Implementation of improved S_Affine cipher algorithm for digital audio cryptography

Ass.Prof.Dr. Saad Abdual azize Adbdual Rahman
saadabdualazize@yahoo.com
Al Mamoon University Collage

Ass.Prof.Dr.Sana Ahmed Kadhim
dr.sanaa.ahmed@uotic.edu.iq
University of information technology and communications

Ass.Prof. Emad Matti Bako
imad.m.bko@almamonuc.edu.iq
Al Mamoon University Collage

Abstract

Cryptography is the best arrangement that compromises the imperative assurance from accidental people. Utilization of encryption and unraveling systems, cryptography may be changed over the information structure its lucid structure to indistinguishable one with the goal that main the planned recipient can peruse the letter and modify it. With these lines, any party may guarantee the authentication of a message that to be transmitted or put away with next to no alteration. The introduced work is worried about a specific kind of lopsided key cryptography called S_Affine cipher calculation to be utilized for discourse samples encoding and decoding. The discourse samples are scrambled by utilizing S_Affine cipher calculation all through encoding technique applied by the transporter and afterward communicated to the specific receiver. Then, the receiver will recuperate the first sign all through the unscrambling system.

Keywords: Improved S_Affine algorithm, Cryptography, encoding; decoding; audio signals; residual intelligibility.

1- Introduction

Discourse main correspondence became famous within many new applications such as: video chatting, securities exchange, military areas, electronic banking and learning in addition to many different areas so forward. With the ascending of media and organization advancements these days, secrecy and confidentiality are significant aspects through many applications to safeguard information during capacity or transformation. A major method of information shielding from demolition and unauthentic manipulation in weak transmission channels is encoding. The investigation of data security is known as cryptography that is

connected with various angles like validation, uprightness and protection or certainly. The changings in the sets of information in a way disguising or scrambling data is characterized as cryptography. Encryption is used to imparted information or letters to create information or code, in another hand, unscrambling is used to recreate the first information from the code text by means of mystery "key" that is divided among the sender and collector. Discourse audio encoding is created by adding specific commotion with the quiet piece of the discourse audio prior to transmission, and the recipient will extricate the code to recuperate the first discourse audio [1]. Numerous methods are used to organize cryptography frameworks yet the well-known methods of organization types are symmetric and asymmetric cryptography [2]. The one key encryption method which is also called symmetric key encryption is the most known and used strategy which uses a single key for both sides of transmission that might be a phrase, a number, or sequence of words to be applied for both encryption and decryption [3]. Some of generally utilized standard symmetric key cryptography systems are DES (standard information encryption), Triple Information Encryption Calculation (TDEA), AES (High level), Blowfish and RC4 [4].

2- Affine algorithm

Relative code is an expansion of the Caesar figure, which duplicates the plaintext with a worth and add a shift [5][6]. Numerically encryption of plaintext P to deliver ciphertext (input signal) C can be communicated by harmonious capabilities as follows:

$$C = E(P) = (h*p + g) \bmod s \qquad \ldots..(1)$$

Where: s = size of input signal

h = a number that should be generally prime to s (While perhaps is not prime).

g = the quantity of movements (Caesar figure is the specialty of relative code with s=1).

p = plaintext is switched over completely to a number from 0 to s-1 in respect to the input signal.

E (P) = the encoded data.

While the decoding process can be composed utilizing the accompanying condition: $D(t) = h^{-1}(t - g) \bmod s$ where h^{-1} is the multiplicative opposite modulus 's' to meet the accompanying condition.

$$1 = h*h^{-1} \bmod s \qquad \ldots\ldots(2)$$

a multiplicative reverse exists provided that h and s are coprime. Unscrambling capability is the opposite of the encryption capability that can be composed as follows:

$$D(E(P)) = h^{-1}(E(P) - g) \bmod s$$
$$= h^{-1}(((h*p + g) \bmod s) - g) \bmod s$$
$$= h^{-1}(h*p + g - g) \bmod s$$
$$= h^{-1} h*p \bmod s$$
$$D(E(p)) = p \bmod s$$
$$= p$$

3- Modified S_Affine Cipher

The general encoding strategy based on replacing the plain letters with their relative codes, the letters are planned to numerical order(position), encoded utilizing a straightforward numerical capability, then decoded back completely to their original shape(letters). The recipe utilized implies the encoding of each letter to another letter, then be back once more, meaning the technique is basically a standard replacement with a standard overseeing each letter matches to which figure.

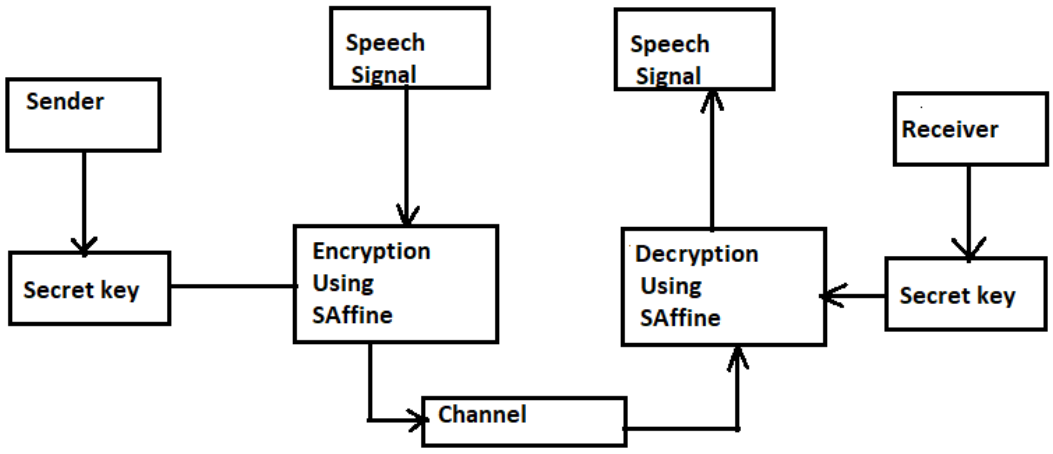

Figure 1: Block diagram for proposed system cryptography

The entire interaction depends on applying modulo y (the number of letters in order utilized). In replacement encoding, each letter in a letter set of length (size) equal "m" are planned to a number within the reach (0 … y-1).

The 'key' for the relative code comprises of 3 numbers, we'll consider them k1, k2 and k3. The accompanying conversation expects the utilization of a 255 input signal(y = 255), is ought to be generally prime to 'y' (which means it shares no common factors with y).

The encoding utilizes particular number juggling to replace the encoded number representation of each letter in a plaintext to a different number that considered as a cipher text letter. The encoding capability for a solitary letter is:

Let $T = k_3^{k1}$, then:

$$C= E(p) = (p + T) *k_2 \bmod y \ldots..(3)$$

y : size of the input

k_1, k_2 and k_3: are secrete keys and should be picked with the end goal that they and y are coprime.

In decoding the ciphertext, the inverse (or converse) capabilities should be applied on the coded text to recover the original text. Yet again the first step is to replace each letter in the ciphertext to its order numbers sequence. The decryption process is:

$$D(C) = P =(C * k_2^{-1}) – T) \bmod y \quad \ldots..(4)$$

K_2^{-1} : secluded multiplicative converse of a modulo m. i.e., it fulfills the condition

$$(k_2*k_2^{-1}) \bmod y = 1$$

Prove:

$P = (C * k_2^{-1}) - T \bmod y$

$P = ((((p + T) *k_2 \bmod y) * k_2^{-1}) - T) \bmod y$

$P = ((p + T) - T) \bmod y$

$P = p \bmod y$

$P = p$

Example:

$k_1 = 6$;

$k_2 = 8$

$k_3 = 13;$

$k_2^{-1} = 157;$

$T = k_1^{k2}$

Encryption:

P = 171

C = ((P + T)* k_3 mod 255)

C = 51

Decryption:

P=(C *k_2^{-1}) - T mod 255)

P = 171

Table 1: implementation of proposed method on speech sample

Y(input signal samples)	convert input sample to decimal number	plain text (P)	cipher text (C)	The signal sample after decryption
٠,١٥٦١٢٧٩٢٩٦٨٧٥٠٠	١٧١	١٧١	١٢٣	٠,١٥٦١٢٧٩٢٩٦٨٧٥٠٠
٠,١٤٠٥٠٢٩٢٩٦٨٧٥٠٠	١٧٣	١٧٣	١١٤	٠,١٤٠٥٠٢٩٢٩٦٨٧٥٠٠
٠,٠٨٧٧٦٨٥٥٤٦٨٧٥٠٠ ٠	١٨٤	١٨٤	١٩٢	٠,٠٨٧٧٦٨٥٥٤٦٨٧٥٠٠ ٠
٠,٠٦٨٢٣٧٣٠٤٦٨٧٥٠٠ ٠	١٨٩	١٨٩	٤٢	٠,٠٦٨٢٣٧٣٠٤٦٨٧٥٠٠ ٠
٠,٠٢٣٨٠٣٧١٠٩٣٧٥٠٠ ٠	٢١١	٢١١	١٩٨	٠,٠٢٣٨٠٣٧١٠٩٣٧٥٠٠ ٠
٠,٠١٤٠٣٨٠٨٥٩٣٧٥٠٠ ٠	٢٢١	٢٢١	١٥٣	٠,٠١٤٠٣٨٠٨٥٩٣٧٥٠٠ ٠
٠,٠٣٢١٠٤٤٩٢١٨٧٥٠٠ ٠	٢٠٥	٢٠٥	٢٢٥	٠,٠٣٢١٠٤٤٩٢١٨٧٥٠٠ ٠
-٠,٠١٩٨٩٧٤٦٠٩٣٧٥٠٠ ٠	٨٧	٨٧	٢٤٦	-٠,٠١٩٨٩٧٤٦٠٩٣٧٥٠٠ ٠
-٠,١٢٤٨٧٧٩٢٩٦٨٧٥٠٠	٤٧	٤٧	١٧١	٠,١٢٤٨٧٧٩٢٩٦٨٧٥٠٠-
-٠,٠٨٣٨٦٢٣٠٤٦٨٧٥٠٠ ٠	٥٧	٥٧	١٢٦	-٠,٠٨٣٨٦٢٣٠٤٦٨٧٥٠٠ ٠
-٠,٠٦٨٢٣٧٣٠٤٦٨٧٥٠٠ ٠	٦١	٦١	١٠٨	-٠,٠٦٨٢٣٧٣٠٤٦٨٧٥٠٠ ٠
-٠,١٣٢٦٩٠٤٢٩٦٨٧٥٠٠	٤٦	٤٦	٤٨	٠,١٣٢٦٩٠٤٢٩٦٨٧٥٠٠-
-٠,١٧١٧٥٢٩٢٩٦٨٧٥٠٠	٤١	٤١	١٩٨	٠,١٧١٧٥٢٩٢٩٦٨٧٥٠٠-

٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠-	٧٥	٤٠	٤٠	-٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠
٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠-	٧٥	٤٠	٤٠	-٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠
٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠-	٧٥	٤٠	٤٠	-٠,١٧٩٥٦٥٤٢٩٦٨٧٥٠٠
٠,١٩٥١٩٠٤٢٩٦٨٧٥٠٠-	٨٤	٣٨	٣٨	-٠,١٩٥١٩٠٤٢٩٦٨٧٥٠٠
٠,٢١٨٦٢٧٩٢٩٦٨٧٥٠٠-	٢٢٥	٣٥	٣٥	-٠,٢١٨٦٢٧٩٢٩٦٨٧٥٠٠
٠,٢٥٣٧٨٤١٧٩٦٨٧٥٠٠-	٢٤٣	٣١	٣١	-٠,٢٥٣٧٨٤١٧٩٦٨٧٥٠٠

4- The presentation estimations

To get to the presentation of the introduced cryptosystem, the normal following measures are utilized: Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR), mean squared error (MSE), the maximum error (MAXERR) and (L2RAT). Brief depiction of these estimations is given underneath[7][8][9][10].

4-1 Signal to Noise Ratio (SNR)

To quantify the twisting value in discourse encryption, the SNR quantifier is quite possibly one of the major well-known measures that is generally utilized for this reason. It is calculated by:

$$SNR_{dB} = 10 * log_{10} \frac{\sum_{i=1}^{P} Si^2}{\sum_{i=1}^{P} [Si - yi]^2} \qquad \text{.....(5)}$$

Where P, is the audio length(number of samples), Si is the input signal sample and yi is the decoded sample.

4-2 Peak Signal_to_Noise Ratio

One of the most popular evaluation metrics is “PSNR” which is used as indicator of the signal quality. The equation that defines PSNR is as follows:

$$PSNR = 10 * \log \left\lfloor \frac{\sum (Q^D - Q^O)^{\,2}}{M^2} \right\rfloor \qquad \text{.....(6)}$$

Where Q^D and Q^O are the deciphered and the original sample, while “M” the audio size.

4-3 Mean Squared Error (MSE)

A factor that measures the error with a transmitted signal is the mean squared error "MSE". It calculated by finding the difference between the original (plain) and the received signal samples (deciphered). MSE is zero when the original and the deciphered signal are identical, and it increases when the system has errors. MSE is defined by the following equation:

$$MSE = \frac{1}{M}\sum_{t=1}^{M}(St - Dt)^{2} \qquad \text{.....(7)}$$

Such that:

- St is the t^{th} original value.
- Dt is the t^{th} decrypted value.
- M is the number of samples.

4-4 Maxerror

It is the compute of the maximum error between the original and the deciphered signal. It is the maximum absolute deviation of the error occurred within the reversion model. The value of "maxerr" is 0 when the original and the deciphered signal are identical while it reaches infinity in the worst case (it reflects the similarity between decrypted and actual signal).

$$MAXERR = \sum_{t=1}^{M} F(Dt)(S_t - S_{t-1}) \qquad \text{.....(8)}$$

4-5 L2RAT

It is the ratio of the squared norm of the signal approximation to the input signal. The value of L2RAT comes near 1.

5- Results and Analysis

The calculation of the performance metrics shows that all values for different audio files were reasonable and satisfied. The proposed method was tested over many audio files and the pre_stated metrics were calculated as shown in table 1 below:

Table 2: Measurements of applying the proposed method on audio files

FILE	SNR	PSNR	MSE	MAXERR	L2RAT
Die	23.92011	84.29855	0.0413785	0.0980346	1.004054
YES	13.76850	75.38130	0.05309455	0.0980346	1.041990
SAMPL3	22.18719	80.664070	0.017809	0.0980437	1.00674

By calculating the correlation between the signal samples during the process, it was clear that the correlation between the fractional samples of the audio file before encryption and after decryption was very strong “corrAu1”. Also the correlation between digitized signal “corrAu2”, which means the encryption/decryption process didn’t corrupt the audio signal. In another hand, the correlation that calculated between the plain and ciphered signal was very weak “corrAu2”, which means the proposed encryption method gave a full confusion to the original signal.

Table 3: Correlation measurements of applying the proposed method on audio files

FILE	corrAu1	corrAu2	corrAu3
Die	0.9983	0.9980	-0.0212
YES	0.9868	0.9789	-0.0037
SAMPL3	0.9961	0.9970	-0.0083

Figures (2,3,4,5) shows the signal within different steps of the proposed method for the same audio file “Yes.au”. Figure 2 shows the original signal and the signal after converting its samples from fractions to integers.

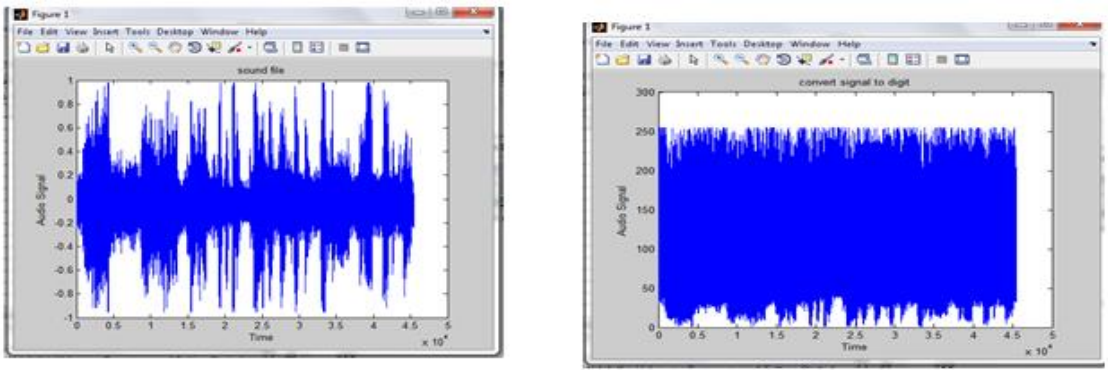

Figure 2: original and digitization of ‘Yes.au’ file

Figure 3 represents the original “plain” signal and the ciphered signal. The correlation between them is very weak as it is clear from the figure below.

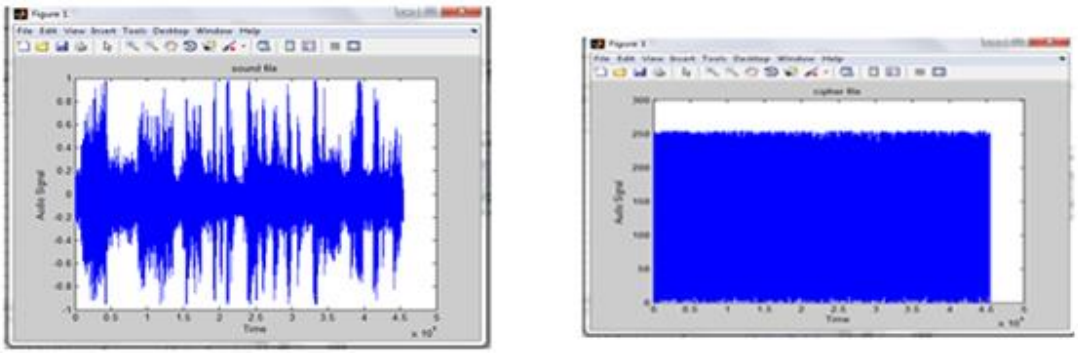

Figure 3: the plain and the ciphered signal of "Yes.au" file

After decryption, the deciphered signal returns to its original shape as it is shown in figure 4 below.

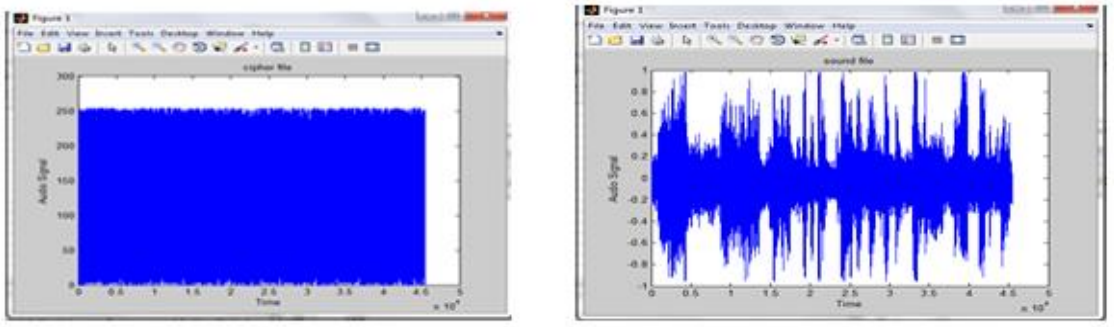

Figure 4: ciphered and deciphered 'Yes.aud' file

Figure 5 below shows the histogram of the original and the deciphered signal of the same file "Yes.au". The samples distribution after decryption returns to the original state, and the signal brought back its main construction.

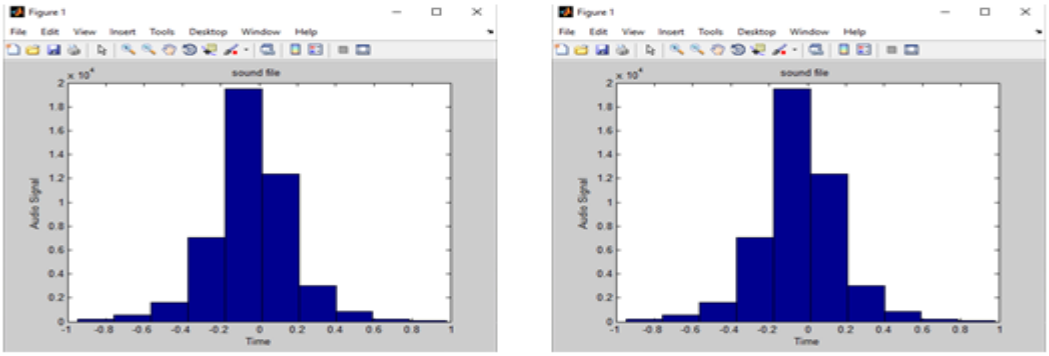

Figure 5: original and deciphered histogram of 'Yes.aud' file

6- Conclusion

Audio encryption is one of the most critical processes involving transmission and media distribution. The difficulty of manipulating audio is mainly came from the sensitive nature of audio and the effect of noise, which is defined as any change that happened to audio caused from outside or from the cryptosystem used. The proposed method "S-affin" proved to have a strong encryption method with maximum diffusion and in the same time recreate the deciphered signal with a maximum

correlation with the original. The results of (SNR, PSNR, MSE, MAXERR and L2RAT) show a very satisfied values.

References

[1]. Hammid, A.T., M.H.B. Sulaiman, and A.N. Abdalla, "Prediction of small hydropower plant power production in Himreen Lake dam (HLD) using artificial neural network", Alexandria Engineering Journal, 2018. Volume 57, issue (1): p. 211-221.

[2]. Sonia Rani, and Harpreet Kaur. (2017) "Technical Review on Symmetric and Asymmetric Cryptography Algorithms", International Journal of Advanced Research in Computer Science volume 8 ,issue (4): p 182-186.

[3]. Prashant Kumar Arya, Mahendra Singh Aswal, and Vinod Kumar. "Comparative Study of Asymmetric Key Cryptographic Algorithms", International Journal of Computer Science & Communication Networks 5 (1): 17-21.

[4]. Sonia Rani, and Harpreet Kaur. "Technical Review on Symmetric and Asymmetric Cryptography Algorithms", International Journal of Advanced Research in Computer Science 2017 volume 8 ,issue 4 , 182-186.

[5] Ihsan, A., Doğan, N. "Improved affine encryption algorithm for color images using LFSR and XOR encryption". Multimed Tools Appl 82, 7621–7637 (2023). https://doi.org/10.1007/s11042-022-13727-w.

[6] Pavel Panchekha, Alex Sanchez-Stern, James R. Wilcox, Zachary Tatlock, "Automatically Improving Accuracy for Floating Point Expressions", Conference: ACM SIGPLAN Notices, June 2015, 50(6):1-11.

[7] Malathi V., Gopinath M.P, "noise deduction in novel paddy data repository using filtering techniques", Scalable Computing: Practice and Experience, ISSN 1895-1767, 2020, Volume 21, Issues 4, pp. 601–610, DOI 10.12694:/scpe.v21i4.1718.

[8] Basaad Hadi , Ahlam Majeed , "Transform Infra- Red Image Using Discrete Wavelet Function", IOP Conf. Series: Materials Science and Engineering 571 (2019) 012114 IOP Publishing doi:10.1088/1757-899X/571/1/012114.

[9] K. Sasirekha, R. Ravikumar, K. Thangave, "Online Signature Denoising using Deep Autoencoder", International Journal of Computational Intelligence and Informatics, Vol. 7: No. 1, June 2017.

[10] Sana Ahmed Kadhim, Saad Abdual Azize abdual Rahman, "A proposed method for encrypting and sending confidential data using polynomials", Global Journal of Engineering and Technology Advances, 8(2), 082–087, (2021), https://doi.org/10.5281/zenodo.5516895.