3

Speech Scrambling Based on Wavelet Transform

Sattar Sadkhan and Nidaa Abbas
University of Babylon
Iraq

1. Introduction

The increased interest in analog speech scrambling techniques are due to the increased visibility and publicity given to the vulnerability of communication systems to eavesdropping of unauthorized remote access (Gersho & Steele, 1984). In wireless communications, including High Frequency (H.F) and satellite communications, it is almost impossible to prevent unauthorized people from eavesdropping unless speech scramblers may be used to protect privacy. Among speech scramblers, analog scramblers are attractive and wide applicable. The conventional analog scramblers manipulate speech signal in the frequency or time domain or both. A typical frequency domain scrambler is the band splitting scrambler, which breaks the speech signal into several sub bands and permutes them. A typical time domain scrambler is the time division scrambler, which breaks the speech signal into short time segments and permutes them within a block of several segments(Sakurai et al., 1984). These conventional analog scramblers cannot provide sufficient security against cryptanalysis because the number of permutable elements in these scramblers is not large enough to provide an adequate number of different permutations due to hardware limitation and processing delays.

To strengthen security, a two-dimensional scrambler which manipulates the speech signal both in the frequency domain and in the time domain was proposed. Regarding other types of scramblers, which can attain a high degree of security, the transform domain scrambler was proposed.

In 1979, Wyner proposed a method, in which the orthogonal transform called a Prolate Spheroidal transform (PSD) was executed on a set of the sampled speech signal. A mathematical basis for using both band splitting and time division, at the same time was presented by F. Pichler in 1983. He showed how an operation which realizes band splitting and time division can be designed, and pointed out that such an operation can be realized by a fast algorithm. The mathematical background is the theory of group-character for finite Abelian Groups and the theory of the General Fast Fourier Transform (GFFT) (Pichler, 1983). Also in 1984 Lin-Shan et. al., presented frequency domain scrambling algorithm, which is an extension of the Discrete Fourier Transform (DFT) scrambler previously proposed. The use of short-time Fourier analysis and filter bank techniques lead to the special feature that the original speech could be correctly recovered while the frame synchronization is completely unnecessary. In 1990 Sridharan et. Al., presented a comparison among five discrete orthogonal transforms in speech encryption systems. The results of the research showed that the Discrete Cosine Transform (DCT) and the Discrete Prolate Spheroidal Transform (DPST) could be used in narrow band systems. The Karhunen Loeve Transform (KLT) and the Discrete Hadamard Transform (DHT) were more suitable where wider bandwidth was available. The DCT turned out to be the best transform with respect to residual intelligibility of the encryption speech and recovered speech quality. The DFT produced results which were inferior to the DCT. The DCT implementation would also offer speed advantage over the FFT (Sridharan et al., 1990).

Original BSS (Blind Source Separation) – based speech encryption system utilizes BSS to perform decryption, but the complexity of BSS algorithms limits the decryption speed and its real-time applications. In 2010 , fast decryption utilizing calculation for BSS-based speech encryption was proposed. The paper analyzed the correlation of speech signals with key signals, and then utilized the correlation calculation to achieve speech decryption. The experiment results showed that correlation calculation decryption nicely simplifies BSS-bsed speech encryption system, largely speeds up the speech decryption, and slightly improves the quality of decrypted speech signals (Guo & Lin, 2010). While Mermoul and Belouchhrani claimed that the interactability of the under-determined BSS problem has been used for the proposal of BSS-based speech encryption has some weakness from cryptographic point of view. In their paper they proposed new encryption method that bypass these weaknesses. Their proposed approach is based on the subspace concept together with the use of nonlinear function and key signals. An interesting feature of the proposed technique is that only a part of the secret key parameters used during encryption is necessary for decryption (Mermoul & Belouchrani, 2010)

(Mosa, et al., 2010) introduced a new speech cryptosystem, which is based on permutation and masking of speech segments using multiple secret keys in both time and transform domain.

In 2000, an automated method for cryptanalysis of DFT-based analog speech scramblers was presented by Wen-Whei and Heng-Iang, through statistical estimation treatments. In the proposed system, the cipher text only attack was formulated as a combinatorial optimization problem leading to a search for the most likely key estimate. For greater efficiency, they also explored the benefits of Genetic Algorithm to develop the method. Simulation results indicated that the global explorative properties of Genetic Algorithms make them very effective at estimating the most likely permutation and by using this estimate significant amount of the intelligibility could be recovered from the cipher text following the attack on DFT-based speech scramblers (Whei & Iang, 2000)

A time-frequency scrambling algorithm based on wavelet packets was proposed by Ajit S. B. Bopardikar (1995) by using different wavelet packet filter banks, they added an extra level of security since the eavesdropper had to choose the correct analysis filter bank, correctly rearrange the time-frequency segments, and choose the correct synthesis bank to get back the original speech signal. Simulations performed with this algorithm give distance measures comparable to those obtained for the uniform filter bank based algorithm(Bopardikar, 1995). In 2005, an analog speech scrambler which is based on Wavelet Transformation and Permutation was proposed by Sattar B. Sadkhan and evaluating the scrambling efficiency through the calculation of distance measures, and takes the effect of the channel noise into consideration (Sadkhan, et al., 2005). In 2007, A Parallel Structure of different wavelet transforms were applied for speech scrambling. The proposed structure provided a good results in comparison with the system implemented in 2005 (Sadkhan, Falah, 2007)

2. Speech scrambling system

Speech Scrambling seeks to perform a completely reversible operation on a portion of speech, that it is totally unintelligible to unauthorized listener. The most important criteria used to evaluate speech scramblers are:

- The scrambler's ability to produce encrypted speech with low residual intelligibility.
- The extent to which the encryption and decryption processes affect the quality of the speech recovered by intended reception; and
- The scrambler's immunity to cryptanalysis attack.

Cryptographers face the problem of designing scrambling systems which distort the very redundant speech signal to the extent that useful information is unable to be recovered. The encryption process must remain secure when subject to the powerful information processing structures of the human auditory system and knowledge-base automated cryptanalytic processes. There are two fundamentally distinct approaches to achieve voice security in speech communication systems: digital ciphering and analog scrambling. In spite of significant progress in digital speech processing technology, analog speech scramblers continue to be important for achieving privacy in many types of voice communication (Gersho & Steele, 1984), due to the desire for secure communication over existing channels with standard telephone bandwidth at acceptable speech quality and reasonable cost. To make the distinction between analog and digital speech encryption devices, the following definitions can be considered. Analog scramblers produce scrambled speech which is analog signal occupying the same bandwidth as the original speech. Analog or digital signal processing may be used to generate this signal. Digital speech encryption systems digitize and compress the input speech in order to obtain a digital representation at a bit rate suitable for the communications channel to be used. The resulting bit stream is encrypted using well-know data encryption techniques. The ability of a digital encryption schemes to compete with the well-established analog scramblers is depend on the quality of the speech compression algorithms used. The speech quality resulting from contemporary compression schemes is rapidly improving (Sakurai, et al., 1984).

Analog speech scrambling experienced a metamorphosis as a result of the development and release of very high speed signal processing hardware. Analog scrambling algorithms which were impractical due to their complex nature are now being implemented in real time using this technology.

One family of analog scramblers that has shown a great deal of promise is the transform domain scrambler. These scramblers operate on speech which has been sampled and digitized. The sampled speech is portioned into frames of equal length, containing N speech samples. A chosen transformation is then performing on each frame to yield a transform vector with N components. Encryption is achieved by permuting these transform components within the vector before the inverse transform is applied to return the components to the time domain. The encrypted time domain frame is transmitted in place of original speech frame (Pichler, 1983).

2.1 Secure speech communication

There are many reasons that make the user hide the meaning of the transmitted speech. Secure speech communication refers to the masked speech communication. Generally, secure speech communication, shown in Fig. 1 , deals with three parts: -

- The first part is transmitter (T_x) which has the ability to produce encrypted speech with low residual intelligibility;
- The second part is receiver (R_x) which recovers the encrypted speech, which is near as possible to the original speech signal.
- The third part is the eavesdropper that attacks the communication system according to many available methods (Lee, 1985).

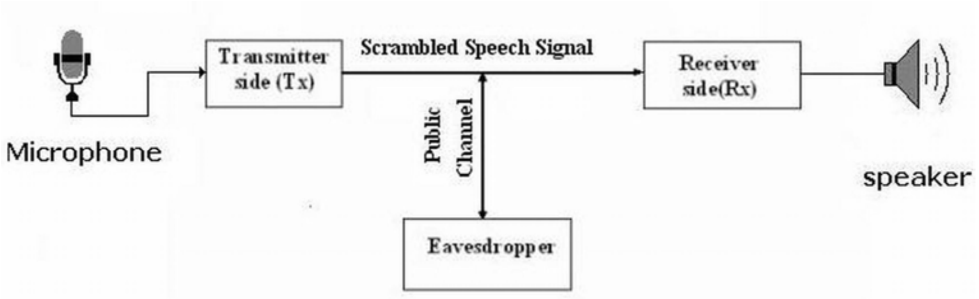


Fig. 1. Block diagram of secure speech communication

The first and second part uses a secure communication channel, while the eavesdropper tries to destroy the security of the communication system. If the system is destroyed, the receiver may lose the ability of getting the transmitted signal. In this case, the first and second parts must try to find another secure algorithm (to be used in the communication system) which is more secure and more difficult to be cryptanalysis by the third part.

The worst case is when the first and second part does not know that the system is destroyed by the third part. The first part wishes to mask or hide the meaning of the transmitted speech where, the second part can recover it without allowing the third part to get any meaningful speech. Almost all speech security systems reduce (at least to some extent) the audio quality of a voice transmission. Security will not be enhanced if the link has been so badly degraded that we have to repeat the same message a number of times.

2.2 Analog speech scrambling

In analog speech scrambling, the only real analog operation is signal transmission, since the signal processing is carried out digitally. Incoming speech signals are digitized using analog to digital converter (ADC) , then processed by a special scrambling algorithm, converted back to analog, and transmitted to a receiver, where they are digitized again, inversely processed (descrambled), and reconverted to analog form for reconstruction, as shown in Fig. 2

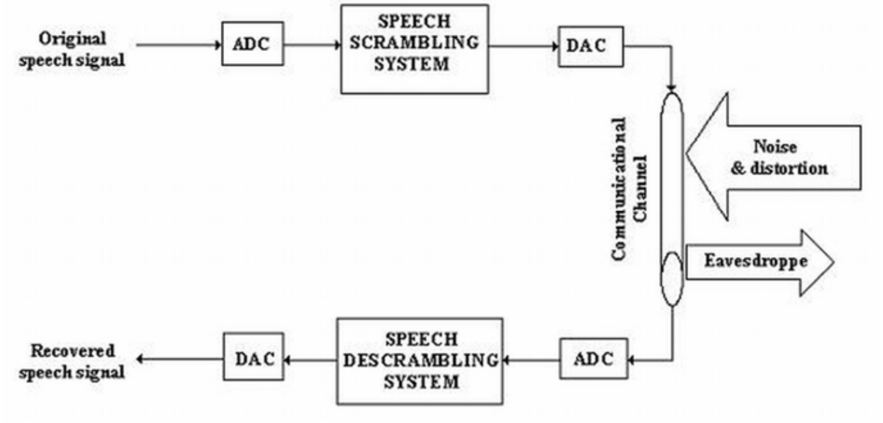


Fig. 2. Block diagram of speech scrambling

There are numerous scrambling methods. The two main processes involve dividing the signal in small time frames and manipulating the frequencies (scrambling in the frequency domain). The descrambler block descrambles the input signal, which must be a scalar or a frame-based column vector. The descrambler block is the inverse of the Scrambler block (Mascarin, 2000). The main attraction of this method arises from the fact that it can be used with the existing analog telephone, H.F., satellite and mobile communication systems, provided the encrypted signal occupies the same bandwidth as the original speech signal (Goldburg, et al., 1991).

3. Wavelet based speech scrambler

The analog scrambling process which employs a transformation of the input speech to facilitate encryption can best be described using matrix algebra. Let us consider the vector *x* which contains **N** speech time samples obtained from A/D conversion process, representing a frame of the original speech signal. Let this speech sample vector *x* be subject to an orthogonal transformation matrix *F* such that:

$$u = F \cdot x \tag{1}$$

This transformation results in a new vector (u) made up of transform coefficients. A permutation matrix is applied to (u), such that each transform coefficient is moved to a new position within the vector given by:

$$v = P \cdot u \tag{2}$$

A scrambled speech vector *y* is obtained by returning vector *v* to the time domain using the inverse transformation F^{-1} where:

$$y = F^{-1} \cdot v \tag{3}$$

Descrambling, or recovery of the original speech vector *x*' is achieved by first transforming *y* back to the transform domain .The inverse permutation matrix P^{-1} is then used to return the transform coefficients to their original position. Finally, the resulting transform vector is returned to the time domain by multiplying by F^{-1}

$$x' = F^{-1} \cdot P^{-1} \cdot F \cdot y \tag{4}$$

The transform domain scrambling process outlined above requires the transform matrix F to have an inverse. One attempts to insure that the scrambling transformation $T=F^{-1}.P.F$ is orthogonal. The inverse transformation T^{-1} will also be orthogonal. This property is useful since any noise added to the scrambling signal during transmission will not be enhanced by the descrambling process as shown in Fig. 3. The scrambled speech sequence is given by (Goldberg, et al., 1993):

$$y = F^{-1} \cdot P \cdot F \cdot x = T \cdot x \tag{5}$$

At most. N elements are able to be permuted in the transform –based scrambling process. It is important to note that for a given sampling frequency. N will determine the delay introduced by the scrambling device. So a tradeoff between system delay and security .N is usually chosen to be equal to 256. Practically, the number of transform coefficients M! possible coefficient arrangements. this restriction stems from the requirement that the scrambled speech should occupy the same bandwidth as the original speech. If the, and Biorthognal wavelets etc… transform components have a frequency representation. those lying outside the allowable band are set to zero and the reminder are permuted. Methods for generating all M! Possible permutations have been addressed (Bopardikar, 1995) . The permutations must carefully screen to ensure that components will undergo a significant displacement from their original position in the vector. In addition, components which were adjacent in the original vector should be separated in the scrambled vector.

If it is assumed during it's passage over the communications channel, a noise component μ is added to y, then we have

$$y' = y + \mu \tag{6}$$

Where y' is the signal observed by the receiver. The inverse scrambling transformation is then applied to y' in order to descramble and recover the original sequence x.

$$x' = T^{-1} \cdot y' \quad = x + T^{-1} \cdot \mu \tag{7}$$

Now since T^{-1} is orthogonal, and hence norm preserving $\|\mu\| = \|T^{-1} \cdot \mu\|$ .This implies that the noise energy is not enhanced as a result of the scrambling process.

3.1 Permutation used in the scrambler

The number of possible permutations of ***N*** elements is ***N!***. However, all of these permutations cannot be used because some of them do not provide enough security.

Let ***P*** be a set of permutations, and let $\boldsymbol{P^{-1}}$ be the set of inverse permutations corresponding to the permutation in ***P***. The set S has to satisfy the requirement that any permutation in ***P*** must not produce an intelligible scrambled speech. It is difficult to evaluate the intelligibility of the scrambled speech signal and the intelligibility of the descrambled speech signal by a quantitative criterion because intelligibility is substantially a subjective matter, as shown in Fig. 3.

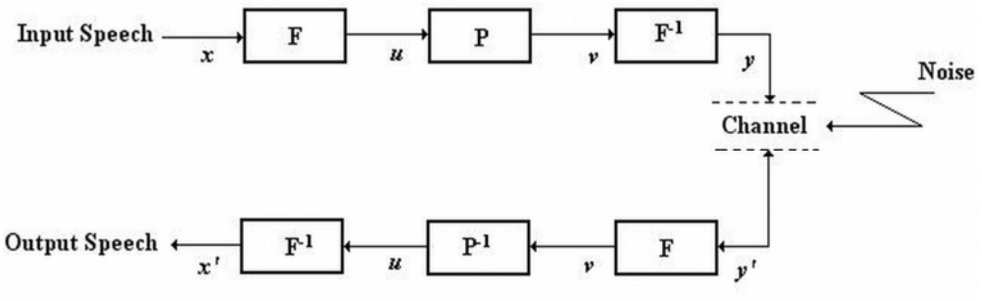


Fig. 3. Block diagram of speech scrambling

A permutation of ***N*** elements can be expressed as (Rao &Homer, 2001).

$$P = \begin{bmatrix} 1,2,3,............N \\ k_1,k_2,k_3,.....k_N \end{bmatrix} \tag{8}$$

by the permutation P on a set of ***N*** elements A=(e_1,e_2,e_3,……,e_N), the element of A at the ***ith*** position is moved to the k_ith position ,namely,

$$P'A = P(e_1,e_2,e_3,.....,e_N) = (d_1,d_2,d_3,.......d_N) \tag{9}$$

Where

$$d_{ki} = e_i \; (i=1,2,3\ldots N) \tag{10}$$

One of the most common and easiest ways of analyzing this is the Hamming distance (HD) between a pair of permutations. The HD of a permutation P is defined as the number of digits which are not coincident in the same position.

3.2 Measures for residual intelligibility and recovered voice quality

Voice quality of the recovered speech and the residual intelligibility of the encrypted speech are usually judged by subjective quality tests. Unfortunately, these tests take much time and labor, and require a large number of trained listeners. Even though intelligibility is a substantially subjective matter, it is possible to use objective tests which are useful, (if not ideal) indicators of intelligibility (Sridharn, et al.; 1991)

The objective measures are useful in indicating the residual intelligibility of encrypted speech and the corresponding quality of recovered speech .

A distance measure is an assignment of a number to an input/output pair of a system. To be useful, a distance measure must posses to a certain degree the following properties :

- It must be subjectively meaningful in the sense that small and large distance must correspond to low and high subjective quality, respectively;
- It must be tractable in the sense that it is possible to mathematically analyze and implement it in some algorithms.

One use of the distance measures is to evaluate the performance of speech scrambling system. The signal-to-noise ratio **(SNR)** and the segmental signal-to-noise Ratio **(SEGSNR)** are the most common time-domain measures (Gray, et al., 1980) · of the difference between original and processed speech signals (scrambled or descrambled speech signals).

3.2.1 Signal-to-Noise Ratio

The signal-to-noise ratio **(SNR)** can be defined as the ratio between the input signal power and the noise power, and is given in decibels **(dB)** as:

$$SNR = 10.\log_{10}\left\{ \sum_{n=1}^{N} X^2(n) \,/\, \sum_{n=1}^{N}\left[X(n) - Y(n)\right]^2 \right\} \text{ (in } \mathbf{dB}\text{).} \tag{11}$$

Where ***N*** is the number of samples, ***X(n)*** is the original speech signal and ***Y(n)*** is the scrambled or descrambled speech signal. The principal benefit of the **SNR** quality measure is its mathematical simplicity. The measure represents an average error over time and frequency for a processed signal. However, **SNR** is a poor estimator for a broad range of speech distortions. The fact that **SNR** is not particularly well related to any subjective attribute of speech quality and that it weights all time domain errors in the speech waveform equality (Gray, Markel 1976) · This can be solved with segmental **SNR**.

3.2.2 Segmental Signal-to-Noise Ratio

An improved version measure can be obtained if **SNR** is measured over short frames and the results are averaged. The frame-based measure is called the segmental **SNR (SEGSNR)** and is defined as:

$$SEGSNR = \overline{SNR(m)} \text{ in } \mathbf{(dB)}. \tag{12}$$

Where $\overline{SNR(m)}$ is the average of ***SNR(m)*** and ***SNR(m)*** is the **SNR** for segment ***m***. The segmentation of the **SNR** permits the objective measure to assign equal weights to load and soft portions of the speech (Yuan, 2003)

4. Results and discussion

In the proposed Wavelet Transform based Speech Scrambling system, (Arabic) messages have been recorded with sampling frequency of 8 kHz as speech files.

At the transmitter, the sampled speech signal is arranged into frames . Each frame contains 256 samples, and then the Wavelet Transformation is performed on each frame. After that, the transform coefficients are permuted before applying the Inverse Wavelet Transform (IWT). The resulting scrambled speech signal is saved in a wave file.

At the receiver, frame by frame of length 256 samples are descrambled and saved in wave file. The proposed scrambled system investigates four types of wavelets: (Haar, db3, sym2 and sym4), each one with three different levels. Two types of tests have been applied to examine the performance of the simulation, these are:

a. Subjective Test: in which the scrambled speech files have been played back to a number of listeners to measure the residual intelligibility, subjectively. For all cases, the judge was that the files contain noise only, which means that the residual intelligibility is very low. The analog recovered speech files have been tested in a similar way to measure the quality of the recovered speech files, the judge was that the files were exactly the same as the original copies.
b. Objective Test: As mentioned earlier, the objective test is a valuable measure to the residual intelligibility of the scrambled speech, and the quality of the recovered speech.

The distance measures indicate the perceptual similarity of the speech recovered following decryption and the original speech. They are also used to quantify the difference between scrambled speech and original speech.

The signal to noise ratio (SNR) and the segmental signal to noise ratio measure (SEGSNR) have been chosen to test the residual intelligibility of the scrambled speech and the quality of the recovered speech for all files. The segmental signal to noise ratio measure (SEGSNR) is an improved version measure of the (SNR).

Generally, these distance measures for all the scrambled speech files are very low (good negative value) which means that the residual intelligibility is very low, and the distance measures for all the recovered speech files are very high (large positive value) which means that the quality of the recovered speech is very high.

Using the relation between estimated PSD (dB/Hz) in relation with frequency of the used speech signals in two cases, as follows:

- To compare the original and scrambled speech.
- To compare the original and descrambled speech.

The wavelet based speech scrambling system have been tested under two states of the simulation, these are:

4.1 Noise free channel simulation

Simulation results of typical experiments with the Wavelet based scrambler, and descrambler for an Arabic word spoken by women's voice '"evenning" are shown in figures (4) to (8), and Tables (1) to (2), using different wavelets and different levels.

Case Study:

Using (Haar) Wavelet , (db3) wavelet, Sym2 and Sym4 wavelet each one will be considered with three different levels for the **Arabic word**'" **evening** ".

Figure (4)shows the waveform, spectrum, and spectrogram of a sample original clear speech signal that represents an Arabic word " evening ".

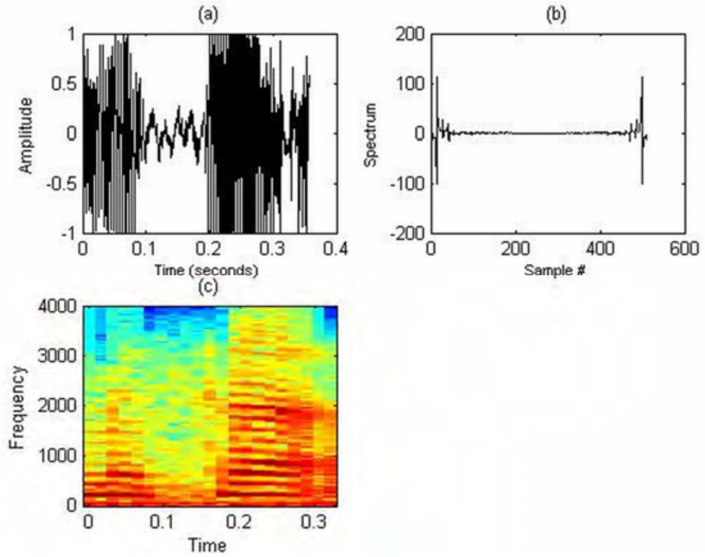


Fig. 4. Original Speech Signal; a) Waveform. (b) Spectrum. (c) Spectrogram.
Using Wavelet Transform (Haar) With Level 1

Figure (5) shows the waveform, spectrum, and spectrogram of the scrambled speech signal, while the comparison of the scrambled speech signal, that resulted from applying a wavelet transform of type (Haar) with a specified level (level 1)is shown in Fig. (6) .

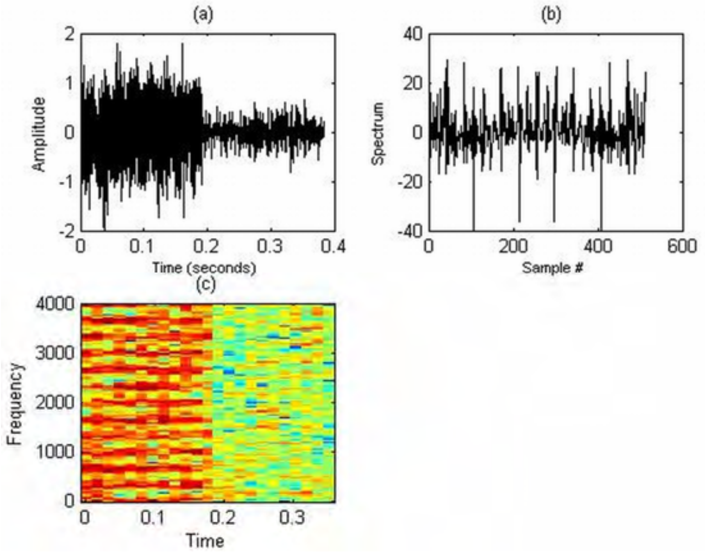


Fig. 5. Scrambled Speech Signal Using Haar Wavelet With Level 1; (a) Waveform. (b)Spectrum. (c) Spectrogram;

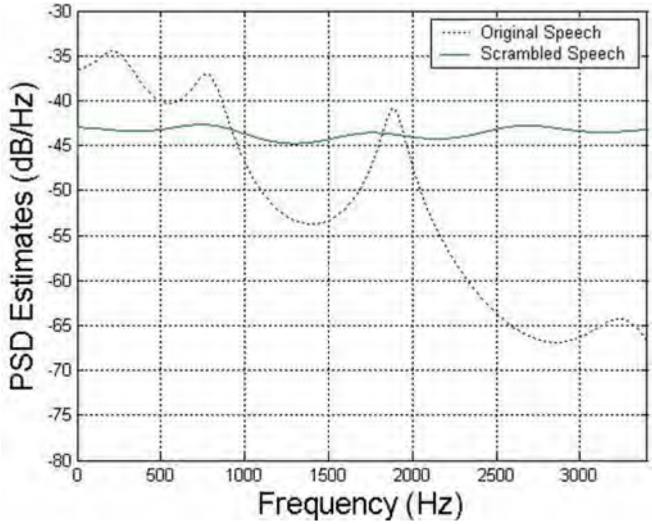


Fig. 6. Comparison between original speech and scrambled Speech using PSD Estimates

Fig. (7) shows the waveform, spectrum , and spectrogram of the resulted descrambled speech signal, while Fig. (8) shows the comparison of the descrambled speech signal and the orignal speech signal.

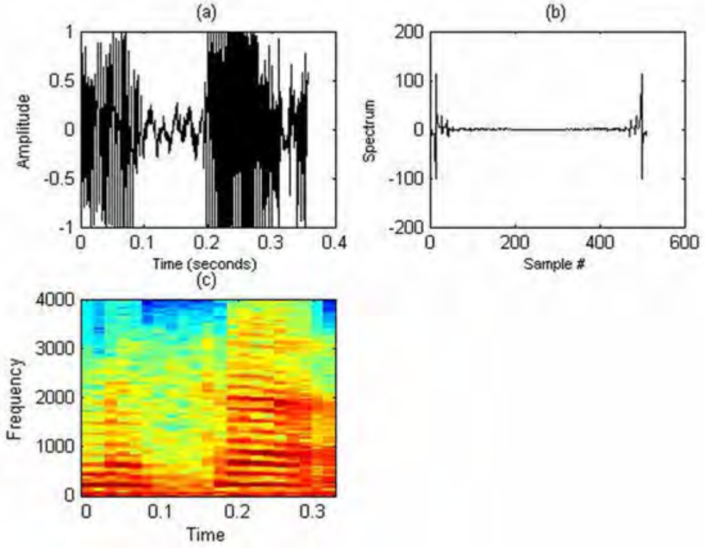


Fig. 7. Descrambled Speech Signal Using Haar Wavelet With Level 1; (a) Waveform. (b) Spectrum. (c) Spectrogram.

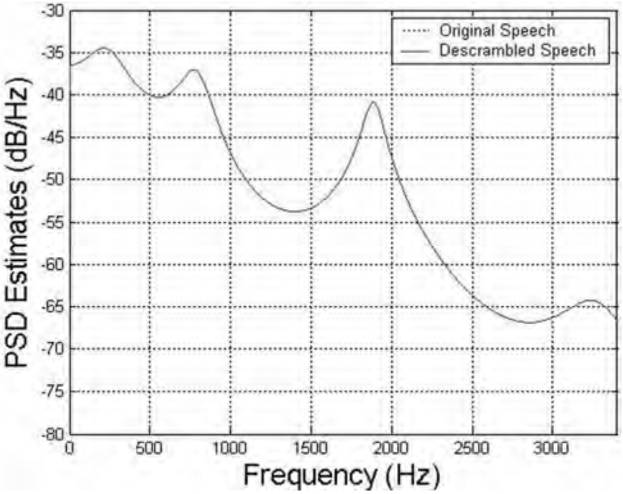


Fig. 8. The Comparison Between Original Speech and Descrambled Speech.

Table (1) shows distance measure (SEGSNRs) for the scrambled speech, while Table (2) shows the (SEGSNRd) distance measure for the descrambled speech for different Wavelets and different decomposition levels.

Level of Decomposition	1	2	3
Haar	-4.8732	-4.0857	-4.0907
Db3	-4.7673	-3.7751	-3.8147
Sym2	-4.8064	-3.7710	-3.6743
Sym4	-4.6620	-3.7125	-3.9071

Table 1. SEGSNRs (dB) for the scrambled speech, for each wavelet with a specific level.

Level of Decomposition	1	2	3
Haar	310.26	305.67	303.12
Db3	15.85	17.46	9.59
Sym2	112.96	18.19	13.91
Sym4	12.88	10.07	13.03

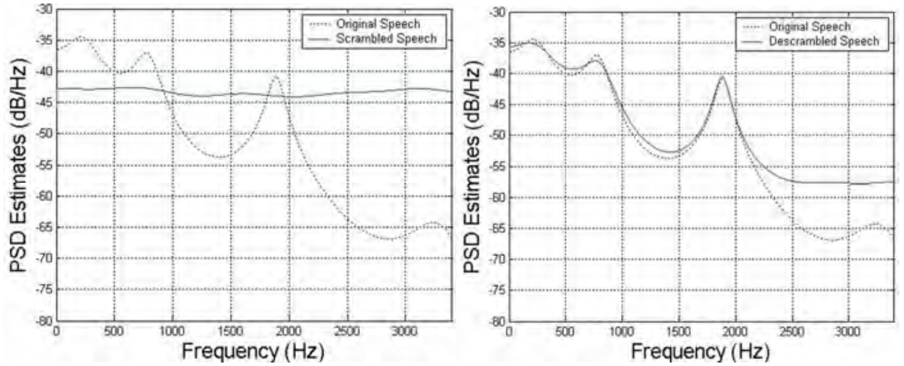
Table 2. SEGSNRd (dB) for the recovered speech, for each wavelet with a specific level.

4.2 Noisy channel simulation

An evaluation of the proposed speech scrambling system with different signal to noise ratios from (5 dB up to 25 dB) was tested.

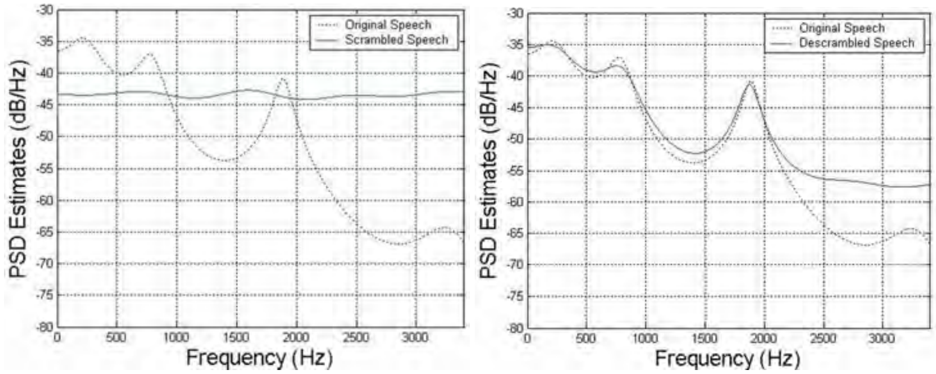
Case study with **SNR = 15 dB**.

The figures (9) to (11), in each one, figure (a) represents spectrogram comparison between original speech signal and scrambled speech signal, while figure (b) represents the comparison between the spectrogram of the descrambled and original speech signal. Both figures are tested under the same level of the chosen Wavelet Transform-Type: Sym2.



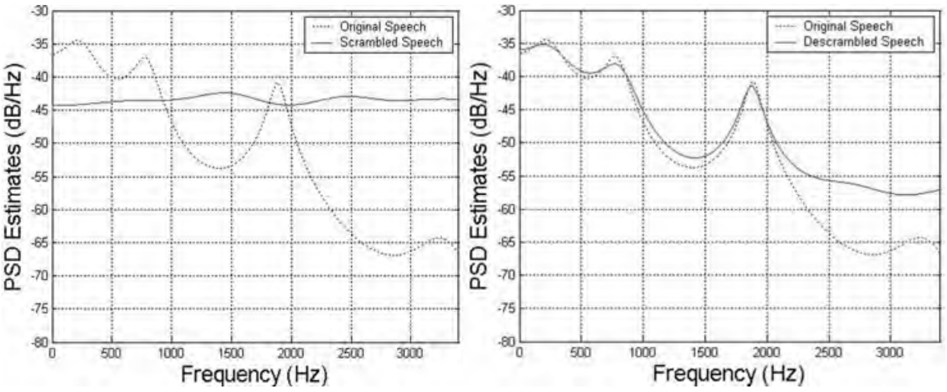
a) b)

Fig. 9. (a) The Comparison between Original Speech and Scrambled Speech Using **sym2 / Level 1**
(b) The Comparison between Original Speech and Descrambled Speech Using **sym2 / Level 1**



a) b)

Fig. 10. (a) The Comparison between Original Speech and Scrambled Speech Using **sym2 / Level 2**
(b) The Comparison between Original Speech and Descrambled Speech Using **sym2 / Level 2**



a) b)

Fig. 11. (a) The Comparison between Original Speech and Scrambled Speech Using **sym2 / Level 3**
(b) The Comparison between Original Speech and Descrambled Speech Using **sym2 / Level 3**

The results from such tests are shown in Tables (3) to Table (8). Table (3) shows the SEGSNRs distance measure for the scrambled speech, and Table (4) shows the SEGSNRd distance measure for the descrambled speech, with SNR = 5 dB. Each two tables corresponding to a specific SNR.

Level of Decomposition	1	2	3
Haar	-5.867	-5.383	-5.292
Db3	-5.711	-5.146	-5.293
Sym2	-5.781	-5.094	-5.017
Sym4	-5.723	-5.220	-5.360

Table 3. SEGSNRs (dB) for the scrambled speech, for each wavelet with a specific level, with SNR = 5 dB.

Level of Decomposition	1	2	3
Haar	3.195	2.852	2.993
Db3	2.530	2.015	1.407
Sym2	2.911	2.481	2.151
Sym4	2.262	1.995	1.204

Table 4. SEGSNRd (dB) for the recovered speech, for each wavelet with a specific level, with SNR = 5 dB.

Level of Decomposition	1	2	3
Haar	-5.007	-4.267	-4.249
Db3	-4.863	-3.950	-4.032
Sym2	-4.898	-3.928	-3.625
Sym4	-4.796	-3.932	-4.112

Table 5. SEGSNRs (dB) for the scrambled speech, for each wavelet with a specific level, with SNR = 15 dB

Level of Decomposition	1	2	3
Haar	13.025	12.852	12.993
Db3	10.073	9.218	7.133
Sym2	12.349	11.001	9.438
Sym4	9.472	8.002	7.707

Table 6. SEGSNRd (dB) for the recovered speech, for each wavelet with a specific level, with SNR =1 5 dB.

Level of Decomposition	1	2	3
Haar	-4.892	-4.111	-4.110
Db3	-4.771	-3.793	-3.846
Sym2	-4.805	-3.784	-3.682
Sym4	-4.672	-3.741	-3.932

Table 7. SEGSNRs (dB) for the scrambled speech, for each wavelet with a specific level, with SNR = 25 dB

Level of Decomposition	1	2	3
Haar	23.088	22.852	22.99
Db3	13.933	13.693	9.184
Sym2	20.304	16.255	12.956
Sym4	12.273	9.824	10.641

Table 8. SEGSNRd (dB) for the recovered speech, for each wavelet with a specific level, with SNR =25 dB.

5. Conclusion

The performance of the Wavelet Transform based speech scrambling system was examined on actual " **Arabic Speech Signals** " , and the results showed that there was no residual intelligibility in the scrambled speech signal. The descrambled speech signal at receiver was exactly identical to the original applied speech waveform. Hence it provides the high security scrambled speech signal and the reconstructed signal was perfect . Some interesting points can be mentioned here:

a. It is clear that (SNRs & SEGSNRs) give small values at any decomposition level, while (SNRd & SEGSNRd), give large values. As the level decreases the system performs better. The absolute low values of distance measures does not necessarily mean a perceptually poor assessments. The distance measures (SNR and SEGSNR) for scrambled/descrambled speech, can in some cases, be used for design purposes as a relative number of intelligibility loss or speech quality.

b. The spectrogram is used because it is a powerful tool that allows us to see what's happening in the frequency and time domains all at once. Thus we can easily see the theory at work here by observing the original signal, it's scrambled version, and the descrambled version. Note that on the scrambled plot it is observed that the order of the frequencies has changed. And, as expected the descrambled version has been correctly decoded to its original form.

c. An evaluation of the speech scrambling system with different power levels of the additive white Gaussian noise was tested. The results proved that as the signal to noise ratio increases, the correspondence between original and descrambled speech increases. Hence, it can be concluded that, the WT algorithm can be implemented to scramble and descramble speech with high efficiency.

d. For real time speech scrambling it is recommended to use a wavelet with a small number of order at a reasonable decomposition level (level **3** decomposition or less), because the number of coefficients required to represent a given signal increases with the level of decomposition (higher wavelet decompositions requires more computation time, which should be minimized for real time speech scrambling) and with the large number of order.

6. References

Bopardikar, A. S. (1995). Speech Encryption Using Wavelet Packets. Indian Institute Of Science.

DeelRe, E.; Fantacci, R. & Maffucci, D. (1989). A New Speech Signal Scrambling Method For Secure Communications: Theory, Implementation, And Security Evaluation. IEEE Journal On Selected Areas in Communications, Vol.7, No.4.

Gersho, A. & Steele, R. (1984). Encryption of Analog Signals a Perspective. IEEE Journal on Selected Areas in Communications, Vol. SAC-2, No.3

Goldburg, B.; Dawson, E. & Sridharan, S. (1991). The Automated Cryptanalysis Of Analog Speech Scramblers. Advances in Cryptology: Proceeding of EUROCRYPT'91, New York: Springer Verlag.

Goldburg, B.; Dawson, E. & Sridharan, S. (1993). Design And Cryptanalysis Of Transform-Based Analog Speech Scramblers. IEEE Journal On Selected Areas in Communications, Vol. 11.

Goldburg, B; Sridharan, S. & Dawson, Ed. (1993). Design And Cryptanalysis Of Transform-Based Analog Speech Scramblers. IEEE Journal On Selected Areas in Communications, Vol. 11, No.6.

Graps, A. (2011). An Introduction to Wavelet. IEEE Computational Science and Engineering, http://www.amara.com/IEEEwave/IEEEwavelet.html

Gray, R; Buzo, A. & Matsuyama, Y. (1980). Distortion Measures for Speech Processing. IEEE Trans. Acoustics, Speech and Signal Proc., Vol. ASSP-28, No. 4.

Gray, A. & Markel, J. (1976). Distance Measures for Speech Processing. IEEE Trans. Acoustics, Speech and Signal Proc., Vol. ASSP-24, No. 5.

Guo, Da & Lin Q (2010). Fast Decryption Utilizing Correlation Calculation for BSS-based Speech Encryption System. Sixth International Conference on Neural Computation.

Lee, L. (1985). A Speech Security System Not Requiring Synchronization. IEEE Communications Magazine, Vol.23, no.7.

Mascarin, A. (2000). Wavelet Toolbox-Featured Product. The Math Works Inc. , Natick, MA, http://www.mathworks.com/products/wavelet.

Mermoul, A & Belouchrani, A. (2010). A Subspace – Based Method for Speech Encryption. Int. Conf. On Information Science, signal Processing and their applications (ISSPA 2010).

Mosa, E.; Messiha, N. & Abd El-Samie, F. (2010). Encryption of Speech Signal with multiple secret keys in time and transform domain. Int. J. Speech Technol., 13.

Pichler, F. (1983). Analog Scrambling By The General Fast Fourier Transform. Department of System Science.

Rao, N. & Homer, J. (2001). Speech Compression Using Wavelets. Electrical Engineering Thesis Project.

Sadkhan, S. B, Khaged, N. H., & Al-Saadi, L. H. (2005). A Proposed Speech Scrambling System Based On Wavelet Transform And Permutation. IEEE Communication And Signal Processing, Vol. 3.

Sadkhan, S.; Falah, N. (2007). A Proposed Analog Speech Scrambler Based on Parallel Structure of Wavelet Transform. M.Sc. Thesis, AlNahrain University, IRAQ.

Sakurai, K; Koga, K. & Muratani, T. (1984). A Speech Scrambler Using The Fast Fourier Transform Technique. IEEE Journal on Selected Areas in Communications, Vol. SAC-2, No.3

Sridharan, S.; Dawson, E. & Goldburg, B. (1990). Speech Encryption In The Transform Domain. Electronics Letters, Queensland Univ. of Technol., Vol. 26.

Sridharan, S.; Dawson, E. & Goldburg, B. (1991). Fast Fourier Transform Based Speech Encryption System. IEE Proceedings-I, Vol. 138, No. 3.

Whei, W. & Iang, H. (2000). The Automated Cryptanalysis of DFT-Based Speech Scramblers. IEICE Transactions on Information and Systems, Vol. E83-D, No.12.

Yuan, Z. (2003). The Weighted Sum of The Line Spectrum Pair for Noisy Speech. M.Sc. Thesis, Department of Electrical and Communications Engineering, Helsinki University of Technology.