Image Complexity and Feature Extraction for Steganalysis of LSB Matching Steganography

Qingzhong Liu[1], Andrew H. Sung[1,2], Jianyun Xu[3], Bernardete M. Ribeiro[4]

[1]Department of Computer Science
[2]Institute for Complex Additive Systems Analysis
New Mexico Tech
Socorro, NM 87801 USA
{liu, sung}@cs.nmt.edu

[3]Microsoft Corporation
One Microsoft Way
Redmond, WA 98052 USA
dennisxu@microsoft.com

[4]Department of Informatics Engineering
University of Coimbra
3030-290 Coimbra, Portugal
bribeiro@dei.uc.pt

Abstract

In this paper, we present a scheme for steganalysis of LSB matching steganography based on feature extraction and pattern recognition techniques. Shape parameter of Generalized Gaussian Distribution (GGD) in the wavelet domain is introduced to measure image complexity. Several statistical pattern recognition algorithms are applied to train and classify the feature sets. Comparison of our method and others indicates our method is highly competitive. It is highly efficient for color image steganalysis. It is also efficient for grayscale steganalysis in the low image complexity domain.

1. Introduction

Steganography is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows of the existence of the message. Steganalysis aims to expose the presence of hidden data. In the past, some steganographic embedding methods, such as LSB embedding, spread spectrum steganography, F5 algorithm, have been very successfully attacked [1, 2, 3]. While other embedding paradigms, such as LSB matching [4] are much more difficult to detect.

The literature does contain a few detector for LSB matching steganography. One of the first papers on detecting of embedding by noise adding is the paper by Harmsen and Pearlman [3], wherein the measure, histogram characteristic function center of mass (HCFCOM), is extracted and a Bayesian multivariate classifier is applied. In [15], Adjacency HCFCOM and Calibrated HCFCOM were presented to improve the probability of detection for LSB matching in grayscale images. Farid and Lyu described an approach to detecting hidden messages in images by using a wavelet-like decomposition to build high-order statistical models of natural images [5]. Fridrich et al. [6] proposed a maximum likelihood estimator for estimating the number of embedding changes for non-adaptive ±K embedding in images. Unfortunately, the ML estimator starts to fail to reliably estimate the message length once the variance of sample exceeds 9 [6]. In [7], correlation features are proposed for image steganalysis, which is efficient for detection of several steganography systems.

In this paper, we extend the features in [7] and present a method for steganalysis of LSB matching steganography based on correlation features and pattern recognition techniques. Additionally, we introduce the shape parameter of generalized gaussian distribution (GGD) of the wavelet domain to measure the image complexity, and study the relation between the performance of detection and image complexity.

2. Image Complexity & Statistical Property

Experiments show that a good PDF approximation for the marginal density of coefficients at a particular subband produced by various types of wavelet transforms may be achieved by adaptively varying two parameters of the GGD [8], which is defined as

$$p(x;\alpha,\beta)=\frac{\beta}{2\alpha\Gamma(1/\beta)}e^{-(|x|/\alpha)^{\beta}} \qquad (1)$$

Where $\Gamma(\bullet)$ is the Gamma function, $\Gamma(z)=\int_0^{\infty}e^{-t}t^{z-1}dt,\ z>0$.

Here α models the width of the PDF peak (standard deviation), while β is inversely proportional to the decreasing rate of the peak. Sometimes, α is referred to as the scale parameter while β is called the shape parameter. The GGD model contains the

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

Gaussian and Laplacian PDFs as special cases, using $\beta = 2$ and $\beta = 1$, respectively.

Generally, an image with high complexity has a high shape parameter to the GGD in the wavelet domain.

3. Feature Extraction

Considering LSB matching steganography mainly modifies the binary bits in the Least Significant Bit Plane (LSBP), we define C1, the correlation between LSBP and the second Least Significant Bit Plane (LSBP2), $M_1(1{:}m,\ 1{:}n)$ denotes the binary bits of the LSBP and $M_2(1{:}m,\ 1{:}n)$ denotes the binary bits of the LSBP2.

$$C1 = cor\ (M_1,\ M_2) \qquad (2)$$

The autocorrelation of LSBP, C(k, l), is defined as:

$$C(k,l) = cor\ (X_k,\ X_l) \qquad (3)$$

where,

$X_k = M_1(1:m-k, 1:n-l);\quad X_l = M_1(k+1:m, l+1:n).$

Different values are set to *k* and *l*, and C2 to C15 is:

C2 = *C(1, 0)*; C3 = *C(2, 0)*; C4 = *C(3, 0)*;
C5 = *C(4, 0)*; C6 = *C(0, 1)*; C7 = *C(0, 2)*;
C8 = *C(0, 3)*; C9 = *C(0, 4)*; C10 = *C(1, 1)*;
C11 = *C(2, 2)*; C12 = *C(3, 3)*; C13 = *C(4, 4)*;
C14 = *C(1, 2)*; C15 = *C(2, 1)*.

The histogram probability density, H, is denoted as $(\rho_0, \rho_1, \rho_2 \ldots \rho_{N-1})$. The histogram probability densities, H_e, H_o, H_{l1}, and H_{l2} are denoted as follows:

$H_e = (\rho_0, \rho_2, \rho_4 \ldots \rho_{N-2})$, $\quad H_o = (\rho_1, \rho_3, \rho_5 \ldots \rho_{N-1})$;
$H_{l1} = (\rho_0, \rho_1, \rho_2 \ldots \rho_{N-1-l})$, $\quad H_{l2} = (\rho_l, \rho_{l+1}, \rho_{l+2} \ldots \rho_{N-1})$.

The autocorrelation coefficients C16 and $C_H(l)$ are defined as follows:

$$C16 = cor(H_e, H_o) \qquad (4)$$

$$C_H(l) = cor\ (H_{l1},\ H_{l2}) \qquad (5)$$

Set *l = 1, 2, 3* and *4*, the features from C17 to C20 are defined as follows:

C17 = $C_H(1)$, C18 = $C_H(2)$,
C19 = $C_H(3)$, C20=$C_H(4)$.

Besides the features mentioned above, we consider the difference between test image and the denoised version. Firstly, the test image is decomposed by haar wavelet. Zero is set to the coefficients in HL, LH and HH subbands, whose absolute value are smaller than some threshold value, *t*. The image is reconstructed according to the inverse wavelet transform. The reconstructed image is treated as denoised image. The difference between test image and reconstructed version is E_t (*t* is the threshold value).

$$C_E(t;\ k,l) = cor\ (E_{t,k},\ E_{t,l}) \qquad (6)$$

where,

$E_{t,k} = E_t(1:m-k, 1:n-l);\quad E_{t,l} = E_t(k+1:m, l+1:n).$

Different values are set to t, k and l, features from C21 to C41 are defined as follows:

C21=C_E(1.5; 0,1); C22=C_E(1.5; 1,0);
C23=C_E(1.5; 1,1); C24=C_E(1.5; 0,2);
C25=C_E(1.5; 2,0); C26=C_E(1.5; 1,2);
C27=C_E(1.5; 2,1); C28=C_E(2; 0,1);
C29=C_E(2; 1,0); C30=C_E(2; 1,1);
C31=C_E(2; 0,2); C32=C_E(2; 2,0);
C33=C_E(2; 1,2); C34=C_E(2; 2,1);
C35=C_E(2.5; 0,1); C36=C_E(2.5; 1,0);
C37=C_E(2.5; 1,1); C38=C_E(2.5; 0,2);
C39=C_E(2.5; 2,0); C40=C_E(2.5; 1,2);
C41=C_E(2.5; 2,1).

When detecting RGB color images, the correlation features across the color channels are explored. Assuming M_{r1}, M_{g1}, M_{b1} are the matrices that stand for the least significant bit planes of red, blue and green channels, respectively, the correlation coefficients C_{rg}, $C_{rb,}$ and C_{gb} are defined as follows, where abs(·) is the function of absolute value.

$$C_{rg} = abs(cor(M_{r1},\ M_{g1})) \qquad (7)$$

$$C_{rb} = abs(cor(M_{r1},\ M_{b1})) \qquad (8)$$

$$C_{gb} = abs(cor(M_{g1},\ M_{b1})) \qquad (9)$$

Similar to (8), $E_{t,c}(c=r,\ g,\ b)$ is the difference between the color channel (r, g, and b) of the test image and reconstructed version. *t* is the threshold value. The correlation features are defined as follows.

$$C_{E_{rg}}(t) = cor(E_{t,r}, E_{t,g}) \qquad (10)$$

$$C_{E_{rb}}(t) = cor(E_{t,r}, E_{t,b}) \qquad (11)$$

$$C_{E_{gb}}(t) = cor(E_{t,g}, E_{t,b}) \qquad (12)$$

4. Experiments and Results

The original images in our experiments are 5000 TIFF raw format digital pictures from Olympus C740, taken in U.S.A, across spring to winter. The original images are 24-bit, 640×480 pixels, never compressed.

In steganalysis of color images, like the method proposed in [5], we cropped the original images into 256×256 pixels in order to get rid of the low complexity parts. The cropped images are categorized according to the image complexity. The image complexity for color images is defined as follows:

$$\beta = (\beta_r + \beta_g + \beta_b)/3 \qquad (13)$$

The variable, $\beta_c (c = r, g, b)$, is the shape parameter of the GGD of the HH subband coefficients, corresponding to red, green, and blue channel.

In steganalysis of grayscale images, the cropped color images are converted into grayscales which are covers. The image complexity for grayscale is measured by the shape parameter of the GGD of the HH subband coefficients.

4.1 Steganalysis of LSB matching

In steganalysis of color images, correlation feature set consists of the following feature elements: C1, C2, C6, C10, C14, C15, C16, C17, $C_E(t;\ \mathrm{k},l)$ ($t \in$ {2.5, 3}; $(k,\ l) \in$ {(0,1),(1,0),(1,1)}) defined in Section 3, corresponding to red, green, and blue channels, 14×3 = 42 features; $C_{E_{rg}}(t)$, $C_{E_{rb}}(t)$, $C_{E_{gb}}(t)$, (t ∈ {1, 1.5, 2}), 3×3 = 9 features; adding Crg, Crb, and Cgb, total 54 features. In addition to extract correlation features motioned above, the Histogram Characteristic Function Center of Mass (HCFCOM) features are extracted according to steganalysis of additive noise modelable information hiding [3]. Additionally, Farid and Lyu described an approach to detecting hidden messages in images by building High-Order Moment statistics in Multi-Scale decomposition domain (HOMMS) [5]. There are 3-dimension features for HCFCOM and 216-dimenstion features for HOMMS.

The experiments on steganalysis of LSB matching steganography in grayscale domain are the same to those for color images, except that the feature sets are different. Correlation feature set consists of the 41 features, C1 to C41, defined in section 3. In addition to compare HOMMS feature set, which consists of 72 features in grayscale domain [5], we extend HCFCOM feature set to the high order moments. HCFHOM stands for HCF center of mass High Order Moments; HCFHOM (*r*) denotes the *r*th order moment. In our experiments, the HCFHOM feature set consists of HCFCOM and HCFHOM(*r*) (*r* = 2, 3, and 4). Additionally, Ker [15] proposed two novel ways of applying the HCF: calibrating the output using a downsampled image and computing the adjacency histogram instead of the usual histogram. In [15], the best discriminators are Adjacency HCFCOM (A.HCFCOM) and Calibrated Adjacency HCFCOM (C.A.HCFCOM), which are compared with our method in steganalysis of grayscale LSB matching steganography.

Since different classifier has different classification performance on different feature set, we apply the following classifiers to each feature set. The classifiers are Fisher Linear Discriminate (FLD), optimization of the Parzen Classifier (ParzenC), Naive Bayes classifier (NBC), Support Vector Machines (SVM), Linear Bayes Normal Classifier (LDC), Quadratic Bayes Normal Classifier (QDC), Bayes Classifier (BC) based on maximal likelihood estimation of Gaussian mixture model, Levenberg-Marquardt trained feed-forward Neural net Classifier (LMNC), and Adaboost algorithm (Adaboost) [9, 10, 11, 12, 13, 14].

Several experiments are done on each feature set using every classifier. The average classification accuracy on each feature set is compared.

4.2 Classification performance

Fig. 1 shows the top two classification accuracy (average values) for each feature set and the corresponding classifiers in steganalysis of color LSB matching steganography. Fig. 1(a) is with the LSBP hiding ratio, 1. It indicates that both correlation features and HCFCOM features are highly efficient. Fig. 1(b) is with the LSBP hiding ratio, 0.5. It shows that correlation feature set is the best in the experiments. Fig. 1(a) and (b) also indicate that the classification performances decrease when the image complexity increases. Especially the performance on HOMMS features decreases obviously and the classification performance is not good when the GGD shape parameter of the HH subband wavelet coefficients is bigger than 1. Fig. 2 shows the top classification accuracy (average values) for each feature set in steganalysis of grayscale LSB matching steganography. Fig. 2 indicates that the probability of correct classification on correlation features is the best, and the probability of correct classification on HOMMS features is the lowest.

5. Conclusions and Future Work

In this paper, we present a scheme for steganalysis of LSB matching steganography based on correlation features and pattern recognition techniques. GGD shape parameter in the wavelet domain is introduced to measure the image complexity. In comparison with HCFCOM and HOMMS for color image steganalysis, and HCFHOM, HOMMS, A.HCFCOM, and C.A.HCFCOM for grayscale steganalysis, our method is highly competitive. It is highly efficient for color image steganalysis and efficient for grayscale steganalysis in the low image complexity domain.

One task in the future is to improve the detection performance on grayscale steganalysis in the high image complexity domain; another is to optimize the feature set. As steganalysis is a topic of increasing interest to the national and cyber security communities, integration of all effective and applicable steganalytic algorithms into a practical tool is usually anticipated.

Acknowledgements

Partial support for this research received from ICASA (Institute for Complex Additive Systems Analysis, a division of New Mexico Tech) and a DoD IASP Capacity Building grant is gratefully acknowledged.

References

[1] A. Ker, "Improved Detection of LSB Steganography in Grayscale Images", *Lecture Notes in Computer Science*, vol. 3200, Springer-Verlag, New York, pp. 97–115, 2005.

[2] J. Fridrich, M. Goljan, D. Hogea, "Steganalysis of JPEG Images: Breaking the F5 Algorithm", *Lecture Notes in Computer Science*, vol. 2578, Springer-Verlag, New York, pp. 310–323, 2002.

[3] J. Harmsen and W. Pearlman, "Steganalysis of Additive Noise Modelable Information Hiding", *Proceedings of SPIE*, vol. 5020, pp. 131–142 , 2003.

[4] T. Sharp, "An Implementation of Key-Based Digital Signal Steganography," *Lecture Notes in Computer Science*, vol. 2137, Springer-Verlag, New York, pp. 13–26, 2001.

[5] S. Lyu and H. Farid, "How Realistic is Photorealistic", *IEEE Trans. on Signal Processing*, 53(2), pp. 845-850, 2005.

[6] J. Fridrich, D. Soukal, M. Goljan, "Maximum Likelihood Estimation of Length of Secret Message Embedding using ±K Steganography in Spatial Domain", *Proc. Of SPIE,* vol. 5681, pp. 595–606, 2005.

[7] Q. Liu, A. Sung, B. Ribeiro, "Statistical Correlations and Machine Learning for Steganalysis", *Adaptive and Natural Computing Algorithms*, Springer-Wien NewYork, pp. 437–440, 2005.

[8] K. Sharifi and A. Leon-Garcia, "Estimation of Shape Parameter for Generalized Gaussian Distributions in Subband Decompositions of Video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 52–56, 1995.

[9] V. Vapnik, *Statistical Learning Theory,* John Wiley, 1998.

[10] M. Schlesinger, V. Hlavac, *Ten Lectures on Statistical and Structural Pattern Recognition*, Kluwer Academic Publishers, 2002.

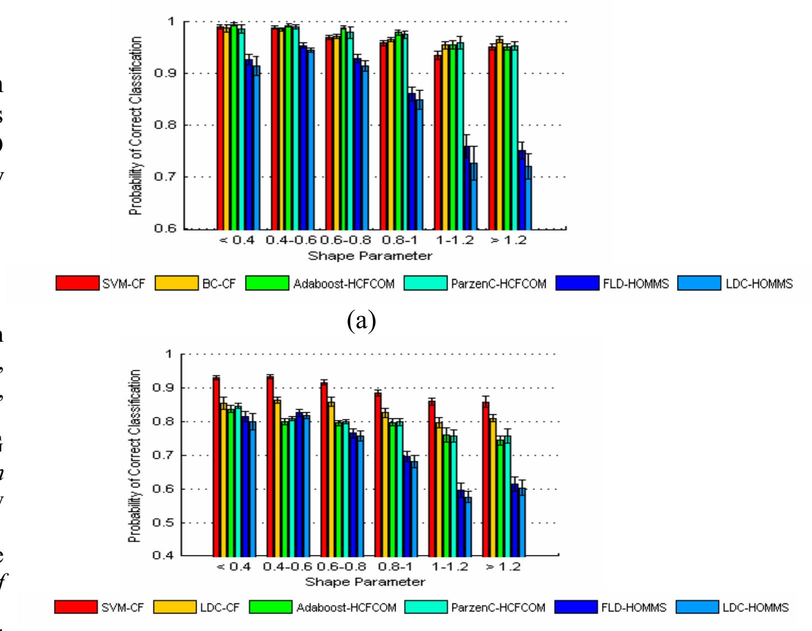
[11] F. Heijden, R. Duin, D. Ridder, D. Tax, *Classification, Parameter Estimation and State Estimation,* John Wiley, 2004.

[12] R. Duda, P. Hart, and D. Stork, *Pattern classification*, 2nd edition, John Wiley and Sons, New York, 2001.

[13] A. Webb, *Statistical Pattern Recognition*, John Wiley & Sons, New York, 2002.

[14] J. Friedman, T. Hastie, and R. Tibshirani. **"**Additive Logistic Regression: A Statistical View of Boosting**".** *The Annals of Statistics*, 38(2):337–374, 2000.

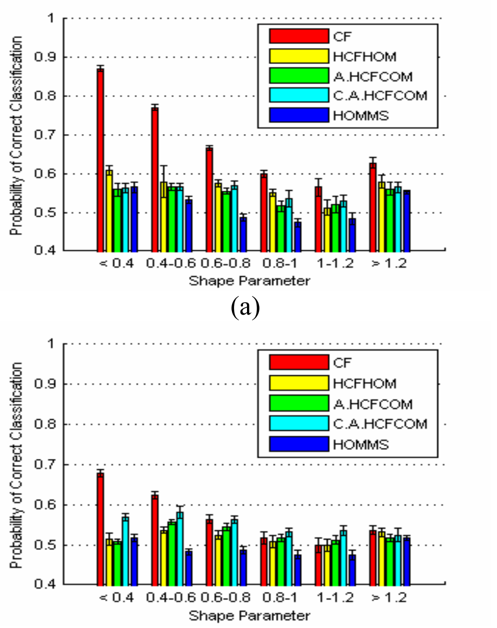
[15] A. Ker: "Steganalysis of LSB Matching in Grayscale Images", *IEEE Signal Processing Letters*, 12(6), pp. 441–444, 2005.



(a)

(b)

Fig. 1 Top two classification accuracies on each feature set. Fig. 1(a) is the test performance on the color steganography with the LSBP hiding ratio, 1. Fig. 1(b) is the test performance on the color steganography with the LSBP hiding ratio, 0.5.



(a)

(b)

Fig. 2 The best classification accuracy on each feature set. Fig.2(a) is the test performance on grayscale LSB matching steganography with the LSBP hiding ratio, 1. Fig. 2(b) is the test performance on the grayscale steganography with the LSBP hiding ratio, 0.5.