A Review on Audio Encryption Algorithms Using Chaos Maps-Based Techniques

Ekhlas Abbas Albahrani¹, Tayseer Karam Alshekly² and Sadeq H. Lafta³,*

¹Department of Computer Science, Mustansiriyah University, Baghdad, Iraq
²Department of Banking and Finance science, Al-Imam Al-A’tham University College, Baghdad, Iraq
³Department of Applied Science, University of Technology – Iraq
E-mail: akhlas_abas@uomustansiriyah.edu.iq; tayseer.karam@yahoo.com; sadeq.h.lafta@uotechnology.edu.iq
*Corresponding Author

Received 14 January 2021; Accepted 06 September 2021; Publication 22 October 2021

Abstract

Due to the quick improvement in digital communications and multimedia applications during recent periods up to the current time, data protection of digital data such as image, audio and video becomes a significant challenge. The security of audio data that transfer through different networks was rated as a preferred research field in the preceding years. This review covers the recent contribution for audio encryption and gives the most evaluations for audio encryption algorithm involving security analysis, computational complexity and quality analysis and their requirements. This paper fundamentally concentrates on displaying the different types of audio encryption and decryption techniques based on chaotic maps. Digital and analog audio algorithms were displayed, discussed and compared with the illustration of the important features and drawbacks. Various digital and audio proposed projects for audio encryption using chaotic maps have been covered, which they showed extreme sensitivity to initial conditions, unpredictability and conducting in a quasi-random manner. A comparison among the proposed

Journal of Cyber Security and Mobility, Vol. 11, 53–82.
doi: 10.13052/jcsm2245-1439.1113
© 2021 River Publishers
algorithms in the key space, chaotic maps sensitivity and statistical analysis were provided.

**Keywords:** Encryption algorithms, audio encryption, audio decryption, chaotic maps.

1 Introduction

Many tasks like wire and wireless communication, researching, teaching, financial procedures, etc... need a helpful and assistant source to perform them easily and quickly, such as in the internet. Security is considered as an important element in the audio communication, voice-over internet protocols, secret voice seminars, and business sections. Audio encryption is a way to immune information in an audio file from parasitical attacks by applying a key (noise) and precise algorithm to the plain text. The security system has to be very secure, fast, and durable to ensure data confidentiality, data high solidity, and data trusted. Continuing in this context, researchers evolved several cryptographic algorithms to be dependent in the evolution of wireless communication methods. Standard symmetric encryption programs like the data encryption standard (DES) and the advanced encryption standard (AES) can achieve a very good level of guaranty. In spite of most of these algorithms are utilized for most data like a binary data, they are not to be perfect for real-time audio encryption due to the following reasons [1–6]:

First: Audio data are estimated to be typically huge and bulky, so if the traditional encryption systems is used to encrypt such bulky data, it acquires critical overhead, excessively and costly real-time multimedia applications and require continuous tasks, for example, cutting, duplicating, bit-rate control or recompression.

Second: The standard symmetric cryptographic schemes have a small key space, which makes them suffer from an assault of a brute force in addition to the large level of redundancy between the specimen, the amplification of the ciphered signal by bandwidth and the decrease of the signal to noise ratio of the output.

Third: These algorithms demand a longer computational period and more computing power, due to the complex permutation process.

Fourth: For many real-life audio applications, light encryption becomes too important in order to save some perceptual data. This is too difficult to be
accomplished by conventional ciphers alone, which in all probability corrupts
the information to produce a perceptually unrecognizable content.

On the other hand, the encryption by asymmetric algorithms is not perfect
for owing to their low processing speed and complexity.

Recently, a great significance was given to the use of chaotic theory
to perform the encryption for audio files [7–10]. Chaos theory is certainly
supposed as the dynamic part of present cryptography, where its methods
have a major point of attention lies on:

First: the arbitrary behavior of chaotic maps can deceive unauthorized people
without a need for a special mechanism for creating it.

Second: the chaotic map evolution time, depending on control parameters
and initial conditions and slight varieties in these amounts, yields a very
extraordinary time evolution. This means one can apply these control param-
eters and initial conditions as keys in cipher system. In addition, the low
cost of chaotic signal makes it appropriate to be used in audio encryption
algorithms [11, 12]. Because of all these chaos theory advantages, now, audio
encryption algorithms based on chaotic maps have been a more interest and
progress.

Literature reviews did not cover purely the work that using chaos method
for audio encryption algorithms. This may be due to the late time for this work
to be done and published. The first and most related work was done in 2012,
where a survey chapter [4] focusing on different chaos-based encryption
algorithms for image, video and audio was published. In 2015, another
review [13] focused on a various audio steganographic & cryptographic
methods. It contained different algorithms for audio encryption or hiding.
The authors explain the symmetric and asymmetric algorithms that have been
used for audio encryption or audio steganography. In 2016, [14] authors
provided a survey study on different audio encryption methods and their
enhancements to achieve different cryptographic security parameters. They
confirmed that the audio data need for faster, less complex and secured
encryption algorithms. In 2018, a review [15] primarily centered on encryp-
tion algorithms for audio data. It exhibited an examination and correlation
of essential encoding norms. Audio Encryption can be done using certain
encryption algorithms like AES, DES and Triple DES (3-Data Encryption
Standard).

This review is focusing in detail on the aspect that were not addressed for
previous chaotic audio algorithms of all types involving their force and weak
points. Analog and digital signal processing were given as an important base for encryption process. All the requirements of these algorithms and security and statistical tests for evaluation such algorithms were displayed. Because of the above reasons and the quick improvement in communications and multimedia applications, it is obvious that the research question is justified to review the chaotic audio encryption.

2 Audio Encryption Requirements and Evaluation Methods

Audio data have special characteristics like great data sizes, high excess, intelligent activities and requirements of continuous reactions. Sometimes the fields of use of audio data have their private necessities like security, constancy of the compression ratio of the audio file, design consistency, transmission mistake resistance and to request of ongoing. Special requirements of audio encryption and their evaluation methods are summarized in the next topics.

2.1 Security Analysis

Audio encryption demands at first acceptable level of security as it is assumed that the chaos guarantees the security of the audio data. This means that the audio encryption should be secure to resist different attacks. Thus, if the encryption algorithm cannot break in an hour, then it might be viewed as a secure algorithm in this application [4]. Encryption security usually includes key space, key sensitivity, perceptual security and its resistance to potential attacks.

1. Key space: the key is the crucial part of each cryptosystem. In general, key space could be gotten by resolving the number of secret keys to the given encryption algorithm. Let $k_i$ indicates a key and $K$ is a limited set of probable keys, then the key space is denoted by $K = \{k_1, k_2, \ldots, k_r\}$, where $r$ is the number of possible keys. An example of this, a 16-bit key will own a key space of $2^{16}$.

2. Perceptual security: during using an algorithm for encryption of an audio data, if one is unable to recognize the encrypted audio, the encryption algorithm is considered as confident in terms of perception.

3. Key sensitivity: a good audio encryption algorithm must be sensitive to the secret key i.e. modifying of a one bit in the key must create a totally unlike ciphered outcome. This sensitivity is known as key sensitivity. As a rule, a key sensitivity in chaotic cipher related to the initial values...
and control variables sensitivities of the chaotic map that was chosen as audio encryption algorithm. The key sensitivity is usually measured by the key sensitivity analysis where the typical test is based on the next steps [11]:

**Step 1.** Firstly, encrypting the original audio file using the chosen secret key. For example, let the key is “Key1=1234567899876543” and the as encrypted resulting audio is called audio X.

**Step 2.** After that, the audio file itself is encrypted by making the little alteration in the secret key i.e. “Key2=2234567899876543” that modify only minimal significant bit of Key1. The outcome audio is called the as encrypted audio Y.

**Step 3.** At last, two as encrypted audios X and Y that encrypted by Key1 and Key2 respectively are now being compared by finding the correlation coefficient \( r_{xy} \) them [11]:

\[
\begin{align*}
 r_{xy} &= \frac{\text{cov}(X, Y)}{\sqrt{D(X) \cdot D(Y)}} \\
\text{cov}(x, y) &= \frac{1}{N} \sum_{i=1}^{N} (X_i - E(X))(Y_i - E(Y)) \\
D(x) &= \frac{1}{N} \sum_{i=1}^{N} (X_i - E(X))^2 \\
E(x) &= \frac{1}{N} \sum_{i=1}^{N} X_i
\end{align*}
\]

Knowing that X and Y are the audio vectors, \( \text{cov}(x, y) \) is covariance between the two files, \( D(x) \) is defined as in Equation (3) below and \( D(y) \) are defined in the same way, N is the total number of the samples, E(x) and E(y) are discrete forms defined as in Equation (4):

If the resulted correlation value is close to zero, this means that the audio encryption algorithm is very sensitive to the secret key.

4. Potential attacks: the good cipher should avoid a following number of common attacks:

- **Brute-force attack:** can be described as the inquiry to break an encryption by attempting every possible key. A lot of time, a cipher
is viewed as secure in the event that it must be broken by brute force. A regular brute force attack includes an exhaustive search for the key, in identical circumstances when a thief experiences every conceivable blend in the lock of safe [16]. Brute-force attack is generally tested by finding the size of the key space. The size of key space should be large enough, to make brute-force attack infeasible.

- **Known-plaintext attack**: if the attacker can catch the ciphertext and its related part of the plaintext, the key can be discovered. Known-plaintext attack is generally analyzed by comparing the premier data and the decrypted one. To make Known-plaintext attack inefficient, the encryption algorithm must be so complex to the extent that the key cannot be discovered even if the ciphertext and an associated piece of plaintext are known.

- **Differential attack**: a good cipher program must have the required feature, which propagates the effect of a single plaintext bits over as much as possible of the cipher text, so as to cover up the statistical texture of the plaintext. This implies that if one makes so little change in the original audio, this can bring about a huge change in the cipher-audio, in turn, the differential attack really loses its effectiveness and becomes practically pointless. The differential attack is based on the test of the contrast between two plaintexts. Three common measures that examine the effect of a little changing in the original audio data is called Mean Absolute Error (MAE), Unified Average Changing Intensity (UACI) and Number of Pixels Change Rate (NPCR).

1. **Mean Square Error (MSE)**: it measures the avalanche effect where the total squared lapse is discovered between two encrypted audios. Let X and Y are two vectors that contain the two audio streams, then MSE between these two audios can be computed as [13,16]:

\[
MSE = \frac{1}{N} \sum_{i=0}^{N-1} [X(i) - Y(i)]^2
\]  

(5)

Where N is vector length X (or Y).

MSE value is usually given in dB by taking 10log for the product of Equation (5). If the MSE value (in dB) is =30 dB, consequently the quality difference between two the audios is obvious and algorithm has good immunity the differential attack.
2. Number of Pixels Change Rate (NPCR): It denotes to the rate of change that occurs in the encrypted audio data when the original audio data changes slightly. It is gauged by the distinctive data stream between two audios. If one supposes the two encrypted audios vectors to be X and Y whose original audios have only slight changes. Also, let a bipolar array D is of the same size as X or Y, then D(i) is defined by X(i) and Y(i) such that:

In case of \( X(i) = Y(i) \) then \( D(i) = 0 \) Else \( D(i) = 1 \)

Thus, NPCR is defined by the following formulas:

\[
\text{NPCR} = \frac{\sum_{i,j} D(i)}{W} \times 100\%,
\]

where \( W \) is the length of X or Y [12].

3. Unified Average Changing Intensity (UACI): It determines the average intensity of the contrasts between the original audio and cipher audio. Let the two encrypted audios vectors be X and Y whose original audios have only slight changes. In this case, UACI is given by [16]:

\[
\text{UACI} = \frac{1}{N} \left[ \sum_{i,j} \frac{|X(i) - Y(i)|}{255} \right] \times 100\%,
\]

Audio encryption algorithm to be a highly secure, it should be safe in perception, excessive key sensitivity, have big key space, and withstand potential attacks.

### 2.2 Computational Complexity

Computational complexity can be defined as the investigation of the parameters, such as time and memory, which are required in order to implement the given computational tasks. The capacity of the audio data, compared to the text data, is too huge. Probably, the great computational complexity be occurred when the audio cryptographic algorithm encrypts all of the audio data bits and in similar importance. Due to the test by the human senses that has high strength to detect the image degradation or sound noise, only efficient encryption of these data and what bound to it can achieve multimedia protection efficiently with high clarity and reduced computational complexity.
The analysis of encryption time is determined by the next three ways:

1. Absolute encryption time: it indicates to the time that assumed to encrypt audio data on a specific operating algorithm that is measured by seconds.
2. Relative encryption time ratio: it indicates to the time ratio of encryption time to the compression time.
3. Computation complexity: relies upon the cost of the cipher based on a chaotic system and the audio data volume to be encrypted.

Considerably, audio encryption algorithms are preferred for real-time applications when its computational cost or absolute time is very small with respect to its compression.

2.3 Quality Analysis

The audio quality is computed by Peak Signal to Noise Ratio (PSNR) and Signal to Noise Ratio (SNR). PSNR test explains modifications in audio value quality of the original audio with respect to the encrypted audio as given in the Equation (8) [17]:

$$\text{PSNR} = 10 \times \log_{10} \left( \frac{M^2}{\sum_{i=0}^{N} (X(i) - Y(i))^2} \right)$$  \hspace{1cm} (8)$$

$X(i)$ represents the premier audio, $Y(i)$ denotes the ciphered audio while $M$ is the maximum audio data. Low value of PSNR means that the ciphered audio file has a high level of noise; as a result, it makes it more resistant to attacks.

SNR test is utilized to find the remaining clarity of the algorithm output and the quality of the restored audio. Usually, algorithm output (encrypted audio) is identified by a small SNR value and indicates to the higher noise compared to that of the algorithm input speech. The perfect quality of decrypted audio is distinguished by elevated SNR value. The value of SNR is measured using Equation (9) [1, 2]:

$$\text{SNR} = 10 \times \log_{10} \left( \frac{\sum_{i=0}^{N} P^2}{\sum_{i=0}^{N} (P(i) - D(i))^2} \right)$$ \hspace{1cm} (9)$$

Here, $P(i)$ is the premier audio and $D(i)$ is the decrypted audio.
3 Chaos-Based Audio Cipher Systems

The proposed audio signals encryption algorithms are mostly classified to digital audio encryption and analog audio encryption.

3.1 Digital Audio Encryption Algorithms

In digital encryption, the analog signal is firstly digitized to produce a data signal at a convenient bit rate. Certain algorithm then encrypts the bit stream. There are different algorithms have been suggested to perform encryption process based on chaos theory in the previous years.

R. Gnanajeyaraman et al. [18] in 2009 proposed a framework of cipher block chaining architecture for audio encryption depending on Chaos-based Look-Up tables that are produced by higher dimensional cat map, known as $N^{th}$D cat map. The $N^{th}$D cat map is shown in the Equation (10):

$$
\begin{bmatrix}
A_{n+1} \\
B_{n+1} \\
\vdots \\
H_{n+1}
\end{bmatrix}
= x
\begin{bmatrix}
A_n \\
B_n \\
\vdots \\
H_n
\end{bmatrix}
\mod 1
$$

(10)

Where:

$$
\begin{bmatrix}
1 & a_{12} & 0 & \ldots & 0 \\
b_{12} & 1 + a_{12}b_{12} & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\times
\begin{bmatrix}
1 & 0 & a_{13} & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
b_{13} & 0 & 1 + a_{13}b_{13} & \ldots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

(11)
Where \( a_{ij} \) and \( b_{ij} \) are integers in \([0,2^{L-1}]\) (\( L \) is the number of bits used by the key stream) and \( m \) is the number of tables. For chaotic behavior, they used \( L = 16 \) and \( m = 8 \). In this audio encryption algorithm, a random numbers are generated in the encryption block. The sequence of these numbers will select the tables for encryption. Based on the selected table, the digital audio value is mapped to the chaotic iteration number. In the encryption process, the previous cipher digit is added with \( n^{th} \) current plain digit that the resultant value will be changed with the value to look up table. The diagram of the Gnanajeyaraman et al. audio algorithm is illustrated in Figure 1.

This proposed algorithm is characterized by its resists to brute-force attack and also to chosen/known-plaintext attacks besides its sensitivity to the initial condition.

Juliano B. Lima et al. [19] in 2015, introduce a new scheme for audio ciphering depending on the cosine number transform (CNT). The definition of CNT requires a finite field cosine function that can be defined through:

- **Definition A**: Let \( p \) an odd prime and \( \zeta \) is a nonzero element in the finite field \( \text{GF}(p) \). The finite field cosine function belonged to \( \zeta \) is

\[
\cos_{\zeta}(x) := \frac{\zeta^x + \zeta^{-x}}{2}
\]

\( x = 0, 1, \ldots, \text{ord}(\zeta) \), \( \text{ord}(\zeta) \) denotes de multiplicative order of an element \( \zeta \) and \( \zeta \) is the minimum positive integer i.e. \( \zeta^l = 1 \mod p \).

- **Definition B**: Let \( \text{ord}(\zeta) = 2N \), then the cosine number transform of the vector \( \mathbf{x} = [x_0, x_1, \ldots, x_{N-1}] \), \( x_i \in \text{GF}(p) \), is the vector
\( X = [X_0, X_1, \ldots, X_{N-1}], \quad X_j \in \text{GF}(p) \) which is gotten by:

\[
X_j := \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} \beta_j x_i \cos \left( \frac{j}{2} \frac{2i+1}{2} \right) \tag{13}
\]

where

\[
\beta_j = \begin{cases} 
1 \sqrt{2} \text{(modp)}, & j = 0, \\
1, & j = 1, 2, \ldots, N - 1. 
\end{cases} \tag{14}
\]

Calculating CNT of a row vector \( x \) may be denoted by the matrix \( X = C \cdot x^T \) where \( x^T \) is the vector \( x \) transpose and \( C \) corresponds to the transform matrix where its element in the \((j+1)\)-th row and the \((i+1)\)-th column is given by:

\[
C_{i+1,j+1} = \sqrt{\frac{2}{N}} \beta_j \cos \left( \frac{j}{2} \frac{2i+1}{2} \right) = \sqrt{\frac{2}{N}} \beta_j \cos \sqrt{\pi}(2i+1),
\tag{15}
\]

\( i, j = 0, 1, \ldots, N - 1. \) It is clear to that the inverse CNT is given by the transformation matrix \( C^{-1} = C^T \).

In the proposed algorithm, a CNT matrix \( C \) is generated and audio data are broken into blocks. Each block with 8 samples is denoted by \( b_n \) where each sample is an integer number involved within the period \([0 – 65535]\). The processed audio block index \( n \) specifies the chosen element of the secret-key utilized in the algorithm. The diffusion operation in this algorithm is performed by taking the audio block in which it overlaps the former encrypted audio block in two specimens. This means that the first two specimens of the premier audio block \( b_n \) are the final two samples of the encrypted audio block \( b_{n-1} \) (just the block \( b_1 \) is left with no overlapping). The power of the CNT matrix \( C \) is determined by the computation of the \( k_n \text{(mod } K) \). Then the matrix \( C^{k_n \text{mod } K} \) is multiplied by the block \( b_n \). This will produce the proviso cipher audio block. That means the computation of CNT of \( b_n \) is iterated by \( k_n \text{(mod } K) \) times. The proposed algorithm steps are illustrated in Figure 2. The decryption process is a reverse of the encryption where the matrix \( C \) is changed by the matrix \( C^{-1} = C^T \) and the blocks are processed from right to left.

Depending on the set of block cipher and chaotic maps, Ekhlas Abbas Albahrani [17] in 2017 proposed an audio encryption scheme. The suggested
algorithm encrypts 625(25 × 25) bytes block size and decrypts it. Every block is inserted to three stages: permutation stage, adding XOR stage and substitution stage. The audio block is permuted in the permutation stage using a chaotic 2D tent map. The tent map is a simple 2D chaotic system, which can be represented as in Equation (16) [20].

\[
    x_{i+1} = \begin{cases} 
        \frac{x_i}{\alpha} & \text{if } x \in [0, \alpha] \\
        (1 - x_i) \frac{1}{1 - \alpha} & \text{if } x \in (\alpha, 1]
    \end{cases}
\]  

Equation (16)

Here, \(\alpha\) is the control variable chosen to be in the period \([0,1]\], and \(x_i\) is the current. It is important to mention here that Tent map has a regular invariable probability density in the range 0–1.

Then the resulting block and the key block are processed by XOR operation. Lastly, the resulting block is changed by a new substitution process that depends on reverse multiplication. An algorithm based on Chebyshev’s polynomial generates the key [18]. These polynomials are described as in Equation (17) [21]:

\[
    x_{i+1} = T_n(x_i) = \cos(n \chi \arccos(x_i)) \\
    y_{i+1} = T_m(y_i) = \cos(m \chi \arccos(y_i))
\]  

Equation (17)
Where \((x_i, y_i) \in [-1, 1]\) and \((n, m)\) are the control variables in the period \([2, \infty]\). The initial values \(x_o\) and \(y_o\) and the parameters \(n\) and \(m\) were used as the key.

The diagram of the Albahrani audio encryption is shown in Figure 3. The proposed algorithm has regular histograms, large key space, small PSNR, small correlation, large entropy and a good resistance to brute-force attack.

Also in 2017, S. J. Sheela et al. [22] suggested a new audio encryption scheme using 2D adjusted Henon map (2D-MHM) and standard. 2D-MHM is utilized to produce the random sequences. The 2D-MHM [23] is described as following:

\[
H(x_k, y_k) = \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} 1 - b_1 \cos(x_k) - b_2 y_k \\ -x_k \end{pmatrix}
\]  

(18)

Where \(b_1\) and \(b_2\) are control parameters and \(x_k\) and \(y_k\) represent the 2D state of the system. 2D Standard map is defined as in Equation (19) [24]:

\[
\begin{cases} 
    x_{i+1} = (x_i + r_x + y_i + r_y) \mod N \\
    y_{i+1} = (y_i + r_y + k \sin \frac{x_i + 1}{2\pi}) \mod N
\end{cases}
\]  

(19)

Where \(K\) is a positive integer and the two parameters \(r_x\) and \(r_y\) are located in the range \([0 \ldots N - 1]\). The Standard map has the largest key space, which works as a good candidate for block permutation.

The confusion property can be achieved by shuffling the position of audio samples based on the generated random sequences. While the diffusion property achieved by changing the value of the audio samples based on
encoding of the DNA and the series created from the standard map. The diagram of S. J. Sheela et al. audio encryption algorithm is shown in Figure 4. This algorithm, as well, has large key space, which is sufficient to counter a brute force attack and it is very sensitive to the secret key so it can resist exhaustive attack.

Krasimir Kordov [25] in 2019 used a classic symmetric models to propose a new audio encryption algorithm with a new pseudo-random generator scheme which is utilized as the basis for efficient encryption of chaotic bit-level permutations and substitutions. The algorithm divide the digital audio data into samples and process them by shifting the sample bits to perform permutation and substituting for the bits values in the sample. The proposed pseudo-random number generator bases on two chaotic maps, chaotic circle map and altered rotation equations. The Circle map is 1D dynamical system usually utilized in cipher due to its chaotic manner. The iterations of the standard circle map are computed by [26]:

$$\theta_{n+1} = \left( \theta_n + \omega - \frac{k}{2\pi} \sin(2\pi \theta_n) \right) \mod 1$$  \hspace{1cm} (20)

Where \( \omega \) represent a constant as an angle in polar prompt of sinusoidal oscillator.
An altered format of rotation equation, which is presented in some literature [22, 23, 27, 28], is applied very little in encryption systems. The used formula is given by:

$$x_{t+1} = -a - (x_t - a) \cos \theta + y_t \sin \theta / r_t$$  \hspace{1cm} (21)

$$y_{t+1} = -x_t r_t \sin \theta - y_t \cos \theta$$  \hspace{1cm} (22)

$$r_t = \sqrt{0.5 (x_t^2 + \sqrt{x_t^4 + 4y_t^2})}$$  \hspace{1cm} (23)

Where $x_t$ and $y_t$ are the initial values, $\theta$ and $a$ are the parameters that generate chaotic manner at $\theta = 2$ and $a = 2.8$, and $r_t$ is the polar resultant of $x_t$ and $y_t$. Figure 5 shows the diagram of the Kordov audio encryption algorithm. A combination of two chaotic maps was used in the proposed algorithm in order to expand the key-space and increases the cryptographic security.

Lastly, R. I. Abdelfatah [29] proposes a new audio encryption method contains three stages with a secret keys for each stage, these stage are:

- **Scrambling stage**: the binary audio bit sequences are scrambling using the SHA512 hash function with input audio as an initial key.
- **Dynamic DNA encoding stage**: the audio specimens are changed and distributed into the cipher audio. Dynamic DNA encoding stage is better than firmed DNA encoding because it performs security development by eight encoding basis as a candidate for only one base. It consists of four various nucleic acid bases: adenine (A), thymine (T), cytosine (C), guanine (G), where C and G, T and A are pairs complementary. The secret
key for this stage is generated using the new proposed pseudo random generator that links three chaotic maps: Chebyshev Equation (17), Sine Equation (24), and Logistic map Equation (25). Sine map is 1D map that can be defined by Equation (24):

\[ x_{n+1} = \mu \sin(\pi x_n) \]  

(24)

Here, \( x \) and \( \mu \in [0, 1] \), the map is chaotic when \( \mu \in [0.87, 1] \).

Logistic map is considered as the simplest chaotic functions that used recently in cipher applications; it is expressed as [30]:

\[ x_{n+1} = r.x_n.(1 - x_n) \]  

(25)

Here, \( x_n \) falls in (0, 1), \( r \) is a constant and falls in the range (0-4). The value of \( r \) specifies the behavior of the logistic map. At \( r = 3.57 \), the iterations be completely chaotic and they are obeyed to the aim of encryption. Therefore, the high value of the variable \( r \), the highly chaotic deterministic discrete-time signal.

• Third stage: this stage used two dynamic DNA algebraic processes, “AND” and “XOR”. The secret key that used in this stage is generated using a pseudo random generator that links three chaotic maps: 1D Henon map Equation (26), 1D Logistic map Equation (25) and 1D Gaussian map Equation (27). 1D Henon map defined as:

\[ x_{n+1} = 1 - u(x_{n+1})^2 + \beta x_n \]  

(26)

Where \( u \) and \( \beta \) are the chaotic parameters and when \( a = 1.4 \) and \( b = 0.3 \), the map is chaotic.

Gaussian map is 1D chaotic map as:

\[ x_{n+1} = \text{Exp}[-r \times (x_n)^2] \]  

(27)

The block diagram of the R. I. Abdelfatah cipher algorithm for audio files is shown in Figure 6.

3.2 Analog Audio Encryption Algorithms

Analog audio cipher has four main types, they are:

• Frequency domain encryption,
• Time domain encryption,
• 2D encryption, which combines frequency domain encryption, time domain encryption, and amplitude encryption.
Since 2014 researchers tried to present different algorithms for encryption analog audio data. In 2014, Saad Najim et al. [31] presented multilevel speech encryption algorithm based totally on confusion and diffusion properties. They used three keys (key1, key2 and mask key) for permutation and substitution operations. These keys are generated using the chaotic logistic map Equation (25).

The speech samples are permuted in two level:

- In the first level, all the speech samples are permuted using key1. The resulted samples are entered into DWT or DCT (they are transforming technique that converts an audio sign from the time domain into the transform domain) for getting the transform coefficients components. The key2 are used to permute the resulted transform coefficients.
- In second level, the speech samples in the time domain are permuted again using the Arnold cat map.

Figure 6 The block diagram of the Rouyat Ismail Abdelfatah audio encryption algorithm [29].
The resulted samples from the first level are masked using the mask key to yield the substitution property. Arnold cat map is a two-dimensional chaotic map defined by Equation (27) [32].

\[
\begin{bmatrix}
    x_{n+1} \\
    y_{n+1}
\end{bmatrix} = \begin{bmatrix} 1 & a \\ b & ab + 1 \end{bmatrix} \cdot \begin{bmatrix} x_n \\
    y_n
\end{bmatrix} \mod N
\] (27)

In this equation, \(x_n\) and \(y_n\) represent the positions of specimens in the \(N \times N\) matrix, \(n = 1\) to \(N - 1\), \(x_{n+1}, y_{n+1}\) denote the new changed position after cat map, and \(a\) and \(b\) are the control variables that have integers values larger than zero.

Figure 7 shows the diagram of Saad Najim et al. audio encryption algorithm. The latter algorithm is again marked by a large key space and multilevel permutation -substitution operation to increase the speech signal security.

**Figure 7** The diagram of Saad Najim et al. audio encryption algorithm.
E. M. Elshamy et al. [33], in 2015, proposed a new audio encryption algorithm utilizing double security stages. In the initial stage, a chaotic method either Baker map or cat map was used to perform the prime level of security. Arnold cat map is a 2D dimensional chaotic map defined through Equation (16).

The modification of the original Arnold transformations that given in Equation (16) is shown in Equations (28) and (29) to produce Arnold transformations sequence.

\[
\begin{pmatrix}
  x' \\
  y'
\end{pmatrix} =
\begin{pmatrix}
  i & i + 1 \\
  1 & 1
\end{pmatrix}
\begin{pmatrix}
  x \\
  y
\end{pmatrix}
\pmod{N}
\tag{28}
\]

Or

\[
\begin{pmatrix}
  x' \\
  y'
\end{pmatrix} =
\begin{pmatrix}
  i + 1 & i \\
  1 & 1
\end{pmatrix}
\begin{pmatrix}
  x \\
  y
\end{pmatrix}
\pmod{N}
\tag{29}
\]

Where \( i \in \{1, 2, 3, \ldots\} \).

The Baker map has two ways: a general Baker map and estimated Baker map, which is an effective way to randomize the elements in a square matrix. Let the estimated map denoted as \( B(n_1, \ldots, n_k) \). The vector \([n_1, \ldots, n_k]\) here is the secret key \( S_{key} \). Let the number of data elements in a row is denoted as \( N \), the chosen secret key is each integer \( n_i \) divides by \( N \), and \( n_1 + \cdots + n_k = N \). The Baker map Equation (30) moves the data piece at the indicators \((q, z)\) to the indicators:

\[
B_{(n_1, \ldots, n_k)}(q, z) = \left( \frac{N}{n_i}(q - N_i) + z \bmod \left( \frac{N}{n_i} \right), \frac{n_i}{N} \left( z - z \bmod \left( \frac{N}{n_i} \right) \right) + N_i \right)
\tag{30}
\]

In the second stage, optical encryption is used based on double random phase encoding (DRPE). The 2nd stage of security which denotes a physical security is so difficult to be attacked. The suggested algorithm divides the native audio data into blocks and reforms them into a 2D array. These blocks are masked with the Arnold cat map or Baker map and then subtract 2 from every value in these blocks to make all values between \((-1)\) and \((1)\). Lastly DRPE is applied. The flowchart of Elsayed M. Elshamy et al audio algorithm is shown in Figure 8.

F.J. Farsanaet et al. [34], in 2019, offered an audio encryption algorithm that using discrete modified Henon map that given in Equation (14) for audio
samples permutation followed by xoring the audio samples with the generated key stream for substitution operation. The keystream of the proposed algorithm is generated based on the modified Lorenz-Hyperchaotic system. Lorenz-hyperchaotic system display chaotic manner for the control variables values at $a = 10$, $b = 8/3$ and $c = 28$. Mathematically, Lorenz-Hyperchaotic system can be modelled as in Equation (32) [34]:

\[
\begin{align*}
    x' &= a(y - x) + w \\
    y' &= cx - y - xz \\
    z' &= xy - bz \\
    w' &= k_i a(y - x) + k_p x
\end{align*}
\]

where the state variables are $x, y, z, w$ and $a, b, c, k_p, k_i$ are the system parameters.
In this algorithm, first, the audio data is compressed by Fast Walsh Hadamard Transform (FWHT) to remove the maintain clarity in the transform domain. After that, the produced file is ciphered through two stages. In the first stage, permutation operation is performed by 2D-MHM to decrease the correlation between neighboring samples. In the next stage, the modified-Lorenz hyperchaotic substitute the samples so that the silent periods in the speech conversation are filling. The diagram of F.J. Farsanaet et al. audio encryption algorithm is shown in Figure 9. The performance analysis of the presented algorithm indicates that it is secure and able to withstand against the brute force attack and differential attack.

A new speech signal scrambling algorithm with low residual intelligibility was proposed by P. Sathiyamurthi et al. [35] in 2020. They constructed a new chaotic map with higher dimensions called 3D Lorenz-Logistic map, which is built by spreading the logical map in 3D Lorenz Map. The 3D Lorenz-Logistic map is modelled as follows:

\[
\begin{align*}
x_i &= a[dy_{i-1}(1 - y_{i-1})] - [dx_{i-1}(1 - x_{i-1})] \\
y_i &= b[dx_i(1 - x_{i-1})] - [dy_{i-1}(1 - y_{i-1})] - [dx_{i-1}(1 - x_{i-1})][dz_{i-1}(1 - z_{i-1})] \\
z_i &= [dx_{i-1}(1 - x_{i-1})][dy_{i-1}(1 - y_{i-1})] - c[dz_{i-1}(1 - z_{i-1})] \\
\end{align*}
\]

Where the initial conditions are \(x_{i-1}, y_{i-1}\) and \(z_{i-1}\) and \(a = 8\) to \(10\), \(b = 15\) to \(30\), \(c = 1\) to \(3\) and \(d = 0\) to \(4\). In this audio algorithm, the sampled sequence of the speech signal is converted from complex values to real and imaginary values by applying FFT. In permutation operation, the \(y\) dimension of 3D Lorenz-Logistic map is applied to change the order of the
real values of speech samples while the imaginary values are permuted using the z dimension of 3D Lorenz-Logistic map. In substitution operation, the reference speech sample provides information about the speech signal such as the speaker’s name, the title of the speech or lecture Organization and etc. It is firstly permuted using x dimension of 3D Lorenz-Logistic map and then substituted for the permuted real values by adding the permuted reference speech and permuted real values. The diagram of P. Sathiyamurthi et al. audio encryption algorithm is shown in Figure 10. Statistical and security analyses show the proposed algorithm is secure and can be used in real time application.

4 Performance Comparison of Audio Encryption Algorithms

Here, the aim is to compare between the audio encryption methods that were previously displayed. In the beginning, it is preferred to make a general
A Review on Audio Encryption Algorithms  

Table 1  Comparison between digital and analog audio encryption algorithms

<table>
<thead>
<tr>
<th></th>
<th>Digital Audio Encryption</th>
<th>Analog Audio Encryption</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The system encrypts audio data with lower residual intelligibility.</td>
<td>The system encrypts audio data without any residual intelligibility.</td>
</tr>
<tr>
<td></td>
<td>It has higher cryptanalytic strength</td>
<td>It has lower cryptanalytic strength</td>
</tr>
<tr>
<td></td>
<td>It has high security because they greatly change the redundancy of the audio.</td>
<td>It has poor security because they do not greatly change the audio redundancy.</td>
</tr>
<tr>
<td></td>
<td>The system requires audio compression or modems</td>
<td>No demand for audio compression technique or modems program in the system.</td>
</tr>
<tr>
<td></td>
<td>The system is sensitive to synchronization errors.</td>
<td>The system is less sensitive to synchronization errors.</td>
</tr>
<tr>
<td></td>
<td>At present, it is considered as the main algorithm for audio encryption</td>
<td>The algorithm is mostly used in analogue telephone, satellite and mobile communications without using a modem.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Chaotic Maps</th>
<th>Key Space</th>
<th>Key Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gnanajeyaraman</td>
<td>N\textsuperscript{th} D cat map</td>
<td>(S = (k_1 * k_2 * k_3 * \ldots * k_8)^n)</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>et al. [18]</td>
<td></td>
<td>(S) increases with the rise of the parameters (k_1 * k_2 * k_3 * \ldots * k_8) or iteration time n.</td>
<td></td>
</tr>
<tr>
<td>Juliano B. Lima</td>
<td>2D cosine number transform (CNT)</td>
<td>(2^{319})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>et al. [19]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elshamy et al. [17]</td>
<td>2D Tent map and Two Chebyshev polynomials</td>
<td>((6.28)^3 + 2^{36})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>S. J. Sheela</td>
<td>The modified Henon map and 2D standard map</td>
<td>(2^{149})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>et al. [22]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krasimir</td>
<td>circle map and A Modified rotation equation</td>
<td>(2^{928})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>et al. [25]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. I.</td>
<td>Chebyshev, Sine map, Logistic map, Henon map and 1D</td>
<td>(2^{348})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>Abdelfatoh [29]</td>
<td>Gaussian map.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saad Najim</td>
<td>Logistic map</td>
<td>(2^{548.11})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>et al. [31]</td>
<td>2D Cat map</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elsayed M.</td>
<td>2D Cat map and 2D Baker map</td>
<td>(S_{key} = \text{secret key vector for 2D Baker + the key of 2D Cat map + two keys for DRPE algorithm})</td>
<td>High key sensitivity</td>
</tr>
<tr>
<td>Elshamy et al. [33]</td>
<td>Lorenz-hyperchaotic system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F.J. Farsanaet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>et al. [34]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. Sathiyamurthi</td>
<td>3D Lorenz – Logistic map</td>
<td>(2^{66})</td>
<td>Low key sensitivity</td>
</tr>
<tr>
<td>et al. [35]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

comparison between the analog and digital audio encryption algorithms. This comparison is illustrated in Table 1. Tables 2 and 3 compares the work of the references \[17–19, 22, 25, 29, 31, 33–35\]. The comparison among them is based on the chosen chaotic maps, when the effects on (key space and the key sensitivity) is shown in Table 2. The statistical and the security analysis
including (correlation coefficient analysis, MSE, PSNR and SNR) are shown in Table 3.

From Table 2, one can conclude that the degree of security determines whether one choose analog or digital method of encryption. Of course, digital audio encryption is more secure, but analog doesn’t need compression and modem which shorten encryption time and that what is required in online encryption methods.

Depending on the fact that the algorithm should have a big key space to resist the brute-force attacks, key space size smaller than \(2^{128}\) is not adequately safe [36]. The results of Table 2 demonstrate two things. First, chaotic maps (2D Tent map, 2D CNT, 2D Chebyshev polynomial 2D Henon map) give clearly better results than others. Second, it is clear that all suggested algorithms withstand against all types of brute-force attacks because the key space is larger than \(2^{128}\) and have high key sensitivity.

The results of Table 3 affirm clearly that all correlation analyses of the proposed algorithms fulfills zero co-correlation and displaying that the intruder is unable to know any valuable data by investing the statistical attack. On the other hand, the PSNR and SNR results are low for all algorithms that imply that the proposed algorithms have a great noise producing them more resistant to attacks.

### 5 Conclusion and Future Work

Cipher chaos method is an important, required and efficient encryption method for audio files security. Encryption of audio data is characterized
by huge size, which affects the chosen proposed algorithm method. Each
cryptographic algorithms of the tens uses various-size of chaotic maps to
perform the task. The complexity of the encryption chaos method affects
the speed depending on audio size. Chaotic algorithms have different level of
security and so they suitable for different applications. The encryption perfor-
mane of those algorithms clarifies that each algorithm has its advantages and
disadvantages. Digital algorithms have lower residual intelligibility, higher
cryptanalytic strength and high security compared to analog algorithms. On
the other hand they require audio compression or modems and they are
sensitive to synchronization errors. All digital and analogs algorithms are
resistive to various attacks, which can be demonstrated by reviewing attacks
results due to the chaos features. At present, chaotic encryption systems are
considered as the main algorithm for audio encryption. The following future
works are preferred to be done to overcome the flaws that exist in digital
and analogue algorithms and to expand the uses of these algorithms in most
applications: Light weight audio algorithms for encryption voice over internet
protocol (VOIB) using the analog algorithms due its higher speed, adding the
compression algorithms for the digital algorithms to be synchronized with
the fast advance in internet applications, and using multi dimensions hyper
chaotic algorithms in by digital algorithms for real time audio applications.

References

encryption of speech signals,” Int. J. Speech Technol., vol. 14, no. 4,

Encryption: Zaslavsky Map as Pseudo Random Number Generator,”

chaos-based cryptosystems,” Int. J. Bifurc. Chaos, vol. 16, no. 8,

survey of chaos based encryption technology”, in book “Multimedia – A
multiclinary Approach to Complex Issues”, edited by Loannis Karydis,
INTECH, March 2012.

electronic health care information using obfuscation and RSA algorithm


A Review on Audio Encryption Algorithms


[27] An improvement on the chaotic behavior of the Gauss Map for cryptography purposes using the Circle Map combination. To cite this article: MT Suryadi et al. 2020 J. Phys.: Conf. Ser. 1490 012045. https://doi.org/10.1088/1742-6596/1490/1/012045


Biographies

Ekhlas Abbas Albahrani Assistant Prof. was born at 1974. She had the PhD Degree in Computer science/Data Security in 2016. The Msc. in Computer science/Data Security in 2001, and Bachelor Degree in Computer science 1996 at the University of Technology/Iraq. Work teams are at University of Technology and Mustansiriyah University, Baghdad, Iraq on Cryptography research field. Academic member from 1996 to 2001 at the University of Technology/Iraq, from 2001 to 2009 at Sebha University and at Mustansiriyah University/Iraq from 2009 to 2021. She participated in different committees such as organizing symposium and conferences in computer Sciences at Mustansiriyah University and IEEE conferences. She participated in different seminars, Training courses and lectures inside Iraq. She reviewed more than 20 articles in data security for Clarivate and Scopus journals. She participated in different organizing and scientific committees of conferences in College of Education at Mustansiriyah University/Iraq. She has about 20 published research papers in Cryptography and Data Security.

Tayseer Karam Alshekly, born in 1991 in Baghdad, B.Sc. 2014 in computer science. MSc. In Computer Science from College of Education at
Mustansiriyah University/Baghdad. She has four published papers in the field of cryptography. Now, she works for Banking and Finance science, Al-Imam Al-A’tham University College in Baghdad- Iraq.

Sadeq H. Lafta Assistant Prof. was born at 1972 He had his PhD Degree in Magnetic Material Science in 2016, the Msc. in Laser Physics1998, and Bachelor Degree in Applied Physics 1995 University of Technology/Iraq. His work teams are at University of Technology in semiconductor, nanomaterials and sensor fields, Duisburg-Essen University in nano-magnetic material field and Mustansiriyah University in Cryptography research field. He was an academic member from 1998 to 2009 in Sebha University, researcher at Nanotechnology Centre and Applied Science Department/University of Technology/Iraq from 2010 to 2021. He participated in different committees such as organizing symposium and conferences in Nanotechnology Centre/University of Technology. He participated in revealing the validity of laboratory equipment and suitability for work. He participated in different seminars, Training courses and lectures inside and outside Iraq. He has different social and scientific articles in Applied Science Dep. Website. He reviewed more than 50 articles in different applied physics for Clarivate and Scopus journals. He participated in different organizing and scientific committees of conferences in Nanotechnology Centre Applied Science Dep. At University of Technology/Iraq. He has about 20 published research papers in Applied Physics and Cryptography. He has a Science Day Medal/Ministry of Higher Education/Iraq in 2016 and Inventors Medal/Ministry of Science and Technology/Iraq in 2018.