Steganalysis Based on Differential Statistics

Zugen Liu, Lingdi Ping, Jian Chen, Jimin Wang, and Xuezeng Pan

College of Computer Science, Zhejiang University, Hangzhou 310027, China
lewissy2005@yahoo.com.cn

Abstract. Differential statistics were proposed in this paper to disclose the existence of hidden data in grayscale raw images. Meanwhile, differential statistics were utilized to improve the algorithm introduced by Fridrich to attack steganographic schemes in grayscale JPEG images. In raw images, to describe the correlation between data and their spatial positions, co-occurrence matrix based on intensities of adjacent pixels was adopted and the use of co-occurrence matrix was extended to high-order differentiations. The *COM*s (center of mass) of *HCF*s (histogram character function) were calculated from these statistics to form a 30-dimensional feature vector for steganalysis. For JPEG files, differential statistics were collected from boundaries of DCT blocks in their decompressed images. The *COM* of *HCF* was computed for each of these differential statistics and statistics from DCT domain so that a 28-dimensional feature vector can be extracted from a JPEG image. Two blindly steganalytic algorithms were constructed based on Support Vector Machine and the two kinds of feature vectors respectively. The presented methods demonstrate higher detecting rates with lower false positives than known schemes.

Keywords: differentiation, co-occurrence matrix, HCF, COM.

1 Introduction

Steganography and digital watermarking have recently emerged as a flourishing research area. Many steganographic software and watermarking algorithms can be downloaded freely from the Internet. People might utilize these tools to communicate secretly with each other. Effective steganalytic techniques are therefore hoped to emerge as means to prevent badmen from covert communicating. Steganalysis aims to expose the presence of hidden messages and, if possible, to extract the hidden messages. Steganalytic techniques can be classified into two categories: targeted and blind. Targeted method is designed to attack specific steganographic algorithm and blind one to conquer various steganographic schemes. Each image is a feature point in the multi-dimensional (M-D) feature space. Blindly steganalytic method has thus become a pattern classification in the M-D feature space.

By using *histogram character function* (HCF) and *center of mass* (COM), Harmsen et al.[1,2] proposed a method to attack cox[3] and piva[4] spread spectrum steganographies in raw images. However, it is found in our experiments that

D. Pointcheval, Y. Mu, and K. Chen (Eds.): CANS 2006, LNCS 4301, pp. 224–240, 2006.

the performance of method in [1] is not good enough for a lot of cover images since it adopts very limited number of features. To attack least significant bit (LSB) matching steganography in grayscale images, Ker[5] introduced two novel ways of applying the HCF: calibrating the output using a down-sampled image and computing the adjacency histogram instead of the usual histogram. In [6], applying HCF and moments of HCF to wavelet sub-bands of gray BMP images, Shi et al. designed a steganalytic method adopting 78-dimensional feature vectors. Compared with [1], its performance has been improved but still not high enough since there are too much relativity among its features calculated using statistical moments of wavelet characteristic function. Lyu et al.[7] proposed to use mean, variance, skewness and kurtosis of coefficients of wavelet sub-bands as features for a general steganalytic method. However, Shi et al.[6] showed the performance of their algorithm outperforms that of Lyu et al.'s method for spread spectrum steganographies in gray BMP images.

Fridrich et al.[8,9] attacked steganographic schemes in gray JPEG files effectively utilizing 21 features from DCT domain and 2 features from sum of differences of intensities distributed at boundaries of DCT blocks. However, Fridrich et al. considered DC coefficients when some features were collected from DCT domain and, the features from spatial domain can not reflect the statistical properties of changes to boundaries of DCT blocks brought by embedded data. Therefore, although Fridrich et al.'s method achieved very good detecting performances for Jsteg[10], F5[11], OutGuess[12] and MB1[13] steganographies, it can not attack MB2[13] and Steghide[14] well. Related researches include Bohme et al.'s work[15] and Lyu et al.'s works[7,16].

In this paper, we introduced differential statistics into steganalysis to attack cox[3] and piva[4] spread spectrum steganographies in raw images and, utilizing differential statistics to improve Fridrich et al.'s methods[8,9] to attack MB2 and Steghide in grayscale JPEG images.

In raw images, the differentiations were computed at pixel-locations in gray BMP image. Firstly, histograms were used to count the frequencies of high order differentiations. Secondly, co-occurrence matrices based on differentiations were calculated at adjacent locations. Finally, *COM* of *HCF* was utilized to calculate features for each of these statistics and, a 30-dimensional feature vector can be obtained from a grayscale raw image.

In JPEG files, firstly, the co-occurrence matrices were used in DCT domain to count the co-occurring state of two AC coefficients at the same locations in original image and its calibrated image. Secondly, in decompressed version of original JPEG image, histogram of intensities distributed at boundaries of DCT blocks were counted. Similarly, histograms of high order differentiations were computed. Finally, *COM*s of *HCF* were utilized to calculate 2 features from each co-occurrence matrix and 1 feature from each histogram. A 28 dimension feature vector can therefore be obtained through these three steps.

Based on the two presented statistical models and Support Vector Machine (SVM), two blindly steganalytic algorithms were constructed. The steganalytic algorithm in grayscale raw images, we call it "***DS***" (*Differential Statistics*)