Speech Encryption Based on Wavelet Transformation and Chaotic Map

Dr. Hala B.Abdul Wahab
Computer Sciences Department, University of Technology/Baghdad.
Sundus I. Mahdi
Computer Sciences Department, University of Technology/Baghdad.

Received on:10/8/2015 & Accepted on:20/1/2016

ABSTRACT

In this paper a new algorithm is presented for speech encryption. It includes dividing the speech signal into overlapped blocks then shuffling those blocks in the time domain. A second permutation is done for the coefficients of the block which generated from wavelet transform by using chaotic key based on Hénon map, and partially encrypting the shuffled online speech signal in a transform domain. The security for the proposed system will depend on different parameters, including secret block sizes, the amount for overlapping along the x-axes and y-axes, permutation key and dynamic encrypted key. By having a new level of security the eavesdropper has to choose the amount of overlapping correctly. Many online speech signal tests demonstrate the validity of the proposed algorithm. The results show that it provides secure approach to real-time speech encryption and at the same time gives high intelligibility for the recovered speech.

Keywords: speech encryption, chaotic map, block shuffling, residual intelligibility, transform domain

INTRODUCTION

The rise in networking and multimedia technology, has increased the need for more security. Voice based communication has been widely used in e-learning, e-banking, teleconferencing, and in the stock market. It is important to provide a high level of security to keep privacy and protect data before transmission [1]. Encryption is the key to securing transmission over vulnerable mediums. The requirements to perform the security needs of digital speech signals have led to the development of good encryption techniques [2]. Traditional symmetric-key algorithms such as the Data Encryption Standard (DES), and Advanced Encryption Standard (AES) "use fixed block sizes with complex permutations and substitutions to give secure output cipher-texts" [3]. The complexity of these algorithms use more resources and thus cannot be applied in real-time due to slow responsiveness. This makes them unsuitable for application on large quantities in real-time, not to mention their high sensitivity to noise [4]. These flaws can be addressed with chaos-based encryption. These techniques provide the needed speed, complexity and security. Chaotic maps are ingrained with properties including mixing, diffusion, and sensitivity (to system parameters and initial conditions). Chaos-based encryption prosperities are similar to the properties of ideal ciphers as confusion, diffusion, and balance and avalanche property. All these benefits have placed chaos-based algorithms for encryption at the top of the interest-radar of those working in the field [5, 6].

M. Ashtiyani [7] in their proposed speech signal encryption use chaotic symmetric cryptography, based on a combination of scrambling and confusion. Chaotic cat map is used for scrambling the speech signal samples and chaos is used for S-box design. Hence, the main contribution of this work is using chaos in both signal diffusion and confusion. Another proposal presented by Zhenjun Tang [8], exploited skew tent map and Arnold transformation to generate dynamic mask matrices, then conducts exclusive "OR" operations between

corresponding elements of each block and a random mask matrix. Mosa, et al., [9] introduced a speech cryptosystem based on permuting and masking speech segment using secret keys in the multi-domain.

Even though encryption algorithms exist, there are restrictions in terms of applied practicality. One example is related to speed, where-by real-time systems necessitate high encryption/decryption speeds while maintaining good security. A new permutation matrix [10], which is generated by two chaotic maps used to diffuse the pixels employs the 2D Hénon chaotic map to perform XOR operation to confuse the grey-level of pixel values in the image. The experiment proves that the key space is large enough and gets ideal security intensity after several iterations.

What is being proposed is an algorithm that sections an inputted speech signal into overlapping blocks after converting a 1 dimensional online speech signal to 2 dimensions. A pseudo-random generator key is used to shuffle the blocks in the time domain first. A chaotic map is employed to generate a chaotic key in order to permute the transformed coefficients for each block. Subsequently the results of each block are partially encrypted in the frequency domain. Many online speech signal tests are used to demonstrate the validity of the proposed algorithm. The results show that it provides an efficient and secure approach to real-time speech encryption. This paper is organized as follows:

Section 1: Introduction.
Section 2: Hénon chaotic map.
Section 3: Wavelet Transform.
Section 3: The proposed algorithm.
Section 4: Experimental Results.
Section 5: Conclusion and Discussion.

HÉNON CHAOTIC MAP

Hénon map, a two-dimensional discrete-time nonlinear dynamical system [11], represented by the state equations which defined as:

$$x_{n+1} = 1 - ax_n^2 + y_n \quad (1)$$

$$y_{n+1} = bx_n \qquad n = 1,2,\ldots \quad (2)$$

Hénon maps redefine any point on the plane (x_n, y_n) as a new point (x_{n+1}, y_{n+1}). Hénon is chaotic in this range, where $a = 1.4$, $b = 0.3$.

Wavelet Transform

The Wavelet transform is a set of tools for converting an inputted signal from the time into the transform domain, to find transform coefficients [14]. The process of encryption is reached through a permutation or XOR of these coefficients. The obtained encrypted coefficients are then converted again into the time domain for transmission. The Haar Wavelet is used to remove any possibility of intelligible pieces which may remain from time domain permutation for the blocks.

The Proposed Algorithm

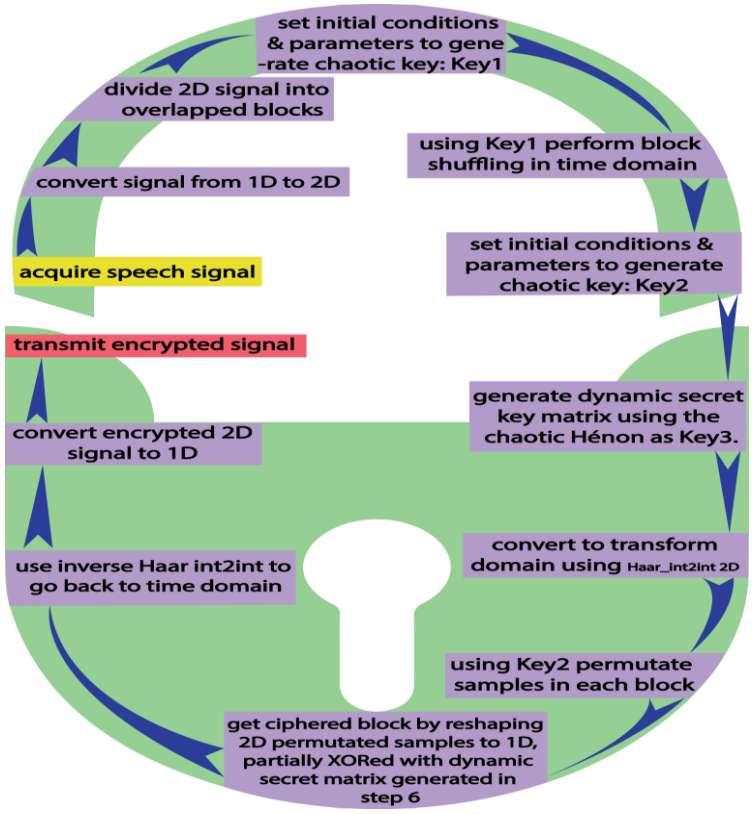
The basic idea of the proposed algorithm is to convert the speech signal from a 1- to a 2-dimensional square matrix that is divided it into square blocks. To make the division secure, overlapping is used where the overlapping portion is between neighboring blocks along the x- and y-axes respectively. A pseudo number generator is used to create a control shuffling key in order to shuffle the overlapped blocks. This achieves a disordered speech signal. This process can enhance the resistance of a ciphered signal against a plain attack. The shuffled speech signal is transformed to a transform domain and the coefficient of the transformed domain is

permutated again by using chaotic key. Partial encryption with a secret dynamic key generated from chaotic Hénon map is applied to obtain ciphered speech. Dynamic encryption keys are generated for each block to enhance the security and robustness of the encryption algorithm with confusion-diffusion structure, and to resist a differential attack. The algorithm summarized in steps.

Input: original speech signal from sound card.

Output: ciphered speech signal

- Step 1: convert input speech signal from 1D into 2D, resize it into a square, and let signal (N, N) be the origin speech, where N are the number of rows which is equal to the number of columns.
- Step 2: divide the 2D signal (N, N) into overlapped blocks of size (b × b) samples and specify the overlapping size. Then calculate the number of blocks.
- Step 3: generate a chaotic control key as Key1 for shuffling the overlapped blocks, ensure no neighboring blocks are allowed in shuffling.
- Step 4: perform block shuffling using Key1 in time domain.
- Step 5: generate a chaotic key as Key2 for permutation.
- Step 6: generate a dynamic secret key matrix by using the chaotic Hénon as Key3.
- Step 7: transform the shuffled block using Haar_int2int 2D into transform domain.
- Step 8: permutate the new coefficients for each block using Key2.
- Step 9: reshape the 2D permutated samples into 1D, partially XORed with the dynamic secret matrix generated in step 6 to get ciphered block.
- Step 10: use the inverse for Haar int2int to go back to time domain.
- Step 11: convert the encrypted signal to 1D.
- Step 12: transmit the encrypted speech
- signal.



Figure(1). Structure of the proposed encryption algorithm

Block Division

Initially, the 2D speech signal is divided into overlapped blocks of the size (b × b). The strategy of overlapping blocks is used to make the algorithm more robust by taking a_x and a_y as the size for overlapping along the x-axes and y-axes correspondingly. The division design is done through setting the block coordinates and counting the number of blocks as n_x, n_y along the x-axes and y-axes correspondingly. Therefore, the total number of blocks (n_{blocks}) is $n_x \times n_y$.
Suppose that the speech signal has (N, N) dimensions and the block side is b. If an integral multiple of blocks can completely cover the signal column i.e. $\mathrm{mod}(N - a_x, b - a_x) = 0$ $n_x = \frac{(N-a_x)}{(b-a_x)}$. If the signal columns can't be covered by integral multiple block, then $n_x = \left\lfloor \frac{(N-a_x)}{(b-a_x)} \right\rfloor + 1$ and padding zeros to the rest of the last blocks.

Block Shuffling Process

In block division, the speech signal is divided into a chunk of overlapped blocks and these blocks are shuffled based on a chaotic key. Better encryption results can be obtained with blocks of smaller size.

$$\begin{bmatrix} b(1,1) & b(1,2) & \dots & b(1,n_y) \\ b(2,1) & b(2,2) & \dots & b(2,n_y) \\ \vdots & \vdots & & \vdots \\ b(n_x,1) & b(n_x,2) & \dots & b(n_x,n_y) \end{bmatrix}$$

Where $b(n_x,n_y)$ denotes a block of speech values. The speech signal is a collections of blocks, block (n) (n=1, 2, n_{blocks}) which is indexed from left-top to the right-bottom.

$$[block(1) \quad block(2) \quad block(n_{blocks})]$$

To make the encryption more confusing and complex, the block shuffling process is put forward and controlled by a chaotic key and the samples are permutated too. Shuffling the overlapped blocks and permutating the samples for each block will distort the speech time envelope.
Small changes in initial conditions can result in major changes in output and this makes chaotic systems attractive to pseudo-random number generators [12]. It is impossible to predict the behavior of the chaotic system even if given partial knowledge of its organization [13].

Sample Permutation Process

The wavelet Haar_int2int 2D is used to transform the shuffled block into a transform domain, the coefficients are permutated using the following procedure:

After specifying initial Hénon map values, the iterated map is used to generate real values representing the chaos sequence. The chaos sequence is sorted in ascending order. The resulting position row is used as a permutation key.

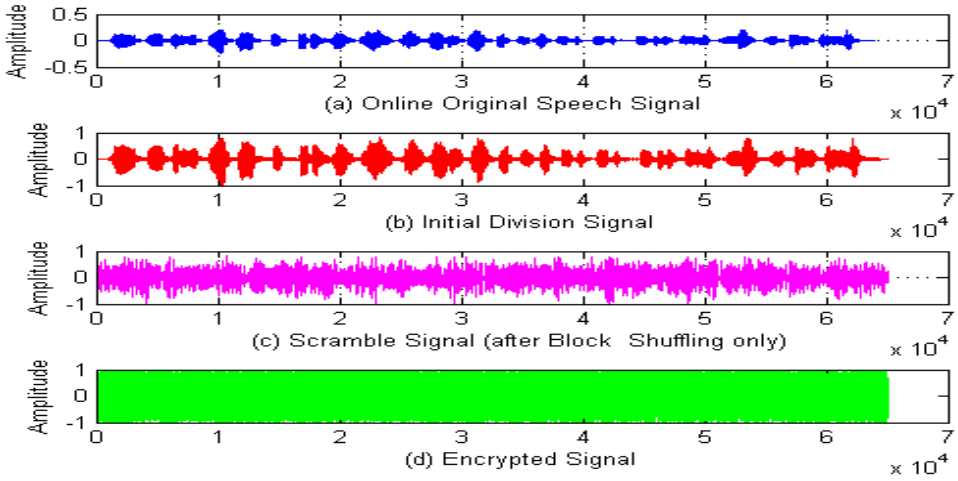
Sample Masking Process

Each sample in the blocks is substituted by XOR with mask key, the mask key is generated by the following procedure:

Set x_0, y_0 and the parameters (a, b) according to equations (1) and (2) to obtain two chaotic sequences $[Z_1(k)]$ and $[Z_2(k)]$. Then change them into integers to get the matrix key as follows:

$$Z_q(k)' = \lfloor abs(Z_q(k) * 2^{15} - 1) \rfloor \quad k = 1,2,\dots n_{blocks}, \qquad q = 1,2$$
$$P(i,j) = Z_1(i,j)' + Z_2(i,j)' \ mod\ 2^{15} - 1\,, \qquad i = 1,2,\dots m \quad j = 1,2,\dots n$$
$$C(i,j) = P(i,j) \oplus Block(i,j)\,, \quad i = 1,2,\dots m \quad j = 1,2,\dots n$$

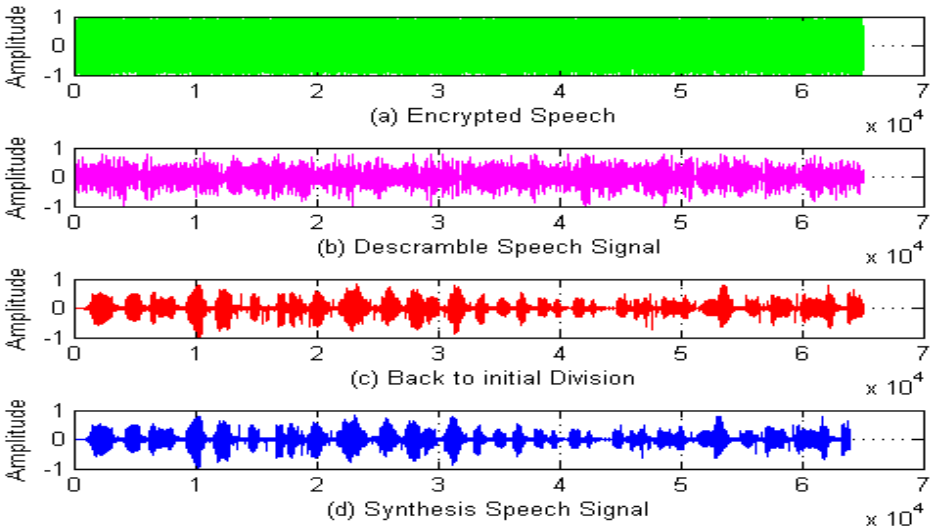
Encryption steps are expressed in Fig. 2.



Figure(2). The online speech signal: (a) Online Original Speech, (b) initial division, (c) Scramble the signal after block shuffling only, (d) Encrypted Signal

Decryption

Steps from the previous operations are applied in an inverse order to the encrypted signal. The continuous signal is then reshaped into a 2D matrix, then divided into overlapped blocks. The blocks are decrypted in the transform domain using dynamic Key3 followed by descrambling the samples for each block in the same domain using the inverse of Key2. A second inverse permutation is applied for the blocks using inverse Key1. Finally the padding is discarded from the last blocks on the right and at



Figure(3). The online speech signal: (a) Encrypted Speech, (b) descrambled speech signal, (c) return to initial division, (d) decrypted speech signal

the bottom. Fig. 3, expresses the decryption process.

EXPERIMENTAL RESULTS

This part is dedicated to present and discuss the test results for evaluating the performance of the proposed algorithm. The simulation has been implemented using MATLAB (R2013a) programming language. In these experiments, the parameters are:

- block side (b = 32)
- overlapping (a_x = 12, a_y = 12)
- shuffling key length is the number of blocks, depends on signal length
- dynamic key is a set of matrices, each matrix is used to partially encrypt a single block
- key length is half the number of block samples

Five objective measures are used to test the quality and any intelligible remnants (residual intelligibility) from the recovered signal. Key space and key sensitive tests are performed as well. Performance tests are applied on the proposed algorithm by using online speech signals with different times set at 4, 8, and 16 seconds.

Key Space

This should be as random and as wide as possible in order to prevent infiltration by dictionary or brute-force attacks. Key space size should be greater than 2^{128}, as this limits computational power in terms of a brute-force occurrence. The suggested algorithm, has the following key parameters:

- 2 initial conditions
- 2 control parameter

permutation Key1

- 2 initial conditions
- 2 control parameter

permutation Key2

- 1 initial conditions
- 2 control parameters

masking key

- Dynamic initial condition (interchangeable with block number which is at least 36)

This can take $(10^7)^{12} \approx 2^{279}$ which is brute-force safe.

KeySensitivity

An algorithm should have high sensitivity to key changes. A minor key change can produce a completely different recovered speech signal.

Quality of Speech Encryption and Decryption

Generally there are objective metrics to assess the residual speech and quality of recovered signal. There are three types:

- Time domain metrics include traditional signal to noise ratio (SNR), and segmental signal to noise ratio (segSNR) as expressed in table 1 and 2.
- Linear predictive coefficients (LPCs) such as log likelihood ratio (LLR) as represented in table 1 and 2.
- Spectral domain metrics such as a comparison between the original and processed signals which can be seen in figure 4.

Table (I). OBJECTIVE METRICS FOR ONLINE SPEECH SIGNAL USING BUILD IN MICROPHONE

Time	4sec	8sec	16sec
Encryption Metrics			
SNR	-28.9124	-28.1379	-25.7558
segSNR	-31.2935	-29.4973	-28.107
FreqSegSNR	-31.1581	-29.3496	-27.9231
Cor	-0.0084	0.0026	-0.0011
LLR	0.5723	0.5149	0.5858
Decryption Metrics			
SNRd	24.8789	29.4015	31.3191
segSNRd	26.9567	26.1297	27.4861
FreqSegSNRd	29.2560	28.3977	29.7614
Cor	0.9983	0.9994	0.9996
LLRd	0.0561	0.0513	0.0441

Table (2). OBJECTIVE METRICS FOR ONLINE SPEECH SIGNAL USING EXTERNAL MICROPHONE

Time	4sec	8sec	16sec(Dell)	16sec(Asus)
Encryption Metrics				
SNR	-28.9104	-27.3669	-26.3422	-22.7172
segSNR	-31.4831	-31.3214	-34.3511	-32.0513
FreqSegSNR	-31.3386	-31.2021	-34.2538	-31.9390
Cor	0.0023	-0.0018	0.0034	-0.0057
LLR	0.7330	0.9588	0.9410	1.4168
Decryption Metrics				
SNRd	28.8354	34.1364	37.8781	46.6324
segSNRd	24.1397	24.4529	22.3698	21.3470
FreqSegSNRd	26.3152	26.2462	24.0711	22.9047
Cor	0.9993	0.9998	0.9999	0.9999
LLRd	0.1486	0.0457	0.0370	0.0314

Each table divides into two halves horizontally; the first halve express encryption metrics and the second one used for decryption metrics

The first half in the two tables represents the lowest values for the SNR and segSNR, FreqSegSNR and highest values of LLR. Moreover the correlation coefficient (Cor) is very low which has a good indicator that having a low residual speech values that demonstrate high security in encryption process.

In the second half for the two tables it would be noticed that the highest values of SNR_d, $segSNR_d$, $FreqSegSNR_d$ and lowest values for LLR_d. In addition to the correlation between samples is almost one which means that the recovered speech has a high quality.

In Fig. 4, contains two different shapes for each speech signal test: the waveform and the spectrogram shape. The waveform shows signal amplitude within the time, while the spectrogram shows different features such as energy and frequency which gives more details than the time domain.

Both waveform and spectrogram shapes represent the process of the encryption/decryption representation.

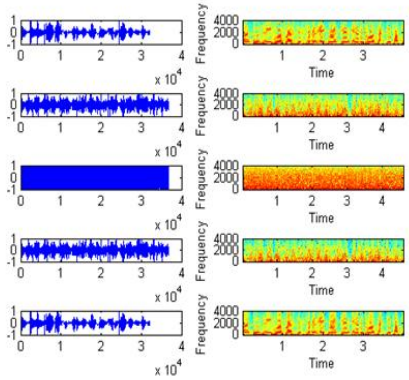
CONCULSION & DISCUSSION

The algorithm for online speech encryption converts an intelligible speech signal to an unintelligible signal using secure multi-domain levels. The aim of this is to limit the approach of the cryptanalyst and to increase the security. Shuffling block processes and scrambling the samples are not enough to remove residual and the silent portions from the speech. So it would give a chance for the trained listener to directly interpret the signal, if any part of it remains intact, therefore the masking process is very important to get the speech completely encrypted.

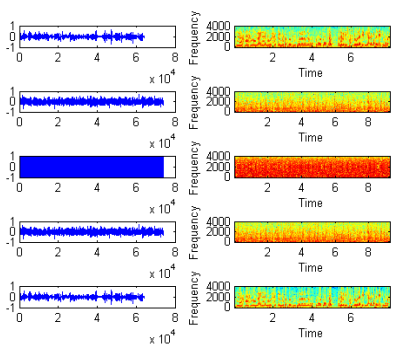
Secret keys control all the permutation and the encryption of the algorithm. This provides enough key space. Moreover the proposal is very sensitive to initial conditions and control parameters of the Hénon chaotic map which makes it robust. The low complexity of computations of block shuffling and generation of secret keys, gives the algorithm a high speed in addition to use a partial encryption.

The spectrogram provides a window into the frequency and time domain at the same time. Thus it is easy to see the theory at work by observing the original signal, the encrypted signal and decrypted version. It is observed that the order of frequencies changes when shuffling the blocks, and is decoded back to the original signal when the inverse permutation version has been applied.

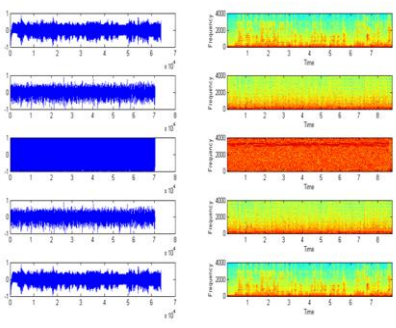
The proposed algorithm outlined in this paper is able to produce efficient speech encryption in a real-time environment based on multiple tests with robust secure.



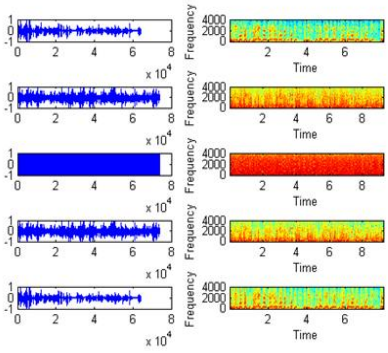
(a) shows online speech signal for 4 seconds



(b) online speech signal for 8 seconds



(c) online speech signal for 8 seconds



(d) online speech signal for 8 seconds

Figure(4). Waveform and spectrogram for speech signal from (a) to (d)

REFERENCES

[1].Zhaopin Su, Guofu Zhang and Jianguo Jiang, Multimedia Security: A Survey of Chaos-Based Encryption Technology, Multimedia - A Multidisciplinary Approach to Complex Issues, Dr. Ioannis Karydis (Ed.), ISBN: 978-953-51-0216-8. (2012).
http://www.intechopen.com/books/multimedia-a-Multidisciplinary-approach-to-complex-issues/multimedia-security-a-survey-of-chaos-based-encryption-technology
[2].Hala B. and Sundus I. ," Modify Speech Cryptosystem Based on Shuffling Overlapping Blocks Technique" International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 4, Issue 2, March-April 2015
[3].National Bureau of Standards ,Data encryption standard. Federal Information Processing Standards Publication no. 46, U.S. Government Printing Office, Washington(1977).
Daemen, J., & Rijndael, V. R. (2001). The advanced encryption standard. Dr. Dobb's Journal, 26(3), 137–139.
[4].Gnanajeyaraman, R., Prasadh, K. & Ramar, D. Audio encryption using higher dimensional chaotic map, International Journal of Recent Trends in Engineering Vol.1 (No.2): 103–107., (2009).
[5].Wang, Y., Wong, K.W., Liao, X. & Chen, G. (2011). A new chaos-based fast image encryption algorithm, Applied Soft Computing Vol.11 (No.1): 514–522.
[6].,M. Ashtiyani, P. Moradi Birgani, S. S. Karimi Madahi, "Speech Signal Encryption Using Chaotic Symmetric Cryptography". J. Basic. Appl. Sci. Res., 2(2)1678-1684, 2012© 2012, TextRoad Publication
[7].Zhenjun Tang & Xianquan Zhang & Weiwei Lan "Efficient image encryption with block shuffling and chaotic map" #Springer Science & Business Media New York 2014
[8].E. Mosa, Nagy.W. Messiha, and O. Zahran "Chaotic encryption of speech signals in transform domains". International conference on computer engineering &systems; 2009. p. 300–5, 14–16. doi: http://dx.doi.org/10.1109/ICCES.2009.5383252.
[9].Osama M. Abu Zaid1 , and Moussa Demba2," A Proposed Encryption Scheme based on Henon Chaotic System (PESH) for Image Security" International Journal of Computer Applications (0975 – 8887) Volume 61– No.5, January 2013
[10].MADHEKAR SUNEEL," Cryptographic pseudo-random sequences from the chaotic Henon map", Adhana Vol. 34, Part 5, October 2009,
[11].M. Francois and D. Defour, "A Pseudo-Random Bit Generator Using Three Chaotic Logistic Maps ", hal-00785380, version 1 – 6 Feb 2013.
[12].M. N. Elsherbeny, M. Rahal, "Pseudo Random Number Generator Using Deterministic Chaotic System ", International Journal of scientific& technology research, Vol 1, Issue 9, October 2012.
[13]Sadkhan Sattar B. and Abbas Nidaa A., "Speech Scrambling Based on Wavelet Transform," in ,"Advances in Wavelet Theory and Their Applications in Engineering"Physics and Technology, edited by: Dumitru Baleanu, InTech, (2012).