The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2210-8327.htm

An audio encryption scheme based on Fast Walsh Hadamard Transform and mixed chaotic keystreams

F.J. Farsana and V.R. Devi
Department of Electronics and Communication, LBS Centre for Science and Technology, Kerala University, Trivandrum, India, and
K. Gopakumar
Department of Electronics and Communication, TKM College of Engineering, Kollam, India

Received 5 August 2019
Revised 4 October 2019
Accepted 7 October 2019

Abstract

This paper introduces an audio encryption algorithm based on permutation of audio samples using discrete modified Henon map followed by substitution operation with keystream generated from the modified Lorenz-Hyperchaotic system. In this work, the audio file is initially compressed by Fast Walsh Hadamard Transform (FWHT) for removing the residual intelligibility in the transform domain. The resulting file is then encrypted in two phases. In the first phase permutation operation is carried out using modified discrete Henon map to weaken the correlation between adjacent samples. In the second phase it utilizes modified-Lorenz hyperchaotic system for substitution operation to fill the silent periods within the speech conversation. Dynamic keystream generation mechanism is also introduced to enhance the correlation between plaintext and encrypted text. Various quality metrics analysis such as correlation, signal to noise ratio (SNR), differential attacks, spectral entropy, histogram analysis, keyspace and key sensitivity are carried out to evaluate the quality of the proposed algorithm. The simulation results and numerical analyses demonstrate that the proposed algorithm has excellent security performance and robust against various cryptographic attacks.

Keywords Audio encryption, Fast Walsh Hadamard Transform, Modified Henon map, Modified Lorenz hyperchaotic system

Paper type Original Article

Author contribution: Farsana F J, Devi V R and Gopakumar K have equally contributed in proposing the ideas, discussing the results, and writing and proofreading the manuscript. The authors have implemented the algorithms and carried out the experiments.

Declaration of Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Publishers note: The publisher wishes to inform readers that the article "An audio encryption scheme based on Fast Walsh Hadamard Transform and mixed chaotic keystream" was originally published by the previous publisher of *Applied Computing and Informatics* and the pagination of this article has been subsequently changed. There has been no change to the content of the article. This change was necessary for the journal to transition from the previous publisher to the new one. The publisher sincerely apologises for any inconvenience caused. To access and cite this article, please use Farsana, F.J., Devi, V.R., Gopakumar, K. (2019), "An audio encryption scheme based on Fast Walsh Hadamard Transform and mixed chaotic keystream", *Applied Computing and Informatics*. Vol. ahead-of-print No. ahead-of-print https://10.1016/j.aci.2019.10.001. The original publication date for this paper was 11/10/2019.

Applied Computing and Informatics
Vol. 19 No. 3/4, 2023
pp. 239-264
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1016/j.aci.2019.10.001

1. Introduction

Voice based communication becomes prominent in several areas such as military, phone banking, confidential voice conferencing, education etc. With the increasing need for secure speech communication, data encryption protocols are critically important for storage and transmission of sensitive information over exposed systems. Unlike text and message signals, adjacent samples of voice signals are highly correlated and slowly time-varying. Moreover, the presence of redundant and unvoiced samples in audio signal demands the need for efficient compression techniques in the transform domain. Therefore, the conventional cryptographic algorithms are poorly suited for speech encryption. With the advancement of security level, chaotic-systems play a significant role in developing cryptographic algorithms. Claude Shannon introduces two basic elements for secure cryptographic algorithms [1]. These elements are confusion (permutation) and diffusion (substitution) operations. In confusion operation, data samples are permuted according to some specific key parameters to destroy the local correlation between adjacent samples. While in diffusion operation, data samples are substituted with pseudo random numbers (PRNs) generated by some entropy sources, to change the sample values. Both these operations eventually strengthen the complex relationship among plaintext, ciphertext and symmetric key parameters. While developing symmetric key encryption algorithm, designers utilize substitution- permutation network as the basic structural element [2]. Chaos theory plays a significant role in developing encryption algorithms, due to its inherent properties such as topological transitivity, ergodicity, sensitive dependence on initial conditions and deterministic pseudo randomness. These eventually satisfy the basic requirement for theoretical cryptography. Most of the chaos based algorithms such as multiple iterations of chaotic map [3,4], bit-level scrambling approaches [5,6] bit-level confusion methods [7] and hybrid key methods [8] were designed based on the above mentioned properties. Also, single and multiple round of permutation-then-diffusion without substitution-box (PT-DWOS), DNA encoding methods are introduced based on the reproducibility and deterministic nature of chaotic functions, since the process can be repeated for the same function and same initial conditions. To improve the complexity and randomness of encryption scheme various chaotic systems are introduced e.g. discrete time [9], continuous time [10], hyper chaotic [11] and time delayed systems [12]. Cascading of different chaotic systems [13], iteratively expanding lower dimensional chaotic map to higher dimension [14], parametric perturbations of chaotic trajectories [15] are the common methodologies followed by the designers to develop algorithm based on chaos theory. Furthermore, chaos theory have been incorporated in many conventional cryptographic approaches like S-box design [16], RC5 stream cipher [17] and elliptical curve cryptography [18] to strengthen the security of the encryption processes. But these methods are flawed with limited keyspace and computational complexity. Recently, researchers have attempted to develop computationally efficient and unconditionally secure chaotic-quantum algorithms suitable for cloud and Internet of Things (IoT) environments [19–22].

Several audio encryption algorithms have been introduced to provide secure data transmission. Among these techniques voice encryption algorithm based on chaos theory are considered to be effective to handle with redundant and bulky audio files. An overview of speech encryption algorithm based on chaos theory is discussed here after. Long Jye Sheue et al. [23] proposed a speech encryption algorithm based on fractional order chaotic systems. It is based on two-channel transmission method where the original speech signal is encoded using a nonlinear function of the Lorenz chaotic system. Moreover, they analyzed the conditions for synchronization between fractional chaotic systems theoretically by using the Laplace transform. Mosa et al. [24] introduced an algorithm based on permutation and substitution of speech segments using chaotic Baker map. They used Discrete Cosine Transform (DCT) to remove the residual intelligibility in order to compress the signal.

Maysaa abd ulkareem et al. [25] proposed a method based on logistic map and blowfish encryption algorithm. They employed partial encryption method by wavelet packet transform for splitting the raw signal to improve the speed of encryption process. Halto et al. [26] presented a hybrid chaotic speech encryption algorithm in which Arnold cat map is utilized for permutation and logistic map for substitution operation. They used Discrete Cosine Transform (DCT) for compressing the audio signal to minimize residual intelligibility. In [27] Halto et al. presented a method where the Lorenz system generates the keystream for substitution operation and Rossler chaotic system for permutation process. Elashy at al. [28] proposed a two level audio encryption algorithm based on chaotic Baker maps and double phase random coding. In the first level it utilizes Baker map and in the second level it utilizes optical encryption using double random phase encoding (DRPE) for providing physical security which is hard to break. Sathya murthi et al. [29] introduced an algorithm based on chaotic shift keying. In [29], audio signals are sampled and its values are segmented into four levels, and then the samples are permuted using chaotic systems such as Logistic map, Tent map, Quadratic map, and Bernoulli's map. Sheela et al. [30] proposed an algorithm based on two-dimensional modified Henon map (2D-MHM) and standard map. They introduced Hybrid Chaotic Shift Transform (HCST), and deoxyribonucleic acid (DNA) encoding rules to enhance the security level. Aissa Belmeguenai et al. proposed a method, where only relevant part of the speech segment undergoes encryption process [31]. Animesh Roy et al. [32] presented an algorithm based on audio signal encryption using chaotic Henon map and lifting wavelet transform. Z. Habib et al. presented a paper based on amplitude scrambling and Discrete Cosine Transform coefficient scrambling. In [33], they designed a permutation network using TD-ERCS chaotic map. Some of the constraints observed in the above mentioned literature are summarized as follows:

1. Inefficient compression method to remove the unvoiced data segments.
2. Suggested Permutation methods are not strong enough to break the correlation between adjacent samples in the audio file.
3. Lower Dimensional chaotic maps have periodic window problems like smaller chaotic range and non-uniform distribution.
4. In substitution-permutation rounds, permutation matrix and keystream generated from the chaotic function depends only on the initial condition and control parameters of the chaotic map.

To overcome the above mentioned drawbacks, we propose an audio encryption algorithm based on chaotic maps based on modified Henon map and modified Lorenz-hyperchaotic system. To overcome the first drawback, the algorithm employs a signal compression mechanism by Fast Walsh Hadamard Transform (FWHT), which reduces the sample size for further encryption process by removing higher order coefficients. Unlike Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT), FWHT has excellent energy compression properties. Furthermore, rectangular basis functions of FWHT can be realized more efficiently in digital circuits rather than the trigonometric basis functions of the Fast Fourier Transform. The second problem can be eliminated by modified Henon map, which weakens the strong correlation between adjacent coefficients. The samples are shuffled through the strong permutation matrix generated with the modified Henon map. The output data is then diffused by XOR-ring the permuted coefficients with keystream generated by the modified Lorenz-hyperchaotic system. The encryption process in higher dimensional space eliminates periodic window problems such as limited chaotic range and non-uniform distribution. It also extends the keyspace and consequently enhances the security. To eliminate the fourth drawback, a dynamic keystream selection mechanism is introduced.

If the initial conditions remain the same, the introducer can easily acquire the keystream by known chosen plain text and chosen ciphertext attacks. This possibility can be prevented by dynamic keystream generation, in which the keystream generated will be relevant to audio segments. Therefore different audio segment generates completely different keystream which eliminates chosen plain text and chosen ciphertext attacks.

The rest of this paper is organized as follows: The preliminary studies of the proposed audio encryption algorithm are presented in Section 2. Theoretical framework of the proposed approach is given in Section 3. Numerical simulations and performance evaluations are discussed in Section 4. Comparison of proposed work with other state-of-art is discussed in Section 5, followed by conclusion in Section 6.

2. Preliminary studies

In this section, we describe mathematical models of chaotic maps used for encryption, i.e., parametric perturbated Lorenz-hyperchaotic system and modified Henon map. Periodic, quasi-periodic, chaotic and hyper-chaotic behavior of parametric perturbated Lorenz map is discussed by means of bifurcation diagram. One dimensional signal compression by FWHT is also discussed.

2.1 Hyperchaotic system

Higher dimensional chaotic systems show distinct advantages over lower dimensional chaotic system due to its complex dynamic random behavior. In this work, we adopt a parametric perturbated Lorenz- hyperchaotic system [34]. Lorenz system shows chaotic behavior for the control parameters, $a = 10$, $b = 8/3$ and $c = 28$ [35]. Parametric perturbation in Lorenz system may be given to all of the parameters *a, b & c* or on selected parameters. In the proposed system (1) the control parameter '*a*' is selected for parametric perturbations by adding a PI controller in the feedback path of the Lorenz system.

$$\begin{aligned} x^{\cdot} &= a(y-x)+w \\ y^{\cdot} &= cx-y-xz \\ z^{\cdot} &= xy-bz \\ w^{\cdot} &= ki\ \ a(y-x)+kp\ \ x \end{aligned} \tag{1}$$

where x, y, z, w are the state variables and a, b, c, k_p, k_i are the system parameters. k_p and k_i are control parameters of the PI controller. Parametric perturbation changes the three dimensional autonomous system to non-autonomous system (1), which is equivalent to a four dimensional hyperchaotic system. When the control parameters, $a = 10$, $b = \frac{8}{3}$, $c = 10$, $k_p = -3.6$, $k_i = 5.2$, and initial conditions (1,1,1,1), then the Lyapunov exponent obtained are $L_1 = 0.01000$, $L_2 = 0.421905$, $L_3 = -0.326781$, $L_4 = -13.385272$. Since more than two Lyapunov exponents are positive, the system is hyperchaotic. The evolution of periodic, chaotic, and hyperchaotic attractors in this system can be generated by varying k_i $[-20, 15]$ by fixing all other parameters constant. Figure 1 illustrates the bifurcation diagram and Lyapunov exponents of the modified system.

2.2 Henon map dynamical system

Henon map was proposed by Michel Henon in 1976 as a comprehensible approach of the Poincare map that results from the solution of complex Lorenz equation [36]. Modified Henon map was developed to improve the complex dynamic behavior and bifurcation range [30]. In this map, x_n^2 term of the seed map is replaced with nonlinear term $\cos(x_n)$. Modified Henon map can be mathematically modeled as follows:

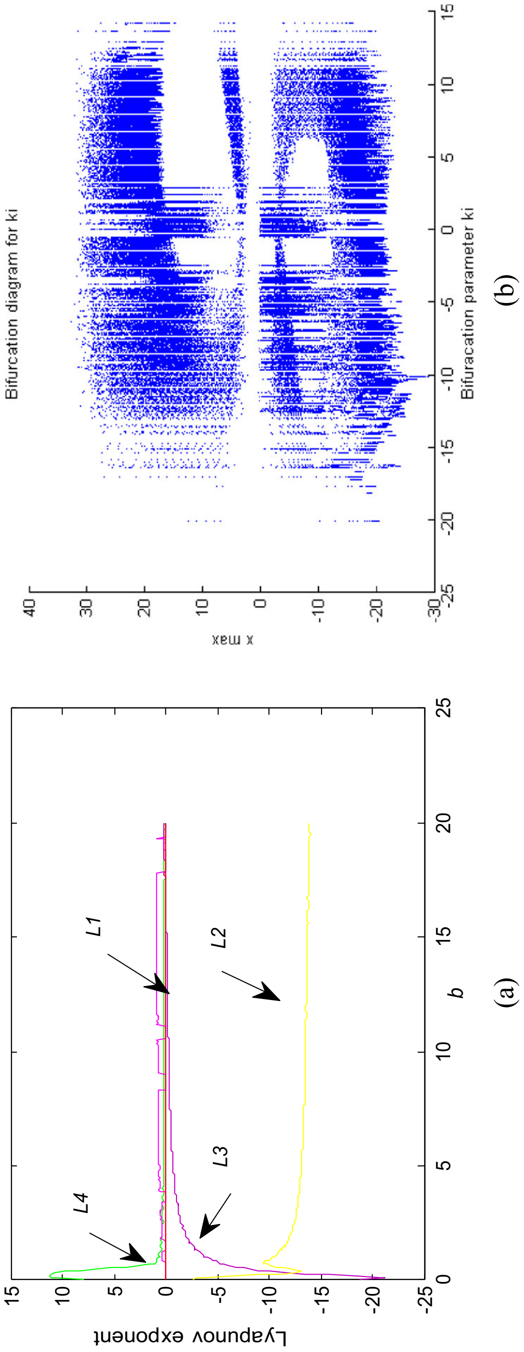


Figure 1.
(a) Lyapunov exponents b) Bifurcation diagram for k_i [−20, 15].

$$\begin{aligned} x_{n+1} &= 1 - \mathrm{acos}(x_n) - by_n \\ y_{n+1} &= -x_n \end{aligned} \tag{2}$$

In the proposed method, the system (2) is taken as a computational tool to generate permutation matrix. The data set resulted from dynamic system (2) for some specific values of a and b are used to generate permutation matrix for encryption process. This dynamical system (2) takes a point (x_n, y_n) in the two dimensional plane and map this point to a new point given by (x_{n+1}, y_{n+1}). The map is dependent on two bifurcation parameters $a(> 0)$ and $b(> 0)$. Moreover, the parameter b measures contraction rate of the 2D quadratic Henon map which is independent of x_n and y_n. Figure 2 depicts the bifurcation diagram of both the systems. Bifurcation diagram of modified Henon map shows wider range of output distribution for the control parameter a compared to its seed map. Therefore, modified henon map increases the range of permutation operation compared to its seed map.

2.3 Fast Walsh Hadamard Transform (FWHT)

Walsh-Hadamard transform is used in different applications, such as data compression, processing of speech and image signals, coding and communications. It is an orthogonal transformation that decomposes input signals into rectangular waveforms called Walsh functions. The Walsh functions forms a system of orthogonal functions and have only two values +1 and –1. This transformation is computationally simple since it has no multiplication and division operations. To implement FWHT of order $n = 2^m$ requires only $nlogn$ addition and subtraction. Generally, Fast Walsh Hadamard transformation follows the recursive definition of symmetric Hadamard matrix. Let H be the Hadamard of order $n = 2^k$ described as follows:

$$\begin{aligned} &H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix} = H_2 \otimes H_{2^{k-1}} \\ &k = 1, 2, 3, \ldots\ldots\ldots\ldots\ldots\ldots, \\ &H_1 = 1 \end{aligned} \tag{3}$$

$$when\ \ k = 1, H_2 = H_2 \otimes H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$when\ \ k = H_4 = H_2 \otimes H_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix} \tag{4}$$

where $\otimes$ denotes Kronecker product.

The Walsh transform is modeled as in Eq. (5) for one dimensional signal.

$$W(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\left[\prod_{i=0}^{n-1}(-1)^{b_i(x)b_{n-1-i}(u)}\right] \tag{5}$$

where x and u are independent variable represented in n bits. The binary representation of x and u can be written as:

$$(x)_{10} = (b_{n-1}(x)b_{n-2}(x)\ldots b_0(x))_2$$

$$(u)_{10} = (b_{n-1}(u)b_{n-2}(u)\ldots b_o(u))_2$$

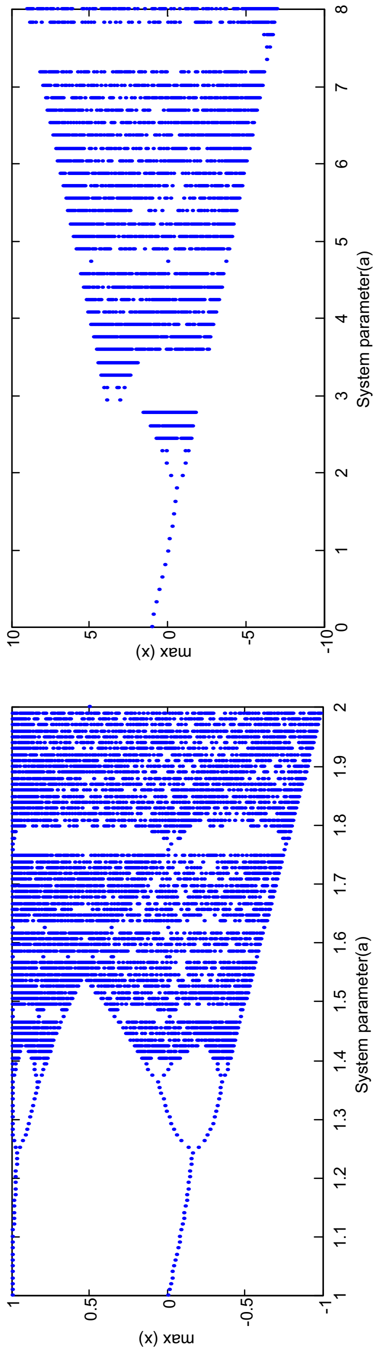


Figure 2.
(a) Bifurcation of diagram Henon map (b) Bifurcation diagram of modified Henon map.

$$\text{with } b_i(x) = 0 \ \textit{or} \ 1 \ \text{for} \ i = 0, \ \cdots.n-1 \tag{6}$$

In this transformation, Walsh kernel forms an array of matrix having orthogonal rows and columns. Therefore, both the forward and inverse transformations are identical operations except for the constant multiplicative factor of 1/N for 1D signal.

$$f(x) = \sum_{u=0}^{N-1} W(u) \left[\prod_{i=0}^{N-1} (-1)^{b_i(x)b_{n-1-i}(u)} \right] \tag{7}$$

3. Proposed method

The overall idea of the encryption process is depicted in Figure 3. The input speech signal is initially compressed by FWHT followed by permutation operation. Then, generate keystream for substitution operation relevant to characteristics of plain speech signal. The various steps in encryption process are systematically demonstrated as follows:

3.1 Encryption process

3.1.1 Audio signal compression by FWHT. Initially, we divide the input audio file into frames, each with N samples. Assume the message signal is $m = \{m_1, m_2, ..m_i, \cdots m_N\} i = 1, 2, \cdots\cdots N$. We compress the audio signals based on Eq. (5) and reduce the sample space by discarding the higher order coefficients (8). Original signal, compressed signal with numerical values are displayed in Figure 4.

$$\begin{aligned} &m_w = FWHT(m) \\ &m_w(N - L \ : \ length(m_w)) = 0; \ \ length(m_w) = N \end{aligned} \tag{8}$$

where $m_w(P, \ 1)$ is the compressed data signal and P is the sample size of the compressed signal.

3.1.2 Permutation process. Prior to permutation process, the audio samples are reshaped $m_w(P, \ 1)$ into two dimensional vector space $m_w(Q, R)$. In this step permute the audio samples depending on the random matrix generated from the modified Henon map by performing some transformation on the original data samples that produce ciphered audio samples $C_1(Q, R)$, with uniform histogram. The minimum number of iterations for Henon map should be greater than P^2 to completely permute the data samples. A chaotic system shows gradual evolution from Periodic to quasi-periodic and to chaotic regime by slowly varying the control parameters. Since there is a slow transition from periodic to chaotic, there may be a chance to produce periodic or redundant samples for the first few iterations. In this scenario, the first few iterations in the permutation process seems fairly close together, hence it should be discarded. Therefore, the total number of iterations is $p^2 + 1000$. *Algorithm 1* describes the entire permutation process in detail. Iterate the data samples as follows:

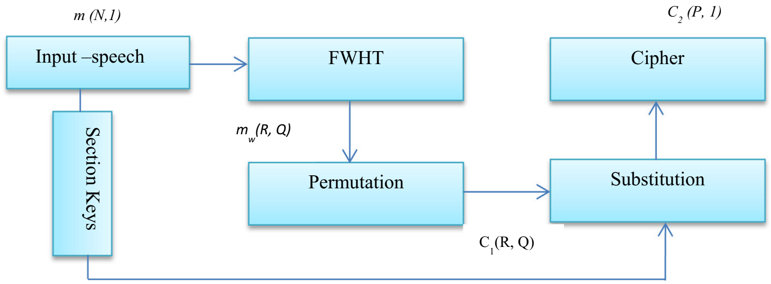


Figure 3.
Block diagram of the proposed algorithm.

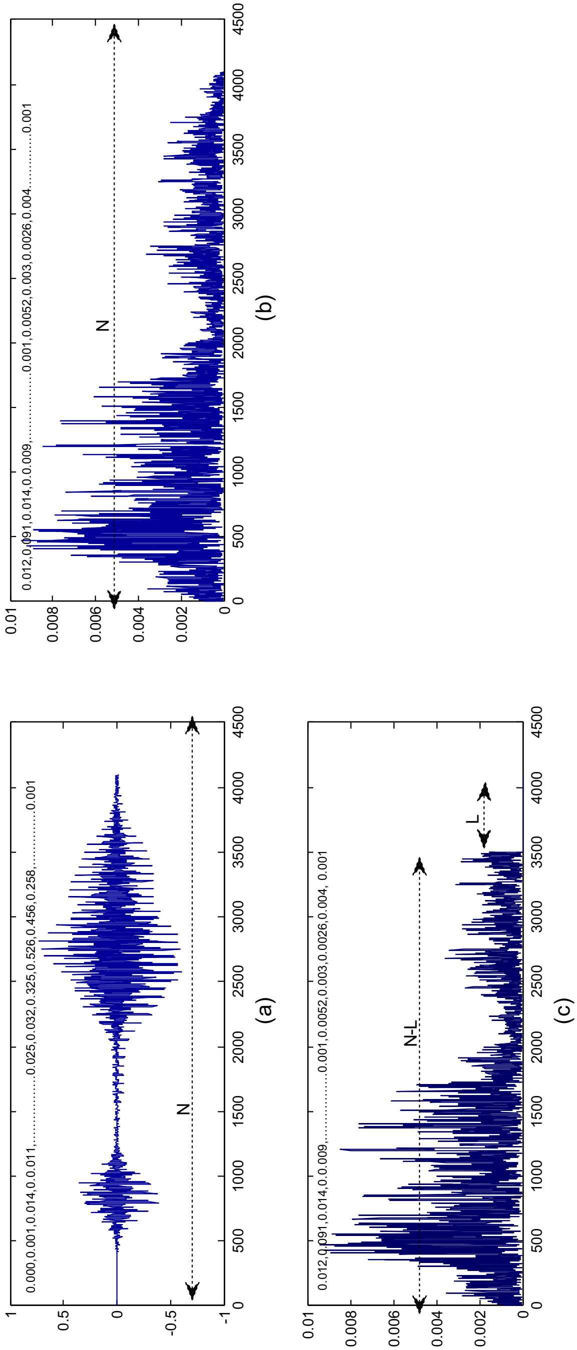


Figure 4.
(a) Original speech (b) Compressed speech signal (c) Compressed signal with zeroing out the coefficients.

$$\begin{aligned} x_i + 1 &= mod\big((1 - a\cos x_i) + by_j),\ Q)\ i \in (1, Q) \\ y_{j+1} &= mod((x_i), R)\ j \in (1, R) \end{aligned} \tag{9}$$

where (x_i, y_j) is the initial position of the data sample and (x_{i+1}, y_{j+1}) is the first iterated position. Figure 5 displays the values of original and permuted data samples.

Algorithm 1. Permutation of speech samples

```
Input: Audio files, system parameters of modified Henon map
  Output: permuted speech samples
Set audio file with frequency 8 KHz for a duration of 0.5 sec
  and system parameters a = 3.58, b = 0.56
Read and segment the audio file
%%Compression, Compress the signal and Zeroing out the
  coefficient.
$m_w = wht(audio.wav)$
$m_w(N - l : length(m_w)) = 0;$
%%permutation
Reshape($m_w, Q, R$)
for i = 1:Q do
  for j = 1:R do
    $x_i + 1 = 1 - a * cos(x_i) + by_j$
    $y_j + 1 = bx_i$
  end for
end for
```

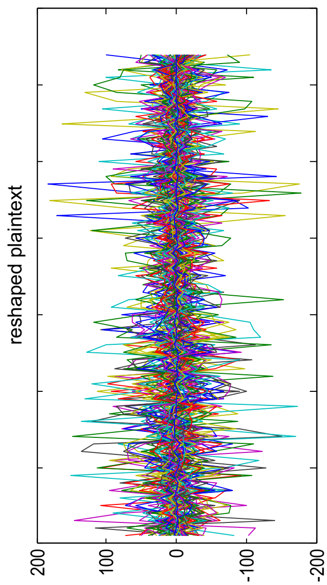
3.1.3 Dynamic keystream selection mechanism. In order to make initial conditions dependent on each other and be sensitive on plain audio data, the initial conditions are derived from Eq. (10) and by performing basic arithmetic operations as in Eq. (11). The sub keys so obtained are the initial conditions $x(0), y(0), z(0), w(0)$ for the hyperchaotic system (1). In each round of operation the keys are updated, and it will avoid the possibility of various differential attacks.

$$\begin{aligned} key(1) &= \frac{1}{N/2}\sum_{i=0}^{N} m_w(2i+1)i = 1, 2, \cdots .N \\ key(2) &= \frac{1}{N/2}\sum_{i=0}^{N} m_w(2i) \end{aligned} \tag{10}$$

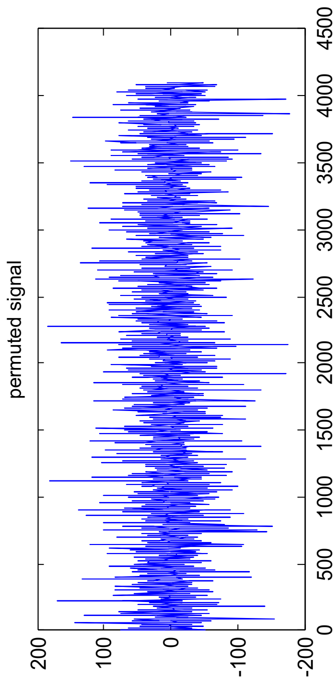
$$\begin{aligned} subkey(1) &= key(1) + key(2), mod1 \\ subkey(2) &= key(1) - key(2) \\ subkey(3) &= key(1) \times key(2) \\ subkey(4) &= key(1)\tilde{A} \cdot key(2) \end{aligned} \tag{11}$$

After computing the initial conditions $[\mathrm{x}(0), \mathrm{y}(0), \mathrm{z}(0), \mathrm{w}(0)]$ for the modified Lorenz hyperchaotic system, generate the keystream by iterating the system (1) by Runge-Kutta method. Convert the four hyper-chaotic sequences into integer sequences $\{x_i^*\}, \{y_i^*\}, \{z_i^*\}, and\ \{w_i^*\}$ as follows:

$$\begin{aligned} x_i^* &= mod\big((abs(x_i) - floor(x_i))10^{14}, 1\big) \\ y_i^* &= mod\big((abs(y_i) - floor(y_i))10^{14}, 1\big) \\ z_i^* &= mod\big((abs(z_i) - floor(z_i))10^{14}, 1\big) \\ w_i^* &= mod\big((abs(w_i) - floor(w_i))10^{14}, 1\big) \end{aligned} \tag{12}$$



(a)



(b)

-0.948	7.4009	67.077	.....	20.27
-2.993	9.0728	-13.18	.....	-2.693
1.307	-17.25	-56.17	.....	7.4009
...	.....	.....	.....	....
-2.706	40.646	-15.92	...	-20.27
7.4009	67.077	7.4009	67.077	-26.60
9.0728	-13.18	9.0728	-13.18	0.0396

(c)

-20.27	-26.60	-15.92	.....	40.646
-2.693	0.0396	-1.148	.....	67.077
7.4009	0.950	-4.465	.....	-13.18
...	.....	.....	.....	40.646
0.948	-9.238	-1.109	...	
40.646	34.646	54.646	9.646	100.646
67.077	-2.706	67.077	-2.706	67.077

(d)

Figure 5. (a) Original speech signal in 2D vector space (b) Permuted speech signal (c) Original data samples in 2D vector space (d) Permutated sample values.

x_i, y_i, z_i, w_i indicate the i^{th} element of keystreams. The $abs(x)$) returns the absolute value of x and $floor(x)$ returns the largest integer less than or equal to x. Generate the normalized keystream for substitution operation as follows:

$$X = \mathrm{XOR}\left(x_i^*, y_i^*, z_i^*, w_i^*\right)$$
$$key\_norm = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{13}$$

where *key_norm* is the normalized keystream generated. X_{max}, X_{min} are the maximum and minimum values present in the array X. The sequence of operation is elaborated in *Algorithm 2.*

Algorithm 2: Dynamic keystream Generation

```
Input: compressed signal coefficient, set the system parameter as
   a = 10; b = 10/3; c = 28; kp = –3.5; ki = 5.2. Time step = 0.005
Output: Pseudo random number sequences
%% initial conditions for hyperchaotic system
for i = 1:N
    $key(1) = \frac{2}{N} sum(2i + 1)$
    $key(2) = \frac{2}{N} sum(2i)$
end for
%%%sub key generation
$subkey(1) = key(1) + key(2), mod1$
$subkey(2) = key(1) - key(2)$
$subkey(3) = key(1) \times key(2)$
$subkey(4) = key(1)\tilde{A} \cdot key(2)$
%%% keystream generation
%%set the initial conditions for hyperchaotic system as x(0)
   = subkey(1); y(0) = subkey(2) z(0) = subkey(3); w(0) = subkey(4)
for k = 1:N
    $fx = a * y(k) - a * x(k) + w(k)$
    $fy = -x(k) * z(k) + c * x(k) - y(k)$
    $fz = x(k) * y(k) - b * z(k)$
    $fw = kp * a * (y(k) - x(k)) - kix(k)$
    %%update the values of x, y, z, w
    $x(k + 1) = x(k) + dt * fx$
    $y(k + 1) = y(k) + dt * fy$
    $z(k + 1) = z(k) + dt * fz$
    $w(k + 1) = w(k) + dt * fz$
    $t(k + 1) = t(k) + dt$
end for
%% Convert the sequences into integer sequences
for i = 1:N
    $x_i = mod((abs(x_i) - floor(x_i))10^4, 1)$
    $y_i = mod((abs(y_i) - floor(y_i))10^4, 1)$
    $z_i = mod((abs(z_i) - floor(z_i))10^4, 1)$
    $w_i = mod((abs(w_i) - floor(w_i))10^4, 1)$
end
for k = i:N
    X = XOR(x(i),y(i),z(i),w(i))
end for
%%Normalized key stream generation
key_norm=$X - X_{min}/X_{max} - X_{min}$
```

3.1.4 Substitution operation. After first level encryption process (permutation operation), the two dimensional data samples $(C_1(Q,R))$ are reshaped to one dimensional data samples $(C_1(P,1))$. Then, generate keystream for substitution operation based on *Algorithm 2* and eliminate first few samples of the keystream since the samples are redundant. Substitution operation is then carried out by XOR-ing the permuted samples with the keystream generated using Eq. (12). Figure 6(a) and (b) show the speech samples after p^2+1000 permutation operation and speech signal after substitution operation respectively. Pseudo code for substitution operation is given as in *Algorithm 3*.

$$C_2=\sum_{i=0}^{P}C_1(i)\cdot \mathrm{key\_norm} \tag{14}$$

```
Algorithm 3. Substitution operation
Input: permuted audio file, dynamic keystream
   Output: encrypted speech file
1: for i = 1:p do
2: $C_2 = xor(C_1(i).key\_sub)$
3: end for
```

3.2 Decryption process

The procedure of decryption process is just reverse of the encryption process. The decryption process can be performed easily by means of the pre-shared keys. The decryption process can be briefly described as follows*:*

Step 1: Generate the same keystream bits according to the steps *3.1.3* in the encryption process.

Step 2: Xoring the samples $C_2(P,1)$ with the keystreams.

Step 3: Do the inverse permutation process according to the step *3.1.2*.

Step 4: Take Inverse Fast Walsh Hadamard Transform (IFWHT).

4. Simulation result and analysis

The proposed algorithm is simulated on a classical computer with MATLABR 2013a (version) software. Different voice samples of male and female audio files with sampling rate of 8000 samples/sec are selected for the test. We evaluated the performance of this algorithm through various statistical and differential analyses.

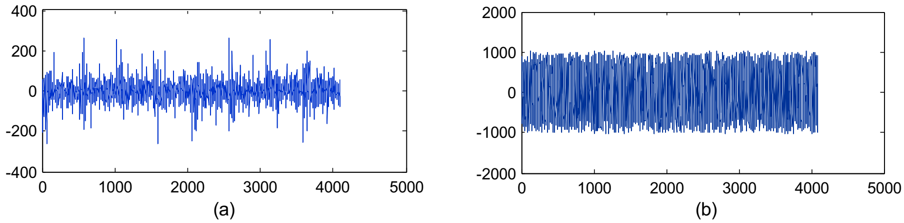


Figure 6. Encrypted speech signal after second level of encryption (a) Permuted speech signal after p^2+1000 rounds of iterations (b) Speech signal after substitution operation.

4.1 Correlation analysis

Correlation analysis is a statistical method to evaluate the performance of cryptographic algorithm over various statistical attacks [37]. Correlation coefficient analysis measures the mutual relationship between similar segments in the plain and encrypted audio file. A secure data encryption algorithm converts original data into random-like noisy signal with low correlation coefficient. Low correlation coefficient indicates the narrow correlation between original and encrypted speech files. In this work Correlation analysis is carried out for both Henon map and modified Henon map. Correlation coefficient (r_{xy}) between original and encrypted audio samples are calculated and listed in Table 1. Correlation coefficient between original data samples in the same audio file, and original and encrypted data samples are also calculated and given in Table 2. Analysis shows that there is an improvement of permutation operation since the correlation coefficients of modified Henon map is smaller than its seed map. Figure 7(a) shows the scatter plot diagram of original speech signal. Randomized nature of speech signal after permutation and substitution operation is illustrated in Figure 7(b) and (c). Correlation coefficient can be calculated as follows:

$$r_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{\frac{1}{N_S}\sum_{i=1}^{N_s}(x_i - E(x))(y_i - E(y))}{\sqrt{\frac{1}{N_S}\sum_{i=1}^{N_s}(x_i - E(x))^2}\sqrt{\frac{i}{N_S}\sum_{i=1}^{N_s}(y_i - E(y))^2}} \tag{15}$$

$$\sigma_x, \sigma_y \neq 0$$

where $E(x)$ and $E(y)$ are mean and σ_x, σ_y are the standard deviation of the encrypted and decrypted speech signal.

4.2 Signal to noise ratio (SNR)

The signal to noise ratio is one of the straight forward methods to validate the performance of data encryption algorithm. SNR measures the noise content in the encrypted data signal.

Table 1.
Correlation coefficient (r_{xy}).

Audio files	Henon map	Modified Henon map
Audio.wav1	0.00163	0.000991
Audio.wav2	0.00451	0.001359
Audio.wav3	0.00856	0.000651
Audio.wav4	0.00421	0.001569
Audio.wav5	0.00131	0.001865
Audio.wav6	0.00969	0.002634
Audio.wav7	0.005624	0.001695

Table 2.
Correlation coefficient (r_{xy}).

Audio files	Original	Encrypted
Audio.wav1	0.98153	0.000991
Audio.wav2	0.96421	0.001359
Audio.wav3	0.98816	0.000651
Audio.wav4	0.97461	0.001569
Audio.wav5	0.99141	0.001865
Audio.wav6	0.98979	0.002634
Audio.wav7	0.97562	0.001695

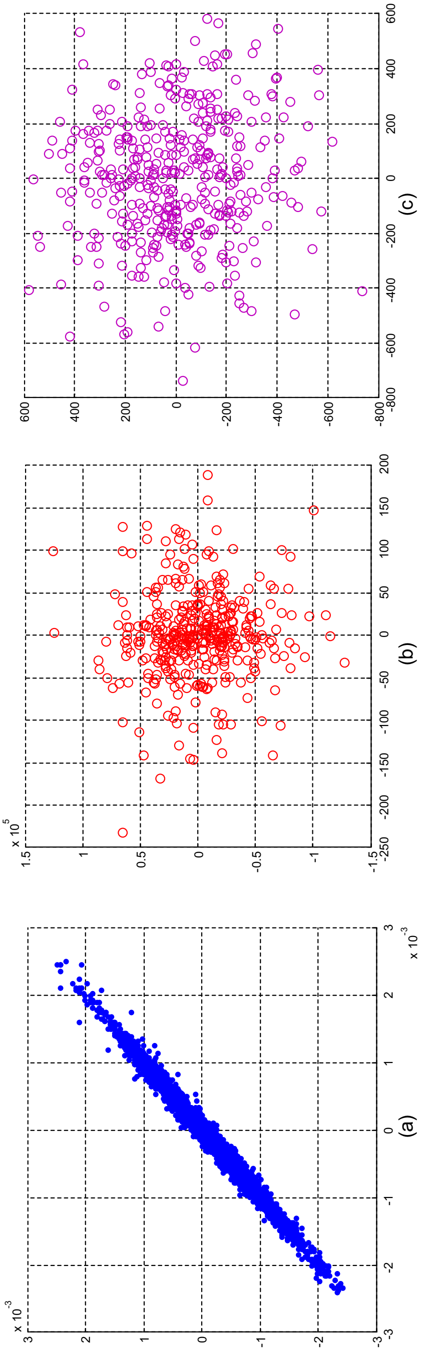


Figure 7.
Scatter plot diagram of (a) Original signal (b) Encrypted signal after permutation (c) Encrypted signal after substitution.

Cryptanalyst always try to increase the noise content in the encrypted signal so as to minimize the information content in the encrypted data [38]. Figure 8 shows the original and encrypted speech signal. Figure 8(c, d) illustrates the randomized nature of encrypted signal after permutation and substitution operation. The SNR values of encrypted audio files are calculated based on the Eq. (16) and given in Table 3.

$$SNR = 10 * log10 \frac{\sum_{i=1}^{N_s} x_i^2}{\sum_{i=1}^{N_s} (x_i - y_i)^2} \tag{16}$$

4.3 UACI and NSCR analysis

In data encryption, resistace to differential attacks is generally analyzed through the NSCR (number of samples change rate) and UACI (unified average changing intensity) tests [39]. In this analysis two different speech segments are encrypted with same keystreams, where the original speech segments are differed by one sample space. Then the encrypted speech segments are compared by the number of sample change rate (NSCR) and the unified average changing intensity (UACI). Both these parameters can be expressed as follows:

$$NSCR = \sum_i \frac{D_i}{N} \times 100\% i = 1, 2 \cdots .N$$

$$UACI = \frac{1}{N} \left[\frac{\sum_i x_i - x_i^{'}}{65535} \right]$$

where

$$d_i = \int \begin{matrix} 1, x_i \neq x_i^{`} \\ 0, otherwise \end{matrix} \tag{17}$$

c_i and $c_i^{`}$ denotes the the audio samples at ith position of the encrypted speech samples and N corresponds to the length of the speech segments. The upper-bound for NSCR and UACI are 100% and 33.3% respectively. For a secure encryption scheme these parameters should be close to the upper bound ideal values. NSCR and UACI values of the proposed algorithm is calculated and listed in Table 4. The results show that the values obtained through proposed algorithm is considerably closer to ideal values.

4.4 Spectral entropy

Spectral entropy measures the randomness in both encrypted and original speech signal. Its measurement is based on the assumption that the spectrum of meaningful speech segment is correlated than the noisy signal [1]. The spectral measurement compares the entropy where the amplitude component of the power spectrum is taken as a probability parameter in entropy calculation. The amount of information can be calculated as the negative of entropy or the negative logarithm of probability. Thus, meaningful speech segments show low entropy since it contains organized data samples. However the encrypted speech signals have high entropy and large spectral peaks similar to noisy signal. The entropy E_i can be measured as follows:

$$E_i = \sum_n PSD_n(f_i) \log(PSD_n(f_i)); i = 1, 2, 3 \cdots, n \tag{18}$$

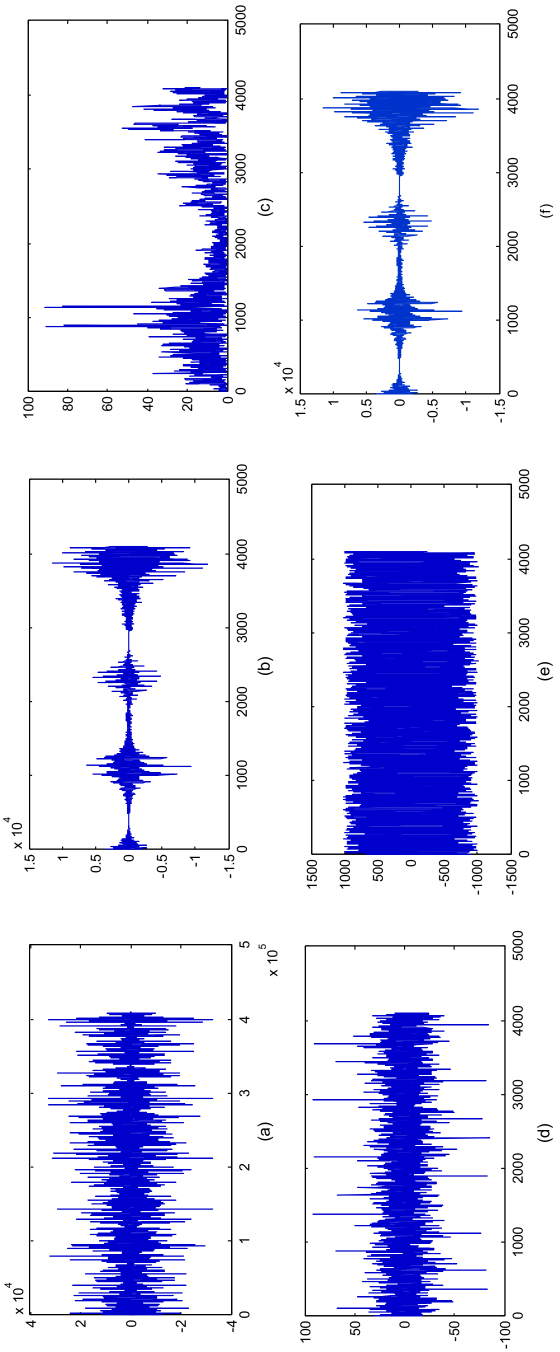


Figure 8.
(a) Original signal (b) Compressed signal (c) Data samples after permutations (d) Data samples after substitution (e) Decrypted signal.

where PSD_n is the normalized power spectrum and f_i is the frequency of the signal. Irregularities of amplitude in original and encrypted signals are shown in Figure 9.

4.5 Keyspace and key sensitivity analysis

In the proposed algorithm, system parameters of modified Henon map $(a = 3.58, b = 0.56)$, system parameters of modified Lorenz-hyperchaotic system $(a = 10, b = 3.33, c = 28, k_p = -3.5, k_i = 5.2)$ and initial conditions of the hyperchaotic system $(x(0), y(0), z(0), w(0))$ constitutes the keyspace. If the precision of each system parameter and initial condition is set to 15 decimal points, the key space of the proposed algorithm is $(10^{15})^6 = 2^{548.11}$. It is sufficiently large enough to resist the brute force attack. Key sensitivity is the essential quality for any good data encryption algorithm, which make sure that the security level of the algorithm against the brute-force attack [40]. It means that, a small variation for any key parameter bring an apparent change in both encrypted and decrypted speech signal. The effect of variation in keyparameter on encryption process is verified by encrypting the signal with slightly different initial conditions. The simulation result shows that the slight variations in keyparameter will result in completely different encrypted signal. Figure 10 shows the encrypted signals with two different initial conditions. To evaluate the key sensitivity of decrypted signal, encrypt the speech file with one fixed secret key then decryption is performed with slightly different keys. The resulting speech files decrypted with wrong keys apparently looks different and reveals no information. Figure 11(a) shows the decrypted speech signal with correct key. Figure 11(b, c, d, e) shows the decrypted signal with slight variations in the initial conditions in the range of 10^{-15}.

4.6 Histogram analysis

Histogram analysis is one of the accurate methods to evaluate the quality of encrypted speech signal. Since a practical encryption algorithm is likely to encrypt original speech file into random like noise, it is desirable to obtain an encrypted speech file with equally probable sample values. Therefore, the encrypted speech furnishes no information that would facilitate

Table 3.
Signal to noise ratio.

Audio files	SNR (dB)
Audio.wav1	−133
Audio.wav2	−154
Audio.wav3	−124
Audio.wav4	−110
Audio.wav5	−135
Audio.wav6	−137
Audio.wav7	−144

Table 4.
Differential analysis.

Audio files	NSCR	UACI
Audio.wav1	99.9997	33.33
Audio.wav2	99.9999	33.28
Audio.wav3	99.9998	33.34
Audio.wav4	99.9996	33.25
Audio.wav5	99.9992	33.31
Audio.wav6	99.9993	33.29
Audio.wav7	99.9989	33.30

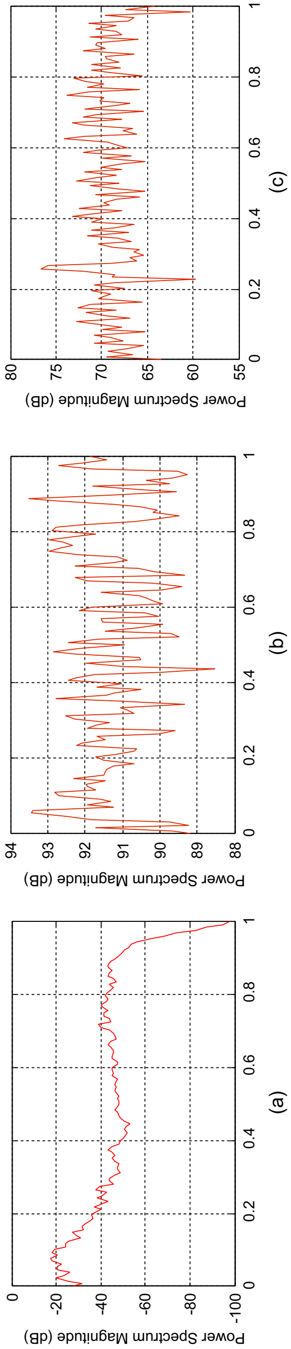


Figure 9.
Power Spectral Density (a) Original speech signal (b) Permuted speech signal (c) Substituted speech signal.

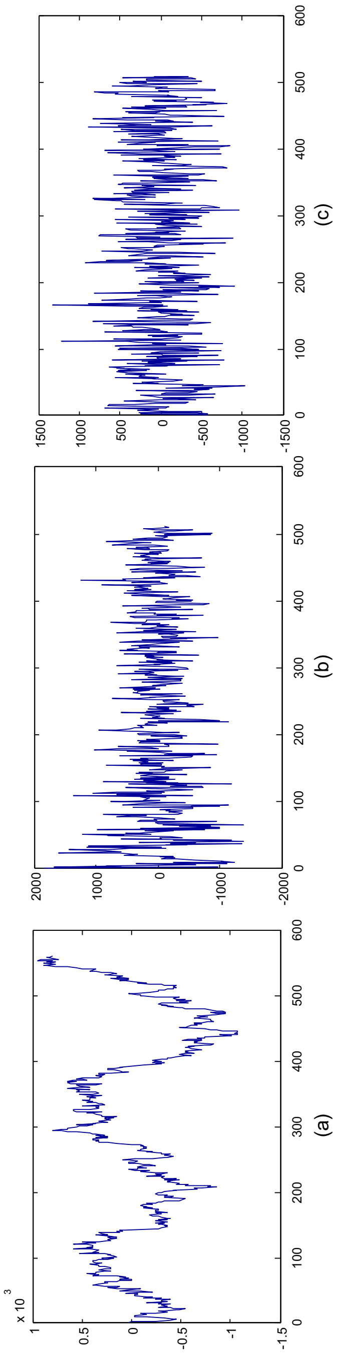


Figure 10.
Key sensitivity on encryption process (a) Original speech signal (b) Encrypted speech signal for key $x0=0.413$, $y0=-0.931$, $z0=0$, $w0=0.825$ (c) Encrypted speech signal for $x0=0.913$, $y0=-0.131$, $z0=0.123$, $w0=0.925$.

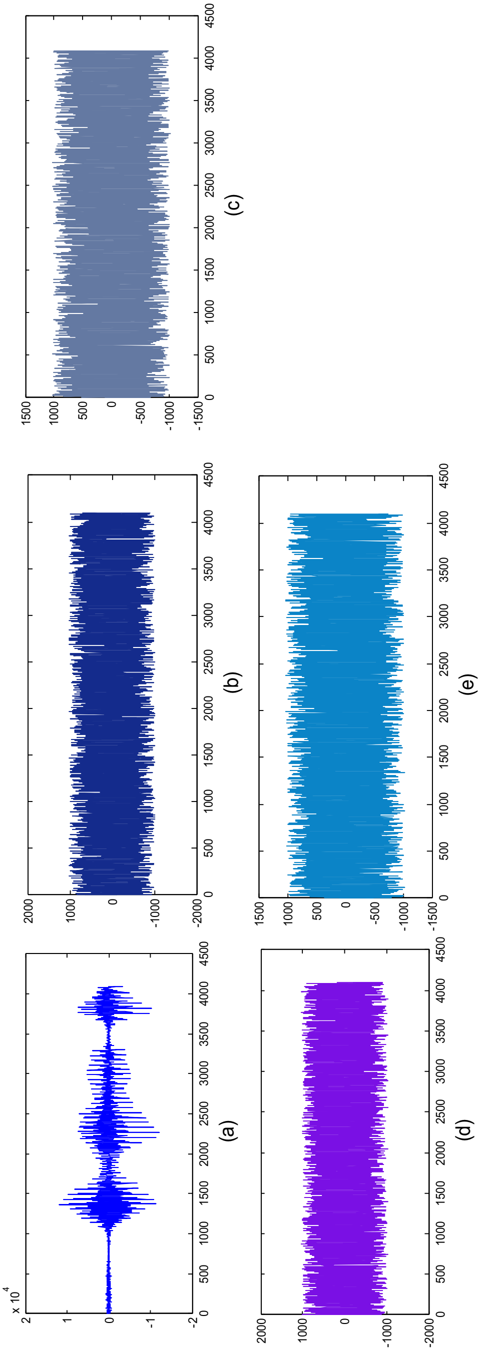


Figure 11.
Key sensitivity on decryption process (a) decrypted signal with correct key, decrypted signal with incorrect key (b) $x_0 + 10^{-15}$ (c) $y_0 + 10^{-15}$ (d) $z_0 + 10^{-15}$ (e) $w_0 + 10^{-15}$.

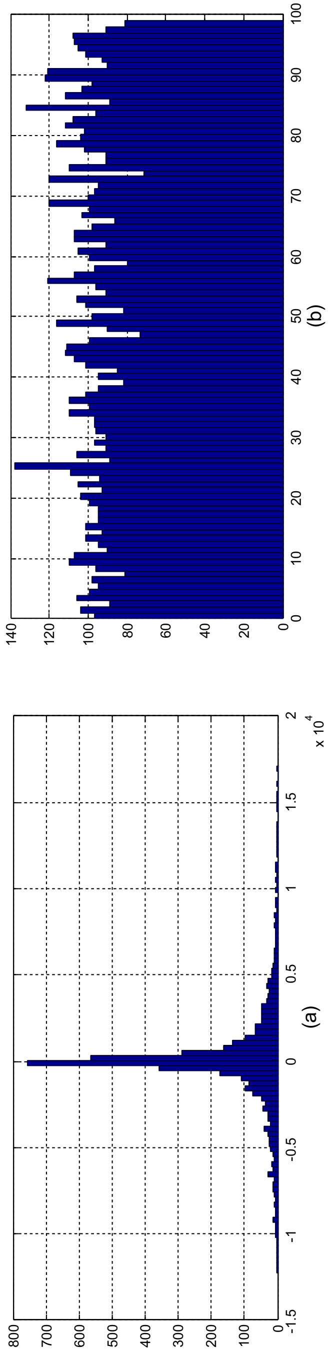


Figure 12.
Histogram of (a) Original speech signal (b) Encrypted speech signal.

the possibility of any statistical attacks on the encrypted domain. Histogram of both encrypted and original speech signal is illustrated in Figure 12. Figure 12(b) displays the uniformly distributed histogram of encrypted speech signal, which indicates the randomness of encrypted speech signal. From the histogram, it is clear that the proposed algorithm is highly secure against various statistical attacks.

4.7 Computational complexity

Big-oh Notation is a unified way to express the complexity of an algorithm. In classical computation, computational complexity is evaluated by the elementary operations involved in the encryption process. Since speed of an algorithm depends on the target computer processor, it is difficult to estimate the exact runtime of an algorithm. Big-oh notation measures the execution time of an algorithm in terms of the input array size and the nature of arithmetic operations. In the proposed algorithm, the encryption process consists of $p^2 + 1000$ round of permutation operations and a single round of substitution operations. Computational complexity of permutation operation is $O(n^2)$ time or quadratic time, since the time complexity of the operation grows quadratically with respect to input array size n. However, the computational complexity of substitution operation is independent of input array size, since this process takes single step to complete the operation irrespective of array length n. Thus the computation complexity of substitution round is $O(1)$. Computational complexity of entire process can be expressed as $O(n^2) + O(1) = O(n^2)$.

5. Comparison with existing works

The proposed speech encryption algorithm differs from other methods, in terms of data compression, permutation and substitution operations. Therefore comparison of proposed method with other state-of-the-art approaches is difficult. However, we have analyzed various quality metrics such as key length, keyspace, signal to noise ratio (SNR), NPCR, UACI and correlation coefficient between original and encrypted signals and tabulated in Table 5. We have compared our proposed algorithm with advanced encryption standard (AES), Data Encryption Standard (triple DES), algorithm based on quantum chaotic system [16] and an algorithm based on substitution-permutation chaotic network [20]. Speech encryption algorithm based on Zaslavsky map [8], TD-ERCS chaotic map [33], multiple chaotic shift keying [29] and a non-chaotic [31], method are also considered for comparative analysis. The size of the proposed method's key space is greater than 2^{540} (Section 4.5). It is clear from the simulation results (Table 3) that the encrypted speech signal contains more noise content than in original speech signal. Correlation coefficient (r_{xy}) is evaluated to be almost zero (Table 2). Also NSCR and UACI analysis validates the capacity of the proposed algorithm

Method	Key length	Key space	CC	SNR	NPCR	UACI
AES	128, 192, 256	2^{128}, 2^{192}, 2^{256}	0.00971	−1.4461	99.6032	33.3218
Triple DES	168	2^{168}	0.1704	−0.250	99.5985	33.3419
Ref. [31]	521	$>2^{512}$	0.0011	−10.6357	99.6399	33.8085
Ref. [29]	533	2^{744}	0.0233	−33.7464	99.9982	33.1197
Ref. [8]	477	2^{477}	0.0029	−23.89	99.8725	33.1160
Ref. [16]	>264	$>2^{624}$	0.0491	−44.8	99.6399	33.1085
Ref. [20]	2^{107}	2^{107}	0.0321	−54.89	99.6317	33.2782
Ref. [33]	398.66	$2^{398.66}$	8 0.016 12,830	−130	99.6521	33.2122
[Prop: Meth]	548.11	$2^{548.11}$	0.0.0009	−133	99.9989	33.3421

Table 5. Quality metrics comparison of various encryption methods.

against various differential attacks. From this analysis, we can found that proposed algorithm shows considerable improvement in almost all the encryption quality metrics.

6. Conclusion

In this paper, a novel approach for speech encryption algorithm based on parametric perturbated Lorenz-hyperchaotic system and modified Henon map is introduced. Modified Henon map maximized the permutation operation compared to its seed map, which eventually decreased the correlation between original and encrypted data samples. Selection of hyperchaotic system eliminated weak chaotic trajectories and smaller chaotic ranges, which is commonly observed in lower dimensional chaotic system. Due to the hyperchaotic nature, the proposed system has more keyspace, which protects the proposed algorithm against various statistical attacks like brute force attack. Moreover, dynamic keystream generated with hyperchaotic system eliminated the possibility of differential attacks. Furthermore, Fast Walsh Hadamard Transform (FWHT) improved the efficiency of algorithm by reducing the computational complexity while doing the compression process. Various simulations and numerical analysis have been carried out on classical computer to evaluate the performance of the proposed algorithm. Finally, we have made a comparison of proposed algorithm with chaotic, non-chaotic and standard encryption algorithms. From the comparative study, it can be concluded that the proposed algorithm shows improvement over some of the existing algorithms and it is an excellent choice for voice encryption in practical applications.

References

[1] C.E. Shannon, Communication theory of secrecy systems, Bell Syst. Tech. J. 28 (1949) 656–715.

[2] C. Li, When an attacker meets a cipher-image in 2018: A year in review, in arXiv:1903.11764, 2019, [online] Available at: https://arxiv.org/abs/1903.11764.

[3] J.X. Chen, Z.L. Zhu, C. Fu, H. Yu, L.B. Zhang, A fast chaos-based image encryption scheme with a dynamic state variables selection mechanism, Commun. Nonlinear Sci. Numer. Simul. 20 (2015) 846–860.

[4] G.D. Ye, J. Zhou, A block chaotic image encryption scheme based on self-adaptive modeling, Appl. Soft. Comput. 22 (2014) 351–357.

[5] X.L. Chai, An image encryption algorithm based on bit level Brownian motion and new chaotic systems, Multimed. Tools Appl. 76 (1) (Jan 2017) 1159–1175.

[6] C.L. Fu, B.B. Lin, Y.S. Miao, X. Liu, J.J. Chen, A novel chaos-based bit- level permutation scheme for digital image encryption, Opt. Commun. 284 (23) (2013) 5415–5423.

[7] W. Zhang, K.W. Wong, H. Yu, Z.L. Zhu, A symmetric color image encryption algorithm using the intrinsic features of bit distributions, Commun. Nonlinear Sci. Numer. Simul. 18 (3) (2013) 584–600.

[8] F.J. Farsana, K. Gopakumar, A novel approach for speech encryption: Zaslavsky map as pseudo random number generator, in: Proceedings of International Conference on Advances in Computing and Communications, ICACC 2016, 2016, pp. 816–823.

[9] C. Li, G. Luo, Ke Qin, C. Li, An image encryption scheme based on chaotic tent map, Nonlinear Dyn. 87 (2017) 127–133.

[10] X. Zhang, Z. Zhao, Chaos-based image encryption with total shuffling and bidirectional diffusion, Nonlinear Dyn. 75 (2014) 319–330.

[11] B. Norouzi, S. Mirzakuchaki, A fast color image encryption algorithm based on hyper-chaotic systems, Nonlinear Dyn. 78 (2014) 995–1015.

[12] S. Lakshmanan, M. Prakash, C.P. Lim, R. Rakkiyappan, P. Balasubramaniam, S. Nahavandi, Synchronization of an inertial neural network with time-varying delays and its application to secure communication, IEEE Trans. Neural Networks Learn. Syst. 29 (2018) 195–207.

[13] L. Teng, X. Wang, J. Meng, A chaotic color image encryption using integrated bit-level permutation, Multimedia Tools Appl. 77 (2018) 6883–6896.

[14] Y. Luo, R. Zhou, J. Liu, S. Qiu, Y. Cao, An efficient and self-adapting colour-image encryption algorithm based on chaos and interactions among multiple layers, Multimedia Tools Appl. 77 (20) (2018) 26191–26217.

[15] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H. Wang, S.W. Baik, Secure surveillance framework for IoT systems using probabilistic image encryption, IEEE Trans. Ind. Inf. 14 (2018) 3679–3689.

[16] Akram Belazi, Majid khan, Ahmed A. Abd El-Latif, Safya Belghith, Efficient cryptosystem approaches: S-boxes and permutation-substitution based encryption, Nonlinear Dyn. 87 (1) (2017) 337–361.

[17] Mohamed Amin, Ahmed A. Abd El-Latif, Efficient modified RC5 based on chaos adapted to image encryption, J. Electron. Imaging 19 (1) (2010) 013012.

[18] Ahmed A. Abd El-Latif, Li Li, Xiamu Niu, A new image encryption scheme based on cyclic elliptic curve and chaotic system, Multimedia Tools Appl. 70 (3) (2014) 1559–1584.

[19] Ahmed A. Abd El-Latif, Bassem Abd-El-Atty, Muhammed Taha, Robust encryption of quantum medical image, IEEE Access 6 (2018) 1073–1081.

[20] Ahmed A. Abd El-Latif, Li Li, Ning Wang, Qi Han, Xiamu Niu, A new approach to chaotic image encryption based quantum chaotic system, exploiting color spaces, Signal Process. 93 (11) (2013) 2986–3000.

[21] Ahmed A. Abd El-Latif, Bassem Abd-El-Atty, M. Shamim Hossain, Md. Abdur Rahman, Atif Alamri, B.B. Gupta, Efficient quantum information hiding for remote medical image sharing, IEEE Access 6 (2018) 21075–21083.

[22] Ahmed A. Abd El-Latif, Bassem Abd-El-Atty, M. Shamim Hossain, Samir Elmougy, Ahmed Ghoneim, Secure quantum steganography protocol for fog cloud Internet of Things, IEEE Access 6 (2018) 10332–10340.

[23] Long Jye Sheu, A speech encryption using fractional chaotic systems, Nonlinear Dyn. 63 (6/7) (June 2011) 103–108.

[24] E. Mosa, N.W. Messiha, O. Zahran, F.E. Abd El-Samie, Chaotic encryption of speech signal, Int. J. Speech Technol. 14 (2011) 285–296.

[25] Maysaa abd ulkareem, Iman Qays Abdul jaleel, Speech encryption using chaotic map and blowfish algorithm, J. Basrah Res. 39 (2), (2013).

[26] S.N.Al. Saad, E. Halto, A speech encryption based on chaotic map, Int. J. Comput. Appl. 93 (4), (2014).

[27] E. Halto, D. Shihab, Lorenz and Rossller chaotic system for speech signal encryption, Int. J. Comput. Appl. 128 (11), (Oct 2015).

[28] E.M. Elshamy, El-Sayed M. Rabaie, H.S. El-Sayed, Efficient audio cryptosystem based on chaotic maps and double phase random coding, Int. J. Speech Technol. 14 (2015) 285–296.

[29] P. Sathiyamurthi, S. Ramakrishnan, Speech encryption using chaotic shift keying for secured speech communication, EURASIP J. Audio Speech Music Process., Springer Open Access (2017), available at: http://dx.doi.org/10.1186/s13636-017-0118-0.ISSN:1687-4722, ISSN: 1687-4722, Online First.

[30] S.J. Sheela, K.V. Suresh, Deepa knath Tandur, A novel audio cryptosystem using chaotic maps and DNA encoding, J. Comput. Networks Commun, 2017.

[31] Aissa Belmeguenai, Zahir Ahmida, Salim Ouchtati, A novel approach based on stream cipher for selective speech encryption, Int. J. Speech Technol. 20 (July 2017) 685–698.

[32] Animesh Roy, A.P. Misra, Audio signal encryption using chaotic Henon map and lifting wavelet transforms, Eur. Phys. J. Plus (2017).

[33] Z. Habib, J.S. Khan, J. Ahmad, M.A. Khan, F.A. Khan, Secure speech communication algorithm via DCT and TD-ERCS chaotic map, in: Electrical and Electronic Engineering (ICEEE) 2017 4th International Conference on. IEEE, 2017, pp. 246–250.

[34] V.R. Devi, Farooq Farsana, K. Gopakumar, Hyperchaos generated from 3D chaotic systems using PI controller, AIP Conf. Proc. 1859 (020030), (2017).

[35] Edward Norton Lorenz, Deterministic non-periodic flow, J. Atmos. Sci. 20 (2) (1963) 130–141.

[36] M. Henon, A two-dimensional mapping with a strange attractor, Commun. Math. Phys. 50 (1) (1976) 69–77.

[37] J. Ahmed, S.O. Hwang, A secure image encryption scheme based on chaotic and affine transform, Multimedia Tools Appl. 75 (2016) 13951–13976.

[38] B.T. Bosworth, W.R. Bernecky, J.D. Nickila, B. Adal, G.D. Carter, Estimating signal-to-noise ratio (SNR), IEEE J. Oceanic Eng. 33 (2008) 414–418.

[39] F.A. Khan, J. Ahmed, J. Ahmad, A novel image encryption based on Lorenz equation, Gingerbreadman chaotic map and S8 permutation, J. Intell. Fuzzy Syst. 33 (2017) 3753–3765.

[40] Ecrypt, II Yearly Report on Algorithms and Keysizes; 2010. available at: http://www.ecrypt.eu.org/documents/D.SPA.13.pdf.

Corresponding author
F.J. Farsana can be contacted at: talk2cryptofarsana@gmail.com