An Efficient Audio Encryption Scheme Based on Finite Fields

Dawood Shah1, Tariq Shah1, Muhammad Mazyad Hazzazi2, Muhammad Imran Haider1,3, Amer Aljaedi4, Iqtadar Hussains

1Department of Mathematics, Quaid-i-Azam University Islamabad
2Department of Mathematics, College of science, King Khalid University, Abha, 61413, Saudi Arabia
3Department of Mathematics, Gomal University D.I Khan Pakistan.
4 College of Computing and Information Technology, University of Tabuk, Tabuk 71491, Saudi Arabia

*Corresponding author: Dawoodshah254@gmail.com. *Corresponding author: Dawoodshah254@gmail.com.

Finite fields are well-studied algebraic structures with enormous efficient properties which have applications in the fields of cryptology and coding theory. In this study, we proposed a lossless binary Galois field extension-based efficient algorithm for digital audio encryption. The proposed architecture hired a special type of curve in the diffusion module which depends on efficient elliptic curve arithmetic operations. So, it generates good quality pseudo-random numbers (PRN) and with slight computational efforts, it produces optimum diffusion in the encrypted audio files. For the confusion module, a novel construction mechanism of block cipher has been employed which includes prominent arithmetic operations of binary Galois field inversion and multiplication operations. The suggested scheme generates multiple substitution boxes (S-boxes) by using a higher-order Galois field. Thus, the replacement with multiple S-boxes generates effective perplexity in the data and provides additional security to the ciphered audio. The investigational outcomes through different analyses and time complexity demonstrated the ability of the technique to counter various attacks. Furthermore, as a consequence of a rapid and simple application of the binary finite field in hardware and software, the proposed scheme is more appropriate to be applied for data security.

INDEX TERMS Galois field; Elliptic curve; Random numbers generator; S-box; Audio encryption scheme

I. INTRODUCTION

In recent decades, due to the speedy development in science and digital technologies, the role of multimedia data in social life has been increased. Multimedia data are used in various fields such as education, engineering, mathematics, art, advertisement, military, medicine, scientific research, and many more. This excessive growth of multimedia data boosts the importance of multimedia data processing tools and digital documentation. This access to multimedia data through the internet has created inappropriate prospects which are hazardous for the confidentiality and integrity of the multimedia data. To encounter these threats, the domain of multimedia data security gains broad attention. A considerable number of algorithms have been established to protect personal information over open networks. The most prominent field to provide security is cryptography which can be further divided into symmetric and asymmetric key methods. Some of the prominent algorithms such as data encryption standard (DES) [1], international data encryption standard (IDES) [2], triple data encryption standard (TDES) [3], advanced encryption standard (AES) [4], and RSA are widely used for the security purposes and considered as well protected and reliable. Since multimedia data contains a large amount of data that is highly correlated, therefore, the only dependency on the algorithms like AES, RSA, and DES is not good enough for multimedia data security. Since multimedia data contains a large amount of data that is highly correlated, therefore, the only dependency on the algorithms like AES, RSA, and DES is not good enough for multimedia data security. In literature, a considerable number of encryption methods for the security of multimedia data have been introduced. For instance, the encryption algorithm for a digital image depending upon chaotic systems and finite algebra of Galois fields is given in [5-14]. In addition, more complex algebraic structures such as elliptic curves are widely used for digital image security [16,17]. Recently, Hua, Zhongyun, et al. introduced novel color image encryption schemes that are based on orthogonal Latin squares and parallel compressive sensing and adaptive thresholding sparsification [17,18]. The audio files contain massive data capacity and are somehow different from the
other multimedia data. Therefore, there should be a separate algorithm for the protection of digital audio.

**Literature Review**

In the literature, numerous digital audio encryption algorithms are presented. De Martin and Servetti [19] proposed an encryption algorithm for the encryption of telephonic speech relying based on the perception method in 2002. The author recommended two techniques for the encryption of partial speech. The first scheme was envisioned to have a high bit of rate and low-security capability. Consequently, the cryptanalysis could easily reveal the ciphered speech. But, the second scheme is considered to encrypt additional bitstream, thus provides more security to the ciphered audio. Thorwirth et al. [20] gave an algorithm for the selective encryption technique of perceptual audio coding based on the standard compression in which the author’s main focus was on examining the encryption of the encoded MP3 files. Subsequently, Servetti et al. [21] proposed an MP3 audio selective partial encryption algorithm; the suggested algorithm has considerably low time complexity and also preserves the contents of the audio information but unfortunately compromises the quality of the original audio sequences to preserves the perceptual information. Next, in 2004, Bhargava et al. [22] proposed four fast encryption algorithms for MPEG video, where a key is used to randomly change the sign bits of the Discrete Cosine Transform (DCT) coefficients and/or the sign bits of motion vectors. These schemes put on a small overhead to the MPEG codec. Grange et al. [23] introduced a new framework that relies on randomized arithmetic coding for the security of multimedia data. In the recommended framework, the security purpose of multimedia data was achieved by producing some randomness in the arithmetic coding process. In 2008, Yan et al. [24] introduced progressive multimedia data security by scrambling audio data in a compressed domain. In the proposed scheme, the secrete MP3 audio was twisted via a shared secret key before transmission. However, Au and Zhou in [25] showed that the Yan scheme is conquerable against key search attacks. In [26], Neto and Lima presented an encryption scheme for digital audio that relies on cosine number transform. The encryption procedure recursively applies to a block of uncompressed audio data and uses simple overlapping to select the block and produce diffusion in the encrypted data.

1.1 Motivation

The majority of these audio encryption techniques have a deficiency of cryptanalysis and insufficient security estimations were achieved to confirm the permanency of these cryptosystems to counter the malicious attacks. For this reason, a strong algorithm is required to enhance the security of audio data against different attacks. Moreover, the development of cryptographic applications on hardware attempted to take benefit of the comfort implementation of Galois fields to boost the performance and abbreviate the costs. These properties of finite fields attract us towards the development of one new algorithm for digital audio data security based on a Galois field.

1.2 Our Contribution

In this manuscript, we designed a novel lossless audio data encryption scheme based on arithmetic operations of an elliptic curve over a finite field \( \mathbb{Z}_p \) and binary Galois field \( GF(2^n) \). The basic aim of this scheme is to provide a strong algorithm to ensure authentication and integrity. The arithmetic operations in the elliptic curve are performing efficiently, so in the begun of the proposed scheme special type of curve based on the elliptic curve, operations are used to generate a good quality sequence of random numbers. The generated sequence is subsequently used to defuse the matrix of the audio data. The confusion module of the scheme is executed through multiple substitution boxes having higher nonlinearity. The experimental results demonstrate the efficiency of the proposed scheme against various attacks.

The rest of this paper is structured as: We introduced the basic notion of the elliptic curve and finite extension field in section 2. The methodology of the proposed encryption technique is presented in section 3. Section 4 represents the simulation and performance results of the proposed scheme. In the last section, we concluded the findings.

**II. PRELIMINARIES**

A. CONSTRUCTION OF GALOIS FIELD

Let \((R, +, \cdot)\) be a commutative ring with identity. An ideal \(I\) is a subring of the ring \(R\) satisfying the condition \(aI \subseteq I\) for every element \(a\) in the ring \(R\). An ideal is said to be a maximal ideal if it does not properly contain in any other proper ideal of \(R\). \(F[w]\) is a polynomial ring in one indeterminate \(w\) having coefficients from the field \(F\). The ring \(F[w]\) is, in fact, a Euclidean domain and hence a principal ideal domain (PID). A polynomial \(h(w)\) in \(F[w]\) is said to be irreducible if it cannot be written as the product of non-unit polynomials in \(F[w]\). Accordingly for a finite field \(F_q\) and maximal ideal \(< h(w) >\) generated by \(n\) degree primitive irreducible polynomial \(h(w) \in F_q[w]\), the quotient ring \(F_q[w] / < h(w) >\) is a field known as the Galois field, an extension of the field \(F_q\) and it is denoted as \(GF(q^n)\). The nonzero elements of \(GF(q^n)\) forms a multiplicative group known as Galois cyclic group.

B. ELLIPTIC CURVE
An elliptic curve over a finite field $F_p$ is a plot of equation solutions $E: y^2 = x^3 + ax + b \pmod{p}$, where $a, b \in F_p$, satisfy the equation $(4a^3 + 27b^2) \neq 0 \pmod{p}$. All these points (solutions) with point of infinity (neutral element) $O$ form an abelian group, which is denoted by $E(F_p)$. The formation process of the group is as under.

C. ELLIPTIC CURVE ARITHMETIC

Let $P_1 = (u_1, v_1)$ and $P_2 = (u_2, v_2)$ be any two points lie on the graph. Define $P_1 \oplus P_2 = (u_3, v_3) = R$ as under.

i. If $P_1 \neq P_2$ with $u_1 \neq u_2$, then

$$(u_3, \ v_3) = \left(\lambda^2 - u - u_2 \mod p, \ \lambda(u_1 - u) - v_1 \mod p\right) \quad (1)$$

And

$$\lambda = \frac{v_2 - v_1}{u_2 - u_1} \mod p \quad (2)$$

ii. If $P_1 \neq P_2$ with $u_1 = u_2$ but $v_1 \neq v_2$, then $P_1 + P_2 = O$.

iii. If $P_1 = P_2$ with $v_1 \neq 0$, then

$$(u_3, \ v_3) = \left(\lambda^2 - u - u_2 \mod p, \ \lambda(u_1 - u) - v_1 \mod p\right) \quad (3)$$

And

$$\lambda = \frac{3u^2 + a}{2v} \mod p \quad (4)$$

iv. If $P_1 = P_2$ with $v_1 = 0$, then $P_1 + P_2 = O$.

v. Furthermore, define

$$P + O = P \text{ for all } P \text{ on } E$$

On the above footprints, one can easily show that $E(F_p)$ is an abelian group with an identity element $O$.

D. SINGULAR POINT

Let $(u, v)$ be a point on affine curve $f(x,y) = 0$ over field $K$. Then the point $(u, v)$ is said to be a singular point of the curve $f(x, y) = 0$ if both partial derivatives $\frac{\partial f}{\partial u}$ and $\frac{\partial f}{\partial v}$ vanish at $(u, v)$.

The following theorem is from [2].

**Theorem 2.2.1.** Let $E^{\text{ns}}(F_p)$ be the set of non-singular points on $E_{F_p}^a$ with $y^2 = a$ for some $\gamma \in F_p$ against a curve $E_{F_p}^a: y^2 = x^3 + ax$ over a finite field $F_p$, with $0 \neq a \in F_p$.

Then the homomorphism

$$\varphi_{\gamma}: E^{\text{ns}}(F_p) \to F_p^*$$

is defined as

$$\varphi_{\gamma}(u, v) = \frac{\gamma + v}{v - \gamma}; \text{ and } \varphi_{\gamma}(O) = 1 \quad (6)$$

is an isomorphism.

III. AUDIO ENCRYPTION SCHEME

The audio technology is used to store, manipulate, reproduce, and generate the sound using the arrays of the audio signals encoded in digital format. Digital audio also refers to the sample of discreet sequences, which are choosing from the audio wave format. The digital audio data is virtually consisting of discreet sockets that indicate the amplitude of the wave of digital data. In this study, we manipulate the discrete sockets of the digital audio and encrypt the original content of the original audio. The proposed encryption technique is planned to protect the uncompressed digital audio integer 16 (int16) format. We represent the matrix set of the plain audio by $A$ of dimension $M \times N$ for $N \in [1, 2]$. In the next subsection, we discuss step by step procedure of the encryption scheme in detail.

A. PROPOSED RANDOM NUMBER GENERATOR

The generation of random numbers plays a significant role in various multimedia data security applications. The debauched research comes up with numerous random number generation schemes. An elliptic curve is also the widely used generation of random numbers. In general, the elliptic curve-based random number generation procedure utilizes group law and the arithmetic operation of the elliptic curve. In this section, we are giving an efficient scheme for the generation of random numbers based on the elliptic curve operation. The proposed scheme generates distinct random numbers with enough long periods. At the begin of the encryption procedure, generate a sequence of distinct pseudo-random numbers with a long period greater than the length of the audio data. Along this select a large prime $p$. Then generate the curve $E_{\alpha}^p: y^2 = x^3 + ax \mod p$ through brute force technique. Subsequently, use the following map to transmute points of the curve $E_{\alpha}^p(u, v)$ into the field $F_p$.

$$\varphi_{\gamma}: E_{\alpha}^p \to F_p$$

Defined by

$$\varphi_{\gamma}(u, v) = \frac{v + yu}{v - yu} \quad (7)$$

Where $\gamma \in F_p^*$ is the squared element such that $y^2 = a$ and $(u, v)$ is the element of the curve $E_{\alpha}^p$. The map $\varphi_{\gamma}$ is the isomorphism between $E_{\alpha}^p$ and $F_p$ by Theorem 2.2.1. The consequential set $\varphi_{\gamma}(E_{\alpha}^p)$ is a sequence of random numbers in the field $F_p$. Afterward, use the sequence of random numbers to shuffle the matrix $A$ and get a new data set $A_2$. In this study, we fixed the elements $\alpha = 2$ and $p = 99991$ and generate a sequence of random numbers by using the above procedure. The generated sequence is then analyzed by the NIST test, the results are tabulated in Table 4.

I. MULTIPLE S-BOXES CONSTRUCTION SCHEME

The S-box plays a significant role in symmetric key cryptography. In general, S-box uses in the substitution module of the cryptosystem and produces confusion in the cipher data. Therefore, the confusion creating the capability of the cryptosystem depends on the quality of the S-box. Since audio contains a large amount of data. So, in the proposed cryptosystem we used multiple S-boxes to produce more randomness in the encrypted data. To construct multiple S-boxes, we introduced a novel S-box construction scheme based on Galois field $GF(2^n)$. The traditional S-box construction schemes are based on the finite field of order 256. However, the proposed construction scheme for
multiple S-boxes generations is based on the Galois field of order greater than 256. Here we discussed the general idea of the construction scheme. Initially, define a bijective map from the Galois field \(GF(2^n)\) onto \(GF(2^m)\). The mapping is defined as follows.

\[
S: GF(2^n) \rightarrow GF(2^m) \\
h \mapsto \hat{v} \left( \frac{x(h) + \hat{y}}{2} \right) + \hat{u} 
\tag{8}
\]

In equation (8), \(x, \hat{v}, \hat{y}\) and \(\hat{u}\) are the elements of the Galois field \(GF(2^n)\). Afterward, define an inclusion map from the Galois field \(GF(2^n)\) onto \(GF(2^m)\). The mathematical representation is given as follows.

\[
I_k: GF(2^n) \rightarrow GF(2^m) \\
I_k(\sum_{i=0}^{n} a_i x^i) = \left( \sum_{i=0}^{m} a_i x^i \right) I_k \text{ if } i \leq m - 1 \\
0 \text{ if } i > m - 1
\tag{9}
\]

Where \(k \geq 2\) and \(n\) is strictly greater than \(m\). The composition map \(I_k \circ S\) generates \(m \times m\) S-box. With this process, one can generate \(n - m\) number of S-boxes.

**II. PROPOSED ALGORITHM**

**Step 1.** Generate a binary matrix \(M\) having dimension \(M \times N\) to identify the location of the negative integers in the matrix of the original audio.

\[
M_{ij} = \begin{cases} 
-1 & \text{if } A_{ij} < 0 \\
1 & \text{if } A_{ij} \geq 0 
\end{cases}
\tag{12}
\]

Where \(A_{ij}\) indicates the sample of the audio data at \((i,j)\) position. The aim of generating binary matrix \(M\) is to specify the position of the negative samples.

**Step 2.** Select a prime number \(p > M \times N\) and generate a sequence \(\sigma\) of random number via the proposed random number generator we have discussed in section 3.2.1. Afterward, reduce the length of the sequence and then use the new sequence and shuffle the matrix of the original audio. The equation is given as follows.

\[
\delta_i = \begin{cases} 
\sigma_i & \text{if } \sigma_i \leq MN \\
1 & \text{if } \sigma_i > MN 
\end{cases}
\]

\[
A_{ij} = A_{\delta_i \delta_j}
\tag{15}
\]

Where \(\delta_i, \delta_j\) denote the position of the integer value \(A_{\delta_i \delta_j}\) in the newly shuffled matrix \(A_e\). The waveform and the spectrogram graph of the shuffled audio are shown in Fig 1(b) and Fig 2(b) respectively. From the figures, one can observe that the permutation step caused optimum disruption in the plain audio.

**Step 3.** Next, we will convert the entries of the matrix \(A_e\) using the absolute function from the set in the range \([-2^{12}, 2^{12} - 1\) to the elements of the Galois field \(GF(2^{15})\). Consequently, get a new matrix \(A_G\).

**Step 4.** Subsequently, convert the elements of the Galois field \(GF(2^{15})\) to the elements of the Galois field \(GF(2^8)\) and Galois field \(GF(2^7)\) by using the following map.

\[
\psi: GF(2^{15}) \rightarrow GF(2^8) \times GF(2^7)
\]

Defined by

\[
\psi \left( \sum_{i=0}^{14} a_i x^i \right) = \left( \sum_{i=0}^{7} a_i x^i, \sum_{i=8}^{14} a_i x^i \right) \tag{16}
\]

Where \(x_i \in \{0, 1\}\), by using the above mapping split data in the matrix \(A_G\) into two matrices \(A_p^1\) and \(A_p^2\) containing elements of the Galois fields \(GF(2^8)\) and \(GF(2^7)\) respectively, to reduce the time complexity.

**Step 5.** Divide the blocks \(A_p^1\) into four subblocks. Then generate four \(8 \times 8\) S-boxes using the proposed S-box construction method, which we have discussed in subsection 3.2.2. Afterward substitute each subblock with a different S-box and then combine all the subblocks. Similarly, divide the block \(A_p^2\) into four subblocks and generate four \(7 \times 7\) S-boxes using the proposed S-box construction method. Then substitute each subblock with a different S-box and combine all the substituted subblocks. Consequently, get new blocks \(A_s^1\) and \(A_s^2\).

**Step 6.** Combine the resultant matrices \(A_s^1\) and \(A_s^2\) using the inverse map of the map, which we have discussed in step 4. The inverse map is given as follows.

\[
\psi^{-1}: GF(2^8) \times GF(2^7) \rightarrow GF(2^{15})
\]

Defined by

\[
\psi^{-1} \left( \sum_{i=0}^{7} a_i x^i, \sum_{i=8}^{14} a_i x^i \right) = \sum_{i=0}^{7} a_i x^i + \sum_{i=1}^{6} a_{i+8} x^{i+8} \tag{17}
\]

As a result of the above map, we get a new matrix \(A_s1\) containing elements of the Galois filed \(GF(2^{15})\).

**Step 7.** Then mask each element of the matrix \(A_s1\) and produce more diffusion in encrypted audio. Firstly, generate a sequence \(\rho\) of the nonrandom number of lengths \(M \times N\). Subsequently, use mode operation and convert the elements...
of the sequence into the elements of the Galois field $GF(2^{15})$.

$$A_{s2}(i,j) = (A_{s1}(i,j) + (\rho(i,j))^{-1}) \quad (18)$$

Where $A_{s1}(i,j)$ and $\rho(i,j)$ are the elements of the Galois field $GF(2^{15})$ and $(i,j)$ signify integer position in the matrix. Because of the equation (18) get a new matrix $A_{s2}$.

Step 8. Eventually, use the binary matrix $M$ and convert the entries of the matrix $A_{s2}$ from the Galois field $GF(2^{15})$ into the set of integers sixteen $\{-2^{15}, 2^{15} - 1\}$. The mathematical representation is given as follows.

$$A_E(i,j) = \begin{cases} 
A_{s2}(i,j) & \text{if } M(i,j) = 1 \\
-A_{s2}(i,j) & \text{if } M(i,j) = -1 
\end{cases} \quad (19)$$

The resultant matrix is then converted into the Audio file which is our ciphered audio file. Our new proposed encryption technique is functional to many audio files of various sizes and different characters. The waveform of the encrypted audio is given in Fig. (1). From the figure, it is evident that the waveform of the encrypted is uniform. Accordingly, the proposed technique is proficient in safe the actual content of the original audio. The decryption process of the scheme is the same as the encryption.

### IV. SECURITY ANALYSIS

It is mandatory for a standard encryption scheme to counter different kinds of attacks that try to hit the confidentiality, integrity, non-repudiation, and authentication of the data. Here, we evaluate the strength and robustness of the proposed technique against different malicious attacks. These all analyses are performed by using MATLAB 2019(b) on a personal computer. To scrutinize our scheme, we amalgamate different audio models with music, speech, and other characters and encrypt these models with our proposed technique using multiple keys. The wave version of plain, encrypted, and decrypted files is represented in Fig.1. From Fig.1 it is obvious that the amplitude of original and encrypted audios has no resemblance with each other as the encrypted audios are uniform in nature. This depicts that audio is properly encrypted. In addition to this, the waveform of decrypted and the original audio is also similar as made known in Fig. 1 (d). In upcoming sections, the proposed scheme undergoes different analyses which include histogram analysis, keys space analysis, key sensitive analysis, and Correlation.

![FIGURE 1 Waveform of the (a) original Audio (b) Permuted Audio. (c) Encrypted Audio (d) Decrypted Audio](image)

#### A. SPECTROGRAM ANALYSIS

To perform the spectral analysis of sound, it is recommended to use spectrogram analysis. This analysis is demarcated as two-dimensional graph and different colors represent its third dimension. It is considered as the pictorial illustration of the frequency of the spectrum that fluctuates concerning time. The third-dimension color identifies the amplitude or loudness of the sound at a precise time. The low amplitude
is specified by using red and blue colors whereas the bright color indicates the stronger amplitude. The results of the spectrogram analysis of our encryption scheme are given in Fig. 2. The spectrogram graphs of original and encrypted audio files are represented in Fig. 2(a) and Fig. 2(c) respectively. The audio file is effectively encrypted which is evident from the uniformity of the spectrogram graph of the encrypted audio files. This encrypted audio file has a strong amplitude and an altogether different spectrogram from the original audio.

![Spectrogram Graphs](image)

**FIGURE 2:** Spectrogram Graph of (a) Original Audio (b) Permuted Audio (c) Encrypted Audio, (d) Decrypted Audio

### B. HISTOGRAM ANALYSIS

To assess the quality of any encryption scheme against statistical attacks, it is recommended to perform histogram analysis. It is most likely that cryptosystems change the original information into noise and generate randomness in the data. It is observed that in an efficient cryptosystem most likely the encrypted data does not offer any information which helps to decipher the encrypted data free from the requirement of the confidential key. In such cryptosystems, the original data is encrypted with similar possible values.

Figure 3 represents the outcomes of histogram analysis of our encryption scheme. The histogram of the original audio is graphically represented in Fig. 3(a) and Fig. 3(c) and the histogram of the cryptographed audio is made known in Fig 3(b) and Fig 3(d). One can see that the original audio signal histogram is haphazard and heading towards a single point, but the histogram of the encrypted audio file is uniform. It concludes that our technique is shown strength to counter any statistical outbreak and it’s extremely hard to extract info from the encrypted information.
C. CORRELATION ANALYSIS

The correlation coefficient is one of the analyses which are performed to evaluate the ability of any cryptosystem to fight against various statistical attacks. As data is strongly correlated in multimedia applications so, a robust cryptosystem must intrude on the correlation among the segment of the data. In this analysis, the focus is to observe the correlation between identical sections of the data. The correlation coefficient is given by:

\[ \gamma_{uv} = \frac{cov(p, q)}{\sqrt{D(p)D(q)}} \]  

(20)

Where

\[ cov(p, q) = \frac{1}{P} \sum_{i=1}^{P} p_i - \mathcal{E}(p) (q_i - \mathcal{E}(q)) \]  

(21)

\[ D(p) = \frac{1}{P} \sum_{i=1}^{P} (p_i - \mathcal{E}(p))^2 \]  

(23)

And

\[ \mathcal{E}(p) = \frac{1}{P} \sum_{i=1}^{P} p_i \]  

(24)

Where sample at \( i_{th} \) the position is signified by \( p_i \) and \( q_i \) indicates the equivalent adjacent sample. Commonly, correlation analyses of the data are performed for horizontal, vertical, and diagonal directions but as our scheme is dealing in audio data so for the single string data only the horizontal direction is taken for correlation analysis. The outcomes of the correlation analysis are shown in Table 1. It indicates that the original audio correlation is equivalent to 1, which depicts the sections in the audio data have a strong correlation. On the other hand, the correlation analysis for the ciphered audio is nearly a value of 0, i.e., the proposed technique analytically intrudes the correlation of the audio segment. Correlation analysis of the original and the encrypted audio is represented in figure 4. It establishes that our scheme gradually minimizes the intercorrelation of the audio file. For this reason, our proposed technique is robust against malicious statistical attacks.

D. INFORMATION ENTROPY

For coded information, the amount of uncertainty is measured by using information entropy analysis. The entropy is directly proportional to the rate of uncertainty i.e., higher uncertainty in encrypted audio files depicts that it has the higher entropy. We can represent entropy as

\[ H = - \sum_{k=0}^{L} P(k) \log_2 P(k) \]  

(25)

Where \( \mathcal{L} \) directs the grayscale value of the audio file and \( P(k) \) implies the probability of the presence of the grey-value \( k \). For our case, the audio file has a value of 16 in correspondence to the theoretical value of \( H \). So, the cryptosystem is considered to be well-secured if the information entropy of the ciphered file is exactly 16. We examine our new proposed scheme by using information entropy analysis and the outcomes are organized in Table 2. It is obvious from the Table that the information value of our

---

**Table 1: Correlation results of various Audio**

| No | Audio                  | Plain Audio | Ciphered Audio |
|----|------------------------|-------------|----------------|
| 1  | Audio                  | 0.9945      | -0.0081        |
| 2  | Animal sound.wav       | 0.7317      | -0.0033        |
| 3  | Alarm sound.wav        | 0.8368      | -0.0039        |
| 4  | Applause sound.        | 0.9962      | 0.0011         |
| 5  | Bells sound.           | 0.9924      | -0.0031        |
| 6  | Birds sound.wav        | 0.9933      | -0.0029        |
| 7  | Female sound.wav       | 0.9886      | -0.0010        |
| 8  | 44100 Hz tone.wav      | 0.9464      | 0.0017         |
| 9  | Male sound.wav         | 0.9523      | 0.0030         |
| 10 | Machine sound.wav      | 0.9935      | -0.0040        |
| 11 | Music sound.wav        | 0.9847      | -0.0081        |
| 12 | Ref. [27]              | 0.001699    |                |
| 13 | Ref. [28]              | 0.0119      |                |
| 14 | Ref. [29]              | 0.0263      |                |

---

**FIGURE 3:** Histogram Analysis (a) Histogram of the Original audio. (b) Histogram of the encrypted audio (c) Histogram of the original music sound (d) Histogram of the encrypted audio sound.

**FIGURE 4:** Correlation results (a) Correlation results of original image (b) Correlation results of the encrypted image.
We gauge the proposed audio encryption technique in terms of NPCR and UACI and the outcomes are shown in Table 3. Table 3 predicts that the proposed technique tends to negate differential attacks.

### F. SIGNAL TO NOISE RATIO (SNR)

In [58-59], the signal quality is measured with the help of the Signal to Noise Ratio (SNR). The signal will be more than noise once the value is bigger than 0 dB. The SNR can easily be calculated provided that host and encrypted audio files are available. The formula for the SNR is given as:

$$SNR = 10 \cdot \log_{10} \left( \frac{\sum_{j=1}^{N_s} u_j^2}{\sum_{j=1}^{N_s} (u_j - v_j)^2} \right)$$

(29)

where $N_s$ represents the number of samples. Moreover, $u_j$ and $v_j$ are the trials of the host and encrypted audio samples. Table 4 depicts the SNR outcomes of the tested audio files. The negative value of the SNR indicates the strength of the scheme. Table 4 indicates that our scheme has improved negative SNR and hence has a better resistance against malicious attacks.

### G. PEAK SIGNAL TO NOISE RATIO (PSNR)

In order to calculate the mean squared error for two vectors namely $U$ and $V$ can be computed by using the formula:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (U[i] - V[i])$$

(30)

If $U$ shows the host audio file and $V$ represents its coded audio file, the PSNR can be obtained as follows:

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right)$$

(31)

where $MAX$ is the maximum value of the stream. The PSNR values for the encrypted tested audio files are computed and listed in Table 4. It is noted that the values are small. Lower values of PSNR is desired for encrypted audio files as it refers to the high level of noise in the encrypted audio files and so strong resistance against attacks.
H. ROOT MEAN SQUARE (RMS) & CREST FACTOR (CF) VALUE

For an audio signal, the average amplitude value is calculated by using the Root mean square (RMS). The RMS is alike standard deviation provided that the input signal mean value equal to zero and it is calculated as follows:

\[ \text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |A_i|^2} \]  

(32)

The ratio of the peak values to the effective value is named as Crest factor (CF) and this is the waveform parameter. Its main purpose is to find the extremeness of the peaks in a waveform minimum possible value. CF of ratio 0 dB indicates no peaks as DC. Higher CF means peaks.

It is given as follows:

\[ CF = 20 \log_{10} \frac{|\text{Peak}|}{\text{RMS}} \]  

(33)

For the proposed algorithm, Table 5 depicts the RMS and CF values for the tested values. From the table, it is clear that values of RMS and CF of the coded audio files are nearly equal to 0.61 and 4.4 respectively. This demonstrates that there is no statistical connection between the host audio files and the corresponding coded audio files.

Table 3: SNR and PSNR Analysis

| No. | Audio                      | RMS          | CF           |
|-----|----------------------------|--------------|--------------|
| 1   | Audio                      | 0.6985       | 4.331        |
| 2   | Animal sound.wav           | 0.6873       | 4.543        |
| 3   | Alarm sound.wav            | 0.7132       | 5.543        |
| 4   | Applause sound.            | 0.6732       | 4.987        |
| 5   | Bells sound.               | 0.6451       | 4.675        |
| 6   | Birds sound.wav            | 0.6832       | 4.790        |
| 7   | Female sound.wav           | 0.7654       | 4.087        |
| 8   | 44100 Hz tone.wav          | 0.6456       | 4.054        |
| 9   | Male sound.wav             | 0.6953       | 4.786        |
| 10  | Machine sound.wav          | 0.7358       | 4.342        |
| 11  | Music sound.wav            | 0.6423       | 4.761        |

V. CONCLUSION

In this manuscript, we offered a lossless audio encryption technique that depends on the arithmetic operation of the elliptic curve and Galois field. Initially, we introduced a novel random number generator scheme, which is used to generate quality random numbers and passed all the NIST tests successfully. The generated random sequence is then used to shuffle the original audio data set. In the confusion phase of the idea, a new S-box construction scheme is deployed, which generates multiple S-boxes without much computational effort. The S-boxes are then used to substitute the shuffled audio. The substitution with multiple S-boxes produced optimum confusion in the encrypted and make capable the scheme robust against differential attacks. The scheme was thoroughly securitized over various simulation analyses. The results of the simulation experiment evidenced that the proposed scheme is secure against various cryptanalysis methods. Accordingly, the proposed scheme is secured and suitable for audio encryption applications.

Conflict of interest: The authors have no conflict of interest

REFERENCE

[1]. PUB, FIPS. "Data Encryption Standard (DES)." FIPS PUB (1999): 46-3.
[2]. Basu, Sandipan. "International data encryption algorithm (IDEA)—a typical illustration." Journal of global research in Computer Science 2.7 (2011): 116-118.
[3]. Barker, Elaine, and Nicky Mouha. Recommendation for the triple data encryption algorithm (TDEA) block cipher. No. NIST Special Publication (SP) 800-67 Rev. 2 (Draft). National Institute of Standards and Technology, 2017.
[4]. Daemen, Joan, and Vincent Rijmen. "Reijndael: The Advanced Encryption Standard." Dr. Dobb's Journal: Software Tools for the Professional Programmer 26.3 (2001): 137-139.
[5]. Naseer, Yasir, Dawood Shah, and Tariq Shah. "A novel approach to improve multimedia security utilizing 3D mixed chaotic map." Microprocessors and Microsystems 65 (2019): 1-6.
[6]. Alghafis, Abdullah, Hafiz Muhammad Waseem, and Majid Khan. "A hybrid cryptosystem for digital contents
confidentiality based on rotation of quantum spin states." _Physica A: Statistical Mechanics and its Applications_ (2019): 123908.

[7]. Naseer, Yasir, Tariq Shah, and Dawood Shah. "A novel hybrid permutation substitution base colored image encryption scheme for multimedia data." _Journal of Information Security and Applications_ 59 (2021): 102829.

[8]. Arshad, Usman, et al. "An efficient image privacy scheme based on nonlinear chaotic system and linear canonical transformation." _Physica A: Statistical Mechanics and its Applications_ (2019).

[9]. Shah, Dawood, Tariq Shah, and Sajjad Shaukat Jamal. "A novel efficient image encryption algorithm based on affine transformation combine with linear fractional transformation." _Multidimensional Systems and Signal Processing_ (2019): 1-21.

[10]. Tanveer, Muhammad, et al. "An Efficient Image Privacy-Preserving Scheme Based On Mixed Chaotic Map and Compression." _International Journal of Image and Graphics_ (2021): 2250020.

[11]. Khan, Majid, and Hafiz Muhammad Waseem. "A novel image encryption scheme based on quantum dynamical spinning and rotations." _PloS one_ 13.11 (2018).

[12]. Shah, Dawood, and Tariq Shah. "A novel discrete image encryption algorithm based on finite algebraic structures." _Multimedia Tools and Applications_ 79.37 (2020): 28023-28042.

[13]. Waseem, Hafiz Muhammad, and Majid Khan. "Information confidentiality using quantum spinning, rotation and finite state machine." _International Journal of Theoretical Physics_ 57.11 (2018): 3584-3594.

[14]. Waseem, Hafiz Muhammad, Majid Khan, and Tariq Shah. "Image privacy scheme using quantum spinning and rotation." _Journal of Electronic Imaging_ 27.6 (2018): 063022.

[15]. Haider, Muhammad Imran, et al. "Block cipher’s nonlinear component design by elliptic curves: an image encryption application." _Multimedia Tools and Applications_ 80.3 (2021): 4693-4718.

[16]. Khalid, Ijaz, et al. "A Novel Scheme of Image Encryption Based on Elliptic Curves Isomorphism and Substitution Boxes." _IEEE Access_ (2021).

[17]. Hua, Zhongyun, et al. "Color image encryption using orthogonal Latin squares and a new 2D chaotic system." _Nonlinear Dynamics_ (2021): 1-18.

[18]. Hua, Zhongyun, et al. "Visually secure image encryption using adaptive-thresholding sparsification and parallel compressive sensing." _Signal Processing_ 183 (2021): 107998.

[19]. Servetti, Antonio, and Juan Carlos De Martin. "Perception-based partial encryption of compressed speech." _IEEE Transactions on Speech and Audio Processing_ 10.8 (2002): 637-643.

[20]. Thorwirth, N. J., et al. "Security methods for MP3 music delivery." _Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No. 00CH37154)_ . Vol. 2. IEEE, 2000.

[21]. Servetti, Antonio, Cristian Testa, and Juan Carlos De Martin. "Frequency-selective partial encryption of compressed audio." _2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)_ . Vol. 5. IEEE, 2003.

[22]. Bhargava, Bharat, Changgui Shi, and Sheng-Yih Wang. "MPEG video encryption algorithms." _Multimedia Tools and Applications_ 24.1 (2004): 57-79.

[23]. Grangetto, Marco, Enrico Magli, and Gabriella Olmo. "Multimedia selective encryption by means of randomized arithmetic coding." _IEEE Transactions on Multimedia_ 8.5 (2006): 905-917.

[24]. Yan, Wei-Qi, Wei-Gang Fu, and Mohan S. Kankanhalli. "Progressive audio scrambling in compressed domain." _IEEE Transactions on Multimedia_ 10.6 (2008): 960-968.

[25]. Zhou, Jianguo, and Oscar C. Au. "Security and efficiency analysis of progressive audio scrambling in compressed domain." _2010 IEEE International Conference on Acoustics, Speech and Signal Processing_. IEEE, 2010.

[26]. Lima, Juliano B., and Eronides F. da Silva Neto. "Audio encryption based on the cosine number transform." _Multimedia Tools and Applications_ 75.14 (2016): 8403-8418.

[27]. Kordov, Krasimir. "A Novel Audio Encryption Algorithm with Permutation-Substitution Architecture." _Electronics_ 8.5 (2019): 530.

[28]. Sathiyanamurthi, P., and S. Ramakrishnan. "Speech encryption using chaotic shift keying for secured speech communication." _EURASIP Journal on Audio, Speech, and Music Processing_ 2017.1 (2017): 20.

[29]. Lima, Juliano B., and Eronides F. da Silva Neto. "Audio encryption based on the cosine number transform." _Multimedia Tools and Applications_ 75.14 (2016): 8403-8418.