International Journal of Security and Its Applications
Vol. 12, No. 1 (2018), pp.21-36
http://dx.doi.org/10.14257/ijsia.2018.12.1.02

Cover Selection for More Secure Steganography

Reza Esfahani[1*], Zynolabedin Norozi[2] and Gholamreza Jandaghi[3]

[1]Information Technology and Communication Faculty, Imam Hossein Comprehensive University, Tehran, Iran*
[2]Information Technology and Communication Faculty, Imam Hossein Comprehensive University, Tehran, Iran
[3]University of Tehran, Iran
[1]resfahani@ihu.ac.ir, [2]znorozi@ihu.ac.ir, [3]jandaghi@ut.ac.ir

Abstract

The lack of a cover bank with minimum detectability against various steganalysis methods is one of the major concerns in the field of secure communication. This paper introduces a method for preparing cover images with appropriate security. To this end, some important features of the image such as contrast, energy, and the use of fuzzy logic are considered by selecting the Harris threshold level. First, using each cover image, some images are obtained with different contrasts and a constant Harris threshold level; besides, in order to extract the above-mentioned features, the "Gray Level Co-occurrence Matrices" (GLCM) are calculated. Then, using the fuzzy logic, different security levels are presented for images. The images with different security levels could be stored in different image banks. Afterwards, to achieve lower detectability a stegangraphy against a steganalyzer technique, the difference were obtained between features of clean and stego images, and then to evaluate the performance of each feature, "Area Under the Curve" (AUC) values were calculated using the steganalyzer to the "Fisher Linear Discriminant" (FLD) classification. Actually, the cover bank is formed through cover classification based on an index pre-steganography security level.

Keywords*: Co-occurrence matrices, fuzzy logic, Harris corner detector, Fisher's linear discriminant classifier*

1. Introduction

Security of steganography is influenced by different factors including the type of cover image, method of selecting a part of the cover to apply the changes, the way of embedding the message, and the number of changes caused by message embedding. The type of the image and selecting a part of it such as edges, or even the feature-points near the edges or corners to apply the changes is among the criteria that can be defined on the co-occurrence matrices. This leads to minimize the changes caused by embedding the message. The uncertainty of the fuzzy logic would create a range to define further conditions and rules. Tuning the parameters is one of the issues that must be considered. Since during data insertion, there exists no agreement between the capacity and the message size, there should be a method for selection a few covers. It is necessary to prepare a cover bank with an appropriate security level for a certain application before steganography. The issue of selecting an appropriate cover[2] is of great interest and importance in steganography. Some research papers are focused on selecting the

Received (August 23, 2017), Review Result (December 19, 2017), Accepted (January 4, 2018)
* Corresponding Author
[2] In the appropriate cover, there is an appropriate balance between the parameters of capacity (embedding rate), security and robustness.

ISSN: 1738-9976 IJSIA

appropriate cover considering the image texture as in [1]. It seems that statistical criteria would help more if it is taking into account. On the other hand, the maximum capacity, provided that it is balanced with security in steganography, is one of the objectives of designing the algorithms, as presented in [2] and [3]. However, to the best of our knowledge no one has discussed the issue of leveling the cover security before data embedding. The use of fuzzy logic in data hiding can be seen in some papers such as [4] in which, at the FIS output, sensitivity of the image edge for data hiding is discussed. Kharrazi et. al. discussed the issue of cover selection with three scenarios of no knowledge, partial knowledge, and full knowledge of the steganalyzer with the aim of maximum embedding rate and minimum detectability against steganalysis [5]. In this paper, the appropriate image are extracted. First, contrast and energy features of near some of the corners and edges of images have been calculated. Then leveling the cover security would be investigated using the fuzzy logic. To this end, statistical criteria are also considered as well as image texture. In the proposed method, regarding the importance of the co-occurrence matrices, the criteria capable of expressing the similarity of these matrices are used. Using the features of co-occurrence matrices and fuzzy logic, we categorized two forms VHS (Very High Security) and HS (High Security). Two banks 5000 images will created each for VHS and HS. In this paper the cover bank is formed through cover classification based on an index such as pre-steganography security level. Now, to achieve lower detectability against steganalyzer system such a bank requires an aggregated evaluation using “Area Under the Curve” (AUC) value of each feature and “Fisher Linear Discriminant” (FLD) classification.

The rest of the paper is organized as follows. In section two, the required concepts and definition such as fuzzy logic, required criteria, classifier, and Harris detector are presented. In the third section, the proposed method in selecting appropriate cover is presented while in section four the experimental results have been demonstrated. Finally section five concludes the paper.

2. Concepts

In this section, the concepts and definitions that have been used in this paper will be presented briefly.

Fuzzy logic

Unlike classical set theory that security has assessed in binary format (Secure and Not secure), sometimes you have need the different security levels. The fuzzy logic has been used for security leveling in several papers such as [1], [4] and [6]. The fuzzy logic relies on the theory of the uncertainty sets or, namely, the fuzzy sets. In this paper, a range of different security levels have been defined using to the uncertainty for estimating security value of any image. The Mamdani method is simple and intuitive, commonly is used in easy problems with two inputs and one output. The output of Mamdani method offers higher accuracy. In order to apply the fuzzy rules for system controlling, The Mamdanis Fuzzy Inference System (FIS) will be used.

Indistinguishability and Security

The system transparency represents the lack of significant difference before and after data embedding in the cover media. The two criteria of SSIM[3] and PSNR[4] of images are good measures for transparency since they can be eventually used to evaluate the images [7], [8]. From security viewpoint, not only there should be no significant difference between the cover and stego media, but they must be very close to each other from

[3] Structural SImilarity Measure
[4] Peak Signal to Noise Ratio

statistical features. This is highly critical for features mostly used by steganalysis algorithms. To measure this fact, the subject of advantage of a warden in performing any detection before and after embedding a message in the cover can be examined as follows [9] and [10].

$$Adv_{C,S} =: \Pr[W^C = 1] - \Pr[W^S = 1] \quad (1)$$

$Adv_{C,S}(W)$ indicates the "advantage of the warden" in distinguishing the stego signal from the cover one. In [11], indistinguishability was considered equivalent to security. Security implies the statistical significance caused by the distortion of message embedding. In many studies, a steganography system is called $secure - \varepsilon$ if the relative entropy between the cover and stego distribution is at most equal to ε. The signals C and S are considered as cover and stego signals; if the warden W can perform an activity on C and S with probability of $Pr[W^C = 1]$ and $Pr[W^S = 1]$, respectively and if the Equation (2) is established in a way that the value of ε is very small, then the signal C would not be distinguishable from the stego signal S; furthermore, when $\varepsilon \to 0$ such that $\varepsilon \cong 0$, then the full indistinguishability and, as a result, the perfect security will be achieved:

$$\Pr[W^C = 1] - \Pr[W^S = 1] < \varepsilon,\ if\ \ \varepsilon \cong 0\ \Rightarrow\ \Pr[W^C = 1] = \Pr[W^S = 1] \quad (2)$$

As previously stated, a steganographic method is called perfect secure if $\varepsilon = 0$. Such a definition will result in numerous problems. Such a definition of security will be appropriate for the sequence of random bits, while it is not efficient for the covers such as natural images; because, there is a high dependency between the signal samples and these dependencies can be used even when the distribution keep unchanged as much as possible which causes to design an appropriate steganalyzer. The indistinguishability value shown in Equation (3) with SoI_s[5] is limited in the range of [0,1], where the value of 1 indicates the maximum similarity between the two images.

$$Idistinguishability\ \therefore\ SoI_s(C_{n\times m}, M_k)\ \longrightarrow 0, SoI_s(C_{n\times m}, M_k) \leq 1\ if\ \Pr[W^C = 1] = \Pr[W^S = 1] \Longrightarrow\ SoI_s(C_{n\times m}, M_k) = Perfect\ Secure \quad (3)$$

Indistinguishability or $SoI_s(C_{n\times m}, M_k)$ indicates the security level. Moreover, $C_{n\times m}$ and M_k indicate the cover image with dimensions of $n \times m$ and the secret message with length of k, respectively. Also, according to Equation (2), the perfect security can be observed here as the Equation (3).

AUC criterion

In this paper, the AUC[6] criterion, which indicates the area under the ROC[7] curve, was used [12]. Regardless of the fact that the result of testing a steganography is a numerical value and how different it is from the previous and next steganography results, the use of the area value under the curve would facilitate the operation. Depending on the changes in the shape of the steganographies abundance curve as well as the normality or abnormality of the curves, the area under the ROC curve will be independent from various forms of the studied population and there will be only one significant parameter, which will be the area under the curve as well. The ROC diagram is a method to investigate the efficiency of the classifiers. Thus, the values of AUC[8] will indicate the test status [13].

[5] SoI_s (Security of Image - Stego)
[6] Area Under Curve
[7] Receiver Operating Characteristic
[88] 0.9-1 = excellent, 0.8-0.9 = good, 0.7-0.8 = fair, 0.6-0.7 = poor, 0.5-0.6 = fail

If the test is carried out ideally, AUC will be equal to 1, but when the noise is taken into account, the FP[9] and FN[10] results will influenced and also the AUC will be reduced as well.

Classification and feature extraction

Classification is considered as one of the important branches of artificial intelligence, which refers to putting a sequence of unknown features in a previously known class. Feature means a description of the objects, the aim of which is to classify them. The purpose of extracting the features is to reduce the image data[11] and achieve better performance. In this paper, the features of color, features resulted from the co-occurrence matrices, FFT[12] and FWT[13] of image are used for detection. The color features included the mean, variance, skewness, and kurtosis, and the texture features included entropy, energy, contrast, and correlation [14], [5]. Most of the steganographic methods lead to changes in the low-value bit planes, thus eliminating the valuable bits would not lead to the loss of information related to the message signal. For example, for embedding the Jsteg[14] algorithm [15], more than 90% of the changes caused by the message embedding have occurred in the first four bits. The main content of the images is located within the valuable bits, meaning that the low-value bit planes contain the noise information, high frequencies of the image, and information of the embedded message. To classify images, different classifiers can be used [16]. The FLD is used to reduce the dimension [17], since FLD selects the best feature among several features and then reduces the multiple dimensions to a single one [18].

Harris detector

Boundaries of the objects usually cause the changes in the pixels intensity, and edge detection is mainly used to identify these changes. One of the features of the edges is their lower sensitivity to the brightness changes compared to the color features, and thus it seems more appropriate to embed the secret message in the edges [19]. The algorithms that track the objects boundaries commonly consider the edges as a feature and representative of the object. Near some of the corners and edges of the image, there are prominent points known as feature points. Detectors of these feature points find the prominent points in the normal image. Harris corner detector is one of the detectors of the feature points used for three-dimensional reconstruction [20], [21], [22].

3. Appropriate Cover Image

Regarding the growth and efficiency of the steganalyzers, random selection of a cover medium for steganography reduces the originality and scientific nature of the process; thus, selecting a cover medium from a favorable data bank through a scientific process can be useful for promoting the steganographic methods and, similarly, reinforcing the design against different steganalyzers. Subsequently, using the contrast, fuzzy logic, and then the bank analysis of the obtained images, the cover image will be obtained with desirable security for information embedding.

[9] False Positives
[10] False Negatives
[11] by certain features or the features of each segmented area such as color, texture, or shape in the image
[12] Fast Fourier Transform
[13] Fast Wavelet Transform
[14] LSB and DCT steganography

3.1. The Proposed Method

Figure (1) shows the procedure of the proposed method. From each photography sample, several pictures are taken. However, for time-saving, it has been used from images of the well-known BOSS bank; then, from each image, the images with a constant Harris threshold and different contrasts are made. The values of the colors used in an image reveal that some of the colors might be used in mass in the image, and some others less. Increasing the contrast equalizes the use of various colors and the above-mentioned significant difference will no longer exist. One of the major tasks in processing the images is to increase the images resolution in order to achieve more accurate visual interpretation and vision. There exist several techniques for achieving this objective, the most important of which is the image contrast and filtering. First, different images of an image with different contrasts are considered, and the appropriate image is extracted from them and, accordingly, a bank of images will be developed. Then, selecting the cover image based on the steganalysis will be performed; in other words, those images that have less detectability against various steganalysis techniques will be selected for steganography.

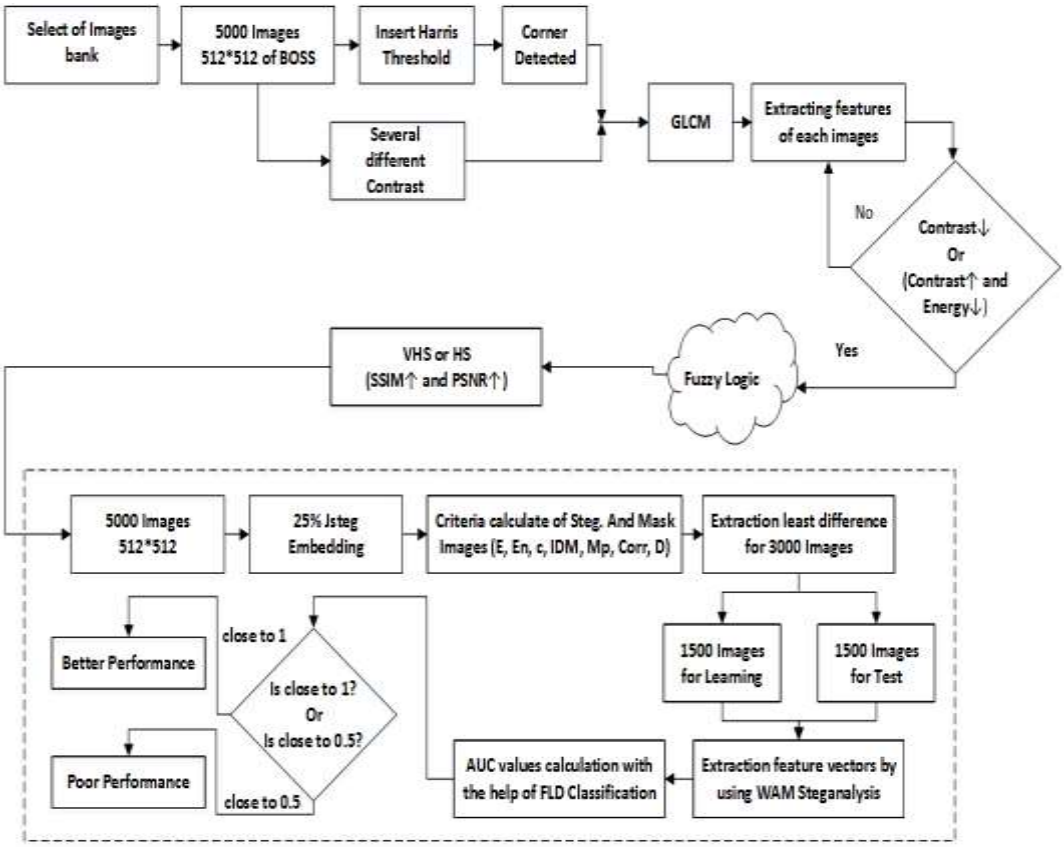


Figure 1. Generic Block Diagram of the Proposed Method

3.1.1. Selecting Cover Images with High Security

First, a set (initial bank) of the images is prepared. Next, the features of the images are extracted using the co-occurrence matrices, and then the bank of images is sorted based on the features of contrast and energy through the fuzzy logic, so that some of the corners and edges of the image (the feature points resulted from Harris detector) have the appropriate SSIM and PSNR of image during steganography at the close prominent points, and the images can be extracted from or maintained in the bank using a threshold level of contrast and energy (Figure 2).

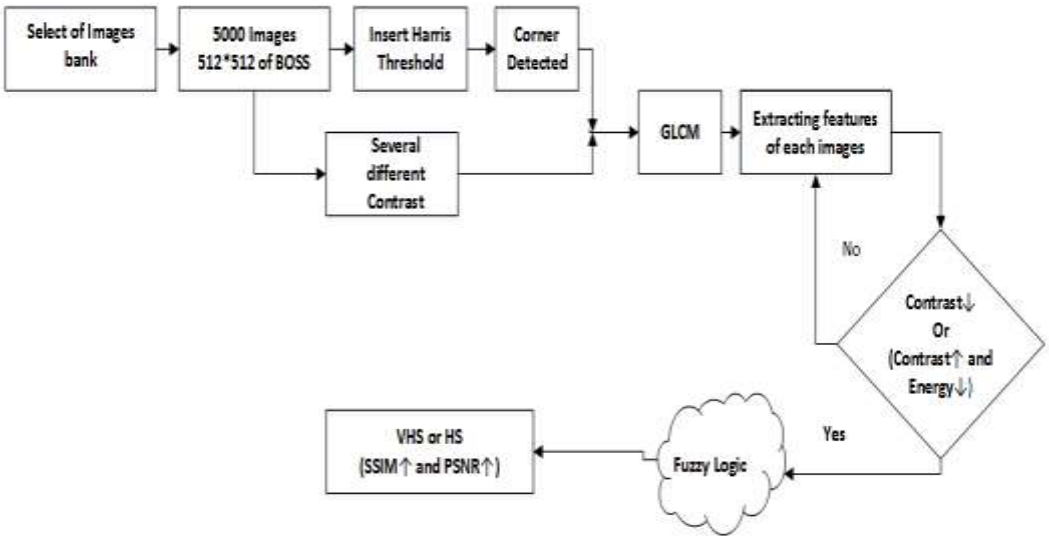


Figure 2. Selecting Cover Image with Appropriate Security Using the Contrast and Energy Criteria

Regarding the main task of contrast, which is the image resolution and transparency, it is very important to examine the contrast in steganography; because, high contrast, lonely, means high resolution of the image, the transparency and indistinguishability of which is very difficult to preserve so that even a minor change will be observable visually or statistically. Thus, the low contrast for a cover image will be appropriate for steganography. However, in case of high contrast and low energy, the changes will not be visible. The contrast defined with Equation (4) is obtained using the "Gray Level Co-occurrence Matrices" (GLCM). i and j refer to the brightness levels of the two adjacent pixels, and $p(i,j)$ indicates the 2D histogram in GLCM. In fact, $\sum_i \sum_j p(i,j)$ is the percentage of the paired pixels, and their brightness levels are different by $(i-j)$. Therefore, the smaller the value of $|i-j|$, the smaller value of the contrast $(c \downarrow)$ we have.

$$C = \sum_i \sum_j p(i,j)(i-j)^2 \ , 0 < p(i,j) \leq 1 \ , if \ |i-j| \to 0 \Rightarrow c \downarrow \tag{4}$$

Now, if the information is embedded as LSB in some adjacent pixels with close brightness levels, it will not be distinguishable due to little difference in the brightness of the adjacent pixels. In general, if C_c and C_s indicate the contrast of the cover image and the contrast of the stego image, respectively, then, according to Equation (2), we have:

$$If \ \ c_c \downarrow \therefore \Pr[W^{C_C} = 1] \cong \Pr[W^{C_S} = 1] \tag{5}$$
$$\Rightarrow \ Adv_{C,S}(W^C) =: \Pr[W^{C_C} = 1] - \Pr[W^{C_S} = 1] = \varepsilon \cong 0 \Rightarrow SoI_c \uparrow$$

Equation (5) shows that in case that the clean image has low contrast, the "advantage of the warden" in distinguishing has been resulted indistinguishably. This means that distinguish the cover and stego image is limited. The term SoI_c[15] refers to the security level of the clean image, and ↑ indicates the high security and very high security levels[16] . But the value of $p(i,j)$ will be also effective in terms of being large or small. In order to observe the value of $p(i,j)$, it is better to examine the energy criterion (E), which is

[15] SoI_c (Security of Image - Clear)
[16] Image security levels are determined to minimum value of 35 dB for PSNR and maximum value of 1 for SSIM

obtained by the gray level co-occurrence matrices (Eq.6). The energy specifies the images softness, and if all the pixels have equal brightness levels of *k*, then we will have *E*=1, according to *p (k, k)* = 1. Therefore, the more uniform the image area, the more uniform the distribution and the lower the value of *E*, and the smaller value of *E*, which means that $p(i,j)$ is small as well.

$$E = \sum_i \sum_j (p(i-j))^2 \ , \ if \ \ p(i,j) \downarrow \Rightarrow E \downarrow \tag{6}$$

Now, the bigger the $|i-j|$, the higher the contrast ($c \uparrow$) :

$$C = \sum_i \sum_j p(i,j)(i-j)^2 \ , \ if \ |i-j| \gg 0 \Rightarrow c \uparrow \tag{7}$$

$$If \ \ c_c \uparrow and \ E_c \downarrow \therefore \Pr\left[W^{(C_C,E_C)} = 1\right] \cong \Pr\left[W^{(C_S,E_S)} = 1\right] \tag{8}$$

$$\Rightarrow \ Adv_{C,S}\left(W^{(C,E)}\right) =: \Pr\left[W^{(C_C,E_C)} = 1\right] - \Pr\left[W^{(C_S,E_S)} = 1\right] = \varepsilon \cong 0 \Rightarrow SoI_c \uparrow$$

Equations (5) and (8) are equivalent. With regard the fuzzy logic, different weights has been attributed them ("Very High Security" for Eq. 5 and "High Security" for Eq. 8).

According to Figure 3, it has been considerd two inputs and one output for FIS. The inputs and output have Guassian form. The first input of FIS is the contrast in the range of $0 \leq Cnt \leq (S-1)^2$ where *S* is the size of gray level co-occurrence matrices and it's value is 8. The second input is the energy (*E*) where $0 \leq E \leq 1$. Also, The output is the security level of the cover image (*Sec*) and has five different values between 0 and 1 ($0 \leq Sec \leq 1$).

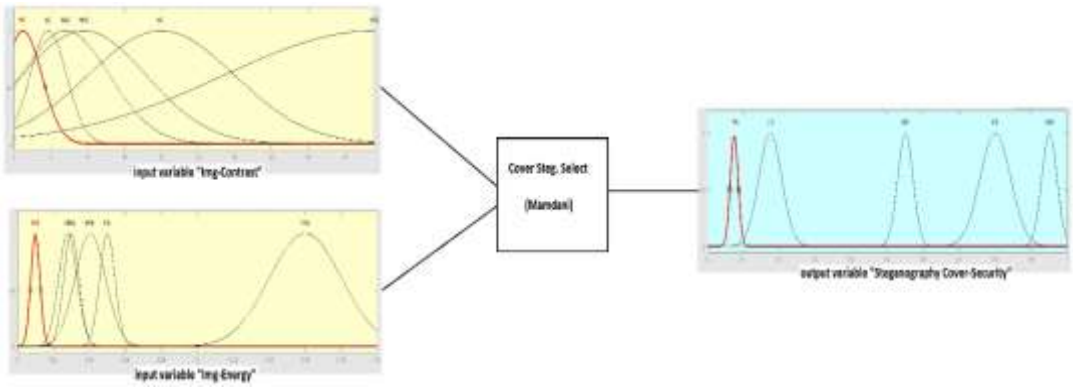


Figure 3. Fuzzy Inference System (FIS) of the proposed Method with two inputs (Contrast and Energy) and One Output (Security) Gaussian Curve Membership Function

As seen in the last column of Table (1-c), the probability of each output has been obtained regardding to the rules of inference resulted from simulation (including 20 rules) (Figure 4). For example, the probability of the VHS (very high security) level in this method has been calculated with regard to the number of the VHSs counted from among the 20 rules of inference as $\Pr[SoI_C(C_{n\times m}, M_k)] = \Pr[VHS] = \frac{4}{20}$. In other words, the security level of the clean image with VHS weight has a probability of $\frac{4}{20}$.

Table 1. FIS Inputs and Outputs

(a) Input 1 : Contrast levels of cover image			
No	Contrast Weight	Contrast Range	Summary
1	Very High Contrast	$22.7 < Cnt \leq 49$	VHC
2	High Contrast	$9.5 < Cnt \leq 22.7$	HC
3	Middle High Contrast	$4.7 < Cnt \leq 9.5$	MHC
4	Middle Low Contrast	$2.8 < Cnt \leq 4.7$	MLC
5	Low Contrast	$1.2 < Cnt \leq 2.8$	LC
6	Poor Contrast	$0 < Cnt \leq 1.2$	PC

(b) Input 2 : Energy levels of cover image			
No	Energy Weight	Energy Range	Summary
1	Very High Energy	$0.15 < Eng \leq 1$	VHE
2	High Energy	$0.06 < Eng \leq 0.15$	HE
3	Middle High Energy	$0.04 < Eng \leq 0.06$	MHE
4	Middle Low Energy	$0.01 < Eng \leq 0.04$	MLE
5	Low Energy	$0.005 < Eng \leq 0.01$	LE
6	Very Low Energy	$0 < Eng \leq 0.005$	VLE

(c) Output : Security levels of cover image				
No	Security Weight	Security Range	Summary	Pr.
1	Very High Security	$0.8 < Sec \leq 1$	VHS	4/20
2	High Security	$0.5 < Sec \leq 0.8$	HS	6/20
3	Middle High Security	$0.1 < Sec \leq 0.5$	MHS	3/20
4	Low Security	$0.07 < Sec \leq 0.1$	LS	3/20
5	Poor Security	$0 < Sec \leq 0.07$	PS	4/20

1. If (Img-Contrast is PC) then (Steganography-Security is VHS) (1)
2. If (Img-Contrast is LC) then (Steganography-Security is HS) (1)
3. If (Img-Contrast is MLC) then (Steganography-Security is HS) (1)
4. If (Img-Contrast is VHC) and (Img-Energy is LE) then (Steganography-Security is VHS) (1)
5. If (Img-Contrast is VHC) and (Img-Energy is VLE) then (Steganography-Security is VHS) (1)
6. If (Img-Contrast is VHC) and (Img-Energy is MLE) then (Steganography-Security is HS) (1)
7. If (Img-Contrast is VHC) and (Img-Energy is MHE) then (Steganography-Security is MS) (1)
8. If (Img-Contrast is VHC) and (Img-Energy is HE) then (Steganography-Security is LS) (1)
9. If (Img-Contrast is VHC) and (Img-Energy is VHE) then (Steganography-Security is PS) (1)
10. If (Img-Contrast is HC) and (Img-Energy is VLE) then (Steganography-Security is VHS) (1)
11. If (Img-Contrast is HC) and (Img-Energy is MLE) then (Steganography-Security is HS) (1)
12. If (Img-Contrast is HC) and (Img-Energy is LE) then (Steganography-Security is HS) (1)
13. If (Img-Contrast is HC) and (Img-Energy is HE) then (Steganography-Security is MS) (1)
14. If (Img-Contrast is HC) and (Img-Energy is MHE) then (Steganography-Security is LS) (1)
15. If (Img-Contrast is HC) and (Img-Energy is VHE) then (Steganography-Security is PS) (1)
16. If (Img-Contrast is MHC) and (Img-Energy is VHE) then (Steganography-Security is HS) (1)
17. If (Img-Contrast is MHC) and (Img-Energy is HE) then (Steganography-Security is MS) (1)
18. If (Img-Contrast is MHC) and (Img-Energy is MLE) then (Steganography-Security is LS) (1)
19. If (Img-Contrast is MHC) and (Img-Energy is LE) then (Steganography-Security is PS) (1)
20. If (Img-Contrast is MHC) and (Img-Energy is VLE) then (Steganography-Security is PS) (1)

Figure 4. 20 Rules of FIS Output

Figure (5) demonstrates the steganography security estimation graph based on the contrast and energy inputs. The images with appropriate contrast and energy that have high security (VHS and HS[17]) would be selected for the bank of cover images.

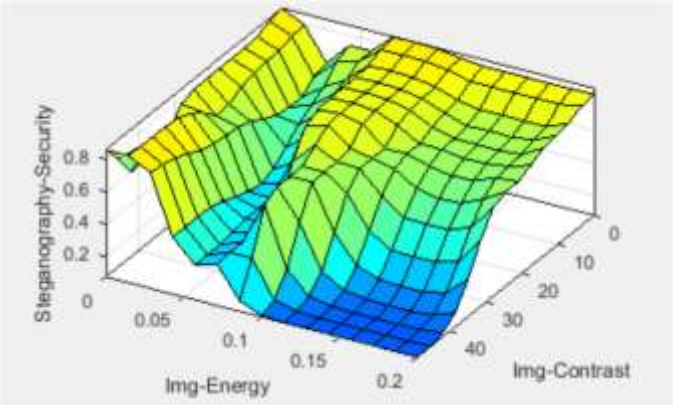


Figure 5. Steganography Security Estimation Graph Based On the Contrast and Energy Inputs

3.1.2. Selecting Cover Image Based On Steganalysis

To improve the security of the steganography method, the images can be used for embedding so that a low detectability against various steganalysis techniques is achieved. In [5], the following criteria have been used for cover selection:

- Number of Modification
- Mean Square Error (MSE)
- Prediction Error
- Watson criteria
- Structural Similarity Measure (SSIM)

In the present paper, regarding the importance of the correlation of the values of the image pixels, the feature extraction has been used along with the co-occurrence matrices. The co-occurrence matrices represent in some way the image probability distribution, which is a good criterion for cover selection. In fact, the cover image is selected so that the co-occurrence matrices representing the statistical distribution are close to each other as much as possible. The co-occurrence matrices represent the relationship of gray levels of the points in an image. This matrix is based on the conditional second-order estimation of the probability density function. For each image, four co-occurrence matrices of M1, M2, M3, and M4 are calculated in different directions of 0, 45, 90, and 135 degrees. In this method, after calculating the gray level co-occurrence matrices of the image, the features are extracted. According to several tests, the following criteria that are capable to explain the similarity of these matrices have been used:

- Energy

$$E = \sum_i \sum_j (p(i,j))^2 \quad (9)$$

- Entropy

$$E_n = -\sum_i \sum_j p(i,j) log\, p(i,j) \quad (10)$$

- Contrast

$$C = \sum_i \sum_j p(i,j)(i-j)^2 \quad (11)$$

- Inverse difference moment[18]

[17] According to table 1-C, is “High Security”
[18] namely, the homogeneity that was extracted as the features of each image at the beginning

$$IDM = \sum_i \sum_j \frac{(p(i,j))^2}{|i-j|}, \quad i \neq j \tag{12}$$

- Maximum probability

$$MP = \max(p(i,j)) \tag{13}$$

- Correlation

$$Corr = \sum_i \sum_j \frac{(i-\mu)(j-\mu)p(i,j)}{\sigma^2} \tag{14}$$

- Divergence criterion: This measure can be used in steganography for quantizing the statistical changes, which is the distance between the co-occurrence matrices $C^d(X)$ (cover image) and $C^d(S)$ (steg. image).

$$D\left(C^d(X), C^d(S)\right) = \sum_i \sum_j C^d_{i,j}(X) log\left(\frac{C^d_{i,j}(X) - \sum_j C^d_{i,j}(S)}{\sum_j C^d_{i,j}(X) - C^d_{i,j}(S)}\right), i, j \epsilon y \tag{15}$$

It should be noted that, not only the majority of the steganalysis methods can be expressed based on the co-occurrence matrices, but also the cover evaluation can be performed based on the steganalysis, which is the basis of our task in this section [23]. The main problem in using these matrices is their very large size. For instance, for the 8-bit gray levels, this size is equal to 65636, so they cannot be directly used as the features. Therefore, the classification is eventually performed using AUC and FLD.

Since most of the steganographic method lead to some changes in the low-value bit planes, eliminating the valuable bits would not result in the loss of signal-related information; for instance, for the Jsteg embedding, more than 90% of the changes have occurred as a result of the message embedding in the first four bits. Hence, following the proposed method, by considering the effect of the above-mentioned criteria on selecting the appropriate cover with regard to the experimental results of the Jsteg embedding on the LC images and HC-LE images, the procedure was performed as in Figure (6):

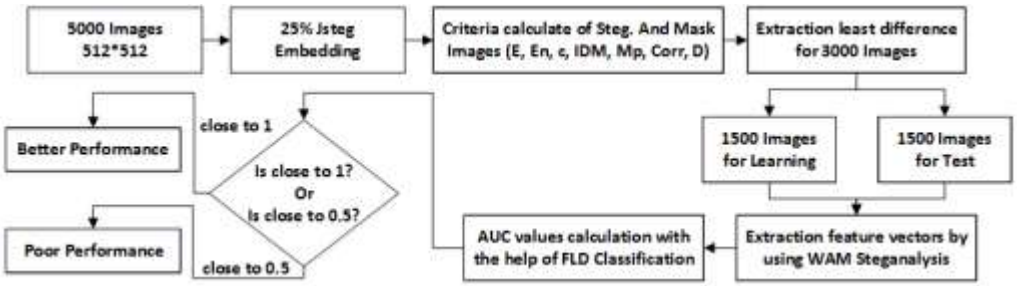
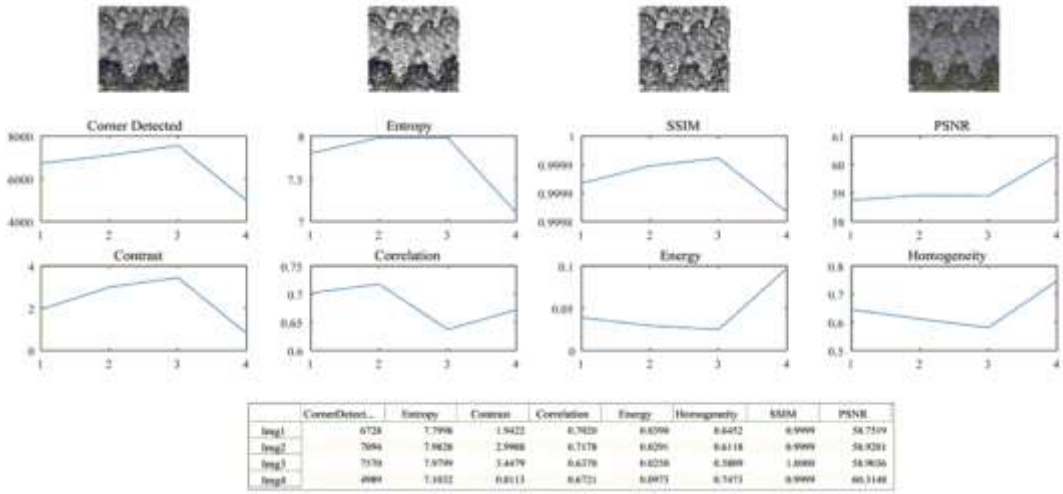


Figure 6. Classifying Images using FLD Classification and Calculating the Value of AUC

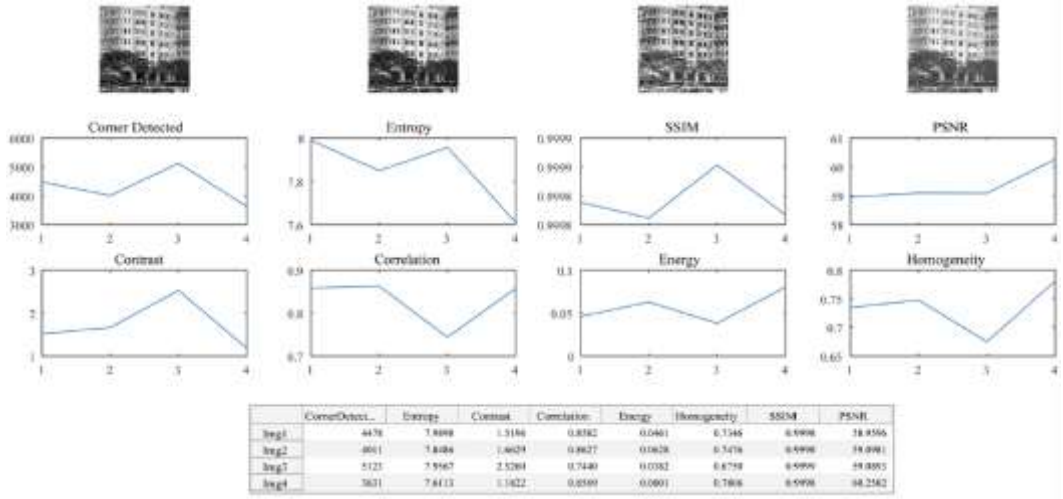
3.2. Experimental Results

First, the experimental results of "Selecting cover images with high security" section have presented. In a practical experience, three images were selected, for each of which four contrasts were considered. The Harris threshold level was considered as 1000. In each contrast, the number of the prominent points of the edges (feature points) would be different. By changing the Harris threshold level, the number of the feature points and other parameters would be changed; furthermore, increasing the number of the feature points would lead to the increase in the embedding rate as well as the changes in each Harris threshold level. Consequently, the security level would be changed. Finally, the above operation has been applied on 5000 images of the BOSS bank.



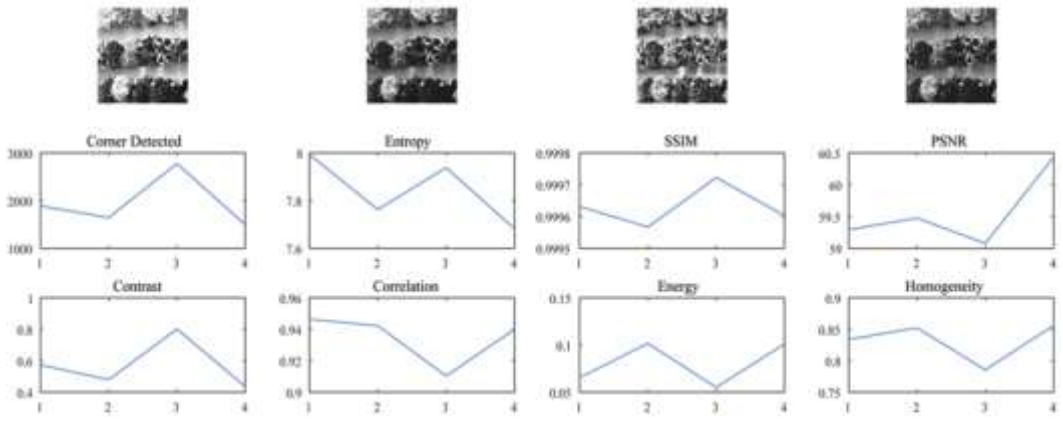
	CornerDetect...	Entropy	Contrast	Correlation	Energy	Homogeneity	SSIM	PSNR
Img1	6728	7.7998	1.9422	0.7020	0.6599	0.8452	0.9999	58.7519
Img2	7694	7.9828	2.9968	0.7178	0.6291	0.8118	0.9999	58.9281
Img3	7170	7.9799	3.4479	0.6370	0.6258	0.7889	1.0000	58.9036
Img4	4989	7.1032	0.8113	0.6721	0.8975	0.7473	0.9999	60.3140

(a)



	CornerDetect...	Entropy	Contrast	Correlation	Energy	Homogeneity	SSIM	PSNR
Img1	4478	7.8998	1.5186	0.8582	0.6461	0.7346	0.9998	58.9566
Img2	4011	7.8486	1.4629	0.8627	0.6628	0.7478	0.9998	59.0081
Img3	5123	7.9567	2.5289	0.7440	0.6758	0.6758	0.9999	59.0893
Img4	3631	7.6413	1.1622	0.8569	0.6801	0.7806	0.9998	60.2582

(b)



	CornerDetect...	Entropy	Contrast	Correlation	Energy	Homogeneity	SSIM	PSNR
Img1	1888	7.9932	0.5734	0.9463	0.8653	0.8330	0.9996	59.2929
Img2	1641	7.7621	0.4787	0.9422	0.8697	0.8522	0.9996	59.4715
Img3	2777	7.9382	0.8813	0.9165	0.8550	0.7854	0.9997	59.0742
Img4	1490	7.6813	0.4349	0.9399	0.8852	0.8350	0.9996	60.4309

(c)

Figure 7. Results of the Quite Similar Behaviors of Image Features (a) 1, (b) 2, and (c) 3; with Different and Sorted Contrasts Based On Contrast and Energy Features of Each Image

As seen in the tables in Figure (7), the entropy and feature points (prominent points near the edges) as well as the other parameters have been changed with regard to the changes in the type of the image contrast. The diagram behaviors of each criterion in different contrasts of each image are almost similar. To compare the method, using the Jsteg algorithm, the information was embedded in the cover images with capacity of 25%, and the last two columns of each table are related to the PSNR and SSIM parameters of image. Although by reducing the contrast, the values of both PSNR and SSIM parameters of image are increased, but in case of a sudden increase in the contrast and reduction in the energy, the two PSNR and SSIM parameters of image will be still increased. Thus, at this stage, the images are sorted based on the LC[19] or HC-HE[20] in each image. According to the 5000 used images of the BOSS bank, two banks 5000 images will created each for LC and HC-LE.

In this part, Experimental results of the effect of the given measures on selecting the appropriate cover ("Selecting cover image based on steganalysis" section) have provided. According to the flowchart in Figure (6), that be applied once on the LC bank and once on the HC-LE bank, the administrative procedure was performed as following:

1. First, 5000 images (512×512) of the bank prepared in the first step of the proposed method were used;
2. The Jsteg embedding was performed on all the images with rate of 25% (0.25 bits per pixel);
3. All the mentioned criteria are calculated;
4. Difference of the criteria related to the cover image and the image containing steganography are calculated;
5. Values of the previous paragraph are arranged in an ascending order, and then 3000 cover images as well as the stego image related to the minimum difference of the criteria are selected;
6. 1500 images are used for training and the other half for testing;
7. From these images, the feature vectors are extracted using WAM steganalysis;
8. Using the FLD classification, the AUC values are calculated for each mode. It should be noted that the larger the AUC values of a classifier, the more favorable the efficiency of that classifier. As previously mentioned, the closer to 0.5 the AUC value, the poorer the efficiency of the steganalyzer;
9. In the next step, 3000 images are selected randomly for classification, and the AUC values are calculated;
10. The obtained results are shown in Tables (2) and (3):

Table 2. AUC Values Obtained for WAM Steganalyzer Regarding to the Given Criteria in Jsteg Embedding (LC bank)

Criteria of selecting cover image	AUC value
D	0.643
En	0.712
MP	0.758
Corr	0.792
IDM	0.8
C	0.81
E	0.813
Random Select	0.827

[19] Low Contrast
[20] High Contrast - Lowe Energy

Table 3. AUC Values Obtained for WAM Steganalyzer Regarding to the Given Criteria in Jsteg Embedding (HC-LE bank)

Criteria of selecting cover image	AUC value
D	0.643
En	0.742
MP	0.773
Corr	0.817
IDM	0.821
C	0.83
E	0.847
Random Select	0.859

As shown in Table (2) and (3), in case of selecting the cover randomly, then we will have $AUC_{LC} = 0.827$ and $AUC_{HC-LE} = 0.859$, meaning that the sender has embedded his secret message with no knowledge of the cover image. However, with an overlook at the values of the other criteria, the decreased efficiency of the WAM steganalyzer compared to the random selection mode would be perceivable. The best criterion in the above table is divergence (D) that has the lowest values; thus, considering the divergence criterion would increase the secure cover selection in the Jsteg (Eq. 16).

$$\begin{aligned} If\ AUC^{D}_{(WAM,LSB)} \downarrow \therefore\ & \Pr\left[W^{AUC^{D}_{(WAM,LSB)_C}} = 1\right] \cong \Pr\left[W^{AUC^{D}_{(WAM,LSB)_S}} = 1\right] \\ & \Longrightarrow Adv_{C,S}\left(W^{AUC^{D}_{(WAM,LSB}}\right) \\ & =: \Pr\left[W^{AUC^{D}_{(WAM,LSB)_C}} = 1\right] - \Pr\left[W^{AUC^{D}_{(WAM,LSB)_S}} = 1\right] = \varepsilon \cong 0 \\ & \Longrightarrow SoI_C \uparrow \end{aligned} \tag{16}$$

Regarding the correlation between the adjacent pixels, considering up to 3 adjacent pixels for extracting the features would lead to the considerable improvement in the detection results. Comparing the proposed method on different image bases shows that the applied features have appropriate stability relative to the input image, while the methods such as WAM and SPAM lack such stability, indicating the reliability of these features.

4. Conclusion

An appropriate cover selection for steganography is an important issue in this field. In this paper, using co-occurrence matrices, contrast and energy of the near some corners and edges of the cover image (with Harris constant threshold level) was investigated. With regard the fuzzy logic, security level of each cover was determined. Then, the images with different security levels stored in different image banks. Finally, using FLD, each cover bank was assessed for Jsteg and WAM focusing on the features, Energy, Entropy, Contrast, Inverse difference moment, Maximum probability, Correlation and Divergence.

The secured images bank will cause the security parameter to be considered before the start of steganography embedding. We evaluated two banks with security level, VHS and HS. The results showed that some steganalyzers such as WAM has low chance against some steganographic methods such as JSteg when we focused on the mentioned features.

The results of this paper can be developed for different situations such as other features of images, other security levels and other steganalyzers and steganographic methods.

Reference

[1] S. Nazari and M.-S. Moin, "Cover Selection Steganography via Run Length Matrix and Human Visual System", **(2013)**.
[2] 2. W.-C. Kuo, Y.-H. Chen and C.-T. Chuang, "High-capacity steganographic method based on division arithmetic and generalized exploiting modi_cation direction", Journal of Information hiding and Multimedia Signal Processing.(ISSN 2073-4212) Ubiquitous International, vol. 5, no. 2, **(2014)**.
[3] 3. H. Sajedi and M. Jamzad, "Secure cover selection steganography", in International Conference on Information Security and Assurance, Springer, vol. 326, **(2009)**, pp. 317.
[4] A. M. E. Moghadam, "A high quality image steganography scheme based on fuzzy inference system", **(2013)**.
[5] M. Kharrazi, H. T. Sencar and N. Memon, "Cover selection for steganographic embedding", in 2006 International Conference on Image Processing, IEEE, vol. 120, **(2006)**, pp. 117.
[6] H. Sajedi and M. Jamzad, "Contourlet-Based Steganography Using Cover Selection", International Journal of Information Security, Springer, vol. 9, no.5, **(2010)**, pp. 337-345.
[7] W. Wang and H. Farid, "Exposing digital forgeries in video by detecting double quantization", in Proceedings of the 11th ACM workshop on Multimedia and security, ACM, vol. 48, **(2009)**, pp. 39.
[8] M. Sahraee-Ardakan and M. Joneidi, "Joint dictionary learning for example-based image super-resolution", arXiv preprint arXiv:1701.03420, **(2017)**.
[9] M. Liskiewicz, R. Reischuk and U. W•olfel, "Security levels in steganography-insecurity does not imply detectability", in Electronic Colloquium on Computational Complexity (ECCC), vol. 22, **(2015)**, p. 10.
[10] J. Fridrich, "Steganography in digital media: principles, algorithms and applications". Cambridge University Press, **(2010)**.
[11] O. Goldreich, S. Goldwasser and A. Nussboim, "On the implementation of huge random objects (preliminary version)", **(2003)**.
[12] M. Vuk and T. Curk, "Roc curve, lift chart and calibration plot", Metodoloski zvezki, vol. 3, no. 1, **(2006)**, pp. 89{108}.
[13] T. G. Tape, "The area under an roc curve", Interpreting diagnostic tests, **(2006)**.
[14] R. Gonzalez and R. Woods, "Digital image processing: Pearson prentice hall", Upper Saddle River, NJ, **(2008)**.
[15] H. Sheisi, J. Mesgarian and M. Rahmani, "Steganography: Dct Coefficient Replacement Method and Compare With JSteg Algorithm", International Journal of Computer and Electrical Engineering, vol. 4, no. 4, **(2012)**.
[16] C. Zhou, X. Wei, Q. Zhang and X. Fang, "Fishers linear discriminant (d) and support vector machine (svm) in non-negative matrix factorization (nmf) residual space for face recognition", Optica Applicata, vol. 40, no. 3, **(2010)** , pp. 693{704}.
[17] J. H. Rivera, Z. Harchaoui and F. De la Torre, "Instance-selecting regularization penalty for supervised image classifcation", **(2010)**.
[18] M. Conos, "Recognition of vehicle make from a frontal view", Master, Czech Tech. Univ., Prague, Czech Republic, **(2006)**.
[19] A. Kaur and S. Kaur, "Image steganography based on hybrid edge detection and 2 k correction method", International Journal of Engineering and Innovative Technology (IJEIT), vol. 1, no. 2, **(2012)**.
[20] C. Harris and M. Stephens, "A combined corner and edge detector", in Alvey vision conference, vol. 15, p. 50, Citeseer, **(1988)**.
[21] D. Simitopoulos, D. Koutsonanos and M. G. Strintzis, "Image watermarking resistant to geometric attacks using generalized radon transformations", in Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on, IEEE, vol. 1, **(2002)**, pp. 85{88}.
[22] P. Singla, "Designing and performance evaluation of an advanced method for corner detection using harris technique".
[23] M. Kharrazi, H. T. Sencar and N. Memon, "Improving steganalysis by fusion techniques: A case study with image steganography", in Transactions on Data Hiding and Multimedia Security I, Springer, **(2006)**, pp. 123{137}.

Authors

Reza Esfahani, he is a PhD. student at IHU. He has B.Sc. in Electrical engineering Communication and M.Sc. in Secure Communication. His research interests include Steganography, Image Processing and Cryptography. He has published several technical papers and participated in many scientific conferences.

Zynolabedin Noroozi, he received the BS and MS degrees in applied Mathematics in 2004 and 2007 from Tehran University. He has Ph.D degree in applied mathematics - cryptography in 2012 from Kharazmi University. He worked as a cryptography and steganography. He has published several technical papers and participated in many scientific conferences.

Gholamreza Jandaghi, he is a professor of statistics at University of Tehran. He has B.Sc. in mathematics and M.Sc. and Ph.D. in biostatistics. His research interests include statistical methodology, quantitative modeling in management, data mining and research methodology. He has published more than 200 papers in his areas of interest