Transactions on Machine Intelligence 4(2) (2021) 70-75

ISSN Online: 2821-1693

Transactions on Machine Intelligence

Journal Homepage: https://tmachineintelligence.ir/index.php/TMI

# Speech Signal Encryption in Time and Frequency Domains Using Chaotic Maps

J. Shirazi[1,*]

[1] Assistant Professor, Department of Electrical Engineering, Gonabad Branch, Islamic Azad University, Razavi Khorasan, Gonabad, Iran

**ARTICLE INFO**

Article History:
Received 24 February 2021
Received in revised form 4 April 2021
Accepted 6 June 2021
Available online 7 June 2021

Keywords:
Steganography, Cryptography, Scrambling, Chaotic Map, Discrete Cosine Transform

**ABSTRACT**

In this paper, with the aim of enhancing the security of speech signal transmission, a method based on encrypting speech signal frames in both time and frequency domains using a chaotic map is presented. To increase the complexity of the encryption, two stages of encryption in the time domain and one stage in the frequency domain were performed on the speech signal. In the first stage, the bytes of the speech signal were altered using random numbers generated from a chaotic map, and in the second stage, the resulting samples were scrambled based on random numbers. Discrete Cosine Transform (DCT) was then applied to the resulting samples, followed by scrambling based on random numbers. To further enhance encryption security, the sequence of random numbers used was periodically altered according to a specific order to increase the complexity of the encryption detection. Various criteria were used to evaluate the employed method, and the results indicate high levels of these criteria for the method. One advantage of the used method is its complexity due to combining time and frequency domains and the multi-stage nature of the encryption.

## 1. INTRODUCTION

Recent research has focused on securing speech communication through encryption and scrambling techniques. Various methods employ chaotic maps for their efficiency and security. One approach uses chaotic shift keying to permute speech samples at different levels, followed by Chen map permutation, providing strong diffusion and confusion [1]. Another system combines permutation and substitution using logistic and Arnold cat maps in time and transform domains, offering low residual intelligibility and high key sensitivity [2]. A multi-step process involving random permutation, chaotic Bernoulli mapping, and pseudo-random binary scrambling has also been proposed [3]. These methods aim to transform intelligible speech into unintelligible forms, protecting against eavesdropping and unauthorized access. Cryptographic algorithms for speech security should provide high-level protection while maintaining audio quality and efficient recovery of original signals [4].

In [5], discrete wavelet transform and sample scrambling based on a chaotic map were used for speech compression and encryption. In [6], discrete cosine transform and chaotic maps were employed for speech

* Corresponding Author: j_shirazi@iau-gonabad.ac.ir
Assistant Professor, Department of Electrical Engineering, Gonabad Branch, Islamic Azad University, Razavi Khorasan, Gonabad, Iran



http://dx.doi.org/10.47176/TMI.2021.70

encryption. In [7], the method of discrete cosine transform and chaotic maps was used for speech encryption, evaluated using various criteria, and its robustness against attacks was assessed. In [8], discrete cosine transform was utilized for speech encryption. Initially, the signal was transformed, and then the transform coefficients were scrambled. Subsequently, the inverse discrete cosine transform was applied, and random numbers were added to the resulting time-domain samples. In [9], discrete wavelet transform and sample scrambling in time and frequency domains based on random numbers from a chaotic map were used for speech encryption.

In this study, a method based on using a chaotic map for generating random numbers, altering speech signal samples in the time domain based on these numbers, scrambling the resulting samples, applying discrete cosine transform to the resulting samples, and scrambling these samples is presented. The second section of the paper discusses various speech encryption methods. The third section presents the proposed method for encryption using a chaotic map, and the fourth section provides an evaluation of the proposed method. Finally, the fifth section concludes with conclusions and suggestions.

## 2. CONVENTIONAL SPEECH ENCRYPTION METHODS

The conventional method for encrypting speech signals involves permuting or scrambling speech signal samples in the time or frequency domain. In the frequency domain, methods such as frequency inversion, frequency shifting, permuting frequency bands, and scrambling samples of discrete signal transforms like Fourier transform, discrete cosine transform, and wavelet transform, as well as using spread spectrum, can be mentioned. In the time domain, methods include time sample inversion, permutation in this domain, and digital amplitude masking. To increase security, time and frequency methods are often combined, known as two-dimensional speech encryption algorithms. Examples include frequency-inversion with time-domain permutation and frequency-band permutation with time-domain permutation.

For scrambling time samples or transformed signal samples, random codes are used. Chaotic maps or functions, which have numerous applications in various sciences, are also used for generating random numbers in encryption. To complicate the encryption, some studies have applied alterations to time or frequency domain samples.

## 3. CHAOTIC MAP AND EMPLOYED ENCRYPTION METHOD

### 3.1. Chaotic Map

In this study, the logistic chaotic map was used to generate random numbers. This map, similar to other chaotic functions, is highly sensitive to the initial values of the map's parameters, and any change in them results in significant variations in the function's behavior. The logistic map is represented by Equation 1, where R and $x_0$ are its parameters. Typically, R is chosen between 0 and 4, and $x_0$ between 0 and 1. The map's output at each iteration is a number between 0 and 1 [10].

$$x_{n+1} = Rx_n(1 - x_n) \tag{1}$$

### 3.2. Employed Encryption Method

Figure 1 illustrates the encryption stages used. Initially, the speech signal was sampled at a rate of 8 kHz and segmented into 30-millisecond frames. In the next stage, for each frame, the bytes were XORed with random numbers generated from the chaotic map, and then the resulting samples were scrambled based on a periodically changing random code. Discrete cosine transform was then applied to the obtained samples, and the resulting values were scrambled based on the periodically changing random code. For each frame, the final resulting samples were considered the time-domain samples of the encrypted signal. To further increase the encryption complexity and thus the security of the encrypted signal, the random number code used for scrambling the XORed samples was periodically changed according to a specific order, ensuring that two consecutive frames did not have the same code.

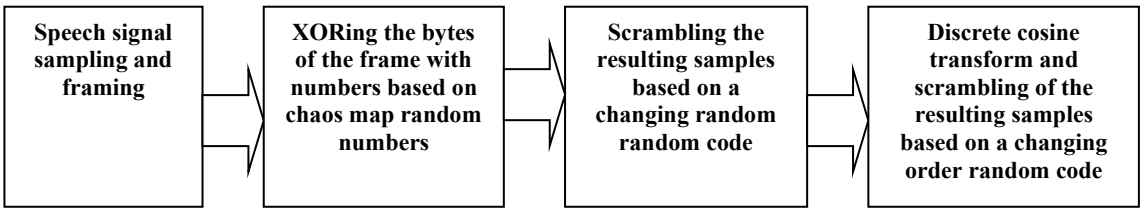


**Fig.1.** Encryption stages

In the receiver, to reconstruct the original speech signal, the inverse operations performed at the transmitter are applied. These stages include segmenting the speech signal at the receiver similarly to the transmitter, restoring scrambled samples to their original positions, applying the inverse discrete cosine transform, restoring scrambled samples to their original positions, and XORing the resulting bytes with the random numbers from the chaotic map used at the transmitter. For this purpose, the receiver must have the initial random number code used at the transmitter, the order of changes in this code, the chaotic map used at the transmitter, and its parameters.

## 4. EXPERIMENTS AND EVALUATION OF THE EMPLOYED METHOD

To perform experiments and evaluate the proposed method, various speech files were used, and for each file, two encrypted and reconstructed speech files at the receiver were saved. In the first stage of evaluation, the auditory quality criterion was used, wherein the encrypted speech signals and the reconstructed signal at the receiver were subjected to auditory tests by several individuals. All results indicated the inability to comprehend auditory information from the encrypted signal, which had a noise-like sound, and complete comprehension of the reconstructed speech signal. Figures 2 to 5 show the results obtained for a sample speech signal.

Figure 2 shows the histogram of the original speech signal and the encrypted signal. In a good encryption, the histogram should show almost uniform amplitude across all encrypted samples. The figure indicates that the histogram of the final encrypted samples is much more suitable compared to the original speech signal samples. The histogram of the samples before the discrete cosine transform is closer to the ideal state.

Figure 3 shows the time-domain waveforms of the original speech signal, the encrypted signal, and the reconstructed signal at the receiver. It can be observed that the original and reconstructed signals are entirely similar in the time domain, and the encrypted signal bears no resemblance to the original signal, exhibiting noise-like behavior.

Figure 4 shows the frequency spectra of the original speech signal, the encrypted signal, and the reconstructed signal at the receiver. It can be observed that the original and reconstructed signals have identical spectral information, while the encrypted signal's spectrum is scrambled, making it impossible to extract information from it.

For quantitative evaluation, the correlation coefficient criterion according to Equation 2 was used. This coefficient indicates the correlation between two signals x and y and ranges between -1 and 1. The closer this coefficient is to 1, the greater the similarity and correlation between the two signals, and the closer it is to zero, the lesser the correlation between the two signals.

$$r_{xy} = \frac{cov(\mathrm{x,y})}{var(x).var(y)} \tag{2}$$

To evaluate, ten diverse speech files were tested. In each stage, the correlation coefficient between the original speech signal and the encrypted signal, and the correlation coefficient between the original speech signal and the reconstructed signal at the receiver were calculated. Table 1 shows the average results of this criterion when the chaotic map parameter Rat the transmitter and receiver is identical and when there is a slight difference. When the chaotic map parameter R at the transmitter and receiver is identical, a correlation coefficient close to zero between the original speech signal and the encrypted signal indicates no correlation between the two signals, and a correlation coefficient of 1 between the original speech signal and the reconstructed signal at the receiver indicates complete

correlation and similarity between the two signals. The third row of the table shows that the proposed method is highly sensitive to the chaotic map parameter R, and a slight change in it at the receiver prevents the original speech signal from being reconstructed at the receiver.

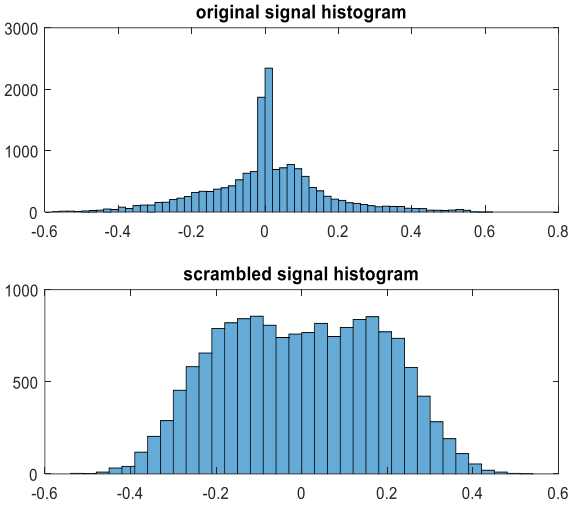


**Fig.2.** Histogram of Original and Encrypted Speech Signals

**Table 1.** Correlation coefficient results

| R Chaotic Map | Signal x | Signal y | Correlation Coefficient |
|---|---|---|---|
| **Equal Sender and Receiver** | Primary Signal | Encoded Signal | -0.0037 |
| **Equal Sender and Receiver** | Primary Signal | Reconstructed Signal | 1.0000 |
| **Different Sender and Receiver** | Primary Signal | Reconstructed Signal | -0.0171 |

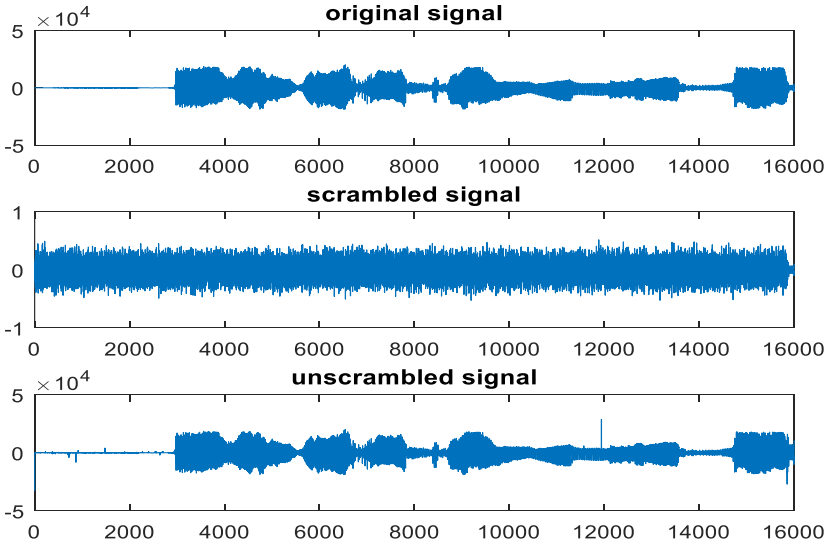


**Fig.3.** Time-Domain Signals of Original, Encrypted, and Reconstructed Speech

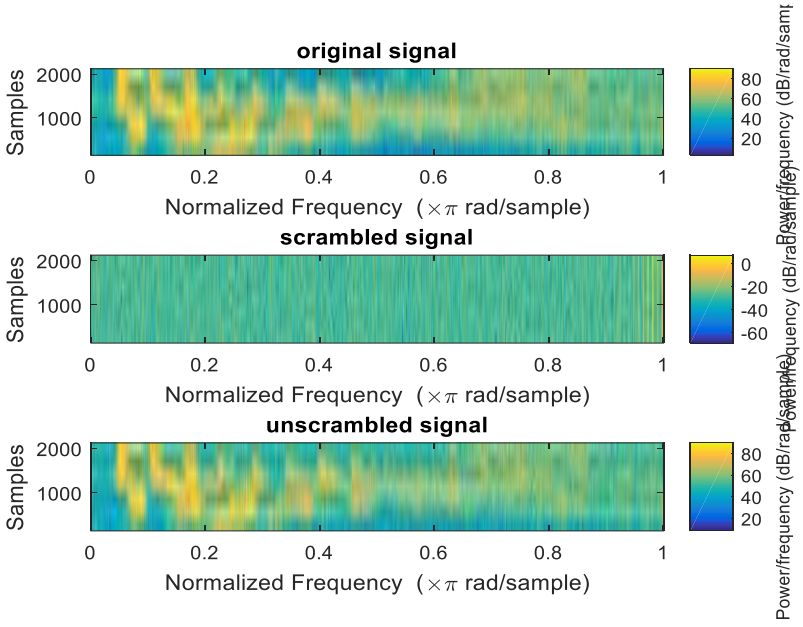


**Fig.4.** Frequency Spectra of Original, Encrypted, and Reconstructed Speech Signals

## 5. CONCLUSION

In this paper, aiming to enhance the security of speech signal transmission, a multi-stage encryption system in both time and frequency domains using a chaotic map was proposed and implemented. After segmenting the speech signal into frames, the data of each frame were XORed with random numbers derived from a chaotic map, and the resulting samples were then scrambled based on a periodically changing random code. The discrete cosine transform was applied to the obtained samples, and the resulting values were scrambled again based on the periodically changing random code. To further increase the encryption complexity and hence the security of the encrypted signal, the random number generator code was periodically altered according to a specific order, ensuring that consecutive frames did not use the same code. Various qualitative and quantitative criteria were used to evaluate the method, and the results indicated high levels of these criteria for the proposed method. Compared to previous research, which performed encryption solely in the time or frequency domain, the proposed method, by combining these two domains and using multi-stage encryption, achieves higher security due to its increased complexity.

### Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

### Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

### Declaration of Interest

The authors declare that they have no competing interests.

**Funding**

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**REFERENCES**

[1] Sathiyamurthi, D., & Ramakrishnan, S. (2017). Speech encryption using chaotic shift keying for secured speech communication. *EURASIP Journal on Audio, Speech, and Music Processing, 2017*. https://doi.org/10.1186/s13636-017-0118-0

[2] Al Saad, S. N., & Hato, E. (2014). A speech encryption based on chaotic maps. *International Journal of Computer Applications, 93*(4). https://doi.org/10.5120/16203-5488

[3] Dhanya, G., & Jayakumari, J. (2017). Speech scrambling based on chaotic mapping and random permutation for modern mobile communication systems. *APTIKOM Journal on Computer Science and Information Technologies.*

[4] R., A., & Chithra, D. P. (2016). A review on cryptographic algorithms for speech signal security.

[5] Enache, F., Deparateanu, D., Oroian, T., Popescu, F., & Vizitiu, I. (2015). Theoretical and practical implementation of scrambling algorithms for speech signals. In *2015 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)* (pp. 49-52). Bucharest, Romania. https://doi.org/10.1109/ECAI.2015.7301160

[6] Hayder, K. Z., Sadiq, A., & Mehdi, Sattar, B. S. (2021). Speech scrambler based on discrete cosine transform and novel seven-dimension hyper chaotic system. *Journal of Physics: Conference Series.*

[7] Zeeshan, H., Jan, S. K., Jawad, A., Muazzam, A. K., & Fadia, A. (2010). Secure speech communication algorithm via DCT and TD-ERCS chaotic map. In *2010 4th International Conference on Electrical and Electronics Engineering.*

[8] Shirazi, J. (2021). Speech signal encryption using discrete cosine transform and sample scrambling in time and frequency domains. In *National Conference on Innovation and New Technologies in Electrical and Computer Engineering*, Khajeh Nasir Toosi University of Technology, September 14.

[9] Shirazi, J. (2021). Speech signal encryption using discrete wavelet transform and chaotic map. In *Fourth National Conference on New Technologies in Electrical and Computer Engineering*, Isfahan, September 30.

[10] Enayatifar, R., Abdullah, A. H., & Isnin, I. F. (2014). Chaos-based image encryption using a hybrid genetic algorithm and a DNA sequence. *Optics and Lasers in Engineering, 56*, 83-93. https://doi.org/10.1016/j.optlaseng.2013.12.003