Multimedia Tools and Applications (2022) 81:6331–6349
https://doi.org/10.1007/s11042-021-11757-4

Speech encryption using hybrid-hyper chaotic system and binary masking technique

P. Sathiyamurthi[1] · S. Ramakrishnan[1]

Received: 26 April 2020 / Revised: 30 September 2021 / Accepted: 22 November 2021 /
Published online: 10 January 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

In this paper, a novel speech encryption algorithm based on hybrid-hyper chaotic system has presented. Instead of using normal chaotic system a hybrid-hyper chaotic system has used for improving the security level of speech communication models. Hyper-chaotic system is highly complex and dynamic system than normal chaotic system where it has more than one positive Lyapunov exponents. Hybrid chaotic system has designed by a disturbed discrete system by another one discrete system. In this algorithm, the input speech signal has compressed by Discrete Cosine Transform (DCT) to reduce the residual intelligibility. The compressed speech signal has permuted by hybrid chaotic system, which has designed using Zaslavsky and Zigzag maps. For substitution process, a reference speech signal has generated by Hidden Markov Model (HMM) speech synthesizer and permuted by using hyper-chaotic system. Masking of encryption signal has done by a masking sequence, which has obtained from the hyper-chaotic system. Our proposed work provides high security for the audio and speech signal over an insecure public network than other traditional speech encryption algorithms based on normal chaotic systems. The betterment of proposed algorithm is proven using the following metrics: key space analysis, key sensitivity analysis, information entropy measure, correlation coefficient analysis, Signal to Noise Ratio (SNR) analysis, subjective evaluation of speech quality, Perceptual Evaluation of Speech Quality (PESQ) analysis, NSCR (Number of Samples Changing Rate) and UACI (Unified Averaged Changed Intensity) analysis have carried out from cryptographic point of view and presented in this paper. The results proof that the proposed speech encryption algorithm ensures appreciable security system with robust encryption and decryption quality.

Keywords Hybrid chaos · Hyper chaotic system · Zaslavsky map · Zigzag map · Permutation · Substitution · Binary masking

✉ P. Sathiyamurthi
sathyamurthi.bit@gmail.com

S. Ramakrishnan
ram_f77@yahoo.com

[1] Department of Information Technology, Dr.Mahalingam College of Engineering and Technology, Coimbatore District, Pollachi, Tamil Nadu, India

1 Introduction

Owing to modern and latest technological enhancement, it is essential to secure the audio and speech signal from eavesdropper over a public wired and wireless network. Desecration may happen in these types of insecure networks for imperative information stealing, dishonest cash, funny activity challenges, terrorism, etc., In order to safeguard the speech data transmitted over a public network, various speech encryption algorithms have been developed such as International Data Encryption Algorithm (IDEA), Advanced Encryption Standard (AES), Data Encryption Standard (DES), Blind Source Separation (BSS) method, Chaotic system based algorithms, etc., [2, 6, 11, 19, 20, 24, 36]. DNA code design based on block quantum algorithm has used in reducing the error in hybridization of large chaotic random number generation. The segmented speech samples have combined and encoded as a DNA's sequences [14, 27]. Quasi groups method is a dominant method wherein this supplements a larger size of samples by changing the amplitudes and this method is appropriate for high volume speech [38]. The sampled speech signal divided into four segments of equal time slot have permuted and substituted using four different chaotic maps [31, 35]. A cyclic elliptic curve (CEC) method and chaotic system generates PRN sequences based on a piecewise nonlinear chaotic map and are mixed with the key sequences derived from CEC points. Asymmetric key cryptography El-Gamal algorithm has implemented for speech signals encryption and decryption [15, 45]. Field Programmable Gate Array (FPGA) hardware realization for all integer, fractional, and mixed-order new Wang chaotic system has proposed to achieve a reasonable reduction in the hardware resources of chaotic systems implementation [8, 18]. Encryption system for speech signals based on circular shifts in row and column using three secret keys has developed [45]. A self-generated true random number based audio encryption system has implemented [41]. To improve the security level for speech signal, X-Tea and RC-6 algorithms have implemented in pipelined manner in terms of memory/area chip, operations used, encryption time etc. [9]. Compared with conventional encryption algorithms, chaotic system based algorithms are highly analogous in various aspects such as robustness of encryption process and decryption quality. Chaotic system based algorithms are in trend for past few decades to achieve secure communication [3, 12, 49]. The main objective of chaos systems in speech algorithms is to generate pseudo-random sequence to scramble the input speech signal [17]. Chaotic systems are more sensitive to its initial conditions and controlling variables. A tiny change in initial conditions and controlling variables makes huge variation in random number generated. These changing parameters have considered as keys in speech encryption algorithms [3, 22, 37].

Chaotic systems are highly preferable for its great permutation and diffusion properties. Some development of research work on chaos based encryption for image, audio, speech and video have discussed in [4, 32, 33, 39, 47, 48]. Substitution-boxes (S-box) based encryption method has constructed using hybrid chaotic maps and Linear Fractional Transform (LFT) to encrypt the necessary parts of sensitive information in Lifting-Wavelet Transform (LWT) frequency domain [13, 23, 25, 29]. Hyper-chaos has proposed for image encryption since hyper-chaos has more than one positive Lyapunov exponents and more complex dynamical behavior than normal chaos. [5, 10, 21, 28, 50] Color image encryption scheme, based on skew tent map and hyper-chaotic system of 6th-order CNN has developed [16, 40] to confuse and diffuse the pixels. Most of the hyper-chaotic systems have effectively designed for image encryption technique. Some researchers have worked on speech encryption using hyper-chaotic system and implemented on hardware [7, 21]. A

Speech scrambling method based on a hybrid chaotic system has developed using Logistic and Henon maps [30].

In this paper, a synchronized hybrid and hyper-chaotic system for speech encryption has introduced. A hybrid chaotic system developed by using Zaslavskey map and Zigzag map have used to generate pseudo random numbers for permuting the speech samples. A five dimensional hyper-chaotic system that is a chaotic oscillator used to produce pseudo random numbers for substituting the scrambled speech samples. A masking sequence has developed from the hyper-chaotic oscillator for masking the encrypted sequence. Existing speech encryption algorithms fail to provide enough security level, lodge extra bandwidth usage and keep noise in the decrypted signal [1, 43, 44, 46]. To overcome these flaws, a novel hybrid-hyper chaotic system and binary masking technique has developed to encrypt the input speech. This speech encryption algorithm provides better security to endure security attacks. The exclusivity of this proposed speech encryption algorithm are:

(i). to scramble the speech signal with high confusion and diffusion
(ii). to protection to cryptanalytic attack
(iii). to reconstruct the input speech with high quality, and
(iv). to minimize the computational complication

The practical applications of the proposed speech encryption algorithm are: Hot links, Military voice communication, Voice over IP, Personal mobile communication and telecommunication services, etc.

The organization of this paper is as follows: Section-2 discusses the proposed hybrid and hyper-chaotic mathematical models, section-3 elaborates the proposed architecture of hybrid-hyper speech encryption technique and masking method of encrypted signal, section-4 presents results and performance analysis, and finally the conclusion is has presented in section-5.

2 Proposed method

The proposed work comprises three levels of encryption schemes. In first level, a hybrid chaotic system has developed by using Zaslavskey and Zigzag maps. The output of the Zigzag map has given to the Zaslavskey as input [46]. The resultant sequence has permuted with the DCT output of the speech signal [20]. In second level, a five dimensional hyper-chaotic system has used to diffuse the output of first level encryption. In this process, a reference speech has used as reference signal to substitution. To generate the reference speech signal a HMM speech synthesizer is used. The hyper-chaotic system has constructed by five variables as five keys. The reference speech has permuted using the hyper-chaotic system and the output has used for substitution with the output of first level encryption. From the 5D hyper-chaotic system, a masking sequence has developed for masking the entire encrypted sequence. As third level of encryption, the masking sequence has used to mask the encrypted sequence.

2.1 Hybrid chaotic system

The hybrid chaotic system has designed using two or more different independent chaotic systems, which may be discrete and/or continuous chaotic system. The hybrid chaotic system increases the complexity in the random number generation and hence becomes difficult to analyze the sequence of random numbers. This is a better approach for a cryptosystem, which has used for secure speech in public network. The advantages of this chaotic system are strong secure key generation, key distribution, authentication, etc. In this paper, we have used two discrete chaotic maps, Zaslavsky map and Zigzag map to achieve better chaotic behavior.

2.1.1 Zaslavsky map

The Zaslavsky Map is a discrete-time nonlinear dynamical system exhibits deterministic dynamic behavior. This map is suitable to build hybrid chaotic map due to its flexible and dynamic characteristics. Equation (1) shows the Zaslavsky map:

$$\begin{aligned} y &= \mathrm{mod}\left(y_{n-1} + v\left(1 + \mu z_{n-1}\right) + \varepsilon v \mu \cos\left(2\pi y_{n-1}\right), 1\right) \\ z &= e^{-r}\left(z_{n-1} + \cos\left(2\pi y_{n-1}\right)\right) \end{aligned} \tag{1}$$

where $\mu = \frac{1-e^{-r}}{r}$; r = 2–5; v = 100–150 and ε = 0–1.

Initial values for random key stream generation are $y_0 = 0.587201561347$ and $z_0 = -0.28432144902$. Equation (1) has iterated for N_s times where N_s is bit stream limit. The pre-processing of generating the key stream is given by eq. (2):

$$x_n = \mathrm{mod}\left(abs\left(\mathrm{int}\left(z_n \times 10^9\right)\right), 1\right) \tag{2}$$

The key stream is normalized by (3):

$$\overline{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{3}$$

where $x_{\min}$ – minimum value of generated key and $x_{\max}$ - minimum value of generated key.

2.1.2 Zigzag map

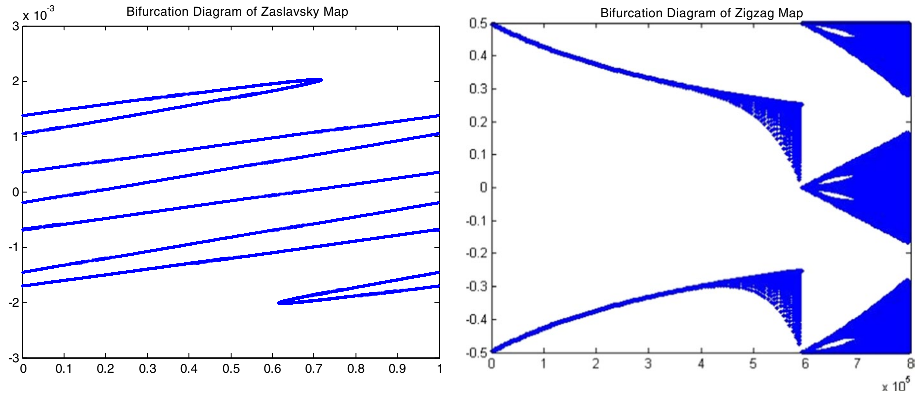
The output values of the zigzag map will alternate between positive and negative values. The absolute value of output of zigzag map is equal to the tent map due to symmetry. The alternation between positive and negative values provides unintelligibility and is helpful in being there execution of non-idealities. The generalized zigzag map has given in the eq. (4):

$$x_{n+1} = \begin{cases} -m\left(x_n + \frac{2}{|m|}\right), -1 < x_n \leq -\frac{1}{|m|} \\ mx_{n,} - \frac{1}{|m|} < x_n \leq \frac{1}{|m|} \\ -m\left(x_n - \frac{2}{|m|}\right), -\frac{1}{|m|} < x_n \leq 1 \end{cases} \tag{4}$$

where m is a real number and m ϵ (−3, 3). In this eq. (5), m = −2 represents the zigzag map. Generating of bits from this map is simple. $|x_n| < 1/2$ represents 0 and $|xn| > 1/2$ represents 1. Without loss of generality, the initial state of the system has been assumed as a slightly positive number, i.e., ×0 = 0+. Critical point falls at m = 3 that provides larger value. This case is at the average of instability, since the system has driven out of the map. In $2 < m < 3$, the system is chaotic and stable. Chaos can be observed for $1 < m < 2$. Since the initial state of the system has assumed as positive, the output value has confined to positive values. If the system gets out of the positive region and goes into the negative region, it will trap in the negative region. For $|m| < 1$, no chaotic behavior has observed. The output of the system will ultimately settle in zero. $-2 < m < -1$ is also a chaotic region for the system. The output in this region alternates between positive and negative values. There is no asymptotic density distribution when the output values are in the positive and negative regions for odd and even time steps respectively. In other words, limfn(x) does not exist as $n \to \infty$. The bifurcation diagram of the Zaslavzky and Zigzag maps have given in Figure 1. These diagrams show the random behavior of the Zaslavzky and Zigzag chaotic maps.

2.2 Hyper chaotic system

Chaos is a highly used system to generate pseudo random numbers for various deterministic non-linear applications. Chaotic systems provide excessive sensitivity to initial conditions and have random like behaviors. Hyper-chaos is a unique chaotic system which has multi direction instability when chaos makes instability in only one direction. Lyapunov exponent plays an effective role in the chaotic property of nonlinear dynamic system. If any one Lyapunov exponent in a nonlinear dynamic system is non-zero and positive then the system becomes a chaotic behavior. Normally the dimension of the nonlinear dynamic system equals to the number of Lyapunov exponent and at least. For example, a 3-dimensional nonlinear dynamic system will have three or more Lyapunov exponent. If three or more Lyapunov exponents are positive and non-zero then the nonlinear dynamic system will produce higher instability. Higher dimensional hyperchaotic system has higher key space, more complex in randomness and more sensitive. This behavior of higher dimension hyper chaos leads to research in cryptographic applications.



(a)

(b)

Fig. 1 Bifurcation diagram chaotic maps. (**a**) Zaslavzky Map, (**b**) Zigzag Map

In this paper, 5-dimensional hyper-chaotic system has used for pseudo random number generation to substitute the speech signal. The 5-D chaotic system is presented in eq. (5):

$$\left.\begin{aligned} x_1' &= -ax_1 + x_2x_3 \\ x_2' &= -bx_2 + fx_5 \\ x_3' &= -cx_3 + gx_4 + x_1x_2 \\ x_4' &= dx_4 - hx_1 \\ x_5' &= ex_5 - x_2x_1^2 \end{aligned}\right\} \tag{5}$$

Where x_1 to x_5 are state variables and a to h are constant parameters. Nonlinear terms are exist in x_1', x_3', and x_5' of this dynamical system (1) i.e. x_1x_2, x_2x_3, and $x_2x_1^2$. We have chosen the initial value for all the five state variables from *[0, 0,0, 0, 0]* to [1] and constant parameters as *$a = 4.3$ to 20, $b = 10.5$ to 50, $c = 10$ to 30, $d = 0$ to 2, $e = 15$ to 25, $f = 20$ to 40* and *$h = 1$ to 10.*

Different dynamical periodic orbit has obtained by varying the constant parameters in a range in addition to hyper-chaotic. The variety of Lyapunov exponents of the system have obtained from the rage of different constant parameters. For simple and efficient mechanism, the constant parameter 'b' has chosen randomly. The constant parameter 'b' rage has taken between [10.5, 50], there are two positive integers have appeared so the system is hyper-chaotic. The bifurcation diagram of the hyper-chaotic and hybrid chaotic maps are illustrated in Figure 2. These diagrams show the random behavior of the hyper-chaotic and hybrid chaotic maps.

3 Proposed encryption and decryption algorithm

The Figure 3 shows the architecture of hybrid-hyper speech encryption technique based on permutation and substitution processes and masking of encrypted sequence developed from hyper-chaotic system.

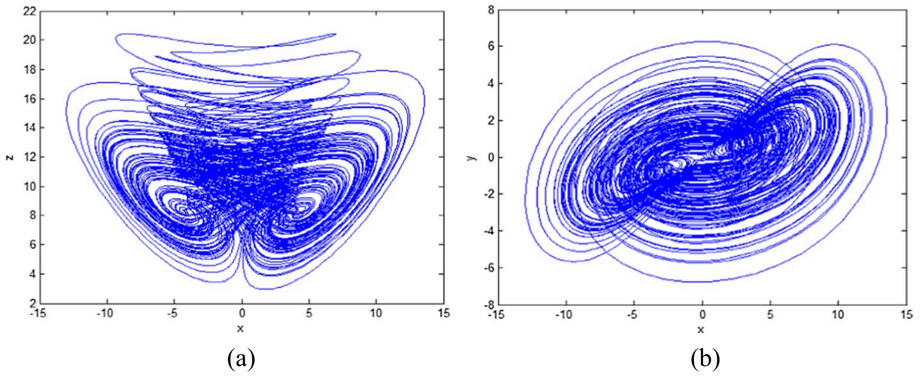


Fig. 2 Bifurcation diagram of hyper chaotic Map and Hybrid Chaotic Map. (**a**) Hyper Chaotic Map, (**b**) Hybrid Chaotic Map

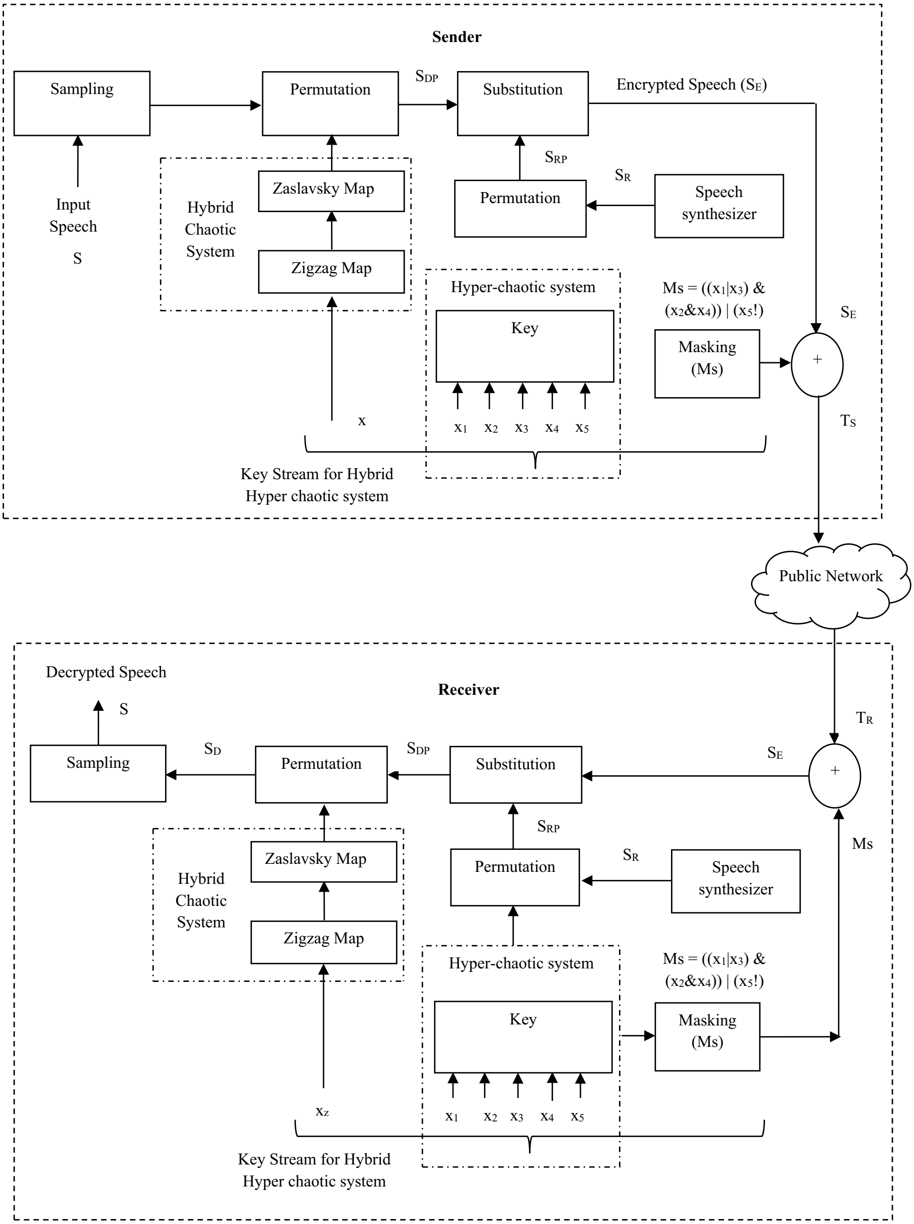


Fig. 3 Block diagram of proposed system

3.1 Encryption algorithm

Step-1: The input speech signal has compressed by DCT to make unintelligibility. Input sequence (S_D) is permuted with the random numbers (Z_k) generated by hybrid chaotic system where the hybrid chaotic system is designed by Zaslavskey and zigzag maps. x_0 of the Zigzag map in equation-4 is given as input to the initial state of the Zaslavskey map in equation-1. Input speech samples have permuted with the random number sequence generated by Zaslavsky map Z_k:

$$S_{DP} = S_D \oplus Z_k \tag{6}$$

where S_{DP} and S_D are encrypted and input speech signal at first level

Step-2: In second level of encryption, a reference speech signal which is generated by HMM speech synthesizer has been used as reference for substitution process. The reference speech have permuted by using the hyper-chaotic system and used for substitution process. Thus the first level encrypted sequence S_{DP} is substituted by the binary sequence generated by the 5-D hyper-chaotic system S_{RP} represented in equation-5 and obtained S_E. Here the output of hyper-chaotic system have permuted using a reference speech S_R that is generated by a speech synthesizer.

Step-3: A masking sequence has developed for masking the encrypted signal by applying three combinations of different logical operation:

$$\mathrm{Ms} = ((\mathrm{x1}|\mathrm{x3})\&(\mathrm{x2}\&\mathrm{x4}))\ |\ (\mathrm{x5}!) \tag{7}$$

Step-4: In third level of encryption, the masking sequence Ms. is *XOR*ed with second level encryption sequence S_E. Finally the encrypted signal T_S is transmitted over the public channel.

3.2 Decryption algorithm

The reverse process of encryption method is the decryption method. Here the original speech signal has recovered without any loss of information.

4 Experimental results and analysis

The performance of proposed crypto algorithm has been measured and analyzed by key space analysis, key sensitivity analysis, correlation coefficient analysis, SNR, Peak Signal to Noise Ratio (PSNR), segmental SNR (SNRseg), frequency-weighed segmental SNR (fwSNRseg) analysis, computational complexity measure, subjective evaluation of speech quality, PESQ analysis, NSCR and UACI analysis. The proposed system has implemented and tested in Matlab. 10 sample speech signals are taken randomly from TIMIT database and are sampled at 8 kHz with length of 3 Sec to 8 Sec and 8000 samples per frame. R2013a in Intel Core-i3 CPU @ 3.30GHZ speed, 4GB RAM, 64bit operating system. The acceptable range of different SNR metrics as per NIST statistical test suite [33] are listed in the Table 1.

4.1 Histogram analysis

Histogram is a pictorial representation (Figure 4) consisting of spikes whose area is proportional to the frequency of a variable and whose width is equal to the class interval. Histogram(X) creates a histogram plot of X. The histogram function uses automatic

Table 1 Acceptable range of SNR as per NIST statistical test suite

S.No	Performance Metrics	Range Level
1	SNR in dB [26]	80 dB and above for chaotic noise content
2	PSNR in dB [2]	20 dB and above for chaotic noise content
3	SNRseg in dB [5]	80 dB and above for chaotic noise content
4	fwSNRseg in dB [5]	400 dB and above for chaotic noise content
5	Correlation Coefficient	Close to 1 for original Vs decrypted signal Close to 0 for original Vs encrypted signal
6	NSCR	99.99
7	UACI	33.33

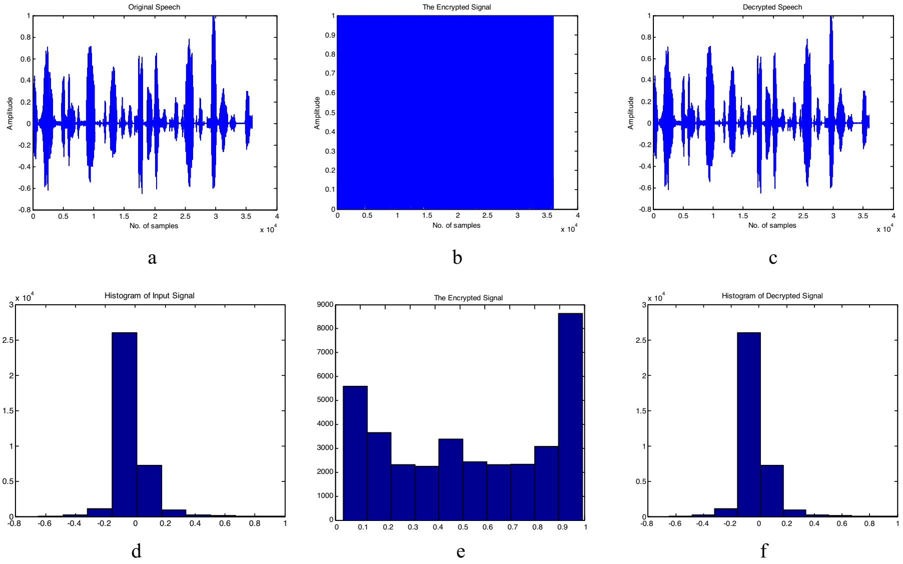


Fig. 4 Plot and histogram of original, encrypted and decrypted speech. **a**. Original Speech, **b**. Encrypted Speech, **c**. Decrypted Speech, **d**. Histogram of Original Speech, **e**. Histogram of Encrypted Speech, **f**. Histogram of Decrypted Speech

binning algorithm that returns the bin with uniform width, chosen to cover the range of elements in X. In a more general mathematical sense, a histogram is a function mi that counts the number of observations that fall into each of the disjoint categories known as bins.

$$n = \sum_{l=1}^{k} m_i \tag{8}$$

where, n is the total number of observations and k be the total number of bins and mi the histogram.

Variance of histogram is analysed as follows:

Table 2 Histogram variance analysis

Speech file (.wav)	Key −1 (%)	x_1(%)	x_2(%)	x_3(%)	x_4(%)	x_5(%)	x_6(%)
Sample-1	11.85	2.23	1.98	4.19	5.58	0.95	2.71
Sample-2	10.92	2.39	1.52	2.94	3.65	1.54	1.93
Sample-3	12.64	2.37	1.29	3.46	4.80	0.99	1.84
Sample-4	18.37	3.45	2.48	4.67	3.72	1.11	2.45
Sample-5	10.51	2.56	1.93	4.48	4.53	1.94	3.09
Sample-6	11.81	3.12	1.03	2.32	3.94	0.92	2.84
Sample-7	10.93	3.87	2.33	3.78	4.76	1.75	1.98
Sample-8	12.64	2.78	1.06	2.65	2.61	1.63	2.31
Sample-9	13.91	3.72	1.96	2.55	2.49	1.34	3.11
Sample-10	14.26	4.07	1.73	4.62	3.73	1.87	1.96

$$Var(x) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2}\left(x_i - x_j\right)^2 \tag{9}$$

where x is the vector of histogram values and x_i and xj are the number of samples equal to i and j respectively. This quantity analysis of each key employ variances of histograms of encrypted speech signals to evaluate the uniformity. If the variance is low means, uniformity of encrypted speech is high and vice versa. Two variances of encrypted speech have calculated by two different secret keys. By varying the secret keys if we have obtained lower uniformity of encrypted speech indicates strong encryption.

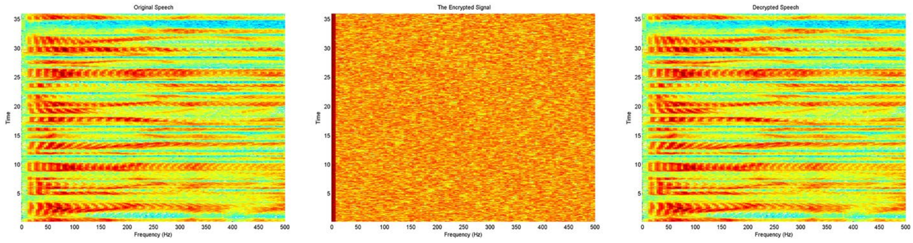
In the Table 2, the variances obtained by two different secret keys x_1 and k and the variances obtained by modifying one parameter of k, x_1, x_2, x_3, x_4, x_5, and x_5. By changing any of the secret key, 10 to 20% fluctuations have obtained. If all the secret keys have changed, then 80–100% fluctuations is possible. It shows the proposed algorithm has good key sensitivity.

4.2 Spectrogram analysis

The spectrograms of the original, encrypted and decrypted speech signals have presented in Figure 5. The spectrograms of the original and encrypted speech signals show that these two signals are completely different with higher encryption quality. The spectrograms of the original and decrypted speech signals are identical to each other. Thus higher decryption quality has proved and this hold good for all the spectrogram analyses.

4.3 Correlation analysis

The auto-correlation function can identify the chaotic system that produces a strong encryption. A useful measure to assess the encryption quality of any cryptosystem is the correlation coefficient between similar segments in the clear signal and the cipher signal. It is calculated as:



(a) (b) (c)

Fig. 5 Spectrogram of original, encrypted and decrypted speech. (**a**). Spectrogram of Original Speech, (**b**). Spectrogram of Encrypted Speech, (**c**). Spectrogram of Decrypted Speech

$$\mathrm{r_{xk}} = \frac{\mathrm{C(x, k)}}{\sqrt{\mathrm{V(x)}}\sqrt{\mathrm{V(k)}}} \tag{10}$$

where C(x, k) is the covariance between the original signal x and the encrypted signal k. V(x) and V(k) are the variances of the signals x and k. The variance V(x) is computed as:

$$V(x) = \frac{1}{N_s}\sum_{i=1}^{N_s}(x(i) - E(x))^2 \tag{11}$$

$$E(x) = \frac{1}{N_s}\sum_{i=1}^{N_s}\Big(x(i) \tag{12}$$

$$C(x, k) = \frac{1}{N_s}\sum_{i=1}^{N_s}(x(i) - E(x))(k(i) - E(k)) \tag{13}$$

where $\mathrm{N_s}$ is the number of speech samples. The low value of correlation coefficient $\mathrm{r_{xk}}$ shows a good encryption quality. The correlation coefficients for the three different encrypted speech samples with the chaotic maps illustrated in Fig. 1 and the encrypted speech signals with proposed method using five different chaotic random number generator are tabulated in Table 2.

It has observed that from these results, the proposed algorithm produces encrypted speech with low correlation between similar segments in the original speech and the encrypted speech, which means that the encryption method gives good encryption results. In this proposed method it is obtained that the correlation co-efficient as 0.998 shows the original speech signal has been permuted almost 100% in decryption process so it is tough to the eavesdroppers to hack the speech signal in channel during transmission.

4.4 SNR analysis

Signal to Noise Ratio (SNR) test is a good estimator for measuring the speech signal intelligibility. The popular time domain metric is the SNR, which has defined as the average of the SNR values of short segments of the output signal as calculated below eq. (14):

$$SNR = 10\log_{10}\frac{\sum_{i=1}^{N_s} x^2(i)}{\sum_{i=1}^{N_s} (x(i) - y(i))^2} \tag{14}$$

where y(i) is decrypted speech signal. If the SNR is closer to zero, the higher is the quality of the decrypted signal.

4.5 Peak signal to noise ratio (PSNR) analysis

Peak Signal to Noise Ratio (PSNR) is ratio between the maximum possible power of original speech signal and the power of encrypted signal [2]. PSNR is a calculation of encryption quality of the original signal [17]. Higher PSNR indicates that the encryption or reconstruction is of higher quality. The PSNR is obtained from eq. (15):

$$PSNR = 10\log\frac{nx^2}{\|x - k\|^2} \tag{15}$$

The quality of decrypted signal of this algorithm is measured using SNR calculation as illustrated in Table 3. The average value of CC, SNR, PSNR, SNRseg and fwSNRseg for the ten decrypted speeches are 0.99299, 122.49 dB, 49.489 dB, 122.905 dB and 55.186 dB respectively with a difference of 0.0087, 2.72 dB, 3.42 dB, 2.72 dB and 2.34 dB. This algorithm brings about good range of SNR value and good quality of decrypted speech.

Table 3 Performance analysis of decryption process

S.No	Speech sample	SNR in dB	PSNR in dB	SNRseg in dB	fwSNRseg in dB	Correlation Coefficient
1.	Sample-1	123.57	50.21	121.25	55.46	0.9875
2.	Sample-2	121.30	48.25	122.92	53.98	0.9942
3.	Sample-3	122.67	50.47	123.50	55.98	0.9934
4.	Sample-4	123.86	51.67	121.58	56.01	0.9929
5.	Sample-5	122.41	49.69	124.15	55.23	0.9961
6.	Sample-6	121.14	50.33	123.24	54.58	0.9962
7.	Sample-7	122.98	49.68	122.98	55.11	0.9895
8.	Sample-8	121.67	48.99	123.52	54.98	0.9942
9.	Sample-9	122.21	49.20	122.84	55.01	0.9941
10.	Sample-10	123.08	50.28	123.07	56.32	0.9918
Average		122.489	49.877	122.905	55.186	0.99299
Difference		2.72	3.42	2.72	2.34	0.0087

Table 4 Robustness Test NSCR and UACI analysis

S.No	Method	Test	Audio files				
			Sample-1	Sample-2	Sample-3	Sample-4	Sample-5
1	Juliano et al. [23]	NSCR in %	99.979	99.912	99.922	99.954	99.931
		UACI in %	33.012	33.031	33.124	33.301	33.219
2	Madain A et al. [25]	NSCR in %	99.925	99.903	99.911	99.936	99.920
		UACI in %	33.049	33.037	33.125	33.271	33.189
3	Peng X et al. [29]	NSCR in %	99.931	99.906	99.929	99.948	99.918
		UACI in %	33.021	33.065	33.099	33.245	33.159
4	P.Sathiyamurthi et al. [33]	NSCR in %	99.979	99.921	99.929	99.961	99.942
		UACI in %	33.048	33.055	33.098	33.298	33.225
5	Proposed Method	NSCR in %	99.981	99.924	99.937	99.962	99.943
		UACI in %	33.059	33.079	33.149	33.318	33.242

Table 5 Key Sensitivity analysis

Chaotic System	Key value	Changed key value	% of difference
Hybrid Chaotic System	x0 = −2	×0 = −2.00001	99.67
Hyper Chaotic System	x1 = 15	×1 = 15.00001	99.48
	×2 = 133.33	x2 = 133.32	99.25
	×3 = 1	x3 = 0.999999	99.37
	x4 = 25	×4 = 25.00001	99.71
	x5 = 40	×5 = 40.00001	99.69

4.6 Robustness test

The robustness test has made for different input signals and the comparison for the number of samples change rate (NSCR) and the unified average changing intensity (UACI) in percentage have examined. The ciphered speech signals have compared by NSCR and UACI [39]. The NSCR is specified by,

$$NSCR = \sum_i \frac{d_i}{l} \times 100\% \tag{16}$$

and

$$UACI = \frac{1}{l}\left[\sum_i \frac{|x_i - x_i'|}{65535}\right] \tag{17}$$

where x and x' are the two ciphered speech signals whose corresponding original signal signals have only one-sample difference; the values of the samples at position i of x and x' are respectively denoted by xi and xi'; l corresponds to the length of the speech vector. The ideal values for NSCR and UACI are 100% and 33.3%. In Table 3,

Table 6 Average and inappropriate correlation coefficient for representative and chosen 100 samples randomly

Data set	Correlation coefficient	PESQ
TIMIT Database	Average	0.971
	Inappropriate	0.923
Recorded speech in lab	Average	0.987
	Inappropriate	0.889

Table 7 Subjective evaluation of encrypted and decrypted speech

Human Listener category	Number of Listener	Number of samples	Average value (in the scale from 0 to 3)	
			Quality of decrypted speech	Quality of encrypted speech
Male less than 20 years old	15	10	3	0
Girls than 20 years old	15	10	3	0
Male staff age between 20 to 40	15	10	3	0
Female staff age between 20 to 40	15	10	3	0.2
Male staff age between 40 to 60	15	10	2.8	0.25
Female staff age between 40 to 60	15	10	2.8	0.2
Scale	Scale for quality of Decrypted / Encrypted speech			
0	Too noisy			
1	Noisy			
2	Speech with noise			
3	Clear speech			

the minimum, maximum and average values of NSCR and UACI, computed from the encryption of four different modified versions of each speech signal are given.

The results are considerably close to the ideal values and independent on the position of the modified sample as compared with the ideal values for NSCR and UACI for robust test are 100% and 33.3% and from the equation ($_{16}$) and ($_{17}$), the performance analysis for robustness test has made. From Table 4, comparison between existing methods and proposed method, depict that the resultant values are significantly nearer.

4.7 Key sensitivity analysis

In this key sensitivity analysis, we have changed the key values for one bit. It demonstrates a dramatically cause of decryption error at the receiver side. The Table 5 clearly proves that a small change in each key provides more than 99% difference in key space. Thus it provides more complex to larger space [33].

4.8 Perceptual evaluation of speech quality (PESQ) analysis

The PESQ analysis is carried out and results are furnished in Table 6. As many as 50 speech samples in different varity from TIMIT database and 50 recorded speech samples in different environments (Total: 100 samples) are taken to find out the average correlation coefficient [33].

4.9 Subjective evaluation

The quality of encrypted and decrypted speech have exposed to suitable evaluation by the known sources in different age population. They agreed that the decrypted speech is same as the original speech and the encrypted speech is noisy. The subjective evaluation [33] of encrypted and decrypted speech listed in Table 7.

4.10 Types of classical attacks

Generally the cryptanalyst knows the whole thing about the cryptosystem other than the secret key and intending to attack the cryptosystem. In cryptanalyzing there are four types of classical attacks: (a) Cipher text only, (b) Known plain text, (c) Chosen plain text and (d) Chosen cipher text. Out of which, chosen plain text attack is the most dominant attack. If a crypto system can withstand this attack, then it can withstand all other attacks.

The proposed method is sensitive to the system parameters and initial key parameters of hybrid chaotic system developed by using Zaslavskey map and Zigzag map ×0, y0, z0 and m. If any one of the key values has changed, the entire output would be dramatically changed. Hence, the proposed method can withstand the chosen plain text attack. The

Table 8 Overall Performance Comparison

S.No	Performance metrics	Exiting methods				Proposed method
		Juliano et al. [23]	Madain A et al. [25]	Peng X et al. [29]	P.Sathyamurthi et al. [33]	
1.	Correlation coefficient (Original Vs Encrypted)	0.0511	0.0686	0.1004	0.0192	0.0104
2.	Correlation coefficient (Original Vs Decrypted)	0.9747	0.9892	0.9806	0.9955	0.9981
3.	SNR in dB	121.19	122.24	121.86	123.02	123.89
4.	PSNR in dB	48.64	49.35	48.98	50.16	51.33
5.	SNRseg in dB	119.45	120.77	119.52	121.20	122.83
6.	fwSNRseg in dB	52.82	54.45	53.08	55.32	55.99
7.	NSCR	98.986	99.025	98.991	99.912	99.994
8.	UACI	32.890	33.115	32.986	33.333	33.338

experimental results proof that the proposed method provides a higher level of security to the speech communication models [42].

4.11 Overall performance comparison

The performance of the proposed speech encryption algorithm has compared with four existing methods. One common speech sample has taken randomly from TIMIT dataset, which has already used in result analysis study. The performance metrics used for this performance comparison are correlation coefficient, SNR, NSCR and UACI. This comparison presented in Table 8 shows that the proposed method provides better results compared with existing methods.

Advanced Encryption Standard (AES) is a digital cryptosystem in which high degree of security can be attained but it is not often used in existing speech communication systems owing to the bandwidth expansion of the encrypted speech and the degradation of the Signal-to-Noise Ratio (SNR) performance. The original information can be easily detected by eavesdroppers from the residual intelligibility such as signal energy, talk spouts, and the original pitch which are present in the encrypted speech. This can degrade the security of the speech cryptosystem. The chaotic based speech scrambler is the best choice to remove the dashes of residual intelligibility in the encrypted speech and to withstand the frequency domain attacks. Chaotic systems have mixing, stretching, folding and sensitivity to key parameter properties. These confusion and diffusion properties make the chaotic systems a worthy choice for constructing a good cryptosystem. [26, 34].

5 Conclusions

In the proposed system, a complex algorithm for secured speech communication is proposed. The proposed algorithm comprising the two chaotic systems: one is hybrid chaotic system built by Zaslavskey and zigzag maps and another one is five-dimensional hyperchaotic system in which a masking sequence is used for masking the encrypted signal and increases the complexity. Since from the result analysis, it, demonstrates large key space, good correlation, NSCR and UACI. Performance analysis reflects that the proposed system is highly complex, secure, and negligibly prone to noise and it can be applied in the real time communication. The system can ensure voice security and it is the finest mechanism. This work covers way for lots of openings, wherein this can be implemented in hardware.

Authorship responsibility This manuscript represents valid work; neither this manuscript nor one with substantially similar content under our authorship has been published or is being considered for publication elsewhere. No copies of any closely related manuscripts are enclosed in the manuscript submission. We confirm that this work is an accurate representation of the trial results. We agree to allow the corresponding author to serve as the primary correspondent with the editorial office and to review and sign off on the final proofs prior to publication. We have participated sufficiently in the work to take public responsibility for the entire content of the manuscript.

The authors are keen and committed for intellectual sharing to extend his frontiers of knowledge in his field of study. Moreover, the carries an unambiguous research focus and maintains that research contributions should serve for the upliftment of the humankind in one-way or other. More importantly, the author is willing to share his knowhow, not for fiscal viewpoint but for the holistic development in his select field of study.

Funding The authors are desirous of publishing research article in the referred journals of repute. This research work is not funded any organization and not supported by any finding schemes.

Declarations

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Intellectual property rights This research work is not under any intellectual property rights.

Research involving human participants and/or animals Though this research work is implemented and evaluated in computer software, this article does not contain any studies with human participants or animals performed by any of the authors for any type of activities such as testing, analysis, etc.,

Conflict of interest The authors declare that there is no actual or potential conflict of interest in relation to this article.

References

1. Abdul DS, Elminaam HM, Kader A, Hadhoud MM (2008) Performance evaluation of symmetric encryption algorithms. IJCSNS Int J Comput Sci Netw Security 8(12):280–286
2. Advanced Encryption Standard (2001) Federal Information Processing Standards Publication, NIST FIPS; 197, Commerce Department, National Institute of Standards and Technology (NIST)
3. Alvarez G, Li S (2006) Some basic cryptographic requirements for chaos-based cryptosystems. Int J Bifurc Chaos 16(8):2129–2151
4. Biswas D, Banerjee T (2016) A simple chaotic and hyper-chaotic time-delay system: designand electronic circuit implementation. Nonlinear Dyn 83:2331–2347
5. Biswas D, Karmakar B, Banerjee T (2017) A hyper chaotic time-delayed system with single-humped nonlinearity: theory and experiment. Nonlinear Dyn 89:1733–1743
6. Daemen J, Rijndael VR (2001) The advanced encryption standard. Doctor Dobb's J 26(3):137–139
7. Elkholy MM, EL Hennawy HM, Elkouny A (2016) Real time implementation of secure communication system based on synchronization of hyper chaotic systems. 33rd National Radio Science Conference
8. Elsafty AH, Tolba MF, Said LA, Madiana AH, Radwan AG (2020) Enhanced hardware implementation of a mixed-order nonlinear chaotic system and speech encryption application. AEU Int J Electron Commun 125. https://doi.org/10.1016/j.aeue.2020.153347
9. Etem T, Kaya T (2020) Self-generated encryption model of acoustics. Appl Acoust 170. https://doi.org/10.1016/j.apacoust.2020.107481
10. Farsana FJ, Gopakumar K (2016) Speech encryption based on four-dimensional hyper chaotic system. Int Conf Data Min Adv Comput. https://doi.org/10.1109/SAPIENCE.2016.7684153
11. Forouzan BA (2010) Cryptography and network security. McGraw Hill Publications
12. Fu C, Lin BB, Miao YS, Liu X, Chen JJ (2011) A novel chaos-based bit-level permutation scheme for digital image encryption. Opt Commun 284(23):5415–5423
13. Gopalakrishnan T, Ramakrishnan S (2017) Chaotic image encryption, with hash keying as key generator. IETE J Res 63(2):172–187
14. Guo Q, Wang B, Zhou C, Wei X, Zhang Q (2017) DNA Code Design based on Block Quantum Algorithm. IEEE Access 5:22453–22461
15. Imran OA, Yousif SF, Hameed IS, Abed WNA-D, Hammid AT (2020) Implementation of El-Gamal algorithm for speech signals encryption and decryption. Procedia Comput Sci 167:1028–1037
16. Kadira A, Hamdulla A, Guo W-Q (2014) Color image encryption using skew tent map and hyper chaotic system of 6th-order CNN. Elsevier – Optik 125:1671–1675
17. Kanso NS (2009) Logistic chaotic maps for binary numbers generations. Chaos, Solitons Fractals 40(5):2557–2568
18. Kohli R, Kumar M (2013) FPGA implementation of cryptographic algorithms using multi-encryption technique. Int J Adv Res Comput Sci Softw Eng 3:112–120
19. Kumar LP, Gupta AK (2016) Implementation of speech encryption and decryption using advanced encryption standard. IEEE International Conference on Recent Trends in Electronics Information

Communication Technology May 2016 Bangalore, India. https://doi.org/10.1109/RTEICT.2016.7808081

20. Leng L, Zhang J, Khan MK, Chen X, Alghathbar K (2010) Dynamic weighted discrimination power analysis: a novel approach for face and palmprint recognition in DCT domain. Int J Phys Sci 5(17):2543–2554
21. Li Y, Wang C, Chen H (2017) A hyper-chaos-based image encryption algorithm using pixel-level permutation and bit-level permutation. Opt Lasers Eng 90:238–246
22. Lian SG, Sun J, Wang Z (2005) A block cipher based on a suitable use of the chaotic standard map. Chaos, Solitons Fractals 26(1):117–129
23. Lima JB, da Silva Neto EF (2016) Audio encryption based on the cosine number transform. Springer Multimedia Tool Appl 75:8403–8848
24. Lin QH, Yin FL, Mei TM, Liang H (2006) A blind source separation based method for speech encryption. IEEE Trans Circ Syst-I 53(6):1320–1328
25. Madain A, Abu Dalhoum AL, Hiary H, Ortega A, Alfonseca M (2014) Audio scrambling technique based on cellular automata. Multimed Tools Appl 71(3):1803–1822
26. Mosa E, Messiha NW, Zahran O, Abd El-Samie FE (2011) Chaotic encryption of speech signals. Int J Speech Technol 14:285–296
27. Nagakrishnan R, Revathi A (2019) A robust speech encryption system based on DNA addition and chaotic maps. Intell Syst Des Appl 940:1070–1080
28. Norouzi B, Mirzakuchaki S, Seyedzadeh SM, Mosavi MR (2014) A simple, sensitive and secure image encryption algorithm based on hyper-chaotic system with only one round diffusion process. Multimed Tools Appl 71:1469–1497
29. Peng X, Cui Z, Cai L, Yu L (2003) Digital audio signal encryption with a virtual optics scheme. Optik Int J Light Electron Opt 114(2):69–75
30. Sadkhan SB, Ali H (2016) A proposed speech scrambling based on hybrid chaotic key generators. Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications
31. Sathiyamurthi P, Ramakrishnan S (2017) Speech encryption using chaotic shift keying for secured speech communication. Eurasip J Audio Speech Music Process Springer 20(2017)
32. Sathiyamurthi P, Ramakrishnan S (2019) Testing and analysis of Chen chaotic mapping for speech cryptography. J Test Eval 47(4):3028–3040
33. Sathiyamurthi P, Ramakrishnan S (2020) Speech encryption algorithm using FFT and 3D-Lorenz–logistic chaotic map. Multimed Tools Appl 79:17817–17835
34. Sathiyamurthi P, Mownesh SS, Suriyavardhan G, Nagaraja S (2021) Speech encryption using Rossler attractor and gingerbread man chaotic maps. J Chengdu Univ Technol 26(8):1–7
35. Slimania D, Merazka F (2018) Encryption of speech signal with multiple secret keys. Int Conf Nat Lang Speech Process Procedia Comput Sci 128:79–88
36. Stallings W (2006) Cryptography and network security principles and practice, 4th edn. Pearson Publications
37. Vaidya PG, He R (1998) Implementation of chaotic cryptography with chaotic synchronization. Phys Rev Lett 57(2):1532–1535
38. Venkat M, Satti K (2006) A quasi group based cryptographic system. Int J Netw Secur 7(1):15–24
39. Wang X, Wang M (2008) A hyper chaos generated from Lorenz system. Phys A-Stat Mech Appl 387(14):3751–3758
40. Wang XY, Xu B, Zhang HG (2010) A multi-array number communication system based on hyper chaotic system of 6th-order cellular neural network. Commun Nonlinear Sci Numer Simul 15(1):124–133
41. Wang X, Zhao J, Liu H (2012) A new image encryption algorithm based on chaos. Opt Commun 285:562–566
42. Wang X, Teng L, Qin X (2012) A novel colour image encryption algorithm based on chaos. Elsevier Signal Process 92:1101–1108
43. Xianyong W, Guan Z-H (2007) A novel digital watermark algorithm based on chaotic maps. Phys Lett A 365:403–406
44. Yang T, Chua LO (1997) Impulsive stabilization for control and synchronization of chaotic systems: theory and application to secure communication. IEEE Trans Circ Syst I 44(10):976–988
45. Ying-Qian Z, Xing-Yuan W (2014) A symmetric image encryption algorithm based on mixed linear – nonlinear coupled map lattice. Elsevier Inform Sci 273:329–351
46. Zaslavskii GM (1978) The simplest case of a strange attractor. Phys Lett A 69(3):145–147
47. Zhang GJ, Liu Q (2011) A novel image encryption method based on total shuffling scheme. Opt Commun 284(12):2775–2780

48. Zhang YQ, Wang X (2015) A new image encryption algorithm based on non-adjacent coupled map lattices. Appl Soft Comput 26:10–20
49. Zheng F, Tian XJ, Song JY (2008) Pseudo-random sequence generator based on the generalized Henon map. J China Univ Posts Telecommun 15(3):64–68
50. Zhu H, Zhao C, Zhang X (2013) A novel image encryption–compression scheme using hyper-chaos and Chinese remainder theorem. Elsevier Signal Process Image Commun 28:670–680

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

P. Sathiyamurthi received the B.E degree in Electronics and Communication Engineering in 2006 and the M.E. degree in Communication Systems in 2008 from the Anna University, Chennai. He has completed PhD degree in Information and Communication Engineering in Anna University, Chennai. He has 12 years of teaching and research experience. He is an Assistant Professor (Senior Scale) in the Department of Information Technology, Dr. Mahalingam College of Engineering and Technology, Pollachi, India. His areas of research include audio and speech signal processing, speech cryptography and internet of things. He has published 23 papers in international, national journals and conference proceedings.

S. Ramakrishnan received the B.E. degree in Electronics and Communication Engineering in 1998 from the Bharathidasan University, Trichy, and the M.E. degree in Communication Systems in 2000 from the Madurai Kamaraj University, Madurai. He received his PhD degree in Information and Communication Engineering from Anna University, Chennai in 2007. He has 19 years of teaching experience and 1 year industry experience. He is a Professor and the Head of the Department of Information Technology, Dr. Mahalingam College of Engineering and Technology, Pollachi, India. Dr. Ramakrishnan is an Associate Editor for IEEE Access and Reviewer of 24 International Journals such as IEEE Transactions on Image Processing, IET Journals(Formally IEE), ACM Computing Reviews, Elsevier Science, International Journal of Vibration and Control, IET Generation, Transmission & Distribution, etc. He is in the editorial board of 7 International Journals. He is a Guest Editor of special issues in 3 International Journals including Telecommunication Systems Journal of Springer. He has published 160 papers in international, national journals and conference proceedings. Dr.S.Ramakrishnan has published a book on Wireless Sensor Networks for CRC Press,USA and two Books on Speech Processing for InTech Publisher, Croatia and a book on Pattern Recognition for InTech Publisher, Croatia and a book on Face Recognition for InTech Publisher, Croatia and a book Computational Techniques for Lambert Academic Publishing, Germany and a book on Modern Fuzzy Control Systems for InTech Publisher, Croatia. He has also reviewed 3 books for McGraw Hill International Edition and 15 books for ACM Computing Reviews. He has guided 9 PhD scholars. His biography has been included in Marquis Whos's Who in the World 2012 & 2016 edition. His areas of research include digital image processing, soft computing, human-computer interaction, wireless sensor network and cognitive radio.