Journal of Advances in Computer Research
Quarterly pISSN: 2345-606x eISSN: 2345-6078
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 7, No. 3, August 2016), Pages: 67-76
www.jacr.iausari.ac.ir

Audio Signal Encryption Based on Permutation Relations and Residue Number System

Fereshteh Ghanbari Adivi✉, Mohammad Mehrnia
M.Sc. Computer Systems Architecture, Dezfoul Branch, Islamic Azad University, Dezfoul, Iran
ghanbary.fereshteh@yahoo.com; qermezkon@gmail.com

Received: 2016/02/28; Accepted: 2016/05/14

Abstract

This paper presents a new audio encryption algorithm based on permutation and residue number system. In the current approach, signal samples are combined with residue number system through permutation; therefore, a specific noise is generated. Then, the audio signal is combined with the noise, generated by the algorithm, to produce a silence signal. For this purpose, the algorithm uses a predefined permutation format and a modular set in the form of $\{2^n-1, 2^n, 2^n+1\}$ to encrypt the main signal. Consequently, the proposed algorithm converts the input signal into a silence signal and increases its security. The signal-to-noise ratio (SNR) was used to evaluate the algorithm. The results indicated very low values of SNR pertaining to audio signals in the proposed method.

Keywords: *Encryption, Decryption, SNR, Residue Number System, Audio Signal, Permutation*

1. Introduction

Voice is the ability to speak or the act of speaking. An audio signal is comprised of a number of bits. This signal contains positive and negative values. Audio communications are increasingly used in a way that they are becoming more vulnerable; therefore, it is really necessary to provide them with more security. Various encryption algorithms are used to provide the security of audio signals. In this regard, there are five general encryption methods such as scrambling time domain, scrambling amplitude, 2D scrambling and the combination of scrambling frequency and time domain. An audio signal is encrypted in digital or analog formats [1].

Generally, the objective of speech encryption is to perform some reversible operations on a piece of speech in a way that it will be totally incomprehensible to an unauthorized listener. Three important criteria for rating encryption algorithms are as follows:

- The ability of algorithm to generate an encrypted speech whose comprehensible residue is little.
- The effectiveness of encryption and decryption process on the quality of retrieved information on the side of the desired recipient.
- The security of algorithm against cryptanalysis attacks

The problem that encryption experts are faced with is to design encryption systems that distort speech signals having so much redundancy that it would be impossible to retrieve useful information from them.

Using a predefined permutation mechanism and a residue number system with modules in such a format as $\{2^n\text{-}1, 2^n, 2^n\text{+}1\}$, the proposed approach encrypts audio signals. The encrypted signal will be like a silence signal with no loss of quality during decryption.

Employing the proposed algorithm, audio files can be encrypted in order to save audio messages and transmit them on the Internet. This system has two prominent advantages. First, the encrypted signal will be heard like a silence signal. The second advantage is that relocating modules, changing their values, and changing the permutation pattern can result in so many states in the encryption algorithm. Using SNR, the proposed algorithm was evaluated. It presented a high level of security with appropriate quality. The current paper is organized as follows: Section 2 presents an introduction to the residue number system. In Section 3, some instances of the previous algorithms are presented on voice encryption. The encryption process is explained in Section 4. The results of analyzing and evaluating the performance of the algorithm are presented in Section 5. Finally, the conclusion is discussed in Section 6.

2. Residue Number System

The residue number system is a non-weighted and unconventional number system that supports parallel computations, the limited distribution of carry, low power consumption, and secure communications. It is used in applications employing addition, multiplication, subtraction, and multiplication by a range of numbers. The residue number system is characterized with a set of modules. If all the modules are co-prime, the system is optimal. The hybrid of other systems can be used to optimize this system so that the selection of modular sets would become easier, and the display range extended. For this purpose, the hybrid of the residue number system, multilevel residue number system, and the Van Hunt residue number system can be employed. Moreover, it is suggested that the redundant residue system be used to establish and increase system security. The residue number system is characterized with a set of modules like $\{m_1, m_2, \ldots\ldots, m_n\}$ in which all the modules are positive integers. If all *n* modules are co-prime, the system has the largest possible display range equal to $[\alpha, \alpha+M)$ in which α is an integer and M is calculated as follows:

$$\boldsymbol{M} = \prod_{\boldsymbol{i}=1}^{\boldsymbol{n}} \boldsymbol{m_i} \tag{1}$$

The integer X ($\alpha \leq X < \alpha + M$) is a unique representation in the residue number system. It is indicated with a set of remainders like $(x_1, x_2, x_3, \ldots, x_n)$ in a way that:

$$x_i = X \ mod \ m_i, \ i = 1, 2, \ldots..., \ n \tag{2}$$

Using the Chinese Remainder Theorem, X is obtained on the remainder set $\{m_1, m_2, \ldots, m_n\}$ as follows:

$$\boldsymbol{X} = \left(\sum_{\boldsymbol{n}=1}^{\boldsymbol{n}} \left(\boldsymbol{x_i N_i}\right)_{\boldsymbol{m_i}} \times \boldsymbol{M_i}\right)_{\boldsymbol{m}} \tag{3}$$

$$M = \prod_{i=1}^{n} m_i \tag{4}$$

$$M_i = \frac{M}{m_i}, N_i = \left(M_i^{-1}\right)_{m_i}, i = 1,2,3,K,n \tag{5}$$

In which $(M_i^{-1})_{m_i}$ is defined as the multiplicative inverse of M_i on module m_i.

Given the characteristics of the residue number system, it is used in computational applications such as digital signal processing system, digital filters, coding, encryption systems, digital communications, ad hoc networks, information storage and retrieval, and error detection and correction. Moreover, since the computations are done separately on remainders in this system, if an error occurs on one of the remainders, others are immune to its impact.

3. Previous Works

Encryption may be performed in either a digital or analog way. In analog encryption, which is also called speech scrambling, the operations are done on the speech samples. In analog scrambling, there is no need for a modem or speech compression to transmit. The quality of retrieved speech does not depend on its language. Using an appropriate interface, these scrambling methods can be easily connected to the available analog channels such as telephone, satellite, and mobile communication links. The digital encryption digitalizes the input speech signal first. Then the digitalized signal is compressed in order to generate a bit sequence with an appropriate bit rate. Using modem, this bit sequence is encrypted and transmitted to the channel.

Analog scrambling is divided into several classes:

- Time scrambling
- Frequency scrambling
- Scrambling with pseudo-noise sequences
- Scrambling with maximal-length sequences
- Scrambling with gold sequences
- Scrambling with pseudo-gold sequences
- Scrambling with Barker sequences
- Scrambling with pseudo-Barker sequences
- Scrambling with Kasami sequences

There are also 7 different styles defined for the cryptography of audio signals.

1- There is no need for cryptography.

2- The quality of speech is not high while the required security level is so low.

3- The required security is low while power consumption and time delay are very low.

4- The required security is higher than Case 3 while more power and time are consumed.
5- The required security is average, and the consumed time and power are high.
6- The security is more than Case 4.
7- The security and time delay are higher.
8- The sender emphasizes on the high security of its audio streaming without considering other factors (power consumption and time delay).

Two instances of audio encryption algorithms based on the abovementioned mechanisms are explained in what follows.

In [1], a four-level method was presented for audio signal encryption. In this method, the output of each level is used as the input of the next. In algorithm [1], changing the formation of bits at the first level, using a random generator at the second level, using it again at the third level, and, finally, using the ascending formation of the amplitude field at the fourth level result in the generation of an encrypted signal like a continuous beep. authors in [3], propose a simple and effective detection algorithm which detects the time-frequency (TF) points occupied by only a single source for each source. The detection algorithm identifies the single source points by comparing the normalized real and imaginary parts of the TF coefficient vectors of the mixed signals, which is simpler than previously reported algorithms. Then we propose a modified similarity-based robust clustering method (MSCM) to estimate the number of sources and the mixing matrix using these detected single source points. Experimental results show the efficiency of the proposed algorithm, especially in the cases where the number of sources is unknown. In [2], the Independent Component Analysis (ICA) of the audio signal was used to propose a method for encryption. In this method, the audio signal is divided into several parts. The signal division process is in a way that each part has an equal number of samples that will be independent from each other so that the successful use of ICA would be guaranteed. In the next step, the correlation matrix is generated, and then each part is encrypted through the process of non-correlation, named whitening in this paper. The whitening process uses the Principle Component Analysis (PCA) of the dimensionality reduction algorithm. the authors in [11], present an RNS algorithm resolving the Closest Vector Problem (CVP). This algorithm is particularly efficient for a certain class of lattice basis. It provides a full RNS Babai round-off procedure without any costly conversion into alternative positional number system such as Mixed Radix System (MRS). An optimized Cox-Rower architecture adapted to the proposed algorithm is also presented. The main modifications reside in the Rower unit whose feature is to use only one multiplier. This allows freeing two out of three multipliers from the Rower unit by reusing the same one with an overhead of 3 more cycles per inner reduction. An analysis of feasibility of implementation within FPGA is also given.

The authors in [12] demonstrate the efficiency of approach using 32 chosen eigenvalues in the key generation algorithm. The eigenvalues are derived from a reference image. Throughout the processes on the image data, we use modular arithmetic to ensure that computations with the resulting RNS become very efficient. Moreover, the approach considers image in a divided matrix domain and finally combines all independent cryptographic operations as encryption is a one-to-one

mapping. This deals with the possibility of having any pixel value ill-stored or wrongly received at the receiver end, without affecting the decryption process. However, the final recovered image will differ by a negligible amount. the authors in [13], present a speech encryption algorithm, which can penetrate compression encoder. The algorithm scrambles the speech signal in frequency domain based on FFT transformation domain in the sender, and it decrypts and reconstructs the signal in the receiving end. The encrypted speech signal still has the characteristics of speech, the residual intelligibility of the speech is low enough and the speech bandwidth does not be widened. Scramble matrix can meet the needs of speech transmission security and achieve one key for one secret communication flexibly. End-to-end communication can avoid the common eavesdropping and theft, and has a relatively broad prospects for development. In [14], a robust Audio steganography technique is proposed by randomizing and dynamically changing the embedding sequence. Advance Encryption standards is used for providing additional security and robustness to the algorithm and tested for the 30 speech files. The addition of cryptography in steganography increases the robustness and introduces a higher level of security since the key is required to decrypt the secret message. To evaluate the quality of the stego files SNR and Correlation coefficients are calculated and an experimental test is also performed over 10 listeners to identify the change in the original audio and cover audio. the authors in [15], presents an analysis of the tradeoff between security level and compression ratio in real-time voice communication. It is hypothesized that the combination of variable bitrate compression and same length encryption will induce vulnerability to traffic analysis. The variation of packet sizes can leak information about the conversation starting with language identification, identifying certain phrases, and reconstructing phonemes. The proposed solution to this problem is to rely on constant bitrate compression or to pad the sent frames to a multiple of 16, 32, or 64 bytes. In [16], a symmetric key is developed which consists of reshuffling and secret arrangement of secret signal data bits in cover signal data bits. In this paper the authors have performed the encryption process on secret speech signal data bits-level to achieve greater strength of encryption which is hidden inside the cover image. The encryption algorithm applied with embedding method is the robust secure method for data hiding.

4. The Proposed Algorithm

This part presents the audio signal encryption based on the permutation of the residue number system as the proposed algorithm. This algorithm can be executed on both stereo and mono signals for encryption. Three different encryption keys are defined in the current approach. The first encryption includes the values of modular components; the second encryption is the formation order of modules; and the third formation key is the permutation pattern. The relations of permutation pattern are multiples of the number of modules. It is obvious that the more the number of modules are, the more the permutation relations become; therefore, the coefficient of signal security soars. However, if the number of modules and the permutation relations of the signal increase, the overhead resulting from their application to the signal becomes heavier; thus, the output signal volume increases. In the proposed algorithm, the permutation relations scramble the elements composing signal. They are scrambled in an organized way due to the predefined relations. Having the encryption keys, the encrypted signal can be decrypted.

A set of modules like $\{2^n\text{-}1, 2^n, 2^n\text{+}1\}$ was used in the current approach. Given the fact that there are 3 members in this set, the permutation relations will have 9 rules. They are closely associated with the number of members in the set of modules. These relations can be the number of modular sets by the power of 2 so that the signal encryption process could be based on them. The relations pertaining to a modular set like $\{2^n\text{-}1, 2^n, 2^n\text{+}1\}$ are demonstrated in the following table.

Table 1: Permutation Relations of the Proposed Algorithm

Output	input
Y[n]	$\boldsymbol{x[n+1]\|_{2^n+1} + (2^n + 1)}$
Y[n+1]	$\boldsymbol{x[n-1]\|_{2^n-1} + (2^n - 1)}$
Y[n+2]	$\boldsymbol{x[n]\|_{2^n} + (2^n)}$
Y[n+3]	$\boldsymbol{x[n+1]\|_{2^n-1} + (2^n - 1)}$
Y[n+4]	$\boldsymbol{x[n-1]\|_{2^n+1} + (2^n + 1)}$
Y[n+5]	$\boldsymbol{x[n]\|_{2^n-1} + (2^n)}$
Y[n+6]	$\boldsymbol{x[n-1]\|_{2^n} + (2^n)}$
Y[n+7]	$\boldsymbol{x[n+1]\|_{2^n} + (2^n)}$
Y[n+8]	$\boldsymbol{x[n]\|_{2^n+1} + (2^n + 1)}$

The definition of the permutation relations should also have special relationships so that the encryption process would be done free of errors. These relationships are as follows:

1- The number of relations should be the number of modules by the power of 2.

2- The remainder of each element of the input signal on all the modules should exist in permutation relations.

In Table 1, the values of x[n] refer to the input signal samples, and the values of Y[n] represent the output signal samples resulting from applying the permutation relations and the residue number system. The relations are applied in a way that the first sample and the last three samples were not changed, and the relations were applied to the rest of the samples. That the first sample and the last three samples were not changed resulted in the correct execution of relation. The encrypted signal is defined according to

$$EncryptedSignal = [x[1]\ Y[1]\ldots Y[n]\ x[end-3]\ x[end-2]\ x[end-1]] \quad (6)$$

Given the regulated scrambling of audio signal samples, signal decryption based on the permutation relations, shown in Table 1, and the modular set will be a quick, efficient process. An instance of signal form change based on the proposed encryption signal can be seen in Figures (1) and (2).

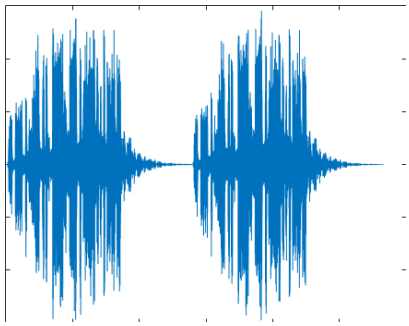

Figure 1. The Main Signal (not Encrypted Signal)

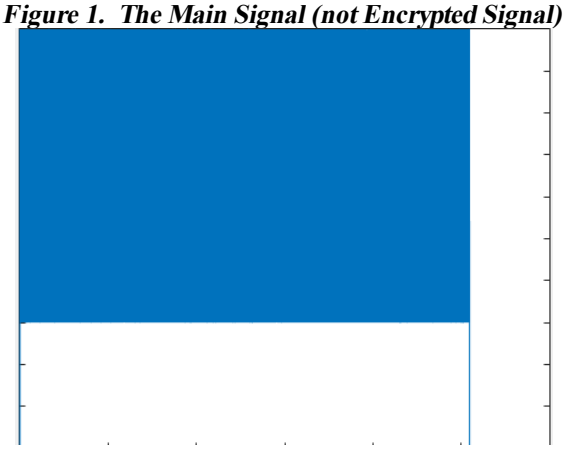

Figure 2. The Encrypted Signal: This Signal is Encrypted by Algorithm and Convert to Silent Signal with same frequency

According to Figures (1) and (2), there is no sign of the main signal samples in the encrypted signal which is like a signal with a constant frequency. Since different encryption parameters are used in the algorithm, it can adapt to every condition related to the transfer medium. Increasing or decreasing the number of modules and changing the number of permutation relations or the type of relations, it can also increase the security coefficient or decrease the overhead resulting from the encryption process. However, it is necessary to establish a balance between the number of modules and permutation relations and the type of relations in order to propose an optimal algorithm. The decryption of audio signals is another problem of audio signal encryption algorithms. Such an algorithm is considered efficient when the decrypted signal has the same quality as the main signal, or the quality is close to the main one.

The quality of decrypted signal is very important in the proposed algorithm; therefore, efforts were made to decrypt the signal with the best quality. Considering the fact that the residue number system was used to decrypt the audio signal, it is necessary that the remaining value of modules be changed to the initial values after permutation through the Chinese Remainder Theorem. The signal decryption process has two general steps:

- Finding the multiplicative inversion for the module used in the encryption algorithm based on

  equ. 3, equu. 4 and equ. 5
- Using Relation equ. 3 and the permutation relations defined in Table 1 to decrypt the encrypted signal

5. Performance Evaluation

There are no absolute measures to evaluate noise, and all noises are measured in a relative way. The signal-to-noise ratio (SNR) is used to measure the signal noise. The values of SNR that are smaller than 12 indicate that the signal has much noise; however, the values above 30 mean that the signal has very little noise.

$$SNR = 10log_{10} \frac{\sum_{N=\infty}^{\infty} s^2(n)}{\sum_{\infty}^{\infty} \left[s(n) | - \hat{n}(n) \right]^2} \quad (7)$$

Matlab was used to simulate the signal encryption and decryption algorithm. The pseudo-code pertaining to the encryption algorithm can be seen in Figure 3:

```
                                                    [Y,Fs]= [Read Audio Signal]
Y=Y(:,1)+Y(:,2);
Y=Y';
Resultn=1;
for n=2:numel(Y)-3
   result(Resultn)=mod(Y(n+1),5) + 5;
   result(Resultn+1)=mod(Y(n-1),3) + 3;
   result(Resultn+2)=mod(Y(n),4) + 4;
   result(Resultn+3)=mod(Y(n+1),3) + 3;
   result(Resultn+4)=mod(Y(n-1),5) + 5;
   result(Resultn+5)=mod(Y(n),3) + 3;
   result(Resultn+6)=mod(Y(n-1),4) + 4;
   result(Resultn+7)=mod(Y(n+1),4) + 4;
   result(Resultn+8)=mod(Y(n),5) + 5;
   Resultn = Resultn + (3^2);
   n=n+3;
end
result = [Y(1) result Y(end-3) Y(end-2) Y(end-1)];
```

Figure 3: Algorithm Source Code in Matlab

In Table 2, SNR was measured for 7 different audio signals that were converted into an encrypted signal through the encryption algorithm. According to this table, the values of SNR are very low, a fact that indicates the high power of the proposed encryption algorithm. This table can also lead to the inference that the proposed algorithm can deal with every number of samples correctly and encrypts the signal at the highest level of security.

Table 2: SNR after Encrypting Audio Signals

Different Signals	The Number of Samples	SNR
Signal 1	114340	**-0.0098**
Signal 2	283136	**-4.2184e-07**
Signal 3	122868	**-5.8517e-05**
Signal 4	2548212	**6.2069e-07**
Signal 5	88848	**-9.0321e-07**
Signal 6	102201	**1.6363e-05**
Signal 7	182050	**-6.73056e-08**

6. Conclusion

In the current approach, an algorithm based on permutation relations and the residue number system was proposed for audio signal encryption. The results indicated that the proposed algorithm had a high level of security and audio quality in decryption. Using the residue number system, a three-member modular set, and 9 permutation relations in this algorithm, the initial signal was encrypted. Therefore, the cryptanalysis was difficult, and the security of audio signal increased. For decryption, it was necessary to take the permutation relations into account first in order to reconstruct the signal. Based on the Chinese Remainder Theorem, the values were then converted into the initial samples of signal. Therefore, a good quality was presented after decrypting the encrypted audio signal.

References

[1] Harjinder Kaur, 2012. A four level speech signal encryption YCSC, vol3, pp 151-153

[2] Sttar B. Sadkhan, Nidaa A. Abbas, 2014, A proposed voice encryption System Base on off-line ica Algorithm, Journal Babylon university/Pure and Applied Sciences/No (7) Vol (22)

[3] Tianbo Dong, Yingke Lei & Jinhshu Yang, 2012, An Algorithm for Underdetermined Mixing Extimation, Elsevier, Neurocomputing(pp-1-9)

[4] Mosa, 2009, Random encryption of speech signal, Computer Engineering & Systems. ICCES International Conference

[5] Mosa, 2009, Chaostic encryption of speech signals in transform domains, Computer Engineering & Systems, ICCES International Conference

[6] Anwar, S.; Alam, S.B.; Rahman, K.M.S, 2010, Information Theory and Information Security (ICITIS), IEEE International Conference on P.446-451

[7] Bashier, E.B.M., 2013, Speech scrambling based on chaotic maps and one-time pad, Computing, Electrical and Electronics Engineering (ICCEEE), International Conference on Pp. 128-133

[8] Patil, S.B., 2015, A robust encryption method for speech data hiding in digital images for optimized security, Pervasive Computing (ICPC), International Conference P.1-5

[9] Kohata, M, 2014, Secure Speech Encryption System Using Segments for Speech Synthesis, 2014, Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), Tenth International Conference, P.264-267

[10] Radha, 2011, Comparative analysis of compression techniques for Tamil speech datasets, Recent Trends in Information Technology (ICRTIT), International Conference on P 712-716

[11] J. C. Bajard, 2015, RNS Arithmetic Approach in Lattice-Based Cryptography: Accelerating the "Rounding-off" Core Procedure, Computer Arithmetic (ARITH) IEEE 22nd

[12] G. K. Armah, 2015, Application of residue number system (RNS) to image processing using orthogonal transformation, 2015, Communication Software and Networks (ICCSN), IEEE International Conference

[13] Yaoyao Chen, 2015, End-to-end speech encryption algorithm based on speech scrambling in frequency domain, Third International Conference on Cyberspace Technology on Pp. 1 – 5

[14] A. Kanhe, 2015, Robust Audio steganography based on Advanced Encryption standards in temporal domain, 2015, Advances in Computing, Communications and Informatics (ICACCI), International Conference on P.p 1449 – 1453

[15] Attie, 2015, Analysis of the tradeoff between compression ratio and security level in real-time voice communication, Applied Research in Computer Science and Engineering (ICAR), 2015 International Conference on, Pp. 1 - 6

[16] S. A. Kulkarni, 2015, A robust encryption method for speech data hiding in digital images for optimized security, Pervasive Computing (ICPC), 2015 International Conference on Pp. 1 - 5