SBHCS: Spike based Histogram Comparison Steganalysis Technique

Sonam Chhikara
Deenbandhu Chhotu Ram University
Murthal, Sonepat(Haryana), India

Parvinder Singh
Deenbandhu Chhotu Ram University
Murthal, Sonepat(Haryana), India

ABSTRACT

Steganography is the art of secretly transferring of data and steganalysis is the art of detecting that hidden data embedded in cover media. In the past years many powerful and robust methods of steganography and steganalysis have been reported in the literature. In this present work, a Steganalysis technique for Histogram-Shifting Based Data Hiding is designed to detect hidden data by using spike generation and template matching. The proposed work analyzes the characteristics of histogram changes during data hiding procedure, and then uses these features to distinguish between stego and original image.The presented work perform the steganalysis in four steps: First, an input image is filtered by using perwitt operator for edge detection. Second, the spike image is divided into 8x8 blocks and then histogram is generated for each block. Third, histogram of each block of stego-image and original image is compared by using 5 similarity measures (norm distance, cosine distance, Euclidean distance, Chi-squared distance, Entropy distance). Fourth, Neural Network (NN) is trained as a classifier to discriminate stego image from original image. Experimental results indicate that the proposed steganalysis method is better than the method proposed by Der-Chyuan Lou et. al.[1] and can effectively detect stego image at low bit rates.

General Terms

Steganalysis, Histogram, Pattern matching.

Keywords

Steganography; Steganalysis; Spikes; Neural Network (NN).

1. INTRODUCTION

Steganography is an art and science of hiding information over the digital media such that data can't be perceived by the user other than intended sender and receiver. Steganography has its greek origin and means conceal writing where "stega" means "covered" from greek word steganos and "nography" means "writing" from greek word graphia. The process of steganography starts by identifying the cover image and the information which is to hidden. This is an ancient art but digital technology give it new direction so that can be hide information in digital images and signals also.

Cryptography and steganography are the two important aspects of

communications. These are closely related topics. Cryptography is a process of converting the message in unreadable form so that it can't be understood by user other than authorized sender and receiver. While steganography hides the information into another information in such a way that presence of hidden message is not known by anyone. In Cryptography, message before encryption is called "plain text" and after encryption is called " cipher text" but in steganography, there is "cover media which result in "stego media" after hiding information into it. These two techniques can be combined for more secure and private communication as it will be more difficult for the steganalyst to find out the encrypted hidden message from the stego media instead of finding out plain message from it.But form the last few years steganography is used more by criminals for sending their important detail, So various steganalysis technique has been introduced for different steganography scheme from last few years.

Steganalysis is a technique of detecting secret messages hidden using steganography. The goal of steganalysis is to collect sufficient evidence about the presence of embedded message and to break the security of its carrier. Thus break the security provided by steganography. The importance of steganalytic techniques that can reliably detect the presence of hidden information in images is increasing. Steganalysis is classified into : Statistical Steganalysis and Signature Steganaysis.
When secret data hides in an image then the statistics of an image altered. Due to add of this secret information in the image, its pixel values change. This change in statistic of the image is used during analysis to detect the secret data.

Statistical steganalysis is further divided into specific and universal statistical steganalysis.

Specific statistical steganalysis include the statistical steganalysis techniques that target a particular steganography hiding technique or its expanded versions:

- ➢ LSB embedding.
- ➢ LSB matching.
- ➢ JPEG compression.
- ➢ Transformation domain, etc.

Universal statistical steganalysis include those statistical steganalysis techniques that are not used for a particular steganography hiding technique. The main idea behind these techniques is to find out some appropriate statistical quantities having 'distinguishing' capabilities. SVM, Neural network, clustering algorithms and other techniques are then used to design the detection model from the experimental data of these tools and techniques.

Signature Steganalysis: Steganography methods hide secret information in images or any other media in such a way that it remain imperceptible to human eye. But hiding information within any cover media using any steganography technique alters the media properties and introduces some form of distortion or some unusual patterns. These patterns and modified properties act as a signature that helps in detecting the existence of hidden message. Signatue Steganalysis is also further divide into Specific and Universal steganalysis.

In Specific Signature Steganalysis, Signatures specific to steganography tools are used to uncover the possibility of secret information which is hidden by that particular steganography tool. These specific signatures automatically discover the steganography tool which is used during message hiding process. For eg. Jpegx, a data hiding steganography tool, embeds the secret information at the end of JPEG files marker and adds a fixed signature of the program between the JPEG end marker and the secret message. The signature is the following hex code: 5B 3B 31 53 00. The appearance of this signature automatically implies that the image contains some secret data embedded using Jpegx data hiding tool. Other stego signature can be seen in [20].

Universal Signature Steganalysis is not related to particular steganographic tool, for eg : Images stored in JPEG format are not good choice for hiding information, as the quantization process during JPEG compression is unique fingerprint with the help of which very small modifications of the cover image can be detected.

Steganalysis finds its use in computer forensics, cyber warfare, tracking the criminal activities over the internet and gathering evidence for investigations particularly in case of anti-social elements . Apart from its law enforcement and anti-social significance steganalysis also has a peaceful application—improving the security of steganographic tools by evaluating and identifying their weaknesses.

Generally, Steganography change the various characteristics of image during data embedding that introduce distortion in original image and its histogram. In many field, it is required to restore the cover image after extracting hidden data from stego image, so various reversible data hiding techniques are introduce which can restore the cover image after extraction of secret data. Different approaches of reversible data hiding include difference expansion (DE) based methods [2], histogram-shifting based methods [3], vector quantization (VQ) based methods [4], and prediction-error based methods [5]. Among these, Histogram-Shifting Based method by Ni et al. [3] is most popular for its efficiency and simplicity.

Histogram Shifting data technique is considered as a variation of LSB. There are various LSB detectors like RS, SPA, DIH etc. are hard to detect that in active manner due to two limitations. Firstly, high detection error at low embedding rate; Secondly, unable to detect the exact embedding location. The experimental results in [3] indicate that the embedding capacity is only about 0.002 bpp for common natural images “Lena” and “Baboon” (both 512 ×512). So, the hidden messages are hard to be detected by existing LSB detectors at such low bit rates. The proposed work analyzes the histogram changing characteristics during data hiding procedure in Ni et al.'s method [3] and model these features by dividing image into 8x8 pixels block and compare the original or stego by similarity measures. A Neural Network is trained as a classifier for discriminating between cover images and stego-images based on these similarity measures. Experimental results indicate that the proposed steganalysis method can effectively detect stego image at low bit rates.

This paper is organized as follows. Section II outlined the related work in steganalysis filed till now. The details of the proposed steganalysis algorithm are presented in Section III. Experimental results and discussions are provided in Section IV, and the conclusion is given in Section V.

2. RELATED WORK

Working in data hiding in images require lot of secure methods so that data can be transferred secretly between two parties, but various steganalysis methods have been developed along steganographic techniques which breaks the security during secret data communication. These steganalysis techniques are developed so that security can be improved of various steganographic tools by evaluating and identifying their weaknesses. This section includes some steganalysis techniques which are designed for different stegnography schemes.

Tao Zhan et. al. [6] proposed Reliable Detection Of LSB Steganography based on the translation coefficients between difference image histogram. Translation coefficients are defined as a measure of the weak correlation between the least significant bit (LSB) plane and the remained bit planes, and then used to construct a classifier to discriminate the stego-image from the carrier image. This technique is used for detection of message as well as amount or information hidden in image. Its performance is better then RS analysis and good computation speed.

Miroslav Goljan et. al. [7] proposed an improved version of ablind steganalysis method proposed by Holotyak et al[8]. and compare it to current state-of-the-art blind steganalyzers. The features for the blind classifier are calculated in the wavelet domain as higher-order absolute moments of the noise residual. This technique use absolute non normalized moments of order 1 to 9. This method is called Wavelet Absolute Moment steganalysis (WAM). This steganalyzer is used for detection of embedding in raster image formats and compared its performance to four previously proposed blind steganalyzers, side information plays an important role in this steganalyzer.

Jing Dong et. al. [9] proposed a Blind Image Steganalysis Based On Run-Length Histogram Analysis. This paper focuses on extracting sensitive features to embedding modification , Statistical moments of characteristic functions of image run-length histogram and its variants are taken as features. SVM is utilized as classifier. The first three moments of the CF of three image RLHs are selected as features. This method has a better performance to an untrained stego-algorithm compared to others. Proposed 36-D feature vector provides clearly better detection accuracy compared with 78-D feature vector and the 108-D feature vector.

T. Pevny et. al. proposed [10] Steganalysis by Subtractive Pixel Adjacency Matrix (SPM). This is a novel approach to steganalysis of various embedding methods by utilizing the fact that the noise component of typical digital media exhibits short range dependences while the stego noise is an independent random component typically not found in digital media. The local dependences between differences of neighboring cover elements are modeled as a Markov chain, whose empirical probability transition matrix is taken as a feature vector for steganalysis. Although the SPAM features were primarily developed for blind steganalysis in the spatial domain, it is worth to investigate their potential to detect steganographic algorithms hiding in transform domains.

Hong Zhao et. al. [11] proposed a Steganalysis For Palette-Based Images. In order to capture the dependencies between adjacent colors, two efficient measurements are introduced. First, three generalized difference images between adjacent colors in horizontal, vertical and diagonal directions are constructed. And then, the first-, second- and third-order absolute moments of characteristic function of three generalized difference images' histograms are extracted. Second, a new feature called color correlogram that can distill the spatial correlation of colors is used to measure the dependencies of neighbor colors. For some algorithms, the classification accuracy is higher than 80% when the embedding rate is not less than 20%.

Zhongwei He et. al. [12] gives a novel approach based on approximate run length for image splicing detection. This scheme first defined approximate run length, which helps achieve higher performance. SVM is used to classify authentic and spliced images, which constructs features by applying the approximate run length on the original source image, its predict-error image, and DWT based reconstructed images. Compared with other methods, proposed approach can achieve a relatively high detection accuracy with far less computational complexity and much fewer features.

Next approach is introduced by Der Chyuan et. al [13] for Steganalysis of HMPD reversible data hiding scheme. HMPD reversible data hiding scheme involves the modification of pixel differences, which introduces artifacts into the pixel difference histograms. Four-way pixel difference features are used to design a specific steganalysis method for detecting HMPD reversible data hiding scheme. Two-class SVM classifier is used to distinguish stego-images from the cover images with an overall accuracy of 98.51%. The SVM classifier is trained with feature sets extracted from 1785 cover images and 10,710 stego-images with different hiding level (binary tree L= 0 to 5). Using a multiclass SVM classifier, an estimator which is capable of estimating the secret key (hiding level) with an accuracy of 99.77% is designed.

Next, J. Kodovsky et. al. [14] introduce Ensemble Classifiers for Steganalysis of Digital Media are built by fusing decisions of weak and unstable base learners implemented as the Fisher Linear Discriminant. The training complexity of the ensemble scales much more favorably allowing the steganalyst to work with high-dimensional feature spaces and large training sets, removing thus the limitations imposed by the available computing resources that have often curbed the detector design like SVM in the past. The ensemble is especially useful for fast feature development when attacking a new scheme. Performance wise, ensemble classifiers offer accuracy comparable and often even better to the much more complex SVMs at a fraction of the computational cost.

J. Kodovsky et. al. [15] further proposed a Rich Model for Steganalysis of Digital Images for building steganography detectors for digital images. The submodels consider various types of relationships among neighboring samples of noise residuals obtained by linear and non-linear filters with compact supports. The proposed framework demonstrated on three steganographic algorithms: HUGO, edge adaptive algorithm by Luo et al., and optimally coded ternary ±1 embedding. For these models G-SVM are trained for all three algorithms. The running time of a G-SVM classifier with 3,300-dimensional features. The proposed scheme is a step towards automatizing steganalysis to facilitate fast development of accurate detectors for new steganographic schemes. Rich models provide a good general-purpose model for various applications in forensics and in universal blind steganalysis.

Der-Chyuan Lou et. al. [1] recently introduce an Active steganalysis for histogram-shifting based reversible data hiding technique analyzes the characteristics of histogram changing during the data embedding procedure, and then models these features into reference templates by using a 1 ×4 sliding window and then use the combinatorial similarity measure to train the classifier. The proposed steganalysis algorithm is mainly composed of four parts: features extraction, classifier training, stego-images detection and embedding locations estimation. This algorithm is highly effective on stego-images detection.and embedding locations estimation at low bit rates.

3. PROPOSED WORK

In this section Spike Based Histogram Comparison Steganalysis technique (SBHCS) for histogram-shifting data hiding technique will be introduced. In this present work, steganalysis is done in four phases:

In first phase, high intensity area in the original and stego-image by using "prewitt operator" and convolution filters will be highlighted. These high intensity areas will help in finding the points in image where the secret data can be hidden .

In second phase, the image is divided into 8x8 blocks so that neighbor pixels of the spikes generated in the first phase can also be checked. Then the histogram of each block is generated, so that the change in histogram can be matched with the original image by using similarity measures in next phase.

Third phase, is for matching the histogram of stego-image's blocks with cover image's blocks. For matching 5 similarity measures: Norm distance, cosine distance, Euclidean distance, Chi-squared distance, Entropy distance are used. Objects are represented in a feature space as: χ=Rn for n dimensional data. A similarity measure s is defined as a function χ×χ=Rn that fulfills the following properties [16]:

1) Maximality: s(P, P)≥s(P, Q) for all P, Q∈χ.
2) Positivity: s(P, Q)≥0 for all P, Q∈χ;
3) Symmetry: s(P, Q) =s(Q, P) for all P, Q∈χ;

In this paper, two n-dimenional feature vector P={p1,p2…..,pn} and Q={q1,q2,...,qn} are used, the similarity between these two vectors is inversely proportional to distance.

S(P,Q) = 1 – d(P,Q) (1)

The following distance function are used to compete the similarity between these feature vectors [17], [18]

1) Norm distance:

$$d_1(\mathrm{P,Q}) = \sum_{i=1}^{n} |\,\mathrm{p}_i - \mathrm{q}_i\,|^2 \qquad (2)$$

2) Cosine distance:

$$d_2(\mathrm{P,Q}) = \sqrt{\sum_{i=1}^{n} |\,\mathrm{p}_i - \mathrm{q}_i\,|^2} \qquad (3)$$

3) Entropy distance:

$$d_3(\mathrm{P,Q}) = \frac{\sum_{i=1}^{n} PQ}{\sqrt{\sum_{i=1}^{n} P^2}\sqrt{\sum_{i=1}^{n} Q^2}} \qquad (4)$$

4) Norm distance:

$$d_4(\mathrm{P,Q}) = \sum_{i=1}^{n}\left[\frac{P\ln P + Q\ln Q}{2} - \left(\frac{P+Q}{2}\right)\ln\left(\frac{P+Q}{2}\right)\right] \qquad (5)$$

5) Chi-squared distance:

$$d5(\mathrm{P},\mathrm{Q}) = \sum_{i=1}^{n} \frac{\left(P-Q\right)^{2}}{\left(P+Q\right)^{2}} \qquad (6)$$

A lower value of S(P, Q) indicates less similarity between the feature objects. Similarity S(P,Q) can be calculated by combining above given distance measures:

$$S\left(P,Q\right) = \prod_{i=1}^{5}\left(1-d_{i}\left(P,Q\right)\right) \qquad (7)$$

In fourth phase, the above measures are used to train the neural network as a classifier to discriminate the stego-image from the cover image.

3.1 Algorithmic Steps For SBHCS:

Algorithm 1: Input stego-image on which spikes will be generated

```
• Adjust the brightness and contrast of image by normalization.
• Set the threshold minimum and maximum value.
• For each pixel, compare its value with threshold values
• IF (PXi>=thresholdmin and PXi<=thresholdmax)
  {Set PXi=White}
• ELSE
  {Set PXi=black}
• END IF
• FOR iter =1 to maxiteration
➢ Smoothen the image by applying Gaussian filter and convolution filter by setting convolution window of size sigma=2.5 .
➢ Identify the Sharp Edge Points over the Image
➢ Identify the ROI over the image ROI=findpixels(spikeareas>threshold)
➢ Sigma=Sigma-0.5
• END
• Setup the Statistical Measures over the images such as Standard Deviation.
• Remove the Unrequired Blocks over the Image by morphological operators and get the image with spike area.
```

Algorithm 2: Input spiked original image I and spiked test image I` by using algorithm 1

```
• For j=1 to Length(I)
➢ Extract the 8x8 blocks from I
  Img=ExtractImage(I, x-4, y-4, 8, 8)
➢ Extract the 8x8 blocks from I`
  Img1=ExtractImage(I, x-4, y-4, 8, 8)
➢ Generate histogram for 'img' and 'img1'
➢ Perform similarity match on these histograms by using above given similarity measures from equation 2 to 6.
  Similarty = Similaritymatch(Hist1,Hist2);
➢ Training=Training U Similarity
• END
• Implement Neural Network (Training)
```

3.2 Classifier Training

Steganalysis is a two class classification problem. A standard classification mainly includes two steps: training and testing. In training, the extracted feature vectors from training data are fed into a classifier to train a model. In second step, such a trained model is used to classify testing data into its belonging class.

Many classifiers such as artificial neural network (ANN), Fisher's linear discriminant (FLD), Bayesian discriminant analysis and SVM can be used to discriminate between cover-images and stegoimages [19]. The proposed scheme uses a Neural Network as a classifier. In training, each image's high intensity area will be find by edge detection then the histogram of 8x8 blocks of each image are compared by similarity measures. The extracted histograms of 8x8 blocks are used to train the model for NN training. The trained model will be used for classifying in the test phase.

4. SIMULATION & ANALYSIS

4.1 Simulation

To compare the SBHCS with the technique provided by Der-Chyuan Lou et. al.[1] and to check the effectiveness of the proposed method, 40 epochs are taken and set the desired goal at mean squared error of 1e-3. Table 1 shows different simulation parameters and their associated values.

TABLE 1: Simulation parameters for the steganalysis

Parameter Name	**Values**
Epoch	50
Threshold	20
Network model	NN(Neural Network)
Goal	1e-3
Gradient	Variable

.

4.2 Performance Evaluation

MATLAB is used as a simulation tool to analyse and evaluate the performance of the proposed technique. In simulation study, Neural Network is trained as a classifier. The stego image is passed in the classifier and NN set maximum 50 epochs for testing the image. The 8x8 block's histogram is compared with cover image during NN testing phase by using similarity measures given above from equation (2) to (6). The Mean Squared Error (mse) is set at 1e-3 to detect the hidden data. SBHCS technique achieves its desired goal of 1e-3 in 8 epochs.

Figure 1 shows one example of an image taken for testing this technique. This image is a stego image which has some secret message hidden inside it which is to be detected by this technique.

Fig 1: Example of a stego image.

Figure 2 shows the highlighted area which represents those pixels which may contain some hidden messages. These pixels are helpful for finding hidden message because every 8 neighbor pixles are also checked for secret message.

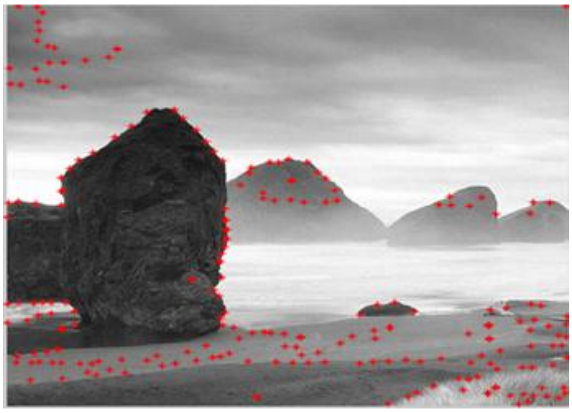

Fig 2 : Image with highlighting the high intensity areas.

The performance of the technique provide by Der-Chyuan Lou et. al.[1] and the SBHCS is compared over 8 epochs and shown in Figure 3. This shows that the SBHCS is better in performance then work given in [1].

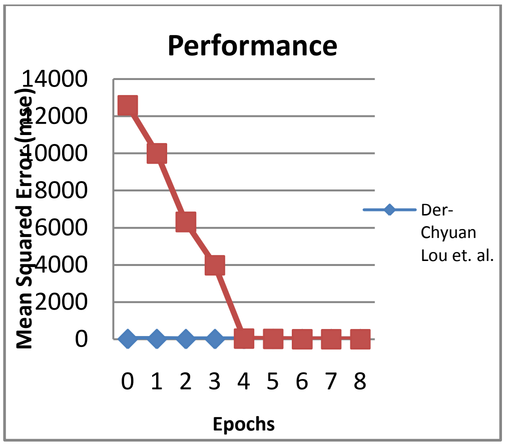

Fig 3: Comparison of Performance b/w SBHCS & [1].

Figure 4 shows the training state and comparison of Gradient between two techniques. The observation shows that proposed technique has less gradient than the previous technique which means that the proposed technique effectively detect the data.

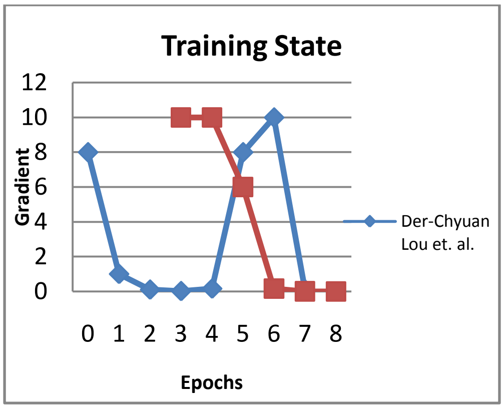

Fig 4: Gradient comparison b/w SBHCS & [1].

Figure 5 shows the comparison of mean value between these two techniques. It observed from the graph that the mean value of both techniques is almost same with a slight difference.

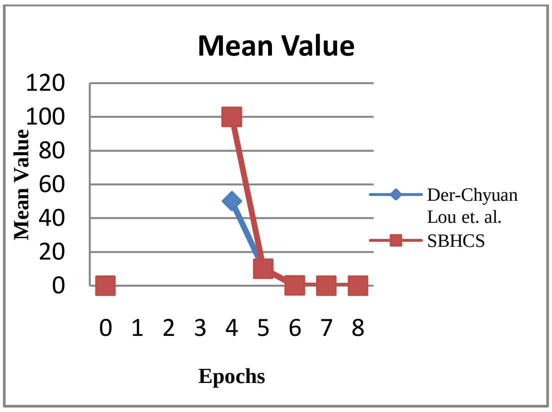

Fig 5: Mean Value v/s epochs for SBHCS & [1].

5. CONCLUSION

Steganalysis is an art of detecting secret data from a cover image hidden using steganography. In this paper, steganography and steganalysis are discussed for basic concepts these techniques. Additionally it include a review of various steganalysis techniques for different steganography methods. After the review the proposed steganalysis scheme Spike Based Histogram Comparison **Steganalysis** (SBHCS) for histogram-shifting based reversible data hiding methods. To accomplish this, first the edges or high intensity area in the image are detected then divide the whole image into 8x8 blocks after that histogram features of each block that originate in the embedding procedure are investigated, and compare the histogram of cover and stego image by similarity measures. Discrimination between cover images and stego-images was performed by the combinational similarity measures and the trained Neural Network (NN) classifier. Experimental results show that SBHCS is highly effective on stego-images detection and embedding locations estimation at low bit rates.

Fig 6: Examples of image set used in testing.

6. REFERENCES

[1] Der-Chyuan Lou et. al. “Active steganalysis for histogram-shifting based reversible data hiding”, Elsevier, Optics Communications 285 , 2012, pp.2510–2518.

[2] J. Tian, IEEE Transactions on Circuits and Systems for Video Technology 13 (8)(Aug. 2003) 890.

[3] Z. Ni, Y.-Q. Shi, N. Ansari, W. Su, IEEE Transactions on Circuits and Systems for Video Technology 16 (3) (March 2006) 354.

[4] C.-C. Chang, C.-Y. Lin, IEEE Transactions on Information Forensics and Security 1 (4) (Dec. 2006) 493.

[5] D.M. Thodi, J.J. Rodriguez, IEEE Transactions on Image Processing 16 (3) (March 2007) 1057.

[6] M. Goljan and R. Du, Reliable Detection of LSB Steganography in Grayscale and Color Images**,** *Proc. of the ACM Workshop on Multimedia and Security*, Ottawa, Canada, October 5(2001), pp. 27-30.

[7] M. Goljan and T. Holotyak, *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII*, vol. 6072, San Jose, CA, January 16-19-2006, pp. 1-13.

[8] T. Holotyak, J. Fridrich, S. Voloshynovskiy, “Blind Statistical Steganalysis of Additive Steganography Using Wavelet Higher Order Statistics,” 9th IFIP TC-6 TC-11 Conference on Communications and Multimedia Security, LNCS vol. 3677, Springer-Verlag, Berlin, 2005,pp. 273–274.

[9] Jing Dong , Tieniu Tan , “Blind Image Steganalysis Based

On Run-Length Histogram Analysis”, IEEE image processing conference, 2008 ,pp.2064-2067.oct,2008.

[10] T. Pevný and P. Bas, *IEEE Trans. on Info. Forensics and Security*, vol. 5(2),2010, pp. 215–224.

[11] Hong Zhao, Hongxia Wang, Muhammad Khurram Khan, “Steganalysis for palette-based images using generalized difference image and color correlogram”,Elsevier,Signal Processing ,vol.91 ,2011,pp. 2595–2605.

[12] Zhongwei He, Wei Sun, Wei Lu, Hongtao Lu, “Digital image splicing detection based on approximate run length” ,Elsevier, Pattern Recognition Letters ,vol.32, 2011, pp.1591–1597.

[13] Der-Chyuan Lou, Chen-Hao Hu, Chao-Lung Chou, Chung-Cheng Chiu “Steganalysis of HMPD reversible data hiding scheme”, Elsevier, Optics Communications ,vol.284 ,2011 , pp.5406–5414.

[14] J. Kodovský and V. Holub, *IEEE Trans. on Info. Forensics and Security*, vol. 7(2), 2012, pp. 432-44.

[15] J. Kodovský, *IEEE Trans. on Info. Forensics and Security*, vol. 7(3), 2012, pp. 868-882.

[16] M.-J. Lesot, M. Rifqi, H. Benhadda, International Journal of Knowledge Engineering and Soft Data Paradigms 1 (1) (2009) 63.

[17] S.-H. Cha, International Journal of Mathematical Models and Methods in Applied Sciences 1 (4) (2007) 300.

[18] D. Weken, M. Nachtegael, E. Kerre, Lecture Notes in Computer Science 2715 (2003) 396.

[19] X.-Y. Luo, D.-S. Wang, P. Wang, F.-L. Liu, Signal Processing 88 (9) (Sep. 2008) 2138.

[20] Tariq Al Hawi, Mahmoud Al Qutayari, Hassan Barada, Steganalysis attacks on stego images using stego-signatures and statistical image properties, in: TENCON 2004, Region 10 Conference, vol. 2, 2004, pp. 104–107.