Secure speech communication algorithm via DCT and TD-ERCS chaotic map

Conference Paper · April 2017

DOI: 10.1109/ICEEE2.2017.7935827

5 authors, including:

Jan Sher Khan
Gaziantep University
25 PUBLICATIONS **353** CITATIONS
SEE PROFILE

Jawad Ahmad
Edinburgh Napier University
81 PUBLICATIONS **889** CITATIONS
SEE PROFILE

Secure Speech Communication Algorithm via DCT and TD-ERCS Chaotic Map

Zeeshan Habib[1], Jan Sher Khan[2], Jawad Ahmad[1], Muazzam A. Khan[3], Fadia Ali Khan[4]
[1] Department of Electrical Engineering, HITEC University Taxila, Pakistan
[2]Department of Electrical and Electronics, University of Gaziantep, 27310 Gaziantep, Turkey
[3]Department of Computer Engineering, National University of Sciences & Technology, Islamabad, Pakistan
[4]Department of Electrical Engineering, Riphah International University, Islamabad, Pakistan
e-mail: jskm893@gmail.com

***Abstract*—Secure communication has always been a demanding area in civil, commercial and particularly in military set up. Robust and time-tested efficient algorithms are needed to have an essential privacy for speech transmission in the telephone networks, radio communication and in the emerging cellular mobile radio systems. In this paper, we present a highly secure speech encryption algorithm based on amplitude scrambling and Discrete Cosine Transform (DCT) coefficients scrambling. The permutation is performed using TD-ERCS chaotic map. The proposed scheme has been implemented successfully on software framework which yielded zero residual intelligibility and high quality of recovered speech. Simulation results reveal that the anticipated algorithm is secure enough and resistant against various security attacks. The authors are reasonably confident of their contribution which can eventually be exploited as software, hardware-software and an integral part of any embedded secure communication system.**

Keywords-DCT; speech encryption; security; cryptography; secure communication

I. INTRODUCTION

Information spying has emerged due to substantial adaptation of speech communication in military, economic and social affairs. Therefore, security and privacy of data becomes important not only in military operations but in civil and business information systems as well. Rapid growth of technology has exposed many threats to secrecy and confidentiality of data. Hence, risk management and assessments is an essential attribute of secure data communication systems [1]. Although the risk factor cannot be eradicated, its affect can still be minimized by implementing data security techniques. The idea of encrypting a message is probably as old as messages had been sent between humans [2]. Most of earlier systems were designed for encrypting text messages; the development of a voice encrypting system is in turn a lot more challenging. Analogue techniques did not suffice to achieve satisfactory results in terms of security before the emergence of digital signal processors. For this reason much of the pioneering work with digital domain has been performed while inventing new digital systems to provide secure voice communication.

Voice encrypting systems are used to guarantee end-to-end speech security in real-time communication systems such as analogue radio, telephone, VoIP, GSM etc. The resources and methods to minimize the effect of threats considerably evolved as highly sophisticated techniques. Many diversified techniques have been developed over the centuries beginning with Caesar cipher all the way to modern standards like Advanced Encryption Standard (AES) [3], [4]. To shield voice communication whether achieved by digital or analogue means, the term scrambling is used to illustrate encryption. Audio scrambling technique encompasses three requirements: inaudibility, robustness and data rate. Embedding is acquired in such a way that scrambled speech signal should be completely inaudible to unauthorized listeners [5]. Therefore, the most significant feature of any speech scrambler is inaudibility. Robustness and data rate are also important but these two attributes cannot be achieved at the same time. Although standardized techniques are not meant to be challenged, yet there is always a room for improvement. Reference [6] presented a speech scrambling algorithm with Hadamard transform in the frequency domain with larger key space. Simulation and results show that it has low residual intelligibility and fair quality of the recovered speech. Transform-based analogue speech scramblers i.e., Prolate Spheroidal Transform (PST), Discrete Walsh Hadamard Transform, Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT) were critically analyzed and evaluated for the suitability of speech encryption [7]. Their results showed that DCT is the best transform to use in transform-based speech encryption. In [8], author described a speech scrambling technique based on Orthogonal Frequency Division Multiplexing (OFDM). The algorithm proposed in [8], has reduced the residual intelligibility and computational complexity. However, this scheme does not change the redundancy of speech, which leads to the intelligibility of the encrypted analog signal.

Speech encryption being an integral part of cryptography is playing a vital role in secure communication. Scrambling can be performed in time-domain or frequency-domain or it can also be done using both dimensions (combination of time domain and frequency domain). In [9], Mathews et al present chaos-based encryption algorithm for the first time. After the novel idea presented by Mathews et al many chaos-based algorithms have been proposed so far. Due to ergodicity, structure complexity, deterministic pseudo randomness, Sensitivity Dependence on Initial Conditions (SDIC) and simple implementation, chaotic maps are employed in image, video and audio encryption schemes [10]. Researchers are using different maps for example, Unimodel Skew-Tent map, Logistic map, Arnold map, Henon map, Standard map and Baker map in the designing of encryption algorithms [4],

978-1-5090-6789-3/17/$31.00 ©2017 IEEE

[10]–[19]. In 2004, Professor Yuan Sheng Lee developed a special two-dimensional chaotic system, based on a physical model of ellipse reflecting cavity, which is known as Tangent Delay Ellipse Reflecting Cavity Map System (TD-ERCS) [20]. Chaotic characteristics and other details of TD-ERCS can be found in our previous research [21]. Security level of any encryption algorithm is measured using the level of residual intelligibility in a scrambled data. Low level of residual intelligibility describes a higher level of security and vice versa [22]. Recovered speech quality is also a demanding attribute of design criteria for encryption algorithms and hence cannot be ignored or overlooked. A variety of encryption schemes have been designed and developed, offering different level of security and recovered speech quality. However a trade-off has to be made between the security level and the quality of the recovered signal [23]. In the preceding literature review, the schemes discussed have paid no or less attention towards software-specific residual intelligibility and recovered speech quality which play an important role in secure communication systems. On the contrary, the software-based approach proposed in this paper has very low residual intelligibility and excellent quality of recovered speech. Moreover, to enhance the security of the speech signal, amplitude scrambling and frequency domain scrambling is performed collectively. The presented scheme is highly compatible in terms of implementation with microprocessors, micro-controllers, Digital Signal Processors (DSP) or any other embedded system.

Section II presents the steps involved in the proposed scheme. Section III elaborates simulated results. Section IV describes the security evaluation of the proposed technique. Finally section V concludes the research work with future recommendations.

II. Proposed Algorithm

A detailed and comprehensive literature survey is carried out in section I, in which different time domain, frequency domain and transform domain encryption techniques have been discussed and analyzed. Problems of secure communication, speech intelligibility of scrambled signal, quality of recovered signal and attempts to improve problem solutions are summarized. This section details with implementation stages and methodology of the proposed scheme.

Purpose of speech encryption is to carry out such operations on a speech signal so that it should be entirely unintelligible to an intruder, even having slightly different key. The main focus of this research is to design and implement a robust secure speech encryption system. In this regard, a new and fast secure speech encryption scheme is proposed, based on DCT and TD-ERCS chaotic maps. The proposed scheme is implemented and finally analyzed using different speech samples. The block diagram of the proposed software-based secure speech encryption algorithm is shown in Fig. 1. The proposed algorithm is designed to provide speech security over a band-limited channel, assumed to pass frequencies in the range zero to 4 kHz. The anticipated scrambler operates on sampled version of speech. After conversion to digital domain, signal is divided into different frames of equal length. Each frame consists of 64 samples. Then DCT is applied on each frame giving 64 transform coefficients in each frame followed by amplitude scrambling of each frame.

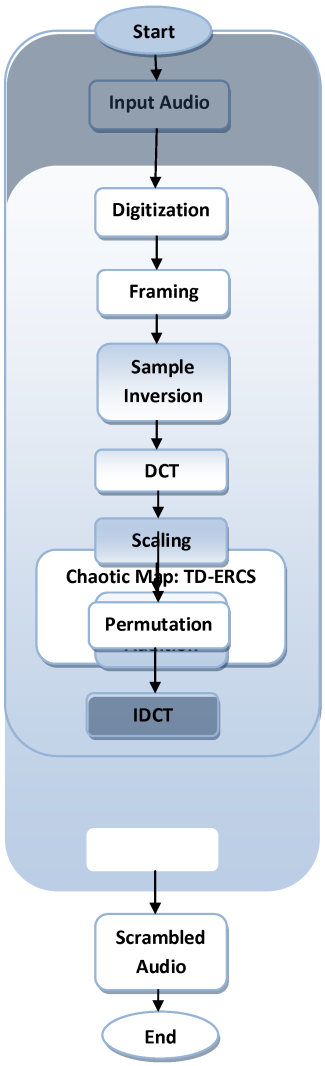


Figure 1. Flow chart of proposed scheme.

Speech data is taken and then passed through the low pass filter (LPF) with cut-off frequency of 3500 Hz to extract only the speech signal part as human voice effective frequency range is between (300-3400) Hz. The data is passed through Analogue-to-Digital (A/D) converter followed by sampling according to the Nyquist criterion. The discrete-time data stream is divided into blocks of 64-bit length for further processing. Each block sample is inverted followed by scaling its DC value down to zero. Then DCT is applied on resulting data bits taken from the input buffer. After applying the DCT, a pseudo random noise vector is added to the resultant DCT coefficients, obtained in frequency domain. Now, frequency domain scrambling is performed using pseudorandom numbers generated via TD-ERCS chaotic map.

Then inverse DCT is applied on this scrambled speech samples and converting it back into time-domain signal. The encryption process is complete now and speech data is ready for trans-

mission over open and insecure channel. To get the intelligible audio signal back, the same operation is carried out exactly in

III. Simulation and Results

Simulation results of the proposed scheme are shown from Figs. 2 to 4. These simulations are carried out using MATLAB 2012*b*. In order to demonstrate the results and simulations, different speech samples are recorded and analyzed. Figure 2 shows the simulation results of the original input signal of a male voice counting digits 0 to 10. Figure 2(a) shows the time domain representation of original input recorded signal. One can clearly see the variation of amplitude representing different digits spoken by the male and in between two successive digits spoken there is a small pause. Figure 2(b) shows the spectrogram of digits file. The spectrogram represents the distribution of frequencies as a function of time, which is basically a time-varying spectral demonstration that represent how the spectral density of a signal alter with time. Time information is represented on the horizontal axis and the vertical axis represents the frequency information, indicating the amplitude of a unique frequency at a particular time, described by the strength or color of every point in the speech. Similarly Fig. 3 gives information about the encrypted speech signal. Figure 3(a) is a time domain representation of encrypted speech signal. In contrast to Fig. 2(a), one can clearly see that there is no significant change in the amplitude of the encrypted signal in Fig. 3(a). This shows that the actual speech signal is encrypted and masked just like white noise giving no information intelligibility. Figure 3(b) presents the spectrogram of encrypted speech signal. From the spectrogram of the scrambled speech signal, it can be seen that frequency is scattered over the entire spectrum. Spectral density is same over the whole time period, which makes it just like white noise data spectrogram. Figure 4 gives the information of decrypted speech signal. Figure 4(a) shows the simulation of time domain decrypted speech signal and 4(b).

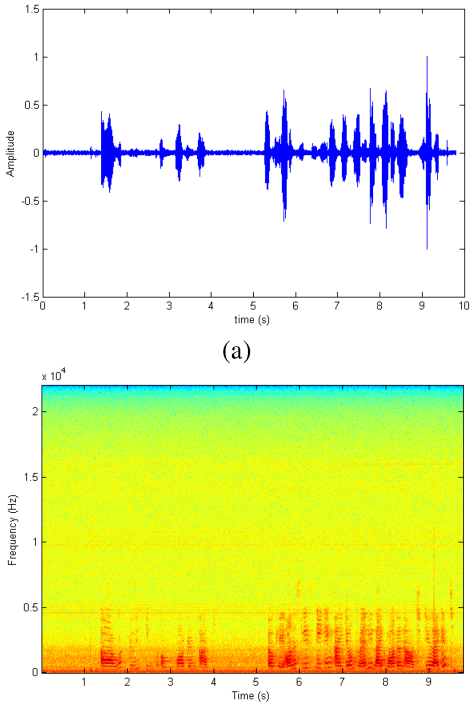


Figure 2. Input speech signal digit file. (a) Input speech signal. (b) Spectrogram of input speech signal.

IV. Evaluation of Proposed Algorithm

The level of security is determined on the degree of residual intelligibility of any encryption scheme. Extensive simulation experiments were conducted using the proposed scheme. These simulations were carried out on different spoken speech data samples including digits, sentences and conversational speech segments. Samples are recorded both by male and female voices and used as input for analysis; their outputs are stored for evaluation of proposed scheme. shows the spectral information of decrypted signal, which are identical to the original speech signal data.

A. Speech Intelligibility Test

Speech intelligibility and quality of recovered speech are largely subjective quantity, so to evaluate the security level of the proposed scheme; several subjective tests were carried out. Furthermore, most widely used intelligibility test is the Diagnostic Rhyme Test (DRT). In DRT, each rhyme group word is presented to the listener and asked to recognize the

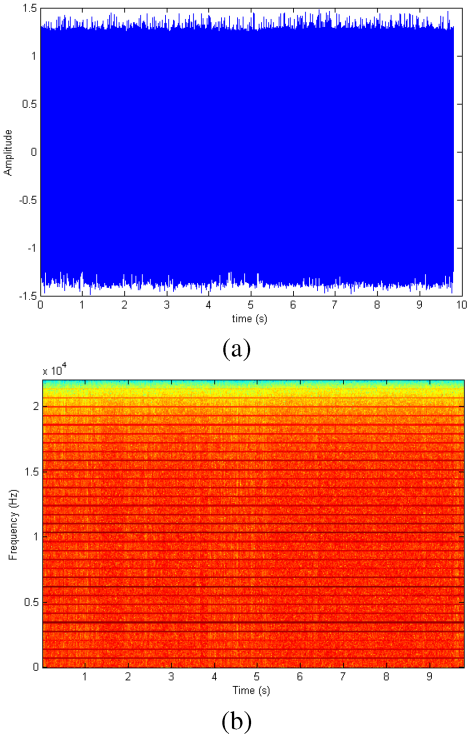
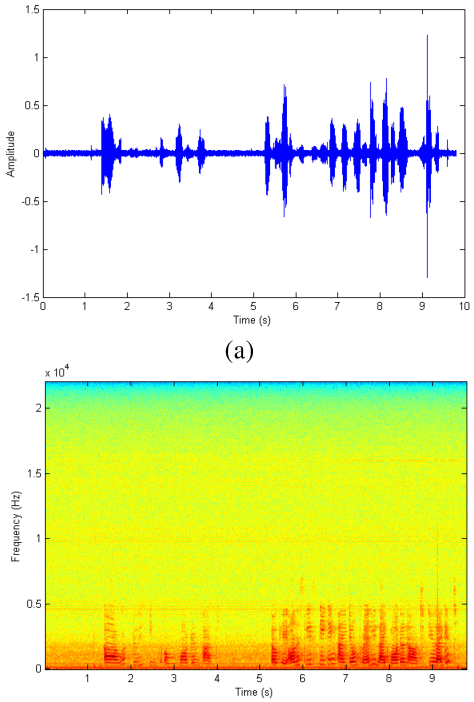


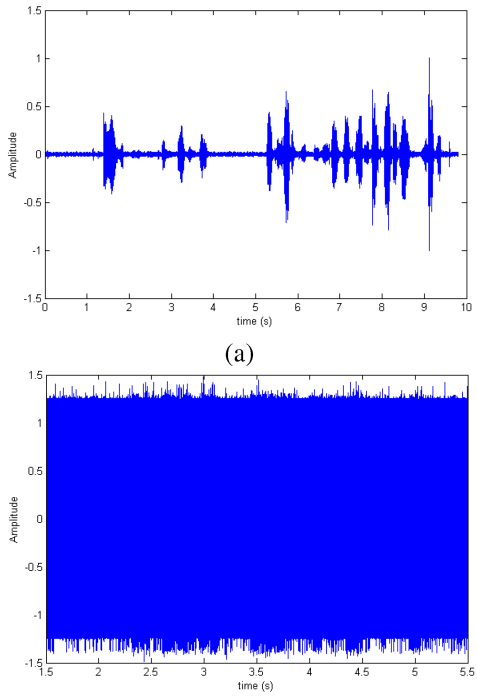
Figure 3. Input speech signal digit file. (a) Scrambled speech signal. (b) Spectrogram of scrambled speech signal.



(a)

(b)

Figure 4. Descrambled speech signal digit file. (a) Descrambled speech signal. (b) Spectrogram of descrambled speech signal.



(a)

(b)

Figure 5. Key Sencitivity Test. (a) Original input speech signal digit file. (b) Recovered speech signal with one-bit key difference.

word. The score of DRT is calculated in terms of correctly identified word Q, which is equal to:

$$Q = \frac{(R-W)}{T} \times 100 \quad (1)$$

In the above Eq. 1, R represents right and W represents wrongly identified answers, whereas T represents total number of words under consideration. Generally, DRT score falls in range of 75 to 95. The intelligibility scores categories are defined as: 95-100 (excellent), 87-95 (good), 79-87 (fair), and 70-79 (poor) and below 70 score is regarded as bad or unacceptable [24]. The average results of tested speech samples are summarized in table I.

B. Correlation

Correlation analysis is one of the most elementary method used in computing the degree of similarity between two quantities [25], [26]. Correlation Coefficient (CC) is one of the useful measure to judge encryption quality of any cryptosystem [27], [28]. The correlation coefficients between adjacent samples of original and encrypted speech signals were tested. Table II compares the correlation coefficients of original and encrypted speech samples. In case of original speech data, the correlation coefficient is 0.982 and for encrypted speech samples -0.028, which are significantly far apart. The value of the correlation coefficient for original speech sample is closer to 1, which shows maximum correlation whereas -0.028 shows that the encrypted speech sample had severely low correlation to the original speech samples.

C. Key Space Analysis

A strong encryption scheme is one which should be more sensitive to the ciphering keys [15]. The key space analysis tests for the proposed encryption algorithm are summarized in the following section.

1) Key Sensitivity Test: A good encryption scheme should be sensitive to the secret keys [29]. Here, we verify the sensitivity of proposed scheme to keys using Matlab 2012*b*. Assume an intruder intercepting both the scrambled and synchronization signals. Suppose the attacker gets an approximate estimate of key, in which there is only a slight mismatch that is one bit only, with the real key. Figures 5 show the sensitivity of the proposed secure communication algorithm to a small mismatch of the key. As one can verify, the recovered speech Fig. 5(b) with wrong key behaves randomly and completely dissimilar from the original speech Fig. 5(a). Undoubtedly, the secret keys are secure enough even if a chosen plaintext/ciphertext attack is adopted.

2) Exhaustive Key Search: For any encryption scheme, the key space is $2k$, where k is the key size in bits. Higher the size of the key better will be the achieved security. Key size should be large enough to resist all types of brute force attacks. An exhaustive key search will take $2k$ operations to succeed. The proposed algorithm uses a 64-bit key length. An attacker would need 264 operations to successfully determine the actual key. Let's assume, an intruder employs a thousand million instructions per second (1000MIPS) computer for guessing the

key via brute force attack, the computational load in years is 584.9424172 years. This is a very long time which is practically infeasible.

TABLE I. SUMMARY OF AVERAGE RESULTS OBTAINED THROUGH DRT

Speech sample	Original	Encrypted	Decrypted
Digits	100%	0%	99.86
Sentences	100%	0%	99.89
Conversation	100%	0%	99.9

TABLE II. CORRELATION COEFFICIENT OF ORIGINAL AND SCRAMBLED SAMPLES

Speech sample	Original signal	Encrypted signal
Digits	0.982	-0.016
Sentences	0.986	-0.022
Conversation	0.974	-0.028

V. CONCLUSION

In this paper, a speech encryption scheme is proposed and implemented. The DCT coefficients of speech signal are permuted via random number generator using TD-ERCS chaotic map. The results based on simulations and cryptanalysis transpires promising results thereby authenticating the level of security of the proposed algorithm. The results which are demonstrated through various Figures, one can see that the encrypted signal has 0% intelligibility. In addition, an excellent quality of descrambled/recovered speech signal can also be observed. It can be easily demonstrated that the frequency domain scrambling of signal frames are equivalent to linear combination of the samples of the corresponding frames in time domain. Consequently, the scrambled speech signal cannot be decrypted by an inverse permutation in the time-domain and this property gives an extra layer of protection against cryptanalytic approaches to the scheme. Also, as the transmission is non-digital, so the proposed technique does not put a stringent requirement on channel bandwidth. In future, symmetric key algorithm can be changed to asymmetric key algorithm.

REFERENCES

[1] A. V. Vanevska and V. Trajkovik, "Secure transport of electronic payment over gprs," in *ICT Innovations 2011*. Springer, 2012, pp. 129–146.

[2] R. Sutton, *Secure communication: applications and management*. John Wiley & Sons, 2002.

[3] J. Khan, J. Ahmad, and S. O. Hwang, "An efficient image encryption scheme based on: Henon map, skew tent map and s-box," in *Modeling, Simulation, and Applied Optimization (ICMSAO), 2015 6th International Conference on*. IEEE, 2015, pp. 1–6.

[4] J. S. Khan, J. Ahmad, and M. A. Khan, "Td-ercs map-based confusion and diffusion of autocorrelated data," *Nonlinear Dynamics*, pp. 1–15, 2016.

[5] E. Mosa, N. W. Messiha, and O. Zahran, "Chaotic encryption of speech signals in transform domains," in *Computer Engineering & Systems, 2009. ICCES 2009. International Conference on*. IEEE, 2009, pp. 300–305.

[6] Y. Wu and P. BOOM, "Speech scrambling with hadamard transform in frequency domain," in *6th International Conference on signal processing*, vol. 1563, 2002.

[7] B. Goldburg, S. Sridharan, and E. Dawson, "Design and cryptanalysis of transform-based analog speech scramblers," *IEEE Journal on Selected areas in Communications*, vol. 11, no. 5, pp. 735–744, 1993.

[8] D. Tseng and J. Chiu, "An ofdm speech scrambler without residual intelligibility," in *TENCON 2007-2007 IEEE Region 10 Conference*. IEEE, 2007, pp. 1–4.

[9] R. Matthews, "On the derivation of a chaotic?encryption algorithm," *Cryptologia*, vol. 13, no. 1, pp. 29–42, 1989.

[10] J. S. Khan, A. ur Rehman, J. Ahmad, and Z. Habib, "A new chaos-based secure image encryption scheme using multiple substitution boxes," in *2015 Conference on Information Assurance and Cyber Security (CIACS)*. IEEE, 2015, pp. 16–21.

[11] Z.-H. Guan, F. Huang, and W. Guan, "Chaos-based image encryption algorithm," *Physics Letters A*, vol. 346, no. 1, pp. 153–157, 2005.

[12] J. Khan, J. Ahmad, and S. O. Hwang, "An efficient image encryption scheme based on: Henon map, skew tent map and s-box," in *Modeling, Simulation, and Applied Optimization (ICMSAO), 2015 6th International Conference on*. IEEE, 2015, pp. 1–6.

[13] S. Lian, J. Sun, and Z. Wang, "A block cipher based on a suitable use of the chaotic standard map," *Chaos, Solitons & Fractals*, vol. 26, no. 1, pp. 117–129, 2005.

[14] A. U. Rehman, J. S. Khan, J. Ahmad, and S. O. Hwang, "A new image encryption scheme based on dynamic s-boxes and chaotic maps," *3D Research*, vol. 7, no. 1, pp. 1–8, 2016.

[15] J. Ahmad, M. A. Khan, S. O. Hwang, and J. S. Khan, "A compression sensing and noise-tolerant image encryption scheme based on chaotic maps and orthogonal matrices," *Neural Computing and Applications*, pp. 1–15, 2016.

[16] V. Patidar, N. Pareek, and K. Sud, "A new substitution–diffusion based image cipher using chaotic standard and logistic maps," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 7, pp. 3056–3075, 2009.

[17] L. Kocarev, "Chaos-based cryptography: a brief overview," *Circuits and Systems Magazine, IEEE*, vol. 1, no. 3, pp. 6–21, 2001.

[18] N. K. Pareek, V. Patidar, and K. K. Sud, "Image encryption using chaotic logistic map," *Image and Vision Computing*, vol. 24, no. 9, pp. 926–934, 2006.

[19] X. Tong and M. Cui, "Image encryption scheme based on 3d baker with dynamical compound chaotic sequence cipher generator," *Signal processing*, vol. 89, no. 4, pp. 480–491, 2009.

[20] S. L.-Y. S. Ke-Hui and L. Chuan-Bing, "Study of a discrete chaotic system based on tangent-delay for elliptic reflecting cavity and its properties [j]," *Acta Physica Sinica*, vol. 9, p. 011, 2004.

[21] J. Ahmad and S. O. Hwang, "A secure image encryption scheme based on chaotic maps and affine transformation," *Multimedia Tools and Applications*, pp. 1–26, 2015.

[22] L. Zeng, X. Zhang, L. Chen, Z. Fan, and Y. Wang, "Scrambling-based speech encryption via compressed sensing," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 1, 2012.

[23] Y. Mao and M. Wu, "A joint signal processing and cryptographic approach to multimedia encryption," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2061–2075, 2006.

[24] J. H. Hansen and B. L. Pellom, "An effective quality evaluation protocol for speech enhancement algorithms." in *ICSLP*, vol. 7, 1998, pp. 2819–2822.

[25] J. Ahmad and F. Ahmed, "Efficiency analysis and security evaluation of image encryption schemes," *International Journal of Video & Image Processing and Network Security*, vol. 12, no. 4, pp. 18–31, 2012.

[26] J. Ahmad and S. O. Hwang, "Chaos-based diffusion for highly auto-correlated data in encryption algorithms," *Nonlinear Dynamics*, vol. 82, no. 4, pp. 1839–1850, 2015.

[27] J. Ahmad, S. O. Hwang, and A. Ali, "An experimental comparison of chaotic and non-chaotic image encryption schemes," *Wireless Personal Communications*, pp. 1–18, 2015, 10.1007/s11277-015-2667-9.

[28] "An experimental comparison of chaotic and non-chaotic image encryption schemes," *Wireless Personal Communications*, vol. 84, no. 2, pp. 901–918, 2015.

[29] M. A. Khan, J. Ahmad, Q. Javaid, and N. A. Saqib, "An efficient and secure partial image encryption for wireless multimedia sensor networks using discrete wavelet transform, chaotic maps and substitution box," *Journal of Modern Optics*, pp. 1–10, 2016.