Review Article on JPEG Steganalysis

Anuradha Singhal¹, Tarun Chander²

¹(Department of Computer Science, Delhi University, India)
²(Department of Computer Science Kurukshetra University, India)

ABSTRACT: Steganalysis is used to detect hidden information in stego object with little or no knowledge about the steganography algorithm. Steganalysis aims to collect the ample evidence about the presence of embedded message. The importance of steganalytic techniques is increasing. Widely used in computer forensics, cyber warfare, tracking the criminal activities over the internet. In this paper we try to review JPEG Image Steganalysis Utilizing both Intrablock and Interblock features Correlations

Keywords – lossless, lossy compression, spatial domain, stego object, transition

I. INTRODUCTION
Steganography is an art of hiding data inside cover medium (text/audio/video) in such a way that the existence of any embedded information is undetectable as opposed to cryptography where the existence of secret communication is known but is non-comprehensible. The word steganography originally came from a Greek word which means “covered writing”. Since information being sent is totally under cover, it has an advantage over cryptography because it does not attract any public attention, and the data may be encrypted before being embedded in the cover medium. It is enhanced cryptography with an added benefit of undetectable communication. In digital media watermarking and steganography are comparable with difference first concerns with increasing integrity of parent message and later concentrates on concealments of message. Steganography can be used to exchange embedded information in a cover object via public communication channel, whereas watermarking is used for copyright protection, checking integrity and tracking legitimate use of a particular software or media file. U.S. officials and experts say steganography is the latest method of communication being used by Osama bin Laden and his associates to outfox law enforcement. Bin Laden, indicted in the bombing in 1998 of two U.S. embassies in East Africa, and others are hiding maps and photographs of terrorist targets and posting instructions for terrorist activities on sports chat rooms, pornographic bulletin boards and other Websites, U.S. and foreign officials say[1].

1.1 JPEG Compression
Image files are the most common cover medium used for steganography. With resolution in most cases higher than human perception, data can be hidden in the “noisy” bits or pixels of the image file. Because of the noise, a slight change in those bits is imperceptible to the human eye, although it might be detected using statistical methods (i.e., steganalysis). One of the most common and naive methods of embedding message bits is Least Significant Bits (LSB) replacement in spatial domain where the bits are encoded in the cover image by replacing the LSB of pixels [2]. JPEG is the most popular image format used over the Internet. The popularity of JPEG is due to its high compression ratio with good visual image quality. The file format defined by JPEG stores data in JFIF (JPEG File Interchange Format), which uses lossy compression along with Huffman entropy coding to encode blocks of pixels. Fig 1 shows the JPEG Compression procedure. First, the algorithm breaks the BMP image into blocks of 8 by 8 pixels (spatial Domain). Then, discrete cosine transformation (DCT) is performed on these blocks to convert these pixel values from spatial domain to frequency domain. These coefficients are then quantized using a quantization table which is stored as a part of the JPEG image. This quantization step is lossy since it rounds the coefficient values. In the next step, Huffman entropy coding is performed to compress these quantized 8 x 8 blocks.

Fig 1 JPEG Compression procedure
Image based Steganography exists both in Frequency and Spatial Domain. Since JPEG transforms spatial data into the frequency domain where it then employs lossy compression, embedding data in the spatial domain before JPEG compression is likely to introduce too much noise and result in too many errors during decoding of the embedded data when it is returned to the spatial domain. These would be hard to correct using error correction coding. Hence, it was thought that steganography would not be possible with JPEG images because of its lossy characteristics. However, JPEG encoding is divided into lossy and lossless stages. DCT transformation to the frequency domain and quantization stages is lossy, whereas entropy encoding of the quantized DCT coefficients (which will be called “JPEG coefficients” to distinguish them from the raw frequency domain coefficients) is lossless compression. Taking advantage of this, researchers have embedded data bits inside the JPEG coefficients before the entropy coding stage. The lossy and lossless stages in a JPEG encoding process are illustrated in Fig 1.

1.2 JPEG Steganography

A generic method to embed message in a JPEG image is to change the DCT coefficient values to reflect the message bits as illustrated in Fig 2. Once the required message bits have been embedded, the modified coefficients are compressed using entropy encoding to finally produce the JPEG stego image. By embedding information in JPEG coefficients, it is difficult to detect the presence of any hidden data since the changes are usually not visible to the human eye in the spatial domain.

Fig 2 JPEG Steganography process

1.3 JPEG STEGANALYSIS

Steganalysis aims to detect the presence of any hidden information in the stego media. Current steganalysis aims to focus more on detecting statistical anomalies in the stego images which are based on the features extracted from typical cover images without any modifications. Steganalysis can be classified into Targeted and Blind Steganalysis. Specific steganalysis, also sometimes known as targeted steganalysis, is designed to attack one particular type of steganography algorithm. The steganalyst is aware of the embedding methods and statistical trends of the stego image if embedded with a known algorithm. This attack method is very effective when tested on images with the known embedding techniques whereas it might fail considerably if the algorithm is unknown to the steganalyst. Blind steganalysis also known as universal steganalysis is the modern and more powerful approach to attack a stego media since this method does not depend on knowing any particular embedding technique. Here Training Data needs to be trained with Stego Algorithms.

II. MARKOV PROCESS (MP) BASED JPEG STEGANALYSIS SCHEME USING INTRA BLOCK CORRELATION

Shi et al. were the first to use Markov model to detect the presence of hidden data in a medium [3]. Their technique is based on modeling the JPEG coefficients as Markov process and extracting useful features from them using intra-block dependencies between the coefficients. Since, the surrounding pixels in a JPEG images are closely related to each other, this correlation can be used to detect if any changes have been made to the coefficients are not. The difference between absolute values of neighboring DCT coefficients is modeled as a Markov process. The quantized DCT coefficients, F(u,v), are arranged in the same way as the pixels in the image. The feature set is formed by calculating four difference matrices from the quantized JPEG 2D array along horizontal, vertical, major diagonal and minor diagonal directions. Four difference arrays are calculated along four directions: horizontal, vertical, diagonal, and minor diagonal, denoted Fh(u,v), Fv(u,v), Fd(u,v), and Fm(u,v), respectively.

\[
F_h(u,v) = F(u,v) - F(u,v+1) \\
F_v(u,v) = F(u,v) - F(u+1,v) \\
F_d(u,v) = F(u,v) - F(u+1,v+1) \\
F_m(u,v) = F(u+1,v) - F(u,v+1)
\]
Transition Probability Matrix is calculated as follows

\[
M_{1h}(i,j) = \sum_{u=0}^{h-1} \sum_{v=0}^{h-1} \left( F(u,v) = i \cdot F(u+1,v) = j \right)
\]

\[
M_{1v}(i,j) = \sum_{u=0}^{v-1} \sum_{v=0}^{v-1} \left( F(u,v) = i \cdot F(u,v+1) = j \right)
\]

\[
M_{2h}(i,j) = \sum_{u=0}^{h-1} \sum_{v=0}^{h-1} \left( F(u,v) = i \cdot F(u-1,v) = j \right)
\]

\[
M_{2v}(i,j) = \sum_{u=0}^{v-1} \sum_{v=0}^{v-1} \left( F(u,v) = i \cdot F(u,v-1) = j \right)
\]

In order to reduce the computational complexity, they used a threshold of [-4, +4], i.e., any coefficient outside the range were converted to -4 or +4 depending on the value. This range leads to a probability transition matrix of 9 x 9, which in turn will produce a total of 81 x 4 = 324 features using the four difference matrices. Fig 3 depicts the two stages in the conversion of a vertical difference matrix.

A threshold of [0, 2] was considered for simplicity and illustration of this process. Fig 4 Depicts Conversion of a Vertical difference matrix to transition probability matrix. Also State Transition graph is shown.

Markov process greatly improves the detection rate of algorithms Outguess[6], F5[5] and MB1[7](Model Based Steganography)

III. JPEG STEAGANALYSIS UTILIZING INTER BLOCK CORRELATION

Similar to the intra-block technique four difference matrices are calculated which results in four probability transition matrices across horizontal, vertical, major and minor diagonals [4]. Elements of TPM’s associated with the horizontal, vertical, main diagonal and minor diagonal difference mode 2-D arrays are given by as illustrated

For Restoring intra-block dependencies we keep track of all the horizontal and vertical transitions. The transitions are stored in bits. One coefficient change can lead to multiple dependency changes. Then a set of coefficient which would restore all
those dependencies is sought out as illustrated in Fig 5.

![Fig 5 Restoring Intra Block Dependencies](image)

### IV. CONCLUSION

To utilize the intra block correlation, difference JPEG 2-D arrays are obtained from the JPEG 2-D array. Threshold of these difference 2-D arrays, applying Markov process to model these threshold difference 2-D arrays, we collect all the entries of the transition probability matrices associated with those Markov processes as intra block features.

To utilize the inter block correlation, from each mode 2-D array; we obtain difference mode 2-D arrays. After threshold these difference mode 2-D arrays and applying Markov process to model these threshold difference mode 2-D arrays, we collect all the entries of those “averaged” transition probability matrices associated with those Markov processes as inter block features.

### REFERENCES


