Encryption of speech signal with multiple secret keys

Dalila Slimani
LISIC Lab. Telecommunications Department
USTHB University
P.O.Box 32 El Alia 16111 Bab Ezzouar, Algiers
Algeria
slimanidalila@usthb.dz

Fatiha Merazka
LISIC Lab. Telecommunications Department
USTHB University
P.O.Box 32 El Alia 16111 Bab Ezzouar, Algiers
Algeria
fmerazka@usthb.dz

***Abstract*— In this paper, we proposed an encryption system for speech signals based on circular shifts in row and column. This cryptosystem uses three secret keys. The original key is generated, randomly, using a pseudo noise sequence generator, and the two other keys are obtained by using the main key. The encryption system also uses Discrete Cosine Transform (DCT) or the Discrete Sine Transform (DST) to remove the signal intelligibility. Moreover, the performance of the proposed algorithm is also estimated by correlation coefficient quantity.**

Keywords—Speech signal, Encryption, Decryption, circular shifts, DCT, DST.

I. INTRODUCTION

Speech is one of the most fundamental forms of human communications. Today, due to communications technology, we can sent and received any speech files, but notwithstanding its numerous benefits [1]. To protect the content of speech signal during communication, introduction of the specific encryption systems is usually a must. Due to some inherent features of speech like high data redundancy, the encryption of speech is different from that of texts; therefore it is difficult to handle them by traditional encryption methods [2].

Analog speech encryption is completed using various procedures in the time [3] or frequency domains [4], such as random substitution of data, inversion or transformation of the spectrum [5] [6], or insertion of dummy data between replaced or inverted data. [7]. The aim of analog speech encryption is to ensure that the encrypted speech is unintelligible and that the properties of the encrypted speech, such as the amplitude distribution or spectral bandwidth, are similar to those of a speech signal, because this makes it possible to transmit the encrypted speech over an ordinary speech transmission channel. On the other hand, it is not robust against decryption attacks, as described in [8], which reported that these classical methods can be decrypted by trying all possible permutations of frequency band substitutes or inversions. Therefore, we present an analog speech encryption method based chaotic, which is more robust against decryption attacks and can be used for practical secure communication.

Encryption by chaotic maps is widely used in image processing due to its random-like performance and its sensitivity to initial conditions in addition to its high confusion property [9][10]. In this paper, we attempt to implement the concepts of permutation used in chaotic encryption for speech signals, but with cyclic shifts based on secret keys.

Our cryptosystem has the advantages of the high degree of security, and the small implementation time.

Speech encryption tries to perform a completely reversible operation on speech by analog scramblers or digital encryption devices to be totally unintelligible to any unauthorized listener. Digital encryption is more secure than analog, but it needs a complex implementation and a large bandwidth for transmission. Thus, in the case of limited bandwidth channels, analog scramblers are better [11][12].

In image and text encryption, if plaintext elements are permutated and substituted to give cipher text, the output can't be recognized although small portion of original plaintext remain intact. In speech encryption the problem is that if small portion of original plaintext (clear speech) remains intact it allows a trained listener to directly interpret the scrambled speech. Our objective here is to present an encrypted signal without residual intelligibility.

Our work focuses on the speech signal encryption using a multi key cryptosystem. For this, the paper is organized as follows: In Section 2, we introduce the encryption system and explain its different steps. In section 3, the decryption process is given. The results of tests and comparisons are presented in Section 4. Finally, we conclude this paper in Section 5.

II. PROPOSED ENCRYPTION ALGORITHM

Our cryptosystem is based on the permutation of the segments of the speech signal using three secret keys. The permutation process uses a circular shift (in row and column) calculated from the bits of the encryption keys as shown by Figure 1.

The encryption system also uses substitution to different values by DCT or DST to remove the signal intelligibility [13].

The input speech signal segments are divided and reshaped to fixed size blocks which their size depend on the secret key. The encryption algorithm steps are as follows:

- *Step 1 : Framing and reshaping into 2-D block*
- *Step 2 : Circular shifts (in row and column)*
  - *1 st Round*
    - *Step 3 : Generation the main key K1*
    - *Step 4 : Permutation with a main key K1*
    - *Step 5 : Generation of the mask M1*
    - *Step 6 : addition of mask M1*
  - *2nd Round*
    - *Step 7: DCT or DST*
    - *Step 8 : Generation of second key K2*
    - *Step 9 : Permutation with a second key K2*
    - *Step 10 : Generation of the mask M2*
    - *Step 11 : addition of mask M2*
  - *3rd Round*
    - *Step 12 : Inverse Discrete Cosine Transform (IDCT) or Inverse Discrete Sine Transform (IDST)*
    - *Step 13 : Generation of third key K3*
    - *Step 14 : Permutation with a third key K3*
- *Step 15 : Reshaping into 1-D format.*

The algorithm steps are given below.

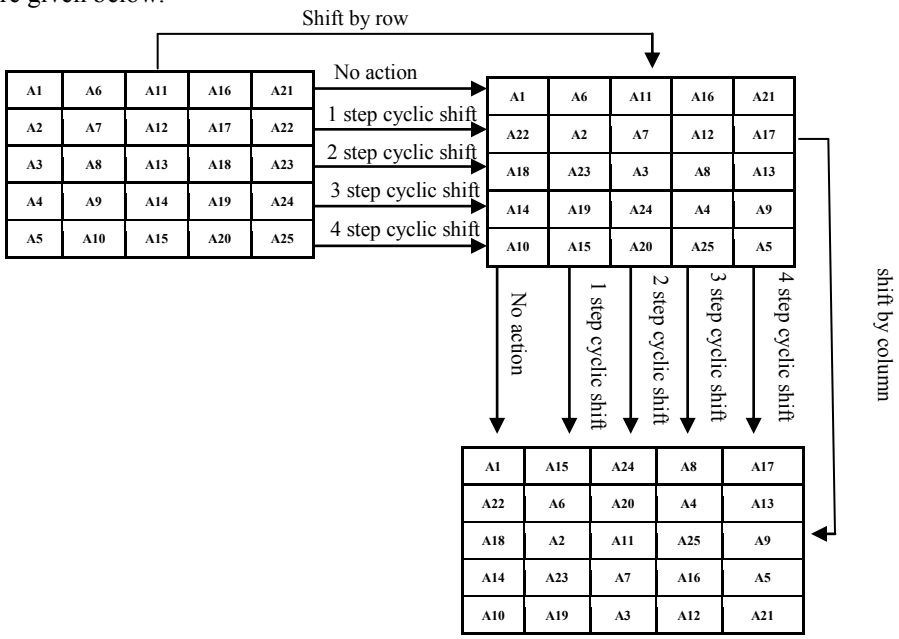


Figure1: Circular shift by row and column

A. *Framing and reshaping into 2-D blocks.*

The initial step of the encryption algorithm is the sampled speech signal and the recorded discrete form with amplitudes between -1 and 1.

The series of samples are grouped in square blocks of width equal to the length of the secret key.

B. *Circular shifts*

In this step, square blocks are transformed by circular shift. We use a circular shift by row and column. An example is given in Figure 1.

C. *Generation of secret keys.*

The encryption algorithm uses three secret keys. The original key is generated, randomly, using a pseudo noise sequence generator. The second key is the inverse of the original key. The third key is generated from the original key by dividing it into two halves and reversing the two halves.

For example if K1 represent the original key:

K1 = 0011 0001

The second key K2 is obtained by inversing K1:

K 2 = 1100 1110

The third key K3 is generated from K1:

K3 = 0001 0011

D. *Permutation with a secret key.*

We also use a circular shift by row and column for permutation process. The secret keys control the permutation process:

- If the key bit is 1: the row or column is shifted by (index of row or column -1) steps.
- If the key bit is equal to 0: the row or column remains unchanged.

Figure 2 show an example of permutation with secret key by row and column.

E. *Generation of mask.*

The encryption algorithm uses two masks M1 and M2 which are respectively generated using a circular shift of the keys K1 and K2.

Figure 3 show an example to generation of mask.

The application of masks M1 and M2 are used to encrypt non permuted portions of the signal to increase the security of the cryptosystem.

F. *Addition of mask.*

The mask generated from the encryption key is then added to the block. Thereafter, we use the subtraction of 2 for any value greater than 1.

G. *Discrete Cosine Transform (DCT) or Discrete Sine Transform (DST).*

To remove the intelligibility of speech signal, we used the Discrete Cosine Trans-form (DCT) or the Discrete Sine Transform (DST) after applying the mask.

Substitution by DCT and DST technique has many advantages over substitution by other discrete transforms as it is produced scrambled speech with lower residual intelligibility, and enabled the recovery of high-quality speech after decryption [5].

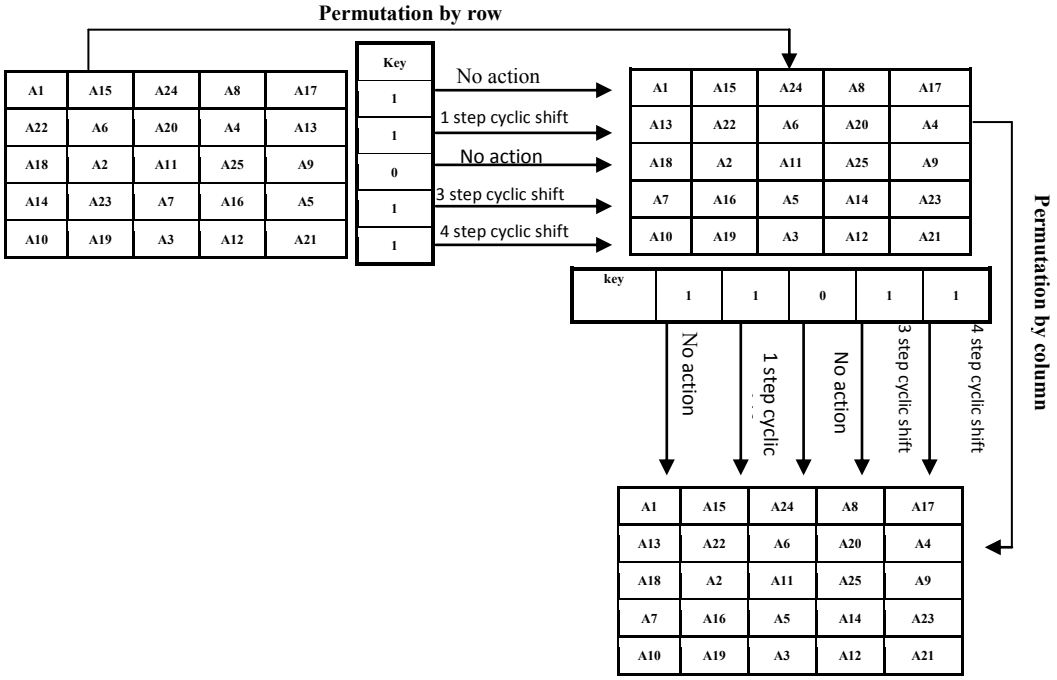


Figure2: Row and column permutation steps

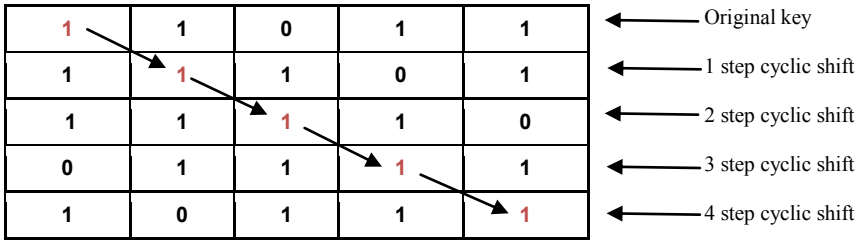


Figure3: *Generation of mask*

III. PROPOSED DECRYPTION ALGORITHM

The decryption algorithm uses inverse operations to encrypted message to retrieve the original message. Decrypt the encrypted speech as the following steps:

- *Step 1 : Generation of three secret keys K1, K2 and K3*
- *Step 2 : Framing and reshaping into 2-D block*
- *Step 3 : : Inverse permutation with a third key K3*
- *Step 4: DCT or DST*
- *Step 5 : Generation of the mask M2*
- *Step 6 : Subtraction of mask M2*
- *Step 7 : Inverse permutation with the second key K2*
- *Step 8 IDST or IDCT*
- *Step9 : Generation of the mask M1*
- *Step10 : Subtraction of mask M1*
- *Step11 : Inverse permutation with the main key K1*
- *Step 12 :inverse Circular shifts (in row and column)*
- *Step 13 : Reshaping into 1-D format.*

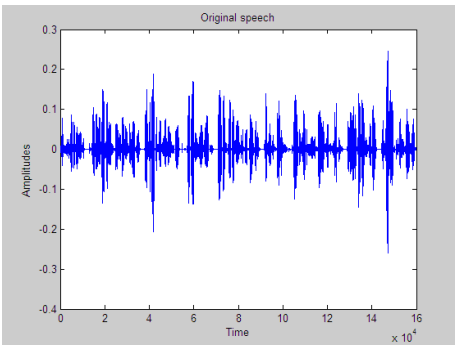
In the decryption process, we used the subtraction between the block and the mask and the addition of 2 for any value below -1 to guarantee the correct reconstruction of these sample values.

IV. EVALUATIONS AND RESULTS

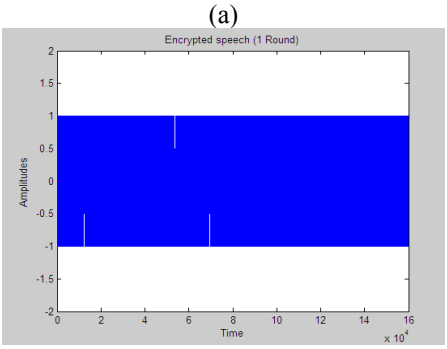
To test our encryption algorithm, we used a speech signal from TIMIT database [14]. It is encrypted with three different methods:

- In the first method the signal is encrypted in the time domain using only a single round.
- The second method uses the cryptosystem with three rounds and DST.
- The third also uses three rounds with DCT.

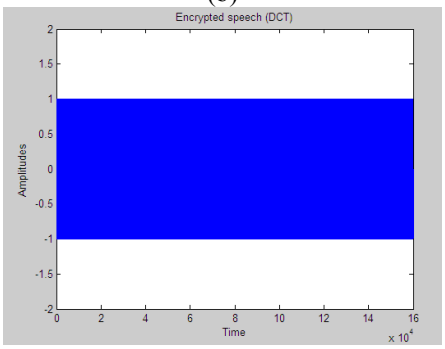
The results of encrypted speech with different methods are shown in Figure 4. It is clear that the encrypted speech with the DCT and DST encryption is visibly similar to the white noise.



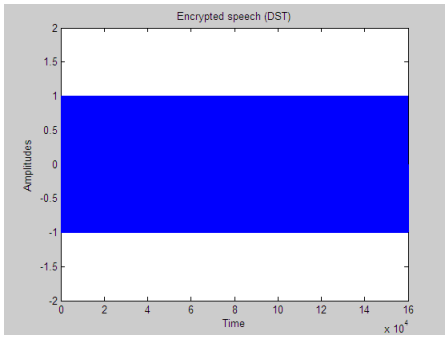
(a)



(b)

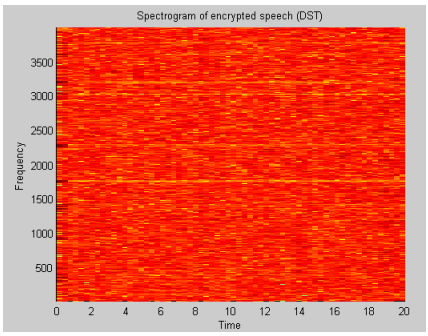


(c)



(d)

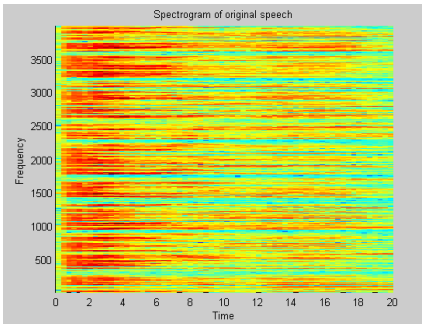
Figure4. Encryption of the speech signals
(**a**) Original signal. (**b**) 1 Round encryption.
(**c**) With DCT encryption. (**d**) With DST encryption



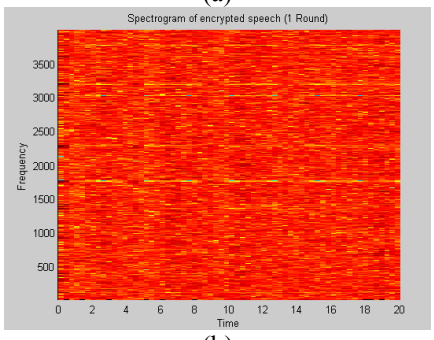
(d)

Figure5. Spectrograms of speech signal
(a) Original signal. (b) 1 Round encryption.
(c) With DCT encryption.(d) With DST encryption

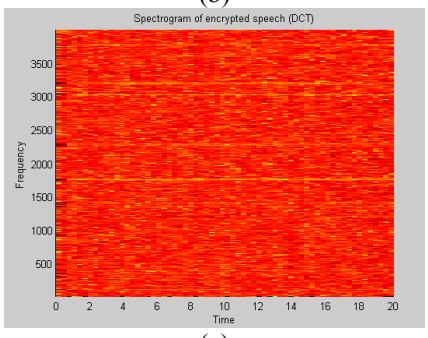
Figures 5 and 6 respectively represent the spectrograms and the histograms of the original signal and the encrypted signal with the three used methods.



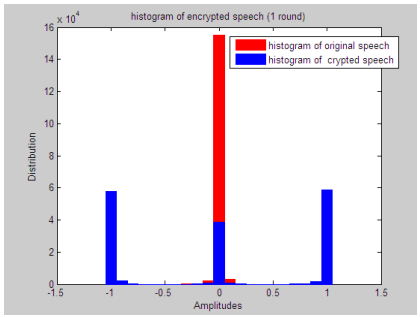
(a)



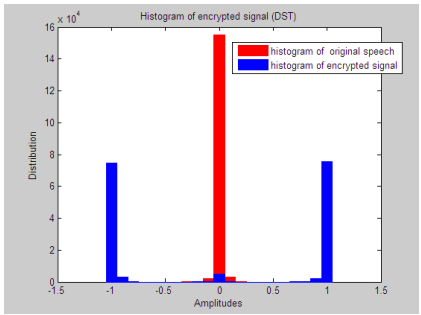
(b)



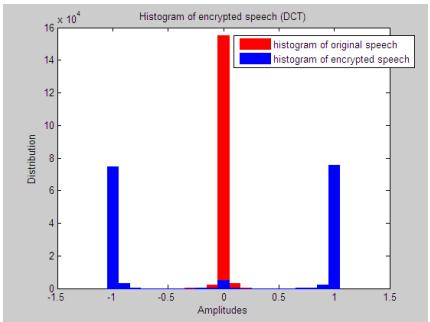
(c)



(a)

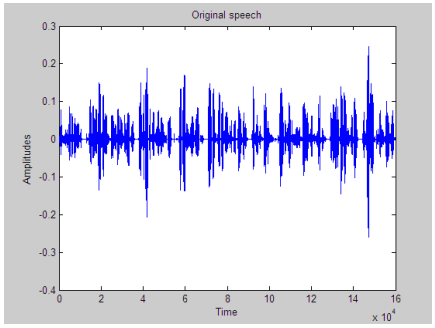


(b)

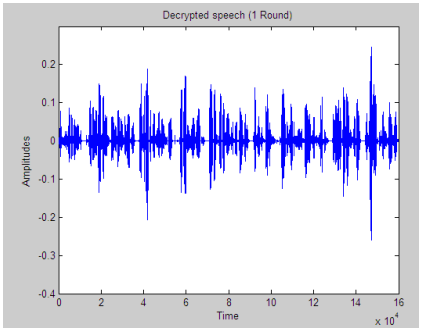


(c)

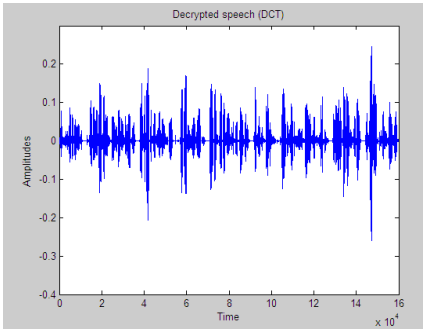
Figure 6. Histograms of encrypted speech and original speech.
(**a**) 1 Round encryption. (**b**) With DCT encryption. (**c**) With DST encryption



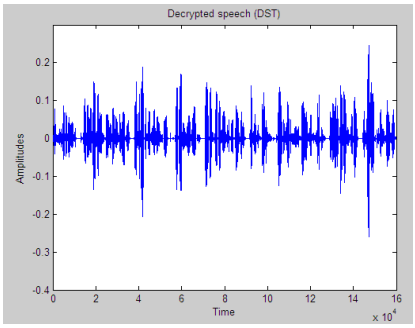
(a)



(b)



(c)



(d)

Figure 7. Decrypted speech signals.
(a) Original signal (b) 1 Round Decryption
(c) With DCT Decryption (d) With DST Decryption

Figure 7 shows the original signals and the signals obtained by the decryption process. From results we can notice those restored signals are comparable to the original.

We can notice that the spectrogram of the speech signal encrypted with the DCT and DST method are similar to white noise.

In order to evaluate the quality of the encryption algorithm, the correlation coefficient is used. To calculate the correlation coefficients between the original signal x and the encrypted signal y the following equation is used:

$$r_{xy} = \frac{c_v(x,y)}{\sqrt{D(x)}\sqrt{D(y)}}$$

where: $c_v(x,y)$: covariance between x and y

$D(x)$: variance of x

$D(y)$: variance of y

The low value of the correlation coefficient indicates a good encryption quality.

We use three different encryption keys KA, KB and KC. TABLE I gives the correlation coefficients between the original signal and the encrypted signal by the three methods.

TABLE I. CORRELATION COEFFICIENTS BETWEEN THE ORIGINAL AND ENCRYPTED SPEECH SIGNAL

Method / Key	1 Round	DCT	DST
KA	0.0287	-0.0412	-0.0412
KB	0.0114	-0.0790	-0.0790
KC	0.0043	-0.0224	-0.0224

We can notice that there is a low correlation between the original signal and the encrypted signal with different used keys. This means that different encryption keys give a good encryption quality. The encryption algorithm is very sensitive to the secret key.
We also note that we have obtained the same value of the correlation coefficients using the DCT and DST.
The use of the DST or DCT improves the quality of the encryption; this can be observed by the correlation coefficient which is considerably higher for the encryption with one round.
The permutation process uses a circular shift. We also use the correlation coefficient for compare the encryption quality using:

- circular shift in row
- circular shift in column
- circular shift in row and column

The results obtained for the different methods are given in TABLE II.

TABLE II. CORRELATION COEFFICIENTS BETWEEN THE ORIGINAL AND ENCRYPTED SIGNAL FOR DIFFERENT CIRCULAR SHIFT

Method / Circular shift	1Round	DCT	DST
in row column	0.0618	-0.0230	-0.0230
In column	0.1697	-0.0099	-0.0099
in row and column	0.0287	-0.0412	-0.0412

From these results, we can see always that there is a low correlation between the original signal and the encrypted signal. But the use of the circular shift in row and column gives a best quality of encryption.

In TABLES III and IV, we respectively represent the values of the encryption and decryption time of two different files: male and female persons.

TABLE III. COMPARATIVE EXECUTION TIMES (IN SECONDS) OF ENCRYPTION ALGORITHMS

Encryption time (seconds)			
Speech signal	**1 Round**	**DCT**	**DST**
Male	15.466214	44.804881	37.049633
Female	16.459817	40.510601	36.076020

TABLE IV. COMPARATIVE EXECUTION TIMES (IN SECONDS) OF DECRYPTION ALGORITHMS

Decryption time (seconds)			
Speech signal	**1 Round**	**DCT**	**DST**
Male	17.700650	46.853220	38.023252
Female	15.409765	41.266836	41.882731

From the results it is easy to observe that encryption and decryption method with one round has an advantage over other algorithms in terms of execution times.

We can notice that the time of encryption and decryption with DCT or DST is higher. This can be explained by the use of 3 rounds instead of one round.

The time of encryption and decryption is also depending on the size of the used file. The time of encryption and decryption are inversely proportional to the file size.

Figures 8 and 9 show the variation in time of encryption and decryption of two files used.

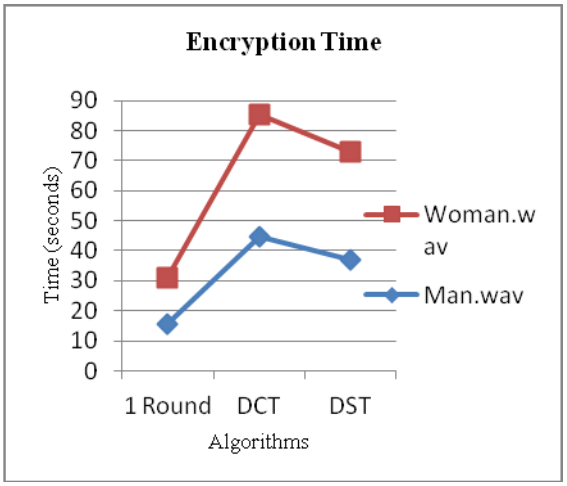


Figure 8: Encryption time

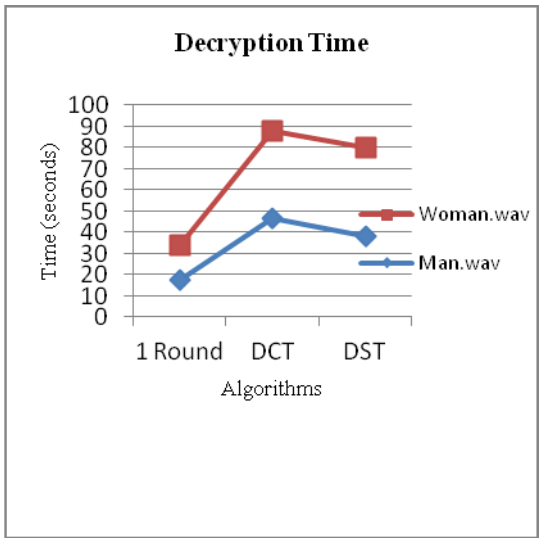


Figure 9: Decryption time

Finally, encrypted speech signals that completely different from original signal is generated. The proposed system uses the simple substitution and fast permutation process. The aim of the proposed system is to change the characteristics of speech signal and eliminate the intelligibility of the encrypted signal. The security aspects of the proposed system depend on different parameters including secrets keys, permutation process and variable number of rounds.

V. CONCLUSION

In this paper, we presented an encryption and decryption algorithms for speech signal using three secrets keys. The proposed algorithm is based on the permutation of the

segments of the speech signal. The permutation process uses a circular shift (in row and column) calculated from the bits of the encryption keys. The encryption system also uses DCT or DST to remove the signal intelligibility.

To test the encryption algorithm, we used two speech signals (male and female speakers). These signals are encrypted with three different methods: The first method is encrypted signal in the time domain using only a single round. The second method uses the cryptosystem with three rounds and DST. The third also uses three rounds with DCT. By using the correlation coefficients between the original signal and the encrypted signal can compared the quality of proposed algorithm. Experimental analysis of the algorithm shows that the algorithm is stronger and more secure. For future work we can suggestion to replace the secrets keys by chaotic keys.

REFERENCES

[1] G. Manjunath and G.V. Anand, “Speech encryption using circulant transformations. In: IEEE Int. Conf. Multimedia and Expo”., vol. 1, pp. 553–556, 2002

[2] Long Jye Sheu, “A speech encryption using fractional chaotic systems”, Nonlinear Dynamics, vol 65, pp. 103-108, 2011.

[3] H. C. Baker and F. C. Piper, “Secure Speech Communications, Academic Press“, 1985.

[4] L. Lee, S., Chou, G. C., & C. S. Chang,. “A new frequency domain speech scrambling system which does not require frame synchronization“. IEEE Transactions on Communications, vol. 32, pp. 444– 456, 1984.

[5] B.Goldburg, S.Sridharan, and E. Dawson, "Design and cryptanalysis of transform- based analog speech scramblers", IEEE Journal of Selected Areas on Communication, vol. 11, pp. 735- 743, 1993.

[6] A. Jameel, M.Y. Siyal and N. Ahmed, “Transform-domain and DSP based secure speech communication system”. Elsevier, Microprocessors and Microsystems”, vol. 31, pp. 335–346, 2007.

[7] A. Matsunaga, K. Koga, and M. Ohkawa, “An Analog Speech Scrambling System Using the FFT Technique with High-Level Security,” IEEE Journal of Selected Areas in Communications, vol. 7(4), pp. 540–547, 1989.

[8] B Goldburg, E Dawson and S Sridharan, “The Automated Cryptanalysis of Analog Speech Scramblers,” Advances in Cryptology EUROCRYPT 91, Lecture Notes in Computer Science, vol. 547/1991,pp. 422-430, 1991.

[9] J. Fridrich, “Symmetric ciphers based on two-dimensional chaotic maps“. International Journal of Bifurcation and Chaos, vol. 8(6), pp. 1259–1284, 1998.

[10] Emad Mosa, Nagy W. Messiha, Osama Zahran, Fathi E. Abd El-Samie, “Chaotic encryption of speech signals”, International Journal of Speech Technology, vol. 14(4), pp. 285-296, 2011.

[11] de Andrade, J.F.; de Campos, M.L.R.; Apolinario, J.A., "Speech privacy for modern mobile communication systems," Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pp.1777-1780, 2008

[12] S.Rajanarayanan and A. Pushparaghavan, “Recent developments in signal encryption- A critical survey”; International Journal of Scientific and Research Publications, vol. 2(6), pp. 1-7, 2012.

[13] E. Mosa, N.W. Messiha, O. Zahran and F.E. Abd El-Samie, “Encryption of speech signal with multiple secret keys in time and transform domains”, vol;. 13, pp. 231-242, 2010.

[14] NIST,Timit Speech Corpus, NIST 1990.