5-2006

Multiple Masks-based Pixel Comparison Steganalysis Method for Mobile Imaging

Sos S. Agaian
University of Texas at San Antonio

Gilbert L. Peterson
Air Force Institute of Technology

Benjamin M. Rodriguez
Air Force Institute of Technology

Recommended Citation

Sos S. Agaian, Gilbert L. Peterson, and Benjamin M. Rodriguez "Multiple masks based pixel comparison steganalysis method for mobile imaging", Proc. SPIE 6250, Mobile Multimedia/Image Processing for Military and Security Applications, 625005 (2 May 2006); https://doi.org/10.1117/12.666385

Multiple Masks Based Pixel Comparison Steganalysis Method for Mobile Imaging

Sos S. Agaian*[a], Gilbert L. Peterson[b] and Benjamin M. Rodriguez[b]
[a]Multimedia and Mobile Signal Processing Laboratory
The University of Texas at San Antonio, Department of Electrical and Computer Engineering
[b]Department of Electrical and Computer Engineering
Graduate School of Engineering and Management, Air Force Institute of Technology

ABSTRACT

Steganalysis has many challenges; which include the accurate and efficient detection of hidden content within digital images. This paper focuses on the development of a new multi pixel comparison method used for the detection of steganographic content within digital images transmitted over mobile channels. The sensitivity of detecting hidden information within a digital image can be increased or decreased to determine if slight changes have been made to the digital image for the target of blind steganalysis. The key thought of the presented method is to increase the sensitivity of features when alterations are made within the bit planes of a digital image. The differences between the new method and existing pixel comparison methods are; multiple masks of different sizes are used to increase the sensitivity and weighted features are used to improve the classification of the feature sets. Weights are also used with the various pixel comparisons to ensure proper sensitivity when detecting small changes. The article also investigates the reliability of detection and estimation length of hidden data within wireless digital images with potential for military applications emphasizing on defense and security.

Keywords: Steganography, Steganalysis, Feature Selection, Support Vector Machine Classifier, Mobile Imaging

1. INTRODUCTION

Several steganography methods exist for various forms of embedding secure data within digital images (stego images) which are shared between a sender and a receiver. Many embedding tools are available over the internet as freeware [1]. This makes the situation of detecting (steganalysis) the malicious information for law enforcement extremely difficult. As technology progresses new practical forms of communication are developed, each providing malicious users the opportunity to further exploit the transmission of digital data. In [2], Agaian, et al. show new techniques in sending secure communication over mobile devices using digital images.

The ability to apply steganography within mobile devices greatly increases the effectiveness of this particular type of covert communication. In addition, the new application has helped validate the art and aides in the establishment of steganography as a practical approach for transmitting sensitive material. The implementation of mobile steganography has demonstrated that the development and implementation of custom applications can be accomplished without the consent or knowledge of the manufacturer and service provider. While this has a significant impact on the transmission of classified data the same methods can be used for opposing reasons. With this said, this paper focuses on the development of a new multi pixel comparison steganography detection method used for digital images transmitted over mobile channels. The presented technique uses multiple masks to generate features and weights them to be sensitive enough to discriminate between clean and stego images.

The following section discusses related work in steganography and steganalysis. In Section 3, the novel feature selection methods used to determine if steganographic content has been embedded within the lower bit planes of the image are described. Section 4 describes the classification methods used to determine if the input images are stego images. In Section 5 we show experimental results and findings for two subsection 1) images developed with the Nokia

mobile device and 2) commonly found digital image taken with digital cameras. The conclusion is discussed in the final section of the paper.

2. RELATED WORK

Johnson, et al. have given an investigation into steganography and introduced some characteristics of steganographic software that identify signs of information hiding [1]. During that time detection techniques applied to steganography had not been devised beyond visual analysis. While this was one of the first insights into steganography, technology has allowed the growth and further development of this versatile science. Agaian, et al. have shown this to be true, in [2] they introduce a new application for steganography, mobile devices. While the ability to apply steganography to mobile devices has greatly increases the effectiveness of covert communication and has established steganography as a practical approach for transmitting sensitive material, this poses a problem when used for unethical purposes.

Many methods are used in conventional detection of covered images stored on PCs, networks, or transmitted over the Internet. Similar techniques can be created to determine if hidden information is embedded within mobile images. However, there are limitations and constraints that limit the use of previously developed techniques.

Statistical pixel comparison methods such as Modified Pixel Comparison by Agaian, et al. [3], Sample Pair Analysis by Dumitrescu, et al. [4], LSB Steganography by Fridrich, et al. [5] and Westfeld and Pfitzmann [6], use a form of pixel comparison, statistical measure and a comparison of altering the LSB of the input image to determine if steganographic content exists within the images. The methods have difficulty detecting hideen information within mobile images when modifications are made to the image. The problems arise when trying to make a comparison by altering the LSB since the LSB of a mobile image is predicted by the system upon receiving the image, i.e. the LSB is not used in mobile image transmission and display.

Targeted methods also are unable to detect steganographic content within mobile images due to the nature in which the detection methods are developed. In [7] and [8], these methods were developed for detection of hidden information embedded with F5 algorithm. In [9] Fridrich, et al. developed a detection method which focuses on Outguess. Problems with just translating these methods to the mobile image domain are that the embedding methods must be developed for mobile applications with memory constraints being considered. The same considerations would have to be taken into account for the development of the detection methods.

Feature based methods developed by Fridrich [10] and Farid 11] have been used to detect hidden information within JPEG images. While these methods have shown remarkable results they depend on a large data base of images to train for classification. This poses a serious problem when developing a detection method for mobile images due to fact that the characteristics of the images produced by specific mobile cameras differ widely.

In [12] and [13] McBride and Peterson developed a true blind method for detecting steganography. Lyu and Farid developed a one-class detection method that can also be considered a blind classifier [14]. While these methods use databases of images for training they are greatly simplified by eliminating the need for training with stego images which makes for a blind classifier of common and future developed stego programs. These methods can be considered for mobile image detection with modifications made to the training of clean images for the specific mobile platform.

In [15, 16, and 17], a description of the Support Vector Machine and it's uses for pattern recognition are investigated. Farid, et al. [14] used a one- class Support Vector Machine for a blind classifier of clean images. In [15], Lyu and Farid extract low-dimensional statistical feature vectors from statistical properties of audio signals for representation and use by a non-linear support vector machine for classification. The properties of the feature selection make the SVM ideal for classification of steganographic content.

This section discussed related work in detection methods and the use of the SVM for classification. While the mentioned methods have served platforms and benchmarks for steganalysis new technology has made these methods inefficient for mobile communication. In the next sections the proposed method demonstrates a new steganalysis technique designed specifically for use in mobile image communications.

3. FEATURE SELECTION

This section presents a new multi pixel comparison for steganalysis used in mobile communications. Difficulties arise when the last two least significant bit-planes of mobile images are predicted by the mobile platform. The presented method alleviates the need to consider the predicted bit-planes by considering the overall amplitude of the image pixels. Below the general algorithm is shown for the presented method.

General Algorithm for Multi Pixel Comparison

Input: Input an image to be analyzed for stego information.
Step 1: Identify the pixel set to be analyzed.
Step 2: Map the pixels into a vector representation.
Step 3: Determine the central pixel, creating pixel various sets
Step 4: Calculate the statistical average for the pixel set
Step 5: Give the pixels a weight class and calculate the statistical average adjacent pixel sets
Step 6: Determine if stego exists with the decision making process
Output: An image showing the stego locations from the received image along with an estimated amount of steganographic content per color channel.

If I is an input image of size $M \times N$, the rows and columns denoting the number of adjacent pixels surrounding a center pixel at the pixel location i, j. The masks used can be of various sizes for the number of compared pixels. As an example a simple mask consist of the following set of analyzed pixels:

$$\begin{bmatrix} x_{i,j} & x_{i,j+1} \\ x_{i+1,j} & x_{i+1,j+1} \\ x_{i+2,j} & x_{i+2,j+1} \end{bmatrix}$$

These pixels can be represented as:

$$\begin{bmatrix} x_{i,j} & x_{i,j+1} \\ x_{i+1,j} & x_{i+1,j+1} \\ x_{i+2,j} & x_{i+2,j+1} \end{bmatrix} \Rightarrow \begin{bmatrix} x_{i,j} & x_{i+1,j} & x_{i+2,j} & x_{i,j+1} & x_{i+1,j+1} & x_{i+2,j+1} \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix} = X \ .$$

By mapping the 3×2 block into a one dimensional 1×6 vector provides a means to easily measure statistical changes before and after embedding. This allows the structure of a set of masks to be applied to the pixels $[x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$. Two classes of masks may be used which are $M_1 = [m_1\, m_2\, m_3\, m_4\, m_5\, m_6]$ and $M_2 = [m_1\, m_4\, m_5\, m_2\, m_3\, m_6]$ where $m_i \in (-A, 0, A)$, $A = 2^n$, $n = 0, 2, 3$. Note, that by rearranging the values of M_1 the second mask M_2 is generated.

The two classes of masks generate three sets of sub masks each which are added to the pixel region being analyzed. The three sub masks are generated with the use of the A values. The series of weights are assigned according to the distance from the current central pixel at the current location i, j for both of the arranged pixel vectors, as follows;

$$W = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 & w_5 \end{bmatrix}$$

We apply the weights to our modulus operation between the pixels and corresponding mask values. The weight vector only consists of five values since 5 calculations are made between adjacent pixels. The weighted statistical average of adjacent pixel pairs within a comparative mask and block is defined as follows:

$$\hat{f}_{i,j}^{n}(M_k, W) = \frac{1}{\hat{n}-1}\sum_{\hat{i}=2}^{\hat{n}}\left(w_{i-1}\times\left|(m_1 \oplus x_1)-(m_{\hat{i}} \oplus x_{\hat{i}})\right|\right)$$

$$F_n = \frac{1}{NM}\sum_{i,j}\hat{f}_{i,j}^{n}(M_k, W)$$

$$F_n = \{f_1, f_2, ..., f_n\}$$

where, $\oplus$ is the modulus operation of adding to the pixel value of one bit plane only, k represents the mask used, $\hat{n}$ is the number of pixels in the mask, n represents the number of features (n = 7 in our case) and i, j subscripts represent the pixel location throughout the image.

The first six features extracted represent statistics over any modifications made to selected six bit planes and the seventh feature is calculated with no modifications made to the bit planes. These features are then used in discriminating between clean images and those containing hidden messages.

4. CLASSIFICATION

To improve classification of clean and stego images over using just the raw features, one must combine the features using a classification. A popular and powerful technique is to use a Support Vector Machine (SVM). This section describes the SVM and the kernel used for classification.

The support vector machine (SVM) is a classification algorithm that provides state-of-the-art performance in a wide variety of application domains, including handwriting recognition, object recognition, speaker identification, face detection, text categorization, time-series prediction, gene expression profile analysis, and DNA and protein analysis. SVM is a nonlinear generalization of the Generalized Portrait algorithm developed in Russia in the sixties. As such, it is firmly grounded in the framework of statistical learning theory, or VC theory, which has been developed over the last three decades by Vapnik, Chervonenkis [16] [17] and others. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Given a training set of instance-label pairs (x_i, y_i), i = 1,…,l where the support vector machines (SVM) [18] require the solution of the following optimization problem:

$$\min_{\omega,b,\xi}\frac{1}{2}\omega^T\omega + C\sum_{i=1}^{l}\xi$$

$$\text{subject to } y_i\left(\omega^T\phi(x_i)+b\right)\geq 1-\xi_i$$

where $\xi_i \geq 0$.

Here training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function ϕ. The SVM algorithm then finds a linear/nonlinear separating hyperplane with the maximal margin in this higher dimensional space. The margin is can be defined as the maximal distance between classes. $C > 0$ is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) = \phi(x_i)^T\phi(x_j)$ is called the kernel function. There are several kernels used in practice and new kernels are continuously proposed by researchers. For our research the following kernel is used:

$$\text{Radial Basis Function (RBF): } K(x_i, x_j) = e^{\left(-\gamma\left\|(x_i - x_j)\right\|^2\right)}, \gamma > 0$$

The next section, demonstrates mobile image steganography classification using a two-class SVM given a set of clean and stego image features and a one-class SVM with only clean image features. The two-class support vector machine classification technique used for detection of clean images,and separate them from images with embedded information using the radial basis function kernel. The input features calculated using the method in Section 3 are the attributes and

values used in the SVM's training algorithm, where the properties of these input data are matched with the features. After training, the SVM is used to predict feature vectors from new image instances as either clean or containing embedded information. A one-class SVM detection method, also considered as a blind classifier, is also used with a database of clean images for training. This greatly simplifies the classification method eliminating the need for training with stego images which makes the classifier a true blind steganalysis method of stego programs.

5. EXPERIMENTAL REULTS AND FINDINGS

In this section the comparisons of detection with an image data set of 100 color JPEG images taken with a Nokia 6620 camera phone are presented. While no other know cell phone detection methods exist, comparisons between the presented method and other detection methods are also presented using an image data set generated with Nikon D100 and Canon EOS Digital Rebel cameras.

5.1 Detection of Nokia Stego Images

The image set taken with the Nokia 6620 camera phone were mid-quality 640x480 Symbian JPEG images. For testing purposes the image data set consisted of 100 clean images, 100 images with 3% stego modification and 100 images with 8% stego modification. The image data is embedded using a simple DCT coefficient method with an embedding rate percentage based on the image size. The embedding technique is a modification of algorithm discussed in [2].

Figure 1: Sample set of images from the set used to detect Nokia clean images from Nokia stego images.

Two-class analysis was performed using 5-fold cross validation using both clean and stego sample instances, where the training set consists of clean images and stego images. Testing was conducted with a set of images consists of 20% clean image dataset and 20% stego image set. Table 1 shows the results from these experiments using 3% and 8% embedding rates. The percentage of true positive (TP) is the average of detection accuracy for clean images when clean images are analyzed. The true negative (TN) represents the average of detection accuracy for detecting stego image when in fact a stego image is present.

Table 1 shows the accuracy of detection for the 3% data not as good as that of the 8% but this is expected because the modifications to an image with 8% embedded data cause a weaker correlation between the stego image features and the clean image features. The opposite holds for the correlation between the features for clean image features and features for 3% stego images. What is interesting is that there is no statistical difference between the embedding procedure which means that the features and classification method are robust to variations in embedding amounts.

Table 1: Detection Accuracy of Nokia Clean and Stego Images [2]

Stego (3% Embedded Data)		Stego (8% Embedded Data)	
TP = 94.75± 3.68%	TN = 91.0± 8.21%	TP = 94.5± 3.37%	TN = 96.0± 2.23%

The blind classification analysis was performed using 5-fold cross validation in a blind manner, where the training set consists only of clean images and the testing set consists of 20% of the clean image dataset, and 20% of the stego image sets. Table 2 shows the accuracy of detection for the 3% stego image data and the 8% stego images. As previously mentioned classification accuracy of the 8% embedded data is higher than the 3% image data due to the correlation between stego image features and the clean image features.

Table 2: Blind Detection Accuracy of Nokia Clean and Stego Images [2]

Stego (3% Embedded Data)		Stego (8% Embedded Data)	
TP = 87.8± 4.34%	TN =92.4± 7.01%	TP =95.6± 4.02%	TN =94.67± 7.67%

5.2 Detection of Commonly Used Stego Images

The embedding algorithm used in the Nokia 6620 camera phone uses different techniques than methods in [19] and [8] since the camera phone predicts some of the bit-plane. To make comparisons with existing steganalysis techniques would cause a bias towards the presented method, so the presented method will be used to detect stego images that RS Steganalysis [5] and "Steganalysis using color wavelet statistics and one-class support vector machines" [14] are capable of detecting.

When using the same detection method for LSB embedding the detection method is compared with RS Steganalysis [5]. Table 3 show the accuracy of detection for Multi Weighted Masks Detection vs RS Steganalysis when all 100 images have been tested with 0%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20% and 25% of steganographic content using a random embedding method. The error of classification is shown in Figure 2, where the presented method maintains accuracy within 2% up to an embedding percentage of 5% and increases in error by 5% at 25% embedded information while RS Steganalysis reaches 7% error in classification with only 5% embedded information and continues a moderate linear increase in error for larger percentages of hidden information.

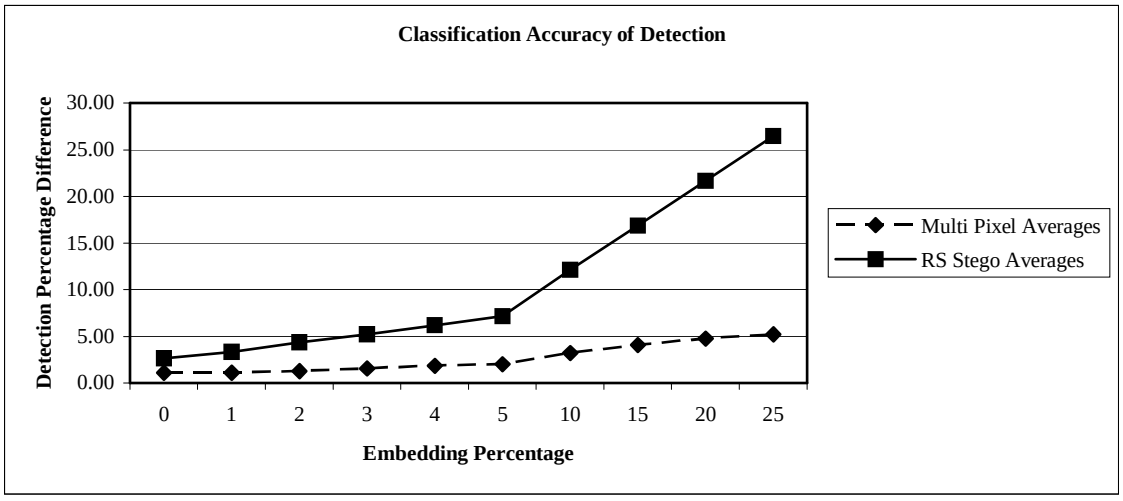

Figure 2: Error of actual detection percentage

Table 3: Detection Percentage Averages for Color Images

Multi Pixel Average	1.10	2.07	3.28	4.55	5.80	7.01	13.18	19.07	24.75	30.20
RS Steog Averages	2.61	4.31	6.32	8.23	10.16	12.15	22.09	31.84	41.66	51.43
Percent Stego	0	1	2	3	4	5	10	15	20	25

Table 4 shows the results from a 2 class SVM classifier using feature extraction developed by Lyu and Farid [14]. These results are based on feature extraction from Matlab code provided by Farid and the combination of the SVM with the RBF kernel. The table shows the average classification accuracy of detection for "Steganalysis using color wavelet statistics and one-class support vector machines." The number of images used are; 100 clean images tested against 100 stego images with varying in modifications caused by embedding method of 1%, 2%, 3%, 4%, 5% and 10% alterations to the images. The testing was conducted using 5-fold cross validation, 80% clean images and 20% stego images. The feature reduction shown was done with a technique developed by Bauer, et al. [20].

Table 4: Detection Averages for Wavelet Based Features

	Classification Accuracy using 2 Class SVM (Radial Basis Kernel)							
Detection Percentage	90.0%	89.2%	90.3%	89.6%	90.3%	89.8%	85.7%	74.4%
Number of Features	72	60	48	36	24	12	6	3
Multi Pixel Features	7 Features 94.6% Classification Accuracy							

CONCLUSION

In this article feature extraction for both clean images sets and dirty (stego) images was presented. It has been shown that the presented features are able to determine if images from mobile picture phones contain steganographic content or if normal bit-plane images contain steganographic content.

In the experimental results two cases are shown:

1) Analysis was performed on mid-quality 640x480 Symbian JPEG images with the test set of n clean images and n stego images out of the set of 100 images each. The results of the 5-fold cross validation using the two-class classifier shown in Table 1 yield results of 94.75% (std 3.6%) TP and 91% (std 1.6%) TN when clean and 3% stego images are compared and 94.5% (std 2.7%) TP and 96% (std 4.47%) TN when clean and 8% stego images are compared.

   The blind classification using a one-class classifier shown in Table 1 yield results of 87.8% (std 4.34%) TP and 92.4% (std 7.01%) TN when clean and 3% stego images are compared and 95.6% (std 4.02%) TP and 94.6% (std 7.67%) TN when clean and 8% stego images are compared

2) Analysis when compared to RS Steganalysis [5]; the presented method determines the amount of steganographic content with an accuracy within 2% maintained up to 10% randomly embedded information and within 5% accuracy when embedding between 15% and 25% hidden information, RS Steganalysis reaches 7% error in classification with only 5% embedded information and continues a moderate linear increase in error for larger percentages of hidden information

   Analysis when compared to "Steganalysis using color wavelet statistics and one-class support vector machines" [14]; a two class classification was performed to determine the accuracy in detection of 94.6% using only 7 features and the duplicated method developed in [14] has a classification accuracy of 90% using 12 features.

The presented feature extraction is also extended to other lossy compression embedding techniques by comparing DCT coefficients.

ACKNOWLEDGMENTS

This research was partially funded by the Center for Infrastructure Assurance and Security and the US Air Force. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Air Force, Department of Defense or the U.S. Government. We would additionally like to express our appreciation to June Rodriguez for the contribution of a multitude of digital images for analytical support.

REFERENCES

1. N.F. Johnson and S. Jajodia, “Steganalysis: The Investigation of Hidden Information”, IEEE Information Technology Conference, 1998, URL: http://www.jjtc.com/pub/it98a.htm
2. S. S. Agaian, D. Akopian and S. D. D’Souza, “Wireless Steganography”, SPIE Symposium on Electronic Imaging, San Jose, CA, 2006
3. S. S. Agaian, B. M. Rodriguez and G. Dietricht, “Steganalysis Using Modified Pixel Comparison and Complexity Measure”, SPIE Symposium on Electronic Imaging, San Jose, CA, 2004
4. S. Dumitrescu, X. Wu and Z. Wang, “Detection of LSB Steganography via Sample Pair Analysis”, LNCS 2578, 355-372, 2003.
5. J. Fridrich, M. Goljan and R. Du, “Detecting LSB Steganography in Color and Gray-Scale Images,” Magazine of IEEE Multimedia Special Issue on Security, October-November 2001, pp. 22-28
6. A. Westfeld, and A. Pfitzmann, “Attacks on Steganographic Systems”, Proceedings of the 3rd Information Hiding Workshop, Dresden, Germany, September 28-October 1, 1999, pp. 61-75.
7. J. Fridrich, M. Goljan, and D. Hogea, “Steganalysis of JPEG Images: Breaking the F5 Algorithm”, 5th Information Hiding Workshop, Noordwijkerhout, The Netherlands, 7–9 October 2002, pp. 310-323.
8. A. Westfeld, “F5a steganographic algorithm: High capacity despite better steganalysis,” 4th International Workshop on Information Hiding, 2001.
9. J. Fridrich, M. Goljan, and D. Hogea, “Attacking the OutGuess”, Proceedings of the ACM Workshop on Multimedia and Security 2002, Juan-les-Pins, France, December 6, 2002.
10. J. Fridrich, “Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes,” Proceedings of the 6th Information Hiding Workshop, Toronto, Canada, May 23-25, 2004.
11. H. Farid, “Detecting Hidden Messages Using Higher-Order Statistical Models”, International Conference on Image Processing (ICIP), Rochester, NY, 2002.
12. B. McBride and G. L. Peterson, “Blind Data Classification using Hyper-Dimensional Convex Polytopes,” Proceedings of the 17th International FLAIRS Conference, pp 520-525, 2004.
13. B. McBride and G. L. Peterson, “A new blind method for detecting novel steganography”, Digital Investigation, Vol 2, 50-70, 2005.
14. S. Lyu and H. Farid, “Steganalysis using color wavelet statistics and one-class support vector machines,” SPIE Symposium on Electronic Imaging, San Jose, CA, 2004.
15. M Johnson, S Lyu and H Farid, "Steganalysis of Recorded Speech", SPIE Symposium on Electronic Imaging, San Jose, CA, 2005.
16. V. Vapnik and A. Lerner. Pattern recognition using generalized portrait method. Automation and Remote Control, 1963
17. Scholkopf, Bernhard; Smola, Alexander J. “Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond”, The MIT Press, 2003
18. C.-W. Hsu, C.-C. Chang, C.-J. Lin. “A practical guide to support vector classification”, http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
19. P. Sallee, “Model-based steganography,” International Workshop on Digital Watermarking, Seoul, Korea, 2003.
20. K. W. Bauer, S. G. Alsing and K. A. Greene, “Feature screening using signal-to-noise ratios”, Neurocomputing 31 (2000) pp. 29 - 44

*sos.agaian@utsa.edu; phone 1 210 458-5939; fax 1 210 458-5947