Breaking Predictive-Coding-Based Steganography and Modification for Enhanced Security

Guangjie Liu, Yuewei Dai, and Zhiquan Wang

Automation Department, Nanjing University of Science and Technology, Nanjing, 210094, China

Summary
The predictive-coding-based (PCB) steganography can embed a large amount of bits into the code stream of lossless compression with high imperceptibility. However, based on two elaborately chosen statistical features, the proposed steganalytic method can easily find the presence of a secret message with small error probability. To enhance the scheme's security, a modified one is proposed, which preserves the prediction errors' distribution by choosing the optimum adjustment parameter. Experimental results prove that the modified scheme can provide near-perfect security in Cachin's definition and defeat the steganalytic method proposed by ourselves.

Key words:
Steganalysis, Steganography, Information Security, Predictive Coding

1. Introduction

In recent years, with the development of the data communication in internet and wireless network, more and more information is frequently transmitted as digital forms including text, image, audio, video and other media. And information hiding and cryptography have become two significant topics of computer science due to the increasing demand of information security.

In cryptographic systems, message are protected by all kinds of encryption techniques such as DES or RSA[1]. The encrypted secure message, the ciphertext, is sent by the sender over the public insecure channel, the internet, GSM for examples. at the receiver end, the ciphertext is decrypted according to the corresponding key. the main disadvantage of the cryptographic system is that the ciphertext looks meaningless, and when the attacker can control the transmission of messages, he can interrupt the transmission or make more careful checks on the data from the sender to the receiver.

Different from the encryption, information hiding involves embedding secret data into other cover media with minimal perceivable degradation. Digital watermarking and steganography are two important branches of information hiding. Different from watermarking which aims to protect the copyright or content of multimedia, steganography is the art and science of hiding information such that its presence cannot be detected by attackers. The secret message is hidden inside a larger message, referred to the cover message, which can be transmitted without arousing any suspicion. The resulting received message containing the hidden secret content is referred to as the stego message. A number of techniques have been proposed for hiding message in digital media [2,3]. On the contrary, the main goal of attack on the steganographic systems, termed steganalysis, is to detect the presence of hidden data, and there are also many steganalytic techniques have been developed [2,3,4,5,6].

To ensure the security of the steganographic system, there are also some secure steganographic schemes having been proposed. One class of them is to design the hiding artifices to resist the known attacks. For example, Wu and Tsai [7] proposed a PVD steganographic scheme, which can resist the RS attack, Zhang and Wang [8] analyzed Wu's scheme and made modification to resist the histogram analysis, Westfeld[9] improved the conditional LSB algorithm and proposed a secure and high capacity scheme called F5, and Yu et al. [10] proposed a secure steganographic scheme against χ^2-analysis and RS attack. The other class aims to keep the distribution of stego data identical to that of stego data, and to achieve the security in Cachin's meaning [11], which means the relative entropy between the cover data and stego data is zero. For example, Eggers [12] proposed a steganography with the histogram preserving in JPEG coefficients, and Sallee [13] proposed a model-based steganographic scheme, Zhang[14] proposed a steganography for BMP images with least histogram abnormality.

The PCB steganography [15] proposed by Yu and Chang provides an efficient method to hide a large amount of secret bits into a still image by modifying the prediction errors. Due to the use of uniform quantization embedding rule, the abnormal prediction errors distribution caused by data hiding provides enough evidences to make steganalysis. To enhance security, a modified scheme is proposed. In Section 2, the PCB steganography is briefly reviewed. Section 3 describes the steganalytic method based on two statistical features. Section 4 presents the

Manuscript revised January 2006.

modification to the PCB method with experimental verification. And Section 5 concludes the whole paper.

2. Simple Review of PCB Steganography

In PCB steganography [15], the secret message bit is sequentially embedded into the prediction error value, which is the difference between the current pixel value and the current prediction pixel value in the image prediction stage. Then the modified values are entropy coded for removing the statistical redundancy of prediction error codes. Yu's method can be easily performed to embed and extract message bits.

Note that the predictor used in PCB is the modified MED predictor as Eq. (1) shows. The predictor is employed to estimate the prediction pixel values $\hat{x}$, and the prediction error values e is modified to e' to embed secret data.

$$\hat{x} = \begin{cases} \min(x_l, x_u) & if\ x_{lu} \geq \max(x_l, x_u) \\ \max(x_l, x_u) & if\ x_{lu} \leq \min(x_l, x_u) \\ (x_l + x_u)/2 \end{cases} \quad (1)$$

For embedding h message bits, b, the embedding rule can be described as:

$$e = x - \hat{x} \quad (2)$$

$$v = \begin{cases} 2^h \left\lfloor |e|/2^h \right\rfloor + b, & \left|(2^h \left\lfloor |e|/2^h \right\rfloor + b) - e\right| \leq \\ & \left|(2^h \left\lfloor |e|/2^h \right\rfloor + b) + e\right| \\ -2^h \left\lfloor |e|/2^h \right\rfloor - b, & \text{otherwise} \end{cases} \quad (3)$$

$$e' = \begin{cases} v & v + x \in [0, 255] \\ v - 2^h & v + \hat{x} > 255 \\ v + 2^h & v + \hat{x} < 0 \end{cases} \quad (4)$$

Here, h stands for the number of embedding bits per pixel in the host image. When h = 1, the embedding rule can be taken as LSB substitution. Eq. (3) is to determine whether the modified prediction error or it's opposite value is chosen to hide secret data. And Eq. (4) is to ensure the embedding pixel value does not exceed the range [0, 255]. The data extraction is just to read the h LSBs from the prediction errors of the stego image.

3. Steganalysis against PCB Embedding

In this paper, for simplicity, we just analyze PCB steganography with h equal to 1. To the case that h = 2, the steganalytic method is similar. Let the distribution of prediction errors of cover image and stego image be P_c and P_s respectively. Generally, in a natural image, the prediction errors' distribution is approximately Gaussian. The histogram of prediction errors distribution for a test image Lena with size 512×512 is shown in Fig. 1(a). Fig. 1(b) gives the histogram of prediction errors distribution of stego Lena.

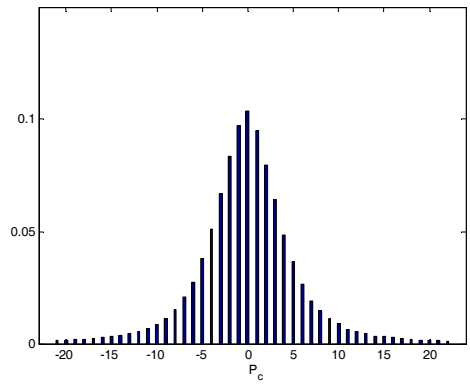
The effect of PCB embedding on the distribution is analyzed as follows. The secret data bits to be embedded can be viewed as a random bit stream because of the compression and encryption before embedding. According to Eq.s (2)-(4), when all image pixels are embedded, the relationship between P_c and P_s can be analyzed as

$$P_s[1] = \frac{1}{2}\left(P_c[0] + P_c[1]\right) \quad (5)$$

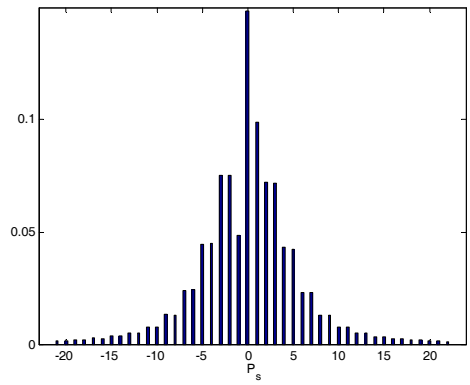
$$P_s[0] = \frac{1}{2}\left(P_c[0] + P_c[1] + P_c[-1]\right) \quad (6)$$

$$P_s[-1] = \frac{1}{2}P_c[-1] \quad (7)$$

$$P_s[2i] = P_s[2i+1] = \frac{1}{2}\left(P_c[2i] + P_c[2i+1]\right), \quad i = \pm 1, \pm 2, \ldots \quad (8)$$



(a)



(b)

Fig.1 the distribution of prediction errors of cover and stego Lena (a) prediction errors distribution of cover Lena, (b) prediction errors distribution of stego Lena

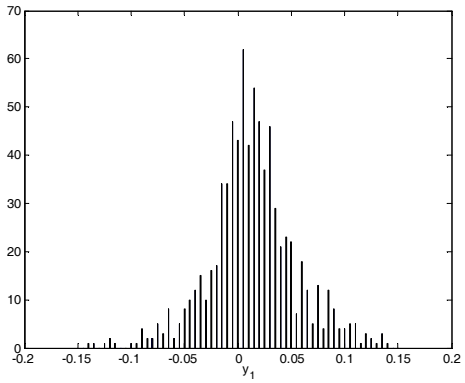
Seen from Fig. 2(b), the abnormal distribution reveals the presence of the secret message. To make the steganalytic detection, two statistical features are given as Eq. (9) shows to

$$y_1 = \frac{P[-1]-P[1]}{P[-1]+P[1]+c}$$
$$y_2 = \sum_{i=1}^{R}\left|\frac{P[2i]-P[2i+1]}{P[2i]+P[2i+1]+c}\right| + \sum_{i=-1}^{-R}\left|\frac{P[2i]-P[2i+1]}{P[2i]+P[2i+1]+c}\right| \quad (9)$$

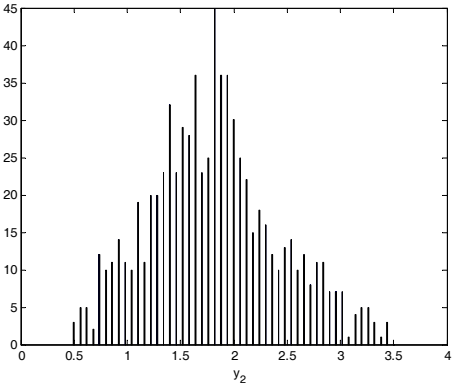
Here, [-*R*, *R*] is the statistical range of the prediction errors distribution, and *c* is a very small number to avoiding the nominator to be divided by zero. In this paper, *R* is chosen to be equal to 150 because the number is large enough to cover almost all of the prediction errors of natural images. According to the relationship between P_c and P_s as equations (5)-(8) show, the feature y_1 of a stego image should have a smaller negative value comparing with a natural image, while the feature y_2 should have a smaller positive value. Furthermore, y_1 indicates the symmetry of the distribution of "1" and "-1", and y_2 stands for the pair values phenomenon caused by semi-LSB embedding rules, which is like discussions in [16], so we can think the two stochastic variables are independent on each other here.

Assume the detection thresholds for y_1 and y_2 to be t_1 and t_2 respectively. For a given image *I*, when $y_1 < t_1$ and $y_2 < t_2$, it is considered that *I* contains secret message embedded by PCB steganography.

To avoid larger detection error probability, the two thresholds should be chosen carefully. Therefore, the experiment is performed to analyze the distribution of the two statistical features for natural images. 1000 gray images sized 480×480 are used in our experiment, and the distributions of y_1 and y_2 are depicted in Fig.2.



(a)



(b)

Fig. 2 The distribution of two statistical features. (a) The distribution of y_1, (b) The distribution of y_2

According to the central limit theorem, it is appropriate to assume that y_1 and y_2 obey Gaussian distribution. Based on the experimental results, we have

$$y_1 \sim N(\mu_1, \sigma_1^2), \mu_1 = 0.0033, \sigma_1^2 = 0.0018$$
$$y_2 \sim N(\mu_2, \sigma_2^2), \mu_2 = 1.9164, \sigma_2^2 = 0.4994 \quad (10)$$

Here, So the false alarm probability of the steganalytic detection with threshold being t_1 and t_2 is equal to

$$P_{FA} = \Phi\left(\frac{t_1-\mu_1}{\sigma_1}\right)\Phi\left(\frac{t_2-\mu_2}{\sigma_2}\right). \quad (11)$$

Here, Φ is the normal cumulative distribution function. Choosing t_1 = -0.2 and t_2 = 0.4, the false alarm probability is just equal to 1.32×10^{-8}.

Based on the detection method, four typical images are tested. Table 1 gives the values of y_1 and y_2 for these four images.

Table 1: The value of statistical features of four typical images

Statistical features	*Lena*	*Baboon*	*Jet*	*Peppers*
y_1	-0.3413	-0.3307	-0.3377	-0.3159
y_2	0.1897	0.2328	0.2795	0.2892

Seen from Table 1, the proposed steganalytic method can successfully detect the presence of the secret messages.

4. Modification to PCB Steganography

4.1 Cachin's Security Measure

It has been shown in the above that the key for steganalysis is the existence of abnormal distribution introduced by PCB steganography. And essentially, the proposed steganalytic method is the statistical analysis. It is true that all statistical steganalytic methods are based on the design of statistical tests that can be used to distinguish original cover data from the stego one.

In [11], Cachin proposed an information-theoretic model

that allows to quantify the security of steganography in terms of the decision error probabilities of hypothesis testing. As already mentioned by Cachin, because the adversary might exploit the information that is not used in the design of steganographic method, such a formal security notion has to be used carefully. When considering a fair game, we assume that both the sender and attacker exploit the same statistical features of the cover data. In our case, the statistical feature is the distribution of prediction errors. So it is reasonable to adopt Cachin's security measure for steganography.

The security measure of Cachin's security is based on the relative entropy (KLD, Kullback-Leibler Distance) [17] between the cover data and the stego data. And it is defined as

$$D(P_c \,\|\, P_s) = \sum_{x \in X} P_c[x] \log \frac{P_c[x]}{P_s[x]} \tag{12}$$

with the convention that $0 \log \frac{0}{P_c[x]} = 0$ and $P_c[x] \log \frac{P_c[x]}{0} = \infty$. Here, X is the support set of x. We note that the KLD is not symmetric. Hence, D $D(P_c[x] \,\|\, P_s[x])$ is different from $D(P_s[x] \,\|\, P_c[x])$.

For the prediction errors, X contains all possible values contained in [-255, 255]. And the value of relative entropy is always non-negative and is zero i.f.f. $P_c[x] = P_s[x]$. The steganographic system is called ε-secure if $D(P_c[x] \,\|\, P_s[x]) \le \varepsilon$, if $\varepsilon = 0$, the steganographic system is called perfectly secure.

Note that the Cachin's theory just give the security measure in statistical means. Despite that the embedding distortion is quite large, the steganographic system is considered perfectly secure if the relative entropy between cover and stego image is equal to zero. It also means that we can achieve statistical security by sacrificing some perceptibility quality and keeping the capacity unchanged.

4.2 The proposed scheme with prediction errors distribution preserved

To achieve the perfect security of the PCB system in Cachin's definition, measures have to be taken to preserve the distribution of prediction errors. For this purpose, a secure PCB steganographic scheme with the distribution preserved is proposed. The data embedding process is described in Fig.3.

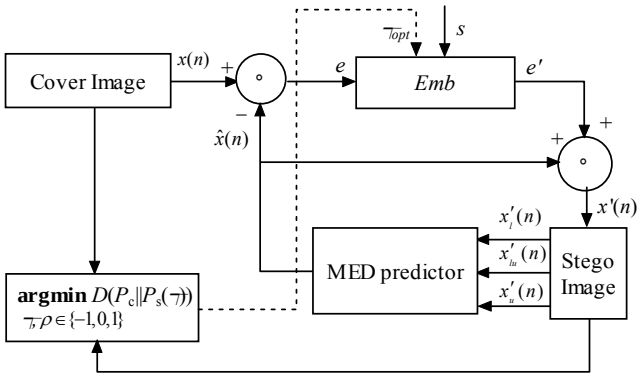


Fig. 3 The PCB steganographic scheme with the distribution preserved

Different from the embedding method used in Eqs. (2)-(4), after data hiding in the prediction error e, the modified prediction error v is adjusted by Eq. (13)

$$e' = Emb(e, b, \rho) = v + 2\rho$$
$$\rho \in \{-1, 0, 1\} \tag{13}$$

where, the parameter ρ is used to adjust the modified error v. From Eq. (13), we can see the adjustment does not change the LSB of v, so the secret data bit will not be destroyed. And the adjustment amplitude is quite small, so the effect on perceptibility quality from adjustment is not large.

For each $\rho \in \{-1, 0, 1\}$, the corresponding prediction errors' distribution of data embedded image is computed. The optimal adjustment parameter ρ_{opt} is chosen according to the following equation.

$$\rho_{opt} = \arg\min_{\rho, \rho \in \{-1,0,1\}} \left[D(P_c \,\|\, P_s(\rho)) \right] \tag{14}$$

Then the ultimate modified error, e', is determined by

$$e' = Emb(e, b, \rho_{opt}) \tag{15}$$

It should be noticed that the additional examination like Eq. (4) is also needed to keep the modified pixel value belonging to the range of [0, 255]. For example, if the prediction error e=15, and the secret bit is 0, after the calculation and comparison, the The optimal adjustment parameter ρ_{opt} is chosen to be -1, so the ultimate modified prediction error is equal to 12.

But we can see that in this secure hiding scheme, for a given prediction error, the relative entropy should be calculated three times, which will be very time consuming. but, just a simple technique can resolve this question. From every calculation of relative entropy, we just need computer the decrease or increase in some distribution component, and make some modification on the memorized relative entropy value.

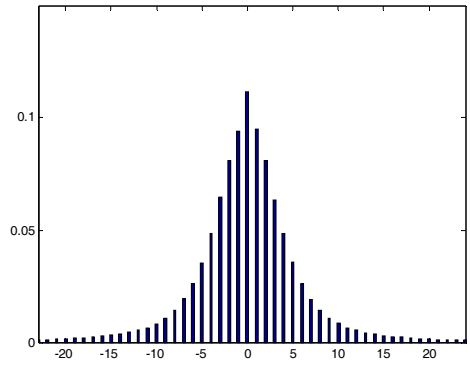
4.2 Experimental Results

To evaluate the performance of the modified scheme, four standard test images, "Lena", "Baboon", "Jet" and "Peppers", all with size equal to 512×512, are used in our experiments. The first experiment is taken to compare the relative entropy produced by the original method with modified method. Table 2 gives the comparison between the original and the modified method.

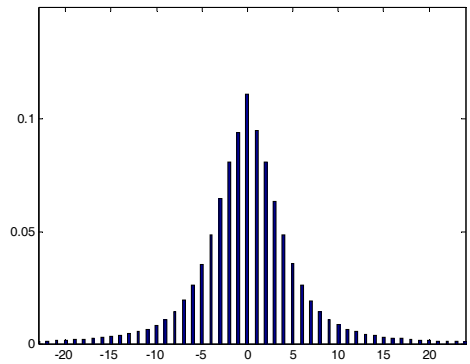
Table 2. The relative entropy of the original method and the modified method

Two methods	*Lena*	*Baboon*	*Jet*	*Peppers*
Original method	0.0403	0.0188	0.0534	0.0305
Modified method	4.5×10^{-5}	0.8×10^{-4}	4.8×10^{-5}	1.0×10^{-5}

From Table. 2, we can see the modified method preserves the prediction errors distribution commendably, while the original method causes larger distribution change. The prediction errors distribution histograms before and after embedding using the modified method, as shows in Fig. 4, also exhibit the excellent distribution preservation characteristics.



(a)



(b)

Fig.4 the distribution of prediction errors of cover and stego Lena (a) prediction errors distribution of cover Lena, (b) prediction errors distribution of stego Lena produced by the modified method

Then, the second experiment is performed to examine whether the addition of preservation step can help the original scheme to escape the statistical analysis proposed in Section 2. Table 3 gives the values of y_1 and y_2 for each of the four images. From Table 3, we can see that all computed features are larger and far away from the detection thresholds, it means the statistical analysis is invalid to our modified scheme.

Table 3. The value of statistical features of four typical images produced by the modified method

Statistical features	*Lena*	*Baboon*	*Jet*	*Peppers*
y_1	-0.0036	-0.0047	-0.0098	0.0104
y_2	2.0017	1.0574	2.1505	2.0026

4. Conclusion

Due to the use of uniform quantization embedding rule in PCB steganography, the prediction errors distribution of the stego image exhibits evidently artificial alteration. By two elaborately chosen statistical features, the secret message hidden by PCB steganography can be easily detected with very small error probability.

According to the statistical security theory, we introduce an additional optimal adjustment step into the system to preserve the prediction errors' distribution. The distribution preservation operation can help the system achieve near-perfect security in Cachin's meaning. Experimental results reveal the validity of the modified method. And in this way, the statistical analysis method proposed by us is also defeated.

Acknowledgment

This work was supported by the National Natural Science Foundation of China through the grant number 60374066, the Province Natural Science Foundation of China through the grant number BK2004132.

References

[1] S.A. Vanstone, A.J. Menezes and P. C. Oorschot, "Handbook of applied cryptography". CRC Press, 1996.

[2] P. Moulin and R. Koetter, "Data-Hiding Codes", Proceedings of the IEEE, vol. 93, no.12, pp. 2083-2126, 2005.

[3] S. Katzenbeisser and F.A.P. Petitcolas, "Information hiding techniques for steganography and digital watermarking", Artech House Books, 2000

[4] J. Fridrich, M Goljan, R. Du, "Detection of LSB steganography in color and grayscale images", IEEE Multimedia, vol. 8, no. 4, pp.22-38, 2001

[5] H. Farid, L. Siwei, "Detecting hidden messages using higher-order statistics and support vector machines", Proceedings of the 5th International Workshop on Information Hiding, Lecture

Notes in Computer Science, vol. 2578, pp.340-354, 2002
[6] I. Avcibas, N. Memon, B. Sankur, "Steganalysis using image quality metrics". IEEE Trans. on Image Processing, vol. 12, no. 2, pp.221-229,2003
[7] D.C. Wu, W. H. Tsai, "A steganography method for images by pixel-value differencing". Pattern Recognition Letters, vol. 24, no. 9/10, pp.1613-1626, 2003
[8] X.P. Zhang and S.Z. Wang, "Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security", Pattern Recognition Letters, vol. 25, no. 3, pp.331-339, 2004
[9] A. Westfeld, "F5— A steganography algorithm: high capacity despite better steganalysis", Proceedings of the 4th Information Hiding Workshop, Lecture Notes in Computer Science, vol.2137, pp.289-302, 2001
[10] J.J. Yu, J.W. Han, K.S. Lee, S.C. O, S. Lee, I.H. Park, "A secure steganographic scheme against statistical analysis", Proceedings of International Workshop on Digital Watermarking, Lecture Notes in Computer Science, vol. 2939, pp.497-507, 2004
[11] C. Cachin, "An information-theoretic model for steganography", Information and Computation, vol. 192, no. 1, pp.41-56 2004
[12] J.J. Eggers, R. Bauml, B. Girod, A communication approach to image steganography, Security and Watermarking of Multimedia Contents V of SPIE Proceedigns, vol. 4675, pp.26-37, 2002.
[13] P. Sallle, "Model-based steganography", Proceedings of International Workshop on Digital Watermarking, Lecture Notes in Computer Science, vol. 2939, pp.154-167,2004.
[14] X.P. Zhang, S.Z. Wang, K.W. Zhang, "Steganography with least histogram abnormality", Lecture Notes in Computer Science, vol.2776, pp.395-406, 2003.
[15] Y.H. Yu, C.C. Chang, Y.C. Hu, "Hiding secret data in images via predictive coding", Pattern recognition, vol. 38, no. 5, pp. 691-705, 2005
[16] A. Westfeld and A. Pfizman, "Attack on steganographic systems", Proceedings of the 3rd Information Hiding Workshop, Lecture Notes in Computer Science, vol.1768, pp.61-75, 2000
[17] T.M. Cover, J.A. Thomas, "Elements of information theory", John Wiley & Sons, 1991

Guangjie Liu received the B.E. and M.E. from Nanjing University of Science and Technology in 1998, 2002 respectively. He now is reading for Dr. Eng. degreee in Autoamtion Department of NUST. His research interest includes watermarking, steganography, steganalysis and multimedia encryption.

Yuewei Dai received the M. E., and Dr. Eng. degrees from Nanjing University of Science and Technology(NUST) in 1987 and 2002 respectively. He works in Automation Department in NUST till now from 1987, and now is the professor. His research interest includes automation control, information technology and information security including data encryption and information hiding. He is a director of System Engineering Academy of Jiangsu Province in China.

Zhiquan Wang received the B.E. degree from Harbin Industry Univeristy(HIT) in 1962. He has been working in NUST from 1962. and been a Professor at Automation Department of NUST since 1991. His research interest includes control theory, fault-tolerance control, fault detection and diagnosis in complex large system, Chaos control, information security. He is the director of Automation Academy in China and the deputy-president of Jiangsu Auotmation Academy.