A Markov Process Based Approach to Effective Attacking JPEG Steganography

Yun Q. Shi, Chunhua Chen, and Wen Chen

New Jersey Institute of Technology
Newark, NJ, USA 07102
{shi,cc86}@njit.edu

Abstract. In this paper, a novel steganalysis scheme is presented to effectively detect the advanced JPEG steganography. For this purpose, we first choose to work on JPEG 2-D arrays formed from the magnitudes of quantized block DCT coefficients. Difference JPEG 2-D arrays along horizontal, vertical, and diagonal directions are then used to enhance changes caused by JPEG steganography. Markov process is applied to modeling these difference JPEG 2-D arrays so as to utilize the second order statistics for steganalysis. In addition to the utilization of difference JPEG 2-D arrays, a thresholding technique is developed to greatly reduce the dimensionality of transition probability matrices, i.e., the dimensionality of feature vectors, thus making the computational complexity of the proposed scheme manageable. The experimental works are presented to demonstrate that the proposed scheme has outperformed the existing steganalyzers in attacking OutGuess, F5, and MB1.

1 Introduction

Internet has become an important communication channel since the 90's of the last century, through which emails, speeches, images, and videos are easily transmitted and shared. With image steganography, covert communication through the Internet can also be conducted.

Steganography is the art and science of *invisible* communication, which is to conceal the very existence of hidden messages. Images have many attributes making themselves suitable for steganography. For instance, images can convey large payloads. Some steganographic method can accomplish a steganographic proportion exceeding 13% of the image file size [1]. Due to the non-stationarity of images, image steganography is hard to attack. Especially, the frequent interchange of digital images nowadays makes image steganography very promising.

Recently, research in the field of JPEG (Joint Photographic Experts Group) steganography has become active as JPEG images are used popularly. Many steganographic techniques operating on JPEG images have been published and become publicly available. Most of the techniques in this category modify the LSB (least significant bit) of the JPEG coefficients, which are the outcomes of block-wise two-dimensional (2-D) discrete Cosine transform (DCT) followed by quantization using JPEG quantization table.

J. Camenisch et al. (Eds.): IH 2006, LNCS 4437, pp. 249–264, 2007.

In this paper we look at three modern and most advanced steganographic methods, i.e., OutGuess [2], F5 [1], and model-based steganography (MB) [3].

OutGuess constructs a universal steganographic framework, which embeds hidden data using the redundancy of cover images. For JPEG images, OutGuess preserves statistics of the JPEG coefficient histogram. Two measures are taken to reduce the change on cover images introduced by data embedding. Before embedding, OutGuess identifies the redundant JPEG coefficients which have least effect on the cover image and will be modified if necessary during data embedding. It also adjusts the untouched coefficients during the embedding procedure to preserve the original histogram of the JPEG coefficients after embedding.

F5 was developed from Jsteg, F3, and F4. F5 takes two main actions to increase the security against steganalysis attacks: straddling and matrix coding. Straddling scatters the message as uniformly as possible over the cover image to equalize the change density. With matrix embedding, F5 improves the embedding efficiency (the number of bits embedded per change of JPEG coefficients). Generally speaking, the smaller the embedding message size is, the larger the embedding efficiency of F5 is.

In general, the hidden data may be uncorrelated to the cover image, which is utilized by many steganalysis algorithms to attack the data hiding algorithms. MB embedding tries to make the embedded data correlated to the cover image. This is realized by splitting the cover image into two parts, modeling the parameter of the distribution of the second part given the first part, encoding the second part using the model and to-be-embedded message, and then combining the two parts to form the stego image. In embedding method MB1 ([3]), which operates on JPEG images, a modified generalized Cauchy distribution (MGCD) is used to model the JPEG mode histogram. The embedding procedure keeps the lower precision version of the JPEG mode histogram unchanged.

To attack steganography, some steganalysis schemes have been proposed. There are two categories, i.e., specific and universal steganalysis [4]. Specific steganalysis focuses on detecting some particular steganographic tool and has good performance on this steganographic tool if well designed. Universal steganalysis yet tries to steganalyze any steganographic tool, known or unknown in advance.

Farid proposed a universal steganalyzer based on image's high order statistics in [5]. Quadrature mirror filters are used to decompose the image into wavelet subbands and then the high order statistics are calculated for each high frequency subband. The second set of statistics is calculated for the errors in an optimal linear predictor of the coefficient magnitude. Both sets of statistical moments are used as features for steganalysis. It can achieve generally better detection rate than random guess for universal steganographic methods.

In [6], Shi et al presented a universal steganalysis system. The statistical moments of characteristic functions of the given image, its prediction-error image, and their discrete wavelet transform (DWT) subbands are selected as features. All of the low-low wavelet subbands are also used in their system. This steganalyzer can provide a better performance than [5] in general.