Robust and Secure Speech Based on Elliptic Curve Cryptography for WB G.722.2

Messaouda Boumaraf, Fatiha Merazka

Robust and Secure Speech Based on Elliptic Curve Cryptography for WB G.722.2

Messaouda Boumaraf [1] and Fatiha Merazka [2]

[1, 2] LISIC Laboratory, Telecommunications Department, USTHB University, Algiers, Algeria
boumaraf.messa@gmail.com
fmerazka@usthb.dz

Abstract. Elliptic Curve Cryptographic (ECC) systems are now increasingly used in protocols using public key cryptography, as it allows to significantly reduce the size of the keys used compared to other cryptographic systems such as RSA (Ron Rivest, Adi Shamir and Leonad Adelman). ECC is based on Elliptic Curve (EC) theory, for creating faster, smaller, and more efficient cryptographic schemes. This paper presents an encryption scheme for AMR-WB ITU-T G.722.2 speech based on ECC for securing transmitted speech signals. The implementation of ECC is carried out by transforming the coded speech into an affine point on the elliptic curve, over a finite field. The proposed cryptosystem is evaluated by including waveform and spectrogram analysis, Enhanced Modified Bark Spectral Distortion (EMBSD) and Wideband Perceptual Evaluation of Speech Quality (WB-PESQ) for objective testing and Mean Opinion Score (MOS) for its subjective counterpart in order to confirm the efficiency of our proposed method in terms of security and privacy.

Keywords: Elliptic Curve Cryptography (ECC), Speech Encryption and Decryption, EMBSD, AMR-WB G.722.2.

1 Introduction

Cryptography is a crucial field in information security as it allows for secure communication, regardless of its essence, such as video, image, audio sound in particular speech through open networks like the internet.

Cryptography can be divided into two major categories, symmetric-key cryptography [1-3] and asymmetric-key cryptography [4-5]. In symmetric-key cryptography, a single key is shared and used for both encryption and decryption However, in asymmetric-key cryptography a pair of keys is involved: a public key to encrypt data and a private key to decrypt data. Symmetric encryption is simple and fast, but its main problem is the fact that the two communicating parties must exchange the key in a secure way. AES (Advanced Encryption Standard), DES (Data Encryption Standard), IDES (international data encryption Algorithm), and 3DES are familiar examples of symmetric-key cryptography [6–10]. Asymmetric-key cryptography is more secure

given that keys can be shared via an unsecure channel. RSA, Digital Signature Standard (DSS) and Diffie-Hellman exchange methods are well-known examples [11-14].

Today, researchers envision replacing the symmetric and asymmetric encryption by Elliptic Curve Cryptographic (ECC) public key encryption [15-21], because it provides the same level of security as the widely used RSA but with reduced key size. In 1985, Koblitz and Miller proposed the use of elliptic curves in public key cryptography [22]. In [23],the authors present the emergence of Elliptic curve cryptography as a preferred cryptographic scheme using minimal computational resources is quite discernible. Its extension to the domain of image encryption using various mapping techniques has been analyzed in this paper. The efficiency of an encryption scheme based on ECC shall depend upon the appropriateness of mapping technique used to map pixels onto the Elliptic curve. In [24] authors presents the ECC for public key exchange and Digital signature operation. ElGamal PKE suffers from data expansion problem and extra computation for embedding integer represented message to specific coordinate satisfying the elliptic curve. The improved version solves the problem associated with ElGamal PKE with a good execution speed for a public key encryption scheme. The analysis results, the strength of ECDLP and the improvement made with ElGamal PKE make the proposed method a strong and reliable public key audio encryption scheme.

Given that the security of ECC encryption is more secure than that of traditional approaches, we propose in this paper a secure speech communication approach based on elliptic curve cryptography for WB G.722.2. The general process of the proposed cryptosystem architecture involving both encryption and decryption is illustrated below in Fig. 1.

The remainder of this paper is organized as follows. In section 2, The description of the AMR-WB G.722.2 standard is introduced. Section 3 gives a very brief description of the ECC technique, which has a direct relation with our contribution. Simulations and interpretations are presented in section 4. Finally, the conclusion is provided in section 5.

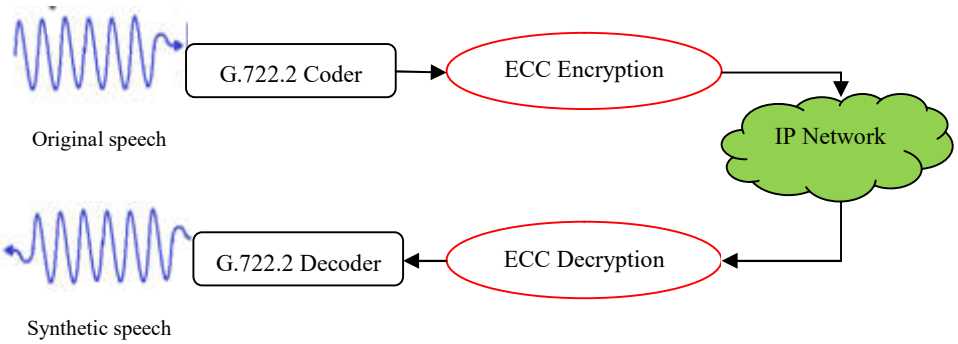


Fig.1. Diagram block of our proposed scheme cryptosystem

2 The AMR-WB G.722.2 codec

The G.722.2 AMR-WB standard is used as an internet wideband speech audio encoder for VoIP applications with an audio band of 50 – 7000 Hz instead of narrowband speech coders wire line quality of 300–3400 Hz. The increased bandwidth improves intelligibility and audio quality resulting in an improved speech quality in 3GPP LTE network.

The AMR-WB bitrates are 6.60, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05 and 23.85 kbps corresponding to encoded blocks of 132, 177, 253, 285, 317, 365, 397, 461 and 477 bits which are represented by modes 0, 1, 2, 3, 4, 5, 6, 7 and 8 respectively. The bitrates may be changed at any 20 ms frame boundary and the sampling rate is 16 000 samples/s. G.722.2 utilizes an integrated Voice Activity Detector (VAD); includes a packet loss concealment algorithm, support Discontinuous Transmission (DTX) and Comfort Noise Generation (CNG) features for increased efficiency [25].

The AMR-WB G722.2 includes six parameters to represent speech. Fig. 2 shows the structure of the frame to encoded blocks of 477 bits (23.85 kbps).

			Sub-frame 1					Sub-frame 2					Sub-frame 3					Sub-frame 4				
Header	VAD-flag	ISP	Pitch delay	LTP-filtering	Algebraic code	Gain	HB-energy	Pitch delay	LTP-filtering	Algebraic code	Gain	HB-energy	Pitch delay	LTP-filtering	Algebraic code	Gain	HB-energy	Pitch delay	LTP-filtering	Algebraic code	Gain	HB-energy
3	1	46	9	1	88	7	4	6	1	88	7	4	9	1	88	7	4	6	1	88	7	4

Total bits: 477 bits

Fig.2. The structure of the frame of the AMR-WB codec in mode 8 [26]

3 Elliptic Curve Cryptography

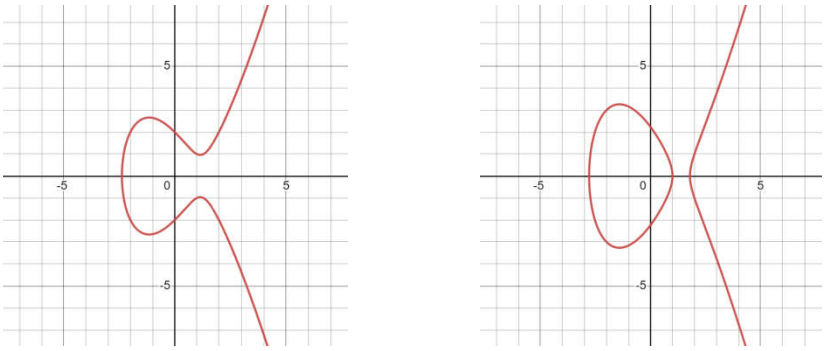
The ECC method was proposed in the mid-'80s by Victor Miller of IBM and Neal Koblitz of the University of Washington [22]. Generally, on R, the elliptic curves E will be considered as the set of pairs (x, y) defined over a field F (p) that satisfy a specific mathematical equation known as the Weierstrass normal form given as follows:

$$E: y^2 = x^3 + ax + b\ (mod\ p) \tag{1}$$

Where a and b are real elements of F that satisfy (2), p is a large prime number, and the set $(x, y) \in R^2$. Each choice of numbers a and b yields a different elliptic curve, and with the requirement that the discriminant is non-zero. $i.\ e.$

$$\Delta = 4a^3 + 27b^2 \neq 0 \ (mod\ p) \tag{2}$$

An elliptic curve is cubic equation that may assume different shapes on the plane, depending on the value of a and b as shown in Fig. 3. As it can be seen and verified, elliptic curves are symmetric about the x-axis.



(a) $E_1: y^2 = x^3 - 4x + 4$ (b) $E_2: y^2 = x^3 - 6x + 5$

Fig.3. Examples of elliptic curve graph

The points on curve form an abelian group, and the *base point* is G also known as the *generator of primitive element*, is predetermined point (X_G, Y_G) on elliptic curve that everyone uses to compute other points on the curve. All the points that satisfy the *ellipse formula (1)* with $a = 1$, $b = 1$ and $p = 89$, are represented in Fig. 4. So, an elliptic curve defined over a finite field has a finite number of points, which is denoted as N. However, the number of points in a group represents the *order of the group*.

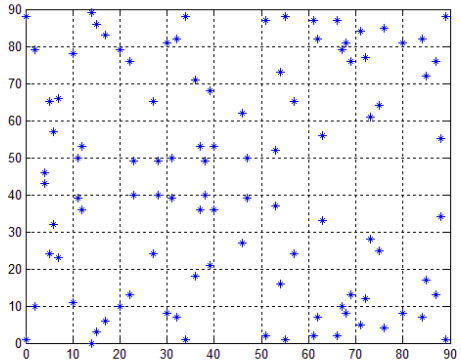


Fig.4. Set of points of elliptic curve $E_{89}(1,1): y^2 = x^3 + x + 1 \ mod\ 89$

Because of the particularity of the elliptic curve encryption algorithm, there are several operations are involved such as; point addition, point doubling, point subtraction and scalar point multiplication [23].

Several phases are involved in ECC to secure speech communications and encrypt transmitted coded speech. These phases are used and listed as follows: Diffie-Hellman key exchange protocol, ECC encryption algorithm and ECC decryption algorithm.

3.1 Diffie-Hellman key exchange protocol

The Diffie-Hellman key exchange protocol well-known ECDH (Elliptic Curve Diffie-Hellman) agrees on a secret key between two parties communicating over an insecure channel. The ECDH algorithm describes the process. To encrypt a speech data user A (U_A) and user B (U_B) must:

1) Choose an elliptic curve $E_p(a, b)$.
2) Take an affine point $G(X_G, Y_G)$ with large order n (n must be a large prime number and a divisor of N, where N is the number of points of elliptic) that lies on the curve.
3) Each user generates their public/private key pair:
   - **User *A* keys generation:**
     Select the private key k_A ($k_A < n$), using the base point G to calculate the public key $P_A = k_A * G$.
   - **User *B* keys generation:**
     Select the private key k_B ($k_B < n$), using the base point G to calculate the public key $P_B = k_B * G$.
4) Compute the shared key k:
   We have, $k_{AB} = k_A * P_B = k_A * k_B * G$ and $k_{BA} = k_B * P_A = k_B * k_A * G$
   Then $k = k_{AB} = k_{BA}$

Algorithm ECDH // key exchange protocol

{
// U_A and U_B legitimate users
$U_A = \{k_A, P_A\}$ // key pair for U_A
$U_B = \{k_B, P_B\}$ // key pair for U_B
// U_B send the public key to U_A over an insecure channel;
$U_A = \{k_A, P_{A,}, P_B\}$
// U_A send the public key to U_B over an insecure channel;
$U_B = \{k_B, P_{B,}, P_A\}$
Compute the shared key $k = k_A * k_B * G$
}

3.2 ECC encryption algorithm

Before encryption process, the coded speech data S is mapped into a point P_s on elliptic curve. For encryption, user A (U_A) encrypt P_s in cipher speech C_s using formula (3). Then, the point C_s will be sent over network to the user B (U_B).

$$C_s = (C_1, C_2) = \{kG\ ,\ P_s + kP_B\} \tag{3}$$

3.3 ECC decryption algorithm

After receiving the information C_s, user B (U_B) perform decryption to compute P_s using formula (4).

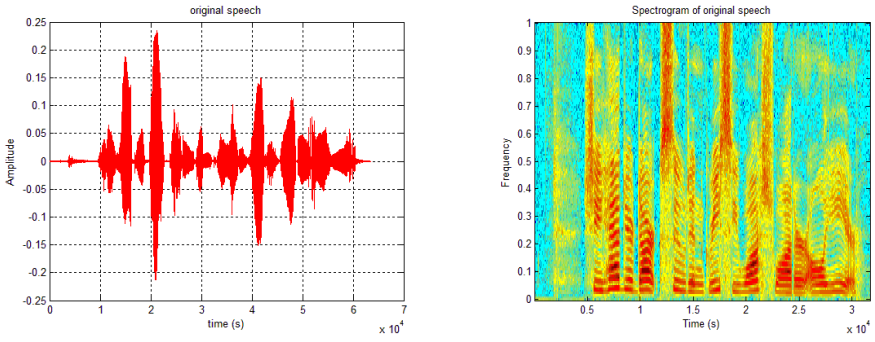
$$P_s = C_2 - k_B C_1 \tag{4}$$

So, the receiver gets the same point P_s, that should be decoded to coded speech.

4 Encryption Simulation

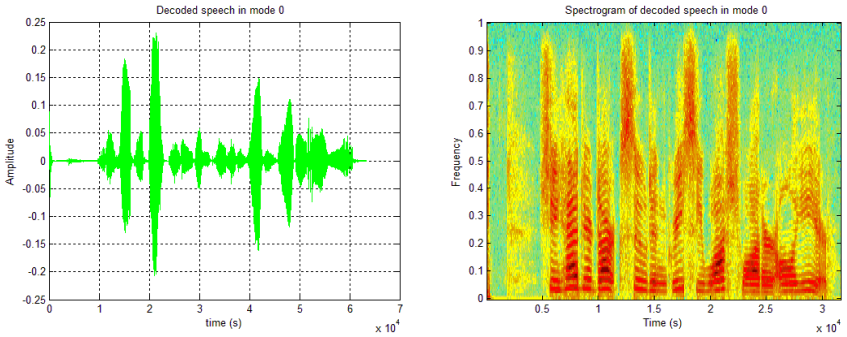
In order to design an ECC encryption system, several experiments are carried out to test the ECC encryption efficiency of the presented wideband speech cryptosystem. The proposed cryptosystem is applied to different G.722.2 speech samples with different sizes. The analysis of the proposed scheme' performance is performed by including waveform display, spectrogram display, EMBSD [27] and WB-PESQ [28] for objective quality assessment and MOS [29] for subjective.

In the experiments, we focus on the quality of encrypted and decrypted speech signals. Before encryption, we represent the original and decoded speech signals as shown in Fig.5 and Fig. 6 respectively in both time and frequency domains. The waveforms representation (Fig.5-a and Fig. 6-a) and the spectrograms display (Fig.5-b and Fig. 6-b) of the original and decoded speech seems identical. Note that we have presented the waveform and the spectrogram of only a single speech file due to space requirements.



(a) (b)

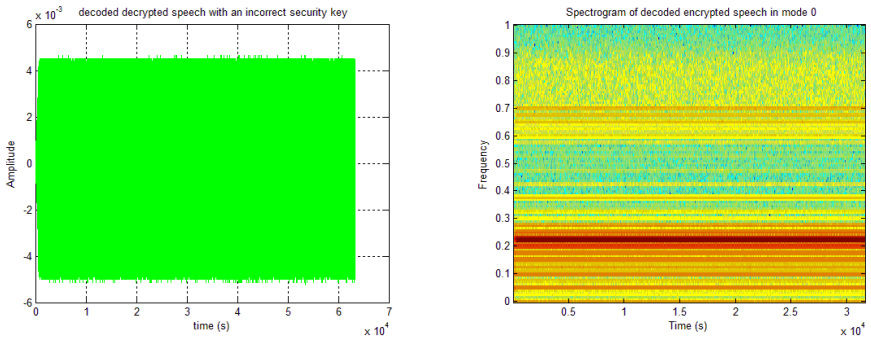
Fig.5. (a) Original speech (b) Its spectrogram



(a) (b)

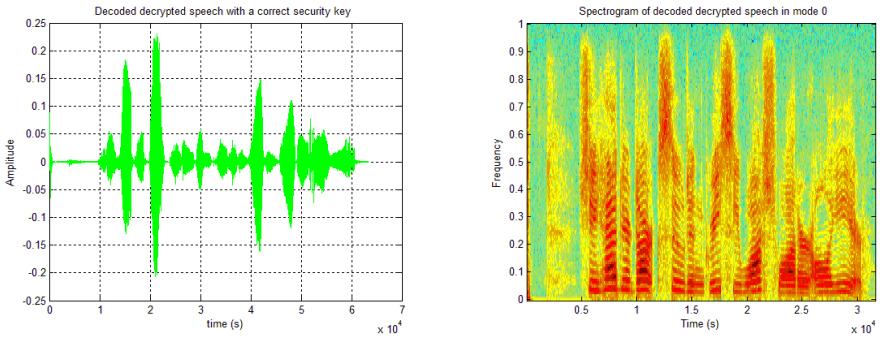
Fig.6. (a) Decoded speech in mode 0 (b) its Spectrogram

The speech file extracted from TIMIT database [30] is coded in different modes using AMR-WB G.722.2 CS-ACELP. The resulting bit streams are encrypted by the ECC technique. Fig. 7-a shows the ECC encrypted speech decrypted and decoded with an incorrect security keys and its Spectrogram (see Fig. 7-b). We can see from these figures that the encrypted speech signals are similar to white noise, which signifies the unintelligibility of the encrypted speech signal. However, the decoded and decrypted speech with a correct security key (see Fig. 8-a) and its spectrogram (see Fig. 8-b) show their similarity with G.722.2 CS-ACELP at the expense of doubling the size of the encrypted file.



(a) (b)

Fig.7. (a) Decoded decrypted speech with an incorrect security keys (b) its Spectrogram



(a) (b)

Fig.8. Encryption using mode 0 of WB-G722.2: (a) Decoded decrypted speech with a correct security key (b) its spectrogram

The evaluation of wideband speech coding quality includes the EMBSD and PESQ measures as objective tests and MOS as a subjective test for three speech files extracted from TIMIT database.

The EMBSD tool gives the comparison between two speech files, it indicates a 0 value for two identical speech files and a greater value as the distortion increases. Fig. 9 shows the EMBSD values of the AMR-WB codec in mode 2 as well as the decrypted speech files with a correct security key; that are near zero and indicate their good quality. However, the values of speeches decrypted with an incorrect security key are significantly larger than zero, proving the efficiency of the encryption schemes.

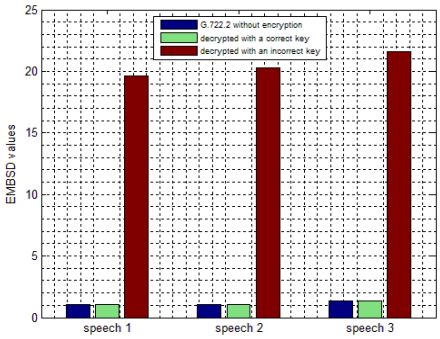


Fig.9. EMBSD values of G.722.2 codec used in mode 2 and of decrypted speeches with a correct / an incorrect key

WB-PESQ and MOS assessments give the comparison between two speech files, WB-PESQ values lie in the range -0.5 to 4.5 and give a score of 4.5 for two identical files. However, MOS scores lie in the range 1 to 5 (1: Unsatisfactory, 2: Poor, 3: Fair, 4: Good, 5: Excellent). Fig. 10 and 11 show the WB-PESQ and MOS results of G.722.2 codec used in mode 2 and of decrypted speech files with a correct / incorrect security key. We can see that the WB-PESQ and MOS values for the encoded speech files using AMR-WB G.722.2 and decrypted decoded speeches with a correct security key are close to 4, which indicate their good quality. In return, when using the decryption of the decoded speeches, the values are close to one, corresponding to their bad quality and proving the efficiency of the used encryption scheme.

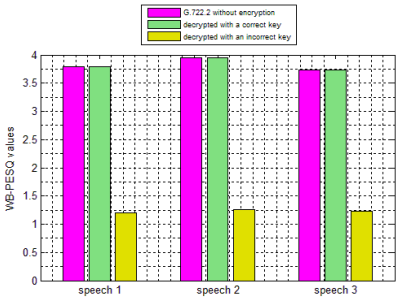


Fig.10. WB-PESQ values of G.722.2 codec used in mode 2 and of decrypted speeches with a correct / an incorrect key

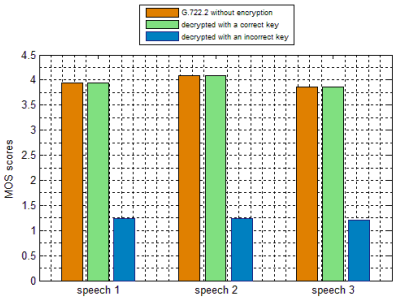


Fig.11. MOS scores of G.722.2 codec used in mode 2 and of decrypted speeches with a correct / an incorrect key

5 Conclusion

In this paper, an encryption scheme for AMR-WB ITU-T G.722.2 speech based on ECC for securing transmitted speech signals have been implemented. ECC has been shown to offer an RSA- grade security with smaller key size.

During the steps involved in implementation, the ECDH generation of public and private keys is proven to be safe and secure. In wireless communication, and more specifically, in speech encryption, the obtained results show the effectiveness, in terms of security and confidentiality of speech data, at the expense of a doubled speech file size resulted at the point mapping step of EC.

Therefore, the major advantages of ECC lie in its simplicity and high security, with the disadvantage of increased file size. To meet this challenge and as future directions, a hybrid cryptosystem combining ECC method for generating security keys with conventional symmetric encryption standards such as AES (Advanced Encryption standard) to encrypt and decrypt the speech data is proposed.

References

1. KHRISAT, M., Zaini, H. G., & AlQadi, Z. (2021). Simple, Qualities, Efficient, and Secure Method to Encrypt Voice Signal. *International Journal of Computer Applications*, *183*(7), 25-29.
2. Mua'ad, M., Aldebei, K., & Alqadi, Z. A. (2022). Simple, Efficient, Highly Secure, and Multiple Purposed Method on Data Cryptography. *Traitement du Signal,* 39(1), 173-178.
3. Kulkarni, S. A., & Patil, S. B. (2015, January). A robust encryption method for speech data hiding in digital images for optimized security. In *2015 International Conference on Pervasive Computing (ICPC)* (pp. 1-5). IEEE.
4. Ambika, D., & Radha, V. (2012). Secure Speech Communication–A Review. *International Journal of Engineering Research and Applications*, *2*(5), 1044-1049.
5. Mohammed, S. J., & Taha, D. B. (2022, March). Performance Evaluation of RSA, ElGamal, and Paillier Partial Homomorphic Encryption Algorithms. In *2022 International Conference on Computer Science and Software Engineering (CSASE)* (pp. 89-94). IEEE.
6. Van der Elst, V., Wilssens, R., Jocqué, J., Sennesael, J., Verhaevert, J., Van Torre, P., & Rogier, H. (2022, March). Platform for Multi-User Channel-Based Encryption of Speech Communication with AES on 2.45 GHz. *In 2022 16th European Conference on Antennas and Propagation (EuCAP)* (pp. 1-5). IEEE.
7. Mua'ad Abu-Faraj, A. A., & Hyari, Z. A. (2022). A DUAL APPROACH TO DIGITAL AUDIO SIGNAL CRYPTOGRAPHY. *Journal of Southwest Jiaotong University, 57*(1).
8. Al-Hazaimeh, O. M., Abu-Ein, A. A., Al-Nawashi, M. M., & Gharaibeh, N. Y. (2022). Chaotic based multimedia encryption: a survey for network and internet security. *Bulletin of Electrical Engineering and Informatics*, *11*(4).
9. Mokhnache, S., Daachi, M. E. H., Bekkouche, T., & Diffellah, N. (2022). A Combined Chaotic System for Speech Encryption. *Engineering, Technology & Applied Science Research, 12*(3), 8578-8583.
10. Patil, P., Narayankar, P., Narayan, D. G., & Meena, S. M. (2016). A comprehensive evaluation of cryptographic algorithms: DES, 3DES, AES, RSA and Blowfish. *Procedia Computer Science*, 78, 617-624.
11. Abouelkheir, E., & El-Sherbiny, S. (2022) Enhancement of Speech Encryption/Decryption Process Using RSA Algorithm Variants. *Human-centric Computing and Information Sciences, 12 (2022)*
12. J. I. Okonkwo, G. O. Ozor, and F. A. Okoye, "Performance analysis of RSA algorithm for audio data security in communication networks," *International Journal of Latest Technology in Engineering, Management & Applied Science*, vol. 8, no. 9, pp. 48-52, 2019.
13. Intila, C., Gerardo, B., & Medina, R. (2019, February). A study of public key 'e'in RSA algorithm. *In IOP Conference Series: Materials Science and Engineering* (Vol. 482, No. 1, p. 012016). IOP Publishing.
14. Al-Ghamdi, S., & Al-Sharari, H. (2022, February). Improve the security for voice cryptography in the RSA algorithm. *In 2022 International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1-4). IEEE.
15. Guruprakash, J., & Koppu, S. (2020). EC-ElGamal and Genetic algorithm-based enhancement for lightweight scalable blockchain in IoT domain. *IEEE Access*, *8*, 141269-141281.
16. Lara-Nino, C. A., Diaz-Perez, A., & Morales-Sandoval, M. (2018). Elliptic curve lightweight cryptography: A survey. *IEEE Access*, 6, 72514-72550.
17. Vagle, J. L. (2000). A gentle introduction to elliptic curve cryptography. *Cambridge (MA): BBN Technologies*.

18. Deepthi, P. P., Nithin, V. S., & Sathidevi, P. S. (2009). Implementation and analysis of stream ciphers based on the elliptic curves. *Computers & Electrical Engineering*, *35*(2), 300-314.
19. Priyanka, S., & Hemalatha, B. (2016). Speech Data Encryption and Decryption Using Elliptic Curve Cryptography". *International Journal of Research in Computer Science*, *3*(1), 48-53.
20. Hureib, E. S., & Gutub, A. A. (2020). Enhancing medical data security via combining elliptic curve cryptography and image steganography. *Int. J. Comput. Sci. Netw. Secur.(IJCSNS)*, *20*(8), 1-8.
21. Subashri, T., Arjun, A., & Ashok, S. (2014, July). Real time implementation of Elliptic Curve Cryptography over a open source VoIP server. *In Fifth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.
22. Miller, V. S. (1985, August). Use of elliptic curves in cryptography. *In Conference on the theory and application of cryptographic techniques* (pp. 417-426). Springer, Berlin, Heidelberg.
23. Jasra, B., Saqib, M., & Moon, A. H. (2021, April). Mapping Images Over Elliptic Curve For Encryption. *In 2021 6th International Conference for Convergence in Technology (I2CT)* (pp. 1-5). IEEE.
24. Khoirom, M. S., Laiphrakpam, D. S., & Tuithung, T. (2021). Audio encryption using ameliorated ElGamal public key encryption over finite field. *Wireless Personal Communications*, 117(2), 809-823
25. Wideband coding of speech at around 16 kbps using Adaptive Multi-RateWideband (AMR-WB). *ITU-T Standard G.722.2* (2003)
26. Boumaraf, M., & Merazka, F. (2021). Secure speech coding communication using hyperchaotic key generators for AMR-WB codec. *Multimedia Systems*, *27*(2), 247-269.
27. Yang W.: 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
28. ITU-T Rec. : Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs, *ITU-T Recommendation P.862.2* (November 2005)
29. ITU-T. Mean opinion score (MOS) terminology Recommendation P.800.1 (July 2006)
30. NIST,Timit Speech Corpus, NIST (1990)