ISSN- 2277-1956

Neighboring Joint Density and Markov Process Based Approach for JPEG Steganalysis

Arun R [1], Nithin Ravi S [2], Thiruppathi K [3]
[1 2 3] TIFAC CORE in Cyber Security, Amrita Vishwa Vidyapeetham
[1,2,3] IVth Semester M.Tech Cyber Security Student
[1] csarunr@gmail.com, [2] nithinravi535@gmail.com, [3] ktirupathi1988@gmail.com

Abstract- Steganalysis is the method used to detect the presence of any hidden message in a cover medium. A novel approach based on feature mining on the discrete cosine transform (DCT) domain, markov process based approach for modeling the difference JPEG 2-D arrays, machine learning for steganalysis of JPEG images which prevents cross validation is proposed. The neighboring joint density and absolute neighboring joint density features on both intra-block and inter-block are extracted from the DCT coefficient array. For markov process based approach, difference JPEG 2-D arrays along horizontal, vertical and diagonal directions are modeled using markov model. 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. After the feature space has been constructed, it uses SVM like binary classifier with cross validation for training and classification. The performance of the proposed method on different Steganographic systems named F5, Pixel Value Differencing, Model Based Steganography with and without deblocking, JPHS, Steghide etc are analyzed. Individually each feature and combined features classification accuracy is checked and concludes which provides better classification.

Keywords – Steganography, Steganalysis, Markov, DCT, PVD, MB1, MB2, F5, JPHS, Steghide

I. INTRODUCTION

Steganography is the art of passing information through apparently innocent files in a manner that the very existence of the message is unknown. Steganalysis is the art of discovering hidden data in cover objects. As in cryptanalysis, we assume that the steganographic method is publicly known with the exception of a secret key. The method is secure if the stego-images do not contain any detectable artifacts due to message embedding. In other words, the set of stego-images should have the same statistical properties as the set of cover-images [1]. If there exists an algorithm that can guess whether or not a given image contains a secret message with a success rate better than random guessing, the steganographic system is considered broken.

Steganalysis is the art and science to detect whether a given medium has hidden message in it. Also steganalysis can serve as an effective way to judge the security performance of steganographic techniques. Steganalysis can be mainly classified into two-Blind Steganalysis and Targeted Steganalysis [2]. Targeted Steganalysis are designed for a particular steganographic algorithm. Blind Steganalysis are schemes which are independent of any specific embedding technique are used to alleviate the deficiency of targeted analyzers by removing their dependency on the behavior of individual embedding techniques. To remove their dependency a set of distinguishing statistics that are sensitive to a wide variety of embedding operations are determined and collected. These statistics are taken from both cover and stego images and are used to train a classifier, which is latter used to distinguish between cover and stego images. Hence, the dependency on a specific embedder is removed using these statistics.

Universal steganalysis is composed of feature extraction and feature classification. In feature extraction, a set of distinguishing statistics are obtained from a data set of images. In feature classification the obtained distinguishing statistics from both stego and cover images are used to train a classifier and finally the trained classifier is used to classify an input image as either being a stego image which carrying a hidden data or a clear image. The above statistics are obtained by observing general image features that exhibit strong variation under embedding.

In [3] Fridrich proposed a universal steganalysis scheme specially designed for JPEG steganography. A set of 23 distinguishing features from the block discrete cosine transform(BDCT) domain and spatial domain is proposed. In [4] Shie et al. presented a new universal steganalysis secheme in which all the 324 features are calculated directly from the quantized DCT coefficients. The Markov process is applied to modeling the difference of JPEG 2D arrays along horizontal, vertical and diagonal directions so as to utilize the second order statistics for steganalysis. In [4] Fu et al. presented another universal JPEG steganalysis scheme totally based on quantized DCT coefficients which extracted 200 features. The Markov empirical transition matrices are used to exploit the magnitude correlations between BDCT coefficients in both intra and inter block. By extending the feature set in [3] and applying calibration to the Markov features a new JPEG steganalysis scheme is developed by Penvy et al. [5] with 274 features. In [6] Qingzhong et al. proposed a new approach based on feature mining on the discrete cosine transform (DCT) and

IJECSE
International Journal of Electronics and Computer Science Engineering
WWW.IJECSE.ORG

machine learning for steganalysis of JPEG images. The neighboring joint density features on both intra and inter block are extracted from the DCT coefficient array and the absolute coefficient array.

.In this paper we propose a new steganalysis scheme to attack some latest developed JPEG steganographic schemes. In this features from transition probability matrix of difference JPEG 2D array and neighboring joint probabilities of the DCT coefficients on intra and inter block are extracted. Here Markov random process is used to model the difference JPEG 2D array along different directions. Transition probability matrix can be used to characterize the Markov process. Neighboring joint density and absolute neighboring joint density of intra and inter block are used.

The paper is organized as follows. In the next section we explain the Markov process based features. In section 3 neighboring joint density features on both intra and inter block features are explained. In section 4 proposed method is explained followed by the experimental results of the proposed steganalysis method in section 5. And paper is concluded in section 6.

II. MARKOV FEATURE

Consider a JPEG image in which 8 x 8 block discrete cosine transform is applied. This 2D array consisting of all of the 8 x 8 block DCT coefficients which have been quantized with a JPEG quantization table and have not been zig-zag scanned, runlength coded and Huffman coded. Taking the absolute value for each DCT coefficients results in a JPEG 2D array. Then the difference JPEG 2D array is formed along horizontal, vertical and diagonal directions [7].

Let F(u,v) represents the JPEG 2D array of a given image where u ϵ [1,S_u], v ϵ [1,S_v] and S_u is the size of JPEG 2D array in horizontal direction and S_v in vertical direction. The difference arrays are generated by

$$F_h(u,v) = F(u,v) - F(u+1,v) \quad (1)$$
$$F_v(u,v) = F(u,v) - F(u,v+1) \quad (2)$$
$$F_d(u,v) = F(u,v) - F(u+1,v+1) \quad (3)$$
$$F_{md}(u,v) = F(u+1,v) - F(u,v+1) \quad (4)$$

where u ϵ [1,$S_u - 1$], v ϵ [1,$S_v - 1$] and $F_h(u,v)$, $F_v(u,v)$, $F_d(u,v)$, $F_{md}(u,v)$ denotes the difference array in the horizontal, vertical, main diagonal and minor diagonal directions respectively.

A.Transition Probability Matrix –

The difference JPEG 2D array is modeled using Markov random process. The one step transition probability matrix can be used to characterize the Markov process. For reducing complexity thresholding technique is applied. A threshold value T is used which results in a transition probability matrix of dimensionality (2T+1) x (2T+1). The elements of the four difference matrices associated with the horizontal, vertical, main diagonal and minor diagonal difference JPEG 2D arrays are given by

$$p\{F(u+1,v)=n|F(u,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m,F(u+1,v)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m)} \quad (5)$$

$$p\{F(u,v+1)=n|F(u,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m,F(u,v+1)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m)} \quad (6)$$

$$p\{F(u+1,v+1)=n|F(u,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m,F(u+1,v+1)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u,v)=m)} \quad (7)$$

$$p\{F(u,v+1)=n|F(u+1,v)=m\}=\frac{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u+1,v)=m,F(u,v+1)=n)}{\sum_{v=1}^{S_v-1}\sum_{u=1}^{S_u-1}\delta(F(u+1,v)=m)} \quad (8)$$

where m ϵ {-T,-T+1,….0,…T}, n ϵ {-T,-T+1,….0,…T} and

$$\delta(F(u,v)=m, F(u,v+1)=n) = \begin{cases} 1, & \text{if } F(u,v)=m \;\&\; F(u,v+1)=n \\ 0, & \text{otherwise} \end{cases}$$

So we have (2T+1) x (2T+1) for each of the four transition probability matrices and therefore 4 x (2T+1) x (2T+1) elements as feature vector for steganalysis. If we take the threshold as T=4 we have 9 x 9 elements in each of the four transition probability matrices and therefore total 4 x 81=324 elements. Here we take reduced Markov features as average of transition probability matrices in each direction, so 81 elements serve as features for steganalysis.

III. NEIGHBORING JOINT DENSITY FEATURES

The dependency between compressed DCT coefficients and their neighbours is explained in [5]. The information hiding will modify the neighboring joint density. When messages are embedded in the compressed DCT domain in JPEG images by any of the steganographic algorithms the DCT neighboring joint density probability density is affected which will gives a way for steganalysis. The modification of joint densities as a result of data embedding is shown in [6].

A.Feature Extraction –

The neighboring joint features are extracted on intra-block and inter-block from the DCT coefficient array respectively.

From the DCT coefficient array the neighboring joint density of intra block and inter block features are extracted as shown below. Let F denote the compressed DCT coefficient array of a JPEG image, consisting of M × N blocks (i = 1, 2.. M; j = 1, 2.. N). Each block has a size of 8 × 8. The intra-block neighboring joint density matrix on horizontal direction NJ_{1h} and the matrix on vertical direction NJ_{1v} are constructed as

$$NJ_{1h}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{8}\sum_{n=1}^{7}\delta(C_{ijmn}=x, C_{ijm(n+1)}=y)}{56MN} \quad (9)$$

$$NJ_{1v}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{7}\sum_{n=1}^{8}\delta(C_{ijmn}=x, C_{ij(m+1)n}=y)}{56MN} \quad (10)$$

where C_{ijmn} stands for the compressed DCT coefficient located at the m^{th} row and the n^{th} column in the block F_{ij} ; δ = 1 if its arguments are satisfied, otherwise, δ = 0; x and y are integers. For computational efficiency, we define NJ_1 as the neighboring joint density features on intra-block, calculated as follows:

$$NJ_1(x,y) = \frac{NJ_{1h}(x,y) + NJ_{1v}(x,y)}{2} \quad (11)$$

Here the values of x and y are in the range of [−6, +6], so NJ_1 has 169 features. Similarly the inter-block neighboring joint density matrix on horizontal direction NJ_{2h} and the matrix on vertical direction NJ_{2v} are constructed as follows:

$$NJ_{2h}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M}\sum_{j=1}^{N-1}\delta(C_{ijmn}=x, C_{i(j+1)mn}=y)}{64M(N-1)} \quad (12)$$

$$NJ_{2v}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M-1}\sum_{j=1}^{N}\delta(C_{ijmn}=x, C_{(i+1)jmn}=y)}{64(M-1)N} \quad (13)$$

We define NJ_2 as the neighboring joint density features on inter-block, calculated as follows:

$$NJ_2(x,y) = \frac{NJ_{2h}(x,y) + NJ_{2v}(x,y)}{2} \quad (14)$$

Similarly, the values of x and y are in the range of [−6, +6] and NJ_2 has 169 features

Hence we extract 169 features from both neighboring joint density of intra and inter block. So totally 338 features are extracted from neighboring joint density of inter and intra block DCT array.

From the absolute DCT coefficient array the neighboring joint density of intra block and inter block features are extracted as shown below. The intra-block neighboring joint density matrix on horizontal direction $absNJ_{1h}$ and the matrix on vertical direction $absNJ_{1v}$ are given by:

$$absNJ_{1h}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{8}\sum_{n=1}^{7}\delta(|C_{ijmn}|=x, |C_{ijm(n+1)}|=y)}{56MN} \quad (15)$$

$$absNJ_{1v}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{7}\sum_{n=1}^{8}\delta(|C_{ijmn}|=x, |C_{ij(m+1)n}|=y)}{56MN} \quad (16)$$

ISSN 2277-1956/V1N3-1020-1025

and the neighboring joint density features absNJ_1 on intra-block, calculated as follows:

$$absNJ_1(x,y) = \frac{absNJ_{1h}(x,y) + absNJ_{1v}(x,y)}{2} \quad (17)$$

Here the values of *x* and *y* are in the range of [0, 5], so absNJ_1 has 36 features. Similarly the inter-block neighboring joint density matrix on horizontal direction $absNJ_{2h}$ and the matrix on vertical direction $absNJ_{2v}$ are constructed as follows:

$$absNJ_{2h}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M}\sum_{j=1}^{N-1}\delta(|C_{ijmn}| = x, |C_{i(j+1)mn}| = y)}{64M(N-1)} \quad (18)$$

$$absNJ_{2v}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M-1}\sum_{j=1}^{N}\delta(|C_{ijmn}| = x, |C_{(i+1)jmn}| = y)}{64(M-1)N} \quad (19)$$

We define $absNJ_2$ as the neighboring joint density features on inter-block, calculated as follows:

$$absNJ_2(x,y) = \frac{absNJ_{2h}(x,y) + absNJ_{2v}(x,y)}{2} \quad (20)$$

Similarly, the values of *x* and *y* are in the range of [0, 5] and $absNJ_2$ has 36 features.
Hence we extract 169 features from both neighboring joint density of intra and inter block, 36 features from both absolute neighboring joint density of intra and inter block. So totally 410 features are extracted from neighboring joint density DCT array.

IV. Feature Based JPEG Steganalysis using markov process and Neighboring Joint Density Based Features

By combining the features obtained from the Markov process and the neighboring joint densities, a new feature based JPEG steganalysis scheme is proposed. From the Markov process 81 features and from neighboring joint densities 410 features are extracted and totally 491 distinguishable statistics are extracted for better steganalysis.

After the features are extracted from both stego and clear images it will be given to SVM like binary classifier for training. After the training is completed the features from test images are given for classification.

V. Experimental Results

.Five hundred and eighty five natural images were collected and these color images span a range of indoor and outdoor scenes and typically are 256 x 256 pixels in size. Another five hundred and eighty five stego images were generated by embedding messages of various sizes into the cover images. The payload corresponding to 100%, 75%, 50%, 25%, 20% and 10% of total cover capacity. The total cover capacity is defined to be the maximum size of a message that can be embedded by the embedding algorithm. Messages were embedded using F5, Model Based Steganography (MB1 and MB2), Pixel Value Differencing (PVD), JPHS and Steghide algorithms.

Individually each feature set is used for steganalysis and the combined one is also used. Markov features, Reduced Markov features, Neighboring joint density of intra block, Neighboring joint density of inter block, Absolute neighboring joint density of intra block, Absolute neighboring joint density of inter block, Combined Neighboring joint density of intra and inter block, Combined Absolute neighboring joint density of intra and inter block, Combined all neighboring joint density features, Combined Markov and neighboring joint density feature and finally Combined Reduced Markov and neighboring joint density features are extracted and used for steganalysis of all the above mentioned stego algorithms. Features will be extracted from each images yielding to a 324, 81, 169, 169, 36, 36, 338, 72, 410, 734 and 491 feature vector respectively. These features are used to train the linear SVM classifier separately. The performance of the classifier was tested using 250 test images which contain 25 cover and 25 stego images for F5, Model Based Steganography (MB1 and MB2 each), Pixel Value Differencing (PVD) respectively, 15 cover and 15 stego images for JPHS and 10 cover and 10 stego images for Steghide algorithm.

ISSN 2277-1956/V1N3-1020-1025

Table 1 to Tables 11 shows the classification accuracy of the feature based steganalysis by above mentioned features. All the individual features and the combined features are used for steganalysis. All the neighboring joint density and reduced markov process features will combined and used for feature based steaganalysis which will give better result when compared to other feature based steganalysis. Different payload can be embedded and used for steganalysis. While for lower payload also this feature based steganalysis gives better results than other features. Strong steganographic algorithms like steghide and JPHS will also gives better result in these features than other.

Table 1. Classification accuracy of Markov process features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	100
F5	25-50-75-100	90
MB1	25-50-75-100	80
MB2	25-50-75-100	60
JPHS	10-20	50
Steghide	10-20	50

Table 2. Classification accuracy of Reduced Markov process features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	92
MB1	25-50-75-100	72
MB2	25-50-75-100	56
JPHS	10-20	53.33
Steghide	10-20	65

Table 3. Classification accuracy of Neighboring joint density of intra block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	100
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	88
JPHS	10-20	70
Steghide	10-20	95

Table 4. Classification accuracy of Neighboring joint density of inter block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	96
F5	25-50-75-100	94
MB1	25-50-75-100	100
MB2	25-50-75-100	96
JPHS	10-20	63.33
Steghide	10-20	90

IJECSE
International Journal of Electronics and Computer Science Engineering
WWW.IJECSE.ORG

Table 5. Classification accuracy of Absolute neighboring joint density of intra block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	94
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	92
JPHS	10-20	76.66
Steghide	10-20	100

Table 6. Classification accuracy of Absolute neighboring joint density of inter block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	90
MB1	25-50-75-100	100
MB2	25-50-75-100	96
JPHS	10-20	70
Steghide	10-20	100

Table 7. Classification accuracy of Combined neighboring joint density of intra and inter block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	98
JPHS	10-20	73.33
Steghide	10-20	100

Table 8. Classification accuracy of Combined absolute neighboring joint density of intra and inter block features.

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	100
JPHS	10-20	66.66
Steghide	10-20	100

Table 9. Classification accuracy of Combined all neighboring joint density features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	100
MB1	25-50-75-100	100

ISSN 2277-1956/V1N3-1020-1025

MB2	**25-50-75-100**	**98**
JPHS	**10-20**	**70**
Steghide	**10-20**	**100**

Table 10. Classification accuracy of Combined markov and all neighboring joint density features

Algorithms	Payload	Classification Accuracy (%)
PVD	**25-50-75-100**	**100**
F5	**25-50-75-100**	**100**
MB1	**25-50-75-100**	**100**
MB2	**25-50-75-100**	**92**
JPHS	**10-20**	**63.33**
Steghide	**10-20**	**95**

Table 11. Classification accuracy of Combined reduced markov and all neighboring joint density features.

Algorithms	Payload	Classification Accuracy (%)
PVD	**25-50-75-100**	**100**
F5	**25-50-75-100**	**100**
MB1	**25-50-75-100**	**100**
MB2	**25-50-75-100**	**100**
JPHS	**10-20**	**73.33**
Steghide	**10-20**	**100**

IV.CONCLUSION AND FUTURE WORK

From the above experiments we concluded that the combination of all neighboring joint density and reduced markov features used steganalysis will gives better result when compared with other features. For strong steganographic algorithms like steghide which uses graph theoretical approach for embedding this feature based steganalysis performs better detection. The results of this paper demonstrate that, with judicious and sophisticated feature mining, it is possible to simultaneously achieve faster detection time, and higher detection performance for JPEG image steganography.

The future work is to do the feature selection by ranking the feature vector using some ranking algorithms and the optimum features has to be discovered out. These optimum features can reduce the miss classification. Feature selection can also be applied using projection pursuit algorithms to improve the detection efficiency. More embedding schemes can be used to analyse the features efficiency.

V. REFERENCE

[1] Gireesh Kumar T, Jithin R, Deepa D Shankar, (2010) " Feature Based Steganalysis using Wavelet Decomposition and Magnitude Statistics", *International Conference on Advances in Computer Engineering*.

[2] Deepa.D.Shankar, Gireeshkumar T, (2010) "Feature Based Classification System for Image Steganalysis", *International Conference on Computer Communications and Networks(CCN-10)*.

[3] J. Fridrich, (2004) " Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes", *Information Hiding, 6th International Workshop,LNCS 3200* ,pages 67-81.

[4] D.Fu, Y.Q.Shi and D.Zou, (2006) "JPEG steganalysis using empirical transition matrix in block DCT domain", *International Workshop on Multimedia Signal Processing* , Victoria, BC, Canada.

[5] T.Pevny and J.Fridrich, (2007) " Merging Markov and DCT features for multi-class JPEG steganalysis", *Proceedings of SPIE Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents IX*, volume 6505, pages 650503-1 to 650503-13

[6] Qingzhong Liu, Andrew H Sung, Mengyu Qiao, (2011) " Neighboring Joint Density-Based JPEG Steganalysis", *ACM Transactions on Intelligent Systems and Technology*. volume 2, No 2, Article 16.

[7] Yun.Q.Shi, C.Chen and W.Chen 2006, A Markov process based approach to effective attacking JPEG steganography. *Information Hiding, 8th International Workshop.*

ISSN 2277-1956/V1N3-1020-1025