Diyala Journal of Engineering Sciences

Journal homepage: https://djes.info/index.php/djes

ISSN: 1999-8716 (Print); 2616-6909 (Online)

Performance Comparison between RSA and El-Gamal Algorithms for Speech Data Encryption and Decryption

Sura F. Yousif*

Department of Chemical Engineering, University of Diyala, 32001 Diyala, Iraq

ARTICLE INFO

Article history:
Received October 4, 2022
Accepted February 24, 2023

Keywords:
Encryption/Decryption
Speech signal
Asymmetric cryptosystem
RSA algorithm

ABSTRACT

This article presents a performance comparison of two known public key cryptography techniques namely RSA (Rivest–Shamir–Adleman) and El-Gamal algorithms to encrypt/decrypt the speech signals during transferring over open networks. Specifically, this work is divided into two stages. The first stage is enciphering-deciphering the input speech file by employing the RSA method. The second stage is enciphering-deciphering the same input speech file by employing the El-Gamal method. Then, a comparative analysis is performed to test the performance of both cryptosystems using diverse experimental and statistical analyses for the ciphering and deciphering procedures like some known speech quality measures: histogram, spectrogram, correlation, differential, speed performance and noise effect analyses. The analyses outcomes reveal that the RSA and El-Gamal approaches are efficient and adequate for providing high degree of security, confidentiality and reliability. Additionally, the outcomes indicate that the RSA speech cryptosystem outperforms its counterpart the El-Gamal speech cryptosystem in most of ciphering/deciphering speech performance metrics.

1. Introduction

Security of the transmitted data across telecommunication channels represents a major challenge. Speech messages can be transferred via many means like private and public wireless networks or telephone networks. A mong all methods of communication, speech is the most common way of human interaction. Speech communication is a form of verbal or oral communication. It is efficient, naturalistic, flexible and simple; therefore, speech communication is utilized in most practical activities of our daily life like banking, e-learning, commerce, politics, education, and other fields. Further, certain attitudes require the transportation of confidential speech data including military and diplomatic telecommunication through times of peace and war. Ensuring speech security is quite difficult because the speech file includes a lot of redundant data in comparison to the videos, digital images and text messages [1]. Traditional techniques are vulnerable to the intruder attacks with the evolution of vigorous computers. The fast and increased developing of data transmission over the Internet and sharing networks requires strong and reliable security methods to provide privacy and to prevent the illegal access to the transmitted message content. Among many solutions, encryption is employed to safeguard the secret data during transferring in insecure channel [2]. Encryption algorithms convert the multimedia data from its readable form (plaintext) to invisible form (cipher text) to increase the security and secrecy. Decryption operation is applied to restore the

* Corresponding author.
E-mail address: sura.fahmy@yahoo.com
DOI: 10.24237/djes.2023.16112

original message at the receptor [3]. Generally, encryption can be categorized into two major sorts: symmetric key encryption and asymmetric key encryption. In the symmetric key encryption, one unique key is shared privately between the sender and recipient for the encryption and decryption processes. Data Encryption Standard (DES), Advanced Encryption Standard (AES), Rivest Cipher 4 (RC4) and Blowfish algorithms are examples of symmetric key encryption. In the asymmetric key encryption or public key encryption, two different keys are employed: public and private. The public key is known to everyone and it is utilized for the encryption, while the private key is kept secret and it is utilized for the decryption. RSA, El-Gamal and Diffie Hellman algorithms are examples of asymmetric key encryption. In the public key cryptosystems, anyone can encrypt the message, but only the person who possesses the corresponding private key can decrypt it [4]. The asymmetric key encryption solves the problem of key distribution because there is no necessity to share the secret key between the communicating participates. Furthermore, the asymmetric key encryption depends on trap door or one-way functions. These mathematical functions can be simply calculated in one direction, while in the inverse direction, they are very difficult to solve unless the secret key is found. Hence, the asymmetric encryption scheme can provide more security than the symmetric scheme because it employs two keys [5].

Because of the complexity of encryption algorithms, they need to be utilized upon flexible platforms to meet the real-time speech encryption demands [1, 6]. Moreover, there is a risk of speech data leakage in the transmission operation. For this reason, speech file encryption is of great importance and many researchers focus on speech cryptographic mechanisms and they proposed several speech cryptosystems that based on different techniques for secure speech communications. For instance, the study in [6] combines chaotic maps and k-means clustering technology for ciphering the speech files. Two permutation steps are utilized in this system. The first step depends on binary representation shuffling mechanism, whereas the second step relies upon k-means principle. The introduced cryptosystem is assessed via various speech quality metrics. The input speech data is compressed in [7] to ensure the signal quality. Then, the compressed information is encrypted by utilizing chaotic map and Fuzzy means method to obtain the final cipher signal. Several chaotic systems are adopted in [8] to achieve the encrypting process at the sender side. Hashing and blowfish algorithms are also employed to increase the speech system security. The authors in [9] designed a chaotic speech encryption approach that based on three stages. The speech samples are scrambled in the first phase followed by implementing Deoxyribonucleic Acid (DNA) code in the second phase in order to flip the sample bits. Substitution process is executed in the eventual phase to accomplish the encryption procedure. The method is validated by utilizing a variety of measurements. The researchers in [10] developed a speech cryptographic model relies on voice over Internet protocol and chaos concept so as to protect the data throughout transferring. The chaotic map generates the key stream to encipher each speech data in the packet. The suggested model is evaluated via sundry encryption/ decryption performance criteria. Chen chaotic system and fast Walsh Hadamar technology are integrated in [11] to encode the speech signal. The audio content in the method is converted into rectangular functions. After that, shuffling/diffusion architecture is carried out to realize the encrypting operation. Modified chaotic map is designed in [12] by merging two classical chaotic maps. The random sequence produced by the new map is used to perform the confusion-diffusion encryption structure upon the input speech file. The plain speech sample is encrypted/decrypted in [13] via elliptic chaotic map. The map is created firstly based on its initial-control parameters. Then, the gained sequences from the chaotic map are employed for encrypting the speech message. A cryptographic scheme is studied in [14] to encipher the digital speech information. The data are scrambled by rearranging the files as cubic sample of six sides. The next step is applying two different maps: Gingerbread and

Hénon chaotic maps to attain the encryption stage. Many standard statistical tests are performed to measure the method quality. The audio and speech signals are fused in [1] to encrypt the audio frames during communication. After this, the data is encrypted via chaotic mapping by utilizing two layers: substitution and permutation to acquire the ultimate ciphered signal. Various assessment tests are performed to quantify the cryptosystem performance.

Based on the above literature review, it is obvious that the researchers proposed many solutions and suggestions in order to protect the important and confidential speech files during the transmission. However, adopting chaotic maps or chaotic maps merged with other techniques is not always the best solution for speech encryption. Speech cryptosystems based on low dimension chaotic maps suffer from weak security and little key space. On the other hand, cryptosystems based upon high dimension chaotic maps have some flaws like increasing the implementation cost and computation complexity, and decreasing the ciphering speed. Moreover, many chaotic encryption approaches can be attacked by some cryptographic analyses [15, 16]. Therefore, a simple method should be developed to encrypt speech signal for meeting the real-time application requirements.

So as to conquer the above shortcomings, an effective method for speech data encryption/decryption using asymmetric key algorithms is introduced in this article. This scheme is partitioned into two parts. The first part deals with encrypting and decrypting the input speech signal via the RSA technique, whereas the second part deals with encrypting and decrypting the input speech signal via the El-Gamal technique. After executing the RSA and El-Gamal mechanisms on the equivalent speech signal samples, a comparative analysis between the two algorithms is presented based on several different statistical and experimental encryption/decryption analyses, such as common quantitative measures, histogram, spectrogram, correlation coefficient, differential, speed performance and noise influence tests. The numerical and visual outcomes confirm that the methods can be used to transmit data securely with high degree of secrecy. Besides, the RSA mechanism gives better outcomes in comparison with the El-Gamal outcomes in most situations. The main contributions of this work are: (1) Analyzing the two models in order to measure their ability to protect speech data. (2) Evaluating the two encryption techniques performance based on standard speech criteria. (3) Finding the suitable technique to encrypt/decrypt the speech data through simulation.

This article is arranged as follows, Sections 2 and 3 explain the public key systems RSA and El-Gamal encryption /decryption mechanisms, respectively, while Section 4 introduces the presented work. The performance measurements are presented in Section 5 followed by the simulation outcomes for both RSA and El-Gamal algorithms in terms of encryption and decryption stages in Section 6. Finally, the conclusions based on the given results are discussed in Section 7.

2. RSA algorithm

In 1977, Ronald Rivest, Adi Shamir and Leonard Adelman were the first who invented the most widely asymmetric key cryptosystem known as RSA algorithm. It is an encryption and authentication cryptosystem that has been employed since that time in many cryptographic applications such as e-mail security, banking, e-commerce and digital signature on the Internet. The security of this algorithm depends on the difficulty of finding prime factors of large integers. RSA operation consists of three main stages: key generation, encryption and decryption processes which are illustrated briefly as follows [17].

2.1. Key generation

1. Select two large prime integer numbers p and q to calculate the modulus n using the formula $n = p \times q$.
2. Calculate $\emptyset$ using the formula $\emptyset(n) = (p-1) \times (q-1)$.
3. Select an integer e which is the public exponent, such that $\gcd\big(e, \emptyset(n)\big) = 1$, where gcd refers to greatest common divisor function between two numbers.

4. Calculate d which is the private exponent, such that $e \times d = 1 \, mod \, \emptyset(n)$, where mod symbolizes to the modulus operation or the reminder after division.

Hence, (n, e) represents the public encryption key, while (n, d) represents the private decryption key [18].

2.2. Encryption/Decryption Processes

Let m be a message that wanted to be encrypted, then the encrypted message c is calculated via the public key (n, e) using the equation: $c = m^e mod \, n$. To extract the original message m, the received encrypted message c is decrypted via the private key (n, d) using the equation: $m = c^d mod \, n$ [19].

3. El-Gamal algorithm

El-Gamal algorithm provides an alternative method of RSA for asymmetric key encryption. Taher El-Gamal was the first who describe this algorithm in 1984. It has been used in Guard software, free GNU Privacy, PGP recent variations and other different cryptosystems. The security of this algorithm depends on the difficulty of calculating discrete logarithms of large prime numbers. If the same plaintext is encrypted using this cryptosystem, then a different cipher text is obtained in each time of encryption. El-Gamal operation can be described in three main steps: key generation, encryption and decryption processes which are explained briefly as follows [3].

3.1. Key Generation

1. First, select a random prime number p and two other random numbers x and g, such that both of them are less than p.
2. Calculate y using the formula: $y = g^x mod \, p$.

Thus, (p, g, y) represents the public key which can be shared between a group of users, while x represents the private key which should be kept secret [20].

3.2. Encryption/Decryption Processes

To encrypt a message m, firstly, a random integer number k is selected, such that k is relatively prime with $(p - 1)$. Secondly, the cipher text pairs (c_1, c_2) is calculated using the equations: $c_1 = g^k mod \, p$ and $c_2 = (y^k \times m) \, mod \, p$. Finally, the cipher text (c_1, c_2) is transmitted to the recipient. To decrypt the cipher text, pair (c_1, c_2), the private key x is employed to recover the original message m using the equation: $m = \frac{c_2}{c_1^x} \, mod \, p$ [3].

4. Presented work

This speech cryptosystem is divided into two parts: (1) Speech ciphering/deciphering process using RSA algorithm. (2) Speech ciphering/deciphering process using El-Gamal algorithm, these two parts are discussed in this section with details.

4.1. Speech Ciphering/Deciphering Operation Using RSA System

Step 1: Generate the public key (n, e) by the transmitter according to Section 2.1.

Step 2: Reshape the input one-dimensional speech signal $A\,(i)$ into two-dimensional signal $B\,(i, j)$.

Step 3: The speech samples obtained from Step 2 are altered by utilizing the following formulas:

$$C_1\,(i, j) = floor\,(B\,(i, j) \times 10^{14}) \quad (1)$$

$$C_2\,(i, j) = mod\,(C_1(i, j), 256) \quad (2)$$

where $floor$ represents the round function to the nearest integer number.

Step 4: Encrypt $C_2\,(i, j)$ according to Section 2.2 by implementing the public key to get $D_1\,(i, j)$ as shown:

$$D_1(i, j) = mod(\,(C_2\,(i, j))^e, n) \quad (3)$$

Step 5: Convert $D_1(i, j)$ into one dimensional signal $D_2(i)$, where D_2 represents the final cipher speech signal which will be transferred to the receptor.

Step 6: Generate the private key $(n, d)$by the receiver according to Section 2.1.

Step 7: Reshape the encrypted signal $D_2(i)$ into two-dimensional signal $E_1\,(i, j)$.

Step 8: Decrypt $E_1\ (i,j)$ according to Section 2.2 by employing the secret key to acquire $E_2\ (i,j)$ as described below:

$$E_2(i,j) = mod\ ((E_1\ (i,j))^d, n) \quad (4)$$

Step 9: The gained decrypted speech samples from Step 8 are restored by applying the equation:

$$F_1(i,j) = 256 \times floor\ \left(\frac{E_2(i,j)}{10^{14}}\right) \quad (5)$$

Step 10: The signal $F_1(i,j)$ is transformed into one dimensional signal $F_2(i)$, where F_2 represents the original reconstructed speech signal.

4.2. Speech Ciphering/Deciphering Operation Using El-Gamal System

Step 1: Generate the public key by the sender according to Section 3.1.

Step 2: Convert the plain one-dimensional signal $A\ (i)$ into two-dimensional signal $B\ (i,j)$.

Step 3: Modify the resultant speech samples from Step 2 by applying Equations (1) and (2) to produce $C_1\ (i,j)$ and $C_2\ (i,j)$, respectively.

Step 4: Encrypt $C_2\ (i,j)$ according to Section 3.2 by carrying out the public key to acquire the cipher pair (D_1, D_2).

$$D_1 = mod(g^k, p) \quad (6)$$

$$D_2(i,j) = mod\ (C_2(i,j) \times y^k, p) \quad (7)$$

Step 5: Transform $D_2(i,j)$ into one dimensional signal $D_3(i)$, where (D_1, D_3) represents the cipher text pair which will be send to the recipient.

Step 6: Produce the secret key by the receptor according to Section 3.1.

Step 7: Reshape the ciphered signal $D_3(i)$ into two-dimensional signal $E_1\ (i,j)$.

Step 8: Decrypt $E_1\ (i,j)$ according to Section 3.2 via the secret key to get $E_2(i,j)$ as:

$$E_2(i,j) = mod((E_1(i,j)/D_1^x), p) \quad (8)$$

Step 9: Recover the decrypted samples from Step 8 by utilizing Equation (5) to generate $F_1(i,j)$.

Step 10: The original retrieved signal F_2 is obtained by converting $F_1(i,j)$ into one dimensional signal $F_2(i)$.

5. Performance metrics

To assess the cryptosystem performance, a number of common quantitative measures are employed for both encrypted and decrypted speech signals using the RSA and El-Gamal cryptosystems. These measures are Signal to Noise Ratio (SNR), Segmental Signal to Noise Ratio (SNRseg), Segmental Spectral Signal to Noise Ratio (SSSNR), Frequency Weighted Segmental Signal to Noise Ratio (fwSNRseg), Log Likelihood Ratio (LLR) and Bit Error Rate (BER). These metrics are explained as follows.

5.1. Signal to Noise Ratio (SNR)

This metric is defined as:

$$SNR\ (dB) = \ 10 \times log_{10} \frac{\sum_{i=1}^{L} x_i^2}{\sum_{i=1}^{L} [x_i - y_i]^2} \quad (9)$$

where L represents the number of speech samples, while $x_i\ and\ y_i$ represent the original and encrypted speech signals, respectively [2].

5.2. Segmental Signal to Noise Ratio (SNRseg)

SNRseg is computed as:

$$SNR_{seg}(dB) = \frac{10}{M} \sum_{m=0}^{M-1} log_{10}\ \frac{\sum_{n=Lm}^{Lm+L-1} x_i^2}{\sum_{n=Lm}^{Lm+L-1} [x_i - y_i]^2} \quad (10)$$

where M represents the number of frames in the speech signal [3].

5.3. Segmental Spectral Signal to Noise Ratio (SSSNR)

Segmental Spectral Signal to Noise Ratio or SSSNR is described as:

$$SSSNR_i(dB) = 10\ log_{10} \frac{\sum_{i=1}^{L} |X_i|}{\sum_{i=1}^{L} [|X_i| - |Y_i|]} \quad (11)$$

where X_i and Y_i represent the DFT of the original and encrypted speech signals, respectively [7].

5.4. Frequency Weighted Segmental Signal to Noise Ratio (fwSNRseg)

fwSNRseg is expressed as:

$$fwSNR_{seg}(dB) = \frac{10}{M}\sum_{m=0}^{M-1}\frac{\sum_{j=0}^{K-1} W(j,m)\ log10\frac{X(j,m)^2}{[X(j,m)-\hat{X}(j,m)]^2}}{\sum_{j=0}^{K-1} W(j,m)} \quad (12)$$

where $W(j,m)$ refers to the weight of the frequency band, $X(j,m)$ and $\hat{X}(j,m)$ are the spectrums of the input and output speech signals, respectively [4].

5.5. Log Likelihood Ratio (LLR)

LLR can be calculated as:

$$LLR = \log\left(\frac{\overrightarrow{a_e}R_o\overrightarrow{a_e^T}}{\overrightarrow{a_o}R_o\overrightarrow{a_o^T}}\right) \quad (13)$$

where a_e and a_o indicate to the LPC vectors of the plain and ciphered or deciphered signals, respectively, whilst R_o represents the autocorrelation matrix of the encrypted or decrypted speech signal [9].

5.6. Bit Error Rate (BER)

This measurement is represented as:

$$BER = 0.5 \times erfc\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (14)$$

where E_b symbolizes to the energy of average bits, N_0 symbolizes to the spectral density of noise [4].

6. Simulation outcomes

In order to quantify and assess the performance of two systems, several tests are carried out. These tests are designed and implemented on Matlab (R2013a), Windows 7, a laptop machine equipped with Processer of Intel Core i3, RAM of 3.90 GB and CPU of 2.40 GHz. The speech signals utilized in all tests are spoken sentences that recorded from different males and females in English language with sampling rate of 16 KHz for each signal and different interval (from 1 to 5 seconds) after eliminating all silence durations from them. The database utilized in this simulation is TIMIT database. This database is designed to supply speech information for acoustic researches, and for the evolution and assessment of automatic voice recognition systems. It includes wide band records of 630 speakers of 8 main American English languages. TIMIT comprises four groups of samples: phonemes, transcripts, audio and word list [21]. The cryptosystem is first implemented using the RSA algorithm. The variable values p and q in this work are set as 3 and 7, respectively, while the value of e is chosen to be 5. Thus, the value of public key (n, e) equals to (21, 5) and the value of private key (n, d) equals to (21, 5). The work is then implemented using the El-Gamal algorithm. The variable values $p, g, x\ and\ k$ are set as 13, 7, 3 and 9, respectively. Hence, the public key value (p, g, y) equals to (13, 7, 5) and the private key value (x) equals to (3). Many values have been tested using Matlab program to generate the keys for both systems: (p, q, e) of RSA and (p, g, x, k) of El-Gamal. These parameters are specified in this simulation because they produce the best encryption result. The results of implementing the two mentioned techniques on the input speech signal are clarified in Figure 1. Figure 1 (a) shows the original signal; Figure 1 (b, c) illustrates the encrypted and decrypted speech signals by employing the RSA, whereas Figure 1 (d, e) illustrates the encrypted and decrypted speech signals by employing the El-Gamal. This figure demonstrates that the ciphered speech signals of the two techniques are quite different from the input speech. In addition, the deciphered signals resulting from applying the two algorithms are identical to the input signal. These visual outcomes prove the high ciphering and deciphering quality of both the approaches.

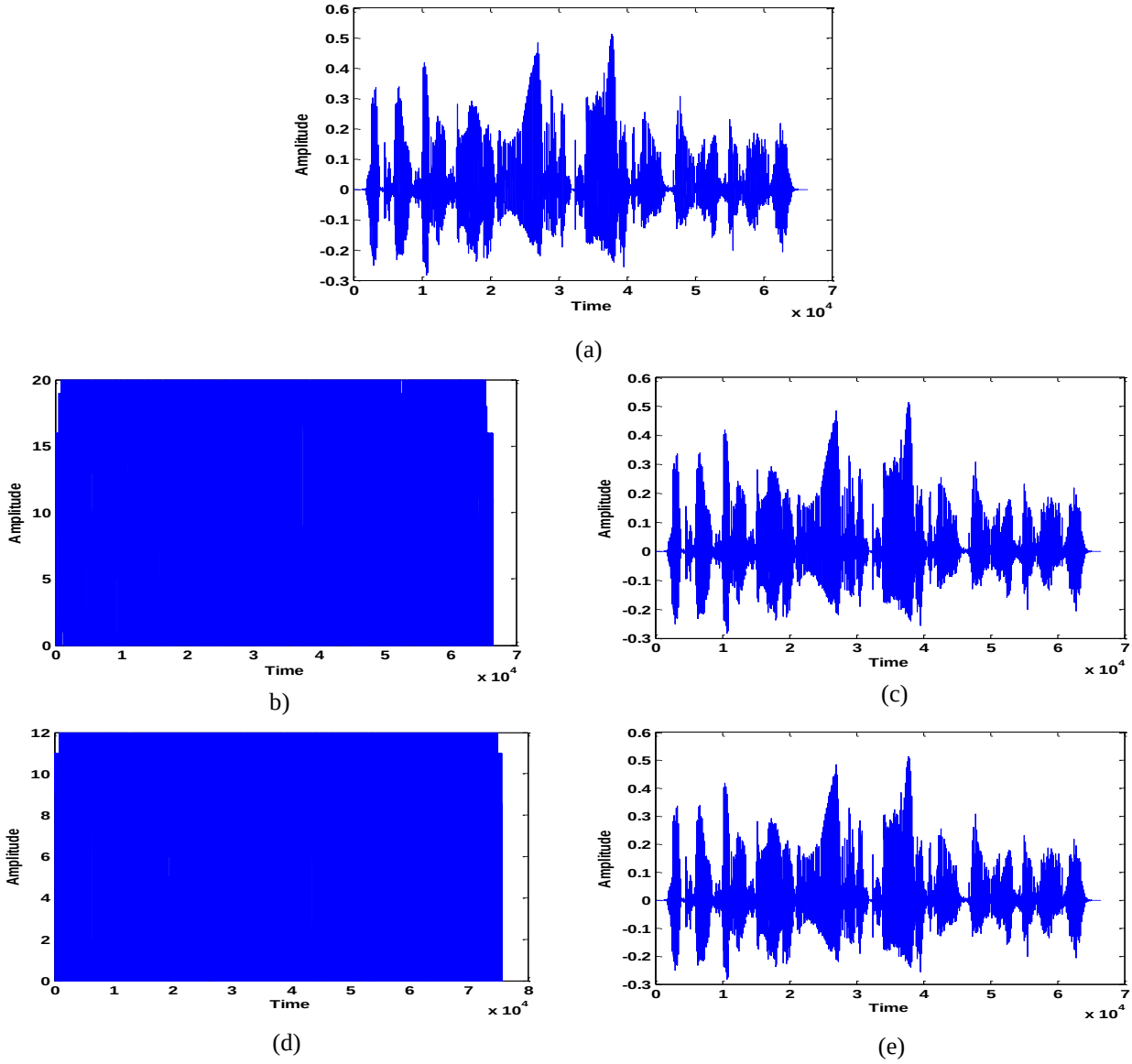


Figure 1. (a) Input speech signal (b, c) Ciphered and deciphered speech signals, respectively using the RSA system (d, e) Ciphered and deciphered speech signals, respectively using the El-Gamal system

6.1. Quality of speech encryption

To assess the quality of speech signal encryption, six quality metrics are used which have been aforementioned: SNR, SNRseg, SSSNR, fwSNRseg, LLR and BER. The quality of encryption is high when the values of LLR and BER increase, whereas the values of SNR, SNRseg, SSSNR and fwSNRseg decrease [1, 7]. The numerical outcomes of the presented system for both RSA and El-Gamal schemes are explained in Table 1. It can be found from this table that LLR and BER scores are high, contrary, SNR, SNRseg, SSSNR and fwSNRseg scores are low for both methods, which means that the encryption quality is high for both systems. But the RSA technique gives lower value results in terms of SNR, SNRseg, SSSNR and fwSNRseg, and higher value results in terms of LLR and BER than the El-Gamal technique. This implies that the RSA ciphering performance is better than that for the El-Gamal method for the same input test speech signals.

6.2. Quality of speech decryption

The same six quality metrics are used to measure the quality of speech signal decryption: SNR, SNRseg, SSSNR, fwSNRseg, LLR and BER. The quality of decryption is high when the values of LLR and BER decrease, whilst the

values of SNR, SNRseg, SSSNR and fwSNRseg increase [3, 4]. The results of the proposed method for both RSA and El-Gamal mechanisms are illustrated in Table 2. It can be noticed from this table that LLR and BER outcomes are low. On the other hand, SNR, SNRseg, SSSNR and fwSNRseg outcomes are high for both systems, which indicate that the decryption quality is high for both cryptosystems. But the El-Gamal approach gives higher value results in terms of SNR, SNRseg, SSSNR and fwSNRseg and lower value results in terms of LLR and BER than the RSA technique. This demonstrates that the El-Gamal deciphering performance is better than that for the RSA method for the same input test speech signals.

Table 1: Results of quality metrics using the RSA and El-Gamal algorithms for encryption

RSA encryption						
File name	**SNR (dB)**	**SNRseg (dB)**	**SSSNR (dB)**	**fwSNRseg (dB)**	**LLR**	**BER**
Signal 1	-38.8024	-42.2134	-20.7970	-40.0409	2.1780	0.9858
Signal 2	-41.0728	-49.2223	-21.9335	-42.6029	1.7553	0.9911
Signal 3	-39.7990	-44.4188	-21.2965	- 39.7569	2.2732	0.9923
Signal 4	-42.2061	-49.2748	-22.2889	-41.1423	1.6969	0.9933
Signal 5	-44.0994	-44.2990	-23.0486	-40.8159	1.6573	0.9932
El-Gamal encryption						
File name	**SNR (dB)**	**SNRseg (dB)**	**SSSNR (dB)**	**fwSNRseg (dB)**	**LLR**	**BER**
Signal 1	-33.5692	-38.7771	-17.9107	-36.7787	2.0128	0.9857
Signal 2	-35.7641	-45.0041	-19.0946	-39.2420	1.7119	0.9910
Signal 3	-34.4144	-37.5670	-18.4378	-36.4386	1.3692	0.9922
Signal 4	-36.6076	-39.6144	-19.3709	-37.6435	1.5968	0.9932
Signal 5	-38.5281	-34.9039	-20.1495	-37.3170	1.2834	0.9931

Table 2: Results of quality metrics using the RSA and El-Gamal algorithms for decryption

RSA decryption						
File name	**SNR (dB)**	**SNRseg (dB)**	**SSSNR (dB)**	**fwSNRseg (dB)**	**LLR**	**BER**
Signal 1	241.4585	244.7265	119.3172	60.7713	1.0608×10^{-15}	0.0579
Signal 2	239.1209	231.1241	118.1779	61.5437	2.8930×10^{-16}	0.0952
Signal 3	240.4719	243.8585	118.7960	61.2856	-4.5323×10^{-15}	0.0967
Signal 4	237.9839	231.5436	117.8404	60.6452	5.7860×10^{-16}	0.0764
Signal 5	236.1555	237.7266	116.9582	60.1367	4.4359×10^{-15}	0.0677
El-Gamal decryption						
File name	**SNR (dB)**	**SNRseg (dB)**	**SSSNR (dB)**	**fwSNRseg (dB)**	**LLR**	**BER**
Signal 1	246.9199	245.7474	122.3737	61.5085	-3.9537×10^{-15}	0.0514
Signal 2	244.4318	239.1378	120.9266	62.4175	-2.4108×10^{-16}	0.0664
Signal 3	246.0546	245.0670	121.6163	62.1142	-5.4002×10^{-15}	0.0881
Signal 4	243.4874	236.1115	120.5347	63.8581	-5.7860×10^{-16}	0.0656
Signal 5	241.7349	244.2878	119.7779	62.0547	4.8216×10^{-16}	0.0475

6.3. Histogram analysis

Histogram is an approximate depiction of the distribution for numerical or categorical information. The speech sample values should have a roughly flat distribution after utilizing the ciphering operation in order to endure the statistical attack [11]. Figure 2 (a) depicts the input signal histogram; Figure 2 (b, c) depicts the encrypted and decrypted signal histograms after employing the RSA technology, and Figure 2 (d, e) depicts the encrypted and decrypted signal histograms after applying the El-Gamal technology. It is obvious from Figure 2 that the consequent histograms after implementing the two algorithms are different

from the input file histogram and the output speech samples possess approximately flat distribution, which confirms the good encryption performance for the two described schemes. Moreover, the decrypted file histograms are similar to their corresponding original file, which proves the good decryption performance for the two schemes. However, by comparing Figures 2 (b) and 2 (d), it can be noticed that the output speech file histogram from the RSA encryption is flatter than the output signal histogram from the El-Gamal encryption. This refers that the ciphering process with the application of the RSA is more efficient than the El-Gamal, whilst the deciphering process is equally reliable and efficient with the utilization of the two methods for the same test speech file.

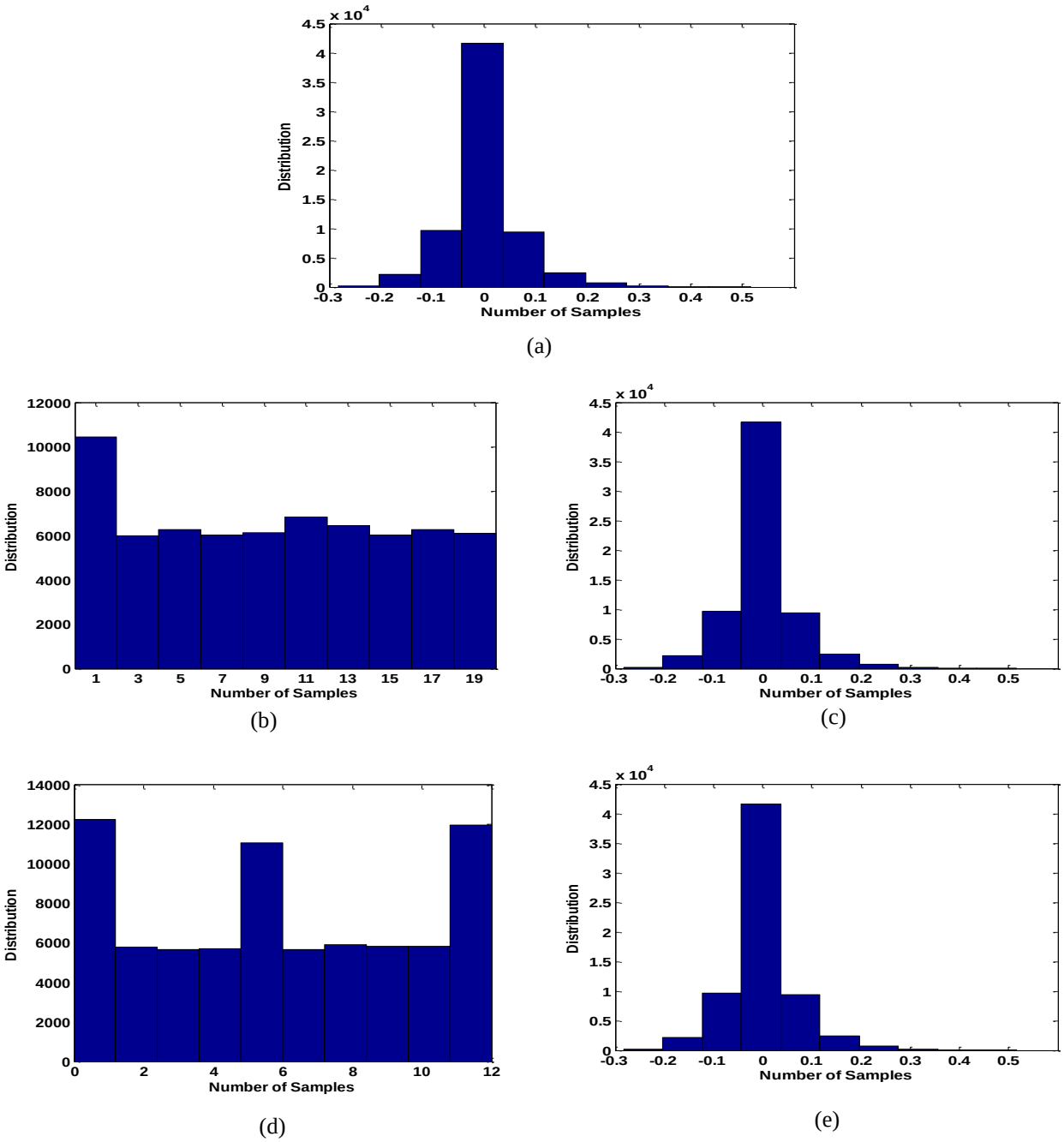


Figure 2. Histogram analysis (a) Histogram of input signal (b, c) Histograms of ciphered and deciphered signals, respectively with the RSA system (d, e) Histograms of ciphered and deciphered signals, respectively with the El-Gamal system

6.4. Spectrogram analysis

Spectrogram points to a visual image of the spectrum for audio signal frequency when it changes with time [12]. Figure 3 (a) exhibits the input file spectrogram; Figure 3 (b) presents the variations in the input file spectrogram after the RSA encryption and Figure 3 (c) presents the recovered file spectrogram after the RSA decryption, whereas Figure 3 (d) shows the changes in the plain file spectrogram after the El-Gamal encryption and Figure 3 (e) shows the restored file spectrogram after the El-Gamal decryption. It is clear that the spectrograms of the cipher signals for the RSA and El-Gamal systems are fully different from the original version spectrogram. Further, the plain and decrypted signal spectrograms are identical. This indicates the high ciphering and deciphering properties of the two audio cryptosystems for the same input audio signal.

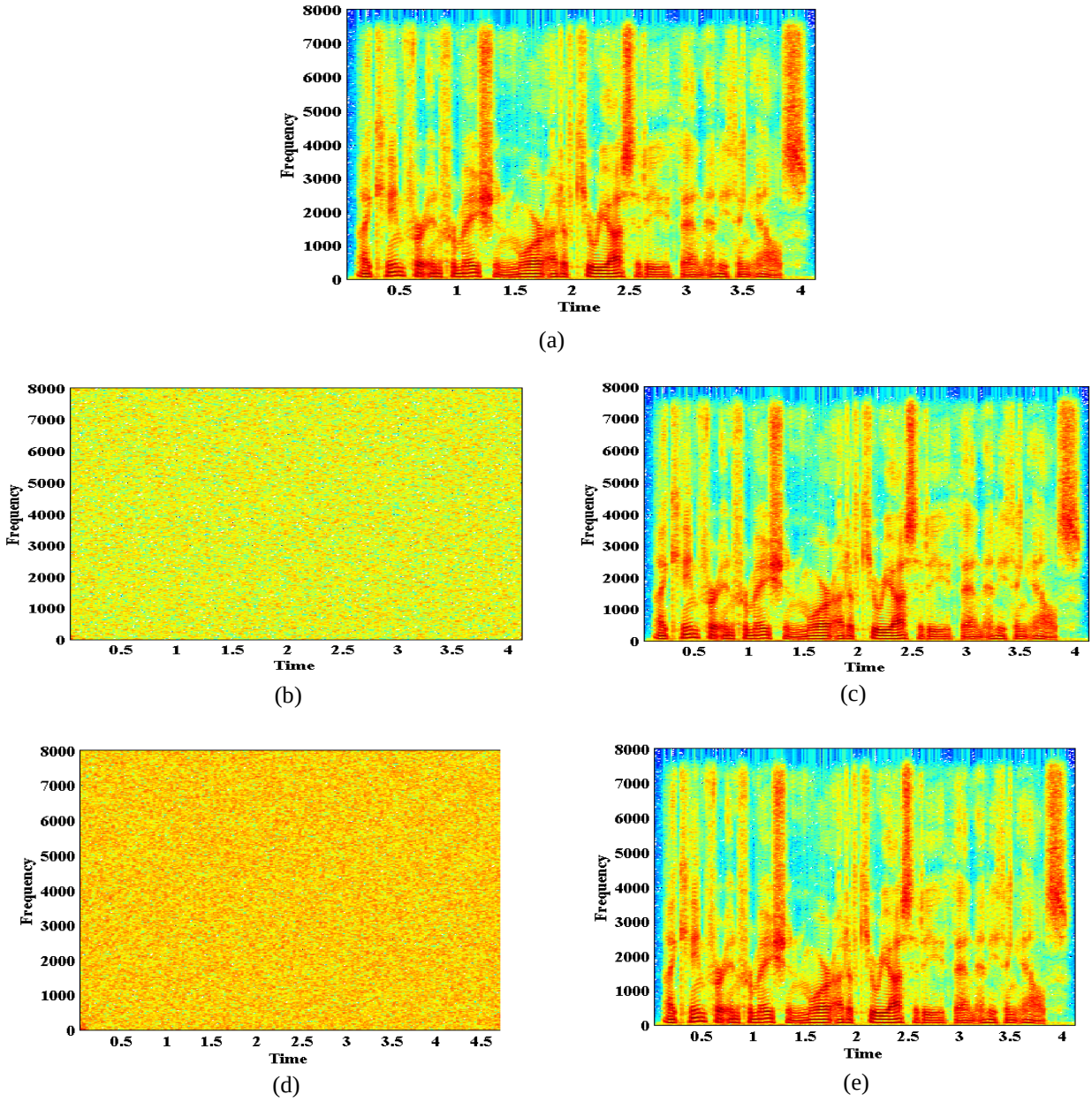


Figure 3 Spectrogram analysis (a) Spectrogram of input signal (b, c) Spectrograms of ciphered and deciphered signals, respectively with the RSA system (d, e) Spectrograms of ciphered and deciphered signals, respectively with the El-Gamal system

6.5. Correlation coefficient analysis

Correlation Coefficient or CC is a major index to assess the ciphering/deciphering quality of speech cryptosystem. If the CC value is zero or near to zero, then the relation between speech samples in the plain and its corresponding cipher signals is weak, this demonstrates a high ciphering effect. Inversely, if the CC value is one or close to one, then the relationship between speech samples in the input and output signals is strong, this refers to a high deciphering effect. This indicator is computed as [13, 14, 19]:

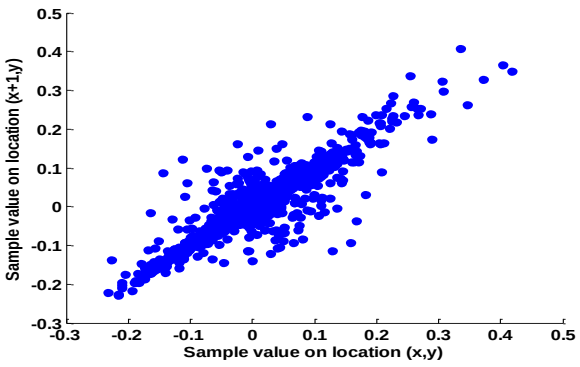
$$CC = \frac{cov\,(x,y)}{\sigma_x\,\sigma_y} = \frac{\sum_{i=1}^{N}(x_i - E(x))(y_i - E(y))}{\sqrt{\sum_{i=1}^{N}(x_i - E(x))^2}\sqrt{\sum_{i=1}^{N}(y_i - E(y))^2}} \quad (15)$$

$$E(x) = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (16)$$

where N symbolizes to the total number of speech samples employed in the calculations, x and y denote the values of speech signals for the input and output files, respectively. Table 3 exhibits the ciphering and deciphering outcomes of both exploited methods for the test speech files. This table reveals that both adopted mechanisms achieve very small CC scores (almost zero) in the encryption process and quite high CC values (one) in the decryption process for all input signals. Hence, the RSA and El-Gamal systems are considered efficient and complicated approaches for ciphering audio samples, and they can produce a deciphered signal that totally corresponds to the original one. Additionally, Figure 4 clarifies the scatter plots of input, encrypted and reconstructed audio files for the RSA and El-Gamal schemes. Obviously, the speech samples in Figure 4 (a) are grouped and centered about the main diagonal. The encrypted speech samples in Figures 4 (b) and 4 (d) possess flat distribution, which reflects the high randomness of speech encoding for the two techniques. On the other hand, the decrypted speech samples in Figures 4 (c) and 4 (e) are like the plain samples, which reflects the high reconstruction ability for the two algorithms. It can be deduced from Table 3 and Figure 4 that both described systems can reduce the CC values between samples in the original signal and increase the CC values to one in the restored signal; thereby RSA and El-Gamal cryptosystems can counter this analysis successfully.

Table 3: Results of correlation analysis using the RSA and El-Gamal algorithms for encryption and decryption

File name	RSA		El-Gamal	
	Encryption	Decryption	Encryption	Decryption
Signal 1	-0.0382	1	0.0202	1
Signal 2	0.0428	1	-0.0364	1
Signal 3	0.0169	1	0.0012	1
Signal 4	-0.0587	1	0.0221	1
Signal 5	0.0136	1	0.0176	1



(a)

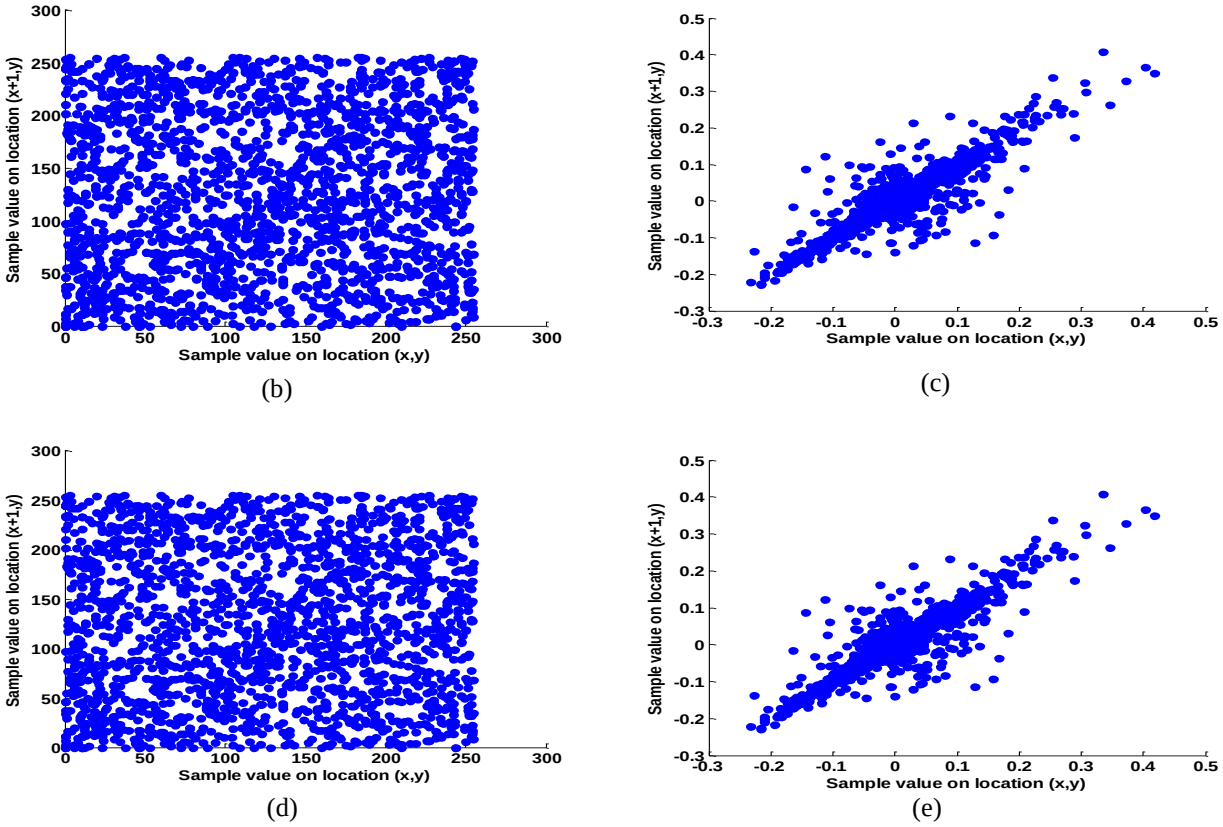


Figure 4. Correlation coefficient analysis (a) Correlation of input signal (b, c) Correlation of ciphered and deciphered signals, respectively with the RSA system (d, e) Correlation of ciphered and deciphered signals, respectively with the El-Gamal system

6.6. Differential analysis

Numbers of Samples Change Rate (NSCR) and Unified Average Changing Intensity (UACI) parameters are generally utilized to analyze the resistance of encryption process against differential attack. Two different signals are enciphered in this test via same keys; the original signals are differ only by one sample. Next, the resultant cipher speech files are compared by applying the NSCR and UACI. These two parameters are given by [11, 19]:

$$NSCR = \frac{\sum_i D(i)}{l} \times 100\% \quad (17)$$

$$D(i) = \begin{cases} 0 & if\ x_1(i) = x_2(i) \\ 1 & if\ x_1(i) \neq x_2(i) \end{cases} \quad (18)$$

$$UACI = \frac{1}{l}\left[\sum_i \frac{|x_1(i) - x_2(i)|}{255}\right] \times 100\% \quad (19)$$

x_1 and x_2 denote the two encrypted files which their input signals differ by one sample, l indicates the overall number of audio samples. The optimal values for NSCR and UACI are 99% and 33%, respectively. A secure ciphering algorithm should possess NSCR and UACI parameters that are close to the idealistic values. The NSCR and UACI outcomes are computed for the test files by executing the RSA and El-Gamal methods, and the obtained results are given in Table 4. Both NSCR and UACI values in this table are near to the optimal values, which clearly reveal that the clear and ciphered signals produced by both systems are totally different. Also, it can be observed that the obtained NSCR and UACI scores for the RSA are better than those scores for the El-Gamal This manifests the high degree of security of the RSA in this analysis in comparison with its counterpart the El- Gamal for encrypting the same different plain signals.

6.7. Speed performance analysis

The speed is an important parameter that must be considered in order to analyze the

cryptosystem efficiency [4]. The hardware configuration utilized for simulation findings has been mentioned in Section 6. The required time (seconds) is computed in this analysis of encryption/decryption for both techniques implementation on the input speech files. Table 5 contains the total computational time results of ciphering and deciphering procedures for the RSA and El-Gamal methods. According to this table, the execution times of encryption/decryption operations for the two cryptosystems are quite short and satisfactory. Further, the encryption/decryption time increases as the input signal length increases for both systems. Besides, the deciphering time consumes more time than the enciphering time for the two schemes. Also, it can be shown that the encryption/decryption times for the El-Gamal cryptosystem are shorter than the corresponding times for the RSA cryptosystem. Table 5 manifests that the El-Gamal system is faster than the RSA system at encryption/decryption for different test plain signals.

Table 5: Results of speed performance analysis for the RSA and El-Gamal algorithms

File name	Encryption time (s)		Decryption time (s)	
	RSA	El-Gamal	RSA	El-Gamal
Signal 1	0.024142	0.017578	0.026980	0.021074
Signal 2	0.041728	0.020996	0.041861	0.021440
Signal 3	0.044199	0.022206	0.044232	0.023730
Signal 4	0.045088	0.026886	0.046233	0.035225
Signal 5	0.051075	0.030539	0.054409	0.036032

6.8. Noise influence analysis

In this analysis, the deciphered speech signal is assessed at the receptor side in the noise existence with various SNR estimations [15, 19]. Additive White Gaussian Noise (AWGN) is added to the ciphered signal with different SNR (5dB-50dB). The noise influence on the performance criterion SNR, SNRseg, SSSNR, fwSNRseg, LLR, BER and CC are computed for the decrypted signal, and the obtained outcomes for the adopted techniques RSA and El-Gamal are represented in Tables 6 and 7, respectively. The larger SNR, SNRseg, SSSNR, fwSNRseg and CC values, and the lower LLR, BER values between the input and reconstructed signals, yields a good deciphering quality. It can be seen from the tables that the SNR, SNRseg, SSSNR, fwSNRseg and CC scores increase, whilst LLR and BER scores decrease as the input SNR of noise increases gradually for both algorithms. This reflects the robustness of the cryptosystems to noise distortion. Furthermore, it can be realized from Tables 6 and 7 that the obtained SNR, SNRseg, SSSNR, fwSNRseg and CC results are always greater, whereas LLR and BER results are always lower for the RSA than those results obtained for the El- Gamal at all input SNR values of noise. According to the outcomes in Tables 6 and 7, it is clear that the RSA method outperforms the El-Gamal method in noise invulnerability by the means of performance metrics.

Table 6: Results of speech quality metrics for the RSA in the presence of AWGN

SNR of noise (dB)	SNR (dB)	SNRseg (dB)	SSSNR (dB)	fwSNRseg (dB)	LLR	BER	CC
5	7.0120	7.6813	8.4353	9.6353	0.8077	0.9565	0.9124
10	9.9870	8.3578	9.0767	11.6486	0.7447	0.8330	0.9534
15	11.7670	11.1683	10.2787	13.6611	0.7175	0.7345	0.9683
20	13.0033	14.0812	11.0615	15.6729	0.6525	0.6022	0.9760
25	14.7739	18.1633	12.0819	16.6684	0.4746	0.5670	0.9806
30	15.4496	19.1889	14.0670	17.6689	0.4326	0.5120	0.9837
35	16.0421	20.1046	15.5707	18.6688	0.3697	0.3222	0.9859
40	17.4144	21.0766	16.0602	20.6806	0.3180	0.2044	0.9877
45	18.1131	22.5502	17.0648	21.6592	0.1983	0.1170	0.9890
50	19.0380	25.2186	19.0170	22.6738	0.1397	0.1000	0.9902

Table 7: Results of speech quality metrics for the El-Gamal in the presence of AWGN

SNR of noise (dB)	SNR (dB)	SNRseg (dB)	SSSNR (dB)	fwSNRseg (dB)	LLR	BER	CC
5	6.0130	6.7404	7.0272	7.8583	0.8435	0.9674	0.8955
10	9.5595	7.8024	8.5539	9.8643	0.7776	0.8930	0.9489
15	11.4849	10.0405	9.4355	11.8633	0.7650	0.7940	0.9572
20	12.7654	13.3442	10.0633	13.8668	0.6830	0.6921	0.9682
25	13.7875	17.7640	11.5555	15.4498	0.5280	0.6673	0.9759
30	14.6275	18.4061	12.9592	17.4421	0.4738	0.5932	0.9778
35	15.2933	19.3448	13.2845	18.4550	0.3963	0.4777	0.9796
40	16.0364	20.4617	14.5850	19.4452	0.3204	0.3219	0.9799
45	17.4263	21.2657	15.8327	21.4517	0.2200	0.2306	0.9803
50	18.1393	23.2586	18.0399	22.4443	0.1771	0.1644	0.9820

7. Conclusions

This work performs a comparative study between the RSA and El-Gamal techniques in order to determine which of the methods is more effective for encrypting speech files. The two schemes are tested and compared via sundry empirical analyses: SNR, SNRseg, SSSNR, fwSNRseg, LLR, BER at encryption and decryption processes, histogram, spectrogram, correlation coefficient, and differential analyses, time for enciphering/deciphering operations, and finally, the effect of noise analysis. It can be concluded from the empirical and visual outcomes that the two speech cryptosystems are robust and can provide a reliable method to encipher and decipher the speech data with high level of confidently, security and privacy. Additionally, the RSA cryptosystem performance is superior to that of the El-Gamal cryptosystem in most analyses by the means of sundry enciphering and deciphering speech quality indicators.

References

[1] Hanaa A. Abdallah and S. Meshoul, "A Multilayered Audio Signal Encryption Approach for Secure Voice Communication", Electronics, Vol. 12, No. 1, pp. 2, 2023.

[2] I. Yasser, Mohamed A. Mohamed, Ahmed S. Samra, and F. Khalifa, "A Chaotic-Based Encryption/Decryption Framework for Secure Multimedia Communications", Entropy, Vol. 22, No. 11, pp. 1253, 2020.

[3] Omar A. Imran, Sura F. Yousif, Isam S. Hameed, W. N. Al-Din Abed, and Ali T. Hammid, "Implementation of El-Gamal algorithm for speech signals encryption and decryption", Procedia Computer Science, International Conference on Computational Intelligence and Data Science (ICCIDS 2019), Vol. 167, pp. 1028–1037, 2020.

[4] Sura F. Yousif, "Secure voice cryptography based on Diffie-Hellman algorithm", IOP Conf. Series: Materials Science and Engineering, 2nd International Scientific Conference of Engineering Sciences (ISCES 2020), Vol. 1076, No. 1, pp. 012057, 2021.

[5] S. Fatima, T. Rehman, M. Fatima, S. Khan, and M. A. Ali, "Comparative Analysis of Aes and Rsa Algorithms for Data Security in Cloud Computing", Engineering Proceedings, Vol. 20, No. 1, pp. 14, 2022.

[6] Amal H. Khaleel and Iman Q. Abduljaleel, "A novel technique for speech encryption based on k-means clustering and quantum chaotic map", Bulletin of Electrical Engineering and Informatics, Vol. 10, No. 1, pp. 160-170, 2021.

[7] Iman Q. Abduljaleel and Amal H. Khaleel, "Speech signal compression and encryption based on Sudoku, fuzzy C-means and threefish cipher", International Journal of Electrical and Computer Engineering (IJECE), Vol. 11, No. 6, pp. 5049-5059, 2021.

[8] G. Kaur, K. Singh, and H. Singh Gill, "Chaos-based joint speech encryption scheme using SHA-1", Multimedia Tools and Applications, Vol. 80, pp. 10927-10947, 2021.

[9] Sarah M. Abdullah and Iman Q. Abduljaleel, "Speech Encryption Technique using S - box based on Multi Chaotic Maps", TEM Journal, Vol. 10, Issue 3, pp. 1429-1434, 2021.

[10] Mahmood K. Ibrahem and Hussein A. Kassim, "VoIP Speech Encryption System Using Stream Cipher with Chaotic Key Generator", Iraqi Journal of Science, pp. 240-248, 2021.

[11] W. Dai, X. Xu, X. Song, and G. Li, "Audio Encryption Algorithm Based on Chen Memristor

Chaotic System", Symmetry, Vol. 14, No. 1, pp. 17, 2022.

[12] S. Mokhnache, M. E.Daachi, T. Bekkouche, and N. Diffellah, "A Combined Chaotic System for Speech Encryption", Engineering, Technology & Applied Science Research, Vol. 12, No. 3, pp. 8578-8583, 2022.

[13] Obaida M. Al-Hazaimeh, A. A. Abu-Ein, Khalid M. Nahar, and Isra S. Al-Qasrawi, "Chaotic elliptic map for speech encryption", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 25, No. 2, pp. 1103-1114, 2022.

[14] Nidaa F. Hassan, A. Al-Adhami, and Mohammed S. Mahdi, "Digital Speech Files Encryption based on Hénon and Gingerbread Chaotic Maps", Iraqi Journal of Science, Vol. 63, No. 2, pp. 830-842, 2022.

[15] A. u. Rehmana, X. Liao, R. Ashraf, S. Ullahc, and H. Wang, "A color image encryption technique using exclusive-OR with DNA complementary rules based on chaos theory and SHA-2", Optik, Vol. 159, pp. 348-367, 2018.

[16] X. Wua, H. Kana, and J. Kurths, "A new color image encryption scheme based on DNA sequences and multiple improved 1D chaotic maps", Applied Soft Computing, Vol. 37, pp. 24-39, 2015.

[17] Sura F. Yousif, "Encryption and Decryption of Audio Signal Based on RSA Algorithm", Vol. 5, Iss. 7, pp. 57-64, 2018.

[18] S. Du and G. Ye, "IWT and RSA based asymmetric image encryption algorithm", Alexandria Engineering Journal, Vol. 66, pp. 979-991, 2023.

[19] Sura F. Yousif, "A new speech cryptosystem using DNA encoding, genetic and RSA algorithms", International Journal of Engineering & Technology, Vol. 7, No. 4, pp. 4550-4557, 2018.

[20] Jumadi M. Parenreng, S. Maulida, and A. Wahid, "E-mail Security System Using El-Gamal Hybrid Algorithm and AES (Advanced Encryption Standard) Algorithm", Internet of Things and Artificial Intelligence Journal, Vol. 2, No. 1, pp. 1-9, 2022.

[21] TIMIT Acoustic-Phonetic Continuous Speech Corpus, National Institute of Standards and Technology (NIST). https://doi.org/10.35111/17gk-bn40, 1993.