Encrypt Audio File using Speech Audio File As a key

Najwan A Hassan, Farah Saad Al-Mukhtar and Esraa H Ali
Baghdad, Iraq, AL-NAHRIAN UNIVERSITY, DEPARTMENT OF COMPUTER SCIENCES.
nah@sc.nahrainuniv.edu.iq, fse@sc.nahrainuniv.edu.iq, eha@sc.nahrainuniv.edu.iq

Abstract in this research, a new method for audio (.WAV) file encryption is propose using a random modification with two secret key applications. The proposed method provided a highly secure audio.WAV files and efficient audio files which have very high-frequency band. The proposed system is to convert speech audio file to text form, this text creates password seed with two keys using hash function, the first key encrypted using proposed algorithm then these keys used to encrypt the original audio (.WAV) file by using Rijndael algorithm.

Keywords. Speech audio file; Encryption; Decryption; Audio; Hash function; Rijndael algorithm Speech Recognition.

1. Introduction
Security and privacy of the data is a big issue among human Bings, for that, the encryption concept is oldest fact as that secret data interchanged among people [1]. Cryptography is a structure of converting any data (audio, text, image, and video) into other types of data that are meaningless unless decryption key is available to protect them from eavesdropper in the internet [2,3]. Cryptography involved two processes; Encryption and Decryption. Asymmetric and symmetric is a two ways Cryptography algorithms turn in. algorithms of Symmetric have similar password secret key for encryption and decryption, this password secret key have to be ambiguity from everyone in the network except two people (sender who send the data and receiver who receive the data) [4]. While, Asymmetric key algorithm used two keys, one of them is a public key for encrypted the other is private key for decrypted [5,9]. Audio files encryption guarantee secure audio file transmitted. With the quick expansion of connection technology preservation of audio files from the eavesdropper and hackers became a critical mission for the expert in technology. For that there is constantly a need of a quickly and safer audio files encryption techniques[6]. Cryptography encryption for Audio is the method of which include the key (noise) to a plain text audio in the same time decryption is the operation of bring out the main plain text back by use the same key[7].

Speech is exceptionally charming broker for human-computer interaction which is hands-free; it is only demand simple hardware for procurement “high quality microphones” it comes at very simple bit rate. Human speech is recognize, essentially uninterrupted connected speech, without wearisome exercise (free speaker), for a vocabulary of suitable complexity (100,000 words) is extremely complex. However, with new algorithm, processes, flow diagram, and method that can process speech signals very easy and realize the text that is talking by a person who talks [8].

In this paper, a new algorithm is proposed to perform encryption audio file (.WAV) by using two secret keys generated from speech audio file (.WAV). The proposed algorithm was tested and implemented.
2. Related Work

[10] Fabian, Michael, Qi Li and Susanne proposed a technique to create a reliable encryption key from the voice of the user when speaking a password, this key is capable to combat code analysis and the person who capture system information regarding to the verification or creation the encryption key.

[11] Dr. Ekhlas a new algorithm proposed for audio file encryption based on a combination of block cipher and chaotic maps. In this algorithm a block of size 625(25 × 25) bytes is encrypted and decrypted, by inputted these blocks into three stages: first stage is Permutation, then Xor adding stage and finally Substitution stage.

[12] Miss. Divya, proposed a method to encrypt an audio for raising the security of data audio that want to transfer on an insecure medium. A five level of encryption applied on the audio that recorded using microphone in real time, which generate a cipher signal routed for the recipient during an insecure line. Then the recipients receive that signal, and in order to retrieve the signal back decrypts the cipher signal.

[13] Monther and Azman suggested a method to create a cryptographic public key from the voice of user while using hand held device. This key used for encrypted the data, which transmit through the open communication channel.

3. Proposed System

A new encryption algorithm for audio files is proposed and presented in this research work. The blocks diagram for the proposed system is shown in figure 1 that performs encryption audio file using Rijndael algorithm the two password keys for this algorithm is resulting from speech audio file in the same way the decrypted audio file can be retrieved as original audio file by using the same two password keys (speech audio file) and a reverse of the proposed method. The proposed encryption algorithm for audio cryptography scheme is quite simple process and safe. The steps for implementation of proposed method are as follows:

Figure 1: the proposed system.
3.1 Generate Password Keys
The generation of the password key in the proposed system is based on hash function. This stage includes the following steps:

**The first step** the sender send speech audio file (.WAV) then converted these speech audio file to text using algorithm of speech recognition then every character in the text take its ASCII code and convert it to binary file as shown in Figure 2.

![Figure 2: The first step.](image)

**Second step** is create two password keys from binary file produce from first step encrypt by using hash function, the hash function is very powerful to convert the string which is binary file with big length to fixed string (512 bits) as length, the result is called digest. Any size and type of data are taken in a hash algorithm, then switch it to a constant length of data. using hash function often to retrieve the data easily at the same time large quantity of data to short string, that is other characteristic of the signature found on hash projects. In proposed system the first key will take specific size (256 bits) that are the result of hash function into byteKey. The 0 To 31 place in initial 256 bit of 512 bit to byteKey. The 32 To 47 will place the next 128 bits into byteIV which is the second key. It’s actually an Initialization Vector (or just ‘IV’), which it is a random sequence of bytes that denies the same plaintext from all the times being transform into the same ciphertext.

**Third step** encrypt the first key (byteKey) by convert every byte to binary then create new key as shown in figure 3.

![Figure 3: Encrypt first key.](image)
3.2 Encrypt audio file

The audio file encryption stage includes the following steps. Entering audio file (.WAV) to encrypt using two keys from the first step as input to Rijndael algorithm. The audio file (.WAV) considered as a stream, and the encryption performed is two keys dependent and data dependent. The encryption algorithm is illustrate as follows:

1. Enter audio file (.WAV) as input to Rijndael function to encrypt it.
2. Enter two password keys to Rijndael function.
3. The first key (bytKey) is the secret key that used for the symmetric algorithm and 256bits is the size of a key.
4. The second key (bytIV) that used for the symmetric algorithm and 128bits is the size of a key.
5. The Rijndael encryption process goes as show in figure 4.
6. Encryption process output is saved as audio file (Encrypt.WAV).

\[
\text{first}_\text{key} = C_1, C_2, C_3, ..., C_{n-1}, C_n \quad \ldots \ldots (1)
\]

Where

\[
C_1 = (K_0 \oplus K_n)
\]
\[
C_2 = (K_1 \oplus K_{n-1})
\]

\[
\vdots
\]
\[
C_n = (K_{n/2} \oplus K_{n/2+1})
\]

\[
C_{n/2} = (K_{n/2} \oplus K_{n/2+1})
\]

3.3 Decrypt audio file

The encrypted file (.WAV) is send to the receiver via secured channel, with the speech file as a key. The decryption algorithm is illustrated as follows:

1. Enter audio file (.WAV) as input to Rijndael function to decrypt it.
2. Enter two password keys to Rijndael function.
3. The first key (bytKey) is the secret key that used for the symmetric algorithm and 256bits is the size of a key.

Figure 4: Rijndael algorithm.
4. The second key (bytIV) that used for the symmetric algorithm and 128 bits is the size of a key.
5. The Rijndael decryption process with two keys.
6. The output of decryption process is saved as audio file (test.WAV) which is exactly same the original file.

4. Results
This section presents the result of a proposed system that described in previous part. The execution of this system is tested using VB.NET. That goes through several steps:
- The first step is to generate keys from speech audio file (.WAV) using hash function and proposed algorithm, for the tested audio file with size (128 KB) the time to generate first key and second key is 0.04097 sec., the autocorrelation test for these keys is 0.372.
- The second step is to generate encrypted file using Rijndael function, to observe the encryption effect on audio file, the plotted audio values before and after encryption. Figure (5) show original tested audio, Figure (6) show the tested audio after encryption.

![Figure (5) original tested audio file.](image)

![Figure (6) tested audio file after encryption.](image)

The correlation coefficient executed in this work for five encrypted audio files. Low correlation coefficient showed a secure data encryption algorithm. Table (1) shows the correlation results of the encrypted files, using the following equation:

\[
\text{Correl} = \left(1 - \frac{\sum_{i=1}^{n} [Y_i - \bar{Y}(x)]^2}{\sum_{i=1}^{n} Y_i^2 - \frac{1}{n} \left[ \sum_{i=1}^{n} Y_i \right]^2} \right)^{\frac{1}{2}}
\]

(2)

Table (1). Correlation Results of the Encrypted Files

| File Name | Correlation coefficient | Size of files |
|-----------|------------------------|---------------|
| Tiger     | 0.0000121149           | 156 KB        |
| Cutafew   | 0.0022673091           | 86.4 KB       |
| Nucase    | -0.001115621           | 72 KB         |
| byebye    | -0.001765              | 84KB          |
| ohyeah    | -0.000752              | 40KB          |
The above table shows low correlation values that indicate the used algorithm in this work difficult to attack and it is secure. The running time is performed for encryption and decryption for five tested audio files by using the algorithm are proposed shown in Table (2)

| File Name | time of Encryption file in second | time of Decryption file in second |
|-----------|----------------------------------|----------------------------------|
| Tiger     | 4.571                            | 4.389                            |
| Cutafew   | 3.736                            | 3.683                            |
| Nucase    | 4.476                            | 4.199                            |
| bybye     | 4.147                            | 4.003                            |
| ohyeah    | 3.921                            | 3.531                            |

From Table (2), can see that a time required for Encrypt file is more than time required for decrypt file. MSE is the measure the total squared fault is founded between the original audio files and encryption audio files.

\[
MSE = \frac{1}{N} \sum_{i=0}^{L} [A(i) - B(i)]^2 \quad \ldots\ldots(3)
\]

PSNR is a measure to describe the modifications for audio value quality of the original audio files and encryption files.

\[
PSNR = 10 \log_{10} \left[ \frac{M^2}{\sum_{i=0}^{N} (X(i) - Y(i))^2} \right] \quad \ldots\ldots(4)
\]

**Table (3)** PSNR and MSE results for encryption tested files

|        | PSNR   | MSE   |
|--------|--------|-------|
| Tiger  | 3.2556 | 0.4725|
| Cutafew| 4.5588 | 0.3500|
| Nucase | 4.1467 | 0.3849|
| bybye  | 4.2737 | 0.3738|
| ohyeah | 4.3413 | 0.3680|

Low values of PSNR indicate the high-level of noise in the encryption audio files, which indicates more resistible to attacks.

5. **Conclusion**

A new encryption algorithm is proposed for audio (.WAV) files using a two secret key by convert speech audio file to text form; this text creates seed with two keys using hash function then encrypts the original audio file by using Rijndael algorithm. The results for all tested audio files shows that the proposed algorithm for audio files is secured because of its, uniform histograms, large keys space, low correlation, and low PSNR, showed that the proposed algorithm for encryption of audio file is a very good choice in same time is a very good security to audio transmission.
6. References

[1] Sheetal Sharma, Lucknesh Kumar and Himanshu Sharma “Encryption of an Audio File on Lower Frequency Band for Secure Communication” International Journal of Advanced Research in Computer Science and Software Engineering Vol. 3, Issue 7, July 2013. p. 79-84.

[2] Suhad Latef, Ban N. Dhannoon, Color Image Encryption using Random Password Seed and Linear Feed Back Shift Register. Journal of Al-Nahrain University, 2011. 14(1): p. 186-192.

[3] Jayaram P, Ranganatha H R, Anupama H S “Information Hiding Using Audio Steganography – A Survey” The International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.3, August 2011. p. 86-96.

[4] Gatiff, Bill, “Encrypting data with the Blowfish Algorithm”, http://www.eetindia.co.in/ARTICLES/2003AUG/EEIOL_2003AUG02_EMS_TA.pdf (accessed January 12th 2015).

[5] Mandal, P.C. (2012), “Superiority of Blowfish Algorithm”, International Journal of Advanced Research in Computer Science and Software Engineering 2(9), September 2012, pp. 196-201.

[6] Manpreet Kaur (12012134) and Ms. Sukhpreet Kaur “Survey of Various Encryption Techniques for Audio Data” International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 5, May 2014. Pp. 1314-1317.

[7] Raghunandhan K R, Radhakrishna Dodmane, Sudeepa K B,Ganesh Aithal “Efficient Audio Encryption Algorithm For Online Applications Using Transposition And Multiplicative Non-Binary System.” International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 6, June – 2013.pp. 472-477.

[8] B. Raghavendhar Reddy,E. Mahender,”Speech to Text Conversion using Android Platform”, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, Vol. 3, Issue 1, January -February 2013.

[9] Sattar B.Sadkhan,Sawsan Kamal,Najwan Hassan ,"Fuzzy based Pseudo Random Number Generator used for Wireless Networks", Journal of Al-Nahrain University, Vol.16(2), july,2013, pp2010-2016.

[10] Fabian Monrose, Michael K. Reiter, Qi Li and Susanne Wetzel, “Cryptographic Key Generation from Voice”, In Proceedings of the 2001 IEEE Symposium on Security and Privacy , May 2001

[11] Dr. EkhlasAbbasAlbahrani, “A New Audio Encryption Algorithm Based on Chaotic Block Cipher”, Annual Conference on New Trends in Information & Communications Technology Applications-(NTICT’2017) 7 - 9 March 2017.

[12] Miss. Divya Sharma, “Five Level Cryptography in Speech Processing using Multi Hash and Repositioning of Speech Elements”, International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 5, May 2012.

[13] Monther Rateb Enayah and Azman Samsudin, “Securing Telecommunication based on Speaker Voice as the Public Key”, IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.3, March 2007.