Secure Speech Transmission Using Chaos Encryption System for AMR-WB Codec

Messaouda Boumaraf, Fatiha Merazka
LISIC Laboratory, Telecommunications Department
USTHB University, Algeria

Abstract

In recent years, the need for secured communication research based on Speech Coding with chaos cryptography has highly increased. Therefore, it is absolutely necessary to have secure systems to protect the transferred data against arbitrary intrusions. So, data encryption is an effective way to meet these requirements. In this paper, we propose a full and a partial speech encryption schemes based on chaos maps for AMR-WB G.722.2 Codec. To increase the security level, we combine chaos maps logistic and Hénon for shuffling and scrambling speech in order to have a hybrid chaotic key generator. The proposed algorithms evaluated with both Perceptual Evaluation of Speech Quality (PESQ) and Enhanced Modified Bark Spectral Distortion (EMBSD) measure confirm the efficiency of our proposed cryptosystem scheme.

1. Introduction

The science of the 20th century has been marked by three major discoveries. The relativity, the quantum mechanics and chaos. The fields of use of chaos theory have touched wide range on several disciplines such as sociology, physics, computer science, engineering, economics, biology.

As a matter of fact, the birth of chaos theory began with works of Henri Poincaré (1854-1912) at the end of the 19th century, by studying the three-body problem in celestial mechanics. It was to determine the trajectory of three interacting body according to the law of universal gravitation. So, Poincaré showed that some dynamical nonlinear systems had unpredictable behaviors. A century later, Edward Lorenz (1917-2008), a meteorological researcher is the official father of chaos theory. In 1961, he used a three-dimensional differential equation system to model the convection of atmosphere. he had just discovered the chaotic behavior of a non-linear system, of which minute differences in the initial conditions of a deterministic system led to completely different and unpredictable results, a phenomenon generally illustrated by the butterfly effect [1].

Later, several models are derived from the Lorenz model; among these, we distinguish a class of chaotic systems called generalized systems, these include several models: The systems of Chen, Chua, Lorenz, Lü etc... Now, some typical new chaotic and hyperchaotic systems appear, such as modified Lorenz system [2], Lü system [3], Rossler [4].

Remember, that Chaos theory has applications in several disciplines. However, our study will focus on the use of chaos to secure information. The security of digital information, such as voice data, is now an important issue for all internet users since the worldwide use of the internet has become very apparent. Therefore, the data between the legitimate users need to be secured before transmission by using encryption methods. With the significant computer's cryptography developments, numerous studies are involved to secure communications [5]. Recently, the amount of research on chaotic cryptography increased more and more in order to improve chaos-based cryptosystems. Chaos-based encryption algorithms are based on different types of chaotic maps whether discrete or continuous maps. Most of them are a combination of two or more chaotic maps to achieve a better security, expanded key space and low complexity.

A chaotic is a non-linear, deterministic system. It presents good properties such as pseudo-randomness, sensitivity to changes in initial conditions, system parameters and aperiodicity which makes it unpredictable. Because of its characteristics, the chaos was used in the encryption system. An adversary is not allowed to find the outputs without any knowledge of the initial values [6].

The introduction of the concepts of cryptosystems is followed by the mapping of the two theories: the chaos and the cryptography. We will highlight the importance of using dynamic chaotic systems in cryptography.

Many chaos-based encryption methods have been introduced during the last decade. The most cited and important chaos-based structure was presented by Fridrich in [7] using substitution and permutation. In [8] Wen has reviewed the dynamic properties of the Hénon map including its fixed points, stability, periodic orbits, and so on. He has also, discussed its physical interpretation. In [9] authors proposed a new image encryption algorithm based on the parameter-

varied logistic chaotic map and a dynamical algorithm. They used the parameter-varied logistic map to shuffle the plain image, and then used a dynamical algorithm to encrypt the image. In [10] authors introduce a speech encryption approach, which is based on permutation of speech segments using chaotic Baker map and substitution using masks in both time and transform domains.

In order to use chaos theory effectiveness in cryptography, the chaotic maps should be implemented such that the entropy generated by the map can produce required confusion (data value permutation) and diffusion (data value modification) architecture proposed by Shannon.

Today, based on some important properties of chaos, such as the unpredictable behavior which can be used in the generation of random numbers. The chaos-based cryptosystems provide several advantages, such as: very high-security level, high speed especially in stream ciphers, computational power, and are easier to implement. These features make them more suitable for large scale-data encryption, such as voice and image. In the other hand, these characteristics can be refined to simulate the characteristics of a white noise or other random signal, which makes chaos a very interesting phenomenon for hiding information signals in order to transmit them in a secure manner. In other words, the encryption of a data by chaos is done by superposing the initial information at the chaotic signal. Afterwards, the data drowned in chaos are sent to a Receiver which knows the characteristics of the chaos' generator. The receiver needs only to subtract the chaos from the data in order to retrieve the information [11,12].

The chaos streams are generated by using various chaotic maps.

In this paper, two chaos-based cryptosystems of transmitted speech security are presented, and the obtained results are discussed.

The remainder sections of this paper are organized as follows. In section 2, an overview of the AMR-WB G.722.2 is introduced. Section 3, our proposed cryptosystem is presented. Simulations and interpretation of obtained results are discussed in section 4. Finally, the conclusion is provided in section 5.

2. Overview of the AMR-WB G.722.2

There are various kinds of speech codec available. These varied are differentiated from each other based from there algorithm, technology, bandwidth, data rates etc...

An adaptive Multi-Rate Wideband (AMR-WB) is an apparent wideband speech audio coding standard enhanced and based on Adaptive Multi-Rate encoding, using a similar methodology as algebraic code excited linear prediction (ACELP). AMR-WB gives improved speech quality due to a larger speech bandwidth of 50–7000 Hz compared to narrowband speech coders which optimized for POTS (Plain Old Telephone Service) wireline quality of 300–3400 Hz. AMR-WB was upgraded by Nokia and VoiceAge and it was first defined by 3GPP [13].

AMR-WB is codified as G.722.2, an ITU-T standard speech codec, formally known as Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB). G.722.2 AMR-WB is the same codec as the 3GPP AMR-WB.

The AMR-WB speech codec contains nine bits rates of 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.6 kbps. These ones are presented by modes 8, 7, 6, 5, 4, 3, 2, 1 and 0 respectively. The bit rate can be changed at any frame boundary of 20 ms. The codec includes Voice Activity Detection (VAD), Discontinuous Transmission (DTX) and Comfort Noise Generation (CNG) features for increased efficiency [13].

The AMR-WB G722.2 uses six parameters to represent the speech and these are shown in Table .1 for bit rate 8.85kbit/s [13].

Table 1. G.722.2 – Bit allocation of the AMR-WB coding algorithm for 20-ms frame

Mode 1 (8.85kbit/s)	VAD-flag					1
	ISP					46
	Pitch delay	8	5	8	5	26
	Algebraic code	20	20	20	20	256
	Gain	6	6	6	6	24
	Total					**177**

3. The cryptosystems

A cryptosystem is pair of algorithms that take a key and convert plaintext to ciphertext and vice versa. The cryptosystems aim is to make the coded data incomprehensible to any curious person different from the legitimate recipient. Fundamentally, cryptosystems can be classified into the following categories: symmetric systems, asymmetric systems and hybrid systems [7].

3.1. Proposed Cryptosystem based chaos

In this work, we employed two kind of encryption full and selective. Both of them use two types of chaotic maps: 2D Hénon map and 1D logistic map. Each one has its own property or characteristic and has its own effect on improving the performance of evolutionary algorithm. In General, the information in secure communication is transmitted through the channel after source encoding, encryption and

channel encoding and modulation, then it will be received by reversing these steps The principle diagram of our proposed cryptosystem is depicted in Figure 1. As shown, firstly, the analog input signal at the transmitter undergoes digital-analog conversion, then, the coder AMR-WB operates on speech frames of 20 ms. At each frame, the speech signal is analyzed to extract the parameters of the CELP model (adaptive and fixed codebooks' indices, gains....). These parameters are encoded. Some or all digital samples are encrypted and transmitted. At the receiver level, we will reverse the encryption process i.e. perform decryption. At the decoder, these parameters are decoded, and speech is synthesized. In the following we will describe Logistic map, Hénon map and then our proposed algorithm.

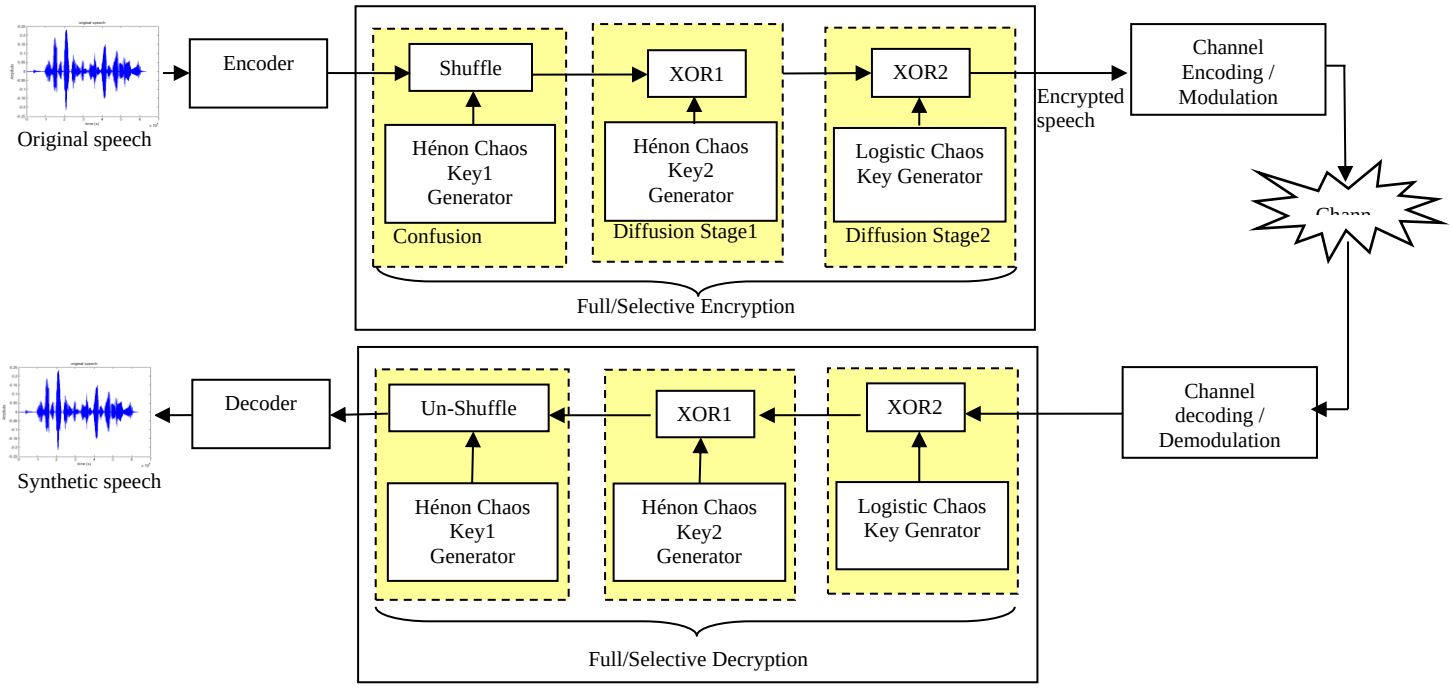


Figure 1. Secure communication based on our proposed scheme

3.1.1. The logistic map. Is mathematically simple, introduced by the biologist Robert M. May in 1976 [14]. It is the discrete time solution of the demographic model analogous to the logistic equation first created by Pierre François Verhulst (1845,1847). The map is a prototypical one-dimensional invertible iterated map represented by the state equations with a chaotic attractor that exhibits complicated behavior. Its recurrence equation is given by [15]:

$$x_{k+1} = \mu x_k(1 - x_k) \quad x_k \in (0,1) \qquad (1)$$

where, μ belongs to the interval [0,4] and this parameter determines the map behavior. When parameter μhas the following range ($3.57 < \mu \leq 4$) it becomes a chaotic map, and x_k belongs to the interval [0,1], knowing that any change in initial value or parameter μ will give various sequences of random and irregular numbers. In our case, we set the value of the control parameter to the value corresponding to $\mu = 4$ and $x_0 = 0.28$.

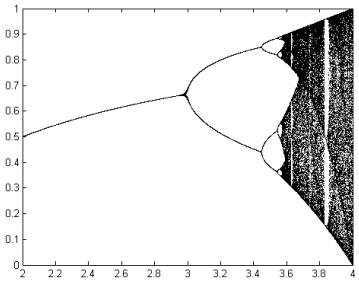


Figure 2. Bifurcation diagram of the logistic map

Remember that the Chaos can be generated by any non-linear dynamic system. Indeed, simple recurrence equations are capable of creating rich chaotic dynamics, if the parameters are well

situated. In many recurrence simple equations, the right choice of these parameters is made by means of the bifurcation diagram and the exponent of Lyapunov. We present the known curves of Bifurcation diagram and the Lyapunov exponent of the map in Figure 2 and Figure 3 respectively.

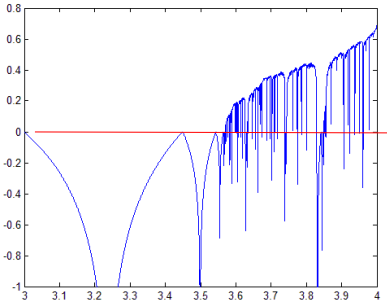


Figure 3. Lyapunov exponent of the logistic map

3.1.2. The Hénon map: is a model proposed in 1976 by the mathematician Michel Hénon. It is a prototypical two-dimensional invertible iterated map [16]. It has a chaotic behavior, and can be expressed as a recurrence of two chaos signals given by [17]:

$$\begin{cases} x_{n+1=}1 + y_n - ax_n^2 & (a) \\ y_{n+1} = bx_n & (b) \end{cases} \quad (2)$$

To get a random sequence, we used for the parameters 'a' and 'b' the following values. $a = 1.2$, $b = 0.1$. with these values, the initial point is $(x_0, y_0) = (0.1, 0.1)$, the sequence of points is obtained by the mapping's iteration and it tends to a strange attractor. The chaotic Hénon map is shown in Figure 4.

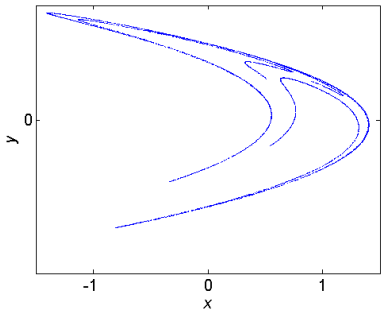


Figure 4. Chaotic behavior of Hénon attractor

3.1. Proposed speech encryption's algorithm

In this subsection, we present the steps of our proposed cryptosystem speech encryption algorithm based on confusion employing Hénon map and diffusion employing both Hénon and logistic map. It is divided in three major steps:

Step1: In the confusion stage, the parameters of frames are shuffled and permuted by using formula (2-a) of Eq. (2). Because, the key xn takes values from interval $]-1.5,1.5[$, as shown in Figure 5-a, we arrange them in an increasing or decreasing order as shown in as shown in Figure 5-b. After sorting, we save the position or the index of each key values then, change the position of speech data according to indexes' keys [16].

Step2: In the diffusion stage 1, the permuted parameters of frames are substituted employing XOR operation using formula (2-b) of Hénon Eq. (2). The keys yn takes a value from the interval $]-1,1.[$, so we calculate:

$$floor(abs(yn) * 106 \; mod \; 65536 \quad (3)$$

The diffusion is performed using XOR operation.

Step3: In the diffusion stage 2, we generate keys by using logistic map. We obtain a series of numbers $k_1, k_2, k_3, \ldots k_n$ in the range [0,1]. Where n is the number of words in the speech to be encrypted. Then, we follow these steps:

- We choose the length of the key with 16 bits because our AMR-WB G.722.2 words are 16 bits. Then, we multiply each number by65536. Let
$$key_i = int(k_i * 65536 + 0.5) \quad (4)$$
- Finally, words of speech data are modified by employing XOR operation.

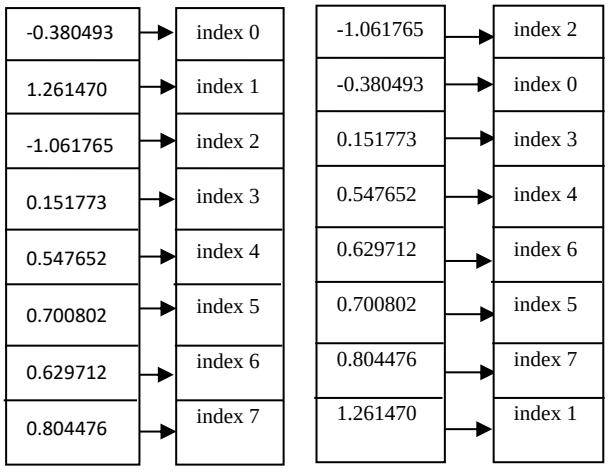


Figure 5 (a): samples of generated keys
(b): keys arranged in an increasing order

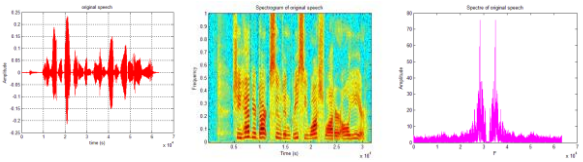
4. Simulation and Results

In this section, we present the simulation setup followed by the obtained results. Several experiments

are carried out to test the encryption efficiency of the presented wideband speech cryptosystem. The quality of both the encrypted and reconstructed signals is assessed for the standard AMR-WB G.722.2.

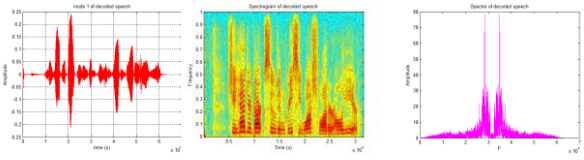
The speech file extracted from TIMIT database [18] and sampled at 16 kHz was encoded using AMR-WB G.722.2 CS-ACELP. The resulting bit streams were encrypted using Hénon and Logistic maps schemes. In the experiments, signal inspection in both the time and frequency domains is done to evaluate the changes between the original and encrypted speech. So, we represented speech signal in different visual representations. First in waveform representation which allow us to show the changes in peak distance over time. Peak distance is more commonly known as amplitude, then in spectrogram representation we perform series of spectral analyses at different times and then display them using a three dimensional display of time, frequency and amplitude. In most cases time is displayed on the X-axis, frequency is displayed on the Y-axis and amplitude is displayed as variations on grayscale darkness of color. Spectrograms of audio can be used to identify spoken words phonetically. Finally, with spectrum representation, we can determine different frequency components present in a speech signal.

Before starting to encrypt our signal, we first represent the original and decoded speech as shown in Figure 6 and Figure 7 respectively. We can see that the original and decoded speech seem identical for waveforms (Figures 6-a and 7-a), spectrograms (Figure 6-b and 7-b) and spectrums (Figure 6-c and 7-c) representations.



(a) **(b)** **(c)**

Figure 6. (a) Original speech, (b) its Spectrogram (c) its Spectrum



(a) **(b)** **(c)**

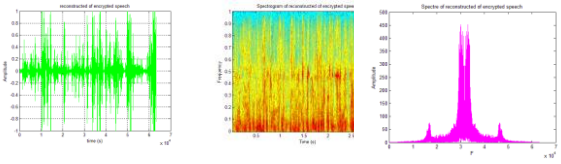
Figure 7. (a) Decoded speech using mode 1(b) its Spectrogram (c) its Spectrum

In our work, we have employed two kinds of encryption full and selective.

4.1. Cryptosystems' representations: full encryption case

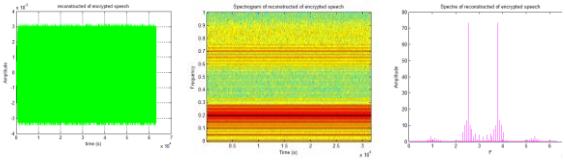
For the full encrypted speech, we have used three kinds of encryption only confusion, only diffusion and both confusion and diffusion given in Figure 8, Figures 9 and 10 respectively.

- By analyzing by analyzing their waveforms representation, we can notice that the use of only confusing (Figure 8-a), the information is present but different from the original speech signal. However, the use of only diffusion (Figure 9-a) or both confusion and diffusion (Figure 10-a) leads to signals that are comparable to a white noise, which indicates that no significant residual intelligibility can be useful for eavesdroppers at the communication channel.
- Now, by analyzing their spectrogram shown in Figure 8-b it appears, that using only confusing the information is also present but is greatly different compared to the original speech signal. nevertheless, the use of only diffusion (Figure 9-b) or both confusion and diffusion (Figure 10-b) the information is absolutely absent.
- In Figure 8-c, it appears that using only confusing, a lot of information are present that refer to the original speech signal. However, the use of only diffusion (Figure 9-c) or both confusion and diffusion (Figure 10-c) the information is absolutely absent.



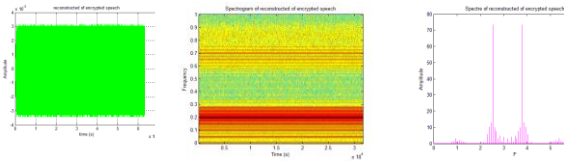
(a) **(b)** **(c)**

Figure 8. (a) Speech encryption using only confusion (Hénon map) (b) its Spectrogram (c) its Spectrum



(a) **(b)** **(c)**

Figure 9. (a) Speech encryption using only diffusion (Hénon & Logistic Maps) (b) its Spectrogram (c) its Spectrum



(a) (b) (c)

Figure 10. (a) Speech encryption using our Proposed scheme b) its Spectrogram (c) its Spectrum

We emphasized that the reconstructed speech signals using the right keys are identical to the original for all cryptosystems full or partial encryption.

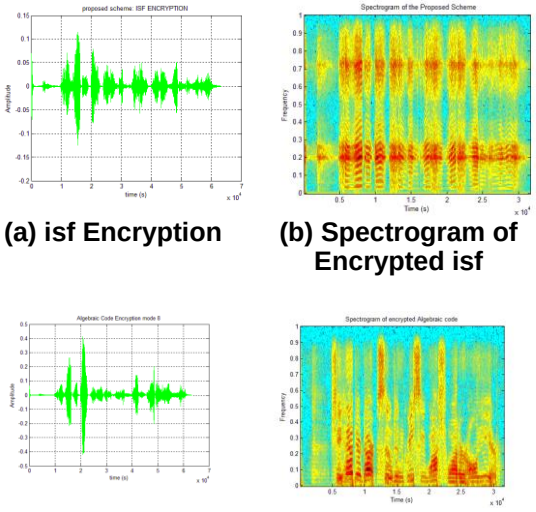
4.2. Cryptosystems' representations: selective encryption case

For the selective encrypted speech, we have used our proposed cryptosystem based on both confusion and diffusion. The parameters (ISF, Algebraic code, Pitch delay and gain) obtained via the coder will be encrypted separately.

From the waveforms (see Figure 11-a) and spectrograms (see Figure 11-b) representations, we can see that in the partial encryption, a signal have been modified with few visible differences from the original speech signal (see Figure 6). But we don't know if it's intelligible or not and we cannot confirm the effectiveness of encryption. However, we need to use an objective measurement to evaluate effectively our cryptosystem.

4.3. Measurement of speech quality

The Enhanced Modified Bark Spectral Distortion (EMBSD) [19] is an objective measurement tool, used to evaluate the efficiency of encryption schemes.



(a) isf Encryption **(b) Spectrogram of Encrypted isf**

(c) Algebraic code Encryption **(d) Spectrogram of Encrypted Algebraic code**

(e) Pitch delay Encryption **(f) Spectrogram of Encrypted Pitch delay**

(g) Gain Encryption **(h) Spectrogram of Encrypted Gain**

Figure 11. Selective Encryption using our Proposed scheme

The obtained results from tests with EMBSD are given in Figure 12 for the full encrypted speech. It is recalled that the EMBSD gives a value of 0 for two identical speech files, and a greater value as the distortion increases. So, we can see that the best values are given for the original speech coder and the worst are given for encryption using both confusion and diffusion, i.e. hybrid chaotic generator and for using diffusion only.

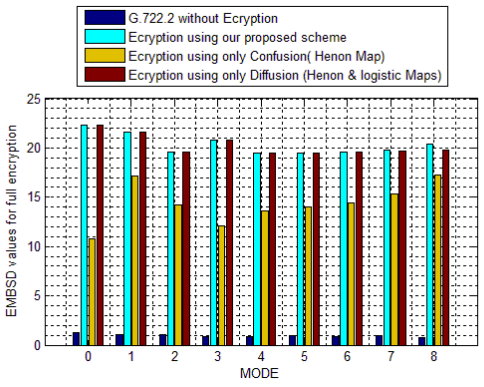


Figure 12. Results with EMBSD evaluation

We have also evaluated the performance of our cryptosystem using the Perceptual Evaluation of the

speech quality (PESQ) [20]. Results are given in Figure 13 for original speech, full encrypted speech using confusion only, encrypted speech using diffusion only and our proposed cryptosystem. It is recalled that the PESQ gives a value of 4.5 for two identical speech files and a less value as the distortion increases. We can see, again, that the best PESQ is given by the original speech coder since it is not encrypted and the worst PESQ values are obtained for encryption using both confusion and diffusion i.e. hybrid chaotic generator and for using diffusion only.

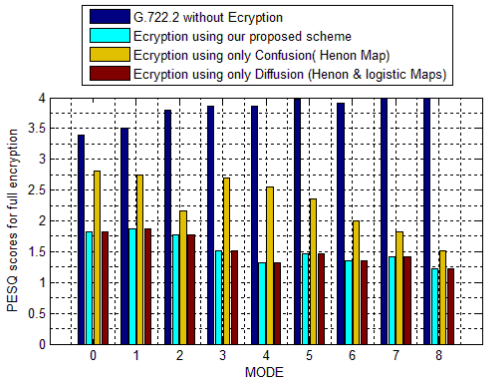


Figure 13. Results with PESQ evaluation

So, both metrics confirm the efficiency of our cryptosystem based on hybrid chaotic maps. They confirm also that confusion does not affect encryption using diffusion with Hénon and logistic maps since results are comparable for encryption with diffusion or with confusion and diffusion.

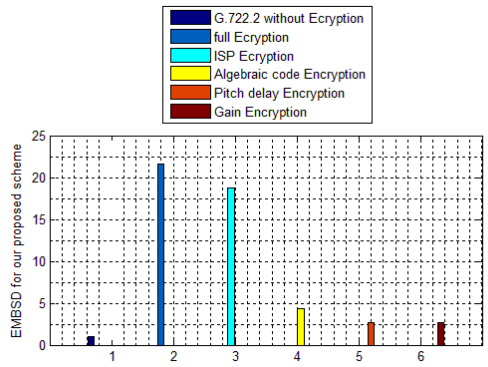


Figure 14. Full and selective encryption results of EMBSD evaluation for our proposed cryptosystem scheme using the codec in mode 1

For partial encryption, we have used our proposed cryptosystem, i.e. using both confusion and diffusion and applied the encryption of some parameters of speech coder in mode 1. So, we use EMBSD tool to evaluate the encryption efficiency as depicted in Figure 14.

We can confirm that full encryption gives the best secure encryption compared to the partial encryption. In the other hand, the use of partial encryption show that the best protection is given for the ISP parameters. With selective encryption, we can know the important bits, and therefore, act on the best parameter to have a best secure and efficient encryption. So, we can conclude, that the selective encryption using ISP parameters compete a full encryption.

4.4. Timeliness of cryptosystem

Tests have been carried out to compute the time required to execute our proposed cryptosystem scheme. To estimate this time, we repeat the execution of each of the 1000000 times (iterations) and then we divide the duration obtained over 1000000 for full and partial encryption.

Results for encryption and decryption using our full proposed cryptosystem for all mode of codec are depicted in Figure 15. Results show that the run time for encryption and decryption is almost the same. We can observe that our cryptosystem is fast and does not exceed hundreds of nanoseconds.

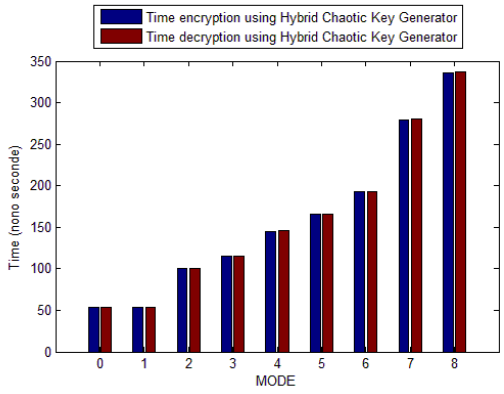


Figure 15. Runtime Per Nano Seconds for full encryption in all mode

Results for encryption and decryption using full and partial proposed encryption and decryption of codec in mode 1 are depicted in Figure 16 and Figure 17 respectively. From these Figures we can see that full encryption or decryption take more time than partial encryption. Comparing partial encryption or decryption with each other, we can notice gain encryption or decryption is the fastest.

We can conclude that partial encryptions or decryptions are more fast than full encryption or

decryption and they do not exceed dozens of nanoseconds.

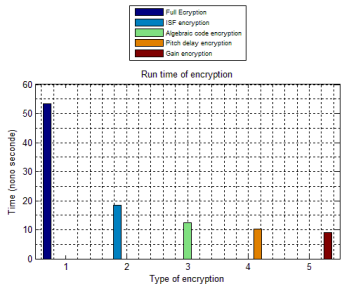


Figure 16. Runtime Per Nano Seconds for full and selective encryption mode 1

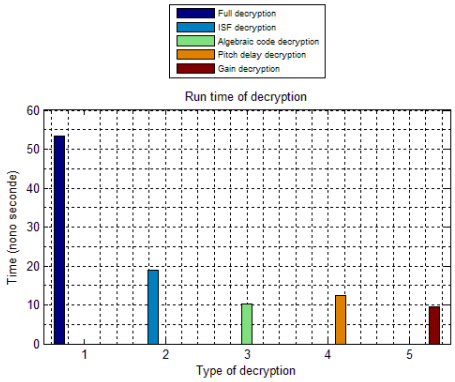


Figure 17. Runtime Per Nano Seconds for full and selective encryption mode 1

5. Conclusion

In this paper, a hybrid chaotic key generator was proposed. It was implemented through two simple chaotic maps to encrypt and decrypt speech for AMR-WB G.722.2 Codec in favor of secure and fast cryptosystem. The encryption scheme produces both confusion and diffusion as Shannon's principle requires. So, our cryptosystem, is designed to shuffle the position of all or some parameters of frame in order to produce permutations followed by using two substitutions. The experimental results and analysis show that, full cryptosystem is the most effective in terms of security. However, this approach has a rather long execution time compared to partial encryption. On another side, the selective schemes permitted us to know the important parameters that must be encrypted. Moreover, it is more efficient in term of execution time. In the future, we propose to use hyperchaos cryptosystem which possesses a large range of parameters which makes the system strong and more secure.

6. References

[1] E. N. Lorenz. "Deterministic nonperiodic flow", *Journal of Atmospheric Sciences*, 20 : 1963, 130-141.

[2] Y. Chen, Q. Yang, "A new Lorenz-type hyperchaotic system with a curve of equilibria", Math. Comput. Simul., 2015, 112:40–55.

[3] A. Chen, J. Lu, J. Lü, S. Yu, "Generating hyperchaotic Lü attractor via state feedback control", Physica A, 2006, 364:103–110.

[4] O.E. Rössler, "An equation for continuous chaos", Phys. Lett. A, 57(5), 1976, 397–398.

[5] W. Chang, "Digital secure communication via chaotic systems", Digital Signal Processing, 19(4), 2009, 693–699.

[6] F. Merazka, "Wideband Speech Encryption based Arnold Cat Map for AMR-WB G.722.2 codec", ICISP 2014: 658-664.

[7] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps", International Journal of Bifurcation and Chaos, 8(06), 1998,1259–1284.

[8] H. Wen, "A review of the Hénon map and its physical interpretations", advance online publication: 30 November 2014.

[9] L. Liu, and S. Miao, "A new image encryption algorithm based on logistic chaotic map with varying parameter",SpringerPlus (2016) 5:289.

[10] E. Mosa, N.W. Messiha, O. Zahran, F. E. Abd El-Samie "Chaotic encryption of speech signals", Int J Speech Technol (2011) 14:285–296.

[11] R. E. BORIGA, A. C. DĂSCĂLESCU, and A. V. DIACONU, "A New Fast Image Encryption Scheme Based on 2D Chaotic Mzps", IAENG International Journal of Computer Science, 30 November 2014.

[12]G. Alvarez and S. Li, "Some basic cryptographic requirements for chaos-based cryptosystems", Int. J. Bifurcation Chaos 16, 2006, 2129.

[13] "Wideband coding of speech at around 16 kbps using Adaptive Multi-RateWideband (AMR-WB) ", ITU-T Standard G.722.2, 2003.

[14] M. Robert. "Simple mathematical models with very complicated dynamics". *Nature.* 1976; 261:459–467.

[15] K. J Aval, M. S. Kamarposhty and M. Damrudi, "A Simple Method for Image Encryption Using Chaotic Logistic Map", Journal of Computer Science & Computational Mathematics, Volume 3, Issue 3,September 2013.

[16] S. B. Sadkhan and H. Ali, "A proposed SpeechScrambling based on hybrid chaotic key generators", Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA) – IRAQ, May 2016: (9-10).

[17] M. Prasad and K.L.Sudha, "Chaos image encryption using pixel shuffling with henon map", Elixir Elec. Engg. 38 (2011) 4492-4495.

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

[19] W. Yang, "Enhanced Modified Bark Spectral Distortion (EMBSD): An Objective Speech Quality Measurement Based on Audible Distortion and Cognition Model", Ph.D Dissertation, Temple University, USA, May 1999.

[20] ITU-T Draft Rec, "Perceptual evaluation of speech quality (PESQ), an objective method of end-to-end speech quality assessment of narrowband telephone networks and speech codecs", May 2000. P.862.

