

## A NEW APPROACH FOR BLOCK BASED STEGANALYSIS USING A MULTI-CLASSIFIER

S.L. Omrani<sup>1</sup> P. Bayat<sup>2</sup>

1. Department of Computer Engineering, Payame Noor University, Iran, [omrani\\_comput@gilan.pnu.ac.ir](mailto:omrani_comput@gilan.pnu.ac.ir)

2. Department of Software Engineering, Gilan Science and Research Branch, Islamic Azad University, Iran  
[peymanebayat@gmail.com](mailto:peymanebayat@gmail.com)

**Abstract-** The aim of the recent block-based steganalysis approaches, is to detect and differentiate cover and stego images. In this research, the discovering capability of the steganography algorithm was used for a sample stego image by designing a multi classifier. This kind of classification was used for the steganalysis of smaller blocks of an image. Because normal images mostly have heterogeneous regions, first, the main image was decomposed into smaller blocks which were similar. Then, these similar blocks were put in the same class. Therefore, several different classes were obtained, for each of which, an appropriate classifier was identified. This approach led to making a decision for identifying the situation of the image blocks that were either cover or stego and identifying the steganography algorithm used in stego images.

**Keywords:** Steganalysis, Steganography, Multi Classifier, Stego, Cover.

### I. INTRODUCTION

Blind steganalysis identifies stego images from cover images without any knowledge about steganography embedding algorithms [1]. Most of previous works in this field have focused on extracting features from images for the purpose of steganalysis using a binary classifier, which identifies stego images from cover images [2, 3, [4]. In the multi classifier scenario, first, blind steganalysis tries to make a decision about the kind of cover or stego of a sample image. If the image is determined as a stego, the multi classifier can determine the used steganography algorithm.

Pevny et al. [4] used 274 merged features for classification. In their research, blind steganalysis classified sample images into 7 different classes, in which stego images were obtained by 6 different steganography algorithms. A multi classifier was arranged by combining several binary classifiers. To classify the sample images into 7 different classes, the "max-wins" strategy (related to binary Support Vector Machine (SVM) classifiers) was used for each pair of classes. Experimental results explained that their approach was effective in terms of classifying the sample images into 7 different obtained

classes. Generally, the performance of blind steganalysis is measured by average detection accuracy as follows [4]:

$$A_{detect} = 1 - P_{error} \quad (1)$$

In this formula,  $P_{error}$  is average error probability. There are two types of errors related to this area, which include false positive (FP) and false negatives (FN). To have a higher detection accuracy, the proposed approach tried to minimize the mentioned block detection errors. When the sample cover image was determined as a stego, it illustrated that a false positive error occurred. In contrast, when a stego image was not correctly detected, a false negative error occurred. Performance of blind steganalysis for a multi classifier can be evaluated through comparing FP and FN errors, as shown in Table 1.

Table 1. Comparing FP and FN errors [4]

Decision	Actual Decision			
	Cover	Stego 1	...	Stego L
Cover	Correct(TN)	Incorrect(FP)	...	Incorrect(FP)
Stego 1	Incorrect(FN)	Correct(TP)	...	Incorrect(FP)
...	...	...	...	...
Stego L	Incorrect(FN)	Incorrect(FP)	...	Correct(TP)

### II. MULTI-CLASSIFIER

A multi-classifier was used to identify which steganography algorithm should be applied for creating the final stego images. If  $I$  were a dataset of images, then  $L+1$  would be the number of images in  $I$  and this set would be obtained as follows:

$$I = \{I_1 = cover, I_2 = stego_1, \dots, I_{L+1} = stego_L\} \quad (2)$$

In this formula, there is one sample cover image and  $L$  stego images in the experimental dataset. Furthermore, the accuracy of final detection was obtained by the average calculation of detection accuracy of the whole existence images in the dataset, consisting of stego and cover images. The following formulate illustrates the final detection accuracy of a multi classifier:

$$A_{detect} = \frac{1}{L+1} (A_{I_1} + A_{I_2} + \dots + A_{I_{L+1}}) \quad (3)$$

where  $A_{I_l}$  is the detection accuracy of the composed multi classifier when real images are assigned to image type  $I_l$  by  $l = 1, 2, \dots, L+1$ .

III. METHODOLOGY

Instead of using one steganography algorithm, the proposed approach used  $L$  different steganography algorithms for creating stego images in testing and training stages. Before applying the major voting rule, for identifying the sample image, two kinds of weights can influence the increasing decision accuracy of each block. All the weights consisted of 1: The weights that depended on different block classes, and 2: The weights that depended on different image types. These two weights were used to achieve an appropriate weight for each block.

For decreasing the dimension of vectors of image features and optimizing the performance of steganalysis algorithm in feature extraction stage, optimal wavelet packet decomposition (WPD) was applied [5]. This stage become after blocking each image, which led to gathering the similar regions of image put in a block. After applying the optimal wavelet packet decomposition on each block, the obtained coefficients were more closed [6]. Figure 1 illustrates the histogram of coefficients after applying the wavelet packet decomposition.

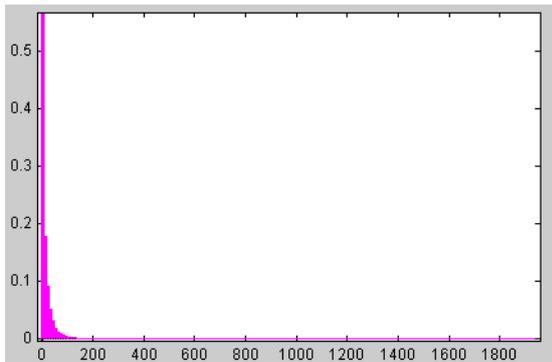


Figure 1. histogram of coefficients after applying WPD

Therefore, after applying the Shannon Entropy Function (SEF) to the achieved coefficients, some more nodes were eliminated from the tree structure and the obtained optimal tree would have less optimal nodes. Thus, feature dimension was decreased in each block.

This solution caused decreased calculation complexity and considerably increased detection accuracy because of choosing the most optimal features. Figure 2(a) and 2(b) illustrate the stage of WPD according to [5]. Figure 3 illustrates the reduction of feature dimensionality after blocking as compared to optimal wavelet packet decomposition [6].

Although there was a direct relationship between increasing the number of feature dimension and calculating complexity, the presence of less optimized features in addition to higher detection accuracy compared with previous approaches, calculation complexity would decrease [6]. Consequently, in the next section, stages of feature extraction about wavelet packet decomposition will be explained in more details.

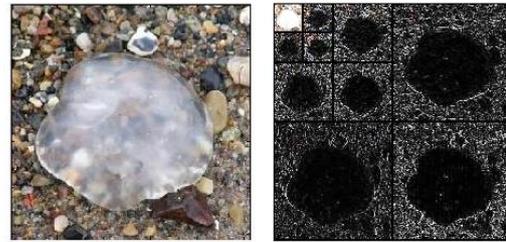


Figure 2. (a) From left to right: Original image, the image after Haar wavelet decomposition [5]

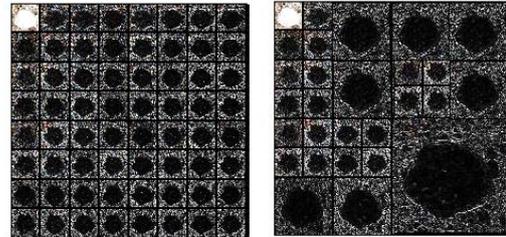


Figure 2. (b) From left to right: The image after complete wavelet packet decomposition and the image after optimal wavelet packet decomposition [5]

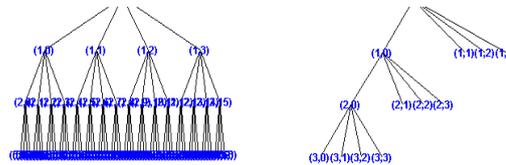


Figure 3. From left to right: the resulted tree from optimal wavelet packet decomposition entropy function before blocking, the same tree after blocking [6]

A. Complete Wavelet Packet Decomposition

Wavelet decomposition decomposes an image into subbands with low and high frequencies. Frequency areas have different resolutions and the analysis of these resolutions is one of the most important concepts in wavelet transform. From this viewpoint, normal wavelet transform with function  $L^2(R)$  is defined according to the following formula [6]:

$$L^2(R) = J \in Z^{\oplus} W_j \tag{4}$$

where  $W_j$  is wavelet space,  $R$  and  $Z$  are two sets of real numbers and integers, and  $L^2(R)$  is the square of integral function in space  $R$ .

In complete wavelet packet decomposition, image coefficients are separated based on various resolutions at different frequencies according to the following steps [5]:

1. First, the three-step Haar wavelet decomposition is applied to the image; according to Figure 4, this action will select 9 detail sub-bands (horizontal  $H_i$ , vertical  $V_i$ , and diametrical  $D_i$ ,  $i=1,2,3$ ) and 3 approximation (low pass) sub bands ( $L_i$ ,  $i=1,2,3$ ). In this stage, these 12 sub-bands are considered in the first set of features.



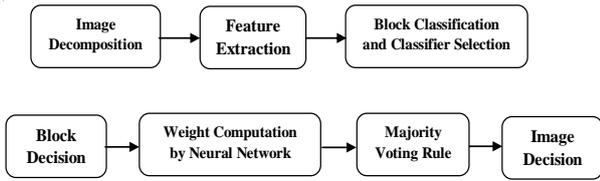


Figure 6. Testing process [11]

### E. Calculating Weights by Neural Networks (NN)

This stage consider before using to major voting rule. Two types of weights are using to make a decision more accurate for each block, including: 1. The weights that depended on different block classes, and 2. The weights that depended on different image types [11]. Both of them were used to determine a bundle of weights to make a decision for each block in terms of identifying cover or stego status. When the cover or stego status was identified for each block by its steganography algorithm, a corresponding weight was allocated according to the class of each block. An appropriate classifier was used to calculate Correct Decision Rate (CDR) for all the  $C$  classifiers. CDR is the weight that will be allocated after identifying the type of an image block. For  $K$  classes where the mentioned blocks are a type of image  $I_l$  by  $l = 1, 2, \dots, L+1$ , the CDR was calculated as follows [14]:

$$CDR_{b=I_l}(K) = P(actual = I_l | decide = I_l) \quad (5)$$

In this formula,  $P(actual = I_l | decide = I_l)$  is the probability of blocks that are decided to be from image type  $I_l$  and are actually from image type  $I_l$ . This measurement was different in terms of detection accuracy and used for correct evaluation of each decision with the actual type of image. Furthermore, detection accuracy could be achieved for images in the training set and in a decision making process. The related weights were assignable to the decision for each block. After achieving  $C$  different classifiers in the training set, it could be used for the same testing set of images to obtain the results of detection accuracy. If the detection accuracy of a special type of an image was low, the next allocated weight to the mentioned type should be increased. In contrast, if the detection accuracy of a special type of an image were high, the next allocated weight to the mentioned type should be decreased.

This routine was performed on a model of a multi-layer perceptron neural network infrastructure [12]. Also, in this paper some of approaches which are based on neural network used as classification optimization. Image features are classified by an unsupervised neural network [13, 14]. The weights for block decision of image type  $I_l$ , were obtained as follows [11]:

$$W_i = I_l = [1 - A_l] + [1 - P_e(decide = I_l | actual \neq I_l)] \quad (6)$$

where  $A_l$ , is the detection accuracy of image type  $I_l$  and  $P_e(decide = I_l | actual \neq I_l)$ , is the error probability making of a decision of image type  $I_l$  when this is not true. The weights relating to the blocks with different classes and image types are shown by [11]:

$$W_{b=I_l}(K) = W_{i=I_l} \times CDR_{b=I_l}(k) \quad (7)$$

where  $W_{b=I_l}(K)$  represents weights for blocks from image type  $I_l$  in the  $K$ th class. These weights are used to identify the importance of block decision using majority voting rule.

### F. Final Decision Using Majority Voting Rule

Assume a  $M \times N$  pixel sample image including  $MN/B^2$  blocks by  $B \times B$  size. Weight of each decision is presented by  $W_{b=I_l}(K)$ , which identifies the importance of the related decision. Therefore, the total number of weighted decision that is made by the mentioned solution is equal to  $MN/B^2$ . After obtaining the value of  $MN/B^2$  as the weighted decision for a sample image, a majority voting rule was adapted to make the final decision on whether a given sample image was a cover or a stego image created from one of the  $L$  mentioned steganography algorithms. The final decision can be made by selecting a cover image or a stego image with a specific steganography algorithm that had the largest sum of weights.

## IV. EXPERIMENTAL RESULT

There are some issues about the setting of implementation environment that are mentioned in the previous related works [11]. UCID [15] and INRIA Holidays [16] databases were used in the implementation of the proposed method. Three ( $L=3$ ) different steganography algorithms were used to embed a secret message in the cover images to create the corresponding stego images: OutGuess (OG) [17], F5 [18], and Model-Based Steganography (MBS) [19]. Some differences existed between the newly presented approach and that of Cho et al.'s research [8]. In this paper, the number of sample blocks was 30000, because this number of experiments showed that the result went to a steady-state behaviour. On the other hand, in this research, the whole set of sample blocks was classified into 16 classes. Furthermore, at feature extraction stage, instead of the Markov and Discrete Cosine Transform (DCT) features, optimal wavelet packet decomposition was applied to each block. Ultimately, the embedding ratio in both of the works was the same.

## V. RESULTS AND DISCUSSION

In this section, the accuracy of steganalysis detection was reported based on blocking for a multi classifier. Table 2 shows a comparison between Pevny et al.'s [4] and Cho et al.'s [11] works and the newly presented approach where used to embedding rate equal to 0.2. For example in this table, the comparison shows that the precision of F5 algorithm in Cho et al. was around 10% less than that in the method by Pevny et al., while the detection precision of the new proposed method was much higher than both of the above methods. As illustrated in Table 2, the precision of Pevny's, Cho's and the proposed methods was 67.61, 58.82, and 71.35, respectively.

This optimization can be explained from different perspectives. The first reason may refer to the advantage of wavelet packet transformation and higher precision of this method in selecting more appropriate features for the classifier, as explained in [8]. The second reason can be related to error detection. As mentioned in introduction, there are two types of errors in the decision-making process which are statistically important: The first one is false-positive (FP) and the second one is false-negative (FN). FP happens when a secret message is detected in the cover image and FN occurs when a hidden message is not recognized in the stego image. Moreover, true-negative (TN) is the correct detection of the absence of hidden message in the cover image and true-positive (TP) occurs when a hidden message is correctly detected in the stego image. Table 1 illustrates the occurrence of these states.

Universal steganalysis algorithm attempts to push the occurrence of these two types of errors to have a minimum rate in order to obtain higher accuracy. Accordingly, comparing FP and FN values could be used to justify the increased accuracy of the proposed method. For example, as illustrated in Table 2, FP error in Pevny et al.'s [4] method was 18.3 on average; in Cho et al. [11], it was averagely 8.95; and, it was 4.51 on average in the proposed method. The proposed method's detection error rate was decrease in comparison to Pevny's and Cho's ethods in the same condition and same ratios.

Also, as illustrated in Table 2, FN error in Pevny et al.'s [4] method had the average of 9.23; in Cho et al. [11], it was averagely 12.25; and it had the average of 6.78 in the proposed method. The proposed method's detection error rate was increase in comparison to Pevny's and Cho's methods; i.e. accuracy of the proposed method was higher than that of the mentioned above methods by the same proportions.

In addition, accuracy of the proposed method for stego images created by OutGuess and MBS algorithms increased by more than 15%. Finally, the overall accuracy increased from 63.63%, 70.93%, and 80.04% in Pevny et al.'s [4], Cho et al.'s [11] and the new proposed method respectively. Figure 7 illustrates the improvement accuracy detection of proposed method in embedding rate 0.2 according to Table 2.

Table 3 compares the proposed method with those of Pevny et al. [4] and Cho et al. [11] with the embedding rate of 0.3. As observed, the accuracy detection for this embedding rate was higher. Classification of testing images into different classes was easier when the length of the embedded hidden messages was longer. Recognition accuracy of all stego images was improved by the proposed method.

Maximum improvement occurred when embedding rate was 0.2 and OutGuess steganography algorithm was applied to embed the message. Recognition accuracy of the proposed method was improved altogether, as compared to that of Pevny et al.'s and Cho et al.'s methods. Comparison of the proposed method with that of Cho et al.'s indicated a significant improvement in both 0.2 and 0.3 embedding rates, which may be because the

blocking procedure had the advantage of decomposing relatively similar blocks in the image into smaller sized blocks. According to the generated classes from the blocks, a variety of classifiers was designed to extract the features of the blocks in different classes.

Furthermore, weights were derived based on the classes of different blocks and, eventually, all the images used in the majority voting rule improved the overall precision of the proposed method.

Table 2. Comparing detection accuracy (embedding rate: 0.2)

Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Detection Accuracy in Pevny et al. [4] 's approach :%63.63</b>				
Cover Image	72.30	16.16	10.13	28.71
Out Guess	2.41	57.14	2.55	5.77
F5	5.70	3.76	67.61	8.05
MBS	19.58	22.94	19.72	57.48
Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Decision Accuracy in Cho et al. [11] 's approach :%70.93</b>				
Cover Image	63.25	0.27	21.93	4.63
Out Guess	2.35	83.03	3.49	8.05
F5	20.39	2.01	58.82	8.72
MBS	14.02	14.49	15.76	78.60
Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Decision Accuracy in proposed approach :%80.04</b>				
Cover Image	75.54	0.07	9.35	4.13
Out Guess	1.01	91.93	0.04	1.02
F5	10.31	0.01	71.53	0.07
MBS	9.03	2.01	3.06	81.17

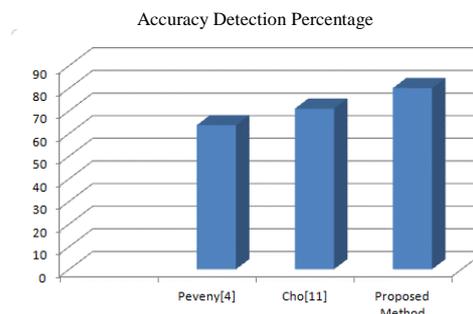


Figure 7. improvement accuracy detection of proposed method by embedding rate 0.2

Table 2. Comparing detection accuracy (embedding rate: 0.3)

Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Detection Accuracy in Pevny et al. [4] 's approach :76.79%</b>				
Cover Image	81.76	6.84	5.90	20.46
Out Guess	1.48	76.73	0.94	3.82
F5	3.55	2.15	78.67	5.70
MBS	13.21	14.29	14.29	70.02
Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Decision Accuracy in Cho et al. [11] 's approach :86.23%</b>				
Cover Image	77.33	0.00	8.58	3.02
Out Guess	1.01	94.90	1.01	4.43
F5	16.30	0.74	86.32	6.17
MBS	5.37	4.36	4.09	86.38
Algorithm's Decision	Actual Decision			
	Cover Image	Out Guess	F5	MBS
<b>Evaluation of Decision Accuracy in proposed approach; 90.40%</b>				
Cover Image	85.01	0.9	2.05	0.16
Out Guess	0.05	94.93	0.04	1.22
F5	8.03	0.01	91.53	0.17
MBS	4.17	1.17	1.06	90.16

## VI. CONCLUSIONS

Since the purpose of steganalysis is to distinguish cover images from stego ones and as this is only a small part of steganalysis, by designing a multi-objective classifier after defining a typical image as stego, in this research also detected the steganography algorithm applied to create the stego image. This type of classification operates according to the results from the steganalysis of smaller blocks. Thus, by using blocking method, images were divided into smaller blocks with relatively similar areas. In addition, by increasing the accuracy of decisions about each block, before applying the majority voting rule, two types of weighting were used for the blocks which consisted of: 1) Weights related to different classes, and 2) Weights associated with different steganography methods. Using these two types of weighting can help in making a more accurate decision on each block. Overall accuracy of the proposed method, as compared with the previous techniques, increased and reached to 90.40%.

## REFERENCES

[1] I.J. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, "Digital Watermarking and Steganography", Morgan Kaufmann Publication, 2007.

[2] J. Fridrich, "Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes", Proceeding ACM International Workshop on Information Hiding, Toronto, Canada, May 2004.

[3] Y.Q. Shi, C. Chen, W. Chen, "A Markov Process Based Approach to Effective Attacking JPEG Steganography", ACM International Workshop on Information Hiding, Old Town Alexandria, Virginia, July 2006.

[4] T. Pevny, J. Fridrich, "Merging Markov and DCT Features for Multiclass JPEG Steganalysis", SPIE Conference Security, Watermarking, and Steganography, Vol. 6505, San Jose, California, Feb 2007.

[5] X.Y. Luo, F. Liu, C. Yang, D. WANG, "Image Universal Steganalysis Based on Best Wavelet Packet Decomposition", Science China Information Sciences, Vol. 53, No. 3, pp. 634-647, Berlin, Germany, March 2010.

[6] L. Omrani, K. Faez, "JPEG Image Steganalysis Using Block Based Optimal Wavelet Packet Decomposition", 6<sup>th</sup> International Symposium on Telecommunications (IST'2012), Tehran, Iran, November, 2012.

[7] X.Y. Luo, F. Liu, D. Wang, "WPD Based Blind Image Steganalysis", Journal on Communications, Issue 9, Vol. 29., No. 10, pp. 173-182, 2008..

[8] Y. Wang, P. Moulin, "Optimized Feature Extraction for Learning Based Image Steganalysis", IEEE Transaction Information Forensic Security, Issue 1, Vol. 2, No. 4, pp. 31-45, 2007.

[9] H. Farid, "Detecting Hidden Messages Using Higher Order Statistical Models", IEEE International Conference on Image Processing, Vol. 2, pp. 905-908, New York, USA, 2002.

[10] S. Cho, B. Cha, J. Wang, C.C. Jay Kuo, "Block-Based Image Steganalysis: Algorithm and Performance

Evaluation", IEEE International Symposium Circuits and Systems, pp. 1679-1682, Paris, France, May 2010.

[11] S. Cho, J. Wang, C.C. Jay Kuo, B. Cha, "Block Based Image Steganalysis for a Multi-Classifer", IEEE International Conference on Multimedia and Expo, pp. 1457-1462, Singapore, 2011.

[12] K.L. Du, "Clustering: A Neural Network Approach", Elsevier Neural Networks, Issue 1, Vol. 23, pp. 89-107, January 2010.

[13] I. Kanellopoulos, G.G. Wilkinson, "Strategies and Best Practice for Neural Network Image Classification", International Journal of Remote Sensing, Issue 4, Vol. 18, pp. 711-725, 2010.

[14] W. Hachicha, A. Ghorbel, "A Survey of Control Chart Pattern Recognition Literature (1991-2010) Based on a New Conceptual Classification Scheme", Elsevier Computers & Industrial Engineering, Issue 1, Vol. 63, pp. 204-222, August 2012.

[15] G. Schaefer, M. Stich, "UCID - An Uncompressed Colour Image Database", SPIE Storage and Retrieval Methods and Applications for Multimedia, Vol. 5307, pp. 472-480, 2004.

[16] H. J'egou, M. Douze, C. Schmid, "Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search", 10<sup>th</sup> European Conference on Computer Vision, pp. 304-317, Marseille, France, October 2008.

[17] N. Provos, "Defending Against Statistical Steganalysis", 10<sup>th</sup> USENIX Security Symposium, Vol. 10, pp. 323-336, Citeseer, 2001.

[18] A. Westfeld, "F5 - A Steganographic Algorithm: High Capacity Despite Better Steganalysis", ACM International Workshop on Information Hiding, Pittsburgh, PA, April 2001.

[19] P. Sallee, "Model Based Steganography", International Workshop on Digital Watermarking, pp. 154-167, Seoul, Korea, October 2003.

## BIOGRAPHIES



**Seyedeh Leila Omrani** received her B.Sc. and M.S.E. degrees in Computer Engineering from IAU of Qazvin in 2007 and 2012, respectively. Currently, she is a Lecturer at Department of Computer Engineering, Payame Noor University, Iran. Her researches are

focused on steganography and steganalysis techniques. She has published several papers and performed some projects in the mentioned area since 2010.



**Peyman Bayat** received his and B.Sc. and M.Sc. degrees from Islamic Azad University, Iran in 1998 and 2003, respectively, and the Ph.D. degree in Computer Engineering Systems from University Putra Malaysia, Malaysia in 2013.