Voice Signal Encryption Scheme Using Transformation and Embedding Techniques for Enhanced Security

Conference Paper · October 2018

DOI: 10.1109/ICISPC44900.2018.9006681

2 authors:

Chithra Pl
University of Madras
39 PUBLICATIONS **35** CITATIONS
SEE PROFILE

Aparna Ramdoss
Shasun Jain College For Women
9 PUBLICATIONS **18** CITATIONS
SEE PROFILE

Voice Signal Encryption Scheme Using Transformation and Embedding Techniques for Enhanced Security

PL. Chithra
Department Of Computer Science
University of Madras
Chennai -600005
e-mail: chitrasp2001@yahoo.com

Aparna R.
Research Scholar (UNOM)
Department Of Computer Applications
S.S.S. Shasun Jain College, Chennai-17
e-mail: aparna.r@shasuncollege.edu.in

***Abstract*—Voice is the most common mode of human interaction. Speech signal plays a vital role in communication and data transmission. It will be very convenient to transfer data in the form of speech to exchange data in easy and faster way. Voice communication also adopted in various fields for easy data exchange. Data flow enormously in open network. Hence it is considered as most vulnerable part where intruders attempt to access secret data that are transmitted across the network. It is essential to devise competent method to enhance data security during voice transmission. In this paper, an efficient technique is proposed for enhanced security in voice signal encryption scheme. The inclusion of transformation and embedding techniques further makes the system very secure for speech signals over open network paths.**

Keywords-voice signal security; speech cryptography; kaiser window; CAT transformation

I. INTRODUCTION

A signal is an action that conveys information about the behavior or attributes of some portent. The use of sounds and words to express or signal others is referred as verbal communication. Sound is a vibration which disseminates as audible wave of pressure through a transmission medium. Sound waves possess frequency. Sound waves with the frequency range 20Hz-20KHz can only heard by humans. Basically voice signal possess analog characteristics and travels in the form of waves. The proposed method deals with data security and signal steganography, the audio has to be converted into cipher and embedding technique should not give any traces of suspicion about the hidden data.

A. Signal Cryptography

Cryptography is the process of converting data into unreadable format before transmission and converting back to original form. This enhances security and protects data from eavesdroppers and hackers. Data is considered as the most valuable asset, so it should be given perfect security. In many fields such as military, financial sector, the data plays a vital role and security is essential. Data has to be transmitted, but at the same time, it should not be disclosed to anybody. Thus, converting the important information into non intelligible data is the preliminary step in secure transmission. As technologies are growing.

B. Signal Stegnography

Steganography is the process of hiding secret information in another file, without leaving any traces of the hidden information. Steganography comprises the concealment of secret information within another message. The best part about steganography is that, the intended secret information does not draw attention to itself as a part of scrutiny. Media files such as audio, video data, that are really large, are best suited for steganographic transmission. If cryptography alone incorporated, encrypted messages, no matter how strong the algorithm to make it unbreakable they are, arouse interest in hackers and may attract them to open it. Thus by combining the cryptography and steganography techniques implements the development of secure system.

Speech samples are taken from OGMLTS-TIMIT corpus, OpenSLR(LibriSpeech ASR) corpus,LDC-IL corpus.

This paper is organized as follows. Section II briefly explains Background work of the proposed work, Section III elaborates Proposed Methodology, Section IV discusses the Detailed Description about the proposed algorithm, Section V shows the Efficiency Analysis and Section VI gives the conclusion of the work.

II. BACKGROUND

Farsana FJ, Gopakumar K devised a method [1] for signal encryption using Zaslavsky map for pseudo random number generation. Cat transformation is used for making the system more secure. P. Sathiyamurthi and S. Ramakrishnan developed a secure method [2] using chaotic shift keying technique wheareas chaotic modulation technique [3] was used by Mahmoud F. Abd Elzaher *et al.* Lorenz map key generator method [4] is incorporated in encryption bymahmood K. Ibrahem, Hussein Ali Kassim. To make the transmission more secure, audio/wave steganography was proposed by Anant Umbarkar *et al.* [5] and Kundu N., Kaur A. [6]. Our previous work elaborates various cryptographic techniques & windowing schemes [7], [8]. One time pad algorithm [9] usage was described by Y. Saleem *et al.* Benefits of using two dimensional chaotic maps [10] in encryption process is illustrated by Alzharaa Mostafa *et al.* Three dimensional chaotic map was used for private key encryption [11]. The concept of speech

978-1-5386-6981-5/18/$31.00 ©2018 IEEE

compression [12] using chaotic sensing with large size key was explained by Al-Azawi & Ali M. Gaze. It is necessary to apply strong cryptographic technique to ensure network security [13]. Chaotic maps and DNA encoding scheme was briefed in the research article by S. J. Sheela *et al.* [14].Role of twin stream cipher algorithm [15] was described by Omar. M. Hammad *et al.* Hence it is identified that, combining encryption & embedding techniques enhances the proposed method by providing two level security to the secret signal.

III. PROPOSED SYSTEM

A. *Algorithm*

Encryption:

Step 1: Digitalization of Speech Signal
Step 2: Reversal of the signal
Step 3: Take DCT (Discrete Cosine Transform) & apply windowing technique to get short-term energy function of the speech signal
Step 4: The one dimensional data stream is converted into two dimensional data stream.
Step 5: Perform Cat map – transformation with the parameters a=b=c=1 and d=2.
Step 6: Apply the approximation factor relevant to the input speech signal
Step 7: Revert back the two dimensional data stream into one dimensional data stream.
Step 8: Embed the resultant with another signal.

Decryption:

Step 1: Digitalization of encrypted signal
Step 2: Remove the embedded signal
Step 3: The one dimensional data stream is converted into two dimensional data stream.
Step 4: Apply the approximation factor.
Step 5: Perform Cat map – transformation with the parameters a=b=c=1 and d=2.
Step 6: Revert back the two dimensional data stream into one dimensional data stream.
Step 7: Take Inverse Cosine Transform (IDCT)
Step 8: Reverse the resultant data stream.

B. *Flow Diagram*

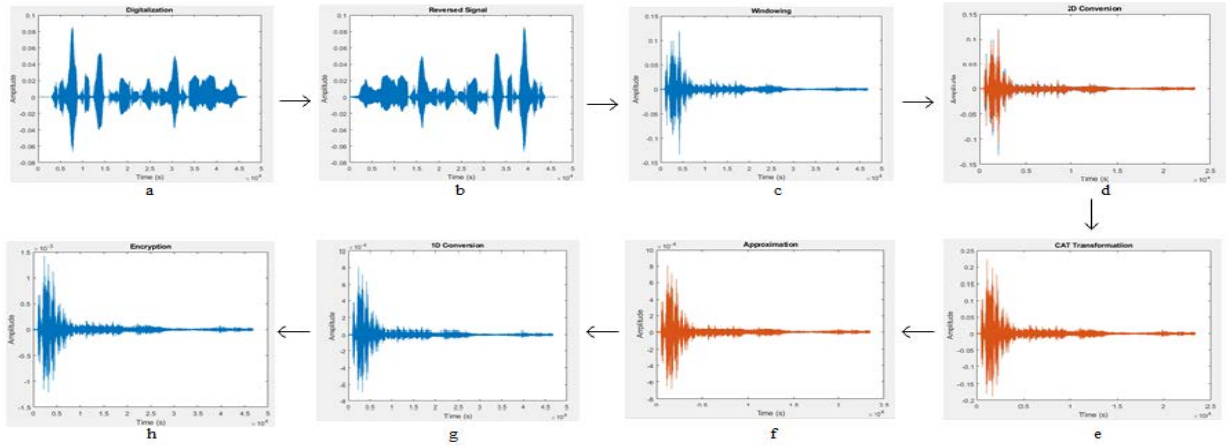

Figure 1. Encryption process a)digitalization b)reversing c)windowing d)1D conversion e)CAT transformation f)approximation g)2D conversion h)embedding.

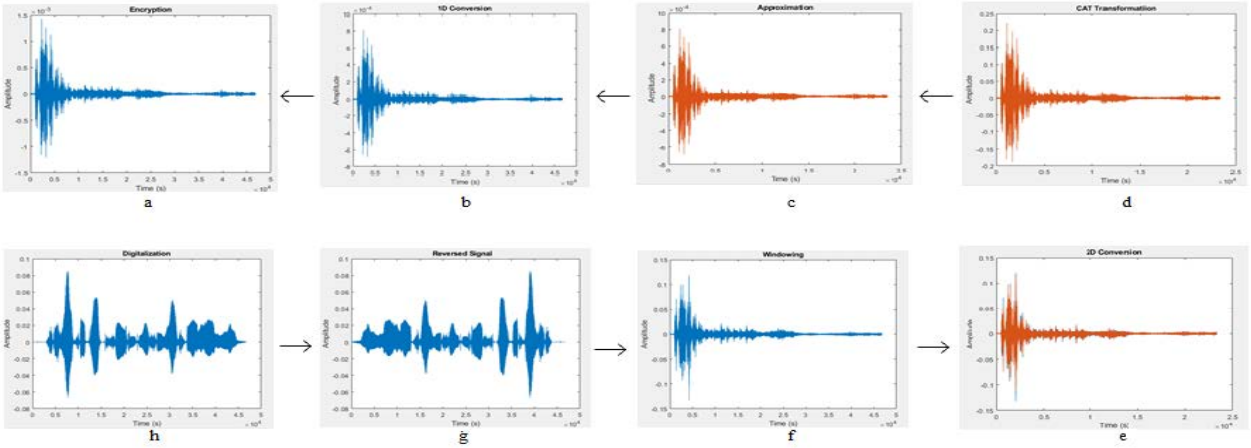

Figure 2. Decryption process a)digitalization b)splitting out secret signal c)1D conversion d)normalization e)CAT transformation f)2D conversion g)IDCT h)secret signal.

IV. Detailed Description

A. Digitalization

The binary representation of speech signal is considered digitalization [9] and it is the basic step in signal processing. The raw analog data has to be converted into digital to do the processing. As the proposed system deals with signal cryptography and signal steganography [5], both the secret signal and the signal to be embedded with should be digitalized.

B. Reversal of the Signal

To enhance the signal security the secret signal should be reversed. It will be comfortable to embed signals of invariable size. It is not possible to restrict the length of the signal as it significantly implies the efficiency of the proposed technique. Hence, to perform the final step of embedding, the signals should be adjusted to same size by adding the minimal values to the shortest signal.

C. Windowing

Most probably the audio signals are very large and many times generate infinite dataset. It is essential to create an efficient subset, which possess all the important information. Windowing is the process of identifying the proper subset of considerably large data. There are many widowing techniques [8] available but choosing the proper window also plays a vital role. Selecting appropriate window is very crucial, as it deals with signal strength. Here, we have used Kaiser Window. The main feature of Kaiser Window is that, it approximates the DPSS window which maximizes the energy concentration in the main lobe which is otherwise very difficult to compute. In multitaper spectral analysis, the DPSS – Discrete Prolate Spheroidal Sequence window is used to maximize the energy concentration and reduces information loss at the edge of the window. For the same reason, Kaiser window is used in the proposed method.

D. Transformation

DCT (Discrete Cosine Transform) is used to encode [10] audio signals. It is done to sequence the consine functions of the signal at different frequencies. The two simple processes in any cryptographic algorithms [8] are confusion and diffusion. Preceding the encryption process the speech signal is converted to frequency domain by discrete cosine transform. Cat transformation is [1] used to improve the confusion and diffusion process. Cat transform parameters are taken such that the determinant of the considered matrix is 1, in order to suit better for the considered work. Here, the cat transform parameters are taken as a=b=c=1 and d=2, which results in expected efficient output. As the speech signal is 1D values, it has to be converted into 2D to perform Cat transformation [1]. Obtained data after the Cat transformation will be in 2D, it should be converted back to 1D.

E. Encryption

Speech signal encryption is the process of converting the original form of speech signal into unrecognized/unintelligible form to improvise the security of communication through an insecure channel. Here, the encryption process is complete with 2D to 1D conversion [1] after applying the normalization factor. This normalization factor plays a vital role as its part is to ensure generating transmittable signal with negligible variations, which in turn does not provide any possibility for suspicion. Hence, it makes difficult for the hacker to sense the secret signal.

F. Embedding

Embedding is a steganographic technique [5] used here to enhance the efficiency of the proposed method. A signal is considered as embed signal and to which the encrypted signal is merged. This technique of embedding the encrypted voice signal will ensure two level security to the secret signal. After applying the identified normalization factor of the signal, it becomes cipher voice which when embed with the original embed signal does not produce and trace of embedded signal. The original embed signal will remain same before and after the embedding of the secret signal, this in turn the advantage of the proposed method. Now, the secret signal is perfectly hidden in the original embed signal and can be sent through any unsecure transmission medium.

G. Decryption

Decryption is the counter technique applied on the receivers end to get back the hidden encrypted voice [9], which has secret message. Encryption algorithm devised for the proposed method produce signal to be transmitted. In the receiver's end counter algorithm should be used to acquire the secret message.

H. SNR Calculation

Signal to Noise Ratio [12] is calculated for the original embed signal after corresponding adjustment and the signal after embedding. Values are compared and tabulated to prove that there is a minimal variation which indirectly highlights the quality of the proposed system. Also the SNR of secret signal and signal at receiver's end is compared to prove that there will be minimal or no loss of data.

I. Correlation before and after Embedding

Cross correlation is a measure to compare the similarity between the signals. Thus, the correlation value [1], [2] depicts the variation between the signal before and after encryption and embedding. Resultant value close to 0 proves there will not be any evidence of secret signal. Also, the correlation between original secret signal [13] and received signal is calculated, which results approximately equal to 1 [1] show the sent any received signals matches. Thus, correlation coefficient [3] values are calculated to check and prove the efficiency of the proposed methodology.

V. Effeciency Analysis

The efficiency of the proposed system is proven with the aid of the following Tables & Figures. Various equations involved in the methodology are also discussed. Equations specified below include Kaiser Window [8] transformation, Normalization calculation & Correlation coefficient calculation. Equation (1) is used by which Kaiser window is applied to bring the values to easily manageable domain.

Normalization factor is calculated, in eq. (2) such that there will not be any trace of secret signal after embedding. Eq. (3) is used to generate the counterpart of normalization factor during decryption process. This adjustment factor plays a vital role in hiding secret signal without any identifiable modification in the original embed signal. Eq. (4) is used to find the correlation between the signals. Spectrogram given in Fig. 1& 2 clearly illustrates the visual representation of the secret signal and encrypted signal. The variations observed show the effectiveness of the cryptographic technique applied.

A. Equations

1) Kaiser Window

$$w[n] = \begin{cases} \frac{I_0\left[\alpha\sqrt{1-\left(\frac{2n}{N-1}-1\right)^2}\right]}{I_0[\propto]}, & 0 \le n \le N\text{-}1 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

2) Normalization Factor

- Encryption

$$f(n) = \begin{cases} n \in N, \frac{SNR(A)}{n} = SNR(B) \\ 0, \quad \text{otherwise} \end{cases} \quad (2)$$

where, f(n) → function to find apt normalization factor during encryption,

A→signal after embedding,

B→Original embed signal

- Decryption

$$g(n) = \begin{cases} n \in N, SNR(A)\ X\ n = SNR(B) \\ 0, \quad \text{otherwise} \end{cases} \quad (3)$$

where, g(n) → function to find apt normalization factor during decryption,

A→Received Signal,

B→Original embed signal

The normalization factor is computed in such a way that there is a negligible or no doubt of hidden secret signal [6] in the original signal which will be transmitted on unsecure [14] medium. Normalization factor is computed based on the comparison of SNR value of the signal before and after embedding.

3) Correlation Analysis

$$\rho(A,B) = \frac{cov(A,B)}{\sigma_A \sigma_B} \quad (4)$$

where,

Correlation 1:

ρ →correlation coefficient, A → *Embed Signal* & B→Embedded Signal

Correlation 2:

ρ →correlation coefficient, A→ *Secret Signal* & B→Encrypted Signal

Table I with Correlation coefficient comparison values between normalized original embed signal and embedded transmitted signal which approximately equals to 1 proves that there will minimal negligible unidentifiable modifications even after hiding the secret message with the embed signal. Table II with Correlation coefficient comparison values between original secret signal and received decrypted signal which approximately equals to 1 shows that the secret message signal reaches the receiver without much modifications.

Our experiment results are tabulated and compared in 3 tables, Table III, Table IV & Table V due to the variations in dataset used. The correlation between secret signals before and after encryption using the existing methods [1], [2] is given in Table III & Table IV. Hence it is proved; Table V with the results of the proposed technique overwhelms the existing methods.

B. Spectrogram

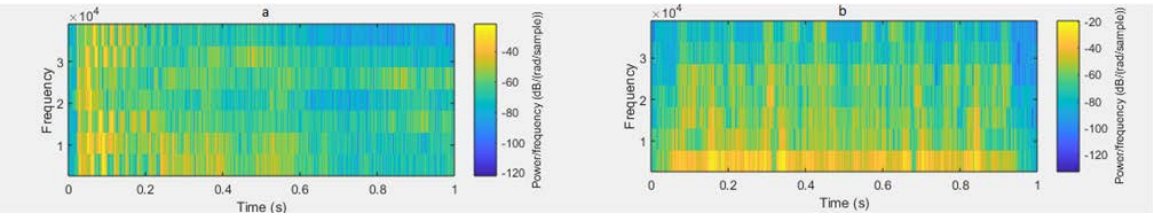

Figure 3. TIMIT corpus spectrogram- a)secret signal b)encrypted signal.

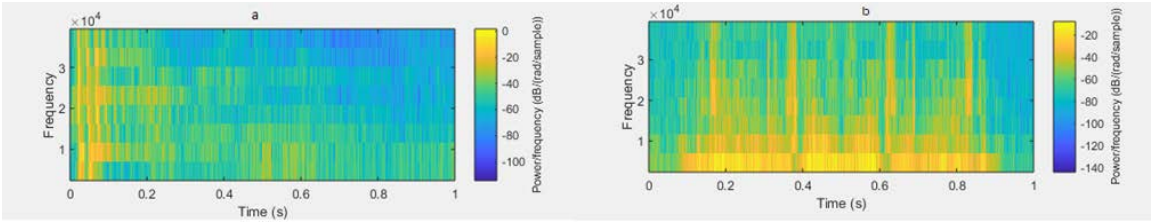

Figure 4. OpenSLR librespeech ASR corpus- spectrogram a)secret signal b)encrypted signal.

C. Tables

TABLE I. SIGNAL NOISE RATIO OF ORIGINAL EMBED SIGNAL BEFORE & AFTER EMBEDDING (CORRELATION BETWEEN ORIGINAL EMBED SIGNAL BEFORE & AFTER EMBEDDING)

#	Signal DB	N (Approx. Factor)	SNR VALUES		Correlation
			Signal after size adjustment	*After Embedding*	
1	TIMIT Corpus	156	-12.8293	-12.8293	1
2	OpenSLR	45	-14.7208	-14.7208	0.9992
3	LDC-IL	275	-16.9355	-16.9355	1

TABLE II. SIGNAL NOISE RATIO OF SECRET SIGNAL & DECRYPTED SIGNAL (CORRELATION BETWEEN SECRET SIGNAL & DECRYPTED SIGNAL)

#	Signal DB	SNR Values		Correlation
		Secret Signal (Original)	*Received Signal (Decrypted)*	
1	OGMLT-Corpus	-10.5502	-10.5417	0.9855
2	OpenSLR	-7.7999	-7.7986	0.9999
3	LDC-IL	-15.3694	-15.3665	0.9980

TABLE III. EXISTING METHOD[1] (CORRELATION BETWEEN ORIGINAL & ENCRYPTED SIGNAL)

#	Sample Files	Correlation
1	A	0.000569
2	B	0.000819
3	C	0.000456

TABLE IV. EXISTING METHOD [2] (CORRELATION BETWEEN ORIGINAL & ENCRYPTED SIGNAL)

#	Sample Files	Correlation
1	Audio1.wav	0.0233
2	Audio2.wav	0.0384
3	Audio3.wav	0.0157

TABLE V. PROPOSED METHOD (CORRELATION BETWEEN ORIGINAL & ENCRYPTED SIGNAL)

#	Sample Files	Correlation
1	TIMIT Corpus	-0.0015
2	OpenSLR	0.00011469
3	LDC-IL	-0.00020515

VI. CONCLUSION

The speech signal carrying secret information is prone to cyber-attack. Hence, the proposed technique provides a strong algorithm to protect the secret message from snooping. Combing the cryptographic & steganographic techniques strengthens the effectiveness of the proposed method. Embed signal acquired after encryption is found to be similar with original embed signal indicates that it doesn't leave any traces of hidden signal. Signal obtained after decryption almost matches with the sent secret signal. Signals from standard corpus such as TIMIT database, OpenSLR – LibriSpeech ASR corpus and LDC-IL are used in the proposed work validates it. Also the comparison of results obtained with existing method proves it to be a better technique.

ACKNOWLEDGMENT

Sincere thanks to the Department of Computer Science, UNOM & S.S.S. Shasun Jain College for consistent support in carrying the research work.

REFERENCES

[1] Farsana F J , Gopakumar K,A Novel Approach for Speech Encryption: Zaslavsky Map as Pseudo Random Number Generator,6th International Conference on Advances In Computing & Communications, ICACC 2016 ,pp. 816-823, 6-8 September 2016.

[2] Sathiyamurthi and Ramakrishnan, Speech encryption using chaotic shift keying for secured speech communication, EURASIP Journal on Audio, Speech, and Music Processing (2017) 2017:20.

[3] Mahmoud F. Abd Elzaher et.al, A Speech Cryptosystem Based onChaotic Modulation Technique,International Journal of Emerging Trends & Technology in Computer Science Volume 4, Issue 2, pp. 64-71,2017.

[4] Mahmood K. Ibrahem , Hussein Ali Qasim, "Implementation of VoIP Speech Encryption System Using Stream Cipher with Lorenz map Key Generator", International Journal of Scientific and Engineering Research , volume 8, issue 7, pp. 533-541,August 2017.

[5] Umbarkar, Dr. Anantkumar & Joshi, Abhijit & Jadhav, Ajay. (2010). Wave steganography,ICCIC.pp. 862-865,2009.

[6] N. Kundu, A. Kaur , "*Audio Steganography for Secure Data Transmission*", International Journal of Computer Sciences and Engineering, Vol.5, Issue.2, pp.124-129, 2017.

[7] R.Aparna ,Dr.P.l.Chithra,"A Review on Cryptographic Algorithms for Speech Signal Security" International Journal of Emerging Trends & Technology in Computer Science(IJETTCS), Volume 5, Issue 5,pp. 84-88, September - October 2016.

[8] R.Aparna,Dr.PL.Chithra,Role of Windowing Techniques in Speech Signal Processing For Enhanced Signal Cryptography,Advanced Engineering Research amd Applications, Chapter 28,Volume V,pp. 446-458,2017.

[9] Y.Saleem,M.Amjad,M.H.Rahman,F.Hayat,T.Izhar,M.Saleem,"Speec h Encryption,Implementation of One Time Pad Algorithm in Matlab" Pakistan Journal of Science,Vol.65, pp 114-118,March 2013.

[10] Mostafa, Alzharaa & Soliman, Naglaa & Abdalluh, Mohamoud & Abd El-Samie, Fathi. " Speech encryption using two dimensional chaotic maps". ICENCO, pp. 235-240,2015.

[11] Farsana F J, Dr.K.Gopakumar, Private Key Encryption of Speech Signal Based on Three Dimensional Chaotic Map, International Conference on Communication and Signal Processing, pp. 2197-2201, April 6-8, 2017.

[12] Maher K. Mahmood Al-Azawi 1 and Ali M. Gaze 1,Combined speech compression and encryption using chaotic compressive sensing with large key size,IET Signal Processing , Volume 12, Issue 2, pp. 214 – 218April 2018.

[13] Forouzan B., "Cryptography and Network Security" special Indian Edition 2007, pp.240-330 Tata McGraw-Hill Publishing Company Limited, New Delhi.

[14] S. J. Sheela, K. V. Suresh, and Deepaknath Tandur, “A Novel Audio Cryptosystem Using Chaotic Maps and DNA Encoding,” Journal of Computer Networks and Communications, vol. 2017, Article ID 2721910, 12 pages, 2017.

[15] Omar. M. Hammad, Hebah H. O. Nasereddin, Abdulkareem. O. Ibadi, Voice Encryption Using Twin Stream Cipher Algorithm, Journal of Advanced Computer Science and Technology Research, Vol.6 No.2, pp. 37-51,June 2016.