IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 8, 2021, accepted September 26, 2021, date of publication October 13, 2021, date of current version October 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3119515

An Efficient Audio Encryption Scheme Based on Finite Fields

DAWOOD SHAH[1], TARIQ SHAH[1], MOHAMMAD MAZYAD HAZZAZI[2], MUHAMMAD IMRAN HAIDER[1,3], AMER ALJAEDI[4], AND IQTADAR HUSSAIN[5]

[1]Department of Mathematics, Quaid-i-Azam University Islamabad, Islamabad 15320, Pakistan
[2]Department of Mathematics, College of Science, King Khalid University, Abha 61413, Saudi Arabia
[3]Department of Mathematics, Gomal University, Dera Ismail Khan 29220, Pakistan
[4]College of Computing and Information Technology, University of Tabuk, Tabuk 71491, Saudi Arabia
[5]Mathematics Program, Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha 2713, Qatar

Corresponding author: Dawood Shah (dawoodshah254@gmail.com)

This work was supported by the Deanship of Scientific Research at King Khalid University through the Research Groups Program under Grant RG.P. 2/150/42.

ABSTRACT 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

The associate editor coordinating the review of this manuscript and approving it for publication was Alessia Saggese.

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. Servetti and De Martin [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, Zhou and Au [25] showed that the Yan scheme is conquerable against key search attacks. In [26], Lima and Neto 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.

A. 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.

B. 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 filed $GF(2^n)$. 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 $\frac{F_q[w]}{<h(w)>}$ 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 $E : y^2 = x^3 + ax + b(mod\ p)$, where a,b $\in F_p$, satisfy the equation $\left(4a^3 + 27b^2\right) \neq 0(mod\ p)$. 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 $P_1 \boxplus P_2 = (u, v_3) = R$ as under.

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

$$(u_3, v_3) = (\lambda^2 - u - u_2 modp, \\ \lambda (u_1 - u) - v_1 modp) \qquad (1)$$

And

$$\lambda = \frac{v_2 - v_1}{u_2 - u_1} modp \qquad (2)$$

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

$$(u_3, v_3) = (\lambda^2 - u - u_2 \ modp, \lambda (u_1 - u) - v_1 modp) \qquad (3)$$

And

$$\lambda = \frac{3u_1^2 + a}{2v_1} mod\ p \qquad (4)$$

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

v. Furthermore, define

$$P + O = P \quad \text{for all } P \text{ on } E \qquad (5)$$

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

Then the homomorphism

$$\varphi_\gamma : E^{ns}(F_p) \rightarrow F_p^* \text{ is defined as}$$

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

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 begun 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 $E^p_a : y^2 = x^3 + ax \ mod\ p$ through brute force technique. Subsequently, use the following map to transmute points of the curve $E^p_a(u, v)$ into the field F_p.

$$\varphi_\gamma : E^p_a \longrightarrow F_p$$

Defined by

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

where $\gamma \in F_p^*$ is the squared element such that $\gamma^2 = a$ and (u, v) is the element of the curve E^p_a. The map φ_γ is the isomorphism between E^p_a and F_p by Theorem 1 The consequential set $\varphi_\gamma(E^p_a)$ 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_s. In this study, we fixed the elements $a = 2$ 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.

B. 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 $GF(2^n)$. The mapping is defined as follows.

$$\begin{aligned} S : GF(2^n) &\longrightarrow GF(2^n) \\ \dot{h} &\longmapsto \dot{v}\left(\left(\dot{x}(\dot{h}) + \dot{y}\right)^{-1}\right) + \dot{u} \end{aligned} \tag{8}$$

In equation (8) $\dot{x}$, $\dot{v}$, $\dot{y}$ and $\dot{u}$ 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.

$$I_k : GF(2^n) \longrightarrow GF(2^m)$$

Define as

$$I_1\left(\sum_{i=1}^{n} a_i x^i\right) = \begin{cases} \sum_{i=1}^{m} a_i x^i, & if\ i \leq m-1 \\ 0, & if\ i > m-1 \end{cases}$$

$$I_2\left(\sum_{i=1}^{n} a_i x^i\right) = \begin{cases} \left(\sum_{i=1}^{n} a_i x^i - \sum_{i=1}^{m-1} a_i x^i\right) x^{-m-1}, \\ \quad if\ m-1 < i \leq 2m-1 \\ 0, \\ \quad if\ i < m-1 ori > 2m-1 \end{cases}$$

$$I_k\left(\sum_{i=1}^{n} a_i x^i\right) = \begin{cases} \left(\sum_{i=1}^{n} a_i x^i - \sum_{i=1}^{(k-1)m-1} a_i x^i\right) x^{-m-1}, \\ \quad if\ m-1 < i \leq km-1 \\ 0, \\ \quad if\ i < m-1 ori > km-1 \end{cases} \tag{9}$$

where $k \geq 2$ and n is strictly greater than m. The composition map $I_i o S$ generates $m \times m$ S-box. With this process, one can generate $n - m$ number of S-boxes.

C. 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.

$$M_{i,j} = \begin{cases} -1, & if\ A_{i,j} < 0 \\ 1, & if\ A_{i,j} \geq 0 \end{cases} \tag{10}$$

where $A_{i,j}$ indicates is 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 σ 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.

$$\delta_i = \begin{cases} \sigma_i, & if\ \sigma_i \leq MN \\ 1, & if\ \sigma_i > MN \end{cases} \tag{11}$$

$$A_{i,j} = A_{\delta_i,\delta_j} \tag{12}$$

where δ_i, δ_j denote the position of the integer value A_{δ_i,δ_j} in the newly shuffled matrix A_s. 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_s by using the absolute function from the set in the range $\{-2^{15}, 2^{15-1}\}$ 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.

$$\psi : GF(2^{15}) \longrightarrow GF(2^8) \times GF(2^7)$$

Defined by

$$\psi\left(\sum_{i=0}^{14} a_i x^i\right) = \left(\sum_{i=0}^{7} a_i x^i, \sum_{i=8}^{14} a_{i-8} x^{i-8}\right) \tag{13}$$

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 × 8 S-boxes using the proposed S-box construction method, which we have discussed in subsection 3.2.2. Afterwar 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 × 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^2 and A_s^2.

Step 6. Combine the resultant matrices A_s^2 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.

$$\psi^{-1} : GF(2^8) \times GF(2^7) \longrightarrow GF(2^{15})$$

Defined by

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

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 ρ 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 $GF(2^{15})$.

$$A_{s2}(i, j) = (A_{s1}(i, j) + (\rho(i, j)^{-1}) \tag{15}$$

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 (15) 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.

$$A_E(i,j) = \begin{cases} A_{s2}(i,j), & if\ M(i,j) = 1 \\ -A_{s2}(i,j), & if\ M(i,j) = -1 \end{cases} \tag{16}$$

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.

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.

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

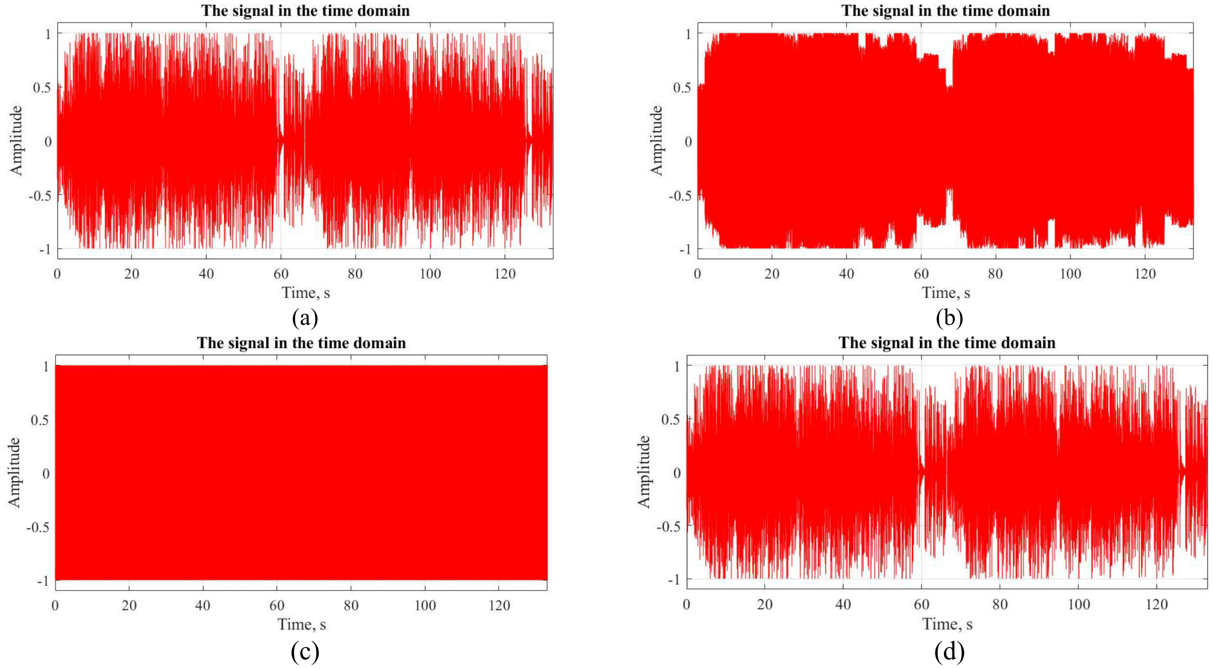
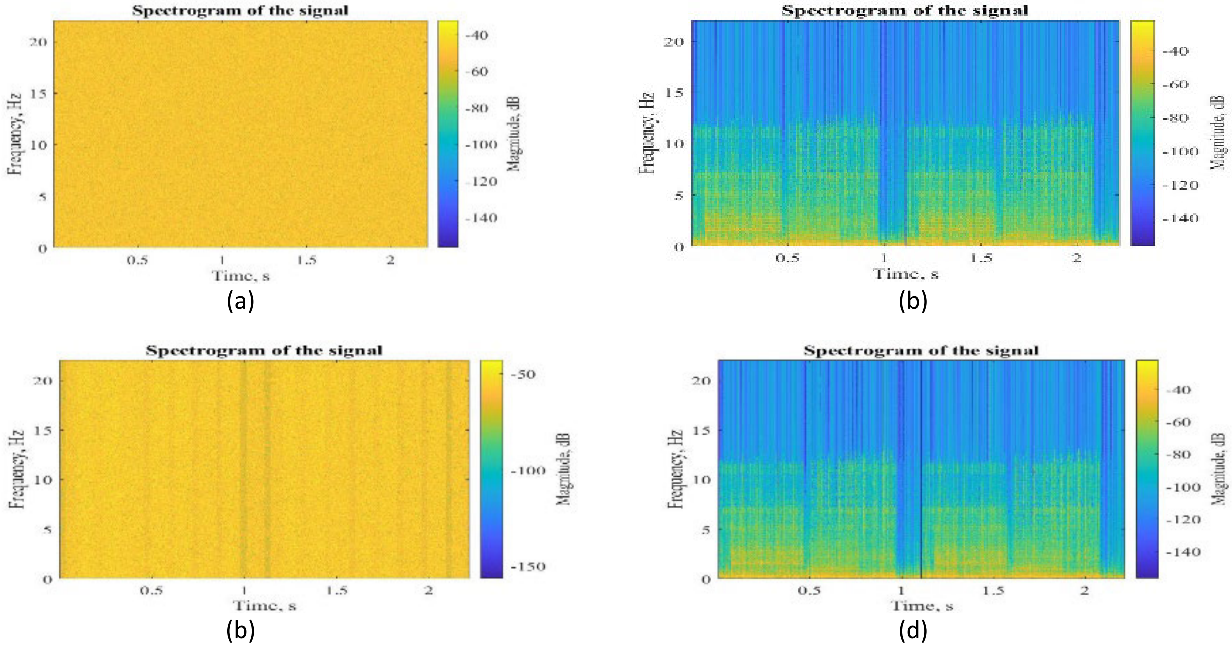


FIGURE 1. Waveform of the (a) original audio (b) Permuted audio. (3) Encrypted audio (d) Decrypted audio.



(a)

(b)

(b)

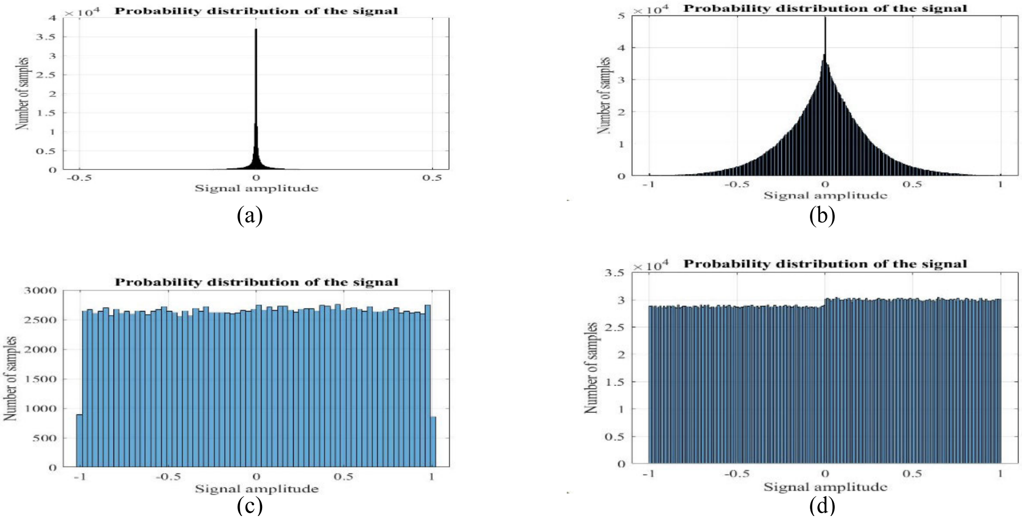
(d)

FIGURE 2. Spectrogram Graph of (a) Original audio (b) Permuted audio (c) Encrypted audio, (d) Decrypted audio.

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



(a)

(b)

(c)

(d)

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.

between identical sections of the data. The correlation coefficient is given by:

$$\gamma_{uv} = \frac{cov(p,q)}{\sqrt{\mathcal{D}(p)\mathcal{D}(q)}} \tag{17}$$

where

$$cov\,(p,q) = \frac{1}{\mathcal{P}}\sum_{i=1}^{\mathcal{P}} p_i - \mathcal{E}\,(p)\,(q_i - \mathcal{E}\,(q)) \tag{18}$$

$$\mathcal{D}\,(p) = \frac{1}{\mathcal{P}}\sum_{i=}^{\mathcal{P}} (p_i - \mathcal{E}(p))^2 \tag{19}$$

And

$$\mathcal{E}\,(p) = \frac{1}{\mathcal{P}}\sum_{i=1}^{\mathcal{P}} p_i \tag{20}$$

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 inter-correlation 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

$$H = -\sum_{k=0}^{\mathcal{L}} \mathcal{P}\,(k)\,log_2\mathcal{P}\,(k) \tag{21}$$

where $\mathcal{L}$ directs the grayscale value of the audio file and $\mathcal{P}(k)$ 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 proposed technique is almost equal to 16 for all ciphered audio and hence formed ideal vagueness in the audio file. So, our scheme has the ability to resist entropy attacks.

E. DIFFERENTIAL ANALYSIS

For differential attacks mostly, we consider two analyses i.e., the number of pixel change rates (NPCR) and Unified Average Changing Intensity (UACI). They calculate the sensitivity regarding the cryptosystem. A quality cryptographic algorithm must have sensitivity so a minor alteration in the original data produces a massive variation in the cipher data. Both NPCR and UACI analysis tend to assess the sensitivity of the cryptosystem. NPCR and UACI can be given as.

$$NPCR = \frac{\sum_{u,v}\mathcal{B}\,(u,v)}{K}\times 100 \tag{22}$$

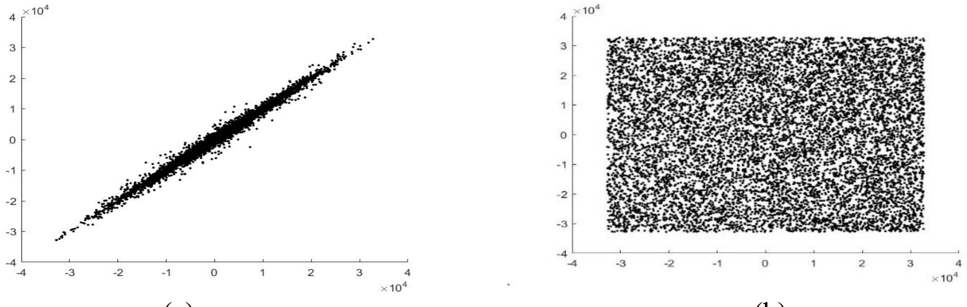
In the above equation, K represent the cardinality of the audio data set and $\mathcal{B}\,(u,v)$ is given by

$$\mathcal{B}\,(u,v) = \begin{cases} 1, & \text{if } \mathcal{A}_1\,(u,v) = \mathcal{A}_2\,(u,v) \\ 0, & \text{if } \mathcal{A}_1\,(u,v) \neq \mathcal{A}_2\,(u,v) \end{cases} \tag{23}$$

UACI can be represented as

$$UACI = \frac{1}{K}\sum_{u,v}\frac{|\mathcal{A}_1\,(u,v) - \mathcal{A}_2\,(u,v)|}{2^K - 1}\times 100 \tag{24}$$

where 2^K designates the order of bits in the audio data set. The satisfactory values of NPCR and UACI rate of the algorithm



(a)

(b)

FIGURE 4. Correlation results (a) Correlation results of original image (b) Correlation results of the encrypted image.

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. Wav	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. [26]		0.0263

are nearly equal to 100 and 33.3333 respectively. We gauge the proposed audio encryption technique by using NPCR and UACI analysis and the outcomes are shown in Table 3. Table 3 predicts that the proposed technique tends to negate differential attacks.

F. SIGNAL TO NOISE RAT IO (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:

$$SNR = 10 * log_{10} \frac{\sum_{j=1}^{N_s} u_j^2}{\sum_{j=1}^{N_s} (u_j - v_j)^2} \tag{25}$$

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.

TABLE 2. Information entropy analysis.

No	Audio	Plain Audio	Ciphered Audio
1	Audio	8.0065	15.4316
2	Animal sound.wav	9.8183	15.5592
3	Alarm sound.wav	13.4401	15.8693
4	Applause sound.	13.4216	15.9388
5	Bells sound. Wav	4.5625	12.2128
6	Birds sound.wav	8.5125	14.9905
7	Female sound.wav	9.8134	15.6663
8	44100 Hz tone.wav	10.6914	15.6024
9	Male sound.wav	14.1688	15.9271
10	Machine sound.wav	14.8475	15.9888
11	Music sound.wav	14.8549	15.8779

TABLE 3. Differential cryptanalysis.

No	Audio	Plain Audio	Ciphered Audio
1	Audio	99.99724	33.1233
2	Animal sound.wav	99.99974	33.456
3	Alarm sound.wav	99.99951	33.2203
4	Applause sound.	99.94041	33.1202
5	Bells sound. Wav	99.9884	33.1344
6	Birds sound.wav	99.99794	33.9205
7	Female sound.wav	99.9940	31.6479
8	44100 Hz tone.wav	99.99728	33.74039
9	Male sound.wav	99.9972	33.0987
10	Machine sound.wav	99.9996	33.67.8
11	Music sound.wav	99.99724	33.1233

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:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (U[i] - V[i]) \tag{26}$$

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

$$PSNR = 10 * log_{10} \left(\frac{MAX^2}{MSE} \right) \tag{27}$$

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:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |A_i|^2} \tag{28}$$

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:

$$CF = 20 log_{10} \frac{|VPeak|}{V_{RMS}} \tag{29}$$

For the proposed algorithm, Table 4 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

TABLE 4. 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. Wav	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

TABLE 5. NIST randomness test.

No	Audio	P-Value
1	Frequency Test (T)	0.9253077508893466
2	Frequency Test	0.347578425321557
3	Run T	0.45321310856174435
4	Longest Run T	0.43428142438827533
5	Binary Rank T	0.7454887332471692
6	Discrete Fourier T	0.12497609962873209
7	Non-Overlapping	0.622298646456104
8	Overlapping	0.1716767122905817
9	Maurer's Universal	0.9253077508893466
10	Linear Complexity	0.347578425321557

there is no statistical connection between the host audio files and the corresponding coded audio files

I. NIST STATISTICAL TEST

For cryptographic applications, we studied the sequence created by the proposed random number generator to assess the random number generator. To examine the randomness of this generated sequence, we first change the random sequence into binary as the NIST test is valid for binary data. The NIST statistical test involves sixteen different tests as presented in Table 5. The generated sequence conceded all the randomness tests, which shows that our proposed technique engenders quality random sequences that have compatibility with different audio encryption applications.

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

REFERENCES

[1] *Data Encryption Standard (DES)*, Standard FIPS PUB 46-3, 1999.

[2] S. Basu, "International data encryption algorithm (IDEA)—A typical illustration," *J. Global Res. Comput. Sci.*, vol. 2, no. 7, pp. 116–118, 2011.

[3] E. Barker and N. Mouha, "Recommendation for the triple data encryption algorithm (TDEA) block cipher," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NIST SP 800-67, Revision 2, 2017.

[4] J. Daemen and V. Rijmen, "Reijndael: The advanced encryption standard," *Dr. Dobb's J., Softw. Tools Prof. Programmer*, vol. 26, no. 3, pp. 137–139, 2001.

[5] Y. Naseer, D. Shah, and T. Shah, "A novel approach to improve multimedia security utilizing 3D mixed chaotic map," *Microprocessors Microsyst.*, vol. 65, pp. 1–6, Mar. 2019.

[6] A. Alghafis, H. M. Waseem, M. Khan, and S. S. Jamal, "A hybrid cryptosystem for digital contents confidentiality based on rotation of quantum spin states," *Phys. A, Stat. Mech. Appl.*, vol. 554, Sep. 2020, Art. no. 123908.

[7] Y. Naseer, T. Shah, and D. Shah, "A novel hybrid permutation substitution base colored image encryption scheme for multimedia data," *J. Inf. Secur. Appl.*, vol. 59, Jun. 2021, Art. no. 102829.

[8] U. Arshad, M. Khan, S. Shaukat, M. Amin, and T. Shah, "An efficient image privacy scheme based on nonlinear chaotic system and linear canonical transformation," *Phys. A, Stat. Mech. Appl.*, vol. 546, May 2020, Art. no. 123458.

[9] D. Shah, T. Shah, and S. S. Jamal, "A novel efficient image encryption algorithm based on affine transformation combine with linear fractional transformation," *Multidimensional Syst. Signal Process.*, vol. 31, no. 3, pp. 885–905, Jul. 2020.

[10] M. Tanveer, T. Shah, A. Ali, and D. Shah, "An efficient image privacy-preserving scheme based on mixed chaotic map and compression," *Int. J. Image Graph.*, Jun. 2021, Art. no. 2250020.

[11] M. Khan and H. M. Waseem, "A novel image encryption scheme based on quantum dynamical spinning and rotations," *PLoS ONE*, vol. 13, no. 11, Nov. 2018, Art. no. e0206460.

[12] D. Shah and T. Shah, "A novel discrete image encryption algorithm based on finite algebraic structures," *Multimedia Tools Appl.*, vol. 79, nos. 37–38, pp. 28023–28042, Oct. 2020.

[13] H. M. Waseem and M. Khan, "Information confidentiality using quantum spinning, rotation and finite state machine," *Int. J. Theor. Phys.*, vol. 57, no. 11, pp. 3584–3594, Nov. 2018.

[14] H. M. Waseem, M. Khan, and T. Shah, "Image privacy scheme using quantum spinning and rotation," *Proc. SPIE*, vol. 27, no. 6, 2018, Art. no. 063022.

[15] M. I. Haider, A. Ali, D. Shah, and T. Shah, "Block cipher's nonlinear component design by elliptic curves: An image encryption application," *Multimedia Tools Appl.*, vol. 80, no. 3, pp. 4693–4718, Jan. 2021.

[16] I. Khalid, S. S. Jamal, T. Shah, D. Shah, and M. M. Hazzazi, "A novel scheme of image encryption based on elliptic curves isomorphism and substitution boxes," *IEEE Access*, vol. 9, pp. 77798–77810, 2021.

[17] Z. Hua, Z. Zhu, Y. Chen, and Y. Li, "Color image encryption using orthogonal Latin squares and a new 2D chaotic system," *Nonlinear Dyn.*, vol. 104, pp. 4505–4522, May 2021.

[18] Z. Hua, K. Zhang, Y. Li, and Y. Zhou, "Visually secure image encryption using adaptive-thresholding sparsification and parallel compressive sensing," *Signal Process.*, vol. 183, Jun. 2021, Art. no. 107998.

[19] A. Servetti and J. C. De Martin, "Perception-based partial encryption of compressed speech," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 8, pp. 637–643, Nov. 2002.

[20] N. J. Thorwirth, P. Horvatic, R. Weis, and J. Zhao, ''Security methods for MP3 music delivery,'' in *Proc. Conf. Rec. 34th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, 2000, pp. 1831–1835.

[21] A. Servetti, C. Testa, and J. C. De Martin, ''Frequency-selective partial encryption of compressed audio,'' in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Apr. 2003, p. 668.

[22] B. Bhargava, C. Shi, and S.-Y. Wang, ''MPEG video encryption algorithms,'' *Multimedia Tools Appl.*, vol. 24, no. 1, pp. 57–79, 2004.

[23] M. Grangetto, E. Magli, and G. Olmo, ''Multimedia selective encryption by means of randomized arithmetic coding,'' *IEEE Trans. Multimedia*, vol. 8, no. 5, pp. 905–917, Oct. 2006.

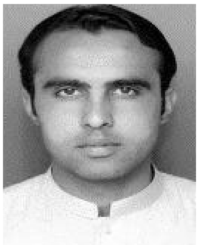
[24] W.-Q. Yan, W.-G. Fu, and M. S. Kankanhalli, ''Progressive audio scrambling in compressed domain,'' *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 960–968, Oct. 2008.

[25] J. Zhou and O. C. Au, ''Security and efficiency analysis of progressive audio scrambling in compressed domain,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2010, pp. 1802–1805.

[26] J. B. Lima and E. F. da Silva Neto, ''Audio encryption based on the cosine number transform,'' *Multimedia Tools Appl.*, vol. 75, no. 14, pp. 8403–8418, Jul. 2016.

[27] K. Kordov, ''A novel audio encryption algorithm with permutation-substitution architecture,'' *Electronics*, vol. 8, no. 5, p. 530, May 2019.

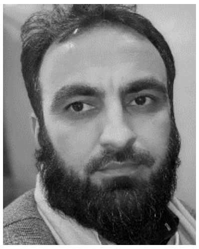
[28] P. Sathiyamurthi and S. Ramakrishnan, ''Speech encryption using chaotic shift keying for secured speech communication,'' *EURASIP J. Audio, Speech, Music Process.*, vol. 2017, no. 1, p. 20, Dec. 2017.



DAWOOD SHAH received the M.Phil. degree from the Department of Mathematics, Quaid-i-Azam University, where he is currently pursuing the Ph.D. degree. His research interests include coding theory, finite fields, and cryptography.

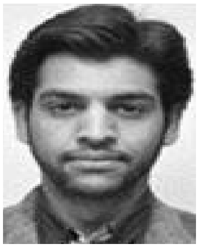
TARIQ SHAH is currently working as a Professor and the Head of the Mathematical Cryptography Group, Quaid-i-Azam University Islamabad, Pakistan. He has introduced number of courses at post graduate and graduate level in different institutions. He is also the founder of mathematical cryptography and designs different structures for the construction of nonlinear component of block ciphers and cryptosystems.

MOHAMMAD MAZYAD HAZZAZI received the Ph.D. degree in mathematics from the University of Sussex, Brighton, U.K. He is currently working as an Assistant Professor with the Department of Mathematics, King Khalid University, Abha, Saudi Arabia. His research interests include coding theory, cryptography, finite geometry, algebraic geometry, and group theory.



MUHAMMAD IMRAN HAIDER is currently pursuing the Ph.D. degree with the Department of Mathematics, Quaid-i-Azam University, Islamabad, Pakistan. He is also working as an Assistant Professor with the Department of Mathematics, Gomal University, Dera Ismail Khan.

AMER ALJAEDI received the B.Sc. degree from King Saud University, Saudi Arabia, in 2007, the M.Sc. degree in information systems security from the Concordia University of Edmonton, Canada, in 2011, and the Ph.D. degree in security engineering from the Department of Computer Science, University of Colorado at Colorado Springs, Colorado Springs, USA, in 2018. He is currently an Assistant Professor with the College of Computing and Information Technology, University of Tabuk. Before that, he was a Senior Research Member with the Cybersecurity Laboratory, Colorado University. His research interests include software-defined networking, network traffic control and monitoring, cloud computing, and cybersecurity. He received multiple research awards from UCCS and SACM for his outstanding research articles.



IQTADAR HUSSAIN received the Ph.D. degree in mathematics specializing in the area of algebraic cryptography, in 2014. He is currently an Assistant Professor with Qatar University. His current research interests include the applications of mathematical concepts in the field of secure communication and cybersecurity, where he has published 63 articles in well-known journals. His H-index score is 23 and I-10 index score is 34. His articles have 1320 Google Scholar citation.