Steganalysis using SPAM and Different Classifiers

Prasoon Jain, Ravi Sharma
IIIT Delhi

***Abstract*—This paper presents a method for detection of steganographic method in spatial domain usiing LSB matching using comparison of results between different classifiers. First, arguments are provided for modeling the differences between adjacent pixels using firstorder and second-order Markov chains. Subsets of sample transition probability matrices are then used as features for a steganalyzer implemented by Support Vector Machines, Fisher Discriminant Analysis, Linear Discriminant Analysis and Fisher Discrminant Analysis(L1 norm). A set of 6000 images from boss database is used form training and testing purposes. We compare the results of using different orders of markov chain and also the result of different classifiers using false positive rate and accuracy.**

I. Introduction

The effect of steganography in spatial domain is equivalent to adding a noise like signal in the cover. A popular method that does this is LSB replacement in which, the lsbs of individual cover elements is replaced by the message bit. The flipping operation in this makes the algorithm easily detectable because of its asymmetry. An addition to this algorithm is LSB matching which randomly increases or decreases the pixel value by (+-1) to match the lsb to message bit. LSB matching is much harder to detect because of this randomness . The heuristic behind embedding by noise adding is based on the fact that noise is superimposed on the cover as an iid signal making the message readable using some key for the detector. This same thing is also used to attack and detect whether the image is steganalyzed or not .
The method in this paper also use the independence of noise signal for steganalysis. We model using the difference between adjacent pixels and show that such deviations are due to steganographic embedding. The steganalyzer is constructed as follows. A filter suppressing the image content and exposing the stego noise is applied. Dependences between neighboring pixels of the filtered image are modeled as a higher order Markov chain. The transition probability matrix is then used as a vector feature for a feature-based steganalyzer implemented using machine learning algorithms.

II. SPAM

The dependences between pixels can be modeled by histograms of pairs, triplets or larger groups of neighbouring pixels. But there are several problems with this stated as follows :-

1) The dimensionality for histogram of pixel pairs for an 8 bit gray-scale image would be very large (256*256=65536).
2) Some histograms bins would be completely irrelevant to us because of there very low probability of occurrence.
3) It is difficult to model pixel pairs as they are influenced by cover also .

Therefore, we intend to work with differences rather than pixel pairs , thus reducing the dimension from 256*256(65536) to 256*2 -1 (511) . But it is still difficult to model using markov chain transition probability matrix as the transition probability matrix would have 511*511 elements which also to large. Further reduction can be achieved by realizing that, the statistical quantities estimated from pixels have to be estimable even from small images. Hence, only pixel pairs with a small difference ,in a range of [-T,T] are relevant for steganalysis.

Now, we explain calculation of spam features that will be used for steganalysis by different classifiers. The transition probabilities along eight directions are computed.The differences and the transition probability are always computed along the same direction.

The calculation starts with calculating the difference matrix . It is calculated by applying respective filter for a specific direction. For example ,

D(i,j)=I(i,j+1)-I(i,j)
D(i,j)=I(i,j)-I(i,j+1)

Now, first order spam features model the difference array using first order markov process for each direction .

M(u,v)= (Pr(D(i,j+1)=u s.t D(i,j)=v))
u,v belongs to [-T, T]

Second order spam features model the difference array using second order markov process for each direction .

M(u,v)= (Pr(D(i,j+2)=u s.t D(i,j+1)=v, D(i,j)=w))
u,v,w belongs to [-T, T]

We separately average the horizontal and vertical and then diagonal matrices to reduce dimensionality and form the two feature sets .

F1 = [Mhorizontal1+Mhorizontal2+Mvertical3+Mvertical4]/4
F2= [Mdiag1+Mdiag2+Mdiag3+Mdiag4]/4

These two are clubbed together to form a final feature set of dimension (2T+1)*(2T+1) for first order and

(2T+1)*(2T+1)*(2T+1) for second order. Now , different classifiers are applied and result is compared.The calculation of the difference array can be interpreted as high-pass filtering with the kernel [-1,1], which is, in fact, the simplest edge detector. The filtering suppresses the image content and exposes the stego noise, which results in a higher SNR.

III. Results

Using T=3 and same stego and non stego images in jpg format.

TABLE I
FIRST ORDER MARKOV SAME T=3

Classifier	Accuracy
SVM	85.2
FDA	91.1

TABLE II
SECOND ORDER MARKOV SAME T=3

Classifier	Accuracy
SVM	89.2
FDA	93.0

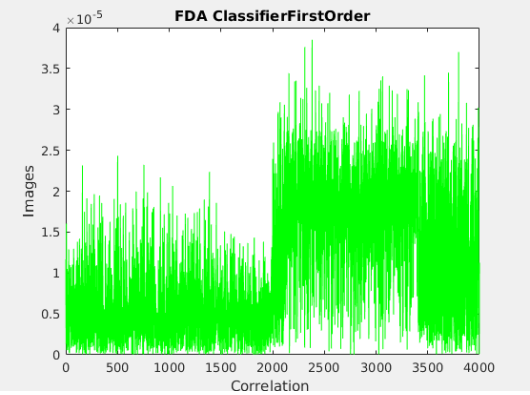

Fig. 1. FDA T=3 Correlation Stego images

Using T=5 and same stego and non stego images in jpg

TABLE III
FIRST ORDER MARKOV SAME T=5

Classifier	Accuracy
SVM	86.0
FDA	93.2

Using T=8 and same stego and non stego images in jpg format.

Using T=8 and different stego and non stego images in jpg format.

Using T=3 and different stego and non stego images in jpg format.

Using T=5 and different stego and non stego images in jpg format.

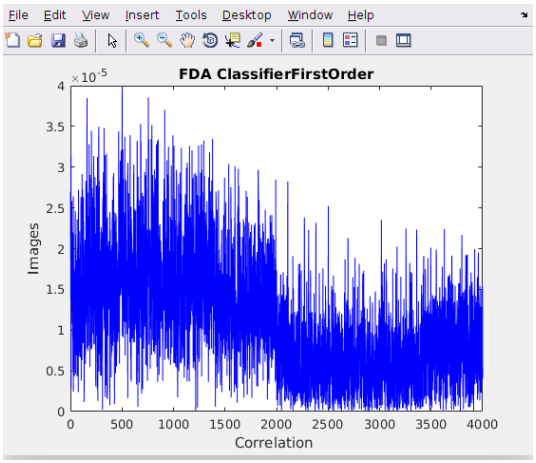

Fig. 2. FDA T=3 Correlation Nonstego

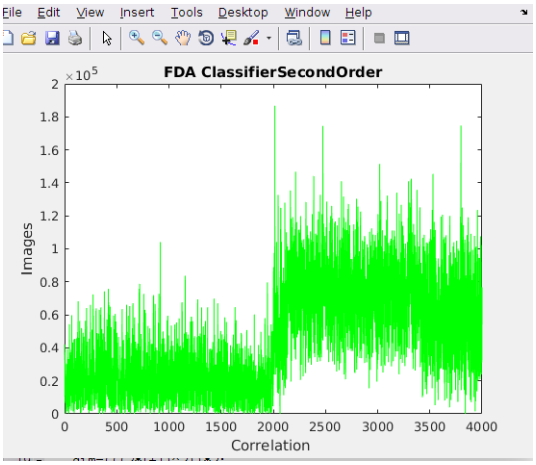

Fig. 3. FDA T=3 Correlation Stego images Second order

IV. Comparisons and Conclusions

It is observed that results for T=3,5 are better than T=8 . We conclude here that using T=8 , we are taking more values which are not relevant . Also, FDA is giving better results as compared to SVM but the computation time of FDA is more than svm. Also, accuracy is greatly influenced by the order of markov features. We believe that this phenomenon is due to the curse of the dimensionality, since first-order SPAM features with T=3 have dimension 98, while first-order SPAM features with T=8 have dimension 578. The contribution to the classification of additional features which may be not relevant for steganalysis is probably not very large and it is outweighted by the increased number of features.Using more T=8 in fda gives better result on different stegoand non stego images while using svm in T=8 gives less accuracy.

References

1) https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=5437325tag

2) http://ieeexplore.ieee.org/document/6750732/

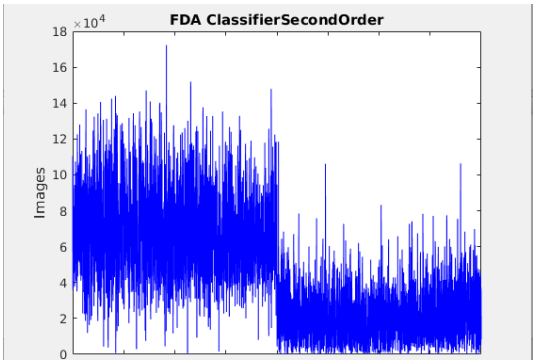

Fig. 4. FDA T=3 Correlation NonStego images Second order

TABLE IV
SECOND ORDER MARKOV SAME T=5

Classifier	Accuracy
SVM	87.9
FDA	93.5

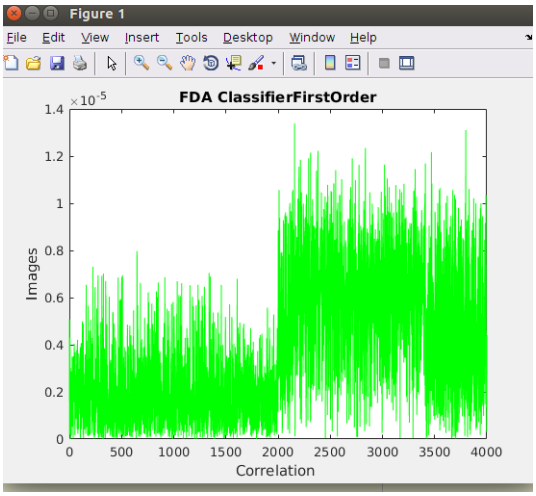

Fig. 5. FDA T=5 Correlation Stego images

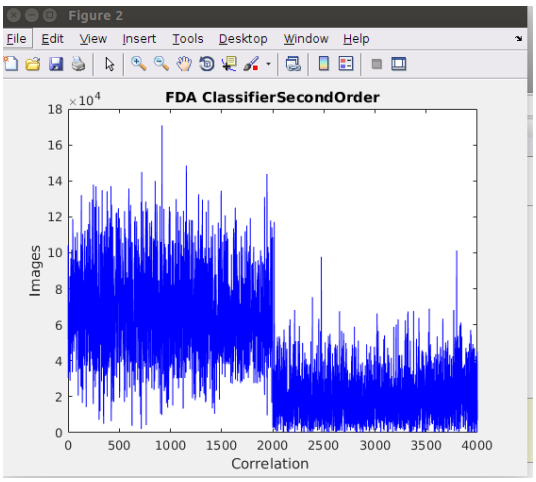

Fig. 6. FDA T=5 Correlation Nonstego

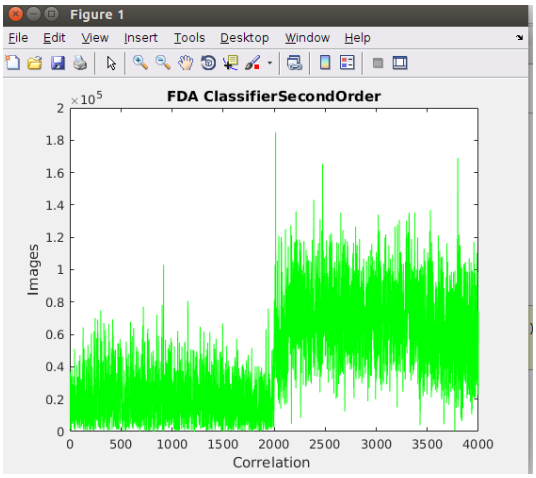

Fig. 7. FDA T=5 Correlation Stego images Second order

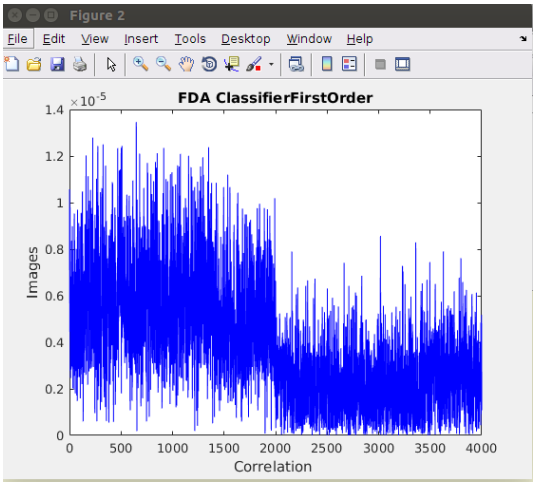

Fig. 8. FDA T=5 Correlation NonStego images Second order

TABLE V
FIRST ORDER MARKOV SAME T=8

Classifier	Accuracy
SVM	84.0
FDA	87.2

TABLE VI
SECOND ORDER MARKOV SAME T=8

Classifier	Accuracy
SVM	85.7
FDA	88.6

TABLE VII
FIRST ORDER MARKOV DIFF T=8

Classifier	Accuracy
SVM	78.6
FDA	97.0

TABLE VIII
SECOND ORDER MARKOV DIFF T=8

Classifier	Accuracy
SVM	85.7
FDA	99.0

TABLE IX
FIRST ORDER MARKOV DIFF T=3

Classifier	Accuracy
SVM	75.5
FDA	95.2

TABLE X
SECOND ORDER MARKOV DIFF T=3

Classifier	Accuracy
SVM	86.7
FDA	96.1

TABLE XI
FIRST ORDER MARKOV DIFF T=5

Classifier	Accuracy
SVM	80.6
FDA	96.9

TABLE XII
SECOND ORDER MARKOV DIFF T=5

Classifier	Accuracy
SVM	86.9
FDA	98.1