give the $\alpha_i(0)$ for an image. Hence subtracting the error from the estimated ratio p will remove the initial bias in the image. Hence the new estimated ratio “$p_{modified}$”will be

$$p_{modified} = p - \varepsilon \tag{17}$$

4 Description of IDIH algorithm

We now describe our detection algorithm

Input: A set of BMP images for detecting.

Output: The embedded ratio estimate $p_{modified}$ for each image.

Step 1. Select one image in the image set;

Step 2. Obtain difference image histogram of the image before (h_i) and after flipping the LSB bit planes to “zero”(g_i);

Step 3. Do from the step 4 to 8 for each value of i= 0,1,2;

Step 4. Calculate the statistical values for the image i.e $\alpha_i = (a_{2i+2,2i+1})/(a_{2i,2i+1})$, $\beta_i = (a_{2i+2,2i+3})/(a_{2i,2i-1})$

and $\gamma_i = g_{2i}/g_{2i+2}$, where the transition co-efficient can be estimated using the following equation

$$\begin{aligned} a_{0,1} = a_{0,-1} &= \frac{g_0 - h_0}{2g_0}, \\ a_{2i,2i} &= \frac{h_{2i}}{g_{2i}}, \\ a_{2i,2i-1} &= \frac{h_{2i-1} - a_{2i-2,2i-1}g_{2i-2}}{g_{2i}}, \\ a_{2i,2i+1} &= 1 - a_{2i,2i} - a_{2i,2i-1}. \end{aligned}$$

Step 5. Obtain the value of "p" from the root of the below equation whose absolute value is smaller.

$$2d_1p^2 + (d_3 - 4d_1 - d_2)p + 2d_2 = 0$$

where $d_1 = 1 - \gamma_i$, $d_2 = \alpha_i - \gamma_i$, $d_3 = \beta_i - \gamma_i$;
Step 6. Calculate the value $\alpha_i(0)$ which represents the estimation of α_i for zero embedded message length using the following equation

$$\alpha_i(0) = \frac{2e_1^2 - 2e_1e_2 - (e_2 + e_3)(1 - e_1)}{2e_1^2}$$

where $e_1 = 1 - p$, $e_2 = 1 - \alpha_i$ and $e_3 = 1 - \beta_i$;
Step 7. Calculate the initial bias 'ε"as

$$\varepsilon = \gamma_i - \alpha_i(0).$$

Step 8. Subtract the error 'ε'from the p to obtain the modified estimation ratio $p_{modified}(i)$.

$$p_{modified}(i) = p - \varepsilon$$

Step 9. The average of $p_{modified}(i)$ for i= 0,1,2 will give the final embedded ratio $p_{modified}$.

5 Experimental Results

We select 150 standard 512 × 512 test images (such as Lena, Peppers and so on). Applying random and sequential LSB replacement to embed the images with the ratio of p= 0, 10%, 20%,..., 90%,100% respectively with 10% increments we created two databases. Then we have use the RS method [7], DIH method [11] and GEFR method [12] to estimate the embedding ratio of secret information respectively. The mask used in the RS method is [1,0; 0,1].

The testing results of the test images got by DIH method and the proposed method (IDIH) are shown in Table 2. The leftmost column in Table 2 is the real embedding ratio, and column "IDIH", "DIH"represent the estimate embedding ratio got by Improve Difference Image Histogram method (proposed method) and Difference Image Histogram Method (DIH) respectively. It can be seen in Table 2 that the estimate precision of IDIH is higher than DIH obviously.

Figure [4] and [5] shows the corresponding plot of the embedded message length to the mean absolute error of the estimated values for random embedding and sequential embedding. Figure [4] indicates that the proposed algorithm (IDIH) algorithm outperforms the other three steganalysis techniques for embedded ratios greater than 40%. Improved Difference Image Histogram algorithm has performance comparable to the RS steganalysis for short messages (when p is smaller than 40%). However, because it is harder to detect smaller messages than large messages, the accuracy of the estimate is far more important for smaller message embedding. The proposed algorithm proves to be effective and reliable when complete range of embedding lengths is considered and compared to the existing algorithms.

Embedding	**Random**		**Sequential**	
ratio(%)	IDIH	DIH	IDIH	DIH
0%	0.3052	1.6855	0.3052	1.6855
10%	14.7804	15.3881	15.6703	16.0368
20%	20.38	20.80	27.98	28.11
30%	20.3764	20.8017	27.9818	28.1124
40%	40.1524	42.9062	44.3258	44.922
50%	48.6793	52.2864	49.7154	48.5228
60%	62.245	63.8	60.5394	56.5979
70%	72.7311	66.67118	69.7919	68.726
80%	84.6388	73.4632	80.8796	72.2582
90%	90.9915	85.8664	84.8516	81.955
100%	98.6088	92.5193	98.6088	92.5193

Table 2: Comparison between IDIH and DIH

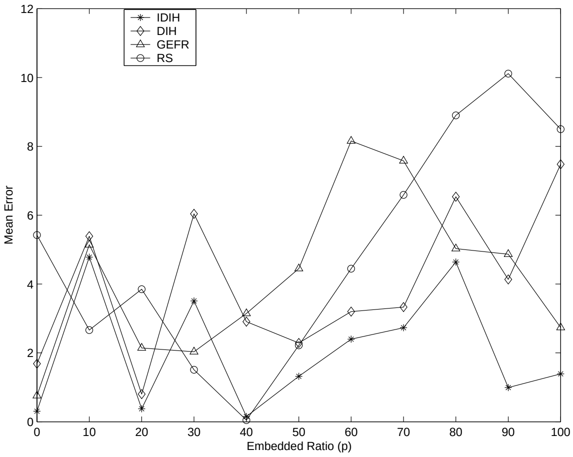

Figure 5: Comparison with other steganalytic techniques for random LSB embedding

In case of sequential embedding (as shown in Figure **??**), the accuracy is much higher than the case of random embedding for the embedded ratios of greater than 40%. It is having a higher performance to all the other steganalytic techniques for entire range of possible embedding lengths.

6 Summary and Conclusions

This paper proposes a new detection algorithm, which is an improved algorithm to the difference image histogram algorithm and performed tests on a group of raw lossless images. Experimental results show that the improved difference image histogram steganalysis method is more accurate and reliable than the conventional difference image histogram method. The proposed algorithm reduces the mean error by 50% for embedding ratios greater than 40% when compared to the DIH algorithm.

References

[1] F.A.P.Petitcolas, R.J.Anderson and M.G.Kuhn, “Information Hiding-A Survey”, *Proceedings of the IEEE*, vol.87(7), *special issue on protection of multimedia content*, June 1999, pp.1062-1018.

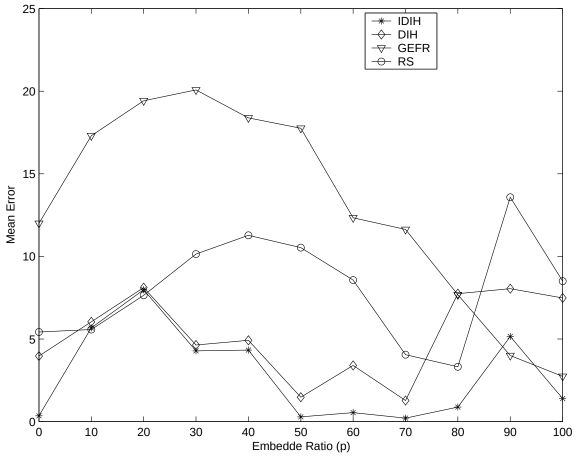

Figure 6: Comp. with other steganalytic techniques for sequential LSB embedding

[2] N.F.Johnson and S.Jajodia, “Steganalysis of images created using current steganography software”, *Lecture Notes in Computer Science*, vol.1525, Springer, Berlin, April 1998, pp.273-289.

[3] N.F.Johnson, and S.Jajodia, “Exploring steganography: Seeing the unseen ”, *IEEE, Computer*, February 1998, pp. 26-34.

[4] R.Chandramouli and K.P.Subbalakshmi, “Current trends in steganalysis: a critical survey ”, *International Control, Automation, Robotics and Vision Conference 2004*, Volume 2, December 2004, pp.964 - 967

[5] Neil F. Johnson and Sushil Jajodia, “Steganalysis: The Investigation of Hidden Information”, *IEEE Information Technology Conference*, September 1998, pp.113-116.

[6] J.Fridrich, M.Goljan and R.Du, “Detecting LSB Steganography in Color and Grayscale Images”, *IEEE*, vol.8(4) *Multimedia*, October-December 2001, pp.22-28.

[7] J. Fridrich, M. Goljan, R. Du, “Reliable detection of LSB Steganography in grayscale and color images”, *Proceeding of ACM, Special Session on Multimedia Security and Watermarking*, Ottawa, Canada, 2001, pp. 27-30.

[8] A.Westfeld, A. Pittzmann, “Attacks on steganographic systems”, *Information hiding, Third International Workshop, IH’99*, Dresden, Germany, 29 September-1 October, 1999.

[9] J.Fridrich, R.Du and M.Long, “Steganalysis of Lsb Encoding in Color Images’, *In Proceedings of ICME 2000*, July-August 2000.

[10] T.Zhang and X.Ping, “Reliable detection of LSB steganography based on the difference image histogram”, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume 3, April 2003, pp.545-548.

[11] Li Zhi and Sui Ai Fen, “Detection of random LSB image steganography”, *Vehicular Technology Conference IEEE* 60^{th} , Vol.3, September 2004, pp.2113 - 2117.

[12] A.Pinet, *SecurEngine professional v1.0*, 2004, http://securengine.isecu qrelabs.com.

[13] CBIR image database, University of washington, available at http://www.cs.washington.edu/research/imagedatabse/groundtruth/.

[14] The USC-SIPI image database, available at: http://sipi.usc.edu/ser vices/databse/Database.html.

Sanjay Kumar Jena
National Institute of Technology
Department ff Computer Science and Engineering
Rourkela, Orissa
India, 769008
E-mail: skjena@nitrkl.ac.in

G.V.V. Krishna
National Institute of Technology
Department ff Computer Science and Engineering
Rourkela, Orissa
India, 769008
E-mail: gvvkrishna@yahoo.co.in

Received: August 18, 2006

Dr. S.K. Jena was born in 28 April, 1954. He received his Ph.D. from Indian Institute of Technology, Bombay and M.Tech from Indian Institute of Technology, Kharagpur. He has joined National Institute of Technology as Professor in the Department of Computer Science and Engineering in 2002. Currently he is working as Professor and Head of Computer Science and Engineering department. He has more than 35 publications in International Journals and conferences. His research areas of interest are Database Engineering, Distributed Computing, Parallel algorithm, Information Security and Data Compression.

Mr. G.V.Vamsi.Krishna was born in 17 June, 1985. He received his M.Tech in Computer Science & Engineering from National Institute of Technology, Rourkela in 2006 and B.Tech in Computer Science & Engineering from Sarada Institute of Science, Technology and Management, Hyderabad in 2004. Currently he is working as Software Engineer in IBM India Pvt. Ltd., Bangalore.