In the Name of God

Journal of
Information Systems & Telecommunication
Vol. 2, No. 4, October-December 2014, Serial Number 8

Indexed in:

Acknowledgement

Table of Contents

A Study on Clustering for Clustering Based Image De-Noising

Hossein Bakhshi Golestani*
Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
h.b.golestani@gmail.com

Mohsen Joneidi
Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, USA
Joneidi@knights.ucf.edu

Mostafa Sadeghii
Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
m.saadeghii@gmail.com

Abstract

In this paper, the problem of de-noising of an image contaminated with Additive White Gaussian Noise (AWGN) is studied. This subject is an open problem in signal processing for more than 50 years. Local methods suggested in recent years, have obtained better results than global methods. However by more intelligent training in such a way that first, important data is more effective for training, second, clustering in such way that training blocks lie in low-rank subspaces, we can design a dictionary applicable for image de-noising and obtain results near the state of the art local methods. In the present paper, we suggest a method based on global clustering of image constructing blocks. As the type of clustering plays an important role in clustering-based de-noising methods, we address two questions about the clustering. The first, which parts of the data should be considered for clustering? and the second, what data clustering method is suitable for de-noising.? Then clustering is exploited to learn an over complete dictionary. By obtaining sparse decomposition of the noisy image blocks in terms of the dictionary atoms, the de-noised version is achieved. In addition to our framework, 7 popular dictionary learning methods are simulated and compared. The results are compared based on two major factors: (1) de-noising performance and (2) execution time. Experimental results show that our dictionary learning framework outperforms its competitors in terms of both factors.

Keywords: Image De-Noising; Data Clustering; Dictionary Learning; Histogram Equalization and Sparse Representation.

1. Introduction

We consider the problem of estimating a clean version of an image contaminated with Additive White Gaussian Noise (AWGN). A general approach to this aim is division of the noisy image into some (overlapping) small blocks, then de-noising of each block and finally obtaining the overall estimation of the clean image by averaging the de-noised blocks. The model is as follows:

$$y_i = z_i + n_i \tag{1}$$

where y_i is the vector form of the ith block of the noisy image, z_i is the vector form of the ith block of the original image, and n_i is a zero-mean AWGN with variance σ^2. Throughout the paper, the blocks are $n \times n$, thus the vector space dimension is n^2.

Image de-noising is still an open problem and numerous methods have been suggested up to now. The methods are based on defining a neighborhood for each block and weighted averaging according to suitable weights. The weights are computed in each neighborhood, as in [1-4] which are some relatively successful approaches. All of them are in the spatial domain. The method in [5] can be considered as same as [1-4], where processing is conducted in frequency domain. This method constructs a three-dimensional matrix by grouping those blocks that are similar (in some senses, e.g. ℓ_2 norm) with a block of the image. Corresponding to each block of the image a group of similar blocks should be found. In this way, a three-dimensional matrix is obtained corresponding to each block. Then, a 3D collaborative signal filtering in the frequency domain is performed which provide a good estimation of the clean version of each block. This method can be considered as the state of the art method of image de-noising; however it suffers from high computational complexity due to local processing. The work in [6] has the same approach and applied filtering in the Principal Component Analysis (PCA) transform domain. Elad and Aharon [7] have suggested a new approach. They have used K-Singular Value Decomposition (K-SVD), which is a dictionary learning algorithm, to produce a global dictionary using the noisy image blocks. This method uses the representation in terms of the dictionary to de-noise image. The estimate of each de-noised block can be estimated by analyzing noisy blocks in this dictionary and applying a sparse recovery algorithm.

Local and global methods have some advantages and disadvantages. A global dictionary can recover general characteristics of an image, which are repeated in its several regions. However, these methods are not able to recover special local textures and details in an image.

---

* Corresponding Author

While local methods indicate higher efficiency in recovering local details of image, they encounter over-learning risk leading from noise learning and incorporating noise into the final result. Deficiency of learning in some regions is another problem of local methods.

In [8], a clustering-based method was suggested. This method produces a local dictionary by clustering feature vectors from all noisy image blocks and conducts de-noising using decomposition of noisy blocks in terms of representatives of the found clusters. Similar to K-SVD, this method is based on dictionary but it uses a local dictionary.

Local patching and similar blocks clustering are effective factors in success of methods including [5], [6] and [8]. Dictionary learning based de-noising methods also perform some type of blocks clustering, for example K-SVD is a generalization of K-means clustering algorithm. So it is necessary to consider the clustering for the de-noising application more closely.

In this paper, we propose an approach for constructing a global dictionary and de-noising based on sparse decomposition of noisy blocks over the dictionary. This global dictionary is constructed by aid of the optimized clustering that will be presented. In the following sections clustering of image blocks is studied with more details in section 2. An analytical comparison between local and global clustering is addressed in section3. Section 4 studies the effect of equalization of data according to their variance in order to have an appropriate clustering. Learning the dictionary is explained in section 5 based on representatives of the found clusters. Section 6 studies applying of de-noising using dictionary. Finally, the local and global methods are evaluated in section 7.

2. Clustering of Image Blocks

In the case of methods including LPG-PCA, KLLD, BM3D ([6], [8] and [5], respectively), grouping of similar blocks is their critical factor of success. So, blocks grouping may has details which should be considered specifically. BM3D and LPG-PCA perform de-noising by clustering of the set of image blocks. K-LLD method performs clustering on feature vector extracted from surrounding blocks (Corresponding to each block). Considering the number of pixels and feature vector dimension, this clustering is of high computational load. In addition to high computational load, unbalanced clustering is one of the problems of global clustering of blocks. This problem is shown in Figure 1.

Assume that in Figure 1-bottom, the goal is to find 2 means. K-means algorithm finds two datacenters indicated by violet circles. These points are not good representatives of the blocks corresponding to the image edges. However, clustering objective function is minimized by these centers. Dense (high number data) correspond to image smooth parts and scattered (low number data) correspond to blocks containing edge or special texture. Traditional clustering algorithms behave

with data corresponding to high energy areas as outlier data. So, these blocks have minor effect on the training by common clustering methods and the final desirable result will not be obtained. To solve the problem, first limitations of clustering-based de-noising methods are examined.

The MSE error lower bounds for image de-noising have been examined in [9] and [10]. This lower bound for one n × ncluster block is calculated as follow.

$$E[\|z_i - \hat{z}_i\|^2] \geq Trace[(J_i + C_z^{-1})^{-1}] \qquad (2)$$
$$C_z = C_y - \sigma^2 I \qquad (3)$$
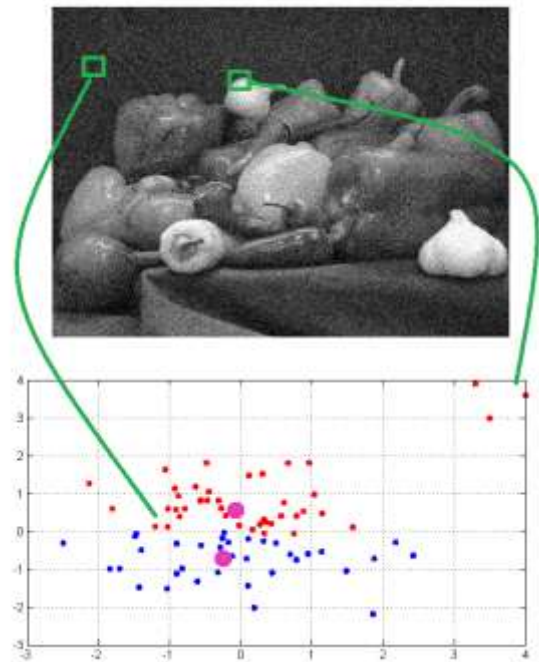
Fig 1. In natural images, number of smooth blocks are more than high energy ones.

where, J_i is the Fisher information matrix and C_z is the estimated covariance matrix for the group of vectors that are similar to ith block. For zero mean Gaussian noise, [10] assumed matrix J_i as follow:

$$J_i = \frac{N_i}{\sigma^2} I \qquad (4)$$

where, N_i is the number of similar vectors of the$i$th block. Assuming that similar vectors for each pixel are of many members and noise level is not high, the right hand of inequality is simplified:

$$\left(J_i + \hat{C}_z^{-1}\right)^{-1} \cong \frac{\sigma^2}{N_i}\left(I - \frac{\sigma^2}{N_i}\hat{C}_z^{-1}\right) \qquad (5)$$

$$E[\|\mathbf{z}_i - \hat{\mathbf{z}}_i\|^2] \geq \frac{\sigma^2}{N_i} Trace(I - \frac{\sigma^2}{N_i}\hat{C}_z^{-1}) \qquad (6)$$

$$E[\|z_i - \hat{z}_i\|^2] \geq \frac{\sigma^2}{N_i}(n^2 - \frac{1}{N_i}\sum_{j=1}^{n^2}\frac{\sigma^2}{\lambda_j}) \tag{7}$$

where λ_j is the jth eigenvalue of covariance matrix of estimated data $\hat{C}_z$:

$$\lambda = eig(\hat{C}_z) = eig(C_y) - \sigma^2 \tag{8}$$

Assuming that the number of similar patches of each block and the noise level is the same for all blocks; thus de-noising bound is related to covariance matrix. High detailed clusters (having high covariance matrix eigenvalues) are more difficult to de-noise. So for blocks corresponding to low complex areas, lower bound will be decreased for MSE of the estimated version and the original image. However the result is predictable; because in smooth areas of an image, a simple averaging can obtain good result but if a block consists of more complexity, specific texture and high variance, would limit de-noising performance. For such blocks, more precise similar block grouping is needed. The more the number of same blocks causes the more appropriate characteristics of grouping. So we suggest that for detailed and textured blocks, more training data should be used.

Let us generalize the concept presented in (2) to clusters (rather than groups for each block). Assume variable i is allocated for clusters rather than blocks in (2). In other words, z_i is a block from the ith cluster and N_i is the number of members of the ith cluster. C_z is the estimated covariance matrix of the ith cluster.

First question that this paper is going to answer is "which blocks should be considered for clustering?" As stated before, using all blocks for clustering not only have high computational load but also leads to unbalanced clustering. Figures 4 and 5 illustrate the idea of equalized clustering. Figure6 is the equalized clustered of Figure 1 providing good properties for de-noising application. Dictionary learning-based methods such as K-SVD decrease training data in a random way to reduce computational load. But as have been seen, removing valuable blocks from training data has negative effect on the de-noising lower bound. In Figure 6, only data corresponding to smooth blocks are removed and the obtained cluster centers are more appropriate for de-noising. In section 3 training data equalization will be studied.

Second question that the paper is going to answer is "how do the clustering?" Now we state the problem of clustering. First we rewrite (2) as follow:

$$E[\|z_i - \hat{z}_i\|^2] \geq \frac{\sigma^2}{N_i}\sum_j\frac{\lambda_j}{\lambda_j + \frac{\sigma^2}{N_i}} \tag{9}$$

Let us write the right side of this inequality for all clusters as a cost function:

$$J(\Omega) = \sum_i\frac{\sigma^2}{N_i}\sum_j\frac{\lambda_j}{\lambda_j + \frac{\sigma^2}{N_i}} \tag{10}$$

Ω is the set of indices of training data that shows membership of the training data to clusters. The problem of the optimum clustering can be stated as follows:

$$\min_\Omega J(\Omega) = \sum_i\frac{\sigma^2}{N_i}\sum_j\frac{\lambda_{ij}}{\lambda_{ij} + \frac{\sigma^2}{N_i}} \tag{11}$$

The above problem is dependent of Eigenvalues of each cluster λ_{ij}, so its computational burden is very high. Thus, exact solution of the problem is not achievable. Eigen values of the clusters corresponding to smooth or constant regions of $\hat{z}$ are about zero so they can be neglected from $J(\Omega)$. So, only high variance blocks affect the cost function.

$$J(\Omega) \cong \sum_{\substack{none\ smooth \\ clusters}}\frac{\sigma^2}{N_i}\sum_j\frac{\lambda_j}{\lambda_j + \frac{\sigma^2}{N_i}} \tag{12}$$

In other words, smooth training data can be ignored in the clustering. At the first glance this simplification just makes the clustering fast but it has an effect on the accuracy of the clustering. In fact, less exploitation of non-important blocks causes in more affection of important blocks in the clustering problem (compare figure 1 and figure 6). Eq. (12) can be interpreted as a hard threshold for selection of blocks in clustering. In the next section variance of blocks will be introduced as a criterion for smoothness and then variance histogram equalization will be presented as the soft threshold version of (12) for selection of data that participate in clustering.

Problem (11) can be viewed from another point of view. The cost function encourages clusters to have a sparse vector of Eigen values. Figure 2 shows how (11) encourages Eigen values to be zero. In other words problem (11) clusters data into low-rank subspaces and guarantees that many of Eigen values will be zero for each cluster.
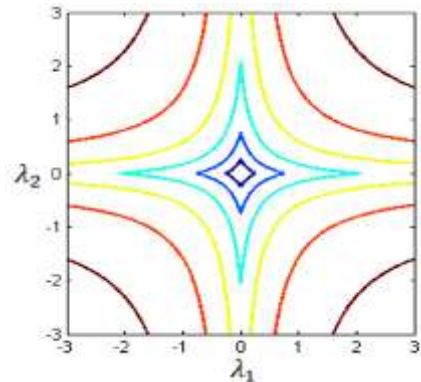
Fig 2. Contour of cost function of (11) for a cluster[1].

---

[1] The figure is contour of $\sum_j\frac{|\lambda_j|}{|\lambda_j| + \frac{\sigma^2}{N_i}}$, as values of λ are positive, figure 2 is true for contour of (11)

High dimensional data that lie in low-rank subspaces have high correlation with each other (see Figure 3). An alternative for subspace clustering may be correlation clustering [11] that has much less computational load. As can be seen in Figure 3, the obtained clusters by correlation clustering lie in a rank-1 subspace that agrees with problem (11) because only one Eigen value of the covariance matrix of this cluster is none-zero. In section 6 simulations has been done by correlation clustering.
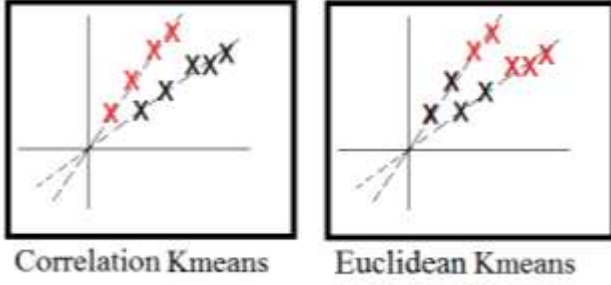
Fig 3. Comparison of correlation clustering and traditional clustering.

3. Global Clustering vs. Local Clustering

A well-known clustering method is the family of K-means clustering algorithms [12], which have been used by K-LLD [8] for image de-noising. K-means clustering algorithm solves the following problem

$$\min_D \sum_{k=1}^{K} \sum_{j \in \Omega_k} \| y_j - d_k \|_2^2 \qquad (13)$$

where, $D = [d_1, \dots, d_k]$. This problem can be written in the following form which is a factorization

$$\min_{D,X} \| Y - DX \|_F^2 , \forall i,j : \| x_i \|_0 = 1, x_i^j \in \{0,1\} \qquad (14)$$

where, Y=[$Y_1, \dots, Y_L$] (L is the number of blocks), x_i is the ith column of X, and x_i^j is the jth entry of x_i. This problem implies that all entries of each x_i must be equal to zero except one of them. The non-zero element is forced to be 1. This restriction does not exist in the so-called gain-shaped variant of K-means [12], which solves the following problem

$$\min_{D,X} \| Y - DX \|_F^2 \ subject\ to \ \forall i : \| x_i \|_0 = 1 \qquad (15)$$

This problem is a K-rank1 subspace (K-lines) clustering. As can be seen in Fig. 4 (b) and (d), the obtained clusters by gain-shaped K-means is in agreement with problem (11). This is because only one eigenvalue of each cluster's covariance matrix is non-zero.

Inspired by the simple approach (15), a suboptimal solution for (11) can be obtained. We propose to construct the proper basis using the obtained cluster centroids and dominant principal components (PCs) of each cluster (generally, natural images are not perfectly lie on rank-1 subspace as in Fig. 4, i.e., thus the proposed dictionary

also contains dominant PCs spanning details of each cluster). Those PCs would be added to the dictionary if their corresponding eigenvalues are greater than noise variance. The noisy image blocks are then de-noised inspired by the framework used in [7]. This leads to a fast and efficient de-noising algorithm (algorithm1). It will be shown in Section 7 that the proposed algorithm outperforms traditional K-SVD.

Algorithm 1 Image de-noising based on gain-shaped K-means
1: **Task** De-noise **Z** from AWGN with variance σ^2
2: Learning K columns of **D** using K-subspace [13]
3: Construct the dictionary by the cluster centroids and most significant PCs
4: **for** $i = 1, \dots, L$ **do**
5:     Sparse code y_i
6:     Estimate z_i by projection on the dictionary
7: **end for**
8: Construct the de-noised image by $\{z_i\}_{i=1}^{L}$

Another approach for clustering is *dictionary learning* in sparse signal representation, which aims to solve the following problem

$$\min_{D,X} \| Y - DX \|_F^2 \ subject\ to \ \forall i : \| x_i \|_0 \leq \tau \qquad (16)$$

K-SVD is a well-known dictionary learning algorithm. Low-rank subspaces found by K-SVD have overlaps. It means that corresponding to each subset of the columns of D, there is a low-rank subspace that K-SVD learns. Data that used the same subset lie on a low-rank subspace but K-SVD learns a very large number of low-rank subspaces for a set of training data such that many of them are empty or low populated (refer to Fig. 5, top). Actually, clusters found by K-SVD include the data that have used the same dictionary columns. Note that these clusters are not guaranteed to be low-rank. In the simulation results we will see that our proposed method based on gain-shaped K-means outperforms K-SVD.

The derived problem (11) describes a suitable global clustering problem, while the state of the art algorithms do not perform global clustering, but instead use local patch-grouping. Translating global clustering to local grouping converts the problem to,

$$G_i = \min_G \| \lambda_G \|_0 \ subject\ to \ |G| \geq \tau, G \in W_i, i \in G \qquad (17)$$

where, G_i is group of blocks corresponding to the ith block, λ_G is the eigenvalues of covariance matrix of G_i and W_i is a window around the ith block. The last constraint implies that the ith block must be member of G_i. An equivalent form of (17) can be stated as,

$$G_i = \max_G |G| \ subject\ to \ \| \lambda_G \|_0 \leq \tau, G \in W_i, i \in G \qquad (18)$$

BM3D, a high performance image de-noising algorithm, implicitly uses (18) in order to perform local grouping. The similarity criterion used in BM3D for

performing local grouping is novel, in which firstly blocks are transformed using an orthonormal transformation (e.g., DCT and DFT), then a projection on a low-rank subspace is performed using hard-thresholding of the coefficients of each block. In the new transformed space, a simple Euclidean distance determines similar blocks with the ith block. Truncated coefficients of the similar blocks with the ith one also lie on a low-rank subspaces near to the ith one, thus many of λ_{G_i} are about zero and the constraint of (18) is satisfied.
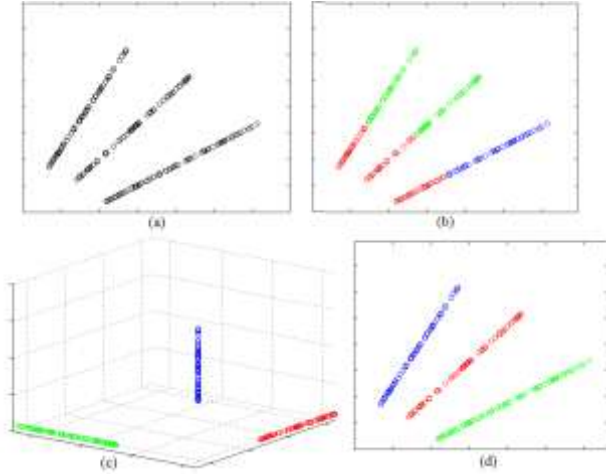
Fig 4. Comparison of clustering in raw data domain and in the sparse-domain transformed data (as used in CSR and LSSC) for some 2D data. (a) Raw data. (b) K-means clustering on raw data (K=3). (c) K-means clustering on sparse-domain transformed data using an over-complete dictionary having 3 atoms. (d) Reconstruction of the data from their sparse representations in (c), in the case of these data Gain-shaped K-means directly results in (d).

The idea behind (18) can be used in another way different from what BM3D has used. These de-noising algorithms first perform grouping using a rough criterion, e.g. Euclidean distance, then in the main de-noising algorithm obtain a low-rank representative for each group and use it. The algorithm suggested by Dong *et al*. (clustering based sparse representation or CSR) [13] which solves the following problem, is an example of these types of algorithms

$$\min_{X,B} \| Y - DX \|_F^2 + \gamma_1 \sum_i \| x_i \|_0 + \gamma_2 \sum_{k=1}^{K} \sum_{j \in G_k} \| x_j - b_k \|_2^2 \qquad (19)$$

where $= [b_k]$, and b_k is the centroid of the kth group. Note that (19) does not optimize the dictionary. In fact, firstly a global dictionary using K-means and PCA is learned which is then used by this problem to simultaneously perform local grouping and sparse coding, in an iterative procedure. The first and second terms in (19) are similar to K-SVD problem, but the last term clusters the sparse-domain transformed data. Figure 5 illustrates the effect of clustering data in the sparse domain rather than the raw data. Contrary to K-SVD, in which the members of a cluster have used one column of D, problem (19) encourages the clustering to put data

that have the same sparse representation (structure) in one cluster.
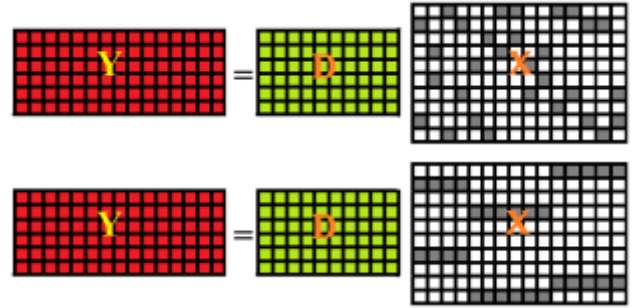
Fig 5: Top: K-SVD approximates data by a union of rank-2 subspaces. No rank-2 cluster can be found. Bottom: Group sparsity constraint on X. There are three rank-2 clusters.

Another local grouping based method is a novel approach, called learned simultaneous sparse coding (LSSC) [14], that simultaneously performs group sparse coding [15] and grouping the similar patches. Group sparse coding implies that the blocks within a group have similar sparse representations, like CSR. This is achieved by jointly decomposing groups of similar signals on subsets of the learned dictionary (as previously explained, K-SVD fails to achieve this goal. See Fig. 5 for comparison). They proposed the following cost function,

$$\min_{X_k} \sum_{k=1}^{K} \| X_k \|_{p,q} \ \ s.t. \ \forall k : \sum_{i \in G_k} \| y_i - D x_i^k \|_2 \le \varepsilon \qquad (20)$$

where, X_k is the coefficient matrix of the kth cluster data, X_j^k is the jth column of X_k, and $\|X\|_{p,q} = \sum_i \|x_{[i]}\|_q^p$, with $x_{[i]}$ the ith row of X. Minimizing $\|X\|_{p,q}$ with $p=1$ and $q=2$ (that is, the ℓ_1 norm of the vector containing the ℓ_2 norms of the rows) implies that the number of engaged rows of X will be limited. In other words, this cost function encourages the data to have the same support of sparse representation in a cluster. As the data in the same cluster can be decomposed by few bases, the rank of the data matrix in the same cluster will be minimized. Thus a solution for (20) tries to minimize (17). i.e, $\Sigma\|X_k\|_{p,q}$ approximates $\|\lambda_{Gk}\|_0$. At the simulation results section, numerical performances of the explained local and global methods are compared, separately.

4. Block Variance Histogram Equalization

For the reasons previously stated some points should be considered. Firstly clusters with different complexities have approximately the same number of members. Secondly, members of complicated clusters should not have high distance from cluster subspace so that covariance matrix eigenvalues would not become high and many of them would be zero. Third, members of high complex clusters should not be neglected for dictionary learning.

Blocks variance is considered as a complexity measure. In natural images, the number of high complex blocks is

lower than low complex blocks. Figure 6 indicates blocks variance histogram of an original image and its noisy version. As can be seen, in the original image, concentration is in lower values of variance and in noisy image concentration is in the point corresponding to noise variance representing smooth blocks of original image. Those blocks that their variances are approximately the same as noise variance are not useful for training. Using these blocks not only increases computational load but also causes unbalance clustering and reduces the effect of important clusters. So their number in final clustering should be reduced. To equalize blocks variance histogram, an equalization transform function must be used. The following function is an example:

$$T(\sigma) \triangleq \begin{cases} \frac{th}{p(\sigma)} & p(\sigma) > th \\ 1 & p(\sigma) < th \end{cases} \qquad (21)$$

where, $p(\sigma)$ is density function of blocks variance probability and th is a threshold. $T(\sigma)$ is the probability of entering a block with variance σ into training data to be used for clustering. Figure 6, indicates an example of the transform function and equalized histogram of noisy image in Figure 7. In this histogram, the effect of blocks with variance 25 is reduced considerably. Figure 8 shows equalized clustering of figure 1.
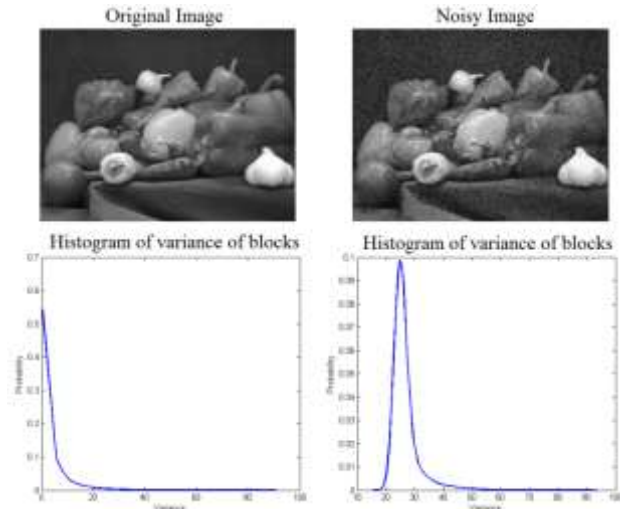
Fig 6. two clear and noisy images with $\sigma = 25$ and their blocks variance histogram.
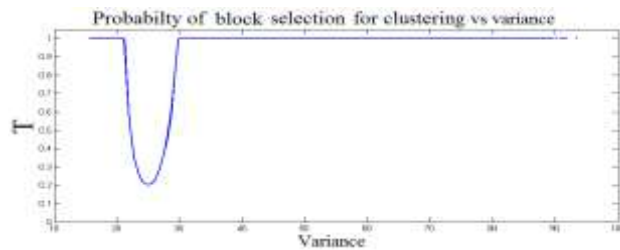
Fig 7. Equalizing transform function

Regarding that the variance of smooth blocks is approximately the same as noise variance. It can be said that there is not valuable information about original image,

and their presence for training not only mislead the clustering algorithm but also have high computational load. Now, subspace clustering should be done on remaining training data which agrees with Eq. (11).
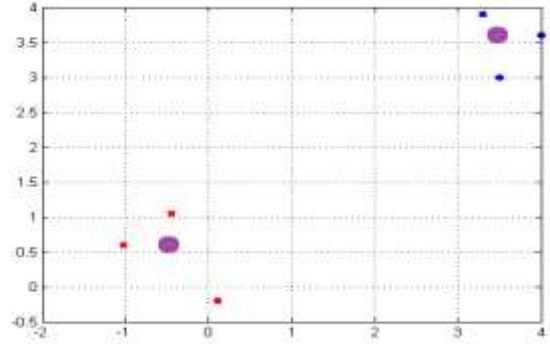
Fig 8. Equalized clustering of figure 1

5. Dictionary Learning

Dictionary learning is performed using the blocks selected in the previous stage. The final dictionary includes P_i dominant principal components from each cluster (equal to non-zero Eigen values of matrix $\hat{C}_z$ explained in Section 2).

In the next stage, SVD transform is derived from covariance of data matrix of each cluster:

$$Y_i Y_i^T = U_i \Lambda_i V_i^T, i = 1 \dots K \qquad (22)$$

where, K is the number of clusters. Singular values on the main diagonal Λ_i are equal to λ_{ij} which are arranged in ascending order by j subscript.

For each cluster, P_i is the number of principal components that will be included in the final dictionary and is obtained by the following equation:

$$P_i = \left(\underset{j}{argmax} \ \lambda_{ij} \left| \lambda_{ij} \leq \sigma^2 \right. \right) - 1$$
$$U_i(:,1:P_i) \in D \qquad (23)$$

The principal components higher than P_i have learned noise for each cluster in matrix U_i. Actually, P_i is the dimension of noise-free data on the ith cluster (or P_i is the rank of subspace that ith cluster lies in it). It means that if the noise power is zero, autocorrelation matrix of ith cluster has only P_i non-zero eigenvalues. In presence of noise, all autocorrelation matrix eigenvalues of each cluster of noisy data will be nonzero; from the component $P_i + 1$ to the end are due to noise. By adding the first principal component to P_i, the dictionary is completed and we can perform denoising by this designed dictionary.

6. Denoising Operation

Usefulness of the union of subspaces model has been proved in many applications of signal processing. As

illustrated in section 2 and 3, this model is appropriate for the analysis of signal de-noising. This model assumes that image blocks are linear combination of few bases of a dictionary:

$$z_i = D\alpha_i \, st \, \|\alpha_i\|_0 \le Th \qquad (24)$$

In the previous section a dictionary was defined. De-noised image should also meet this model whereas noisy image y_i cannot, because in the dictionary learning stage, noise is not trained. In other words, to represent noise, many bases combination should be involved and no sparse representation α_j in equation (25) can be found.

$$\nexists \, \alpha_i \mid y_i = D\alpha_i \, st \, \|\alpha_i\|_0 \le Th \qquad (25)$$

The model must be reformed to model the noise of data:

$$y_i = D\alpha_i + n_i \, , \|\alpha_i\|_0 \le Th \qquad (26)$$

Assuming Gaussian noise with zero mean in this model, MAP estimation for α_j is

$$\hat{\alpha}_i = \min_{\alpha_i} \|y_i - D\alpha_i\|^2 \, st : \|\alpha_i\|_0 \le Th \qquad (27)$$

Optimum threshold is related to P_i of a cluster where y_i belongs to it. This can be replaced by the following problem:

$$\hat{\alpha}_i = \min_{\alpha_i} \|\alpha_i\|_0 \, st : \|y_i - D\alpha_i\|^2 \le \epsilon \qquad (28)$$

where, ϵ is a function of noise variance. Now we can estimate de-noised version by this estimation of sparse coefficients. We just need to project y_i into the nearest low-rank subspace spanned by the columns of the learned dictionary.

7. Simulation Results

In this section, de-noising results of proposed method and some other recent approaches are presented and discussed. First, the global and local methods are evaluated, then a comparison between global and local approaches is presented and finally these methods are compared in term of total execution time.

K-SVD and our simple gain-shaped K-means (proposed method) are presented as global methods. The presented local methods include those introduced in [5], [8], [13], [14], [17] and [18]. Performance comparison of these algorithms can be seen Table 1. We have used the Peak Signal to Noise Ratio (PSNR[1]) as the performance criterion. The PSNR values were averaged over 5 experiments, corresponding to 5 different realizations of AWGN. The variance was negligible and not reported.

Our method is simulated similar to the framework of [7]. Both algorithms have the same amount of error for the training set (depending on the noise variance) but their size of dictionary is different. Table 1 shows that the proposed method surpasses the K-SVD [7] and its results are comparable with the time consuming local methods. As will be tabulated, the execution time of the proposed method is about 70% of K-SVD, 8% of LSSC [14] and 4% of CSR [13]. Recently [16] investigated a comprehensive comparison of different image de-noising methods. They have shown numerically that BM3D, SCR and LSSC studied in this paper have the best results. Figure 9 shows an example of de-noising results by our proposed method.

Table 1. Image de-noising performance of the Global and Local methods in PSNR (dB) for 4 different image and various σ/SNR

Lena				
	σ/SNR	5/34.16	10/28.14	20/22.11
Global	Proposed	38.71	35.60	32.57
	K-SVD [7]	38.60	35.47	32.38
Local	K-LLD [8]	38.01	35.20	32.37
	LSSC [14]	38.69	35.83	32.90
	CSR [13]	38.74	35.90	32.96
	BM3D [5]	38.72	35.93	33.05
	LSC [17]	38.56	35.65	32.54
	SSMS [18]	38.62	35.63	32.30
Barbara				
	σ/SNR	5/34.16	10/28.14	20/22.11
Global	Proposed	38.22	34.68	30.98
	K-SVD [7]	38.08	34.42	30.83
Local	K-LLD [8]	37.26	33.30	28.93
	LSSC [14]	38.48	34.97	31.57
	CSR [13]	38.43	35.10	31.78
	BM3D [5]	38.31	34.98	31.75
	LSC [17]	38.45	34.95	31.29
	SSMS [18]	38.73	35.11	31.25
House				
	σ/SNR	5/34.16	10/28.14	20/22.11
Global	Proposed	39.59	36.54	33.68
	K-SVD [7]	39.37	35.98	33.20
Local	K-LLD [8]	37.63	35.09	32.66
	LSSC [14]	39.93	36.96	34.16
	CSR [13]	39.98	36.88	33.86
	BM3D [5]	39.83	36.71	33.77
	LSC [17]	39.72	36.33	33.23
	SSMS [18]	39.51	36.13	32.77
Boat				
	σ/SNR	5/34.16	10/28.14	20/22.11
Global	Proposed	37.25	33.85	30.52
	K-SVD [7]	37.22	33.64	30.36
Local	K-LLD [8]	35.96	33.16	30.17
	LSSC [14]	37.35	34.02	30.89
	CSR [13]	37.31	33.88	30.78
	BM3D [5]	37.28	33.92	30.87
	LSC [17]	37.16	33.75	30.42
	SSMS [18]	37.09	33.70	30.40

---

[1] PSNR is defined as $10log_{10}(255^2/MSE)$ and measured in dB

(a) Original image

(b) Noisy image

(c) Recovered image by proposed method
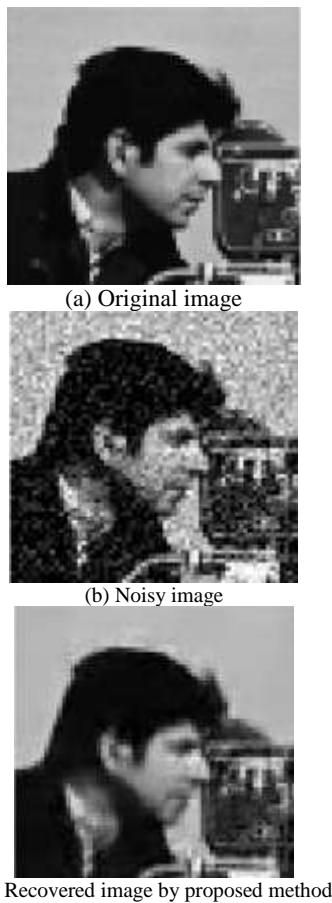
Fig 9. an example of denoising results by our method

In natural images, far away block have generally different patterns, so, using all blocks may result in inappropriate clustering. Moreover, non-overlapped clusters obtained by global methods are not as flexible as the overlapped groups. On the other hand, local grouping assign appropriate groups to each block. Although local methods have better performance, global methods are able to extract salient features of images and use it easily for de-nosing. According to comparison of local and global methods in Table 1, the performance of the proposed global method is just about $0.2dB$ lower than promising local methods (LSSC and CSR), which is not a high difference. However, a common good property of both global and local methods is that they exploit the low-dimensional characteristics of clusters/groups in order to design a suitable de-noising algorithm.

To understand the effect of this method on the dictionary, in the table 2 the results are compared only with K-SVD method, which is global a method like our proposed method. However, in the results of local methods in table 1, the suggested method used about 27%

less blocks for training and the time required for dictionary learning is less than K-SVD method. This table studies the effect of data equalization on K-SVD. As it can be seen equalization improves K-SVD about 0.4dB.

Table 2. comparing the suggested method and K-SVD method. left: K-SVD + Equalization of data, middle: KSVD, right: the proposed clustering

σ/SNR	House			Peppers		
20/22.11	33.29	33.16	33.68	30.89	30.77	31.09
25/20.18	32.37	32.19	32.66	29.82	29.69	29.96
30/18.59	31.40	31.24	31.61	28.95	28.82	29.11
σ/SNR	Lena			Cameraman		
20/22.11	32.55	32.38	32.63	30.14	29.96	30.36
25/20.18	31.42	31.34	31.50	29.10	28.93	29.22
30/18.59	30.59	30.46	30.72	28.16	28.07	28.36

As mentioned, the proposed method is based on dictionary learning and its time efficiency should be compared with other dictionary learning based approaches e.g. [7], [13] and [14]. Table 4 compares the relative execution time of [13], [14], [7] and the proposed method in various image sizes. Our experiments were averaged on 5 different runs carried out on a Personal Computer with a 3.6-GHz AMD 2 Core CPU and 4 GB RAM. As can be seen, the global de-noising methods (KSVD and proposed) are more efficient in term of execution time and our proposed method surpasses KSVD. In fact, dictionary learning running time of proposed method (for identification of K-rank1 subspaces) is about 40% of K-SVD for 20,000 blocks extracted from a 512×512 image, but its overall execution time is about 72% of KSVD.

Table 3. Relative execution time of dictionary learning based methods (in minutes)

Image Size		144×176 (QCIF)	288×352 (CIF)	576×704 (4CIF)
Local	LSSC [14]	1.29	5.31	22.09
	CSR [13]	2.67	11.28	52.87
Global	KSVD [7]	0.14	0.59	2.66
	Proposed	0.10	0.43	1.92

8. Conclusions

Local methods suggested in recent years, have obtained better results than global methods. However by more intelligent training in such a way that first, important data is more effective for training, second, clustering in such way that training blocks lie in low-rank subspaces, we can design a dictionary applicable for image de-noising and obtain results near the state of the art local methods.

As was seen, we have obtained acceptable results by a relatively simple method based on construction of an appropriate global dictionary.

References

[1] D. van de Ville and M. Kocher, "SURE-Based non-local means," IEEE Signal Process. Letter, Vol. 16, No. 11, Nov. 2009, pp. 973–976.

[2] T. Tasdizen, "Principal neighborhood dictionaries for nonlocal means image denoising," IEEE Trans. Image Process., Vol. 18, No. 12, Dec. 2009, pp.2649–2660.

[3] R. Vignesh, B. T. Oh, and C.-C. J. Kuo, "Fast non-local means (NLM) computation with probabilistic early termination," IEEE Signal Process. Letter, Vol. 17, No. 3, Mar. 2010, pp. 277–280.

[4] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel Regression for Image Processing and Reconstruction," IEEE Transactions on Image Processing 16, 2007, pp. 349–366.

[5] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Processing, Vol. 16, No. 8, Aug. 2007, pp. 2080–2095.

[6] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," Pattern Recognition ,Vol. 43, 2010, pp. 1531–1549.

[7] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Processing ,Vol. 15, No. 12, Dec. 2006, pp. 3736–3745.

[8] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," IEEE Trans. on Image Processing, No. 18, Vol. 7, 2009, pp. 1438-1451.

[9] P. Chatterjee, and P. Milanfar, "Patch-Based Near-Optimal Image Denoising," Image Processing, IEEE Transactions on, vol. 21, no. 4, April 2012, pp.1635-1649.

[10] P. Chatterjee, and P. Milanfar, "Practical Bounds on Image Denoising: From Estimation to Information" IEEE Transactions on Image Processing, Vol. 20, No. 5, May 2011, pp. 1221-1233.

[11] C. Bohm, K. Kailing, P. Kroger, and A. Zimek, "Computing clusters of correlation connected objects" In Proc. SIGMOD, 2004, pp. 455-466.

[12] A. Gersho and R. M. Gray, "Vector Quantization and Signal Compression," Springer Press, 1992.

[13] W. Dong, X. Li, D. Zhang and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2011, pp. 457-465.

[14] J. Marial, F. Bach, J. Ponce, G. Sapiro and A. Zisserman, "Non-local sparce models for image restoration," in Proceedings of IEEE International Conference on Computer Vision, 2009, pp. 2272-2279.

[15] J. A. Tropp, "Algorithms for simultaneous sparse approximation," Elsevier Signal Processing journal, Vol. 86, No. 3, 2006, pp. 589-602.

[16] L. Shao, R. Yan, X. Li, and Y. Liu, "From Heuristic Optimization to Dictionary Learning: A review and comprehensive comparison of image denoising algorithms," IEEE Transaction on Cybernetics, Vol. 44, No. 7, 2014, pp. 1001-1013.

[17] A. Adler, M. Elad, Yacov Hel-Or," Probabilistic subspace clustering via sparse representation," IEEE Signal Processing Letter, Vol. 20, No. 1, 2013, pp.63-66.

[18] G. Yu, G. Sapiro, and S. Mallat, "Image modeling and enhancement via structured sparse model selection," In Proceeding of ICIP, 2010, pp. 1641-1644.

Hossein Bakhshi Golestani received the B.Sc. and M.Sc. in Electrical engineering from Ferdowsi university of Mashhad, Mashhad, Iran and Sharif University of Technology, Tehran, Iran, in 2010 and 2012 respectively. His research interests include multimedia signal processing, statistical signal processing, video/speech coding and compression, signal quality assessment and robotics.

Mohsen Joneidi received the B.Sc. and M.Sc. in Electrical engineering from Ferdowsi university of Mashhad, Mashhad, Iran and Sharif University of Technology, Tehran, Iran, in 2010 and 2012 respectively. He is currently working toward the Ph.D degree in department of electrical engineering and computer science at the University of Central Florida, Orlando, USA. His research interests include multimedia signal processing, statistical signal processing, signal quality assessment and compressive sensing.

Mostafa Sadeghi received the B.Sc. and M.Sc. in Electrical engineering from Ferdowsi university of Mashhad, Mashhad, Iran and Sharif University of Technology, Tehran, Iran, in 2010 and 2012 respectively. He is currently working toward the Ph.D degree in department of electrical engineering at the Sharif University of Technology, Tehran, Iran. His research interests include statistical signal processing, sparse representation/coding and compressive sensing.

Facial Expression Recognition Using Texture Description of Displacement Image

Hamid Sadeghi
Department of Electrical Engineering, Amirkabir University of Technology, Tehran. Iran
hamid.sadeghi@aut.ac.ir
Abolghasem-Asadollah Raie*
Department of Electrical Engineering, Amirkabir University of Technology, Tehran. Iran
raie@aut.ac.ir
Mohammad-Reza Mohammadi
Department of Electrical Engineering, Sharif University of Technology, Tehran. Iran
mrmohammadi@ee.sharif.edu

Abstract

In recent years, facial expression recognition, as an interesting problem in computer vision has been performed by means of static and dynamic methods. Dynamic information plays an important role in recognizing facial expression in the image sequences. However, using the entire dynamic information in the expression image sequences is of higher computational cost compared to the static methods. To reduce the computational cost, instead of entire image sequence, only neutral and emotional faces can be employed. In the previous research, this idea was used by means of Difference of Local Binary Pattern Histogram Sequences (DLBPHS) method in which facial important small displacements were vanished by subtracting Local Binary Pattern (LBP) features of neutral and emotional face images. In this paper, a novel approach is proposed to utilize two face images. In the proposed method, the face component displacements are highlighted by subtracting neutral image from emotional image; then, LBP features are extracted from the difference image as well as the emotional one. Then, the feature vector is created by concatenating two LBP histograms. Finally, a Support Vector Machine (SVM) is used to classify the extracted feature vectors. The proposed method is evaluated on standard databases and the results show a significant accuracy improvement compared to DLBPHS.

Keywords: Facial Expression Recognition; Difference Image; Displacement Image; Local Binary Patterns; Support Vector Machine.

1. Introduction

Emotion expression in human face or facial expression is an important way of human emotional social interaction. Psychological studies show that the basic emotions have universal facial expressions in all cultures [1]. There are six basic emotions including anger, disgust, fear, happiness, sadness, and surprise which were proposed in [2]. Different subjects express these emotions differently. However, we can recognize facial expression of an unfamiliar face [3]. Due to various applications, such as human-computer interaction and producing robots with human-like emotions, automatic analysis of facial expression becomes an interesting and challenging problem in pattern recognition and machine vision studies.

Feature extraction plays an important role in the accuracy of recognizing facial expression. Facial expression recognition systems can be divided according to their feature extraction method [4]. Basically, there are three types of feature extraction methods: (1) appearance feature-based; (2) geometric feature-based; and (3) hybrids of appearance and geometric features. Geometric feature-based methods [5-7] use the location and geometric shape of facial components, such as mouth and eyes, to represent a facial image. Appearance feature-based methods [8-11] employ the texture information of facial image [4,11]. In the hybrid methods [12-16], both geometric and appearance features are utilized to represent facial images. Though geometric features have similarity to or more accuracy than the appearance features [17,18], geometric feature extraction generally needs perfect facial fiducial point localization.

In [6,7], a geometric model of 30 fiducial points was proposed and several specific distances were used as facial features. In [5], a subset of the facial expressions was recognized by calculating the correlation functions from some geometric features of the lip regions, such as the relationship between the width and the height of the lips. In some studies [19-22], optical flow-based methods were used to track facial movements in the video sequences. However, optical flow-based methods suffer from the varying illumination and non-rigid facial motions and cannot be used in real-time applications due to their computational complexity [23].

In appearance feature-based methods, Gabor filter [24] was generally applied to the facial images [15,25-27]. However, convolving Gabor filters in different orientations and scales require a huge amount of

---

* Corresponding Author

calculations. In recent years, another type of texture descriptors as Local Binary Patterns (LBP) has been introduced [28], and its different versions have been used for facial appearance feature extraction in both static and dynamic approaches [9-11,15]. A survey of the exiting works on facial image analysis using LBP-based representation can be found in [29]. One important property of LBP is its tolerance against illumination changes [11]. In [11], facial images were normalized by manually labeling the eyes location, and then the face images were represented using LBP features. Volume Local Binary Patterns (VLBP) and Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) were used for dynamic image texture description in [9,10]; as well as the recognition of facial expression. The disadvantage of VLBP is its long feature vector increasing computational cost.

In other categorization, the existing studies can be divided into static and dynamic methods [4]. Static methods use a single frame to recognize facial expression, while in the dynamic methods temporal changes in the video sequence are utilized [4]. Facial expression recognition using LBP presented in [11] is a static approach. However, VLBP and LBP-TOP used in [9,10] are dynamic methods. Psychological studies indicate that dynamic methods provide higher performance than the static approaches [30].

1.1 Database

In this paper, extended Cohn-Kanade database (CK+) [31], that is an extended version of original Cohn-Kanade dataset (CK) [32], is used to evaluate the proposed algorithm. CK+ database was presented in 2010 to overcome the limitations of CK database. The database consists of 123 subjects and all prototypic emotions; in addition, the dataset includes contempt facial expression. Moreover, 68 fiducial points were localized using Active Appearance Model (AAM) in the database [31]. Each data includes a sequence of images starting from neutral face to the peak of its emotion. For instance, Fig. 1 shows

an image sequence with surprise facial expression in CK+ database. In this paper, three peak frames of each sequence were labeled as one of the six basic emotions are used for our experiments. Besides, face regions are localized using AAM feature point localization. Fig. 2 shows some examples of CK+ facial image selected from peak frames.

1.2 A brief overview of this study

As mentioned previously, facial expression recognition using dynamic approaches provide better results than the static methods [30]. Therefore, this paper tries to utilize the dynamic information of facial expression. However, utilizing the whole dynamic information of an image sequence has a huge computational cost. For this reason, this paper compares the emotional image with neutral image to extract suitable facial features. In this study, LBPs are used as facial appearance features due to their accuracy in facial representation [29] and computational simplicity. All of the experiments are person-independent in such a way that the train persons are not present in the testing data. For this reason, a multiclass Support Vector Machine (SVM) classifier is used in different cross validation testing schemes. Support Vector Machine is a powerful classifier which has attracted much attention in pattern recognition and facial expression recognition problems [9-11,27].

The remainder of this paper is summarized as:

In the next section, LBP texture representation is described. The proposed feature extraction method using LBP is described in section 3. Evaluation of the proposed system on standard databases is presented in section 4. Finally, section 5 concludes this paper.

2. Local Binary Patterns

Basic Local Binary Patterns (LBP) operator was

Fig 1. One image sequence in CK+ database [31] with surprise expression. The sequence starts from neutral image (top left) to the peak of surprise emotion (bottom right).

Fig 2. Some samples of original CK database images and extended data in CK+ database are presented in the top and bottom rows, respectively. From top left to bottom right: Disgust, Happiness, Surprise, Fear, Angry, Contempt, Sadness, and Neutral facial expression [31]

introduced as a powerful texture descriptor in [28]. This operator describes each image pixel by an integer number on the interval {0-255}. For each pixel, LBP operator produces 8 labels by thresholding 3×3 neighborhood of the pixel with its gray-level value. The corresponding decimal value of generated binary number by 8 labels is then used to describe the given pixel. Fig. 3 shows the result of applying LBP operator to a pixel.
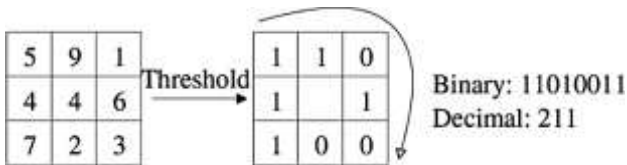
Fig 3. An example of the basic LBP operator on an image pixel [33].

Finally, the matrix of produced LBP codes is defined and the histogram of these codes is given as

$$h_i = \sum_{x,y} S(lbp_k(x, y) = i), \quad i = 1,\dots,n-1 \quad (1)$$

Where lbp_k is the produced LBP code matrix of the kth image; n is the number of LBP codes (in basic LBP operator, n is equal to 256); and function S is defined as
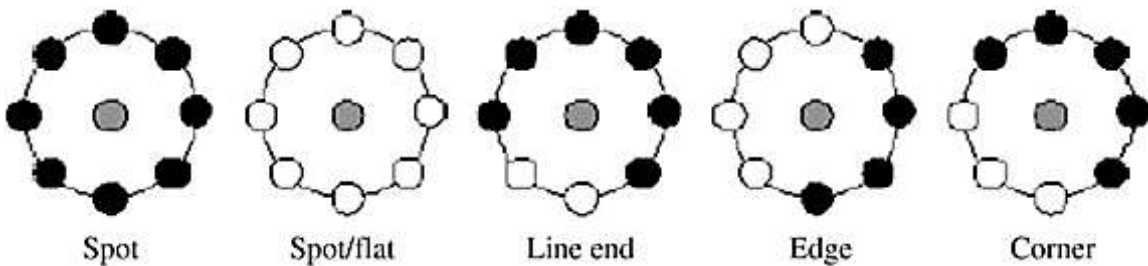
$$S(x) = \begin{cases} 1 & x \text{ is } True \\ 0 & x \text{ is } False \end{cases} \quad (2)$$

The basic LBP operator produces 256 binary patterns. However, it can be shown that a subset of these binary patterns contains more information than others [34]. The most appeared binary patterns in facial images are uniform that contain at most two bitwise transitions from 1 to 0 or vice versa when the binary code string circulates [34]; for example 00011100 and 11111111 binary patterns are uniform. The number of uniform binary patterns in the basic LBP operator is 58; accumulating the non-uniform binary patterns into a single bin yields a 59 bin histogram (LBP^{u2} operator) which can be used as a texture descriptor [34]. LBP^{u2} histogram contains texture information over the image. Fig. 4 shows some examples of uniform binary patterns along with their micro-texture information.

The experiments show that 90.6% of the appeared binary patterns using basic LBP operator in facial images are uniform [36]. Consequently, this paper uses LBP^{u2} operator to represent facial images.

Spot   Spot/flat   Line end   Edge   Corner

Fig 4. The experiments show that 90.6% of the appeared binary patterns using basic LBP operator in facial images are uniform [36].

3. Facial Representation using Displacement Image LBP (DI-LBP)

As mentioned previously, this study uses the dynamic information of facial images. Appearance-based features represent texture information, such as creases, wrinkles, and furrows in the facial image. Intuitively, we can say that a facial expression is a variation in the texture of facial image. This variation is individual in each expression; and it lies in some specific regions. In this paper, the facial regions containing expression-based variation are detected in the video sequence. For this reason, the difference of emotional image and one of the previous frames (here: the first frame of image sequence) is calculated. Then, LBP^{u2} features are extracted from the difference or displacement image (DI). The block diagram of the proposed algorithm is shown in Fig. 5.

At first, three peak frames and the first frame of each sequence are selected from the database. All color images are converted to grayscale image. Next, face region is localized using AAM feature points. Then, all facial images are normalized to 150×110 pixels. In the next
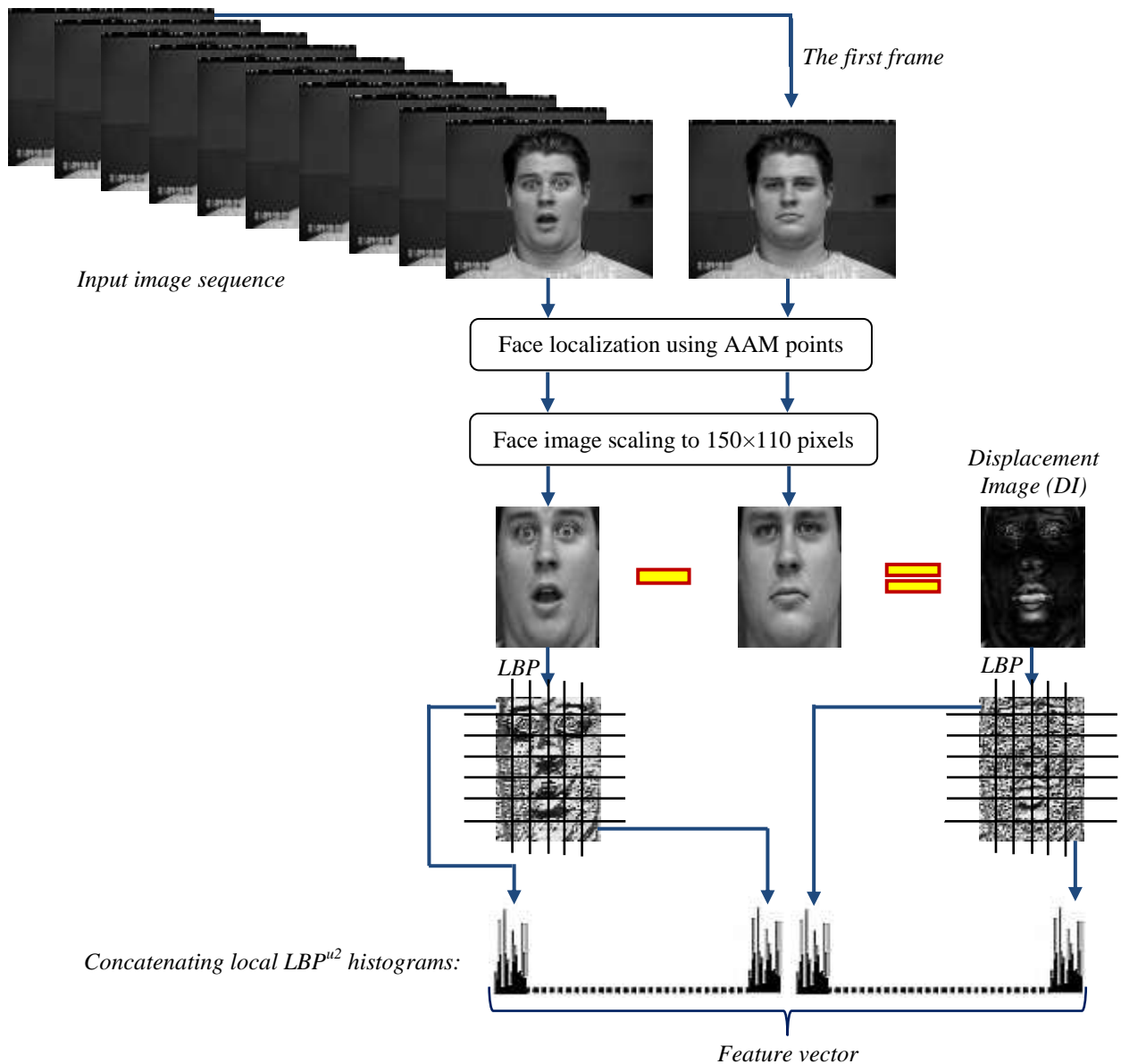
Fig 5. The general framework of proposed facial expression recognition system.

Step, the difference of three peak frames and the first frame is calculated. Then, LBP^{u2} histogram is extracted from difference/displacement image (DI-LBP). Thus, the facial regions being stationary in the image sequence are appeared as flat areas in DI. Accordingly, facial stationary textures are accumulated in the flat area bin of LBP histogram. Holistic LBP histogram cannot represent any indication about location of binary patterns. To overcome

this problem, facial images are divided into 42(6×7) sub-region with the size of 21×18 pixels according to [11]. Then, the LBP^{u2} histograms of these sub-regions are concatenated to a single feature vector with the length of 2478(59×42) to represent facial image. Each sub-region LBP histogram can be defined as

$$H_i = \sum_{x,y} S\{lbp_k(x,y) = i\}.S\{(x,y) \in R^{(j)}\}$$
$$i = 0,\dots,n-1, \qquad j = 0,\dots,m-1 \qquad (3)$$

where R(j) is jth sub-region in facial image; and m is the number of sub-regions (m = 42).

To keep all suitable information of facial image, the LBP histograms of emotional image is calculated in the same way. Finally, this LBP histograms and DI-LBP are concatenated into a final feature vector with the length of 4956(2×2478).

In [37], difference of LBP histograms which are calculated from emotional and neutral frames was used as DLBPHS (Difference of LBP Histogram Sequences) feature vector to represent facial images. One disadvantage of DLBPHS is the elimination of small displacement texture information from LBP histogram by subtracting LBP feature of two images. Small displacements which occurred in a single sub-region of facial image did not affect the LBP histogram of this sub-region. In other words, the LBP histogram is not changed by a small displacement in the facial component. Consequently, subtracting two LBP histograms eliminates the small displacement information of facial image. On the contrary, the effects of these small displacements are highlighted in DI. Calculating LBP histogram from DI can be appropriately used for texture description in facial image.

Another disadvantage of subtracting two LBP histograms in DLBPHS algorithm is the elimination of facial stationary region information. In contrast, in the proposed method, the difference of neutral and emotional images is calculated; thus, facial stationary regions are appeared as flat area in DI. Accordingly, the texture information of these regions lies in flat area bin in the LBP histogram. As a result, feature vectors are inherently normalized. Furthermore, any information is not eliminated from feature vector; and only lies in suitable bins in the LBP histogram.

4. Experiments

To evaluate the performance of the proposed algorithm, Support Vector Machine (SVM) [38] classifier is utilized on CK+ [31] database. For this reason, different testing schemes, including 10-fold cross validation and leave-one-subject-out are used for person-independent classification. To reduce the computational cost, some important regions in facial image can be used for feature extraction. For this reason, mouth and eyes regions are selected in our experiments. This method reduces computational complexity and enhances the accuracy of

the algorithm. A multi-class SVM classifier with linear and polynomial (degree = 2) kernel functions is used to classify the extracted features. Table 1 shows the result of different region selection in facial expression recognition. The confusion matrix of six-expression recognition on CK+ database using SVM (linear) is shown in Table 2. The results of different region selection are compared with [39] in Table 3.

Table 1. 6-class average recognition rate on CK+ database in case of different region selection

SVM kernel	Measure				
Linear	10-fold	91.45	93.72	89.5	79.44
	[1]LOSO	**93.16**	**94.68**	**91.97**	**80.24**
Poly.	10-fold	91.67	94.16	89.72	77.71
	LOSO	92.29	94.68	91.86	79.26

[1] LOSO: leave-one-subject-out

Table 2. Confusion matrix of six-expression recognition on CK+ database using SVM (linear) classifier and leave-one-subject-out testing scheme

Expression	An. (%)	Di. (%)	Fe. (%)	Ha. (%)	Sa. (%)	Su. (%)
Angry	**92.59**	5.19	0	0	2.22	0
Disgust	1.72	**97.70**	0	0.57	0	0
Fear	4.17	2.78	**88.89**	0	0	4.17
Happiness	0	2.42	1.45	**96.14**	0	0
Sadness	8.33	0	0	0	**88.10**	3.57
Surprise	0	1.20	1.20	1.2	0	**96.39**
Average	**94.68 %**					

Table 3. Comparison of the algorithm performance with [39] on CK+ database in case of different facial region selection

	whole face	Best region selection	mouth	eyes
Ours (%)	**93.16**	**94.68**	**91.97**	**80.24**
[39] (%)	88	91.4	81.0	64.8

CK+ database includes contempt facial expression, which makes it more challenging than the original CK database. In this section, 7-class facial expression recognition including contempt and six basic expressions is performed using SVM (polynomial kernel with degree = 2) classifier and leave-one-subject-out cross validation. Table 4 shows the confusion matrix of 7-class facial expression recognition. It can be seen from Table 4 that some classes have higher recognition rate than other ones. It can be due to different numbers of data in the CK+ database. Fig. 6 shows the recognition rate of each class along with the number of training data in such class. According to Fig. 6, in anger, disgust, happiness, and surprise expression, where training data are bigger than other ones, recognition rate is higher than three other expressions.

Table 5 compares the performance of the proposed method to other studies which used SVM classifier on CK+ dataset. In [39], the authors use Manifold based Sparse Representation (MSR) method and show that their method outperforms SVM classifier. As shown in Table 5,

the proposed algorithm has the highest accuracy on the database.

Table 4. Confusion matrix of 7-class recognition on CK+ database

Exp.	An. (%)	Co. (%)	Di. (%)	Fe. (%)	Ha. (%)	Sa. (%)	Su. (%)
An.	**97.44**	0	1.71	0	0	0.85	0
Co.	0	**83.33**	0	0	8.33	8.33	0
Di.	1.72	0	**98.28**	0	0	0	0
Fe.	5	0	0	**85**	5	0	5
Ha.	0	0	1.93	1.45	**96.62**	0	0
Sa.	13.04	0	0	2.9	0	**84.06**	0
Su.	0	2.41	1.20	1.20	0	0	**95.18**
Average	94.41 %						
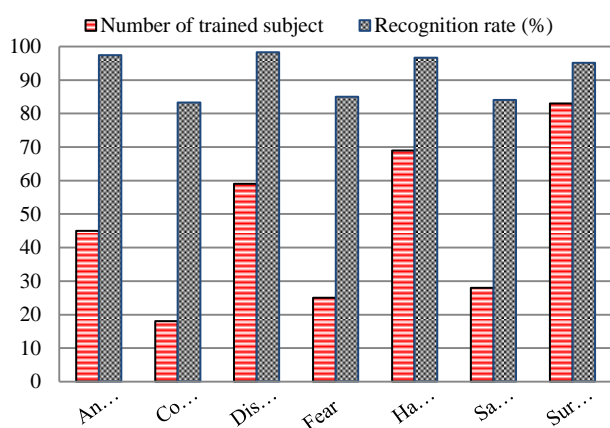
Fig 6. Recognition rate versus number of train data in each class.

Table 5. Comparison of the algorithm performance with the existing work on CK+ database

Method	Classes	Classifier	Cross Validation	Acc. (%)
Ours	6	SVM	leave-one-subject-out	94.68
Ours	6	SVM	10-fold	94.16
Ours	7	SVM	leave-one-subject-out	94. 41
Ours	7	SVM	10-fold	94. 05
[31]	7	SVM	leave-one-subject-out	88.33
[40]	7	SVM	10-fold	90.1
[41]	7	SVM	leave-one-subject-out	82.6
[42]	7	SVM	10-fold	89.3
[39]	7	MSR	leave-one-subject-out	91.4

5. Evaluation on JAFFE database

To provide a fair comparison of the proposed algorithm with DLBPHS [37], we conducted the experiments in the same ways as [37]. In [37], JAFFE database [43] was used in the experiments. JAFFE database consists of 10 Japanese female subjects. Each subject has 3 or 4 images for each basic facial expression

and the neutral face (totally 213 images with size of 256 × 256 pixels). Some samples of JAFFE facial expression images are shown in Fig. 7.

In [37], 2 images from each expression for each subject were selected in training step, and the rest of the images from each expression were used as test images. In this section, the experiments are conducted in the same way as [37]. The confusion matrix of six basic expressions is shown in Table 6. As can be seen in Table 6, sadness expression has lowest recognition rate in six classes. Recognition of sadness expression in JAFFE database is relatively difficult due to the dataset characteristics. We can see this issue in the reported results in [37]. Moreover, we implement the DLBPHS method [37] (which was discussed in section 3) with the same results as [37] on JAFFE database. Then, this method is experimented on CK+ database in a 10-fold cross validation person-independent testing scheme. Table 7 compares the results of our method with DLBPHS on both CK+ and JAFFE databases.

Table 1. Confusion matrix of 6-class expression recognition on JAFFE database

Expression	An.	Di.	Fe.	Ha.	Sa.	Su.
Angry	**90**	10	0	0	0	0
Disgust	0	**100**	0	0	0	0
Fear	0	10	**90**	0	0	0
Happiness	0	0	0	**100**	0	0
Sadness	0	0	20	10	**70**	0
Surprise	0	0	0	0	0	**100**
Average	91.23 %					

Table 7. Comparisons between DLBPHS method [37] and our method

Method	Database	6-calss recognition (%)
Our method	**CK+**	**94.16**
Our method	**JAFFE**	**91.23**
DLBPHS	CK+	82.36
DLBPHS [37]	JAFFE	80.33

6. Conclusions

In this paper, we present an approach for facial expression recognition which uses the dynamic information of neutral and emotional frames for feature extraction. For this reason, the first frame of each image sequence (neutral face) is subtracted from the emotional face instead of subtracting their feature vectors. Then, LBP texture descriptor is utilized to extract efficient facial features. From the experimental results on the standard databases, it can be concluded that feature extraction from difference/displacement image (DI) provides a better accuracy than subtracting LBP feature vectors.

In many dynamic facial expression methods, the whole facial expression image sequence is used. The proposed method can be used in these dynamic systems to reduce the computational cost. Moreover, DI can be computed from emotional face and one of the previous frames (which has low emotion intensity) instead of neutral face for real-world applications in future work.
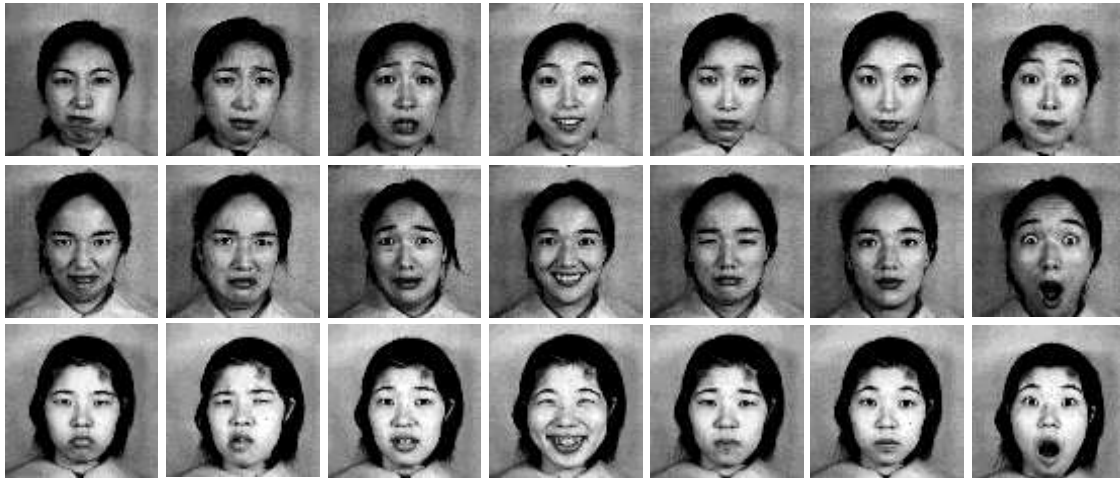
Fig 7. The example images from JAFFE database [43]; each row from left to right: anger, disgust, fear, happiness, sadness, neutral, and surprise.

References

[1] J. Fasel, J. Luettin, "Automatic facial expression analysis: a survey", Pattern Recognition, Vol. 36, No. 1, pp. 259-275, 2003.

[2] P. Ekman, W.V. Friesen, "Constants across cultures in the face and emotion", Journal of Personality Social Psychology, Vol. 17, Vo. 2, pp. 124–129, 1971.

[3] P. Ekman, "Facial expressions of emotion: an old controversy and new findings", Philosophical Transactions of the Royal Society, London, B335:63--69, 1992.

[4] Y. Tian, T. Kanade, J. Cohn, "Handbook of Face Recognition", Springer, 2005 (Chapter 11. Facial Expression Analysis).

[5] N-S. Pai, S-P. Chang, "An embedded system for real-time facial expression recognition based on the extension theory", Computers and Mathematics with Applications, Vol. 61, No. 8, pp. 2101-2106, 2011.

[6] H. Kobayashi, F. Hara. "Recognition of Six Basic Facial Expression and Their Strength by Neural Network". Proc. Int'l Workshop Robot and Human Communication, pp. 381-386, 1992.

[7] H. Kobayashi, F. Hara. "Recognition of Mixed Facial Expression by Neural Network". IEEE International Workshop Robot and Human Communication, pp. 387-391, 1992.

[8] L. Ma, K. Khorasani, "Facial Expression Recognition Using Constructive Feedforward Neural Networks", ", IEEE Transaction on Systems Man Cybernetics, Part B, Cybernetics, Vol. 34, No. 3, 2004.

[9] G. Zhao, M. Pietikäinen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 29, No. 6, pp. 915-928, 2007.

[10] G. Zhao, M. Pietikäinen, "Boosted multi-resolution spatiotemporal descriptors for facial expression recognition", Pattern Recognition Letters, Vol. 30, No. 12, pp. 1117-1127, 2009.

[11] C. Shan, Sh. Gong, P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", Image and Vision Computing, Vol. 27, No. 6, pp. 803-816, 2009.

[12] A. Lanitis, C. Taylor, T. Cootes, "Automatic interpretation and coding of face images using flexible models", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 743–756, 1997.

[13] T. Cootes, G. Edwards, C. Taylor, "Active appearance models", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 23, No. 6, pp. 681–685, 2001.

[14] G. Edwards, T. Cootes, C. Taylor, "Face recognition using active appearance models", Proceedings of the Fifth European Conference on Computer Vision (ECCV), University of Freiburg, Germany, Vol. 2, pp. 581–695, 1998.

[15] R. Xiao, Q. Zhao, D. Zhang, P. Shi, "Facial expression Recognition on multiple manifolds", Pattern Recognition, Vol. 44, No. 1, pp. 107-116, 2011.

[16] L. Zhang, S. Chen, T. Wang, Z. Liu, "Automatic Facial Expression Recognition Based on Hybrid Features", Int. Conf., Future Electrical Power and Energy Systems, Energy Procedia, Vol. 17, pp. 1817-1823, 2012.

[17] M. Valstar, I. Patras, M. Pantic, "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data", IEEE Conf. on Computer Vision and Pattern Recognition-Workshop, Vol. 3, pp. 76-84, 2005.

[18] M. Valstar, M. Pantic, "Fully Automatic Facial Action Unit Detection and Temporal Analysis", IEEE Conf. on Computer Vision and Pattern Recognition-Workshop, p.149, 2006.

[19] Y. Yacoob, L.S. Davis, "Recognizing human facial expression from long image sequences using optical flow", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 6, pp. 636–642, 1996.

[20] I. Essa, A. Pentland, "Coding, analysis, interpretation, and recognition of facial expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 757–763, 1997.

[21] M. Yeasin, B. Bullot, R. Sharma, "From facial expression to level of interest: a spatio-temporal approach", in: IEEE

Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, pp. 922-927, 2004.

[22] J. Hoey, J.J. Little, "Value directed learning of gestures and facial displays", in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, pp. 1026-1033, 2004.

[23] S.-S. Liu, Y. Tian, D. Li, "New Research Advances of Facial Expression Recognition", Proceedings of the Eighth International Conf., Machine Learning and Cybernetics, Baoding, pp. 1150-1155, 2009.

[24] J. Daugmen, "Complete discrete 2d Gabor transforms by neutral networks for image analysis and compression", IEEE Transaction on Acoustic, Speech and Signal Processing, Vol. 36, No. 7, pp. 1169–1179, 1988.

[25] S. Bashyal, G. K. Venayagamoorthy, "Recognition of facial expression using Gabor wavelets and learning vector quantization", Engineering Applications of Artificial Intelligence, Vol. 21, No. 7, pp.1056-1064, 2008.

[26] W. Gu, Ch. Xiang, Y. V. Venkatesh, D. Huang, H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis", Pattern Recognition, Vol. 45, No.1, pp. 80-91, 2012.

[27] L. Zhang, D. Tjondronegoro, "Facial Expression Recognition Using Facial Movement Features", IEEE Transaction on Affective Computing, Vol. 2, No. 4, pp. 219-229, 2011.

[28] T. Ojala, M. Pietikäinen, D. Harwood, "A comparative study of texture measures with classification based on featured distribution", Pattern Recognition, Vol. 29, No. 1, pp. 51–59, 1996.

[29] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey", IEEE Transaction on Systems Man Cybernetics, Part C, Application and Reviews, Vol. 41, No. 6, pp. 765-781, 2011.

[30] J.N. Bassili, "Emotion recognition: the role of facial movement and the relative importance of upper and lower area of the face", Journal of Personality and Social Psychology, Vol. 37, No. 11, pp. 2049–2058, 1979.

[31] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression", Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, pp. 94-101, 2010.

[32] T. Kanade, J. F. Cohn, Y. Tian, "Comprehensive database for facial expression analysis", Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, pp. 46-53, 2000.

[33] T. Ahonen, A. Hadid, M. Pietikäinen, "Face recognition with local binary patterns", European Conference on Computer Vision, pp. 469-481, 2004.

[34] T. Ojala, M. Pietikäinen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp. 971–987, 2002.

[35] A. Hadid, M. Pietikäinen, T. Ahonen, "A discriminative feature space for detecting and recognizing faces", IEEE

Conf., Computer Vision and Pattern Recognition, pp. 797-804, 2004.

[36] T. Ahonen, A. Hadid, and M. Pietik¨ainen, "Face description with local binary patterns: Application to face recognition," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 28, Nno. 12, pp. 2037–2041, 2006.

[37] W. F. Liu, Y. J. Wang, "Expression Feature Extraction Based on Difference of Local Binary Pattern Histogram Sequences", IEEE International Conference on signal processing, ICSP Proceedings, pp. 2082-2084, 2008.

[38] Ch. Ch. Chang, Ch. J. Lin, "LIBSVM : a library for support vector machines", ACM Transactions on Intelligent Systems and Technology, 2012. available at http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf >.

[39] R. Ptucha, A. Savakis, "Manifold Based Sparse Representation for Facial Understanding in Natural Images", Image and Vision Computing, Vol. 31, No.5, pp. 365–378, 2013.

[40] M. S. Islam, S. Auwatanamongkol, "A Novel Feature Extraction Technique for Facial Expression Recognition", International Journal of Computer Science Issues, Vol. 10, No. 3, pp. 9-14, 2013.

[41] S. Yang, B. Bhanu, "Understanding discrete facial expressions in video using an emotion avatar image", IEEE Transaction on Systems Man Cybernetics, Part B, Cybernetics, Vol. 42, No. 4, pp. 980-992, 2012.

[42] A. R. Rivera, J. R. Castillo, O. Chae, "Local directional number pattern for face analysis: face and expression recognition", IEEE Transaction on Image Processing, Vol. 22, No. 5, pp. 1740 - 1752, 2013.

[43] M. J. Lyons, J. Budynek, S. Akamatsu, "Automatic classification of single facial images", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 21, No. 12, pp. 1357–1362, 1999.

Hamid Sadeghi received the B.Sc. and the M.Sc. degrees in Electrical Engineering from Hakim Sabzevari University, and Amirkabir University of Technology (Tehran Polytechnic), Iran, in 2011 and 2013, respectively. Currently, he is a PhD candidate of Electrical Engineering in Amirkabir University of Technology, Tehran, Iran. His research interests are computer vision, machine learning, and pattern recognition.

Abolghasem Asadollah Raie received the B.Sc. degree in Electrical Engineering from Sharif University of Technology, Iran, in 1973 and the M.Sc. and Ph.D degrees in Electrical Engineering from University of Minnesota, USA, in 1979 and 1982, respectively. Currently, he is an Associate Professor with the Electrical Engineering Department of Amirkabir University of Technology, Iran. His research interests are algorithm design and performance analysis, machine vision, sensor fusion, and mobile robots navigation.

Mohammad Reza Mohammadi was born in Qom in Iran, on July 25, 1987. He received BSc and Msc degrees both in Electrical Engineering with rank one from Amirkabir University of Technology (Tehran Polytechnic). He is currently PhD candidate of electrical engineering in Sharif University of Technology. His interests and researches include Machine Vision and Machine Learning.

SIP Vulnerability Scan Framework

Mitra Alidoosti*

Department of Computer Engineering, Iran University of Science and Technology,Tehran. Iran

Alidoosti@comp.iust.ac.ir

Hassan Asgharian

Department of Computer Engineering, Iran University of Science and TechnologyTehran. Iran

Asgharian@iust.ac.ir

Ahmad Akbari

Department of Computer Engineering, Iran University of Science and Technology,Tehran. Iran

Akbari@iust.ac.ir

Abstract

The purpose of this paper is to provide a framework for detecting vulnerabilities in SIP (Session Initiation Protocol) networks. We focused our studies on the detection of SIP DoS related vulnerabilities in VoIP infrastructures because of their generalization. We try to find weaknesses in SIP enabled entities that an attacker by exploiting them is able to attack the system and affect it. This framework is provided by the concept of penetration testing and is designed to be flexible and extensible, and has the capability to customize for other similar session based protocols. To satisfy the above objectives, the framework is designed with five main modules for discovery, information modeling, operation, evaluation and report. After setting up a test-bed as a typical VoIP system to show the validity of the proposed framework, this system has been implemented as a SIP vulnerability scanner. We also defined appropriate metrics for gathering the performance statistics of SIP components. Our test-bed is deployed by open-source applications and used for validation and also evaluation of the proposed framework. The main contributions of this paper are its non-destructive manner in identifying vulnerabilities and incorporating the penetration testing ideas and steps in the overall architecture of our framework. We also defined appropriate metrics that help us to identify vulnerabilities in a black box penetration testing.

Keywords: Vulnerability Scanner; SIP; Denial of Service Attacks; Framework; Evaluation.

1. Introduction

Voice over IP protocols (VoIP) simply enables two devices to transmit and receive real-time audio traffic that allows their respective users to communicate. VoIP architectures are generally partitioned into two main groups: signaling and media [1]. Signaling covers both abstract notions, such as endpoint naming and addressing, and concrete protocol functions such as parameter negotiation, access control, billing, proxying (routing), and NAT traversal [2]. The media transfer aspect of VoIP systems generally includes a simpler protocol for encapsulating data, with support for multiple codecs and content security. A commonly used media transfer protocol is RTP. There exists an RTP profile that supports encryption and integrity protection (SRTP), but it is not yet widely used. The RTP protocol family also includes RTCP, which is used to control certain RTP parameters between communicating endpoints. In spite of the media transport layer of VoIP infrastructures, its signaling layer can accept different signaling like H.323, Skinny and SIP. In this paper we focus on the SIP which is the most widely used protocol in the standard VoIP architectures and next generation networks [3, 4]. Unfortunately, because of the interoperability requirements with the existing telephony infrastructure, its new features, and the speed of development and deployment, VoIP protocols and products

contain numerous vulnerabilities that have been exploited. Most of these vulnerabilities are the result of the complexity of VoIP systems which demonstrates itself both in terms of configuration options and implementation issues. As a result, VoIP systems represent a very large attack surface [1]. So it is expected that security problems arising from design flaws (e.g. exploitable protocol weaknesses), undesirable feature interactions (e.g. the combination of components that make new attacks possible), unforeseen dependencies (e.g. compromise paths through seemingly unrelated protocols), weak configurations, and many other implementation flaws.

Vulnerability scanning is the process of assessing a variety of vulnerabilities across information systems (including computers, network systems, operating systems, and software applications) and allowing early detection and handling of known security problems [5]. A vulnerability scanner can help to identify rogue machines, which might endanger overall system and network security, helps to verify the inventory of all devices on the network [5]. The inventory includes the device type, operating system version and patch level, hardware configurations and other relevant system information. This information is useful in security management and tracking. There are general tools for vulnerability assessment and scanning of some application layer protocols but because of the special

---

* Corresponding Author

vulnerabilities of VoIP architectures, there is no well-known and widely acceptable tool in this field.

Therefore we have proposed a SIP vulnerability scanner framework for evaluating VoIP components against well-known SIP attacks. We focused our studies on the detection of SIP DoS related vulnerabilities in VoIP infrastructures because of their generalization. Although our proposed solution is general and has no assumption about the underlying VoIP component (i.e. user agent devices and proxy servers) but because of our previous experiences, we focused on SIP proxy servers and present its results on our experimental test-bed. The main contributions of this paper are its non-destructive manner in identifying vulnerabilities and incorporating the penetration testing ideas and steps in the overall architecture of our framework. We also defined appropriate metrics that help us to identify vulnerabilities in a black box penetration testing.

The other parts of this paper are organized as follows: The next section reviews the literature and some related works. Our proposed solution is expressed in section 3 and our experimental setup and evaluation of the presented system is defined in section 4. Finally the conclusion is abstracted in section 5.

2. Literature Reviews and Related Works

One of the main approaches to security assessment of computer networks and systems is penetration testing. Penetration testing tools perform a non-destructive attack to check the security status of an organization network and distinguish its vulnerabilities. Generally it has three steps which are done sequentially: discovery, attack and report.

The process of penetration testing contains system analysis to identify the potential vulnerabilities of systems which arises because of misconfigurations or implementation faults. The penetration testing process is categorized in to two broad groups [6]: black box and white box. In black box testing, it is assumed that there is no knowledge about considered network. We selected this approach in our proposed SIP vulnerability scanner. Thus, in discovery step all required information about the given target is collected. This information contains SIP enabled devices and their footprints. Since other steps depend tightly on this step, the pen-tester must take suitable time to complete this phase. In attack step, non-destructive attacks are imposed on the target and according to their effects, being inferred that the target is vulnerable or not. In taking report step, the sufficient report is prepared to notify the organization about available vulnerabilities.

Reference [7] demonstrates VoIP specific security assessment framework to perform automated VoIP specific penetration tests. This framework searches and detects existing vulnerabilities or misconfigured devices and services. This security assessment tool mentions DoS attacks, but flooding attacks are not considered, so could not verify how the behavior of SIP systems may change under system load during flooding attacks. This framework architecture contains three main modules that perform the required tasks such as discovering as much as possible information from the devices in the network, storing and providing all collected information in a usable format and finally launches penetration tests and perform attack actions using gathered information. Other related work for VoIP penetration testing is [8]. It measures the vulnerability of SIP-based VoIP systems during security attacks. It considers some categories of DoS attacks and defines the availability of the system under test (SUT) for its validation that we used its main idea for evaluation in our paper.

In [8], the main focus of availability is on the user interaction during attack times. The ratio of successful call rate during attack's period to pre-attack times is measured; furthermore the re-transmission number of each call is calculated that represents the influence of attacks. Since the main focus of [8] is on attack generation, it is likely to damage the SUT. Therefore we try to solve this problem by considering the potential vulnerabilities of SUT (using the result of discovery phase) and plan the non-destructive attacks based on that.

Reference [9] presents a security management framework for VoIP. In order to estimate the SIP and RTP related security vulnerabilities and threats of VoIP; a fuzzy packet tool is developed. The functionality of the proposed framework defines in XML scenarios. Depends on the physical location of this tool, different tasks can be performed such as man in the middle attacks, user enumeration and password testing for a registration server, ARP injection in order to intercept network traffic or just protocol level fuzzing.

We inspired the steps of our framework from the phases of penetration testing. In other words, we assess the existence of vulnerabilities on a given target by discovering the proper target and plan a proper operation against it for realization. Reference [7] demonstrates an acceptable framework for vulnerability assessment; we get the generality of our framework from it that we customized it based on our knowledge about the VoIP attacks. Reference [8] presents a good idea for measuring the effects of attacks, it divides the duration of the test in 3 parts: pre-attack phase, attack phase and post-attack phase. The perception of this fact that one target is vulnerable to a specific attack or not, is measured by changes of the system behavior to normal users. We use this idea in our framework to figure out the sensitivity of considered platform to attacks. We define three criteria and measure them during pre-attack, attack and post-attack stages. The changes of these criteria during these phases simply detect that our target is vulnerable or not. The details of our proposed framework are defined in the next section. Because of comprehensiveness and importance of DoS flooding attacks in SIP, our main focus in this paper is on them. SIP flooding attacks are reviewed in many papers like [10,11] that we don't review them here.

3. Proposed Approach to SIP Vulnerability Assessment

As it is said in the previous sections, our main goal is to design a system for vulnerability scanning of VoIP systems. The main output of the proposed approach is an evaluation tool for comparing different implementations of SIP components in handling known attacks. We focused on the vulnerabilities that led to denial of service attacks. Since it does not affect the generality of our problem, we limited our studies on SIP proxy servers and three broad classes of SIP DoS attacks but the proposed solution is general and can be extended to any VoIP components and VoIP related vulnerabilities. In other words, we aim to mention whether the given SIP proxy is vulnerable against specific DoS attacks or not. In fact we want to explore the weaknesses of given sip proxy which the attacker misuses them and intrudes the sip proxy.

Our main contribution is to proposed SIP vulnerability scan architecture. The architecture of our proposed solution is shown in Figure 1. It has five main modules that are as follows:

1. Discovery Module
2. Information Model Module
3. Operation Module
4. Evaluation Module
5. Report Module

In the following subsections, the detail design of these modules is described.
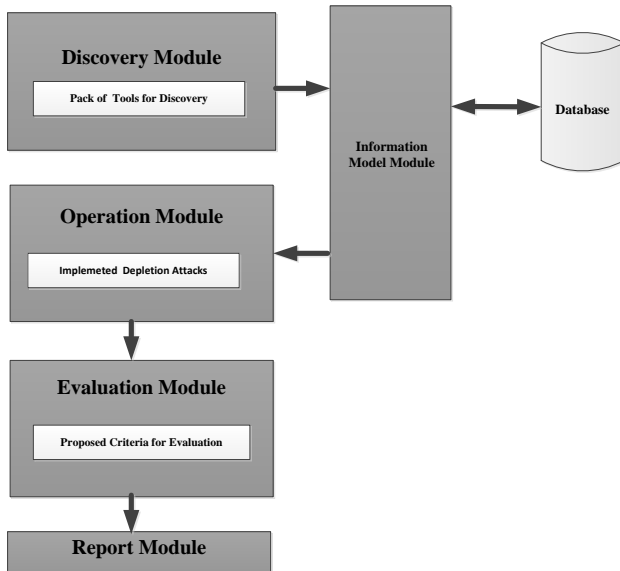
Fig 1. Architecture of the proposed framework

3.1 Discovery Module

The main objective of this module is to recognize the active hosts. Discovery Module is shown in Figure 2.
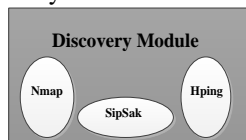
Fig. 2 Discovery Module

This module contains specific tools for fingerprinting the given network. It also provides useful information about the active host such as IP address, open ports, running services, MAC address and also some other information about the specific node in the VoIP network. This module works automatically and finds all active hosts with SIP enabled services. The last step of this phase is recognizing the type of the active host such as being the proxy server or the user agent client. In this module, the Nmap [12] tool is used to identify active hosts in a VoIP network, we configured the SIPSak [13] to discover the type of host and Hping [14] is used to diagnose active and inactive hosts.

3.2 Information Model Module
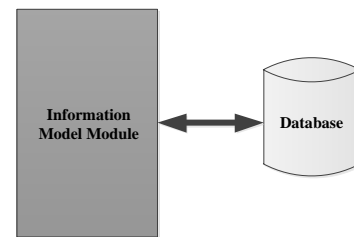
Fig. 3 Information Model Module

The Information Model Module is shown in Figure 3. This module uses some online repositories like NVD [15] to find out the related vulnerabilities of the enumerated hosts by previous module. In this stage by using stored information in the database, potential vulnerabilities within the entity will be discovered. According to recognized vulnerabilities, appropriate attacks are chosen. Type of entity and selected attacks are saved in the database and this information is given to the next module. Therefore this module is responsible for the following two functions:

1. Selecting types of applied attacks for the system under test.
2. Updating the database according to newly discovered vulnerabilities.

In other words, according to information obtained from the previous step and the type of entity, the type of applied attack is selected and the database is updated based on the type of entity and selected attacks.

3.3 Operation Module
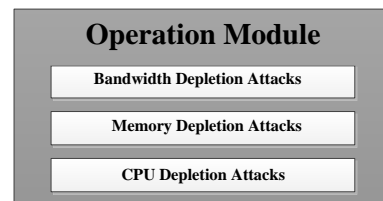
Fig. 4 Operation Module

The responsibility of operation module is to apply selected attacks against the target. This module and evaluation module will start to run simultaneously. By analyzing the underlying traffic through the target, the

evaluation unit can determine whether the target is vulnerable to corresponding attack or not. So in this module the selected attacks are applied to the target and the vulnerable target can be identified by evaluation module. The applied attacks in this module are bandwidth depletion attacks, memory depletion attacks and CPU depletion attacks. The architecture of Operation Module is shown in Figure 4.

Bandwidth depletion attacks by creating a large number of redundant messages try to occupy the bandwidth. The aim of memory depletion attacks is consuming SIP entity's memory so that it is not able to respond to legitimate demands of users. For generating memory attack we produce messages that extend call setup time, so the target is forced to hold call information in its memory until the call is finished, therefore the memory will be occupied longer than usual and the sessions prolong which cause to cease the memory. By producing a certain number of these kind of messages, memory of the target will be occupied and will not have sufficient space for legitimate users. CPU depletion attacks are generated by creating messages that need additional processing to keep the processor busy. So that the target does not have enough time to process messages receiving from other legal entities and users. These kind of attacks usually make by using malformed messages or authentication based attacks.

3.4 Evaluation Module

As stated before, evaluation and operation module start simultaneously. This module assesses passing traffic through the SIP entity, so the vulnerability of the target will be extracted. For detecting vulnerabilities of SIP entity a new metrics are defined. By measuring these criteria, the vulnerability of the target against applied attacks can be diagnosed.

3.4.1 Proposed Scheme to Identify Vulnerabilities

As shown in Figure 5, simulation period was considered as T3 seconds. During the entire simulation period, normal traffic is available between the proxy server and other existing users.
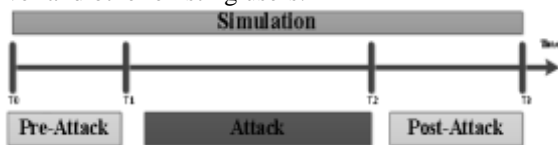
Fig.5 proposed scheme to identify vulnerable hosts

The attack is applied in the time interval [T1, T2]. The defined criteria are measured during time interval [T1, T2] and [T0, T1], [T2, T3]. Time intervals [T0, T1] and [T2, T3] indicate non-attack period and the period [T1, T2] indicates attack period. If the considered value of measured metrics has changed during the attack interval and non-attack interval, it can be concluded that the proxy server is vulnerable to the applied attack.

3.4.2 Evaluation metrics for SIP attack's effects

In this subsection we define some metrics for evaluating vulnerability of the entity against applied attacks. These criteria can help to recognize the existence of the vulnerability.

1. Completion Call Rate

In the attack period due to the heavy traffic to the proxy server, vulnerable server does not have enough resources to create new calls with legitimate users. As a result Completion Call Rate during attack period is decreased than non-attack period. Thus reducing Completion Call Rate during attack period is one of the criteria that determine the vulnerability of the proxy server.

2. Retransmission Call Rate

In attack period due to applying many requests from attacker to the victim, there is not enough time to respond to requests. When the query timeout, retransmission will be performed. So if the server is vulnerable, retransmission call rate will be considerable. The Ratio of retransmission call rate in attack period to non-attack period determines the vulnerability of the proxy server.

3. Response Time

Response time is time interval between sending a request and receiving its response. In attack period due to applying heavy traffic to victim if the proxy server is vulnerable, the response time to legitimate user request will be longer. Prolonging response time in attack period than non-attack period certify the proxy server is vulnerable against applying attack.

4. Call Set Up Time

Call set up time refers to the period of time that a request to establish a call is sent until the call is ended. In attack period due to applying heavy traffic to victim if the target is vulnerable, call set up time will be longer. Prolonging call set up time certify the target is vulnerable against applying attack.

5. Round Trip Time[1]

R.T.T is the time required for a 32-byte packet to travel from a specific source to a specific destination and back again. As stated before in attack period due to heavy traffic the target being busy, so R.T.T become longer if target is vulnerable.

3.5 Report Module

This module is responsible for comparing measured criteria in attack period with non-attack period. If these two values have significant difference, it can be concluded that the SIP entity is vulnerable to applied attack.

4. Experiment Setup and Evaluation

Because of importance and mandatory role of proxy servers in SIP environments, this entity is selected in our

---

[1] R.T.T

experiments. SIP proxy is more vulnerable to security threats especially against Denial of Service attacks.

Our experiment test-bed (as shown in Figure 6) consists of a user agent client (UAC), a SIP proxy server and a user agent server (UAS). The UAC and UAS run SIPp [16] for generating SIP normal traffic with selected parameters. UAC and UAS are connected by proxy server. In fact the connection between UAC and UAS as normal traffic is always available during the simulation period. Since we measure our performance metrics in field just before attack period, we did not have any consideration for normal traffic because it is not important in our experiments. Proposed scanner that is shown as pen-tester in a separate computer in figure 6 applies selected attacks to the proxy server. The pen-tester uses VoIP hacking tools of Linux Backtrack 5 on his computer which run it in a virtual machine in our experiments. Depending on the attack scenario, the proposed scanner sometimes needs a partner to design an attack against the proxy server. This partner is shown as co-pen-tester in figure 6. The generated attacks are applied to the victim proxy server from pen-tester station during the test period, passing traffic through the victim proxy server is captured and defined metrics are measured for captured traffics.

Co-Pen-tester
CPU: Pentium Dual core -2.20GHZ
RAM: 2.00 GB
Back track Linux5 run on VMware
OS: Windows 7

Pen-tester
CPU: Pentium Dual core -2.20GHZ
RAM: 2.00 GB
Back track Linux5 run on VMware
OS: Windows 7

Proxy: Opensips 1.5.0
CPU: Pentium -2.60GHZ
RAM: 1.00 GB
OS: Cent OS 6.2

UAS: SIPp v2.0.1 2007
CPU: Pentium Dual core -2.20GHZ
RAM: 2.00 GB
OS: Windows 7

UAC: SIPp v2.0.1 2007
CPU: Pentium Dual core -2.20GHZ
RAM: 2.00 GB
OS: Windows 7

Fig. 6 Test-bed for evaluating proposed scanner

4.1 Generated attacks in operation module

As stated before, operation module is responsible for applying bandwidth depletion, memory depletion and CPU depletion attacks. As proxy server is selected among SIP entities for our test, so applied attacks should have effects on proxy servers. Therefore applied attacks are:

1. Bandwidth Depletion Attacks

    a. Invite flood attacks

In this type of attack we generate large number of INVITE packets by SIPp tool. In this attack scenario we want to just deplete the bandwidth of the proxy server.

    b. UDP flood attacks

UDP flood attack will produce by Hping tool. A large number of UDP packets are sent to the proxy server, so its bandwidth will be occupied with spurious packets.

2. Memory Depletion Attacks

    a. Brute force attacks

SIP has a session control mechanism in the application layer. The SIP sessions consist of two different concepts: transaction and dialog. Almost all stateful SIP proxies are implemented in the transaction level and for this reason maintains all related statistics of sessions until its expiration. The attacker uses this mechanism to deplete the memory of the proxy server by routing packets to it in a rate which is more than the proxy's capacity. In other word, the pen-tester sends messages for generating call to the victim proxy server but does not responds their responses from the victim, therefore victim proxy server is made to keep the call's information for a longer period of time until time to get the response runs out. So each message leads to occupy memory more than usual. In this case proxy server's memory will be occupied and there is not enough memory to meet demands of legal users. SIPp tool with appropriate scenarios is used to generate such attacks.

    b. SYN flood attacks

In this type of attack pen-tester (attacker) sends many SYN packets to the victim proxy server. The proxy server thinks TCP connection will be established therefore stores calls' information but pen-tester will not answer proxy any more. In this way proxy has to keep calls' information up to its predefined time out of the RFC 3261 (up to 180 seconds). As a result proxy's memory will be occupied by sending a large amount of SYN packets to it and will not have sufficient space for legitimate users. SYN flood attack is generated by Hping tool in our framework.

    c. Incomplete transaction attacks

Pen-tester for applying incomplete transaction attacks needs coworker (co-pen-tester). In order to produce such an attack; pen-tester sends a message (e.g. an INVITE message) to the proxy server and asks him to forward it to co-pen-tester. But co-pen-tester is configured to not respond the proxy server. In consequence proxy server has to keep transaction's information until time runs out. Then proxy sends time-out to pen-tester. There are two possible cases:

Case1: When pen-tester receives time-out from the proxy, sends ACK message to the proxy. By producing a large number of this message (INVITE message), proxy's memory will be occupied and there is not sufficient space to perform other users' requests. This case is shown in Figure 7.
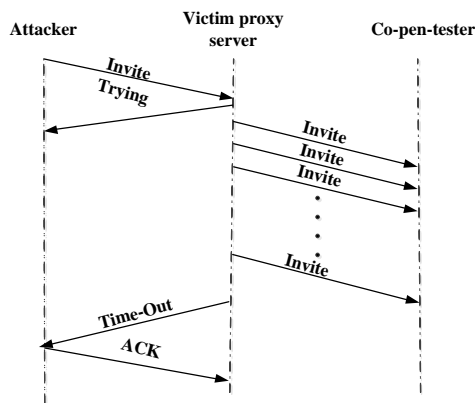
Fig.7 First case of incomplete transaction attacks

Case2: When pen-tester receives time-out from the proxy, does not respond to proxy. So proxy must repeat time-out message until specific time. In this case proxy keeps transaction information in its memory longer than usual. This case is shown in figure 8.
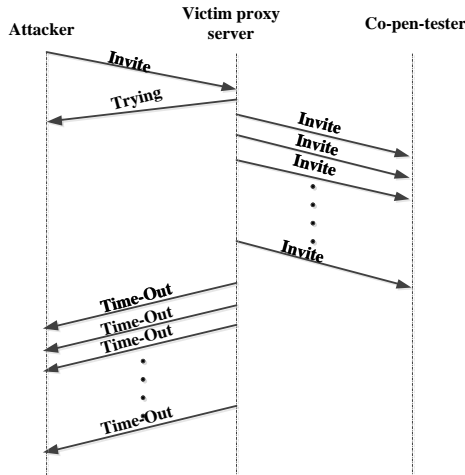
Fig.8 Second case of Incomplete transaction attacks

Both of these attacks are generated by appropriate use of SIPp in UAS and UAC modes.

d. Incomplete transaction with partner

For implementing this attack Pen-tester sends a message for example INVITE message to proxy and wants him to forward it to co-pen-tester. Co-pen-tester is configured to send TRYING messages in order to respond proxy server and made proxy to keep waiting for giving response from it. When time out to respond, victim proxy server sends time-out to pen-tester. But pen-tester does not respond by ACK to proxy. Therefore proxy repeats time-out to pen-tester for a limited time. In this case poor proxy has to keep transaction's information longer than before. This attack is shown in Figure 9.
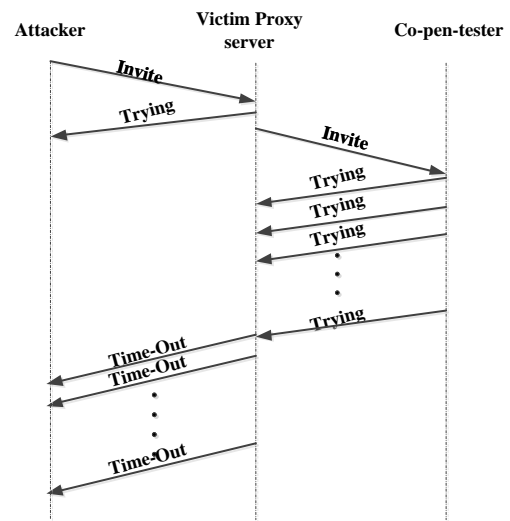
Fig. 9 Incomplete transaction with partner

3. CPU Depletion Attacks

a. ICMP flood attacks

ICMP flood attack sends a large number of ICMP packets to the victim proxy server. A lot of processing power is needed to analyze these packets. Therefore CPU will not have enough time to process requests from legitimate users. A special java application is written for implementation of this kind of attack.

b. Authentication misuse attacks

Most of the SIP servers are configured to authenticate users before their registration. The mandatory authentication mechanism of SIP is HTTP digest method which is based on the challenge and response. The attacker tries to deplete the processing power of SIP proxy by misusing the authentication process. For generating this kind of attack, the pen-tester sends a large number of INVITE messages on victim proxy server. Proxy server for each message designates a random number called nonce (for generating nonce CPU will be involved) and send back both the message and its nonce to pen-tester. The proxy expects from pen-tester for sending message and its nonce to him again. But smarty pen-tester will not do anything! By generating a large number of these messages proxy server' CPU will be busy just to generate random numbers (challenge) and will not have sufficient time to process legitimate users' requests. More details about these attacks can be found in [10].

It should be noted that all the test scenarios are about 30 seconds. During this period, the normal traffic is available on the proxy server and in the time interval [10, 12], the attack scenario is applied to the proxy server. The proposed scanner according to traffic underlying through the victim proxy server and measuring defined criteria, diagnoses the vulnerability of the proxy server in a black box manner.

4.2 Measuring evaluation criteria for analyzing vulnerability of SIP entity

In section 2-4-3 evaluation criteria to diagnose vulnerable SIP entity is expressed. Now this section mentions how to calculate them.

1. Completion call rate in attack and non-attack period

The number of completed calls in a considered window can show the activity ratio of server. For calculating the number of completed calls, those messages should be considered that at least one of the features of source tag, destination tag and call-ID have changed. The reason is that unique triple of [source tag, destination tag, call-id] specifies a dialog. The number of dialogs shows the number of completed calls. Then the average of completed calls in attack and non-attack period are measured. By comparing these two numbers in other words by reducing this number in attack period than non-attack period can be diagnosed vulnerable proxy server against applied attack.

2. Retransmission call rate in attack period and non- attack period

Messages that have same transaction identifiers [source tag, from tag, call-Id, via, CSeq[1]] have been retransmitted. To calculate number of retransmitted calls in traffic through proxy server, those messages that have same five features are counted in both attack period and non-attack periods. Then the average of retransmission calls in both periods are calculated. Increasing this number in attack period than non-attack period indicates the vulnerability of proxy server against applied attack.

3. Response time in attack and non-attack period

SIP INVITE message response time is time interval between SIP INVITE and its RINGING message. For each INVITE message response time is calculated in both attack period and non-attack period. Then the average of them in both periods are measured. By comparing these two numbers the vulnerability of proxy server against applied attack can be identified. If proxy server is vulnerable, response time will be increased in attack period than non-attack period.

4. Call set-up time in attack and non-attack period

Call setup time for INVITE message is time interval between INVITE and its ACK message. These two messages must have same call-ID. In this way we calculate setup time in attack period and non-attack period then we measure the average of call setup time in both attack and non-attack period. By comparing these two numbers vulnerability of proxy server can be identified. Increasing call setup time in attack period than non-attack period indicates that the proxy server is vulnerable to applied attack.

5. R.T.T in attack period and non-attack period

R.T.T is another criterion that is defined to detect vulnerable proxy server. For calculating R.T.T simply can use "ping" command, R.T.T is one of information in ping command. We calculate R.T.T in both attack and non-attack period. Then the average of these time intervals in both periods are calculated. If proxy server is vulnerable, the average of R.T.T in attack period will be increased.

4.3 Simulation results and analysis

We assumed that we want to check the vulnerabilities of the SIP proxy server (OPENSIPS 1.5.0). Therefore proposed scanner is applied to OPENSIPS [17] and the defined criteria are measured. Table 1 shows the results of applying bandwidth depletion attacks on OPENSIPS.

Table 1. The Results of applying bandwidth depletion attacks on OPENSIPS

OPENSIPS				
Metric	Invite Flood		UDP Flood	
	Attack	Non-attack	Attack	Non-attack
Completed Call (CC)	309	362	314	254
Retransmission Calls (RC)	8898	0	400	0
Response time (ms) (RT)	189.13	2.01	33.37	2.91
Call set up time (ms) (CST)	57.37	3.23	39.36	18.36
Round Trip Time (ms) (RTT)	1.3	0.73	9.40	0.74

Shown in Table 1, since there is a significant difference between metrics like response time and call setup time during attack period and normal period, we can conclude that the studied proxy server is vulnerable to bandwidth depletion attacks. It should be said that we assure about the vulnerability of OPENSIPS to bandwidth depletion attack by sending high volume traffic to this proxy server before starting this experiment.

Table 2 shows the results of applying memory depletion attacks on OPENSIPS.

Table 2. The Results of applying memory depletion attacks on OPENSIPS

OPENSIPS								
Metric	SYN Flood		Brute Force Attack		Incomplete Transaction		Incomplete Transaction with partner	
	Attack	Non-Attack	Attack	Non-Attack	Attack	Non-Attack	Attack	Non-Attack
CC	288	297	143	297	685	724	524	767
RC	24	0	5386	0	14766	173	47233	173
RT	58.34	2.25	86.50	4.71	6.23	5.14	40.87	1.86
CST	57.81	3.23	91.46	2.95	429.81	3.25	890.20	4.33
RTT	24.12	0.67	14.77	0.87	1.52	0.84	35.00	10.44

According to Table 2 OPENSIPS is vulnerable to memory depletion attacks. For example call set up time measured in SYN flood attack is increased in attack period than non-attack period. As a result, OPENSIPS is vulnerable against SYN flood attack.

---

[1] Call sequence

Table 3. The Results of applying CPU depletion attacks on OPENSIPS

Metric	OPENSIPS			
	ICMP Flood		Security Checking Attack	
	Attack	Non-attack	Attack	Non-attack
CC	288	297	2204	12196
RC	24	0	6220	0
RT	58.34	2.25	107.37	2.80
CST	57.81	3.23	736.78	5.62
RTT	24.12	0.67	3.14	0.65

According to Table 3 OPENSIPS is vulnerable to CPU depletion attacks because of significant differences between defined metrics in attack period and non-attack periods.

4.4 Validating proposed framework

The purpose of this section is to demonstrate the validity of the proposed framework. As stated before proposed framework's duty is extraction of entity's vulnerabilities. So in other words we want to show that our proposed framework correctly identifies the exist vulnerabilities in an entity. Figure 10 shows steps to prove framework functionality correctness.
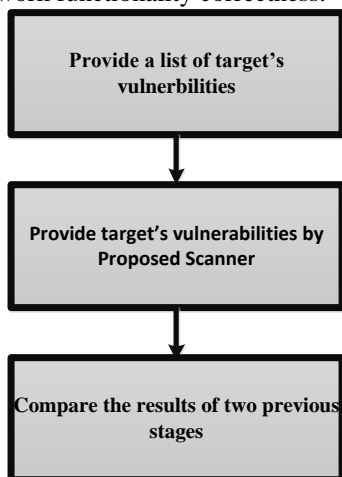
Fig. 10 steps to prove framework functionality correctness

1. Providing a list of target's vulnerabilities

First we should find a list of vulnerabilities of the target. In other words we just answer the question that the proposed system seeks to answer. For providing the list of vulnerabilities we do as follows:

We consider some resources in victim that our attacks can have effects on it. For example for memory depletion attacks, we consider memory. Then we apply the attack that has effects on those resources in a very high rate. If the victim becomes unavailable we can conclude that the victim is vulnerable to that attack. We do this experiment on OPENSIPS 1.5.0 and found out that OPENSIPS is vulnerable to all of our applied attacks. This step is shown in Figure 11.
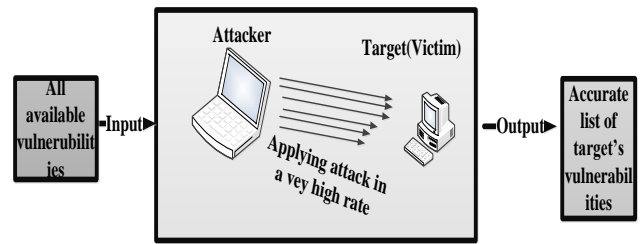
Fig. 11 providing a list of target's vulnerabilities

2. Providing target's vulnerabilities by the proposed scanner

In this step we apply our scanner to our victim to get the list of vulnerability the victim has. We did this experiment in section 4-3 and concluded that our victim (OPENSIPS 1.5.0) is vulnerable to all of our applied attacks. This step is shown in Figure 12.
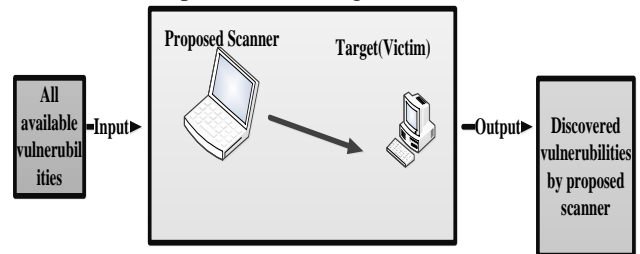
Fig. 12 providing target's vulnerabilities by the proposed Scanner

3. Compare the results of two previous steps

In this step the results of two previous steps are compared. If these results are same, we can conclude that proposed scanner operates correctly. These two results are same so our scanner detects vulnerabilities correctly.

5. Conclusions

Although the main focus of this paper is on VoIP and especially on SIP protocol, the vulnerability analysis framework presented in this paper is general to accommodate a variety of similar protocols. In this paper, a SIP security assessment framework is presented by using the idea of penetration testing. In this context, producing non-destructive attacks is used to identify vulnerabilities in SIP entities. The main idea of this paper includes the identification of vulnerabilities, defining criteria for assessing vulnerability and the generating of non-destructive and controlled attack. To evaluate the proposed idea, a practical VoIP test bed is used.

References

[1] Angelos D. Keromytis, Senior Member, "A Comprehensive Survey of Voice over IP Security Research", IEEE Communications Surveys and Tutorials, Vol 14, No. 2, pp. 514-537, 2012

[2] Angelos D. Keromytis, "Voice over IP: Risks, Threats and Vulnerabilities", In Proceedings (electronic) of the Cyber Infrastructure Protection (CIP) Conference. June 2009

[3] Angelos D. Keromytis, "Voice over IP Security: Research and Practice", IEEE Security & Privacy, Volume: 8 , Issue: 2 , pp 76- 78, March-April 2010

[4] Angelos D. Keromytis, "A Look at VoIP Vulnerabilities", login: Magazine, vol. 35, no. 1, pp. 41 - 50, February 2010.

[5] "An Overview Of Vulnerability Scanners", The Government of the Hong Kong Special Administrative Region, online document: http://www.infosec.gov.hk/english/technical/files/vulnerability.pdf, 2008

[6] Andrew Whitaker, Daniel P. Newman, "Penetration Testing and Network Defense", Cisco Press, November 04, 2005

[7] H. Abdelnur, R. State, I. Chrisment, C. Popi, "Assessing the security of voip services", 10th IFIP/IEEE International Symposium on Integrated Network Management, 2007, pp. 373 – 382

[8] Peter Steinbacher, Florian Fankhauser, Christian Schanes, Thomas Grechenig, "Work in Progress: Black-Box Approach for Testing Quality of Service in Case of Security Incidents on the Example of a SIP-based VoIP Service", IPTComm 2010, 2-3 August, pp. 101-110

[9] H. Abdelnur, V. Cridlig, R. State and O. Festor, "VoIP Security Assessment: Methods and Tools", VoIP MaSe 2006, pp. 28-32

[10] Z. Asgharian, H. Asgharian, A. Akbari, B. Raahemi, "A framework for SIP intrusion detection and response systems", International Symposium on Computer Networks and Distributed Systems (CNDS), pp.100-105, 2011

[11] S. Ehlert, D. Geneiatakis, T. Magedanz, "Survey of network security systems to counter SIP-based denial-of-service attacks", Computers & Security, 29(2), pp. 225-243, 2010

[12] Free Security Scanner for Network Exploration and Security Audits, http://nmap.org/

[13] small comand line tool for developers and administrators of SIP applications, http:// www.voip-info.org /wiki /view/Sipsak.

[14] Active Network Security Tool, http://www.hping.org/

[15] National Vulnerability Database, http://nvd.nist.gov/

[16] free Open Source test tool / traffic generator for the SIP protocol, http://sipp.sourceforge.net/

[17] Open Source SIP proxy/server for voice, video, IM, presence and any other SIP extensions. http://www.opensips.org/

[18] Dorgham Sisalem, John Floroiu, Jiri Kuthan, Ulrich Abend, Henning Schulzrinne, "SIP SECURITY", John Wiley & Sons Ltd, 2009

[19] R. Islam, S. Ghosh, "SIP Security Mechanism Techniques on Voice over Internet Protocol (VoIP) System", International Journal of Computer Applications in Engineering Sciences (IJCAES), Vol I, Issue I, March 2011

[20] A. D. Keromytis, "Voice over IP Security, A Comprehensive Survey of Vulnerabilities and Academic Research", Springer Briefs in Computer Science, 2011

[21] S. McGann, D. C. Sicker, "An Analysis of Security Threats and Tools in SIP-Based VoIP Systems", online document: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.9378, 2006, Last Access: 21-Sept-2012

[22] F. Fankhauser, M. Ronniger, C. Schanes, T. Grechenig, "Security Test Environment for VoIP Research", International Journal for Information Security Research (IJISR), Volume 1, Issue 1, March 2011

Mitra Alidoosti received both her B.Sc. and M.Sc. in Computer Engineering from Computer Engineering Department of Iran University of Science and Technology (IUST) in 2009 and 2012. Currently she is a PH.D candidate in computer engineering. Her research interests are computer network security, penetration testing in computer networks and web pentesting.

Hassan Asgharian received his M.Sc. in Computer Engineering from Amirkabir University of Technology (AUT) in 2009 and He also has got his B.Sc. from Computer Engineering Department of Iran University of Science and Technology (IUST) in 2006. Currently he is a PH.D candidate in computer engineering in IUST and working as a research assistant in the Network Research Group (NRG) of the Research Center of Information Technology (RCIT) of Computer Engineering Department in Iran University of Science and Technology.

Ahmad Akbari received his Ph.D. in Electrical Engineering from the university of Rennes 1, Rennes, France, in 1995. Dr. Akbari is currently an associate professor at Iran University of Science and Technology, Iran. Dr. Akbari works on Speech Processing related research topics (especially speech enhancement) for more than 20 years. His current research interests include Intrusion Detection and Response Systems, VoIP Communications, Next Generation Networks, VoIP and SIP Security and also Data Communications Networks. Dr. Akbari's work has appeared in several peer-reviewed journals and conference proceedings.

Blog Feed Search in Persian Blogosphere

MohammadSadegh Zahedi*

Department of Electrical and Computer Engineering, University of Tehran, Tehran. Iran

sadeghzahedi@ut.ac.ir

Abolfazl Aleahmad

Department of Electrical and Computer Engineering, University of Tehran, Tehran. Iran

aleahmad@ut.ac.ir

Masoud Rahgozar

Department of Electrical and Computer Engineering, University of Tehran, Tehran. Iran

rahgozar@ut.ac.ir

Farhad Oroumchian

Department of Computer Science and Engineering, University of Wollongong in Dubai, Dubai. UAE

farhadoroumchian@uowdubai.ac.ae

Abstract

Recently, user generated content is growing rapidly and becoming one of the most important sources of information in the web. Blogosphere (the collection of blogs on the web) is one of the main sources of information in this category. User's information needs in blogosphere are different from those of general web users. So, it is necessary to present retrieval algorithms to the meet information need of blog users. In this paper, we focus on the blog feed search task. The goal of blog feed search is to rank blogs regarding their recurrent relevance to the topic of the query.

In this paper, the state-of-the-art blog retrieval methods are surveyed and then they are evaluated and compared in Persian blogosphere. Then we introduce our proposed method which is an extension of the voting model that one of the best blog retrieval methods. We have addressed the excessive dependency of the voting model on a list of initially retrieved top N relevant posts by using data fusion methods.

Evaluation of the proposed algorithm is carried out based on a standard Persian weblogs dataset with 45 diverse queries. Our comparisons show considerable improvement over existing blog retrieval algorithms. The results show 67% increase in MAP score for the CompVote data fusion method.

Keywords: Blog Feed Search; Blog Retrieval; Persian Blogosphere; Voting model.

1. Introduction

Emerging Web 2 enables the internet users to share their knowledge in fast and easy manner. This led to a large expansion of the Web content. Weblogs are one of the main technologies of Web 2 that have an important role in content generation on the Web. According to BlogPulse[1] (one of the major blog search engines), there were more than 182 Million Blogs in January 2012. Since the number of blogs is increasing extremely and the scale of the blogosphere has grown dramatically, this phenomenon cannot be ignored by the researchers [1].

Enormous number of blogs and their popularity make answering the users' information needs in the blogosphere context a challenging problem. Users' information needs in the context of the blogosphere are different from those of general web users. Based on a blog search engine query log analysis, Mishne and de Rijke divided blog queries into two categories called context and concept queries [2]. In the context queries, users are looking for named entities with special interest in new events ("find what people say about a given entity"). While in concept queries users are looking for blogs or blog posts related to

one of their topic of interests. For context queries, the users have well defined events that try to access some information and discussions about them. Therefore, these behaviors of users are similar to when they looking for blog posts to read. While for concept queries, the searcher has broader information need, and this is likely to be interested in blogs that they can subscribe to.

In the context of the blogosphere, to answer of context and concept queries, different researches have been done such as: opinion retrieval, topic detection and tracking, top news stories identification, blog post search and blog feed search.

In this paper, we focus on *blog feed search*, also known as *blog retrieval* or *blog distillation* where the goal is to answer the concept queries. The *blog retrieval* task is defined in [3] as follows: "Knowing that users searching in the blogosphere often wish to identify blogs that are about a given topic *X; blog retrieval* algorithms must aim to find those blogs that are principally devoted to the topic *X* over the timespan of the collection at hand".

While blogs share some similar features with traditional web pages, they also have some distinct characteristics in that distinguish blog retrieval from typical ad hoc document retrieval:

---

[1]BlogPulse: http://blogpulse.co/

* Corresponding Author

- Blog retrieval is a task of ranking document collections rather than single documents.
- There exist amount of noise in the blog content and the language of blog posts is informal and conversational
- Blogs have some properties such as post's comment or the time of the post which could be used for blog retrieval

Also in Iran, due to ease of publishing information in weblogs, a large number of internet users are active in this category of social media. According to the latest statistics reported by the popular ranking website (e.g., Alexa[1]), among top 10 websites in Iran, 3 websites belong to blog service providers. Despite the high importance of blogs in Iran, very few studies have been devoted to Persian blogs. In addition, to the best of our knowledge, there exists no research on blog retrieval in Persian blogosphere. Therefore, in this study we try to investigate and compare different blog retrieval methods in Persian blogosphere using *irBlogs* dataset [4].

Most researches in *blog feed search* have been started since 2007, when the organizers of TREC conference initiated a related task. Different methods from similar problems to blog feed search have been proposed like ad-hoc search methods, expert search algorithms or methods from resource selection in distributed information retrieval that are summarized in section 3.

One of the best *blog feed search* methods is *voting model* [5]. The voting model, calculates each blog's score based on a list of initially retrieved top N relevant posts that named R(Q) set, therefore, its accuracy depends on R(Q) set. This method assumes R(Q) contains relevant posts to the query; as a result, if a post is irrelevant in the list or a relevant post is absent in the list, the accuracy of the voting model decreases. To solve this critical problem, in this paper, combined data fusion methods are used to provide the actual relevant posts of R(Q) set.

So, the main contributions of this paper are as follows:

- Cross validation of blog feed search methods using new standard data collection in Persian blogosphere.
- The voting model is extended using combined data fusion methods to provide actual top relevant post in R(Q) set.

The rest of this paper is organized as follows: Persian blogosphere is described in section 2, a general review and classification of blog retrieval algorithms are presented in section 3, our method is presented in section 4 and the experimental setup is discussed in section 5. The results and comparisons are reported in section 6 and finally the paper is concluded in section 7.

2. Persian blogosphere

The exact number of Persian blogs is not known but based on a research in 2005, there exist 700,000 Persian

blogs [6]. Also, the Persian language was introduced as the tenth most widely used language in blog writing by Technorati in 2007 [7]. Currently, Blogger, Persian Blog, Blog sky, ParsiBlog, MihanBlog, Blogfa, IranBlog and Wordpress are known to be the major blogging service providers for Iranian Internet users. Since many Iranian Internet users prefer to be anonymous bloggers, there is no exact statistics about the users' gender, age and population. According to a research by the Harvard University in 2008 [8], most Iranian bloggers are 25-30 years old and among the blog users 24% are women and 76% are men. In [6], the authors used some samples from the blogs to study and analyze the Persian blog contents sociologically. Also, it is claimed that due to the relative freedom of the Persian blogosphere, their content reflects the society's opinion well. Kelly and Etling studied Persian blogosphere from political and cultural point of view [8]. They used automated content analysis tools to cluster Persian weblogs and claimed that the blogs cover four main areas: reformist and political, religious, Persian literature and poetry and blogs with various subjects. The study regarded the blogosphere as the fourth most widely used in the world. Also, in [9] the authors analyzed Persian weblogs from political aspect and conclude that Persian weblogs are a suitable space for Iranian to express their opinion about social matters.

On the other hand, some papers have studied the language style of the Persian blogosphere. A project in MITRE investigated morphological analysis of conversational Persian in weblogs [10]. They could detect many new words that are devised by Persian bloggers. The new words are mainly borrowed from other language like English and French, created by combining English and Persian words or postfixes, or changing already existing Persian words. In [11] the authors analyzed Persian weblogs morphologically and argue that Persian weblogs contain different formal and conversational texts and even the used formal language is different from commonly-used Persian formal language.

3. Related works

In this section, different blog retrieval algorithms will be explained concisely. The following categorization is present for the algorithms proposed previously for the blog retrieval task:

- Resource selection approaches
- Expert finding approaches
- Methods that use structural properties
- Methods that use temporal properties

3.1 Blog feed search using resource selection approaches

If we consider a blog as a collection of posts, then the problem of blog retrieval would be similar to the problem of resource selection in distributed information retrieval [12].

---

[1] Alexa - The Web Information Company: http://www.alexa.com

Elsas et. al., [13] propose large document (LD) and small document models (SD) for blog retrieval. The LD model regards the whole blog as a single document and calculates its relevancy score with regard to any input query. The SD model considers each blog as a collection of posts; after calculating the relevancy score of each post, the blog is scored using a combination of the individual post scores.

Seo & Croft [14] also approached blog distillation as a resource selection problem. Similar to the approach of Elsas et. al., they consider different representations of blogs: Global Representation (GR) and Pseudo-Cluster Selection (PCS). Similar to the LD model the GR model simply treats blogs as a concatenation of their posts. Also, PCS is similar to the SD model of Elsas et al., but it uses a different principle. In PCS, a blog is considered as a query-dependent cluster containing only highly ranked blog posts for the given query.

Lee et. al., [15], present global and local evidences of blog feeds to calculate blog scores, which correspond to the document-level and passage-level evidences used in passage retrieval. They estimated global evidence similar to the SD model of Elsas et al [13] by using all posts within a blog feed. Accordingly, local evidence is estimated similar to the PCS model of Seo & Croft [14] that uses highly relevant posts of a blog feed in response to a given query. They propose a series of methods for evaluating the relevance between a blog feed and a given query, using the two types of evidences.

3.2 Blog feed search using expert finding approaches

The task of expert finding aims at identifying persons with relevant expertise or experience for a given topic. So, if we consider each blog as an individual or expert and consider the content of each blog as expertise of each expert then blog retrieval is mapped into expert finding task. The only difference is that each expert writes only about his/her specialty, but a blogger may write about various subjects.

This connection to the expert search task was explored by Weerkamp et. al., [16] using a language modeling framework. They adopted two expert search models for blog distillation task; in *blogger model*, the probability of a query being generated by a given blog is estimated by representing this blog as a multinomial distribution of terms. In the *posting model*, this probability is estimated by combining the probability of each blog posts generating the query. In both models, the prior probability of choosing a given blog is considered to be uniform as well as the probability of choosing a post given its blog. Macdonald and Ounis [5] propose an approach that is similar to expert finding task. They used different voting models which have been used in [17] for expert finding. Some aggregation methods that they applied to blog distillation are used as our baselines and are discussed in more details later in section 3.

3.3 Blog feed search using structural properties

Some blog feed search methods try to use structural properties of blogs such as inter blog links. In [18] two different graphs are created based on the links which exist in blogs and the link information that is available in individual blog posts. Then they boost post scores using the in-degree of each post and the h-index of each blog. In [19] a graph is created based on blogs, posts and the input query words. Then by performing a random walk on this graph most relevant blogs for each query is extracted. In [20] the effect of content similarity between posts of a blog is investigated and the relevancy score of a post is smoothed based on the connection between the post and the other posts of the blog.

3.4 Blog feed search using temporal properties

Time has been used in blog feed search in various methods. Keikha et. al., [21] used temporal properties of posts in different ways for blog retrieval. They proposed a time-based query expansion method that selects terms for query expansion based on most relevant days for a given topic.

Nunes et. al., used temporal evidence as an extra feature of blogs in addition to their content [22]. They use temporal span and temporal dispersion as two measures of relevance over time, and show that these features can improve blog retrieval. In [23], Keikha et al proposed a framework, named TEMPER, which selects time-dependent terms for query expansion, generates one query for each point of time and calculate a distance measure based on temporal distributions.

Some models are designed to retrieve the blogs that are written about a topic more frequently. Such models show some improvement over the baseline that uses only the content of blogs [24, 25]. MacDonald and Ounis tried to capture recurring interests of blogs over time [5]. Following the intuition that a relevant blog will continue to publish relevant posts throughout the timescale of the collection, they divide the collection into a series of equal time intervals. Then blogs are scored based on their number of relevant posts in different time intervals. Keikha et. al., [26] aim to measure the stability of a blog relevance to a query over time. Their idea is that a blog which has many related posts only during one short period of time, is not highly relevant. Thus they defined the TRS (Temporal Relevance Stability) that scores a blog higher if it has more related posts in more time intervals which is said to possess more stability.

4. Our proposed methods

4.1 Motivation

Voting model has some properties that make it a suitable method for blog feed search. First, the Voting model is efficient methods because it uses individual blog posts indexing.

One of the most important part of blog feed search systems is indexing. In the context of blog feed search, there are three methods for indexing:

- Indexing individual blog posts.
- Indexing a whole blog or homepage of blog. It means concatenating each blog's posts into a single pseudo-document and indexing these pseudo-documents.
- Finally, we can consider hybrid indexing that uses both individual blog posts and single pseudo-document for indexing.

Since the voting model calculates each blog's score based on a list of initially retrieved top N relevant posts, we can use individual blog posts for indexing. So, the individual post index of the voting model allows simple incremental indexing and does not require frequent re-computations of pseudo-documents that are meant to represent an entire blog. Therefore, it has led the Voting model to be a more efficient blog feed search method. Also the Voting model can be used for blog post retrieval.

Information retrieval in the context of the blogosphere usually involves one of the following two main tasks:

- Blog post retrieval: identifying relevant blog posts
- Blog feed search: identifying relevant blogs.

Since the voting model uses individual blog posts score for calculating final score of blog, it can be used also for blog post retrieval.

Finally, the voting model is also an efficient method for blog feed search in Persian blogosphere. NasiriShargh in [27] shows that there exit some differences between Persian weblog and non-Persian language weblog. Due to cultural characteristics of Iranian bloggers, there exists high topical diversity in Persian blogosphere. For example a number of weblogs, that are categorized as "scientific" or "educational", may write about some of the common celebrations such as "Nowruz" or "Yalda" in a separate post. Therefore, a method like large document model or resource selection that considers a whole blog as a document may not have enough accuracy and other models that apply individual post scores such as voting model perform better.

Therefore, in this paper we try to extend voting model to propose efficient blog feed search methods in the Persian blogosphere.

On the other hand, there exist some problems with the voting model: The *voting model* [5], calculates each blog's score based on the posts that are present in the R(Q) set. Therefore, its accuracy depends on R(Q) set. This method assumes R(Q) contains relevant posts of a query; As a result, if a post is irrelevant in the list or a relevant post is absent in the list, the accuracy of the voting model decreases.

In the next section, we extend the voting model to solve this problem. First, the voting method is described and then the proposed solutions for this problem are explained.

4.2 Voting model for blog retrieval

The authors of [5] propose an approach which adopts the voting model of [17] for blog retrieval. It works based on a list of initially retrieved top N relevant posts for a query, named R(Q), which is supposed to be the set of probably relevant posts. Then each blog's score is calculated based on its posts that exist in R(Q). In this model, bloggers are considered as experts in different topics. A blogger that is interested in a particular topic, writes regularly about that topic and it is highly probable that his/her posts are retrieved in response to a related query. In this way, blog feed search can be modeled as a voting process: a post that is retrieved in response to a query is considered as a weighted vote for the expertise of its blogger. In [5], different fusion methods are used to aggregate the weighted votes and finally rank related blogs. *ExpCombMNZ* is the best fusion method in their experiments which is calculated as follows:

$$
\begin{aligned}
Score_{expCombMNZ}(B,Q) \\
= |R(Q) \\
\cap Post(B)| \sum_{p \in R(Q) \cap Post(B)} exp(Content\_Score(p,Q))
\end{aligned} \quad (1)
$$

Where, Post(B) is the set of posts of blog B, $|R(Q) \cap Post(B)|$ is the number of posts from blog B that exist in R(Q) and Content_Score(p, Q) is relevancy score of the content of post P regarding to query Q which is calculated using PL2 retrieval model.

4.3 Proposed method

In order to solve the problem of *voting model*, data fusion methods can be used to improve the accuracy of the R(Q) set. In other words, top 1000 relevant posts are initially retrieved using four different information retrieval models such as BM25 [28], PL2 [28], PL2F [29], LM [30]. Let R(Q)i presents the four retrieved list. Then, we form a new set of posts by using all of the R(Q)i sets and call it *New_R(Q)*. Provided that all R(Q)i sets have no posts in common, the final *New_R(Q)* would contain 1000*4=4000 posts . Then we assign a score to each of the posts in the *New_R(Q)* set using data fusion techniques, and use this new set, named *Optimal_R(Q)* instead of R(Q) in the voting method. But to calculate the score of the posts in *Optimal_R(Q)* set, the following procedure is used:

Since four different information retrieval methods are used to form R(Q)i sets, we initially normalize the scores of the posts in each list in a way that the scores fall in the range of (0-1). Then the final score of each post is computed based on different data fusion techniques. In other words, the results of the four retrieved sets are combined using formulas 3 to 5:

$$
New\_R(Q) = \bigcup_{i=1}^{4} R(Q)_i \quad (2)
$$

$$
Score_{Combvotes}(P,Q) = |P \in New\_R(Q)| \quad (3)
$$

$$
Score_{CombRR}(P,Q) = \sum_{R(Q)_i \in New\_R(Q)} \frac{1}{Rank(P,R(Q)_i)} \quad (4)
$$

$$
\begin{aligned}
Score_{CombMNZ}(P,Q) \\
= Score_{votes}(P,Q) \sum_{R(Q)_i \in New\_R(Q)} Score(P,R(Q)_i)
\end{aligned} \quad (5)
$$

In Formula 3, P represents a post and Q is the user query and $|P \in New\_R(Q)|$ is the number of times that post P appears in $New\_R(Q)$ and for computing the new score of each single post, the total frequency of the post in different $R(Q)i$ sets is used. Each post is scored between 1 and 4; If a post is present in all the $R(Q)i$ sets, it's score is 4. Thus a post will be scored higher if it is retrieved by more retrieval methods.

In formula 4, $Rank(P, R(Q)_i)$ specifies the rank of post P in $R(Q)_i$. In this fusion method, the new score of each post is computed by summing the inverse rank of post P in all $R(Q)_i$ lists. As a result, a post will be scored higher if it is retrieved by more methods or its ranking is higher in any of the methods.

In formula 5, $Score(P, R(Q)_i)$ shows the score of post P in $R(Q)_i$. In order to calculate the new score for each post, the total frequency of the post in all $R(Q)_i$ sets is multiplied by the total score of the post in all $R(Q)_i$ sets. So a post can be scored higher if it is retrieved by more methods and its score is also higher in the lists.

5. Experimental setup

In order to evaluate and compare the presented blog retrieval models, *irBlogs* dataset [4] is used. It is a standard dataset prepared for evaluation of blog retrieval algorithms that includes: (1) a set of blogs together with their posts, (2) a set of standard topics and (3) relevance judgments (ground truth). Table 1 shows some general information about *irBlogs* data set:

Table 1. Statistics of irBlogs collection

Number of blogs	602671
Number of posts	4983365
Average number of posts for a blog	8.2

5.1 Document collection

The document collection of this data set contains permalinks (html version of the blog posts) and feed (xml version). In our experiments the xml version is used.

5.2 Topic set

irBlogs contains 45 manually created queries about different subject categories in TREC standard format, consisting of a title (a few keywords), a description (a few sentences on what the topic is), and a narrative (a short story on which documents should be considered relevant or irrelevant). For our experiments, we are only interested in the title of a topic, which is comparable to a query submitted to a search engine by an end user

The collection also contains manual relevance judgments for each query in a four level scale of: highly relevant, relevant, irrelevant and spam. Table 2 shows the number of relevant and highly relevant blogs for the queries of *irBlogs*. Moreover, the Table 3 compares number of relevant weblogs for topics in *irBlogs*, *TREC2007* and *TREC2008* collections. This table shows

that these collections are similar in terms of the number of relevant weblogs per topic. Note that good topics are those that have enough relevant blogs, so that different algorithms can be compared fairly based on them.

Table 2. Some statistics about relevance judgments of irBlogs

Number of query	**45**
Average high relevant blogs per query	23.28
Average relevant blogs per query	39.04
Number of high relevant blogs	1048
Number of relevant blogs	1773

Table 3. Categorization of topics based on the number of relevant weblogs in Trec and irBlogs data set

Topics with . . .	Trec 2008	Trec 2007	irBlogs
<5 relevant blogs	5	0	1
<10 relevant blogs	11	5	2
<20 relevant blogs	20	12	9
>100 relevant blogs	3	6	11
Between 5 and 100 relevant blogs	27	27	25

5.3 Inverted indexes

The collection is indexed by the Terrier (version 3.5) open source software package. Statistics of the created index are summarized in Table 4.

Table 4: Statistics of irBlogs index

Number of posts	**4746536**
Number of blogs	388994
number of tokens	598659118
size of vocabulary	6019379
Index size	6 GB

5.4 Evaluation and Comparison Criteria

For evaluation of our proposed method, various standard metrics are utilized that are commonly used for comparing blog retrieval algorithms. The metrics are: R-Precision, Mean Average Precision (MAP), Precision at rank 5 (P@5), Precision at rank 10 (P@10) and Normalized Discounted Cumulative Gain (NDCG). Also, in the following comparisons, MAP1 denotes MAP of a run, when only the highly relevant blogs are considered to be relevant and MAP2 denotes the MAP of a run, when both highly relevant and relevant blogs are considered to be relevant. In order to compare the proposed method with other existing methods, some of the best known blog retrieval methods are chosen. The algorithms and their parameters are listed in Table 5:

Table 5. The best blog feed search methods that exist in the literature

Name	Parameters Value	Blog feed Search Method
Blogger	β_{Bloger}= average Blog length	Blogger Model[16]
Posting	$\beta_{posingt}$= average post length	Posting model[16]
Voting	C=1	Voting Model[5]
Local	K=2,α=0.62	Local evidence [15]
Large	Feed Prior=Uniform,μ=2500	Large Document [13]
Resource Selection	M=4,π=1	Resource Selection +Diversity Penalty[14]
Temp	α=0.9	Temporal Evidence[22]

6. Experimental results

Performance of several blog retrieval methods (listed in table 5) are analyzed based on top 1000 retrieved blogs.

6.1 Evaluation and comparison of the state of the art methods

Figure 1 compares already existing blog retrieval methods of table 5 based on MAP2 for TREC 2007 data set. As it is obvious *local evidence* [15], *resources selection* [14] and *large documents model* [13] have the best accuracy. Also, Figure 2 provides a comparison in the performance of those methods based on the measures discussed before using the irBlogs dataset. One can observe that ranking of the methods are almost the same on both collections.

Figure 3 indicates the MAP1 scores of a number of queries for a more in-depth analysis. Figure 3.a indicates that for general queries, with a large number of highly

relevant blogs, such as "learning to cook", "mountaineering training", "women's veil" and "stock market analysis", *large documents* and *resource selection models* perform better. This could be attributed to the fact that there are very few irrelevant posts in these highly relevant blogs. Thus, the *large documents model* that considers the whole blog as a document performs better for such queries. Figure 3.b indicates that for more specific queries, that have a few highly relevant blogs, such as "Computer M.Sc. Entrance exam", "Urmia Lake drying", "Fajr film festival" and "Tehran international book fair", other models that apply local post scores such as the *voting model* or the *local evidence model* perform better. The score of a blog is calculated based on its individual post scores in these models. Since the majority of these blogs are relevant (not highly relevant); so, a method like *large document model* assigns a lower score to these blogs due to the existence of irrelevant posts.
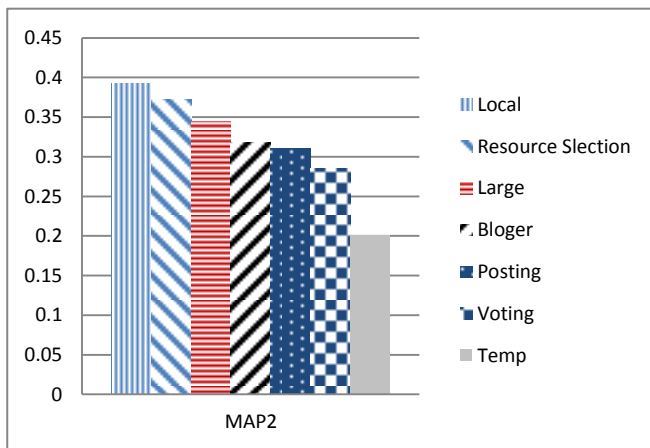
Fig 1. Evaluation results for the existing blog retrieval methods on TREC07 data set
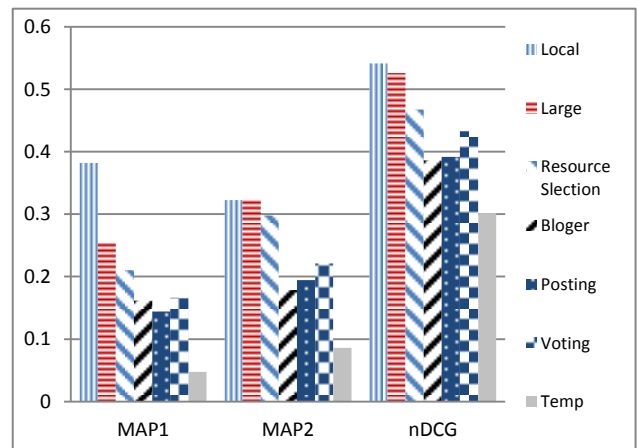
Fig 2. Evaluation results for the existing blog retrieval methods on irBlogs data set

(a)

(b)
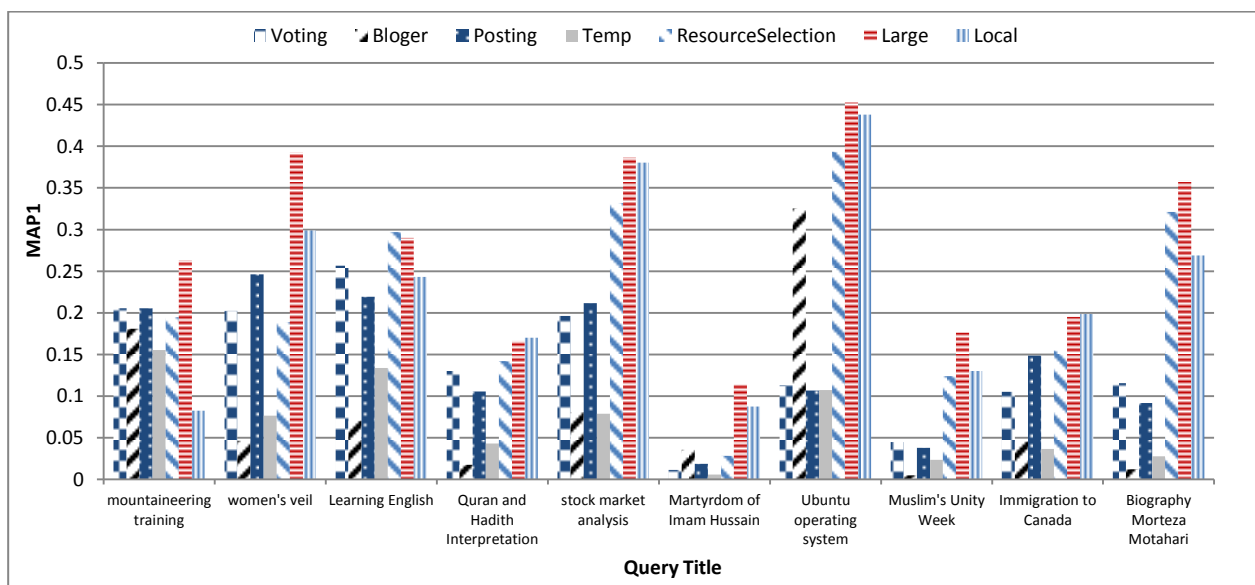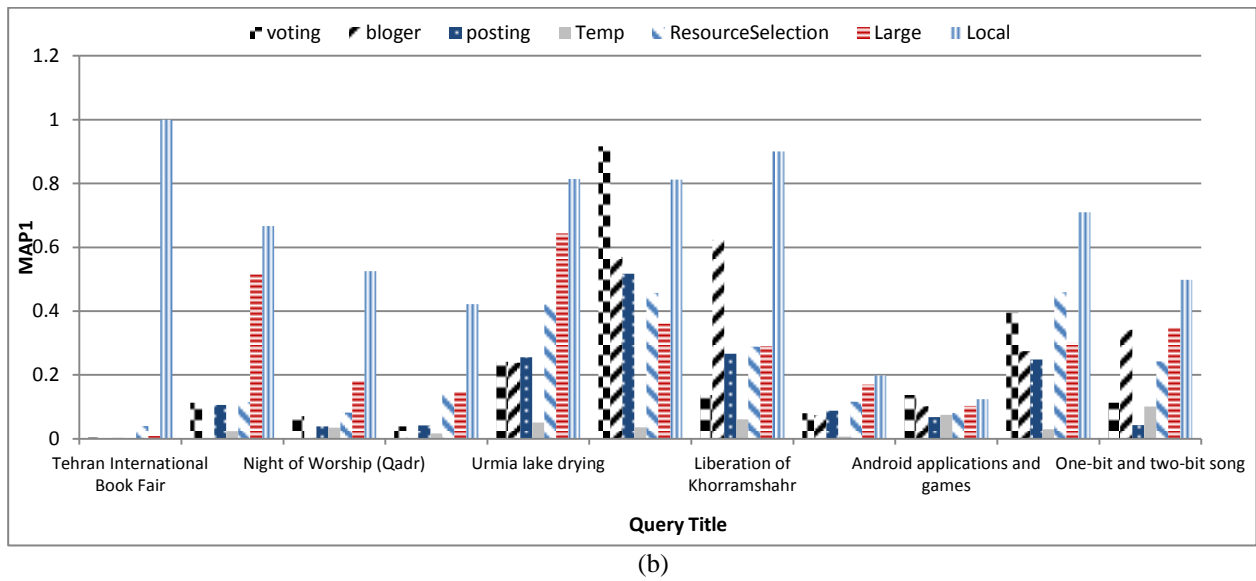
Fig 3. Evaluation results of blog retrieval methods per topic (a) general topics (b) more specific topics of irBlogs

6.2 Evaluation and comparison of the proposed method

Table 6 summarizes the results of the *voting model* for different values of N for initially retrieved top N relevant posts of R(Q) set. As it is shown, by increasing N, the accuracy of the voting method increases but when N reaches near to 4000 the accuracy remains constant and does not increase. The best performance of the *voting method* is obtained with N = 4000 and for the rest of the experiments we have used these optimal value.

Table 6. Accuracy of voting model with different values of N for R (Q) set

	Map1	Map2	NDCG	R_prec	P@5	P@10
N=500	0.157	0.194	0.364	0.244	0.417	0.36
N=1000	0.166	0.221	0.410	0.264	**0.426**	0.375
N=2000	0.171	0.233	0.437	0.273	0.395	0.377
N=3000	**0.174**	0.237	0.445	0.280	0.408	**0.386**
N=4000	**0.174**	**0.239**	**0.450**	**0.281**	0.413	0.382

Figures 4 and 5 show the results of the proposed methods using data fusion for different measures. As it is shown, all the proposed methods outperform the original *voting model* based on the aforementioned measures.

Table 7 shows the improvement percentage achieved by the proposed methods in comparison with the original voting model. As one can observe, *CombVote_fusion* method gained the highest improvement rate.
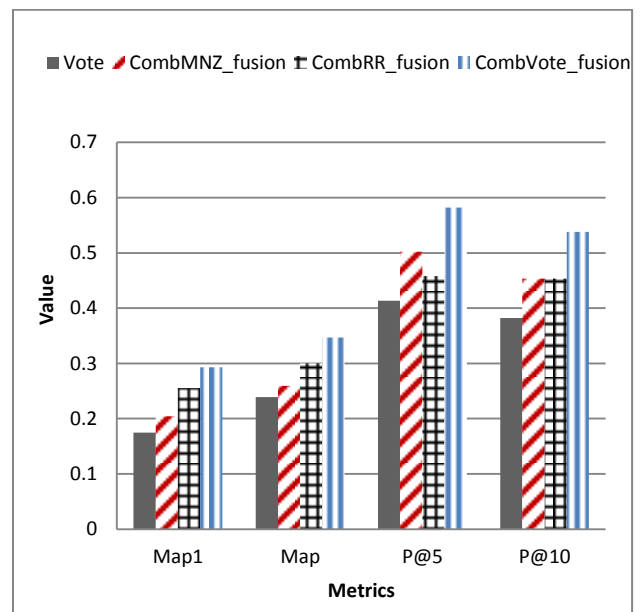
Fig 4. Evaluation results for the proposed methods and original voting model
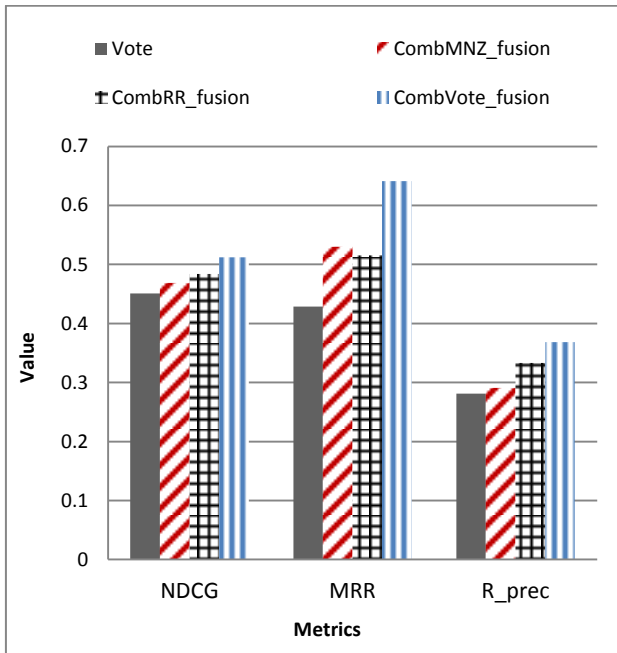
Fig 5. Evaluation results for the proposed methods and the original voting model

Table 7: Improvement percentage of proposed methods compare to the original voting model

	Percent improvement compared with Voting model				
	Map2	Map1	NDCG	P@5	P@10
CombVote_fusion	**44.98%**	**67.50%**	**13.59%**	**40.86%**	**40.71%**
CombRR_fusion	25.29%	45.99%	7.27%	10.76%	18.60%
CombMNZ_fusion	8.44%	16.70%	3.9%	21.50%	18.60%

For an in-depth analysis, Figure 6 and 7 provide per topic comparison of *CombVote_fusion* methods and the original voting model based on the MAP1 measure.

Figure 6 shows the results of the queries for which the proposed method performed better than the voting method. As it can be seen, the proposed method shows significant improvement for most of the queries. Also, Figure 7 indicates queries for which the proposed method shows lower accuracy; even for these queries, the proposed method is comparable with the original voting model. In the latter case, the original voting model turns out to have insufficient accuracy; therefore, the proposed method cannot obtain good accuracy.
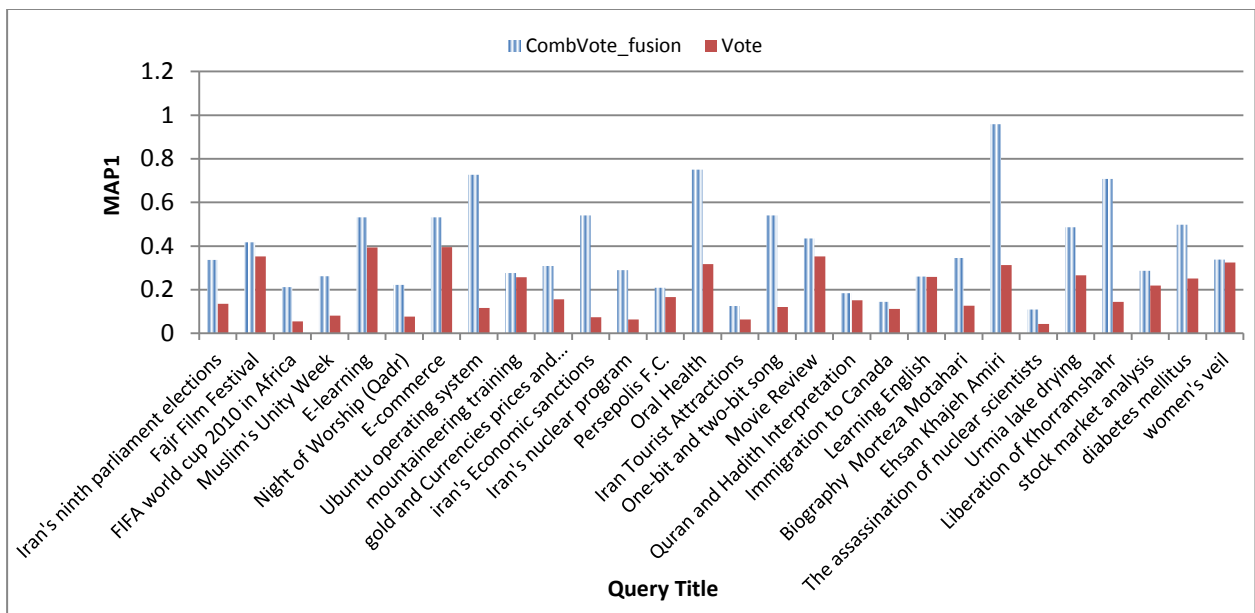
Fig 6. Per topic comparison of *CombVote_fusion* method and the original voting model based on MAP1 scores
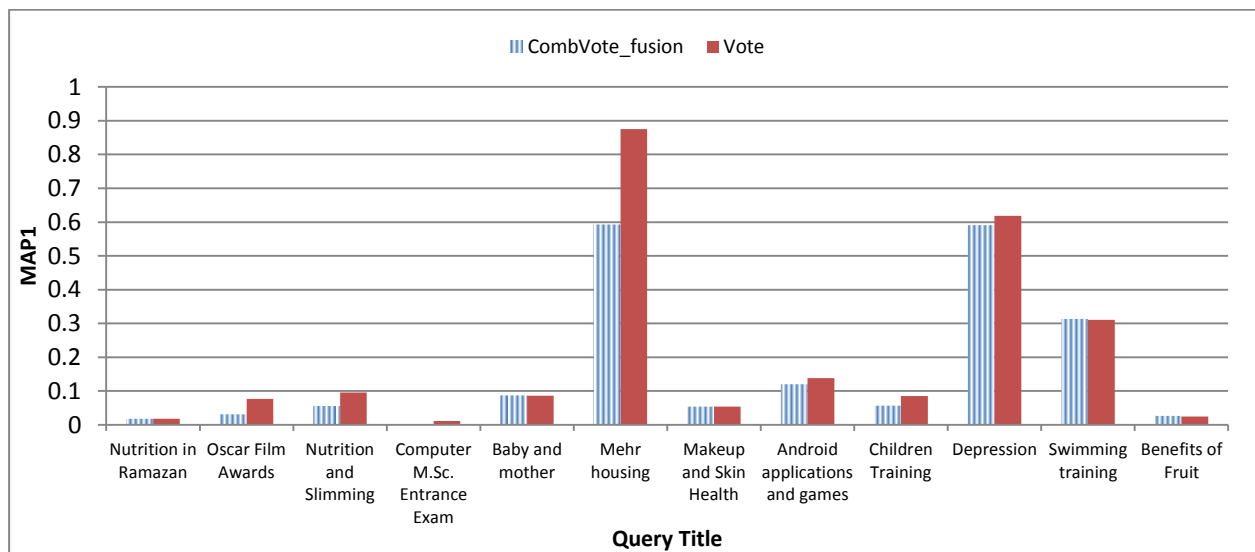
Fig 7. Per topic comparison of *CombVote_fusion* methods and the original voting model based on MAP1

7. Conclusions

In this paper, different blog retrieval methods were compared and evaluated by use of a standard Persian weblogs dataset. The results showed that the best state of the art methods shows very similar accuracy both on English and Persian blogosphere, however their ranking shows a few changes. Also, we can conclude that for general queries, methods such as large document model and resource selection model are better, but for specific queries, local method and the voting model showed better results.

Also, the voting method was developed in this paper; the problem is excessive dependency of the voting model on the R(Q) set that was solved using combined data fusion methods. Our experimental results obtained showed considerable improvement over the voting model and other blog retrieval methods.

References

[1] P. W inn. State of the Blogos pher e, intr o duction, 2008. http://technorati.com/blogging/article/state-of-the-blogosphere-introduction

[2] G. Mishne and M. de Rijke. A study of blog search. In CIR , pages 289{301, 2006.

[3] C. Macdonald, R. L. Santos, I. Ounis, and I. Soboroff, "Blog track research at TREC." In ACM SIGIR Forum, vol. 44, no. 1, pp. 58-75. ACM, 2010.

[4] A. AleAhmad, M. Zahedi, M. Rahgozar, and B. Moshiri, "irBlogs: A Standard Collection for Studying Persian Weblogs , Journal of Language Resources and Evaluation, Springer, submited on December 2013.

[5] C. Macdonald, and I. Ounis, "Key blog distillation: ranking aggregates." In Proceedings of the 17th ACM conference on Information and knowledge management, pp. 1043-1052. ACM, 2008.

[6] N. Alavi, We are Iran: Soft Skull Press, 2005.

[7] "The State of the Live Web, April 2007," http://www.sifry.com/alerts/archives/000493.html.

[8] J. Kelly, and B. Etling, "Mapping Iran′ s Online Public: Politics and Culture in the Persian Blogosphere," Berkman Center for Internet and Society and Internet & Democracy Project, Harvard Law School, 2008.

[9] S. Golkar, "Politics in Weblogs: A Safe Space for Protest," Iran Analysis Quarterly, vol. 2, no. 3, 2005.

[10] K. Megerdoomian, "Analysis of Farsi weblogs," MITRE Corporation, Washington DC2008.

[11] K. Megerdoomian, "Extending a Persian morphological analyzer to blogs," presented at the Zabane Farsi va Rayane [Persian Language and Computers], Tehran, Iran, 2010.

[12] J. Callan, "Distributed information retrieval," Advances in information retrieval, vol. 5, pp. 127-150, 2000.

[13] J. L. Elsas, J. Arguello, J. Callan, and J. G. Carbonell, "Retrieval and feedback models for blog feed search." In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, pp. 347-354. ACM, 2008.

[14] J. Seo, and W. B. Croft, "Blog site search using resource selection." In Proceedings of the 17th ACM conference on Information and knowledge management, pp. 1053-1062. ACM, 2008..

[15] Y. Lee, S.-H. Na, and J.-H. Lee, "Utilizing local evidence for blog feed search," Information retrieval, vol. 15, no. 2, pp. 157-177, 2012.

[16] W. Weerkamp, K. Balog, and M. de Rijke, "Blog feed search with a post index," Information retrieval, vol. 14, no. 5, pp. 515-545, 2011.

[17] C. Macdonald, and I. Ounis, "Voting for candidates: adapting data fusion techniques for an expert search task." In Proceedings of the 15th ACM international conference on Information and knowledge management, pp. 387-396. ACM, 2006.

[18] C. Ribeiro, "FEUP at TREC 2010 Blog Track : Using h-index for blog ranking," Proceedings of   TREC 2010

[19] M. Keikha, M. J. Carman, and F. Crestani, "Blog distillation using random walks." In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 638-639. ACM, 2009.

[20] M. Keikha, F. Crestani, and M. J. Carman, "Employing document dependency in blog search," Journal of the American society for information science and technology, vol. 63, no. 2, pp. 354-365, 2012.

[21] M. Keikha, S. Gerani, and F. Crestani, "Time-based relevance models." In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pp. 1087-1088. ACM, 2011.

[22] S. Nunes, C. Ribeiro, and G. David, Feup at trec 2008 blog track: Using temporal evidence for ranking and feed distillation, DTIC Document, 2008.

[23] M. Keikha, S. Gerani, and F. Crestani, "Temper: A temporal relevance feedback method," Advances in Information Retrieval, pp. 436-447: Springer, 2011.

[24] B. Ernsting, W. Weerkamp, and M. de Rijke, "Language modeling approaches to blog post and feed finding," 2007.

[25] W. Weerkamp, K. Balog, and M. De Rijke, "Finding Key Bloggers, One Post At A Time." In ECAI, pp. 318-322. 2008.

[26] M. Keikha, S. Gerani, and F. Crestani, "Relevance stability in blog retrieval." In Proceedings of the 2011 ACM Symposium on Applied Computing, pp. 1119-1123. ACM, 2011.

[27] Aidin NasiriShargh, "Suggesting and Evaluating a New Content-Based Measure to Find Similarity between Persian Weblogs", MSc. Thesis, Sharif University of Technology, 2009.

[28] G.Amati "Probability models for information retrieval based on divergence from randomness." PhD diss., University of Glasgow, 2003.

[29] C. Macdonald, V. Plachouras, B. He, C. Lioma, and I. Ounis. University of Glasgow at WebCLEF-2005: Experiments in per-field normalisation and language specific stemming. In Proc. of CLEF 2005

[30] Zhai & Lafferty,"A Study of Smoothing Methods for Language Models Applied to Information Retrieval".ACM Transactions on Information Systems, Vol. 22, No. 2, Pages 179—214, April 2004.

Mohammad Sadegh Zahedi received the B.S. degree in Software Engineering (2010) from the University of Zanjan. He is currently M.S. student in Department of Electrical and Computer Engineering, University of Tehran, Iran and member of Database Research Group (DBRG). His area research interests include information retrieval, social network analysis and data mining. His email address is: sadeghzahedi@ut.ac.ir

Abolfazel Aleahmad received the B.S. degree in Computer Engineering (2003) from the Bahonar University of Kerman and M.S. degree in Computer Engineering, major in Information Technology (2008) from the University of Tehran. He is currently Ph. D student in Department of Electrical and Computer Engineering, University of Tehran, Iran. He has published more than 30 journal/conference on his research findings. His area research interests include information retrieval, graph theory, social network analysis. His email address is: aleahmad @ut.ac.ir

Masoud Rahgozar, He received B.Sc. degree on Electronics Engineering from Sharif University of Technology in Tehran in 1979 and, He received Ph.D. on Database Systems from Pierre and Marie Curie University in Paris in 1987. He is Associate Professor in the Department of Electrical and Computer Engineering, University of Tehran, Iran. He has over 19 years of professional career in French software house companies as R&D manager, senior consultant and published more than 100 journal/conference and book chapters on his research findings. His current fields of interests are: Database Systems, designing CASE tools for Object Oriented programming, designing CASE tools for Database normalization and modernization of legacy applications and their environments. His email address is: rahgozar@ut.ac.ir

Farhad Oroumchian received the B.S. degree in Computer Science from National University of Iran(1984) and M.S. degree in Computer Science (1987) from the Sharif University of Technology and Ph.D. Degrees from School of Computer & Information Science, Syracuse University (1995). He is an Associate Professor in the Faculty of Computer Science and Engineering, University of Wollongong in Dubai. His research specialty is in Artificial Intelligence, Information Retrieval and Natural Language Processing. He has developed an intelligent search engine which mimics Human reasoning in order to find relevant documents. He is also a prominent researcher on Persian text processing and retrieval and multi-lingual search engines. He has published more than 90 journal/conference and book chapters on his research findings. He serves on technical/program committees of many international journals and conferences. He is also a member of several Centers of Excellence and Research groups. He has worked as project manager and consultant for many years in U.S.A and Iran. His previous posts include Associate Dean of Faculty Computer Science and Engineering, University of Wollongong in Dubai, Chair of the Department of Software Engineering at the University of Tehran, Manager of IT Services in the Faculty of Engineering, University of Tehran, and Senior Research Engineer TextWise LLC., in Syracuse New York and Database/Computer Consultant for Defelsko Inc, Ogdensburg, New York.

Trust Evaluation in Unsupervised Network: A Fuzzy Logic Approach

Golnar Assadat Afzali
Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran
gafzali@kntu.ac.ir
Monireh Hosseini*
Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran
hosseini@kntu.ac.ir

Abstract

Because of the possibility of anonymity and impersonation in social networks, trust plays an important role in these networks. In social networks, trust can have two aspects: trust of users to social network and trust of users to other users. Pear to pear networks, by eliminating the supervisor roles, besides its benefit in decreasing management costs, have problems in trust and security of users. In these networks, trust evaluation is only related to the trust of pear to other pear and because of the direct relation between pears; each user should know the trust level of other users. However, trust evaluation in pear-to-pear networks (as an unsupervised network), only can be done based on the past relation between pears or trust evaluation of other pears. This kind of trust evaluation cannot give a comprehensive view to pears. In other word, if any pear is not in the friend cycle of a user or friend cycle of user's friends, he will not be able to assign appropriate trust level to this pear. In this research, by using social networks as supervised networks, trust level of each user is evaluated, then by identifying these users in unsupervised networks, appropriate trust level is assigned to them.

Keywords: Trust; Unsupervised Networks; Trust Factors; Fuzzy Logic.

1. Introduction

With increasing growth of using internet, users confront with many problems in security and privacy [1,2,3,4]. Due to the lack of face to face relationships and simplicity of impersonation in these networks, distribution of incorrect information is increased [5,6]. So, considering only shared information of users cannot be a good criterion for measuring trust. Therefore, trust and evaluation of user's trust level is taken into consideration [5,7,8,9,10]. In general, trust is influenced by many factors, such as shared information of user, past interactions with his, positive or negative comments of other users and etc. [11]. Besides, trust in unsupervised networks, -such as pear-to-pear networks- is more sensitive. In these networks, lack of supervision makes tracking user's behavior impossible. On the other hand, because of the nature of activities in these networks, trust is more important. For example, in many cases, users allow others to run programs on their systems; therefore, inappropriate behavior of users can cause serious problems. Many researches have been done on trust in unsupervised networks [12,13,14,15]; nevertheless, in most of them, trust is considered as a static parameter. This means that calculation of user A's trust in time t, would be done based on accessible data about him on that time, such as shared information in his profile, affiliations and membership information of groups. But in real, trust is a dynamic concept and in addition to considering its static aspects, it must be updated during time and based on user's activities.

In this paper, with considering both static and dynamic aspects of trust, the trust level of users would be determined. Then by identifying user on other unsupervised networks, his trust level would be assigned. Therefore, reliable and secure relationships can be established in these networks.

The reminder of this paper is organized as follow: Section 2 provides definition of trust and an overview on related researches in trust. In section 3, after specifying trust factors, related weights are computed. Then, in section 4, evaluated trust in supervised networks is used to predict appropriate trust level of user in unsupervised network. At last, section 5 concludes the paper.

2. Related work

2.1 Trust

Trust is a critical component in human relationships and consequently in social networks [16]. In general definition, trust is a measure of confidence that an entity will behave in an expected manner [17]. Trust-based community is a community that people can share their opinions without any concern about privacy or false judgment of others. Social trust is concerned as a foundation for creating trust-based community in social networks [18]. Social trust is influenced by many factors,

such as past experiences with user, opinion of other users about him, psychological factors and etc. [17]. Difficulty of defining trust and converting it into quantifiable format causes problems in combining trust with algorithms and mathematical analysis [18]. Trust is an asymmetric parameter, means that in a relationship, the trust level of two nodes can be different. Also, trust is a context-aware concept and dynamic [19].

Many researches have been done on trust. PearTrust model [19], tries to determine trust level of users based on three parameters and two adaptive factors. The parameters are received feedbacks from other users about him, transactions and assigned trust level to him by others. The factors are the context of transactions and network environment.

Walter et al. [20], tried to calculate indirect trust between users based on the direct trust between neighbor nodes in social networks. For this, any node can assign a trust level to other nodes in   [-1, 1]. Finally, based on the recommended trust level of any neighbor node and the weight of link between them, user's trust level is determined.

The proposed model of Borbora et al. [10], shows that factors such as shared personal information, node's location within network and other social interactions with nodes in this social network are the most trust influencing factors between users.

New algorithm proposed in Xin et al. model [21], uses indirect trust between users. Indirect trust value is determined depending on direct trust values and trust chain between users that are not neighbors. In this algorithm any node can rate other's trust value in [0, 1].

2.2 Trust in unsupervised social network

Trust in supervised and unsupervised networks has fundamental differences. In unsupervised networks, trust is determined in absolute correct parameters. A file either is impaired or not. A protocol is entirely implemented or not. However, in supervised networks, trust value can have a wide range [22].

In the proposed model of Wang et al. [23], a trust matrix is considered, which any node can rate other nodes. So, any node can easily evaluate each node based on others rates.

Huang et al.[24], emphasize on the role of feedback in building trust between users. In these systems, usually assumed that normal pears can have standard and ideal feedback behaviors. In this paper, instead of direct feedback from users, duration that downloaded file remains in the shared folder is used to determine trust level.

Zhen-wei et al. [25], based on general characteristics of trust between P2P networks (as unsupervised networks) and social networks (as supervised networks), proposed a model to evaluate trust in social networks and utilized it in P2P networks. So, based on user past performance, trust degree is calculated and then assigned to user in P2P network. In this model, static 0-1 view is obvious. However, by using fuzzy logic approach, trust evaluation

would be more précised [26]. For example, it is possible that two users which are in the same trust level have high difference; but with membership degrees, these differences would be more perceptible. In fuzzy logic approach, trust evaluation would be based on a vector of membership degrees related to each trust level.

In [27], Han et al. proposed a topological potential weighted community-based recommendation trust model (IPCommuTrust). In this model, besides considering node's reputation, its status would be effective in determining its trust level. Status of node is based on the shortest path between them.

Yu et al. [28], proposed a model to provide trust evaluation in social network. In this model, trust level of user would be determined based on the referral trust evaluation is given to any requestor user and this can be used beside user's prior experiences. Both direct and indirect trust evaluation will be updated during time.

Li et al. [29], with defining trust as a complex and multi-dimensional parameter, tried to determine these dimensions and effective weights related to them. The main advantage of this model is that trust degree assigned to each user could be changed over time by updating the weights. It should be considered that the mentioned problem in [25], also exits in this proposed model.

The proposed model in [13], contains three factors including quality, popularity, and size of the shared file as trust factors. Then, a fuzzy inference system is used to design P2P reputation management system, which generates the reputation values for users by interacting with other peers, and based on these factors.

In [14], Chen et al., proposed a two-part model containing direct trust and the assigned trust by other users. Once mutual peers recognize each other, they can evaluate other's trust level. In cases which peers don't recognize each other, they could use other's recommendations. This model easily could be hacked. Consider the case that hacker "A" wants to hack user "B". At first step, A gets in the relationship with some friends of B and tries to increase his trust level in their relationship with his trustworthy behaviors. Now, in direct relationship between A and B, because of not existing of mutual recognition, trust evaluation would be done based on mutual friends. So, user A can easily hacks user B.

In our model, trust would be evaluated based on defined trust factors, which each has an appropriate weight, and with appropriate weights for each of them. Due to unawareness of hackers of these factors and their related weights, they could not increase their trust level by fake information. The problem, which exits in many related research studies, is the static consideration of trust. However, trust is a dynamic parameter and should be updated over time. In this paper, trust is considered as a dynamic parameter. So, fake behaviors of users in period of time for getting high level of trust don't have an essential effect on evaluated trust. Table 1 summarizes related works in this manner.

3. Proposed model for computing trust in unsupervised networks

As mentioned, social networks with their virtual relationships increase the importance of trust factors determination. On the other hand, because of not existence of supervisor in unsupervised networks, an integrated view of user's behavior would not be available. To solve this problem, the role of supervisors in supervised networks can be used to evaluate trust level of users and then the results will be used in unsupervised networks. Our proposed model is based on the pattern demonstrated in figure1.
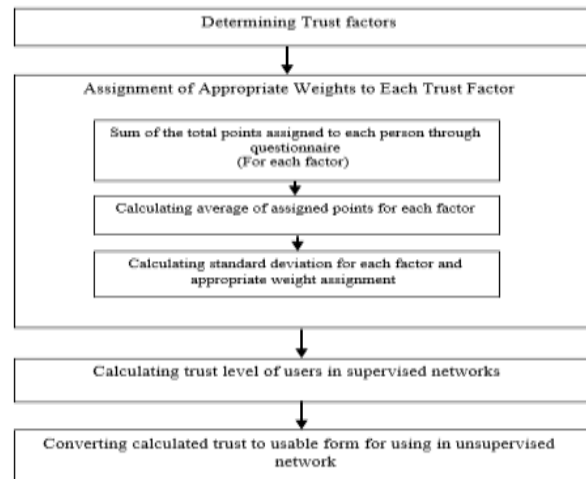

Fig 1. Considered steps of the proposed model

Table 1. Summary of related works

Reference	Publish year	Main purpose	Addressed issues	Not addressed issues
[20]	2009	Determining effective factors in trust	-Node location in network (relation) - Node properties such as personal information, activity related to other users (for example sending message to others)	-Inter-activity (such as updating his profile, etc.) -Behavioral (long-term performance of user interacting with other users)
[13]	2009	Trust evaluation between non-adjacent nodes	-Personal information -Behavioral	-Intra-activity (activity related to other user, such as sending message to others) -Inter-activity
[25]	2010	Trust evaluation between non-adjacent nodes by trust chain between users and the trust of user to his friends	-Reputation (personal information) -Behavioral	-Inter-activity -Intra-activity
[12]	2011	Effective factors in trust	-Personal information -Inter-activity -Intra-activity -Relation	-Behavioral
[29]	2011	Users' feedback for evaluating trust	-Feedback of users, based on retention time of downloaded files in shared folder	- Personal information -Inter-activity -Intra-activity -Behavioral -Relation
[7]	2012	Determining trust level of users in social networks and using it in P2P networks	-Personal information -Behavioral	-Relation -Inter-activity -Intra-activity
[24]	2012	Trust evaluation and updating it by considering dynamic weights for its effective factor	-Feedbacks of other users -Intra-activity -Behavioral -Risk of interaction -Availability of user	-Inter-activity -Relation
[21]	2013	Using fuzzy logic to form trust vector for any shared file by user	- Attributes of shared files (quality, popularity, size of file)	-Personal information -Inter-activity -Intra-activity -Relation -Behavioral
[10]	2013	Trust evaluation using two approach: direct trust (based on past interaction of users), indirect trust (a mutual and trusty node)	-Behavioral	-Personal information -Inter-activity -Intra-activity -Relation
[27]	2014	Trust evaluation using reputation and status of users.	-Relation -behavioral	-Personal information -Inter-activity -Intra-activity
[28]	2014	Trust evaluation based on direct past experience and recommendation from corresponding recommendation peers	-Behavioral	-Personal information -Inter-activity -Intra-activity -Relation
Proposed model		Trust evaluation in supervised network and using it in unsupervised networks	-Personal information -Inter-activity -Intra-activity -Relation -Behavioral	- Feedbacks of other users

3.1 Trust factors determination

Fong et al. [7], in order to determine trust factors, have proposed a hierarchy similar to "trust model" which considered in the model of Gilbert et al.[30]. This hierarchy is included all trust factors in an acceptable level. The main problem in this model is the static view of trust, regardless of the user's past behavior. For this, we will extend this hierarchy with considering behavioral factor. So, our model will be based on the hierarchy similar to figure 3. Each of the factors will be described in detail in the next section.

3.1.1 Analytical factors

3.1.1.1 Relation

In this factor, the relation link between two nodes i and j is regarded as the criterion for rating friends. So, the trust network is plotted. In this network, each relation between nodes will be drawn with a weighted link. This weight can be assigned based on the type of relation between them [7]. For example, in figure 2, node i determines node j as a "close friend". So, the weight of this link can be equal to 5. In this network, two points should be considered:

- Number of weights, which assigned to each group of friends, does not have any effect on solution comprehensiveness. The only notable matter is that all of these weights should be assigned based on the same pattern (similar to table 2).

- Trust evaluation in this network will be done base on the shortest path between two nodes i and j.

Table 2- An example of relation link weighting in trust network

Type of relation	Weight
Family	6
Close friend	5
Friend	4
Co-worker	3
Friend of friend	2
Others	1
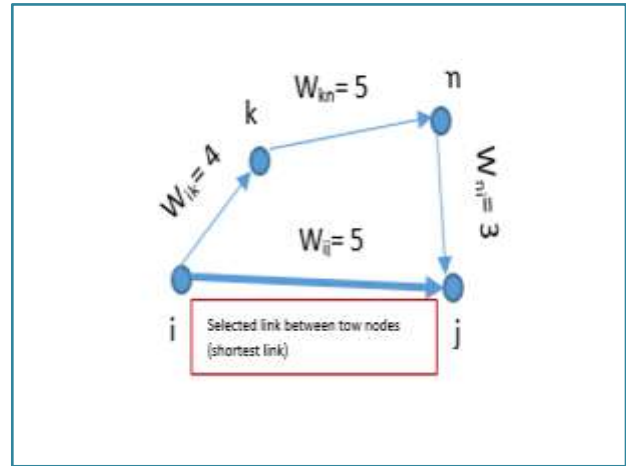
Fig 2. Calcualting relation factor in trust evaluation
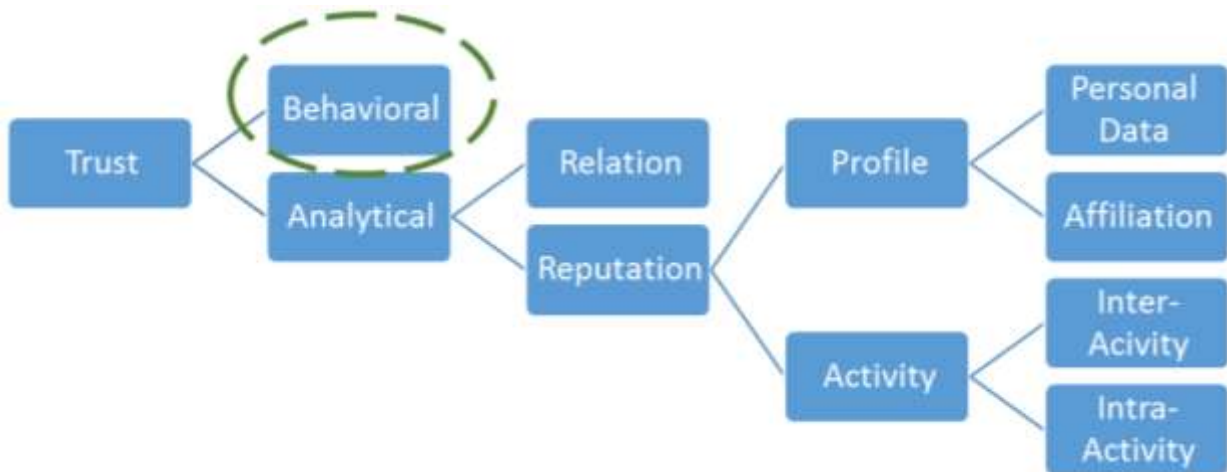
Fig 3. The proposed trust hierarcy

3.1.1.2 Reputation

This factor usually emphasizes on the opinion of others about any user. In social networks, because of the virtual form of relations, this parameter can have an important role. This attribute is related to personal information in user's profile and history of s activity. So, this factor can be studied in two parts: "profile" and "activity"[7].

- Profile:
  This factor mainly focuses on the effect of user's profile on others mindset. This information can present user's affiliation besides personal attributes [7, 22, 25].
- Activity
  User activities are divided in two categories: inter-activity and intra-activity. Inter-activities are related to user's activities such as personal posts,

comments provided on topics and etc. However, intra-activities are activities such as post on other's wall, visiting other's profile and etc. [7, 10].

3.1.2 Behavioral

Behavioral factors considered to add dynamic aspect to analytical factors. The main idea of these factors is related to the dynamic treat of trust. In other word, considering only personal information or user's affiliations cannot be acceptable for trust evaluation; but user activity should be tracked over time [13,20,21,25,29]. This factor consideration, in addition to ensure reasonable trust evaluation, will increase system security. A known attack is that, hacker increases his trust level by providing appropriate values for each trust feature. But, with behavioral factors, besides these features, real behavior of any friend is tracked. So, this kind of attacks can be prevented by using this factor.

3.2 Factors weighting

In previous section, trust factors were determined. Now, the appropriate weights of each factor should be calculated. So, with a questionnaire, some users of social networks were asked to with considering five trust levels, such as "high trusty, trusty, no comment, almost untrusted, untrusted ", assign at least one of their friends to each level and answer to the related questions.

After gathering questionnaire results, weight of each factor is determined. Then, based on formulas 1, 2, and 3, trust degree of users are calculated and finally the range of each trust level is shown by a fuzzy membership function.

Suggested questionnaire has five subdivisions and each of them is related to one of the trust factors. If users which participate in our questionnaire are shown with P_j (j=1,2,..,r) and friends which assigned by users in each trust level are shown with P_{jm}(m=1,2,3,4,5), for each trust factor, appropriate values are assigned to each P_{jm}, based on the answers of P_j in each related subdivision. At first, in order to extract rules about trust factors in any trust level, users with different trust level are considered as an integrated set, regardless of their trust level. To evaluate factor's weight, for all P_{jm}, sum of each factor value are calculated and then the related average is calculated based on formula 1.

$$\lambda i = (\sum_{j=1}^{r} \sum_{m=1}^{5} (Pijm\,))/n \qquad (1)$$

In formula 1, λ_i is the average of factor i, P_{ijm} is the value of factor i for user P_{jm} and n is the number of all P_{jm}.

At last step, by calculated average in formula1, the standard deviation is computed using formula 2.

$$\delta I = \sqrt{\frac{1}{n} * \sum_{j=1}^{r} \sum_{m=1}^{5} (Pijm - \lambda i)^2} \qquad (2)$$

In formula 2, δ_i is the standard deviation of factor i, P_{ijm} and λ_i, similar to formula 1, are user's value and average related to factor i.

Now, based on calculated standard deviation, appropriate weight is assigned to each factor i. In other word, standard deviation can express the importance of each trust factor. More precisely, standard deviation of each factor demonstrates the cohesion of users on this factor. So, if this value is low, it could be concluded that most of users are approximately same in this factor and this weight should be more and if the standard deviation is high, related weight should be low. Table 3, can be used as a criterion for factor weighing.

Table 3. Weight assignment based on standard deviation

δ	W
0 – 0.2	1
0.2-0.4	0.9
0.4-0.6	0.8
0.6-0.8	0.7
0.8-1	0.6

Finally, by formula 3, trust degree of each user is calculated.

$$T_{jm} = \sum_{i=1}^{k} Wi(/\frac{Pijm - \lambda i}{\lambda i}/) \qquad (3)$$

In formula 3, T_{jm} is the trust degree of user P_{jm} and P_{ijm}, λ_i and W_i are user P_{jm} trust value, average of factor i and the weight of factor i, respectively. K is the number of trust factors. (In proposed model, k is equal to 6.)

3.3 Implementation of the proposed method

To collect required information, a questionnaire is considered which can cover all of the trust factors. The questionnaire was distributed among users of Facebook (as a supervised network) and users were asked to with considering the five trust levels (defined in section 3-1), assign one of their fiends to each level and answer to the questions. Finally, based on these results, appropriate trust degree assigned to each friend.

3.3.1 Sample size

In order to ensure the accuracy of sampling and generalizability of results to the other users, sample size should be large enough. There are several methods to determine the sample size. The two most popular methods are Morgan table and Cochran formula. In this research Cochran formula is used. Based on this formula [31], if the population size cannot be determined exactly, the sample size is calculated according to formula 4.

$$n = \frac{pqz^2}{d^2} \qquad (4)$$

In formula 4, n determines sampling size, P is the estimation proportion of an attribute that is present in the population, d is the acceptable sampling error, z is the abscissa of the normal curve of that cuts off an area α at the tails and q is the proportion of the attribute that absence in the population.

In this research, due to not clearing the number of joint members of supervised and unsupervised networks whom are suitable for applying the model on them, each

of p and q values are considered equal to 0.5, d value is considered equal to 0.0.1 and z value is considered equal to 1.96. By these assumptions, based on formula 4, the appropriate sample size would be equal to 96.

3.3.2 Recommended weights for each factor

As explained, the questionnaire was distributed among 96 users. Based on results and formulas, which mentioned in section 2-3 and table 3, results such as average, standard deviation of each factor and finally, the range of trust degree for each trust level, are shown in tables 4, 5 and 6.

Table 4. Calculated avarage of each factor

Feature	Average
Relation	0.366
Personal information	0.633
Affiliation	5.333
Inter-activity	10.055
Intra-activity	10.333
Behavioral	3.111

Table 5. Standard deviation and weight of each factor

Feature	Standard deviation	Weight
Relation	0.277	0.9
Personal information	0.147	1
Affiliation	0.620	0.7
Inter-activity	0.429	0.8
Intra-activity	0.403	0.8
Behavioral	0.233	0.9

Table 6. Range of each trust level

Trust level	Min	Max
Untrusted	317	325
Almost untrusted	323	334
No comment	332	362
Trusty	353	391
High trusty	389	401

3.4 Trust assignment in unsupervised networks

As mentioned, in unsupervised networks, because of the ease of impersonation, not existence of supervisors and the direct and great impact of user behavior on others, trust has an important role in these networks. So, we're going to assign a trust level to users of these networks, based on the evaluated trust in past section. The problem that arises in this matter is that the evaluated trust is in the format of a number and in other networks, this number cannot easily express the trust level of users. To solve this problem, fuzzy logic and especially, membership function can be useful. For this, based on the questionnaire results and formulas in section 3-2, final trust degree of each user is calculated and based on the assigned trust level, range of each trust level is defined and can be shown by a membership function. Figure 4 shows the suggested membership function.

In figure 4, vertical axis represents the membership degree of fuzzy variable and horizontal axis represents the calculated trust degree based on formula 3. In this figure, the membership functions are related to "untrusted", "almost untrusted", "no comment", "trusty", and "highly

trusted", respectively. By these functions, the membership degree of trust variable to each trust level is determined as a vector in the form of [untrusted, almost untrusted, no comment, trusty, highly trusted] and trust evaluation would be based on this vector. For example in figure 4, trust vector of user "A" would be as follow:
$E_a = [0.4, 0.4, 0, 0, 0]$

4. Implication

As an example of proposed solution, trust evaluation and authentication of users can be considered. To do so, the following steps can be considered:
1. User P_1 as a member of network N_1, sends a request to communicate with another user of this network, P_2.
2. User P_2 needs to evaluate trust degree of P_1. So, related information about membership of P_1 in supervised networks, are asked to be sent to P2.
3. User P_1 sends his membership information to P_2.(IP$_1$)
4. User P_2 sends IP$_1$ to "Trust Center (TC)" and requests to evaluate his trust level.
5. Sending membership information of P_1 to all related networks.
6. Specification of considered factors in trust evaluation and determining trust level of user P_1 and forwarding these evaluations to the trust center (TC).
7. Aggregating factors and computing total trust based on the formula 3 and then creating trust vector.
8. Sending result to the P_2 by TC.
9. Accept or reject the request of user P_1.

In figure 5, all of the above steps are shown.

Example of implication:
1. Alice wants to communicate with Bob on Skype (as an unsupervised network). So, she sends her request to Bob.
2. Before any communication, Bob needs to verify Alice and evaluate her trust level. So, he asks her to send her membership information in supervised networks such as Facebook, Twitter and etc.
3. She sends her membership information (such as her ID) of supervised networks, such as Facebook and Youtube.
4. Bob sends Alice's membership information to "Trust Center" and requests to evaluate her trust level.
5. Trust Center sends her membership information to the related networks and asks them to evaluate her trust level.
6. Each of these networks (Facebook and Youtube) sends its calculated trust degree. Beside this, they should send factors which considered for trust evaluation.
7. "Trust Center" aggregates these factors and evaluates final trust based on the received results.
8. Final trust vector will be sent to Bob.
9. Bob decides to accept or reject Alice.

5. Conclusions

Nowadays, social networks have created new challenges in security and trust. Besides, unsupervised networks, because of the not existence of supervisors and high level of access in some cases, such as access to run a program on other user's system, trust has a more important role. Many researches have been done on this problem, but static view of trust in most of them, causes failure in hacker identification. In more detail, because of the lack of supervisors in these networks, observing the behavior of nodes in relation with other nodes cannot be possible. While, in supervised networks, the supervisor can record all of the user behavior and assign appropriate trust level, based on them. In this research, with the purpose of providing a dynamic model for trust

evaluation in unsupervised networks, supervisor roles in supervised networks are used. Based on the behavior of users in these networks and their shared information, appropriate trust degree is assigned to them. After drawing membership function of each trust level, the trust vector of user is determined based on the membership degrees and can be used to evaluate the trust level of him.

The limitation of this research is related to the manual submission of membership information in the supervisor networks for user authentication. In future, works, with the automatic version of this part, analyzing the whole behavior of users in all networks which they are membered in, can be possible and user cannot hide his bad behavior in some networks.
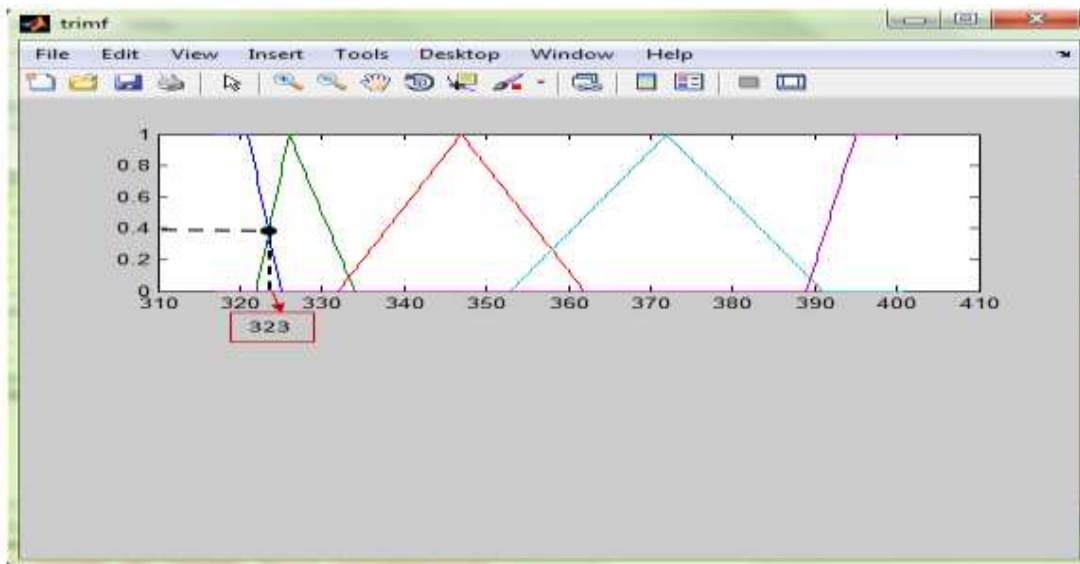

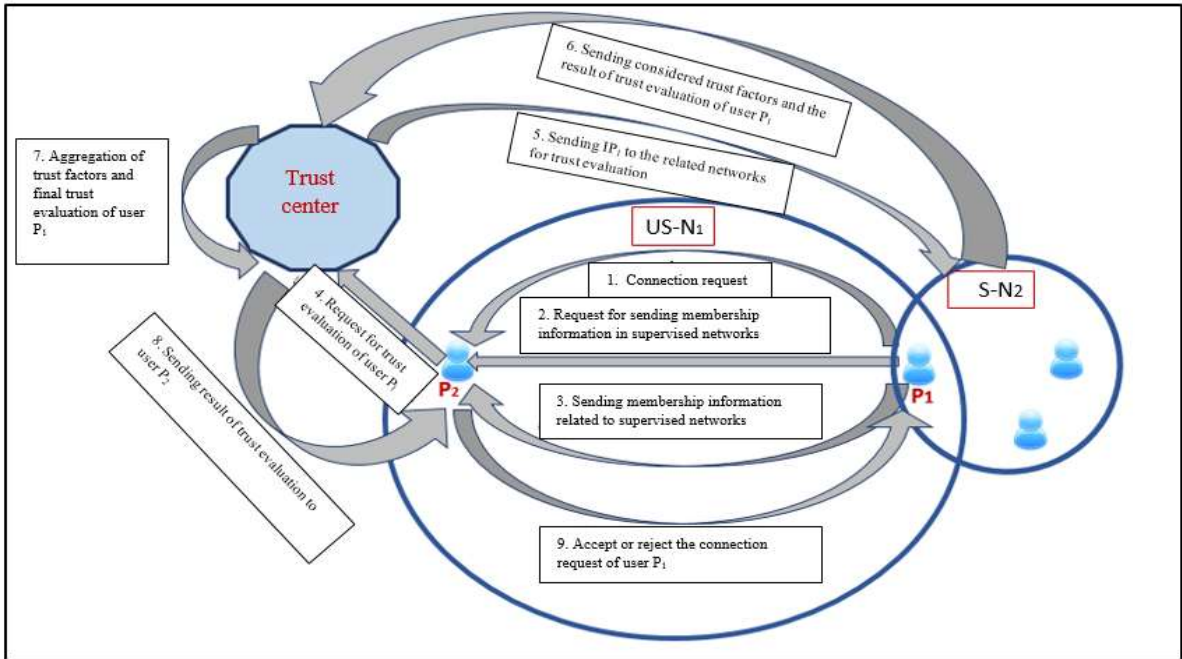Fig 4. The membership function of trust level variable.

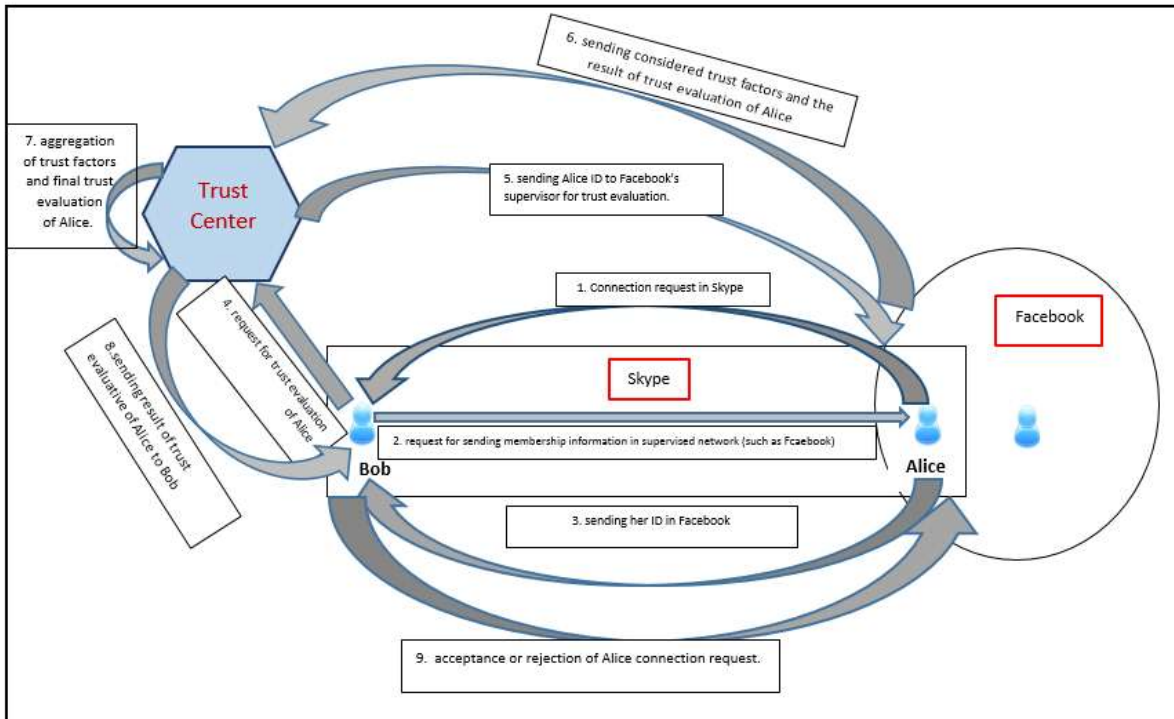Fig 5. The framework for application of the proposed model

Fig 6. A real-world example of the proposed model

References

[1] Adali, S., Escriva, R., Goldberg, M.K. , Hayvanovych, M., Magdon-Ismail, M., Szymanski, B.K., Wallace, W.A., Williams, G., "Measuring Behavioral Trust in Social Networks", Third American Conference on Information System, IEEE, pp 150-152, May 2010.

[2] Wondracek, G., Holz, T., Kirda, E., Kruegel, C., "A Practical Attack to De-Anonymize Social Network Users", IEEE Symposium on Security and Privacy (SP), pp 223-238, May 2010.

[3] Zhang, C., Sun, J., Zhu, X., Fang, Y., "Privacy and Security for Online Social Networks: Challenges and Opportunities", IEEE Network, Vol.24, No.4, pp13-18, 2010.

[4] Pisey, S.H., Ramteke, P.L., Deshmukh, P., Burghate, B.R., "Privacy Access Control Mechanism for Online Social Network", International Journal of Computer Science and Applications, Vol.6, No.2, pp 172-179, April 2013.

[5] Nagy, J., Pecho, P., "Social Networks Security", Third International Conference on Emerging Security Information, System and Technologies, pp321-325, June 2009.

[6] Livingstone, S., Brake, D., "On the Rapid Rise of Social Networking Sites: New Findings and Policy Implications", Children and Society, pp 75-83, Vol. 24, No.1, Jan 2010.

[7] Fong, S., Zhuang,Y., Yu, M., Ma, I., "Quantitative Analysis of Trust Factors on Social Network using Data Mining Approach", International Conference on Future Generation Communication Technology(FGCT), IEEE, pp 70-75, Dec 2012.

[8] Liu, G., Wang, Y., Orgun, M.A., "Social Context-Aware Trust Network Discovery in Complex Contextual Social Networks", Association for the Advancement of Artificial Intelligence, 2012.

[9] Huang, B., Kimmig, A., Getoor, L., Golbeck, J., "A Flexible Framework for Probabilistic Models of Social Trust", 6th International Conference on Social Computing, Behavioral-Cultural Model and Prediction, pp 265-273, 2013.

[10] Borbora, Z.H., Ahmad, M.A., Oh, J., Haigh, K.Z., Srivastava, J., Wen, Z., "Robust Feature of Trust in Social Network", Social Network Analysis and Mining. Vol.3, No.4, pp 981-999, 2013.

[11] Dwyer, C., Hiltz, S.R., Passerini, K., "Trust and Privacy Concern within Social Networking Sites: A Comparison of Facebook and MySpace", International Conference on Intelligence and security informatics, August 2007.

[12] Mahapatra, A., Tarasia, N., "A Fuzzy Approach for Reputation Management using Voting Scheme in Bittorrent P2P Network", International Journal of computer science and information technologies, Vol .2 , No.2, pp735-740, 2011.

[13] Chen, H., Ye, Zh., Liu, W., Wang, Ch., "Fuzzy Inference Trust in P2P Network Environment", International Workshop on Intelligent System and Applications, pp 1-4, 2009.

[14] Biao, C., Zhishu, L., "Computing and Routing for Trust in Structured P2P Network", Journal of Networks, Vol.4, No.7, pp667-674, September 2009.

[15] Sandhya, S., "Trust Management in P2P Networks using Mamdani Fuzzy Inference Systems", International Journal of Computer Application (IJCA), Vol. 66, No.14, March 2013.

[16] Sherchan., W., Nepal, S., Paris, C., "A Survey of Trust in Social Networks", ACM Computing Surveys (CSUR), Vol.45, No.4, August 2013.

[17] Golbeck, J., Hendler, J., "Inferring Binary Trust Relationships in Web-Based Social Networks", ACM Transaction on Internet Technology (TOIT), Vol. 6, No. 4, November 2006.

[18] Bharadwaj, K.K., Al-Shamri, M.Y.H., "Fuzzy Computational Models for Trust and Reputation Systems", Electronic Commerce Research and Application, Vol.8, No.1, pp 37-47, 2009.

[19] Xiong, L., Liu, L., "PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities", IEEE Transaction on Knowledge and Data Engineering, Vol.16, No.7, pp843-857, July 2004.

[20] Walter, F.E., Battiston, S., Schweitzer, F., "Personalised and Dynamic Trust in Social Networks", Third ACM Conference on Recommender Systems, pp 197-204, 2009.

[21] Xin, L., Leyi, S., Yao, W., Zhaojun, X., Wenjing, F., "A Dynamic Trust Conference Algorithm for Social Network", Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), pp 340-346, Oct 2013.

[22] Taherian, M., Amini, M., Jalili, R., "Trust Inference in Web-Based Social Networks using Resistive Networks", Third International Conference on Internet and Web Applications and Services (ICIW), pp233-238, June 2008.

[23] Wang, H.Y., Zhao, Y.F., Meng, X.D., Ji, X.J., "Achieving Quick Trust Transmission in P2P Network Based on Matrix Operation", Manufacturing Engineering and Process II, June 2013.

[24] Huang, Z., Lu, S., Zhang, A., Gu, J., "Impact of Feedback on Trust in P2P Network", Journal of Networks, Vol. 7, No.8, pp 1182-1188, Aug 2012.

[25] Zhen-wei, Y., Qing-guo, S., Jun, L., "A Trust Evaluation Approach for P2P Nodes Based on Trust Computing ", 12th IEEE International Conference on Communication Technology (ICT), pp 1088-1091, Nov 2010.

[26] Zadeh, L.A., "Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic", Fuzzy Sets and Systems journal, Vol.90, pp 111-127, 1997.

[27] Han, Q., Wen, H., Ren, M., Wu, B., Li, S., "A Topological Potential Weighted Community-Based Recommendation Trust Model for P2P Networks", Peer-to-Peer Networking and Applications, June 2014.

[28] Yu.Z., Zhu.J., Shen.G., Liu,H., "Trust Management in Peer- to-Peer Networks", Journal of Software, Vol.9, No.5, pp 1062-1070, May 2014.

[29] Li, X., Zhou, F., Yang, X., "A Multi-Dimensional Trust Evaluation for Large-Scale P2P Computing", Journal of Parallel and Distributed Computing, Vol. 71, No. 6, pp837-847, June 2011.

[30] Gilbert, E., Karahalios, K., "Predicting Tie Strength with Social Media", SIGCHI Conference on Human Factors in Computing System, pp 211-220, 2009.

[31] Cochran, W.G., "Sampling Theory When the Sampling Units Are of Unequal Sizes", Journal of American Statistical Association, Vol.37, No.218, Jun 1942.

Monireh Hosseini holds a PhD from Tarbiat Modares University (TMU). She is currently an assistant professor at the Information Technology Department of Industrial Engineering Faculty at K. N. Toosi University of Technology. Her work deals with customer analytics and network models of customer value. She has teaching experience in Internet marketing, ecommerce strategies and Management Information Systems. She has published a number of research papers in international scientific journals and conference proceedings. She is the highly commended winner of the 2011 Emerald/EFMD Outstanding Doctoral Research Awards and received two scientific awards for the best paper in March 2008 and January 2011.

Golnar Assadat Afzali is a master student of Information Technology Engineering at K. N. Toosi University of Technology. She received her B.S. degree in information technology engineering from Isfahan University of Technology (IUT). Her work deals with network models of trust and information security.

Security Analysis of Scalar Costa Scheme Against Known Message Attack in DCT-Domain Image Watermarking

Reza Samadi*

Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

r.samadi@stu.um.ac.ir

Seyed Alireza Seyedin

Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

seyedin@um.ac.ir

Abstract

This paper proposes an accurate information-theoretic security analysis of Scalar Costa Scheme (SCS) when the SCS is employed in the embedding layer of digital image watermarking. For this purpose, Discrete Cosine Transform (DCT) coefficients are extracted from the cover images. Then, the SCS is used to embed watermarking messages into mid-frequency DCT coefficients. To prevent unauthorized embedding and/or decoding, the SCS codebook is randomized using the pseudorandom dither signal which plays the role of the secret key. A passive attacker applies Known Message Attack (KMA) on the watermarked messages to practically estimate the secret key. The security level is measured using residual entropy (equivocation) of the secret key provided that the attacker's observations are available. It can be seen that the practical security level of the SCS depends on the host statistics which has contradiction with previous theoretical result. Furthermore, the practical security analysis of the SCS leads to the different values of the residual entropy in comparison with previous theoretical equation. It will be shown that these differences are mainly due to existence of uniform regions in images that cannot be captured by previous theoretical analysis. Another source of such differences is ignoring the dependencies between the observations of non-uniform regions in previous theoretical analysis. To provide an accurate reformulation, the theoretical equation for the uniform regions and the empirical equation for the non-uniform regions are proposed. Then, by combining these equations a new equation is presented for the whole image which considers both host statistics and observations dependencies. Finally, accuracy of the proposed formulations is examined through exhaustive simulations.

Keywords: Scalar Costa Scheme; Known Message Attack; Discrete Cosine Transform; Residual Entropy; Watermarking.

1. Introduction

Digital watermarking refers to hiding a verification message into cover or host data in a secure and robust manner. In digital watermarking, the message is encoded to the watermark using a key, and then the watermark is embedded or added into the host. In the receiver side, the message should be estimated without having host signal which is called blind decoding; hence, the host plays the role of the interference in the receiver. Perfect host rejection (interference cancellation) is possible only in Ideal Costa Scheme (ICS) [1] which is optimal but non-practical. Scalar Costa Scheme (SCS) [2] is the down-to-earth implementation of the ICS which approximates the random codebook of the ICS by a set of scalar quantizers. The SCS can be used in the embedding layer in the spatial or transform domain of the image, audio, video or any other documents. The essential assumption on the quantization-based data hiding analysis is *flat-host assumption*. It assumes that the host probability density function (pdf) is uniform in each quantization cell [3]. This is equivalent to infinite Document to Watermark Ratio (DWR).

Watermarking system designers always need to have accurate closed form equation of security and robustness to reliably design such systems. An important characteristic is security level which has been poorly investigated in the literature. In the watermarking security, the purpose of the attacker is to disclose the secret parameters to implement tampering attack. Watermarking security deals with intentional attacks whose aim are accessing the watermarking channel [4], excluding those already encompassed in the robustness category. Such access refers to trying to remove, detect and estimate, write and modify the raw watermarking messages. According to the type of the information the attacker has access to, there are three scenarios, including Known Message Attack (KMA), Watermark Only Attack (WOA) and Known Original Attack (KOA) [5]. Also, there are two information-theoretic [6] and probabilistic [7] frameworks to evaluate the security level of the watermarking techniques. Information-theoretic security analysis of the SCS based on the flat-host assumption has been carried out theoretically in [8], [9]. In these studies, the authors proved that the security level of the SCS is independent of the host statistics when flat-host assumption is valid or the DWR is infinite. These

analyzes are rather general and should be applicable in any spatial/transform domain for any kind of the cover data of image/audio/video/text. However these theoretical results may not be valid in practice due to the implementation considerations. For example these analyses cannot capture the effect of the host statistics. More precisely, work of [10] shows that breaking the SCS is practically feasible with the aid of key guessing and depends on the host statistics. Also, more recent work of [11] shows that the practical security level of the SCS in image watermarking application is much lower than what has been previously derived in [8] and strictly depends on the host statistics.

In this paper, we propose an accurate security analysis of the SCS for the digital image watermarking in Discrete Cosine Transform (DCT) domain. We apply security attack on the SCS in the KMA scenario, and evaluate the information-theoretic security level practically. It is shown in this paper that the image uniformity invalidates the flat-host assumption. Hence, the previous theoretical formulation in [8], derived based on the flat-host assumption, would be no longer valid for the uniform regions. It should be noted that each image is composed of uniform and non-uniform regions. Non-uniform regions do not invalidate the flat-host assumption; hence, it is expected that the results of [8] are valid for the non-uniform regions at least. However, we show there are dependencies between the observations that can be used by the attacker, which have not been considered by the theoretical formulation in [8]. Hence, the previous theoretical formulation in [8] would also not be valid for the non-uniform regions. To provide an accurate reformulation, we propose a theoretical equation for the uniform regions and an empirical equation for the non-uniform regions which considers host statistics and dependencies between the observations. Then, we combine our result to propose a new unified closed-form equation for the security level of the SCS in the image watermarking application. These accurate analyses are crucial when the SCS is used in the complex scenarios or real applications which need high security [12].

The remaining of this paper is organized as follows. Section 2 presents the block diagram, notations and primary definitions of the watermarking in the DCT domain using the SCS encoding. In Section 3, the practical key estimator of the SCS is reviewed and applied to the KMA scenario [8] in the DCT domain and then, the security level stressing is evaluated and compared to the former theoretical result. The theoretical-empirical equation for the security level of the SCS in KMA scenario in the DCT domain image watermarking is presented in Section 4. Finally, Section 5 concludes the paper.

2. Watermarking Using SCS in DCT domain

Block diagram of the watermarking in the DCT domain [13] is shown in Fig. 1. First, the image pixels are subdivided into 8×8 blocks and the DCT is applied on each block. Then, each 8×8 block of the DCT coefficients is ordered in a zigzag mode and the appropriate (mid-frequency) coefficients are selected for embedding. The Illustration of the zigzag ordering and selected mid-frequency DCT coefficients is presented in Fig. 2. In the embedding layer, messages M_j: $j=1,...,M$ are embedded in the selected DCT coefficients (host) X_i: $i=1,...,N$ by using the secret key K. Embedding can be done redundantly in such a way that the whole image is covered. In simple repetition coding and embedding, each message M_j is embedded in $L=\lfloor N/M \rfloor$ different hosts, X_i, resulting the watermarked host Y_i as:

$$Y_i = embedd(X_i, M_j, K), \quad (j-1)L < i \leq jL \qquad (1)$$

After the embedding, both non-selected and selected coefficients are merged and reordered to make 8×8 blocks in the transformed domain. Then, the inverse DCT is applied on each block to create 8×8 blocks in the spatial domain. Finally, these blocks are merged, yielding the watermarked image. This image is open access and ready to travel over the network. We exclude the robustness attacks and any security mechanism other than the secret key. Under the Kerckhoffs' assumption [14], all the parameters and details of the data hiding scheme are known to the attacker, except the secret key. Hence, the attacker conducts the feature extraction on the watermarked image to access the watermarked host.

2.1 SCS for Information Embedding

In this paper, the theoretical model of the information embedding layer between the sender, proposed in [8], is employed. This model is shown in Fig. 3, where the host signal X is independent and identically distributed (i.i.d) scalar feature extracted from the original digital content. The embedder hides an equiprobable message $M \in \{0,\ldots,p-1\}$ in the host by using secret key K yielding watermark W. Then, the watermark is added or embedded into the host, producing the watermarked host Y. The Attacker has access to the watermarked host through the attack channel which produces attacked host Z. Detector receives the attacked host and tries to estimates the embedded message. However, only the legal detector has the secret shared key and is able to estimate the embedded message.
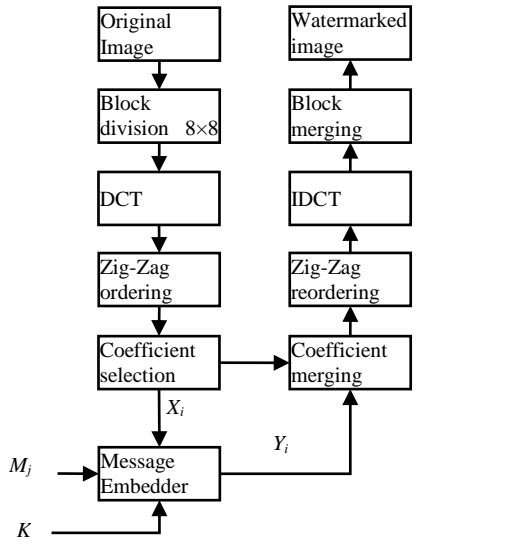
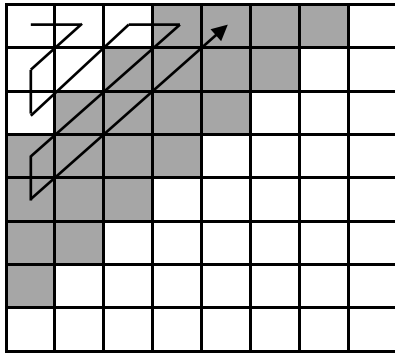Fig. 1. Block diagram of watermarking in DCT domain

Fig. 2. Illustration of Zig-Zag ordering and selected mid-frequency DCT coefficients of Fig. 1.
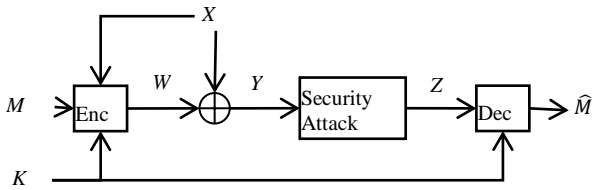

Fig. 3. Theoretical model for additive side-informed watermarking, including security attack and excluding robustness attack/noisy channel

In the SCS, watermarked host in (1) is read as (2), where $e_{M,K}(X)$ is the quantization noise of host over the shifted lattice [15], Q_Δ indicates the uniform scalar quantizer over Voronoi region $[-\Delta/2, \Delta/2]$, Δ is quantization step size and M represents the message symbol to be transmitted. The quantization noise of the host over the lattice Λ is defined as $X mod \Lambda = X - Q_\Delta(X)$.

$$Y = X - \alpha e_{M,K}(X) = (1 - \alpha)e_{M,K}(X) + Q_{M,K}(X) \qquad (2)$$

$$e_{M,K}(X) = X - \Delta\frac{M}{p} - K - Q_\Delta\left(X - \Delta\frac{M}{p} - K\right) \qquad (3)$$

The security of the embedder relies only on the randomization of the codebook via the secret key. The fundamental assumption when analyzing the SCS is that the secret key is distributed uniformly over the Voronoi region. Hence, the error signal $e_{M,K}(X)$ is orthogonal to X,M and the host rejection is possible [2]. In the remainder of the paper, the uniform secret key assumption is employed. The embedding distortion (watermark power) is evaluated as $D_w = \alpha^2\Delta^2/12$. Moreover, the transparency is measured by the host variance to the watermark power ratio $\lambda = \sigma_X^2/D_w$ and DWR is defined as $DWR = 10\log_{10}(\lambda)$.

3. Practical Security Level Evaluation of SCS in KMA Scenario

This paper concentrates on the KMA scenario only, where the attacker has access to the pool of observations with independent messages and corresponding watermarked hosts, all watermarked with the *same* secret key. Attacker's purpose is to disclose the secret key with the aid of the available observations and then perform the tampering attack. The security level is evaluated as the effort of the attacker to reduce his ambiguity about the secret key. In the information-theoretic framework, γ-security level is defined as the number of the N_o observations attacker needs to make $h(K|Y^{N_o}, M^{N_o})\gamma$ [8]. To evaluate the information-theoretic security level of the SCS in the KMA scenario, the residual entropy $h(K|Y^{N_o}, M^{N_o})$ is computed for every N_o. Based on the definitions presented above, the residual entropy has been analyzed theoretically in [8] for $\alpha \geq 0.5$ using flat-host assumption, as follow:

$$h_o(K|Y^{N_o}, M^{N_o}) = log_2(1 - \alpha)\Delta + 1 - \sum_{i=1}^{N_o}\frac{1}{i} \qquad (4)$$

The subscript 'o' indicates (4) is the previous theoretical equation in comparison with the proposed equation in Section 4. In [8], the authors assumed that residual entropy is independent of the host statistics for high DWR. Detailed analysis of the security attack in [8] is reviewed in Section 3.1. In Section 3.2, we practically evaluate the residual entropy of the SCS for information embedding in the DCT domain for some test images with various host statistics. To do this, first the KMA is applied on the SCS, and then the security level is evaluated. The test images and distribution of their selected DCT coefficients (host) are shown in Fig. 4. To compare the shape of hosts' pdf only, variances of the hosts' pdf are normalized to $\sigma_x^2 = 1$.

3.1 KMA on SCS

The best practical key estimator of the SCS in the KMA scenario was proposed in [8], for $\alpha \geq 0.5$ based on the k-means clustering. Feasible value of the secret key in observation index r is derived in (5) where the set $Z(\Lambda) = \{x: |x| \leq (1 - \alpha)\Delta/2\}$ is the scaled version of the Voronoi region. Feasible values of the secret key after N_o observations are the intersection of the sets $\mathcal{D}_r$ as follow.

$$k \in \mathcal{D}_r, \quad \mathcal{D}_r = \left(y_r - \Delta \frac{m_r}{p} - \mathcal{Z}(\Lambda) \right) mod\Lambda \qquad (5)$$

$$\mathcal{S}_{N_o} \triangleq \bigcap_{r=1}^{N_o} \mathcal{D}_r \qquad (6)$$

When the flat-host assumption is valid, the secret key is uniformly distributed on feasible region $\mathcal{S}_{N_o}$ after N_o observations. Hence, the residual entropy of the secret key, when the watermarked host and message are provided, can be written as (7).

$$h(K|Y^{N_o} = y^{N_o}, M^{N_o} = m^{N_o}) = E[log(vol(\mathcal{S}_{N_o}))] \qquad (7)$$

3.2 Implementation of KMA on SCS and Practical Security Level Evaluation

In this paper, any robustness attack is excluded. Hence, the attacker has access to the noiseless watermarked hosts and corresponding messages. Since there are N different hosts, hence we can *independently* implement the KMA on the SCS $Q = \lfloor N/N_o \rfloor$ times in different trials. Then, the accurate approximation of residual entropy is computed by averaging over the trials. We define $\mathcal{S}_{N_o}^q$ as the feasible set of the secret keys in the q-th trial of the practical entropy computation based on the N_o observations. The definitions of the feasible sets $\mathcal{D}_r$ and $\mathcal{S}_{N_o}^q$ in the model used in this paper are given in (8) and (9), respectively. Finally, the practical residual entropy can be approximated by (10).

distribution of selected 1254792 DCT coefficients — test15.bmp

distribution of selected 883938 DCT coefficients — test16.bmp

distribution of selected 880000 DCT coefficients — test27.gif

distribution of selected 2162688 DCT coefficients — test4.bmp

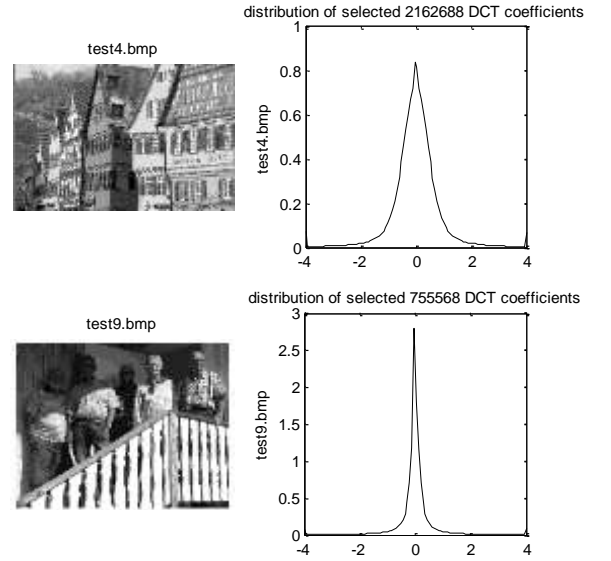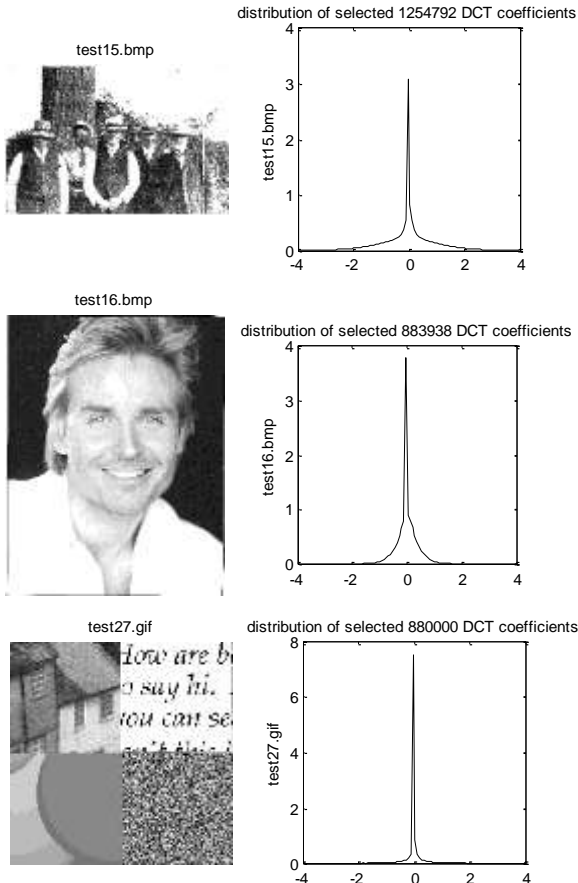distribution of selected 755568 DCT coefficients — test9.bmp

Fig. 4. Used test images and their selected DCT coefficients distribution. In each image, the variance of the coefficients is normalized to 1. Notice the differences in the shape of the distributions.

$$\mathcal{D}_r = \left(y_r - \Delta \frac{m_s}{p} - \mathcal{Z}(\Lambda) \right) mod\Lambda, \quad s = \lceil r/L \rceil \qquad (8)$$

$$\mathcal{S}_{N_o}^q \triangleq \bigcap_{r \in ((q-1)N_o, qN_o]} \mathcal{D}_r \qquad (9)$$

$$h_p(K|Y^{N_o}, M^{N_o}) \cong \frac{1}{Q} \sum_{q=1}^{Q} log(vol(\mathcal{S}_{N_o}^q)) \qquad (10)$$

The subscript 'p' in (10) denotes the practical residual entropy equation. Approximation (10) is resulted by application of the *weak law of large numbers* [16]. As Q increases, the variance of the approximation error tends to zero if the observation sequences, with the length N_0, are independent from each other for each q, which is a valid assumption in our problem. The proof directly follows from the assumption that the watermarked hosts y_r are function of the i.i.d random variables x_r, m_r, which are mutually independent. Moreover, for each q the secret key k is constant.

The residual entropy in (10) is evaluated for the test images of Fig. 4. These results are shown in Fig. 5, compared to the theoretical results for the SCS [8] and Spread Spectrum (SS) schemes [17]. For the sake of the fair comparison, all the test images have embedded with the same embedding parameters, σ_x, α, DWR. It is worth noting that all test images only differ in the shape of the host pdf. It is obvious that the change the host pdf has large impact on the residual entropy, indicating the strong dependency of the security level of the SCS in the KMA scenario on the host statistics.

To clearly show the differences between the empirical and theoretical results, the security level is numerically computed in

Table. 1. We use those values of γ which lead to the security level of $N_o \approx 100$ in the theoretical result for both $DWR = 30, 40dB$. It is easy to note the large gap between the empirical and theoretical evaluations of security level.

For example, previous result suggests $N_o \approx 100$ observations are required for *DWR=40dB* to reduce the residual entropy lower than $\gamma = -8.6$. However, it can be seen that the practical KMA on the "test4.bmp" only needs $N_o=27$ observations to achieve this goal. Therefore, the theoretical result overestimates the security level of the SCS against the KMA and ignores the effect of the shape of the hosts' pdf on it, too.

3.3 on the need of proposing an accurate analysis

The results of this section clearly demonstrate the large gap between the practical evaluation and theoretical result in [8] for the security analysis of the SCS in the DCT domain and the KMA scenario. Theoretical result in [8] overestimates the security level and also ignores the impact of host statistics on it. Therefore, it is necessary to propose an accurate equation for the security level of the SCS.
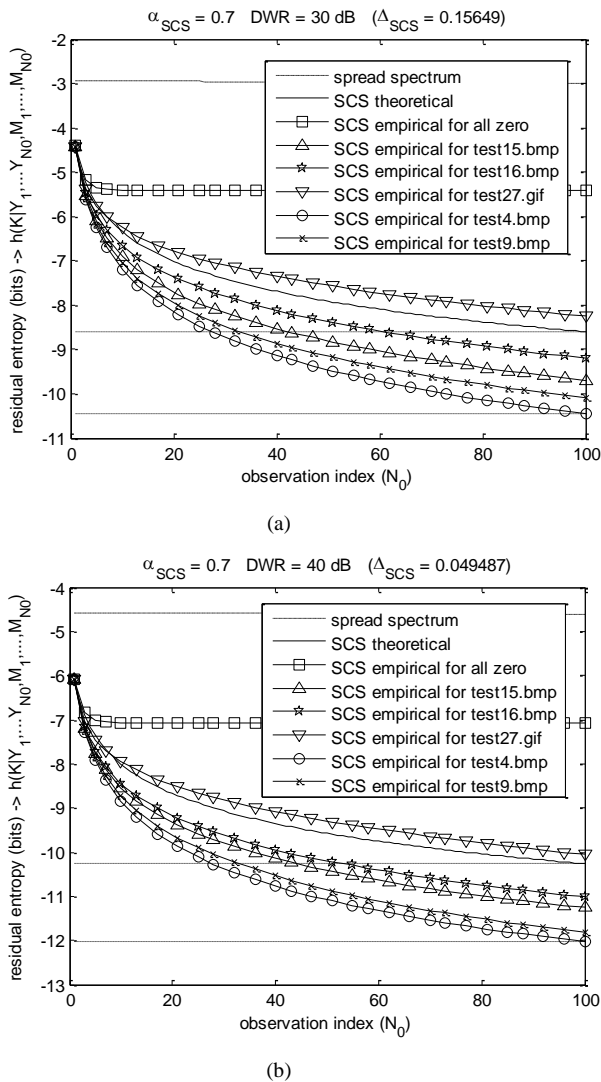

(a)

(b)

Fig. 5. Residual entropy of SCS embedding in DCT domain in KMA scenario for five test images and specified parameters; comparisons are done with theoretical results of SCS and Spread Spectrum. The case of "all zero" is corresponds to uniform (homochromatic) image.

Table. 1. Security level of test images for two values of DWR, γ is chosen in such a way that theoretical result gives security level as 100

Image	Security level No	
	DWR=30dB $\gamma = -8.6$	DWR=40dB $\gamma = -10.26$
theoretical	100	100
test4.bmp	27	27
test9.bmp	32	32
test15.bmp	41	44
test16.bmp	60	52
test27.gif	> 100	> 100
all zero	$+\infty$	$+\infty$

4. Security Analysis of SCS in KMA Scenario

In previous section it was shown that the security levels of the SCS used in DCT domain for different images in the KMA scenario are not identical even for the same embedding parameters. The main goal of this section is to provide an accurate closed-form equation for this security level. To do this, each image is divided into two textured and uniform regions. We prove that the security level for the uniform regions is *constant* for different images and is not a function of the observation index. Also, it will be demonstrated that the security levels of the textured regions are *identical* for different images and are a function of the observation index. We propose a theoretical equation for the security level of the uniform regions and an empirical equation for the security level of the textured regions. Then, these results are combined to derive a unified theoretical-empirical equation for every image. Finally, the accuracy of the proposed equations is verified using a huge set of different test images.

4.1 Security analysis for uniform regions

Uniform regions of each image are adjacent pixels with low brightness variations. The DCT transform of these regions are composed of almost zeros coefficients in low frequencies in the transform domain. For example, the test image *"test27.gif"* in Fig. 4 has much uniform regions. It is clear that for the uniform region, the flat host assumption is not valid anymore. Hence the equation (4) would be no longer valid, too. We prove that the security level of the SCS for uniform regions is stated as follows:

$$h_u(K|Y^{N_o}, M^{N_o}) = log_2(1-\alpha)\Delta - 1 \qquad (11)$$

where, subscript 'u' the proposed theoretical equation in (11) applied to the uniform regions. The proof of (11) is straightforward: Attacker's objective is to estimate the secret key k after N_o observations by evaluating the set $\mathcal{S}_{N_o}$ in (6). Given $\{t_i, m_i\}$, the set $\mathcal{D}_i$ in (5) can be rewritten as follow:

$$\mathcal{D}_i = \left((y_i - \Delta \frac{m_i}{p})mod\Lambda - \mathcal{Z}(\Lambda) \right)mod\Lambda, \tag{12}$$

$$i = 1, \dots, N_o$$

The first step by the attacker is modulo lattice reduction of the observations as $\widetilde{Y}_i = (Y_i - \Delta M_i/p)mod\Lambda$. Under the flat-host assumption, the attacker does not lose any information about the secret key after this modulo reduction. By substituting y_i from (2), the modulo lattice reduced observations is given by:

$$\tilde{Y}_i = ((1 - \alpha)e_{M_i,K}(X_i) + K)mod\Lambda \tag{13}$$

It is clear that the secret key is only concealed by the embedding message and quantization error $e_{m_i,K}(x_i)$. Here, the quantization error plays the role of the host interference. For the uniform regions, the selected mid-frequency coefficients are almost zero $x_i \cong 0$. Hence, after some simplifications, the quantization error for the binary embedding *(p=2)* can be evaluated as follow:

$$e_{m_i,K}(x_i) = \begin{cases} -k, & m_i = 0 \\ \dfrac{\Delta}{2} - k, & m_i = 1 \end{cases} \tag{14}$$

Substituting (14) into (13) and then resulting equation into (12), will yield the following:

$$\mathcal{D}_i = \begin{cases} (k\alpha - \mathcal{Z}(\Lambda))mod\Lambda & m_i = 0 \\ (k\alpha + \dfrac{(1-\alpha)\Delta}{2} - \mathcal{Z}(\Lambda))mod\Lambda & m_i = 1 \end{cases} \tag{15}$$

For $m_i = 0$, the set $\mathcal{D}_i$ is centered on the resized secret key with maximum ambiguity $(1 - \alpha)\Delta$. On the other hand, for $m_i = 1$, the set $\mathcal{D}_i$ is centered on shifted and resized secret key with maximum ambiguity $(1 - \alpha)\Delta$. Intersection of these sets gives the estimated key. However, the sets corresponding to the identical embedding message are identical. Hence, the intersection of just two sets with different embedding messages produces the same result as the intersection of all available sets. Two sets with the different embedding messages are illustrated in Fig. 6. It is easy to see that the intersection of these sets gives the intended result in (11).

4.2 Security analysis for textured regions

Non-uniform regions are called textured regions. These regions are produced by removing the uniform regions from an image. This can be realized by removing nearly zero DCT coefficients in the transform domain. It will be shown that the security level for the textured regions is identical for different images.

After removing nearly zero DCT coefficients, the security levels of the images in Fig. 4 are sketched in Fig. 7. It is interesting to note that the security levels for different textured images are identical for the same embedding parameters. Moreover, it is expected that the security level for textured region should be consistent with the previous theoretical equation of (4). However, as it can be seen in Fig. 7, the practical security level is much lower than that of the previous theoretical equation.

This happens because (4) ignores the dependencies between the observations watermarked with a same secret key. Precisely, authors of [8] assumed that the sets $\mathcal{D}_i$ are independent during the proof of their theorem in [8-Appendix A]. However, from (12) and (13) we can rewrite $\mathcal{D}_i$ as:

$$\mathcal{D}_i = (K + (1 - \alpha)e_{M_i,K}(X_i) - \mathcal{Z}(\Lambda))mod\Lambda \tag{16}$$
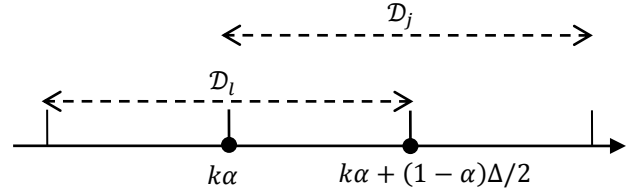
Fig. 6. Illustration of two sets of estimated secret keys for uniform regions with different embedding messages
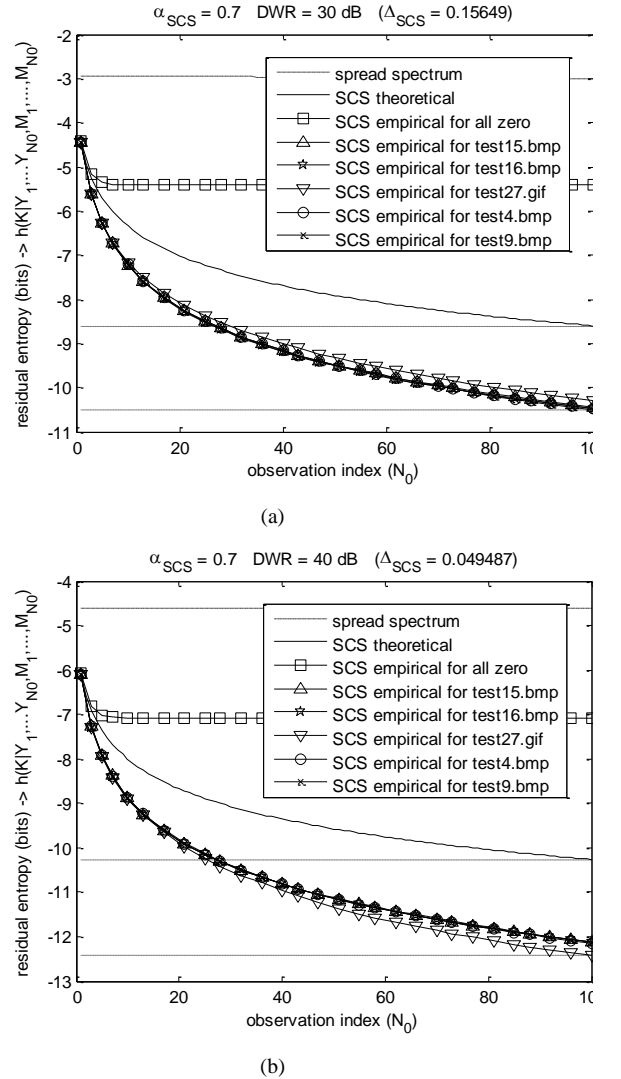

(a)

(b)

Fig. 7. Residual entropy of SCS embedding in DCT domain in KMA scenario for five test images and specified parameters; uniform regions of test images are removed; comparisons are done with theoretical results of SCS and Spread Spectrum. The case of "all zero" is corresponds to uniform (homochromatic) image.

It is easy to see that the sets $\mathcal{D}_i$ are a priori interdependent via the secret key *K*. Indeed, the information about the secret key from these dependencies can be extracted by practical evaluation. The security level is a function of DWR, distortion compensation α, and observation index N_o. Based on the practical evaluations, we intuitively assume that security level is a logarithmic function of the observation index. Hence, the following empirical formulation is proposed for the security level of textured images.

$$h_t(K|Y^{N_o}, M^{N_o}) \cong log_2(1-\alpha)\Delta - 0.96 log_2 N_o + 0.35 \quad (17)$$

To show the accuracy of the proposed empirical formulation, a large set of the well-known image databases with different sizes, resolutions, qualities and other basic characteristics are created as follow:

- USC-SIPI image database [18], composed of textures and miscellaneous images.
- TESTIMAGES archive [19].
- DIP3/e—Book Images [20], composed of images from Digital Image Processing, 3rd ed, by Gonzalez and Woods.
- CIPR Still Images [21], composed of miscellaneous images.

Security level is evaluated for these test images, embedding parameters (α, DWR) and observation index (N_o). To show the superiority of the proposed formulation, we evaluate percentage error between the practical evaluation of security level and the previous/ proposed equations of (4)/ (17). The percentage error between practical evaluation and the previous theoretical equation $EP_{previous,practical}$ is defined in (18). Moreover, the percentage error between the practical evaluation and the proposed empirical equation $EP_{proposed,practical}$ is defined in (19). The indexes 'o', 't' and 'p' in (18) and (19), respectively, indicate the previous theoretical equation, proposed empirical equation and practical evaluation, respectively.

$$EP_{previous,practical}(\alpha.DWR, N_o)$$
$$= \frac{|h_o(K|Y^{N_o}, M^{N_o}) - h_p(K|Y^{N_o}, M^{N_o})|}{h_p(K|Y^{N_o}, M^{N_o})} \quad (18)$$

$$EP_{proposed,practical}(\alpha.DWR, N_o)$$
$$= \frac{|h_t(K|Y^{N_o}, M^{N_o}) - h_p(K|Y^{N_o}, M^{N_o})|}{h_p(K|Y^{N_o}, M^{N_o})} \quad (19)$$

Fig. 8 (a) demonstrates percentage error for $\alpha = 0.71$ as a function of DWR. Moreover, Fig. 8(b) shows the percentage error for DWR = 28dB as a function of α. In both figures, $EP_{previous,practical}$ and $EP_{proposed,practical}$ are marked by square and circular circles, respectively. For the illustration purpose, these figures are averaged over the observation index (N_o). The simulations clearly show that the percentage error of the proposed formulation is below one percent for whole values of the embedding parameters (α. DWR). Hence, the proposed empirical equation of (17) is consistent with the practical evaluation of the entropy, while the existing theoretical analysis of (4)

in the literature leads to the results far from the reality. Hence, it can be concluded that the proposed formulation is remarkably accurate for the textured images or textured regions of any image.

(a)

(b)
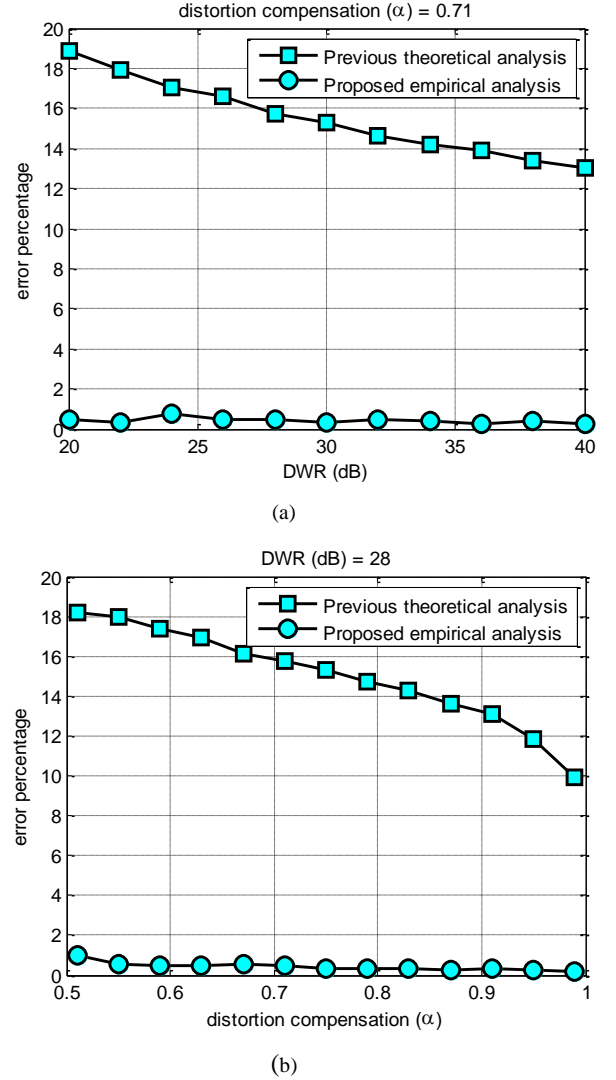
Fig. 8. Percentage error of evaluating residual entropy of SCS embedding in DCT domain in KMA scenario for test images and specified parameters; squares correspond to former theoretical equation; comparisons are done with previous theoretical results of SCS.

4.3 Security analysis for every image

Finally, the results of the uniform and textured regions are combined to propose a unified formulation for the security level of SCS in the DCT domain image watermarking. Comparison of Fig. 5 and Fig. 7 clearly shows that the security level in the DCT domain image watermarking is closely related to the uniformity of the image. Hence, we define a new variable ρ which linearly measures the uniformity of an image. The case $\rho = 1$ corresponds to the complete uniform images while $\rho = 0$ indicates almost textured images. The block diagram of computing uniformity variable ρ is illustrated in Fig. 9. It is reasonable that the unified formulation should be a

combination of (11) and (17). Hence, the new theoretical-empirical formulation is states as follows:

$$
\begin{aligned}
h(K|Y^{N_o}, M^{N_o}) &\cong \rho h_u(K|Y^{N_o}, M^{N_o}) \\
&+ (1-\rho)h_t(K|Y^{N_o}, M^{N_o}) \\
&= log_2(1-\alpha)\,\Delta - 1 \\
&- (1-\rho)(0.96 log_2 N_o - 1.35)
\end{aligned}
\tag{20}
$$

set $u=0$, $t=0$

Coefficient Selection

x_i

$x_i \leq 0.01\Delta$

Yes   No

Uniform regions
$u=u+1$

Textured regions
$t=t+1$
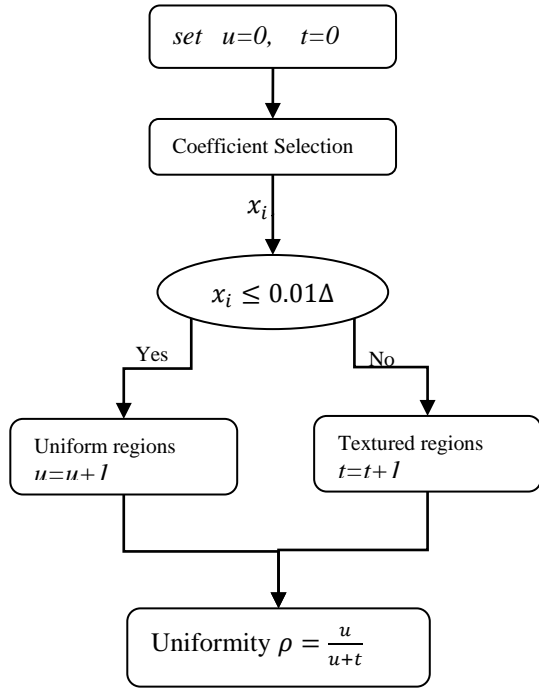
Uniformity $\rho = \dfrac{u}{u+t}$

Fig. 9. Algorithm for evaluating the image uniformity.

The accuracy of the proposed analysis is verified by computing the percentage error of the proposed unified formulation and the practical security evaluation of the test images. The percentage error for several values of the embedding parameters is illustrated in Fig. 10 as a function of the uniformity ρ. Moreover, the average percentage error, computed for all values of the uniformity ρ are sketched in Fig. 11 as a function of the embedding parameters. It can be seen that the average percentage error of the proposed unified formulation is below 4% in extreme cases whereas the SCS's error is more than 27%. This figure demonstrates the accuracy and superiority of the proposed unified formulation.

distortion compensation (α) = 0.71

(a)

DWR (dB) = 28

(b)
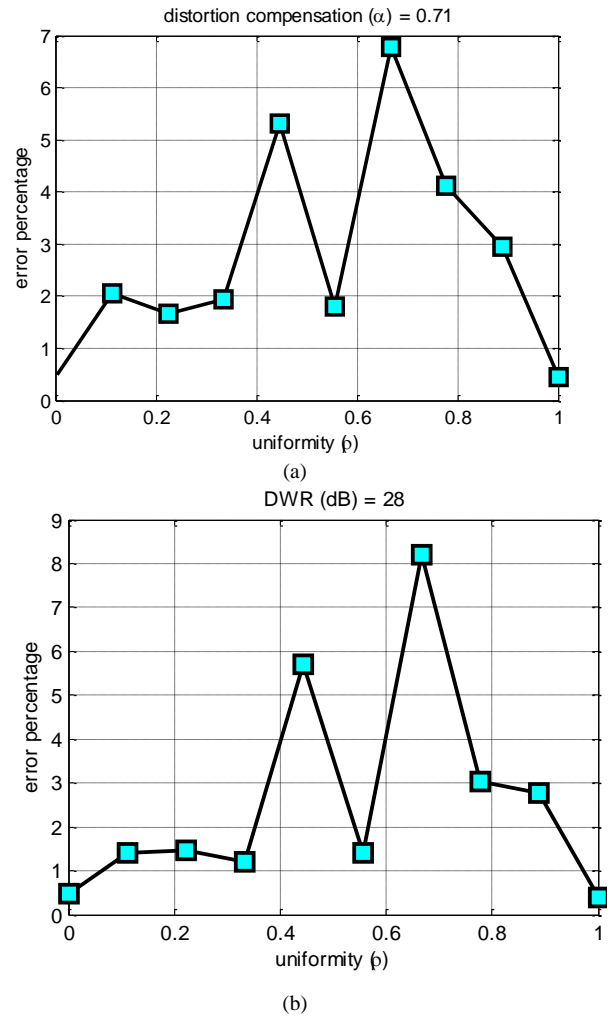
Fig. 10. Percentage error of the proposed unified formulation in comparison with the practical security evaluation of the test images for some embedding parameters.
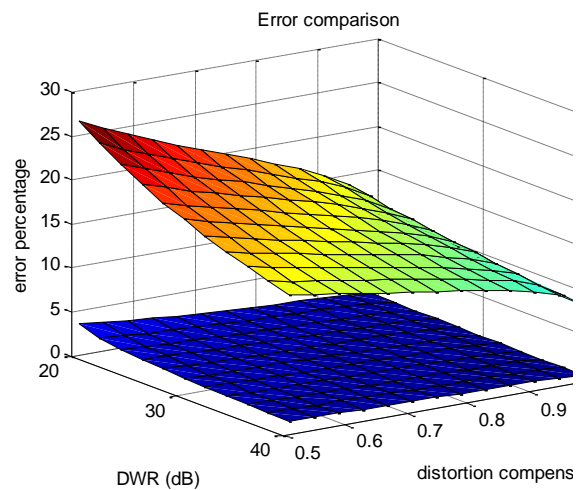
Error comparison

Fig. 11. Percentage error of the proposed unified formulation in comparison with the practical security evaluation of the test images for whole embedding parameters; the lighter sketch corresponds to former theoretical equation in comparison with the practical security evaluation.

5. Conclusions

This paper developed a novel information-theoretic security analysis of the scalar Costa scheme for the digital image watermarking applications. Based on the findings of this study, presence of the uniform regions in the images invalidates flat-host assumption, considered as a founded assumption by previous researches. It was shown that the flat-host assumption results in a large difference between the theoretical and practical security evaluations. To tackle this shortcoming and develop a generic analysis framework for the digital image watermarking, in this paper a theoretical equation was derived for the uniform regions while an empirical formulation was proposed for the textured (non-uniform) regions of the image. Moreover, to tackle another invalid assumption in previous works, this study assumed the observations are a priori dependent. It was demonstrated that assuming independent observation leads to the deviation of the theoretical results from the reality, even in case of non-uniform images. Finally, the theoretical equation for uniform regions and empirical equation for non-uniform regions were unified to produce a single formulation for the whole image. The final unified equation helps watermarking system designers to perform reliably designs and developments.

References

[1] M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439- 441, 1983.

[2] J. Eggers, R. Bauml, R. Tzschoppe and B. Girod, "Scalar Costa scheme for information embedding," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 1003-1019, 2003.

[3] R.M. Gray and D.L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325-2383, 1998.

[4] T. Kalker, "Considerations on watermarking security," *2001 IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 201-206, 2001.

[5] F. Cayre, C. Fontaine and T. Furon, "Watermarking security: theory and practice," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3976- 3987, 2005.

[6] T. Mittelholzer, "An information-theoretic approach to steganography and watermarking," *Proceedings of the Third International Workshop in Information Hiding*, pp. 1-16, 1999.

[7] P. Bas, and T. Furon, "A New Measure of Watermarking Security: The Effective Key Length," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 8, pp. 1306-1317, 2013.

[8] L. Perez-Freire, F. Perez-Gonzalez, T. Furon and P. Comesana, "Security of Lattice-Based Data Hiding Against the Known Message Attack," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 4, pp. 421- 439, 2006.

[9] L. Perez-Freire, and F. Perez-Gonzalez, "Security of Lattice-Based Data Hiding against the Watermarked-Only Attack," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 593-610, 2008.

[10] B.R. Matam, and D. Lowe, "Watermark-only security attack on DM-QIM watermarking: Vulnerability to guided key guessing," *International Journal of Digital Crime and Forensics*, vol. 2, no. 2, pp. 64-87, 2010.

[11] R. Samadi, and S.A. Seyedin, "Security assessment of scalar costa scheme against known message attack in DCT-domain image watermarking," *Electrical Engineering (ICEE), 2013 21st Iranian Conference on*, pp. 1-5, 14-16 May 2013.

[12] F. Chuhong, D. Kundur and R.H. Kwong, "Analysis and design of secure watermark-based authentication systems," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 1, pp. 43- 55, 2006.

[13] J. Hernandez, M. Amado and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: detector performance analysis and a new structure," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 55-68, 2000.

[14] A. Kerckhoffs, "La cryptographie militaire," *Journal des sciences militaires*, vol. 9, pp. 5-38, 1883.

[15] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Transactions on Information Theory*, vol. 42, no. 4, pp. 1152-1159, 1996.

[16] A. Papoulis and U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[17] L. Perez-Freire and F. Perez-Gonzalez, "Spread-Spectrum Watermarking Security," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 2-24, 2009.

[18] USC-SIPI image database. [online]. available: http://www.sipi.usc.edu /database/.

[19] TESTIMAGES archive. [online]. available: http://www.tecnick.com/public/code/cp_dpage.php?aiocp_dp=testimages.

[20] DIP3/e—Book Images. [online]. available: http://www.imageprocessingplace.com/DIP-3E/dip3e_book_images_downloads.htm.

[21] CIPR Still Images. [online]. available: http://www.cipr.rpi.edu/ resource/stills/index.html.

Reza Samadi (M'2011) is graduated from Sharif University of technology in 2006. Now he is pursuing his PhD in Ferdowsi University of Mashhad. His research interest covers information security, physical layer security, and applications of information theory.

Seyed Alireza Seyedin was born in Mashhad. He received the B.S. degree in Electronics Engineering from Isfehan University of Technology, Isfehan, Iran in 1986, and the M.E. degree in Control and Guidance Engineering from Roorkee University, Roorkee, India in 1992, and the Ph.D. degree from the University of New South Wales, Sydney, Australia in 1996. He has been an Associate Professor with the Department of Electrical Engineering, the Ferdowsi University of Mashhad, Mashhad, Iran. His research interest includes image processing, computer vision, signal processing, and pattern recognition. In these fields, specially, he is interested in image analysis, motion detection and estimation in image sequences, autonomous vehicles, and diverse applications of the radon transform.

Tracking Performance of Semi-Supervised Large Margin Classifiers in Automatic Modulation Classification

Hamidreza Hosseinzadeh*
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
hr.hosseinzadeh@srbiau.ac.ir
Farbod Razzazi
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
razzazi@srbiau.ac.ir
Afrooz Haghbin
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
haghbin@srbiau.ac.ir

Abstract

Automatic modulation classification (AMC) in detected signals is an intermediate step between signal detection and demodulation, and is also an essential task for an intelligent receiver in various civil and military applications. In this paper, we propose a semi-supervised Large margin AMC and evaluate it on tracking the received signal to noise ratio (SNR) changes to classify most popular single carrier modulations in non-stationary environments. To achieve this objective, two structures for self-training of large margin classifiers were developed in additive white Gaussian noise (AWGN) channels with priori unknown SNR. A suitable combination of the higher order statistics (HOS) and instantaneous characteristics of digital modulation are selected as effective features. We investigated the robustness of the proposed classifiers with respect to different SNRs of the received signals via simulation results and we have shown that adding unlabeled input samples to the training set, improve the tracking capacity of the presented system to robust against environmental SNR changes. The performance of the automatic modulation classifier is presented in the form of k-fold cross-validation test, classification accuracy and confusion matrix methods. Simulation results show that the proposed approach is capable to classify the modulation class in unknown variable noise environment at even low SNRs.

Keywords: Automatic Modulation Classification; AMC; Tracking Performance Evaluation; Passive-Aggressive Classifier; Self Training; Semi-Supervised Learning.

1. Introduction

Automatic modulation classification (AMC) is the process of identification of the modulation type of a signal in a general non-cooperative environment. It has received significant attention for over two decades now. However, nowadays, there is a vast variety of applications that it is essential to detect and demodulate the signal without given priori information about the received signal [1-3]. In such cases, unlike regular receivers in which the primitive information of the received signals, such as carrier frequencies, frequency bandwidth, bit rate, and modulation type, is available, there is no information available for the received signal, and the receiver is blind. In this case, an automatic modulation classifier as an intermediate step between signal interception and information recovery, helps the receiver employ the correct decoder. The AMC can be used in a wide range of applications including electronic warfare applications, intelligent services systems, spectrum monitoring, signal surveillance, interferer identification and cognitive radio applications [4].

Generally, digital signal type classification algorithms may be divided into two major categories: decision theory based approaches (DTBAs) and feature matching based approaches (FMBAs) [5-6]. DTBA algorithms are based on the received signal likelihood function and hypothesis testing arguments to formulate the classification problem. In contrast, FMBA algorithms usually extract the features form the received signal and determine the membership of signal to each class. The calculation complexity of FMBA algorithms is lower than DTBA algorithms, and they are easy to implement. Our proposed tracking classifier is categorized as FMBA algorithm.

From the published works in AMC classifiers [7-8], it appears clear that unfortunately, most of the previous studies rely on the availability of large labeled datasets to analyze the classifier. However, it is not applicable in most non stationary environments applications. Thus, it is required to deal with both labeled and unlabeled data to train the classifier, simultaneously. In this approach, the classifier is first trained with labeled data and then used to predict the labels of the unlabeled data. A subset of the unlabeled data, together with their predicted labels, is then selected to augment to the labeled data. This training model is called self-training [9-12].

In this paper, we propose two large margin architectures to track the signal-to-noise ratio (SNR)

---

* Corresponding Author

changes of input signals, and then classify the modulated signals. The architectures are developed based on passive-aggressive online algorithm [13] to train the classifier and determine the modulated signal type. Our preliminary result on tracking performance of the proposed classifiers was reported in [14].

The remainder of the paper is organized as follows: in section 2 we provide a description of selected feature extraction and margin based algorithms as the basis of the proposed architectures. Brief information on the development of passive-aggressive algorithm is also provided in this section. In Section 3, we present the proposed architectures for classifier tracking. The simulation results are presented in section 4. Finally, the paper is concluded in section 5 with a summary of our findings.

2. Problem Description

According to the increasing use of digital systems and usage of digital signals in software radio, the research was focused on digital signal types. Considering the changes in message parameters, there are four general digital signal types, M-ASK, M-PSK, M-FSK and M-QAM [15]. The modulation techniques of digital input signals, which are considered in this paper, are FSK2, FSK4, ASK2, ASK4, PSK2, PSK4, PSK8, QAM16, QAM32, and QAM64.

2.1 Feature Extraction

Feature extraction is a crucial part of a pattern recognition system where its aim is to reveal the distinctive properties of an object to be recognized. In this paper, the effective features are considered as a combination of higher order statistics and instantaneous characteristics of digital signal types. The rest of this subsection is devoted to describe these features briefly.

2.1.1 Instantaneous Feature

Instantaneous features are suitable for signals which contain instantaneous phase or instantaneous frequency [16]. In this work, the instantaneous features for classification were selected from the proposed features by Azzouz and Nandi [17-18]. These features were derived from the instantaneous properties of the received signals. Therefore, these features are called as instantaneous features. The instantaneous key features which were used for the proposed tracking algorithm were derived from the instantaneous amplitude $a(t)$, and the instantaneous frequency $f(t)$, of the signal under consideration.

The first feature is the maximum value of the power spectral density of the normalized-centered instantaneous amplitude of the intercepted signal which is formulated as follows:

$$\gamma_{max} = max\left(\frac{|DFT(a_{cn}(i))|^2}{N_s}\right) \qquad (1)$$

Where N_s is the number of the sample in the range and $a_{cn}(i)$ is value of centralized normalized instantaneous amplitude that is defined by

$$a_{cn}(i) \overset{\Delta}{=} a_n(i) - 1, \qquad a_n(i) \overset{\Delta}{=} \frac{a(i)}{m_a} \qquad (2)$$

and m_a is the average value of instantaneous amplitude over one frame, i.e.

$$m_a = \frac{1}{N_s}\sum_{i=1}^{N_s} a(i) \qquad (3)$$

This feature is designed to discriminative between constant envelopes (CE) signals (e.g., FSK and PSK) and non-CE signals (e.g., ASK).

The second feature is the standard deviation of absolute value of normalized-centered instantaneous frequency over non-weak segments of the intercepted signal which is calculated as:

$$\sigma_{af} \overset{\Delta}{=} \sqrt{\frac{1}{L}\left[\sum_{a_{cn}(i)>a_t} f_{cn}^2(i)\right] - \left[\frac{1}{L}\sum_{a_{cn}(i)>a_t}|f_{cn}(i)|\right]^2} \qquad (4)$$

where $f_{cn}(i)$ is the centralized normalized instantaneous frequency and it is defined by

$$f_{cn}(i) = \frac{f_c(i)}{r_b}, \quad f_c(i) - m_f, \quad m_f = \frac{1}{N}\sum_{i=1}^{N} f(i) \qquad (5)$$

where r_b is the bit rate, and a_t is a preset threshold for detecting non-weak samples because instantaneous frequency is very noise sensitive. In this paper, the threshold for detection of non-weak samples is chosen as $a_t = 0.95$ [17]. This feature is designed to discriminative between FSK signals.

2.1.2 Higher Order Statistics (HOS)

The first set of employed statistical features is moments. A moment of a random variable may be defined as:

$$M_{p,q} = E[s^{p-q}(s^*)^q] \qquad (6)$$

Where p is called the moment order and s* stands for the complex conjugation of s.

The second set of employed statistical features is cumulants which is the most widely used feature in this area. The symbolism for pth order cumulants is similar to that of the pth order moment.

$$C_{p,q} = Cum[s, …, s, s^*, …, s^*] \qquad (7)$$

The mentioned expression have *(p-q)* terms of *s*, in addition to *q* terms of *s**. Cumulants may be expressed in term of moments as

$$Cum[s_1, …, s_n] =$$
$$\sum_{\forall v}(-1)^{q-1}(q-1)! \, E\left[\prod_{j\in v_1} s_j\right] … E\left[\prod_{j\in v_q} s_j\right] \qquad (8)$$

where the summation index is over all partition $v = (v_1, \ldots, v_q)$ for the set of indices $(1, \ldots, n)$, and q is the number of elements in a given partition.

Based on fisher discriminant analysis (FDA) [19], we selected a proper set of higher order moment and cumulants as below. FDA represents the capability of the selected features for separation of two predefined classes and is defined by

$$d_{ij} = \frac{(\mu_i - \mu_j)^2}{\sigma_1^2 + \sigma_2^2} \quad i \neq j \tag{9}$$

where μ and σ are mean and variance of these two classes. The important selected statistical features are $M_{41}, M_{61}, M_{84}, C_{40}, C_{61}, C_{63}, C_{80}, C_{82}, C_{84}$. Unfortunately, these characteristics are noise dependent. Therefore, a classifier tracking strategy should be devised to decrease the effect of this dependency, as far as possible.

2.2 Large-Margin Classifier

Suppose to have a binary classification problem and a training set of l labeled samples $\{x_i, y_i\}_{i=1}^{l}$ and u unlabeled samples $\{X_i\}_{i=1}^{u}$ where $x_i \in R^d$ is an input vector describing i^{th} samples and $y_i \in \{-1,1\}$ is its labels. We want to learn a discriminative function f in online and assign the correct label to an unseen new test samples.

2.2.1 Passive-Aggressive Classifier

Passive-Aggressive (PA) algorithm [13] is a maximum margin based learning algorithm, which has been mainly used in online learning. Online PA algorithm, on one hand, modifies the current classifier $w_i. x + b$ in order to correctly classify the current example x_i, by updating the weight vector from w_i to w_{i+1}; and on the other hand, the new classifier $w_{i+1}. x + b$ should be as close as possible to the current classifier $w_i. x + b$. Our idea for tracking the environmental conditions was based on PA algorithm, and we pursued both above ideas at the same time. The vector w_1 is initialized to $(0, \ldots, 0)$. In the time i, the new weight vector w_{i+1} was determined by solving the following optimization problem,

$$w_{i+1} = argmin_w \frac{1}{2} \|w - w_i\|^2$$
$$s.t. \ (w_{pi}. x_i - w_{yi}. x_i) \geq 1 \tag{10}$$

where w_{pi}, w_{yi} are the weight vectors that is produced by using predicted labels and true labels respectively. With a Lagrange multiplier, τ, the stationary points of the

Lagrangian can be computed. Then, we can reformulate the optimization problem as follows:

$$w_{i+1} = w_i + \tau_i x_i; \ \tau_i = \frac{1 - (w_{yi}.x_i - w.x_i)}{2\|x_i\|^2} \tag{11}$$

In order to ensure convergence of the algorithm, Cramer et al. [13] have presented a scheme that employ robust update strategy. This scheme has been obtained by introducing non-negative slack variable into the optimization problem. It is derived with soft-margin classifiers [20]. This makes the optimization problem (10) to the following formulation:

$$w_{i+1} = \min_w \frac{1}{2} \|w - w_i\|^2 + C\xi$$
$$s.t. \ (w_{pi}. x_i - w_{yi}. x_i) \geq 1 - \xi, \ \xi \geq 0 \tag{12}$$

where ξ is slack variable, and c is a penalty parameter.

In real-world problems, to improve the discriminative power of f, the training data are usually mapped into a higher-dimensional feature space via a non-linear mapping $\emptyset(x)$, induced by a kernel function $k(x, x') = \emptyset(x). \emptyset(x')$. In the case of multiclass classification, the prediction output is a vector in $\mathbb{R}^k$ where each element in the vector corresponds to score assigned to the respective label.

These score calculations, have been devised in [21]. The prediction of PA algorithm is set to be the label with the highest score.

3. Proposed Architectures for Tracking Performance Evaluation

In this section, we propose two large margin structures, using the mentioned passive-aggressive algorithm and evaluate the tracking performance and classification accuracy of them. In this study, the SNR of the signals is a priori unknown. The tasks of the proposed tracking algorithms are adaptation to the environment's SNR in addition to detecting the modulation type. In the tracking performance evaluation, it is assumed that the SNR changes are much slower than signal changes. To evaluate the system, we made a synthetic signal which the SNR of its temporal segments were decreased gradually. In the rest of this section, we released two systems for tracking performance evaluation.
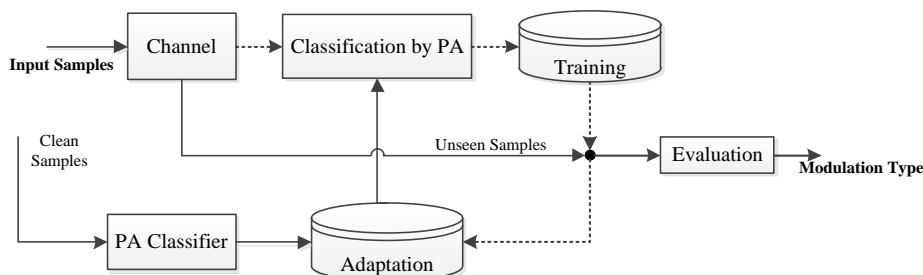
Fig. 1. The overall block diagram of batch learning

3.1   Batch Learning Architecture for Tracking

In this section, we introduce a semi-supervised batch learning architecture. This architecture is semi-supervised since using a reasonable set of recognized samples; it is capable to classify the input samples correctly. The system is first adapted to the unlabeled training samples to have a new adapted trained classification system. In the classification phase, this new classifier is used to classify unknown samples.

The overall block diagram of the batch learning system is presented in Fig. 1. Following the main signal flow in the architecture, we trained a general classifier using the union of gathered features of the clean samples (SNR = ∞). The clean samples are generated noise-free. The unlabeled input signals are randomly divided into s equal size subsets. Subsequently, PA classifiers are used as test labels and then adapted to the new subset by $s - 1$ times. In other words, during the each iteration, one of the s subsets is introduced to this system to classify.

Then, using the predicted samples of this subset, the classifier was adapted.

3.2   Online Learning Architectures for Tracking

In this section, we propose self-training classifier architecture so that the classification result of each sample is determined and evaluated as the sample enters the system. The evaluated samples are then re-used as new training samples. In this study, it is assumed that the label of input samples is not determined.

The block diagram of online learning system is shown in Fig. 2. As depicted, the classifier is first trained with the clean labeled samples. The obtained result is used to classify a new unlabeled input sample that have been entered, and classify them. Then, the unlabeled input samples that are predicted one by one is collected to augment into the labeled samples gradually. Then, the classifier is adapted using this new labeled sample and the result is used to classify the next unlabeled samples every
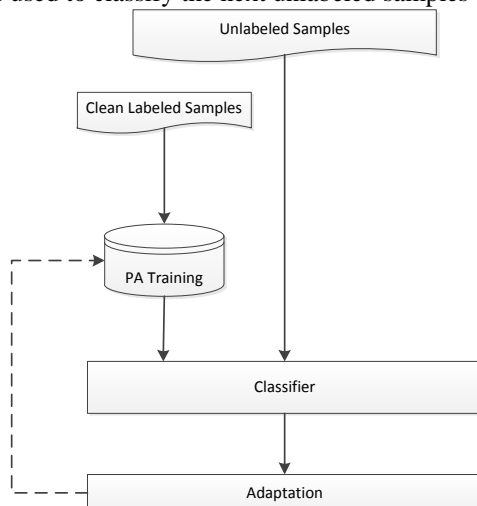
Fig. 2. The overall block diagram of online learning

time. This procedure is repeated until the last unlabeled sample is entered. The technique is presented in Fig. 2.

4.   Experimental Results

This section presents simulation results on the proposed architectures. The radial basis function (RBF) of the form $k(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$ is employed as the kernel function of PA classifier. In practice, the standard method to determine the kernel and misclassification penalty parameters is through the grid search method [22]. Therefore, a grid search technique was used to find the optimal values of these parameters.

The carrier frequency (f_c) was assumed to be 11 MHz, the sampling rate (f_s) was 86 kHz, and symbol rate (r_s) was 10 kHz. We assumed that the carrier frequency had been correctly estimated before or it had been known. Therefore, we only considered complex baseband signals. In addition, it was assumed that the simulated signals were bandwidth limited. The Gaussian noise was added according to 0 dB, 3 dB, 4 dB, 6 dB, 8 dB, 9 dB, 12 dB and 20 dB SNRs. Each signal used in this study was generated using MATLAB. For every signal we generated 300 realizations which created randomly for every trial to ensure results are independent of the considered input samples. Therefore, the number of samples in a segment is 3000. To evaluate the system, we made a synthetic signal with 8 different segments which the SNR of its temporal segments were decreased gradually form 20 dB to 0 dB. This results in 24000 samples in the synthetic signal. We have adopted 3-fold cross-validation procedure on labeled and unlabeled dataset.

We compared the performance of the algorithms on the basis of the classification accuracy. Classification accuracy for experiment is taken as the ratio of the number of samples correctly classified to the total number of samples.

Table 1. Classification rate of proposed batch learning architecture in different SNRs (%)

Modulation classes	SNR			
	0 dB	4 dB	8 dB	12 dB
FSK2	54.2	58.1	81.2	100
FSK4	76.1	93.5	100	100
ASK2	66.7	91.9	100	100
ASK4	58.4	74.0	91.0	100
PSK2	80.3	100	100	100
PSK4	87.1	95.8	100	100
PSK8	81.1	87.8	100	100
QAM16	44.0	61.9	75.5	96.6
QAM32	84.1	81.1	100	100
QAM64	81.0	99.2	100	100
Mean	71.30	84.33	94.78	99.66

Table 2. Classification rate of proposed online learning architecture in different SNRs (%)

Modulation classes	SNR			
	0 dB	4 dB	8 dB	12 dB
FSK2	95.0	100	100	100
FSK4	99.8	100	100	100
ASK2	80.8	94.8	100	100
ASK4	69.6	87.0	100	100
PSK2	92.0	100	100	100
PSK4	71.3	89.1	100	100
PSK8	70.3	88.0	100	100
QAM16	99.8	99.9	98.9	100
QAM32	70.1	87.6	100	100
QAM64	66.5	83.1	100	100
Mean	81.52	92.95	99.89	100

Table 3. Confusion Matrix of proposed batch learning algorithm in SNR=4 dB (%)

True modulations	Predicted modulations									
	FSK2	FSK4	ASK2	ASK4	PSK2	PSK4	PSK8	QAM16	QAM32	QAM64
FSK2	58.1	41.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FSK4	0.0	93.5	0.0	0.0	0.0	0.0	0.0	0.0	6.5	0.0
ASK2	0.0	0.0	91.9	6.6	0.0	1.5	0.0	0.0	0.0	0.0
ASK4	0.3	0.0	15.5	74.1	0.0	10.1	0.0	0.0	0.0	0.0
PSK2	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0
PSK4	0.0	0.0	0.0	0.0	0.0	95.8	0.0	4.2	0.0	0.0
PSK8	0.0	0.0	0.0	0.0	0.0	0.0	87.8	0.0	12.2	0.0
QAM16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	61.9	0.0	38.1
QAM32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	81.2	18.8
QAM64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	99.3

Table 4 Confusion Matrix of proposed online learning algorithm in SNR=4 dB (%)

True modulations	Predicted modulations									
	FSK2	FSK4	ASK2	ASK4	PSK2	PSK4	PSK8	QAM16	QAM32	QAM64
FSK2	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FSK4	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ASK2	0.0	0.0	94.8	5.2	0.0	0.0	0.0	0.0	0.0	0.0
ASK4	0.2	0.0	12.7	87.1	0.0	0.0	0.0	0.0	0.0	0.0
PSK2	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0
PSK4	0.0	0.0	0.0	0.0	0.0	89.2	0.0	10.8	0.0	0.0
PSK8	0.0	0.0	0.0	0.0	0.0	0.0	87.9	6.1	6.0	0.0
QAM16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	0.1
QAM32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.3	87.7	0.0
QAM64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.8	0.0	83.2

Classification accuracy assessments of different classes have been provided by the confusion matrix and accuracy matrix in percentage.

4.1 Tracking Performance Evaluation of Proposed Architectures

In this section, we have evaluated the performance of proposed architectures with different SNR values via simulation. Tables 1-2 show the classification rates of the proposed classifier in 0, 4, 8 and 12 dB SNRs.

From the mentioned results in Tables 1-2, it can be deduced that the performance of classifier in different SNR were generally good. This fact is because of the adaptation ability of the proposed classifiers. However, the performance is slightly degraded in lower SNRs. It should be noted that the previous Offline algorithms have not been able to learn by 300 samples.

As a sample, the confusion matrix is presented at SNR=4 dB to analyze in the confusion of different classes. These results are presented in Tables 3-4.

As it is observed in Table 4, the recognition accuracy of online batch learning architecture for all signals except QAM64 was very good even at low SNRs. It can be seen that there is a tendency for QAM64 modulation to be mostly confused with QAM16 modulation. Because the constellation shape of these classes are very similar.

In addition, the tracking performance of online learning classifier is compared to accuracy in non-variable stationary environment in the same SNR. Numerical results of the accuracy in stationary environment were reported in [23]. This comparison is shown in Fig. 3.

According to Fig. 3, it can be observed that the classification accuracies which were obtained from the proposed tracking methods are close to classification accuracy in stationary environment.
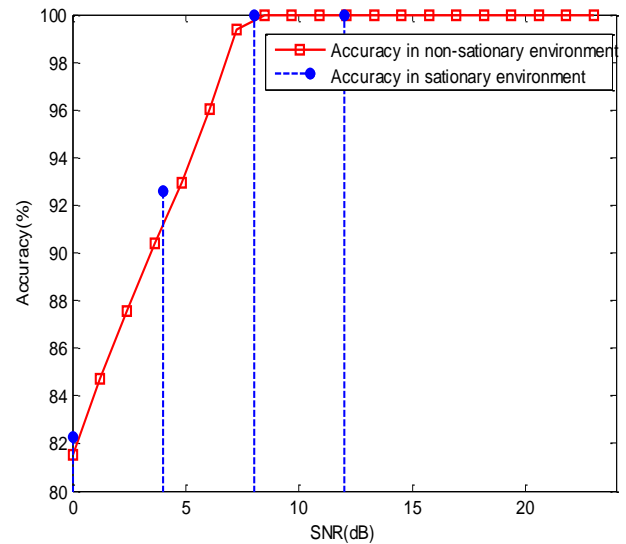

Fig. 3 The curve of accuracy tracking relative to SNR

Table 5. Tracking Performance comparison (%)

SNR	Tracking evaluation of proposed online learning architecture	Tracking evaluation of proposed batch learning architecture	Supervised PA classifier with priori labeled data	Supervised SVM classifier with priori labeled data
0 dB	81.82	80.06	85.06	84.18
4 dB	92.93	88.27	97.21	97.36
8 dB	99.84	99.84	99.84	99.84
12 dB	100	100	100	100
Noise-free	100	100	100	100

4.2 Tracking Performance Comparison

This section, demonstrates the effectiveness of the proposed methods. Therefore, the performance of the proposed tracking architectures were compared to the tracking algorithm that was trained by PA algorithm in the matched SNR (SNR aware mode) with priori labeled data. In addition, this comparison was repeated for SVM algorithm. The results are indicated in Table 5.

As it is observed in Table 5, all of the evaluated algorithms have similar performance in SNR≥ 8. Simulation results show that performance of proposed algorithms deteriorate with decreasing SNR. The tracking performance of online learning classifier is 98% of supervised PA classifier with priori labeled data. Therefore, the proposed algorithm has generally good performance and its accuracy is close to the analysis in the matched SNR.

5. Conclusions

Automatic modulations classification plays a significant role in civil and military applications. In this paper, a new method are developed for classifier tracking in non-stationary environments using a semi-supervised self-trained large margin classifier for classification of most popular single carrier modulations, i.e., FSK2, FSK4, ASK2, ASK4, PSK2, PSK4, PSK8, QAM16, QAM32, and QAM64 which are commonly used in a cognitive radio. Towards this objective, two features are employed that include a selected set of the instantaneous characteristics and higher order statistics of received signal. Simulation results show that the approach is capable to classify the modulation class in unknown variable noise environment at even low SNRs. In addition, the evaluation of the tracking performance shows that the proposed architectures have a good ability to adapt to the environment. The tracking performance evaluation for high frequency band, fading and multipath phenomena in addition to multi-carrier modulation schemes will be investigated as our future work.

References

[1] J. J. Popoola and R. van Olst, "A Novel Modulation-Sensing method," IEEE Vehicular Technology, Vol. 6, No. 3, Sep. 2011, pp. 60-69.

[2] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," IEEE Circuits and Systems, Vol. 2, No. 2, Second Quarter 2009, pp. 27-45

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Select. Areas Commun., Vol. 23, No. 2, 2005, pp. 201-220.

[4] J. J. Popoola and R. van Olst, "The performance evaluation of a spectrum sensing implementation using an automatic modulation classification detection method with a Universal Software Radio Peripheral," Expert Systems with Applications, Vol. 40, No. 3, 2013, pp. 2165-2173.

[5] O. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," IET Commun., Vol. 1, No. 2, Apr. 2007, pp. 137-156.

[6] H. C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," IEEE Trans. Wireless Commun., Vol. 7, No. 8, 2008, pp. 3089-3105.

[7] A. Ebrahimzadeh, and R. Ghazalian, "Blind digital modulation classification in software radio using the optimized classifier and feature subset selection," Engineering Applications of Artificial Intelligence, Vol. 24, No. 1, 2011, pp. 50-59.

[8] Z. Zhu, M. W. Aslam, A. K. Nandi, "Genetic algorithm optimized distribution sampling test for M-QAM modulation classification," Signal Process., Vol. 94, 2014, pp. 264-277.

[9] M. M. Adankon and M. Cheriet, "Help-Training for semi-supervised support vector machines," Pattern Recognition, Vol. 44, No. 1, Sep. 2011, pp. 2220-2230.

[10] X. Zhu, "Semi-supervised learning literature survey", Technical Report 1530, Computer Sciences, University of Wisconsin-Madison, 2008.

[11] X. Zhu and A. B. Goldberg, Introduction to Semi-Supervised Learning, Morgan and Claypool Publishers, 2009.

[12] O. Chapelle, B. Scholkopf, A. Zien (Eds.), Semi-Supervised Learning, MIT Press: Cambridge, MA, 2006.

[13] K. Crammer, O. Dekel, J. Keshet, S. Shalev-shwartz, and Y. Singer, "Online passive-aggressive algorithms," Journal of Machine Learning Research, Vol. 7, 2006, pp. 551-585.

[14] H. Hosseinzadeh, F. Razzazi, A. Haghbin, "Tracking Performance of Online Large Margin Semi-supervised classifiers in Automatic Modulation Classification." In: Proceedings of International Symposium on Telecommunications (IST'2012), 2012, pp. 387-392.

[15] J. G. Proakis, Digital Communications, McGraw-Hill: New York, 2001.

[16] J. Lopatka and P. Macrej, "Automatic modulation classification using statistical moments and a fuzzy classifier," in Proc. Of International Conference on Signal Processing (ICSP'00), Aug. 2000, Vol. 3, pp. 1500-1506.

[17] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types," Signal Processing, Vol. 47, No. 1, Nov. 1995, pp.55-69.

[18] E. E. Azzouz and A. K. Nandi, "Algorithms for automatic modulation recognition of communication signals," IEEE Trans. Commun., Vol. 46, No. 4, 1998, pp.431-436.

[19] R. Duda, P. Hart, and D. Stock, Pattern Classification, Wiley, 2000.

[20] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, Inc.: New York, 1998.

[21] K. Crammer and Y. Singer, "A new family of online algorithms for category ranking," Journal of Machine Learning Research, Vol. 3, 2003a, pp. 1025–1058.

[22] C. C. Burges, "A tutorial on support vector machines for pattern recognition," In Proceedings of Int. Conference on Data Mining and Knowledge Discovery, Vol. 2, No. 2, 1998, pp. 121-167.

[23] H. Hosseinzadeh, F. Razzazi, A. Haghbin, "Online Large Margin Semi-supervised Algorithm for Automatic Classification of Digital Modulations." In: Proceedings of 21st Iranian Conference on Electrical Engineering (ICEE'2013), 2013, pp. 1-6.

Hamidreza Hosseinzadeh received the B.Sc. degree in Electrical Engineering from Shahr-e-Rey branch, Islamic Azad University, Tehran, Iran, in 2009 and the M.Sc. degree in Electrical Engineering from Science and Research Branch, Islamic Azad University, Tehran, Iran, in 2012. He is currently a Ph.D. candidate at department of Electrical and Computer Engineering in Islamic Azad University, Science and Research Branch, Tehran, Iran. His research areas of interest include signal processing, pattern recognition, and communication systems.

Farbod Razzazi received his B.Sc. and M.Sc. in Electrical Engineering from Sharif University of Technology in 1994 and 1996, respectively. He received his Ph.D. from Amirkabir University of Technology (Tehran Polytechnic) in 2003 in Electrical Engineering. He is currently an Assistant Professor in Department of Electrical Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran. His current research interests are pattern recognition methods and their applications in speech recognition systems.

Afrooz Haghbin received her B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2001. She received her M.Sc. degree from Tehran University and her PhD degree from Tarbiat Modares University, Tehran, Iran, all in electrical engineering in 2004 and 2009, respectively. She is currently with the electrical and computer department of Science and Research Branch, Islamic Azad University, Tehran, Iran, as assistant professor. Her research interests include MIMO wireless communications, channel coding, pre-coding, multicarrier modulation and estimation theory.

Joint Source and Channel Analysis for Scalable Video Coding Using Vector Quantization over OFDM System

Farid Jafarian
Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.
farid.jafarian@birjand.ac.ir
Hassan Farsi*
Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran
hfarsi@birjand.ac.ir

Abstract

Conventional wireless video encoders employ variable-length entropy encoding and predictive coding to achieve high compression ratio but these techniques render the extremely sensitive encoded bit-stream to channel errors. To prevent error propagation, it is necessary to employ various additional error correction techniques. In contrast, alternative technique, vector quantization (VQ), which doesn't use variable-length entropy encoding, have the ability to impede such an error through the use of fix-length code-words. In this paper, we address the problem of analysis of joint source and channel for VQ based scalable video coding (VQ-SVC). We introduce intra-mode VQ-SVC and VQ-3D-DCT SVC, which offer similar compression performance to intra-mode H.264 and 3D-DCT respectively, while offering inherent error resilience. In intra-mode VQ-SVC, 2D-DCT and in VQ-3D-DCT SVC, 3D-DCT is applied on video frames to exploit DCT coefficients then VQ is employed to prepare the codebook of DCT coefficients. In this low bitrate video codecs, high level robustness is needed against the wireless channel fluctuations. To achieve such robustness, we propose and calculate optimal codebook of VQ-SVC and optimal channel code rate using joint source and channel coding (JSCC) technique. Next, the analysis is developed for transmission of video using an OFDM system over multipath Rayleigh fading and AWGN channel. Finally, we report the performance of these schemes to minimize end-to-end distortion over the wireless channel.

Keywords: Vector Quantization; Joint Source and Channel Coding; Video Coding; Wireless Channel.

1. Introduction

Source coding is an inseparable component in a video transmission system and is required to ensure manageable transmitted bitrates [1]. H.264/AVC is one of the most popular codecs in use today, and leans on de-correlating transforms, inter/intra prediction, and variable-length entropy coding. Another codec is 3D-DCT video encoder. Although H.264/AVC provides higher coding efficiency than 3D-DCT but 3D-DCT encoder has several advantages compared to H.264/AVC, such as reduced number of operations per pixel because of no motion estimation/compensation required and symmetric encoding-decoding [2].

With increasing interest on real-time video communications over wireless channels, the complex and challenging problem rise up due to the multi-path fading characteristics of the channel [3]. To prevent such error propagation, it is necessary to employ various additional error resilience and error correction techniques [1]. The VQ based encoders/decoders are an alternative which do not use variable-length entropy encoding and have the ability to prevent such an error, also it can be employed to very low bitrate video applications due to its compression performance rather than inter/intra predictive codecs [4]-[6].

The concept of Forward Error Correction (FEC) is to add redundancy bits to the block of video data for the purpose of error detection and correction. In common communication systems, source and channel coding have traditionally been designed independently of each other. This is justified by Shannon's principle which says that there is no performance loss if the two functions are separately treated [7]. However Shannon's theory is an asymptotic result that permits unlimited delay and complexity. In fact this principle relies on the crucial assumption that the source and channel codes can be of arbitrary long lengths. In practice, because of limitations on the computational power and processing delays this assumption does not hold. It is then of benefit to consider the problem of source and channel coding jointly. A review of JSCC can be found in [7]-[9].

In some streaming video applications like Unmanned Ground Vehicle (UGV) and Ground Control Station (GCS) communication scenario [10], due to power restriction, it is required to provide very low bitrate and a robust system to wireless video transmission. To this end, we introduce intra-mode VQ-SVC and VQ-3D-DCT SVC. In the intra-mode VQ-SVC encoder, 2D-DCT and in the VQ-3D-DCT SVC encoder, 3D-DCT is applied on video frames to exploit DCTs coefficients. Then VQ is employed to prepare the codebook of DCTs coefficients.

---

* Corresponding Author

In this very low bitrate video codecs, high level robustness is needed against the wireless channel fluctuations. To achieve such robustness, we propose and calculate optimal codebook of VQ-SVC and optimal channel code rate (CCR) using JSCC technique. Next, the analysis is developed for transmission of video using an OFDM system over multipath Rayleigh fading and AWGN channel.

In section 2, a brief of scalable video coding and VQ is introduced. In Section 3, the optimal codebook of VQ-SVC is presented. Joint source and channel rate optimization, the rate allocation algorithm, is presented in Section 4. Section 5 details simulation setting and in Section 6 simulation results are presented.

2. Summary of scalable video coding and vector quantization

2.1 Scalable Video Coding (SVC)

SVC selects a layered coding scheme to generate multi-layered bit-stream for heterogeneous environments. Scalable video coding is used for various transmission mediums which have various constraints. As today's wireless networks are not only widespread but they also represent a substantial part of larger heterogeneous networks, the support for low-bandwidth and robust transmission is highly desirable. As a general concept, a SVC stream has a base layer and several enhancement layers. As long as the base layer is received, the receiver can decode the video stream. As more enhancement layers are received, the decoded video quality is improved [11].

The SVC design which is an extension of the H.264/AVC [12-13] video coding standard can be classified as a layered video codec. SVC-based layered video coding is suitable for supporting heterogeneous devices with a scalable bit-stream. Such a stream allows for delivering a presentable quality (presentable quality refers to resolution, frame rate and bit rate of a decoded operation point of the scalable video bit stream) of the video depending on the device's capabilities. Paper [14] proposes two SVC structures for high efficiency video coding: one based on multi-loop decoding and the other based on single-loop decoding with High Efficiency Video Coding (HEVC). A detailed overview of SVC can be found in [15]. In [16], Scalable video streaming over fading wireless channels is presented. The authors use PSNR (Peak Signal to Noise Ratio) as a measure of video quality and develop a model to characterize the relationship between the rate and PSNR. The Discrete Cosine Transforms (DCT) based SVC is presented in [17]. The presented approach is applicable to different SNR scalable encoders.

DCT has a strong "energy compaction" feature which means the most signal information tends to be concentrated in a few low-frequency components of the DCT. By increasing the type of DCT, the signal information will be concentrated in fewer low-frequency components of the DCT, which results in more compression. The term "scalability" in this article refers to use DCTs (types: I, II, III, IV…) ([18-19]) for desired level of video compression and consequently accessing to quality and bit rate scalability. It should be mentioned that we use DCT transform in order to maintain acceptable visual quality of the very low bit-rate video transmission.

We use the 3D-DCT SVC in our analysis for a number of reasons. First, the generated embedded bit-stream is highly scalable. Depending on the estimated channel condition, the bit allocation scheme can easily extract varying subsets of the bit-stream represented the video of varying quality. Second, because the codec is DCT based, it can provide low bit rate video stream.

2.2 Vector quantization (VQ)

VQ has been widely used in very low bit rate video transmission because of its simple decoder structure and better rate distortion performance over traditional scalar quantization scheme such as Differential Pulse Code Modulation (DPCM) transform coding [20]. When a set of values is quantized jointly as a single vector, the process is known as VQ. Instead of transmitting a given data vector, a symbol which indicates the appropriate reproduction vector is used. This can result in considerable savings in transmission bandwidth, despite the expense of some distortion. A vector quantizer Q of dimension K and size N is a mapping from a point in k-dimensional Euclidean space into a finite set C containing output or reproduction points that exist in the same Euclidean space as the original point. These reproduction points are known as code-words and these set of code words are called a codebook C with N distinct code-words in the set [21]. Mapping function Q is defined as (1):

$$Q: R^k \to C \tag{1}$$

A codebook containing several code-vectors is used to encode the input vectors. The code-word closest to an input vector, i.e. one having the minimum Euclidean distance with respect to the input vector, is used to represent that input vector. Compression is achieved by transmitting the binary address of the selected code-word in the codebook. Guobin Shen in [22] proposed adaptive VQ with codebook updating based on locality and history for images. Its performance can be improved by exploiting inter index redundancies.

2.3 Intra-mode SVC and 3D-DCT SVC

The DCT has been accepted as an essential part of well-known transform block-based video compression standards, such as JPEG, MPEG1-2-4, and ITU-T H.263. Each basis vector in the DCT domain represents a spatial frequency component of the video frame. As mentioned before, the term *scalability* in this paper refers to quality and rate scalability features using DCT types for required level of video compression. The structure of intra-mode SVC is similar to intra-H.264/AVC with this difference that, in intra-mode SVC, DCT types [18-19] are employed as de-correlating transform, e.g., type-IV DCT is applied on video frames to exploit DCT coefficients.

Employing various types of DCT allows generating scalable video stream. Similarly, the 3D-DCT SVC employs DCT types as de-correlating transform.

Suppose each color component of video frame $\{f(i,t); t = 0,1, …; i = Y, C_r, C_b\}$, where t refers to the frame number and i is color component. Also, suppose L level of compression is desired. Each color component of the video sequence is partitioned into $8 \times 8 \times 8$ pixel blocks that are processed in a zigzag scan order. Each block is transformed by the 3D-DCT for L level, e.g., for $L=1$, type-I 3D-DCT and for $L=2$, type-II 3D-DCT are used. For $L=1$, the type-I 3D-DCT is mathematically as follow:

$$F(u,v,w,i) = c(u)c(v)c(w) \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{N} f(x,y,t,i) \times$$
$$cos\left[\frac{(2x+1)u\pi}{2N}\right] \times cos\left[\frac{(2y+1)v\pi}{2N}\right] \times cos\left[\frac{(2t+1)w\pi}{2N}\right] \quad (2)$$

The inverse of 3D-DCT is:

$$f(x,y,t,i) = \sum_{u=1}^{N} \sum_{v=1}^{N} \sum_{w=1}^{N} c(u)c(v)c(w) \times$$
$$F(u,v,w,i) \times cos\left[\frac{(2x+1)u\pi}{2N}\right] \times cos\left[\frac{(2y+1)v\pi}{2N}\right] \times$$
$$cos\left[\frac{(2t+1)w\pi}{2N}\right]$$

$$c(w) = \begin{cases} \frac{1}{\sqrt{2}} & w = 0 \\ 1, & w > 0 \end{cases} \quad c(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1, & u > 0 \end{cases}$$

$$c(v) = \begin{cases} \frac{1}{\sqrt{2}} & v = 0 \\ 1, & v > 0 \end{cases} \quad (3)$$

2.4 Intra-mode VQ-SVC and VQ-3D-DCT SVC

Conventional wireless video encoders such as H.264/AVC and 3D-DCT employ variable-length entropy coding and predictive coding to achieve high compression ratios. But these techniques render the extremely sensitive encoded bit-stream to channel errors. To prevent error propagation, it is necessary to employ various additional error correction techniques [23]. In contrast, alternative technique, VQ which doesn't use variable-length entropy encoding, have the ability to impede such an error [23].

To demonstrate this ability, in this paper, we apply VQ on the coefficients resulting intra-mode SVC and 3D-DCT SVC and compare the performance (PSNR v.s. channel BER) with H.264/AVC and 3D-DCT. The channel conditions and the OFDM parameters are set according to the simulation setting presented in section 5. The results of our simulation are shown in Fig (2) parts a, b, c and d. From the simulation results it can be observed that VQ based encoders outperform both of the mentioned methods which don't employ VQ.

3. Joint source and channel codebook optimization

As mentioned in introduction section, to optimize the overall performance of video transmission in communication systems, it is necessary to consider the optimization of combination of source and channel

characteristics. A traditional approach in combination of source and channel consists in first matching the source codebook to the source statistics in order to minimize the distortion of the source. Then, the labeling of the source dictionary to the channel code-words is optimized in order to minimize the distortion due to the channel errors. In this section, we analyze the codebook optimization for scalable video coding.

Let's suppose that the streamed coefficients of the video-source to be encoded are real-valued stationary and Ergodic process $\{X,; t = 0,1,….; x \in X,\}$ with zero mean and variance σ_x^2 (DCT coefficients have Laplacian-like probability distributions). The source is needed to be encoded by means of VQ. K-dimensional, M-level VQ and a discrete memoryless channel (DMC) is considered. With input and output alphabet $\psi=\{1, 2, …M\}$, we consider the source encoder as operation S (in fact is the mapping of k-th element X^k of the m-th VQ), channel coder as operation C, channel decoder as C_{de} and source decoder as S_{de}.

The source output is $X_s = \{X_s^1, X_s^2, X_s^3, …\}$ which we derive our distortion model expressions for a binary symmetric channel (BSC) with a given bit error probability. Given the bit error probabilities for any channel (AWGN, Rayleigh fading, etc.) and the fact that the probability of making an error from 0 to 1 is the same as that of 1 to 0, the channel can be represented as a BSC. Therefore, the distortion model presented in this paper can be used to find the distortion curves for any channel that can be represented as a BSC, given that the source coding rate and the bit error rate are known. Our distortion model is independent of modulation type.

Let us define input/output variables and parameters of the coding system. In the system, variables X and $\hat{X}$ represent the source and reconstructed signals respectively, also X_s and $\hat{X}_s$ represent the source encoder output and the source decoder input, X_c and $\hat{X}_c$ the channel encoder output and the channel decoder input respectively. For the source encoder, we have $s: X \to X_s^k, \varphi = \{S_1, S_2, …., S_M\}$ which $s(X) = X_s^k$ if $X \in S_i$. Also for the channel encoder we have $C: X_s^k \to X_c^k, \varsigma = \{C_1, C_2, …., C_M\}$ which $C(X_s^k) = X_c^k$ if $X_s^k \in C_i$.

At the decoder side, for the channel decoder we have $C_{de}: \psi \to \hat{X}_c^k, C_d = \{C_{d,1}, C_{d,2}, …, C_{d,M}\}$, which $C_{de}(\hat{X}_c^k) = \hat{X}_c^k$ if $\hat{X}_c^k \in C_{d,i}$ and also for the source decoder we have $S_{de}: \hat{X}_s^k \to \hat{X}^k, S_d = \{S_{d,1}, S_{d,2}, …, S_{d,M}\}$, which $S_{de}(\hat{X}_c^k) = \hat{X}^k$ if $\hat{X}_s^k \in S_{d,i}$.

Let's suppose that the codebook (reproduction alphabet) be $C_{book}=\{C_{b,1}, C_{b,2}, …., C_{b,j}, …, C_{b,M}\}$ and the source encoder is described in term of a partition $\varphi=\{S_1, S_2, …., S_M\}$. The total average distortion of the coding system is denoted by $D_{s+c}(\varphi, C_{book}, C)$ for L scalable layers. For a layer $l \in L$ we have [24]:

$$D_{s+c,l}(\varphi, C_{book}, C) = \frac{1}{k} \sum_{i=1}^{M} \sum_{j=1}^{M} P\left(\hat{X}_{c,j}^k | X_{c,i}^k\right) \int P_l(x) d_l(x, \hat{x}) dx \quad (4)$$

We assume that the distortion caused by representing the source vector X by a reproduction vector (also called a

code-vector) $\hat{X}$ is given by a non-negative distortion measure $d_l(x, \hat{x})$ and $P_l(x)$ is the k-fold probability density function [24], and the channel is DMC with probability of output $\hat{X}_{c,j}^k$ condition to given input $X_{c,i}^k$, $p(\hat{X}_{c,j}^k \mid X_{c,i}^k)$.

In paper [24], the optimum codebook is calculated for a non-scalable source. While in our paper we calculate the optimum codebook for a scalable source. For the L scalable layers the equation (4) becomes:

$$D_{s+c,l}(\varphi, C_{book}, C) = \frac{1}{k}\sum_{l=1}^{L}\sum_{i=1}^{M}\sum_{j=1}^{M} P(\hat{X}_{c,j}^k|X_{c,i}^k) \int P_l(x)d_l(x,\hat{x})dx \quad (5)$$

It is clear that for a fixed C_i the problem of minimizing the total average distortion is identical to the VQ design problem with a modified distortion measure [24]. Therefore for a fixed C_i and a fixed C_{book} the optimum partition $\varphi^* = \{S_1^*, S_2^*, ..., S_M^*\}$ is such that:

$$S_i^* = \{x: \sum_{l=1}^{L}\sum_{j=1}^{M} p(\hat{X}_{c,j}^k \mid X_{c,i}^k)d_l(x,\hat{x}) \leq$$
$$\sum_{l=1}^{L}\sum_{j=1}^{M} p(\hat{X}_{c,j}^k \mid X_{c,h}^k)d_l(x,\hat{x}), \forall h \neq i\} \;\; i \in \psi \quad (6)$$

Where $d_l(x, \hat{x}) = \|x - \hat{x}\|^2$ is the squared error distortion. Similarly, to find the optimum codebook, the total average distortion should be determined for a fixed C_i and a fixed φ, therefore the equation (5) becomes to:

$$D_{s+c,l}(\varphi, C_{book}, C) = \frac{1}{k}\sum_{l=1}^{L}\sum_{j=1}^{M} P(\hat{X}_{c,j}^k|X_{c,i}^k) \int P_l(x)d_l(x,\hat{x})dx \quad (7)$$

The optimum codebook, $C_{book,j}^* = \{C_{b,1}^*, C_{b,2}^*, ..., C_{b,j}^*, ...C_{b,M}^*\}$, must satisfy $C_{book,j}^* =$arg min E$(d(x, \hat{x}) \mid V = \hat{X}_{c,j}^k)$, $j \in \psi$ where $\hat{X}_j^*$ is equal to $C_{b,j}^*$, and V is used to denote the random variable at the channel output [24].

Let's attempt to solve the unconstrained minimization problem. The best set is determined by setting the partial derivatives of the equation (7) with respect to the $\hat{x}$'s equal to zero, i.e.:

$$\frac{\partial D_{s+c}(\varphi, C_{book}, C)}{\partial \hat{x}} = -\frac{2}{k}\sum_{l=1}^{L}\sum_{j=1}^{M} p(\hat{X}_{c,j}^k|X_{c,i}^k) \int (x-\hat{x})p_l(x)dx = 0 \quad (8)$$

Which, in turn, it implies that

$$\hat{x}_j^* = \frac{\sum_{l=1}^{L}\sum_{j=1}^{M} p(\hat{X}_{c,j}^k|X_{c,i}^k) \int x p_l(x)dx}{\sum_{l=1}^{L}\sum_{j=1}^{M} p(\hat{X}_{c,j}^k|X_{c,i}^k) \int p_l(x)dx} \;, j \in \psi \quad (9)$$

Therefore the optimum codebook can be obtained by (9). We generally refer to the algorithm channel-optimized algorithm VQ (COVQ) which is detailed in [24] for optimizing the codebook.

4. Joint source and channel rate optimization

In this section, we focus on the application of JSCC in video communications, to find an optimal bit allocation between source coding and channel coding based on the Shannon's separation theory [8], [25] and [26].

4.1 Operational rate distortion (ORD) theory

Every lossy data compression scheme and every channel coding technique have only a finite set of admissible mode, suppose for the source mode we have M admissible modes $s = \{s_1, s_2, ..., s_M\} \in S^M$ and for the channel coding modes, $c = \{c_1, c_2, ..., c_M\} \in C^N$. Therefore there is only a finite number of possible rate distortion pairs for any given source and channel coding technique, so ORD is based on this fact [27]. For any given source mode and channel coding technique mode and given channel state information (CSI), we will have a set of Rate-Distortion points ($M{\times}N$ points). For each modulation scheme, probability of error of the channel can be estimated using a given CSI [28-29]. Using this estimated probability of error, experimentally ORD points are calculated [8].

4.2 Optimal bit allocation between source and channel coding using Lagrange multiplier

To solve a problem there are four distinctive steps. First, an appropriate system performance evaluation metric should be selected. Second, the constraints are needed to be specified. Third, a model of the relationship between the system performance metric and the set of adaptation parameters are needed to be established. Finally, the best combination of adaptation parameters that maximize the system performance while meeting the required constraints needs to be identified. Therefor we present the formal approach to formulate the joint source and channel coding problem and provide solution approaches to such a problem. Let S, be the set of source coding parameters and C the set of the channel coding parameter. The general formulation of the optimal bit allocation problem is to minimize the total expected distortion, i.e., provide the best video delivery quality, for the frame(s), given the corresponding bit rate constraint.

$$min_{s\in S^M, \; c\in C^M} E[D(s,c)] \; Min \; E[D(s,c)] \quad (10)$$
$$s.t. \; R(s,c) \leq R_0$$

Where $E[D(s,c)]$ is total expected distortion, $R(s,c)$ the total number of bits used for both source and channel coding, and R_0 is the bit rate constraint for the frame(s). For solving constrained optimization problem Equation (11), which is a convex problem the Lagrange multiplier method is suitable. Lagrange method introduces a variable which is called a Lagrange multiplier that controls the weights of the constraint when added as a penalty to the objective function. In this method constrained problem is converted into an unconstrained problem [8], as:

$$\{S^*(\lambda), C^*(\lambda)\} = arg \; min\{E[D(s,c)] + \lambda R(s,c)\}$$
$$s \in S^M, \; c \in C^M \quad (11)$$

Where $\lambda \geq 0$ is the Lagrange multiplier. The solution of Equation (10) can be obtained by solving Equation (11) with the appropriate choice of the Lagrange multiplier while bitrate constraint $R(S^*(\lambda), C^*(\lambda)) \leq R_0$ is satisfied, [8]. In practice, due to the finite set of source and channel

coding parameters, the objective function and the constraints are not continuous thus the constraint may not be met with equality. In this case, the solution obtained by solving Equation (11) will be the convex hull approximation solution to Equation (10). Let U denote the coding parameter set $S \times C$ and let $u \in U$ be an element of this set, a coding parameter including both source and channel coding parameters. Assuming that is $u^*(\lambda)$ an optimal solution to Equation (11), so we have:

$$D\big(u^*(\lambda)\big) + \lambda R\big(u^*(\lambda)\big) \leq D\big(u(\lambda)\big) + \lambda R\big(u(\lambda)\big), \forall u \in U \quad (12)$$

which can be rewritten as:

$$R(u) \geq -\frac{1}{\lambda} D(u) + \left\{ R\big(u^*(\lambda)\big) + \frac{1}{\lambda} D\big(u^*(\lambda)\big) \right\} \quad (13)$$

This means that all rate-distortion pairs $\{R(u), D(u)\}$ must lie on the upper right side in the rate-distortion plane of the line defined by Equation (13) using an equal sign. $\big\{ R\big(u^*(\lambda)\big), D\big(u^*(\lambda)\big) \big\}$ is on the line and the line has slope $-1/\lambda$. For a different value of λ, the line with slope $-1/\lambda$ will meet at least one optimal solution $\big\{ R\big(u^*(\lambda)\big), D\big(u^*(\lambda)\big) \big\}$.

In this paper we deal with a set of ORD points for each layer which is calculated through the experiment. The distortion for layer l is expressed as $D_{s+c,l}(R_{s+c,l})$ and the overall distortion as:

$$D_{s+c} = \sum_{l=1}^{L} D_{s+c,l}(R_{s+c,l}) \quad (14)$$

R_{s+c} is the bit rate used for source and channel coding for the scalable layer l, $R_{s+c,l}$ is equal to:

$$R_{s+c,l} = \frac{R_{s,l}}{R_{c,l}} \quad and \quad R_{s+c} = \sum_{l=1}^{L} R_{s+c,l} \quad (15)$$

Where $R_{s,l}$ and $R_{c,l}$ are the source and channel rates, respectively, for the scalable layer l. For L scalable layers, the total transmitted bitrate is equal to R_{s+c}. Now we expand discussion for the video, as mentioned the optimization problem in video transmission is:

$$Min\ D_{s+c}(max\ PSNR_{s+c})$$
$$s.t.\ R_{s+c} \leq R_{budget} \quad (16)$$

The solution, $\{D^*_{s+c}, R^*_{s+c}\}$ to the constrained optimization problem Equation (10) is obtained by converting it into an unconstrained problem using Lagrangian optimization function:

$$J_l = D_{s+c}(R_{s+c,l}) + \lambda R_{s+c,l}$$
$$l = 1, 2, \ldots, L \quad (17)$$

The unconstrained problem can now be solved using these Lagrangian cost functions, with the optimal solution being the argument of the following unconstrained minimization problem:

$$\underset{R_{s+c,l}(\lambda), D_{s+c,l}(\lambda)}{Min\ (J_1 + J_2 + \cdots J_L)} \quad (18)$$

To solve this problem, universal rate-distortion characteristic (URDC) is proposed [17]. In this method, universal rate-distortion curves are established using simulations at different source coding rates. These curves

are used to establish operational rate distortion curves, which are then used to specify the value of distortion at different source coding and channel bit error rates. Using these operational rate-distortion curves, the optimal bit allocation and the minimum distortion are obtained, as follows:

Given a set of parameters for the channel (e.g., signal-to-noise ratio), channel coding, and the particular modulation scheme, the probability of bit error, P_b, are calculated for the set of channel coding rates, R_c, of interest. This can be performed using simulations or by theoretical means. It constructs a reference as a performance of the channel coding over the especial channel with the given parameters. Moreover, this channel performance analysis needs to be performed offline and only once [17].

Toward calculating the impact of the errors due to both source coding and channel transmission on a set of data, it is perceived that for a given set of layer source rates, the distortion for a particular layer, $D_{s+c,l}$, given a particular source coding rate, $R_{s,l}$, is a function of the bit error rate. Thus, the rate-distortion function of the layer for a fixed source rate, $R_{s,l}$, is a function of the bit error rate (after channel decoding), P_b. It is then possible to plot a family of $D_{s+c,l}$ versus $1/P_b$ curves given a set of source coding rates of interest, so these are defined as the URDCs of the source [17].

5. Simulation setting

In this study, pixel Common Inter-mediate Format (CIF) resolution video sequences Football and Coastguard video (at 15 and 30 frames per second, respectively) [32] are employed as video sources. Next, the 3D-DCT is applied on $8 \times 8 \times 8$ blocks of video, and intra mode H.264/AVC is applied on frames of videos. Similarly, intra-mode VQ-SVC and VQ-3D-DCT SVC are applied on the video streams and based on the prepared optimum codebook calculated by equation (9) the vectors of data are converted to the symbols which indicate the reproduction vectors. Next, the symbols are coded using Rate-Compatible Punctured Convolutional (RCPC) codes [30]. Punctured convolutional codes are convolutional codes obtained by puncturing some outputs of the convolutional encoder.

Since our focus is on JSCC, we assume perfect channel estimation of each sub-carrier gain and perfect suppression of multipath by the guard interval. In this simulation we use the general OFDM system with FFT size 256, occupied subcarriers 151, Ratio of Cyclic prefix time (guard interval) set as 1/8, and modulation type set as QPSK. The parameters of the simulated channel are presented in Table (1). These parameters are according to the ITU Channel Model for Vehicular Test Environment. Motion causes Doppler shift in the received signal components, the Doppler frequency f_D equals $f_0 v/c$ in which v is the mobile speed, f_c is the carrier frequency

and c is the light speed [31]. In our paper carrier frequency is equal f_c=1.9 GHz and the maximum Doppler frequency is equal f_D=90 Hz.

6. Simulation results

In this section, at first, the compression performance of the proposed intra-mode VQ-SVC and VQ-3D-DCT SVC are compared with that of intra-mode H.264/AVC and 3D-DCT. It is important to note that the results are not presented with a view to establishing which H.264/AVC provides better source compression. Rather, they are intended to show that the proposed VQ based codecs provide compression performance comparable to these other related codecs.

Fig. 1 illustrates the RD curves for two sequences. In Fig. 1(a), for Coastguard, it is observed that proposed intra-mode VQ-SVC has an RD performance similar to H.264/AVC. It is also obvious that at 540-820 Kbps, the proposed intra-mode VQ-SVC and H.264/AVC provide the same PSNR performance. Obviously, at 590-800 Kbps, VQ-3D-DCT SVC and 3D-DCT provide the same PSNR performance. As shown in Fig. 1(b), for Football, the proposed intra-mode VQ-SVC has an RD performance similar to H.264/AVC. In this figure, it can be observed that at 675-930 Kbps, intra-mode VQ-SVC and H.264/AVC result in the same PSNR performance. It is also clear that at 580-1050 Kbps, VQ-3D-DCT SVC and 3D-DCT provide the same PSNR. In Fig. 1(a) and (b), it is also obvious that intra-mode VQ-SVC and H.264/AVC outperform VQ-3D-DCT SVC and 3D-DCT.

As described in section 4, there is an optimal bit allocation in a communication scenario which is defined by the JSCC. As an example, Fig. 3 is presented to show the impact of the optimal bit allocation on the quality of video received by the end user. In the example, Football video is encoded by intra mode VQ-SVC and transmitted over wireless channel (at SNR=15, Mobile speed=50 km/h) in 5 different predefined modes; Mode-1{source code rate (SCR)=380 Kbps, channel code rate (CCR)=0.33}, Mode-2 {SCR=617 Kbps, CCR=0.5}, Mode-3 {SCR=807 Kbps, CCR=0.66}, Mode-4 {SCR=950 Kbps, CCR=0.75} and Mode-5 {SCR=998 Kbps, CCR=0.8}. From the Fig. 3 it can be observed that the received frame-74 in part *d* resulted by setting the Mode-3 as the bit allocation mode, have the highest quality.

Similarly, two simulations are performed for the Football video frames 1-100. The first simulation is performed in the different transmission environments SNR {SNR = 5, 10, and 15 dB} and the second, in the different mobile speeds {V = 30, 50 and 90 km/h}. Related results are presented in Table (2) and Table (3) respectively. The results obtained using repeated experiments and taking the average PSNR. As reported in Table (2), Average PSNR at SNR=15 at the Mode-3, compared with other modes, is maximum (PSNR=24.47 dB). It can be also observed that the mode is changed by reducing the SNR. For instance, at SNR=10 the optimum SCR is equal 617 Kbps and CCR is equal 0.5. Also, at SNR=5 the optimum SCR is equal 380 Kbps and CCR is equal 0.33. In the second simulation, quality of video received by the end user is evaluated at different mobile speeds at SNR=15 dB. As reported in Table (3), Average PSNR at mobile speed (V) 90 km/h at the Mode-1, compared with other modes, is maximum (PSNR=19.39 dB). It can be also observed that the mode is changed by reducing the mobile speed. For instance, at V=50, the optimum SCR is equal 617 Kbps and CCR is equal 0.5 and also at V=30, the optimum SCR is equal 807 Kbps and CCR is equal 0.66 and PSNR is equal to 27.76 dB.

Next, to evaluate the performance of the proposed codecs and schemes over wireless channel, each video stream is compressed at the same bitrate and all utilize the same channel coding rate, i.e. 3/4 rate FEC. As observed in Fig.1, in 800 Kbps, for both sequences, the proposed intra-mode VQ-SVC and H.264/AVC provide the same PSNR (for Coastguard PSNR=39 dB and for Football PSNR=35 dB) and also VQ-3D-DCT SVC and 3D-DCT result in the same PSNR (for Coastguard PSNR=36.3 dB and for Football PSNR=31 dB). Therefore, in this section, Coastguard and Football are encoded at 800 Kbps.

In Fig. 2, parts *a* and *b*, the intra mode VQ-SVC is compared with the intra mode H.264/AVC over Coastguard and Football video respectively, and in Fig. 2, parts *c* and *d*, the VQ-3D-DCT SVC is compared with the 3D-DCT over Coastguard and Football video respectively. The simulation results show that VQ based encoders outperforms both of the mentioned methods which don't employ VQ. For example at 10^{-3}-10^{-2} BER range, intra mode VQ-SVC offer about 3.5 dB PSNR gain for Coastguard and 3.8 dB PSNR gain for Football compared to the best performing intra mode H.264/AVC stream. As well as, at 10^{-3}-10^{-2} BER range the VQ-3D-DCT SVC offer about 2.7 dB PSNR gain for Coastguard and 3 dB PSNR gain for Football compared to the best performing 3D-DCT stream.
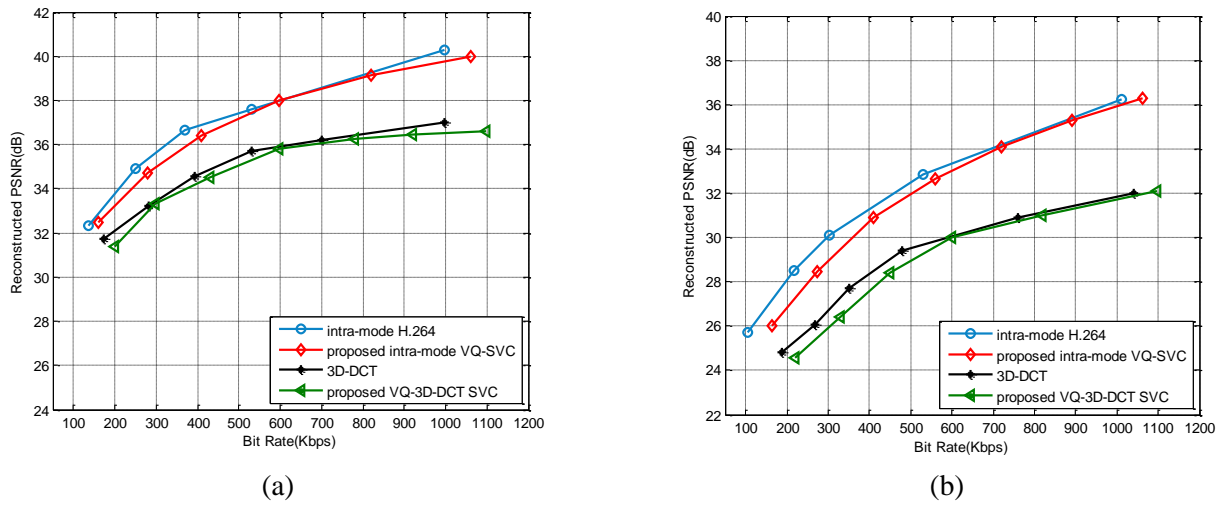
(a)

(b)

Fig. 1. RD comparisons; intra-mode H.264, proposed intra-mode VQ-SVC, 3D-DCT and proposed VQ-3D-DCT SVC. (a) Coastguard sequence. (b) Football sequence.
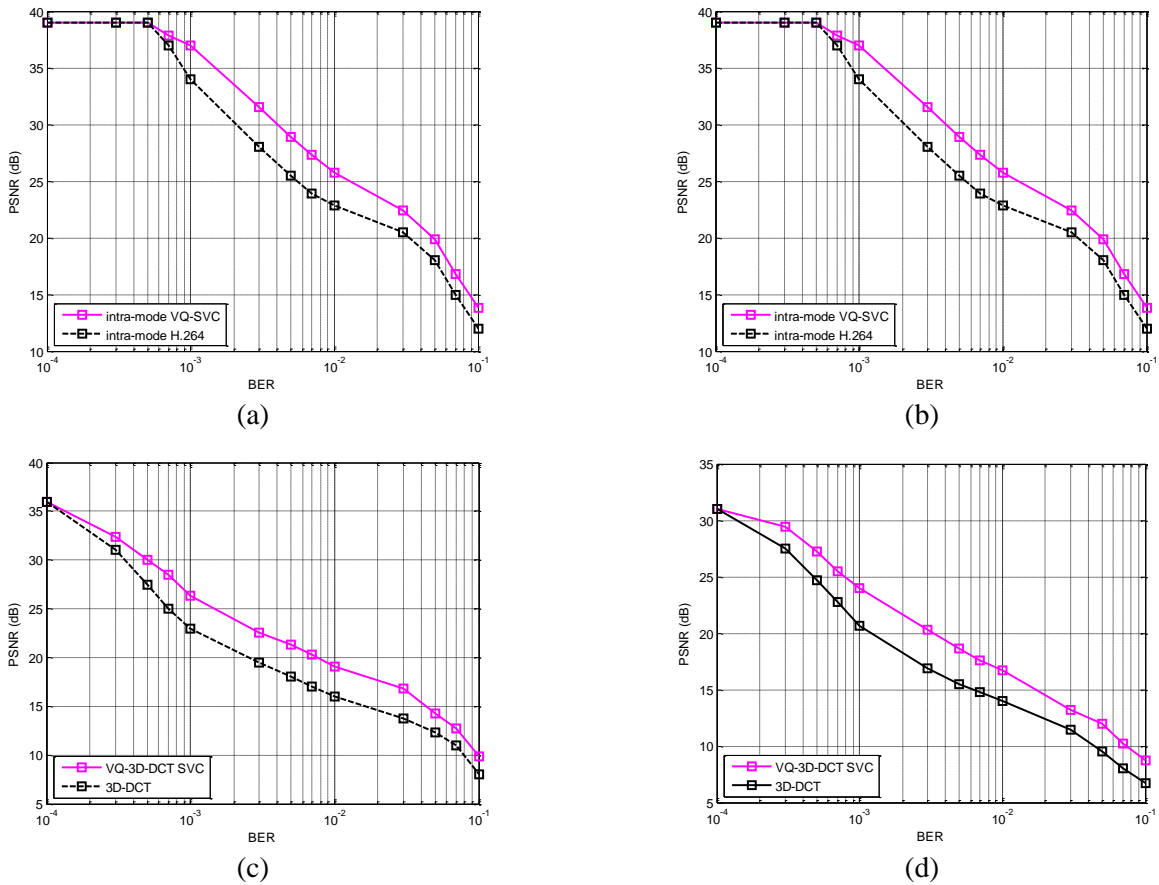

(a)

(b)

(c)

(d)

Fig. 2. Simulation results of the proposed scheme, intra mode VQ-SVC, compared to intra mode H.264: (a) for Coastguard video (b) for Football video. Results of the proposed scheme, 3D-DCT VQ-SVC, compared to 3D-DCT: (c) for Coastguard video (d) for Football video.
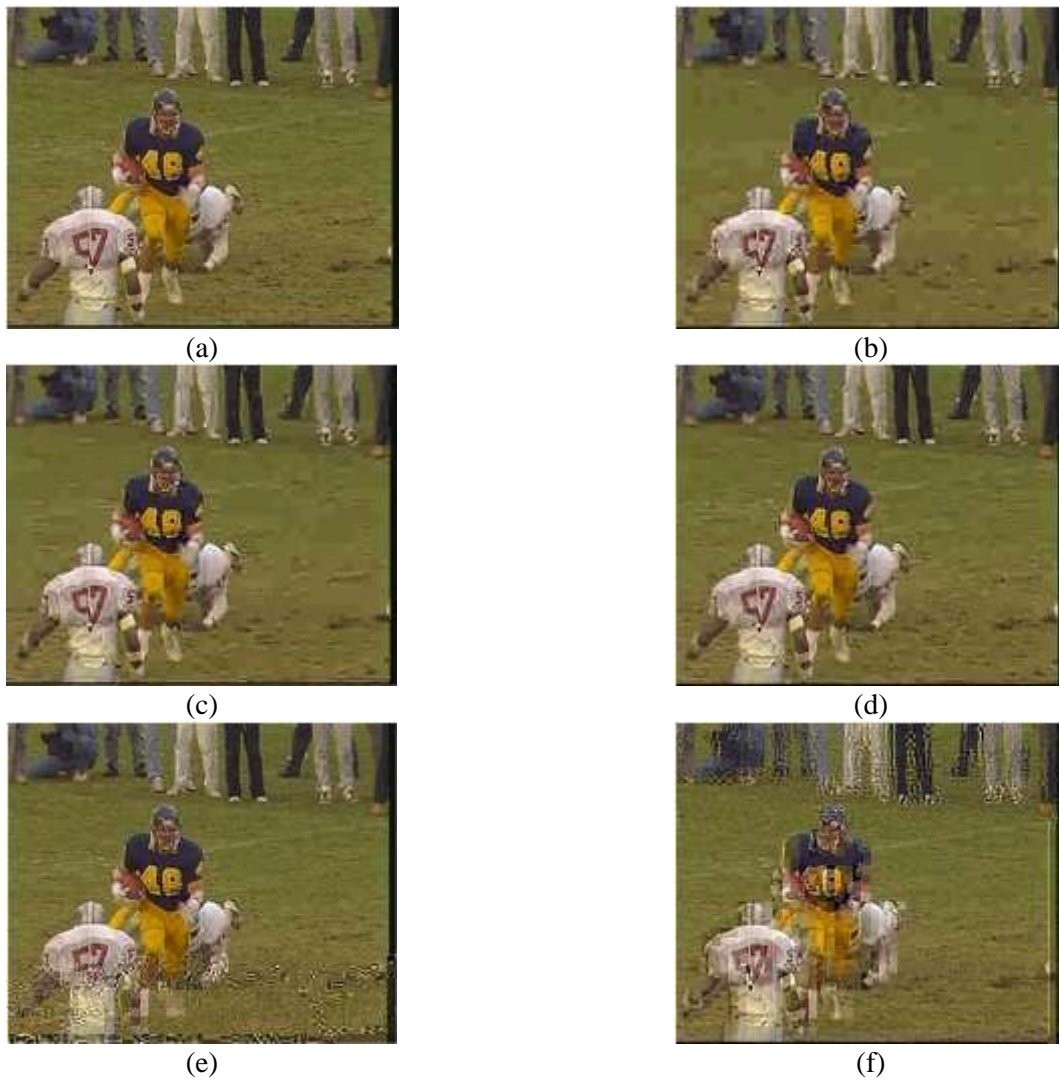
Fig. 3. Simulation Results for channel condition including: SNR=15, Mobile speed= 50 km/h: (a) is the original frame of the Football video, frame number 74, (b) is the received frame in Mode 1, (c) is the received frame in Mode 2, (d) is the received frame in Mode 3, (e) is the received frame in Mode 4, (f) is the received frame in Mode 5.

Table 1. Channel Parameters and channel coding setting

Parameters	Values
Channel model	ITU channel model for vehicular test environment
Relative Tap delay (ns)	0, 310, 710, 1090, 1730 and 2510
Average power (dB)	0.0, -1.0, -9.0 -10.0, -15.0 and -20.0
Doppler frequency (Hz)	90
Channel coding	RCPC [a]; Rates: 1/3, 1/2, 2/3 and ¾

[a] RCPC: Rate-Compatible Punctured Convolutional

Table 2. JSCC simulation results of Football video frames 1-100 for different SNR = 15, 10 and 5 dB

Mode \\ Result	Mode 1: SCR [a] = 380 Kbps CCR [b] = 0.33	Mode 2: SCR = 617 Kbps CCR = 0.5	Mode 3: SCR = 807 Kbps CCR = 0.66	Mode 4: SCR = 950 Kbps CCR = 0.75	Mode 5: SCR = 998 Kbps CCR = 0.8
Average PSNR(dB) at SNR=15	20.39	21.58	24.47	20.81	20.15
Average PSNR(dB) at SNR=10	18.68	20.12	19.72	18.03	17.70
Average PSNR(dB) at SNR=5	17.11	15.89	14.08	12.75	12.03

[a] SCR: source code rate, b CCR: channel code rate

Table 3. JSCC simulation results of Football video frames 1-100 for different mobile speed (v): 90, 50 and 30 km/h

Mode Result	Mode 1: SCR [a] = 380 Kbps CCR [b] = 0.33	Mode 2: SCR = 617 Kbps CCR = 0.5	Mode 3: SCR = 807 Kbps CCR = 0.66	Mode 4: SCR = 950 Kbps CCR = 0.75	Mode 5: SCR = 998 Kbps CCR = 0.8
Average PSNR(dB) at v=90	19.39	17.97	16.38	14.56	14.01
Average PSNR(dB) at v=50	22.09	25.88	23.77	21.81	20.78
Average PSNR(dB) at v=30	23.87	26.20	27.76	26.01	24.43

[a] SCR: source code rate, b CCR: channel code rate

Next, we compare the performance of three proposed schemes; intra mode VQ-SVC with intra mode VQ-SVC using JSCC and with intra mode VQ-SVC using both JSCC and optimum codebook. We evaluate these schemes over Coastguard and Football video. The results are shown in Fig. 4, parts *a* and *b*. From our simulation it can be observed that intra mode VQ-SVC using both JSCC and optimum codebook outperforms the other two schemes; the intra mode VQ-SVC and the intra mode VQ-SVC using JSCC. For example as shown in Fig. 4 part *a*, at 10^{-3}-10^{-1} BER range, intra mode VQ-SVC using both JSCC and optimum codebook offers about 2.5-5dB PSNR gain for Coastguard, also as shown in Fig. 4 part *b*, it offers about 3-5.5 dB PSNR gain for the Football video compared to the intra mode VQ-SVC. Next, we compare the performance of three schemes; 3D-DCT VQ-SVC with 3D-DCT VQ-SVC using JSCC and with 3D-DCT VQ-SVC using both JSCC and optimum codebook. The results are shown in Fig. 4 *c* and *d*. From the simulation it can be observed that 3D-DCT VQ-SVC using both JSCC and optimum codebook outperforms the other two schemes; the 3D-DCT VQ-SVC and the 3D-DCT VQ-SVC using JSCC. For example, as shown in Fig. 4 part *c*, at 10^{-3}-10^{-1} BER range 3D-DCT VQ-SVC using both JSCC and optimum codebook offers about 3 dB PSNR gain for Coastguard and about 2-5 dB PSNR gain for

Football, as shown in Fig. 4 part *d*, compared to the 3D-DCT VQ-SVC.

7. Conclusions

In some wireless streaming video applications, like UGV and GCS system, it is required to provide very low bit rate SVC. We demonstrate that intra-mode VQ-SVC and VQ-3D-DCT SVC outperform intra-mode H.264 and 3D-DCT SVC, respectively. To provide high quality and robust wireless video communication, it is required to employ JSCC technique. We use this technique in two objectives, first to find the optimum codebook for VQ and next to find the optimum channel and source code rates. Since the optimum codebook is not yet calculated for VQ-SVC, in this paper we analyze and apply it on the VQ-SVC. Then we develop the analysis for video transmission over multipath Rayleigh fading channel. The simulation results show that employing both optimal codebook and JSCC for VQ-SVC leads to higher-quality video delivery.
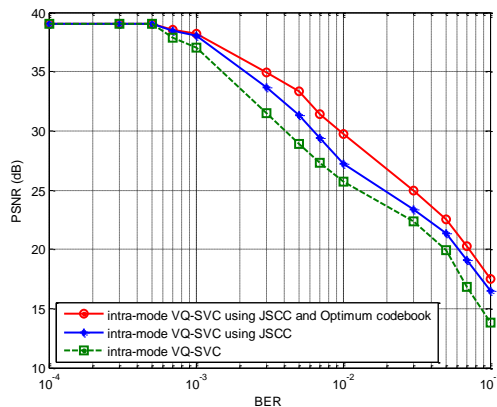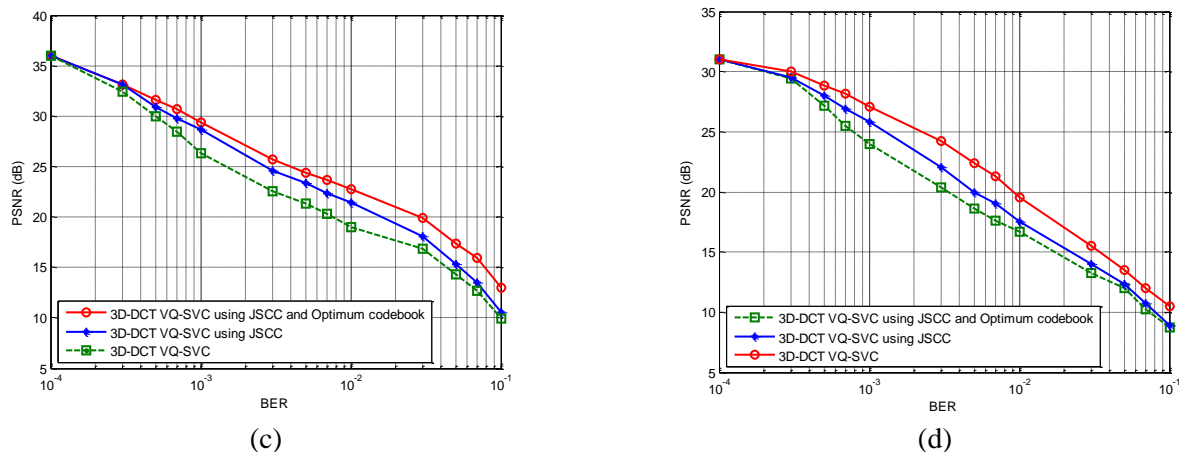
(a)
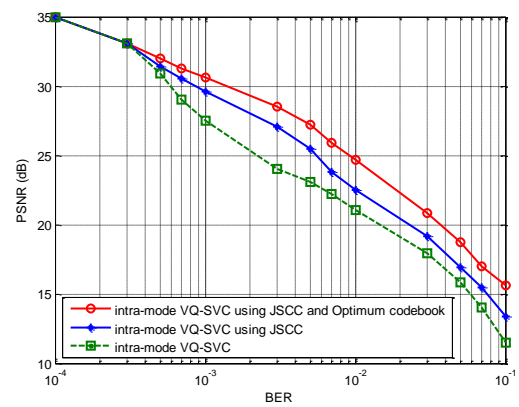
(b)

Fig. 4. Simulation Results of three proposed schemes; intra mode VQ-SVC; intra mode VQ-SVC using JSCC; intra mode VQ-SVC using both JSCC and optimum codebook: (a) for Coastguard video (b) for Football video. Results of three proposed schemes 3D-DCT VQ-SVC; 3D-DCT VQ-SVC using JSCC; 3D-DCT VQ-SVC using both JSCC and optimum codebook: (c) for Coastguard (d) for Football video.

References

[1] H.Wang, L.Kondi, A. Luthra and S.Ci, 4G wireless video communications, New York: Wiley 2009.

[2] S.Saponara, "Real-time and low-power processing of 3D direct/inverse discrete cosine transform for low-complexity video codec", Journal of Real-Time Image Processing Vol.7, No.1, pp.43-53, 2012.

[3] E.Dahlman, S.Parkvall, J.Skold, 4G: LTE/LTE-advanced for mobile broadband, Academic Press, 2011.

[4] A.S.Akbari, P.Bagheri Zadeh, T.Buggy, J.Soraghan, "Multiresolution perceptual and vector quantization based video codec", Multimedia Tools and Applications Vol.58, No.3, pp.569-583, 2012.

[5] R. Krishnamoorthy and R. Punidha, "Low bit-rate multi stage vector quantization based on energy clustered training set", Multimeda Tools and Applications, pp.1-16, 2012.

[6] N.Adami, A.Boschetti, R.Leonardi and P. Migliorati "Embedded indexing in scalable video coding", Multimeda Tools and Applications, pp.105-121, 2010.

[7] A. E.Gamal, Y. H. Kim, Network information theory, Cambridge University Press, pp.336-360, 2012.

[8] F.Zhai, A.Katsaggelos, "Joint source-channel video transmission", Synthesis Lectures on Image, Video, and Multimedia Processing, Vol.3, No. 1, pp.1-136, 2007.

[9] D.Ersson, J.Kron, M.Skoglund, E.G.Larsson, "Joint source-channel coding for the MIMO broadcast channel" IEEE Trans. on Signal Processing, Vol. 60, No.4, pp.2085-2090, 2012.

[10] C.Blumm, C.Heller, R.Weigel, "SDR OFDM waveform design for a UGV/UAV communication scenario" Journal of Signal Processing Systems, pp.11-21, 2012.

[11] H.Sun, A.Vetro, J.Xin, "An overview of scalable video streaming", Wireless Communications and Mobile Computing Vol.7, No.2, pp.172-159, 2007.

[12] ITU-T Recommendation, H.264 advanced video coding for generic audiovisual services, ISO/IEC 14496, 2003.

[13] T.Wiegand, G.J.Sullivan, G.Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Trans. Circuits Syst. Video Technol. , vol. 13, no. 7, pp. 560576, Jul. 2003.

[14] H.Choi, J.Nam, D.Sim, and I.V.Bajic, "Scalable video coding based on high efficiency video coding (HEVC)", IEEE Conf. Communications Computers and Signal Processing, pp. 346-351, 2011.

[15] H.Schwarz, D.Marpe, T.Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard", IEEE Trans. Circuits and Systems for Video Technology, Vol.17, No.9, pp.1103-1120, 2007.

[16] H.Zhang, Y.Zheng, M.A.Khojastepour, S.Rangarajan, "Scalable video streaming over fading wireless channels", IEEE Conf. Wireless Communications and Networking, pp.1-6, 2009.

[17] L.P. Kondi, F.Ishtiaq, A.K.Katsaggelos, "Joint source-channel coding for motion compensated DCT based SNR scalable video", IEEE Trans. image processing, Vol.11, No.9, pp.1043-1052, 2002.

[18] [Online] Available: http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Discrete_cosine_transform

[19] F.Matteo, S.G.Johnson, "FFTW library", 2009, [Online] Available: http://www.fftw.org, a free (GPL) C library that can compute fast DCTs (types I-IV) in one or more dimensions, of arbitrary size.

[20] J.E.Fowler, K.C.Adkins, S.B.Bibyk, S.C.Ahalt, "Real-Time video compression using differential vector quantization," IEEE trans. of circuits and systems for video technology, vol.5, No.1, pp.14-24, 1995.

[21] R. Gray, Vector quantization, IEEE ASSP Magazine, Vol.1, No.2, pp.4-29, 1984.

[22] G.Shen, B.Z.Ming, L.Liou, "Adaptive vector quantization with codebook updating based on locality and history," IEEE Trans. on image processing, Vol. 12, No. 3, pp. 283-295, 2003.

[23] S.M.M.Bokhari, A.R.Nix, D.R.Bull, "Rate distortion optimized video transmission using pyramid vector quantization" IEEE Trans. Image Processing, Vol.21, No.8, pp.3560-3572, 2012.

[24] N.Farvardin, V.Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers", IEEE Trans. Information Theory, Vol.37, No.1, pp.155-159, 1991.

[25] D.Persson, J.Kron, M.Skoglund, E.G. Larsson, "Joint source-channel coding for the MIMO broadcast channel", IEEE Trans. Signal Processing, Vol.60, No.4, pp.2085-2090, 2012.

[26] C.Guillemot, P.Christ, "Joint source-channel coding as an element of a QoS framework for '4G' wireless multimedia", Computer Communications Vol.27, No.8, pp.762-779, 2004.

[27] M.Bystrom, J.W.Modestino, "Combined source-channel coding schemes for video transmission over an additive white Gaussian noise channel", IEEE Journal on Selected Areas in Communications, Vol.18, No.6, pp. 880–890, 2000.

[28] L.Rugini, P.Banelli, "BER of OFDM systems impaired by carrier frequency offset in multipath fading channels", IEEE Trans. Wireless Communications, Vol.4, No.5, pp.2288-2279, 2005.

[29] S.T.Chung, A.J.Goldsmith, "Degrees of freedom in adaptive modulation: a unified view", IEEE Trans. Communications, Vol.49, No.9, pp.1561-1571, 2001.

[30] J.C.Moreira, P.G.Farrell, Essential of error control coding, Wiley: 2006.

[31] T.David, P.Viswanath, Fundamentals of wireless communication, Cambridge university press: 2005.

[32] Available http://media.xiph.org/video/derf/

Farid Jafarian received his B.S. (with highest honors) degree from Isfahan's Sepahan Institute of Science and Technology, Isfahan, Iran and M.S. degree from University of Birjand, Birjand, Iran, in 2008 and 2012 respectively, all in Electrical Engineering. He was with the Research Institute for Information and Communication Technology (ICT), Tehran, Iran. He is currently senior lecturer in Bonyan University. His research interest and contributions are in the areas of communication theory, information theory and their applications to wireless communications (e.g., multiuser MIMO, cognitive radio and cooperative communication), joint source and channel coding (JSCC) and communications signal processing with a focus on video and speech coding. He is also interested in applications of information theory in Bioinformatics with a focus on genome modeling

Hasan Farsi received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 1992 and 1995, respectively. Since 2000, he started his Ph.D in the Centre of Communications Systems Research (CCSR), University of Surrey, Guildford, UK, and received the Ph.D degree in 2004. He is interested in speech, image and video processing on wireless communications. Now, he works as associate professor in communication engineering in department of Electrical and Computer Eng., university of Birjand, Birjand, IRAN. His Email is: hfarsi@birjand.ac.ir