EAI Endorsed Transactions
on _____ Research Article EAI.EU

Crypto Model of Real-Time Audio Streaming Across Paired Mobile Devices

E.O. Ibam[1,*], O.K. Boyinbode[2] and I.O. Ayelabowo[3]

[1]Department of Information Systems, Federal University of Technology Akure, Nigeria
[2]Department of Information Technology, Federal University of Technology Akure, Nigeria
[3]Department of Computer Science, Federal University of Technology Akure, Nigeria

Abstract

The progressive usage of the Internet of Things (IoT) and smartphone ubiquity is inducing attention, importance and interest in leveraging smartphones for deploying and running sophisticated mobile applications. Modern smartphones have good connectivity, a significant amount of processing, and are always with us, making them an ideal candidate for envisioned applications that can serve as a means of ensuring the safety and security of mankind. This research aims at developing a real-time audio encryption application between two or more handheld mobile devices, the communication between devices is made secure by encrypting the audio sent in real-time using Advanced Encryption Standard (AES) 256-bit encryption key while channelling message through a cloud-hosted database (Firebase) that works as a real-time database and performs implicit AES encryption and decryption on its data. Geocoder from the Google Map API library services is used to track the location of the audio sender.

Keywords: Mobile Communication, Mobile Application, Mobile Network, Encryption, Decryption.

Received on 03 09 2019, accepted on DD MM YYYY, published on DD MM YYYY

doi: 10.4108/_____

*Corresponding author. Email: eoibam@futa.edu.ng

1. Introduction

Communication through mobile devices has developed from the use of single cell-phone to an advanced smartphone thus increasing functionality and user interactivity, due to this efficiency and reliability, it can be seen within anyone's reach. In this contemporary day of digital world, importance of networks, its effect and presence can't be neglected. The constant use of digital data in real life applications and its significance craved the need for a new way to ensure safety and usage of electronic gadget possessed by human. Information is power no matter how minute it is and the rapid advances in the development of computer technologies and internet has made the security of information as most important factor in information technology and communication hence the need for a secured information processing system.

The rate of abduction in the world especially in Africa is escalating, recently in Nigeria the recorded average rate of kidnapping and abduction is about three cases per week, while the number of individuals who become victims in these ugly incidences is above seven. In most cases recorded, their abductors will first of all seize their mobile phones and demand for ransom from their relatives. In cases where the victim is injured, they can easily access their phone SIM cards and retrieve the contacts of their close relatives or the last phone number the victim has called. Therefore, there is a need to focus on a way to reduce this threat being posed by hostage takers. In instances where distress occurs, the application could be initiated to take audio input of whatever is going on in the surroundings so as to proffer and effect proper solution to the distress call at an appropriate time.

With the help and advancement of Internet of Things (IoT), communication and interconnection through internet connected device enable user to send and receive

data. This serve as an integrated and helpful way of embedding features that will allow sharing of meaningful audio to the selected contacts and designated receivers of the audio. If Mobile devices can be used for other purposes, why can't it be implemented for the security of mankind, as human security is essential in day-to-day living and within their reach wherever they find themselves located? This idea predominantly comes from the fact that people all over the world are connected through mobile phones that mainly uses audio or video as it main source of communication.

The purpose of this research is to explore and design how a smartphone can be used in the safety and security of mankind along with its usual daily routine usage. It includes all of the fundamental tools and resources needed in order to achieve its development. Android based mobile devices is deployed for testing the features in the application. Other section of this paper contains relevant related works, system model and design, system implementation, conclusion and references.

2. Background and Related Works

Researches have been focused on this area, a brief of some related works that falls within are explained.

[1] implemented an encryption scheme for voice calls, in its implementation, an encryption scheme was provided based on RSA encryption standard that enable users to encrypt a voice call before transmitting it on the mobile network but using an intermediate server as a software to encrypt calls from two parties was not always applicable due to large cost and non-availability of hardware needed.

[2] was able to create a real-time voice encryption system where setup for a one-way communication system is achieved through scrambling voice in real-time with basic digital signal processing operations. However, scrambling of voice signal is not enough to guarantee voice security as the intensity of the different frequencies stays the same even though they are shifted.

[3] developed an android mobile application that allows real-time voice communication through short range local wireless network mainly Bluetooth and Wi-Fi. They were able to create a two-way radio transceiver using an android device that allows allow peers to establish a voice communication provided that devices are in range but the challenge with this method was that android devices need to be discoverable at all time in order to pair devices.

[4] proposed a technique where quaternion-based encryption and decryption of audio signal can be achieved using digital images as a variable key and cover for audio signal, the technique will compose of two processes, the encryption process at the transmitter part and the decryption part at the receiver part. Encryption and decryption are symmetric based and the proposed technique implemented using Matlab simulator.

[5] took a survey of various techniques that can be used for encryption of audio data to suggest a more secure method for audio encryption. From the research, they observed that selective encryption techniques were better than total encryption techniques as it takes less time with degradation of signal and time consumption on MP3 compression is less than total encryption. In the study, they were unable to make modifications into existing algorithms to make audio data more secure.

[6] proposed a technique for encrypting an audio file using combine approach of transformation and cryptography, transformation is used to convert the audio file from time domain to frequency domain using fast Fourier transform modification to encrypt frequency band and RSA for its encryption.

[7] proposed a lightweight encryption scheme for real-time multimedia transmission without loss of security and media quality of service using two block transpositions and XOR operation. The first transformation to be used to generate a key frame in order to seed frame and improve compression, the second transformation with XOR operation is the main encryption process. The outcome of their experimental research result with various MPEG-4 movies shows real-time transmission of the encrypted data without loss of quality of service and states it encrypts faster than AES encryption of MPEG compressed data but the idea of selective algorithm proposed for encryption provides lower usability than the naïve algorithms that encrypts all data.

[8] proposed a system for secured audio data transfer over the internet using steganography. Their idea was to develop an encryption process on secret speech signal data bits level to achieve great strength of encryption, embed the audio file inside the cover image using techniques of steganography and hide the audio as well as text message inside the cover image. To provide additional security noise is added to the audio file before it is hidden in the cover image, audio file encryption is carried out by adding high frequency noise bits at low frequency components of the signal, using least significant bits embedding method to hide the data and XORing secret key with binary data. This same process will be used to retrieve the speech file for a person with the correct secret key as new secret key is generated from the system every time.

[9] proposed five level cryptography in speech processing with Matlab using multi-hash and repositioning of speech elements to increase the security of audio data meant to be transferred on an insecure medium by creating a cipher signal which is routed through an insecure line to the receiving recipient. Microphone converts the sound eave into an electrical wave, the electrical wave is converted into digital audio using pulse code modulation, the speech is sampled and kept in a loop with a key table substitution process that substitute the speech elements with the key table elements that helps in repositioning the speech elements, repositioning is done in such a way that the end result is a cosine waveform and a new file is generated to be sent across the internet. The decryption process follows the inverse of the encryption process.

[10] implemented voice, video and text data through wireless by creating a communication system that allows android based smartphone users to send and receive messages, voice call and video call over the Wi-Fi range which requires neither internet connectivity nor messaging service from the mobile service providers. The base idea behind their system was to unify voice and data onto a single network infrastructure by digitizing voice signals, converting them into IP packets and sending them through an IP network together with the information gotten. It was able to reduce the cost of data transmission and communication within a fixed range providing zero cost communication through Wi-Fi.

[11] carried out a research on secured data transmission using the different security approaches together. Their proposed work follows an approach based on the use of Encryption of data using AES, DES and Blowfish algorithm sequentially on the data and the use of steganography using LSB algorithm to embed the encrypted data file into frames to distract the attention of the attacker.

[12] carried out a research on the design and implementation of encrypted call application on android system using VoIP to transmit voice and convert it from an analog signal to a digital signal, RTP protocols to transmit digital packets over the network line and SRTP to add protection with the use the of AES algorithm and Zimmermann real-time transport protocol to generate key for each call.

[13] conducted a study on the current scenario of audio encryption by demonstrating various techniques for different applications. In his research work, it was found that all audio encryption techniques can be divided into either full or partial encryption, with a good amount of overhead increases with partial encryption and complexity in algorithm.

[14] worked on encryption of an audio file on lower frequency band for secure communication by taking a frequency domain of the wav audio signal for encryption and decryption using partial encryption approach and RSA technique to encrypt the important portion of the audio and DFT to transform it time domain audio signal to frequency domain audio signal.

[15] proposed a method to encrypt audio stream of data in mobile handset by applying chaos using a pair of one-dimensional logistic maps to generate a chaotic encryption sequence for audio transfers between phones.

3. System Model and Design

Observing several researches brought about the idea of creating a real-time system that could be used in scenarios where emergency occurs and there is need to track location of an individual on a search basis and also allow sharing of recorded audio.

The requirement needed in developing the crypto-model audio streaming mobile device are Android studio and Android SDK Tools.

The programming languages used are Java and JavaScript and the backend server implemented using Firebase.

Features incorporated into the application are:

- Authentication and Token generation achieved in registration process for profile creation.
- Selecting contacts for sharing, contacts to be selected are saved on the user mobile phonebook and registered on the mobile application.
- Recording of audio (default of 3-5mins).
- Encoding and decoding of recorded audio.
- Location listener for location updates as recording of audio happens.
- Generating of Url for each encoded recorded audio
- for identification and immediate upload to cloud database.
- Sharing of the uploaded encoded recorded audio among selected contacts.
- Instant service notification on received encoded audio for selected contacts.

The figure below describes the system sequence of operation.

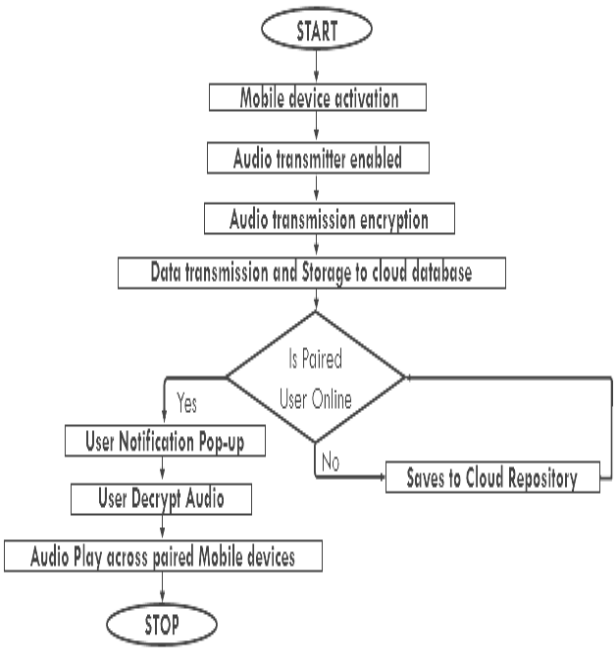

Figure 1. System Flowchart

The system architecture consists of the users (sender and receivers) who are referred to as the clients connected to the cloud-based storage (Database) which serves as the server. The users here are in synchronism with their

mobile devices since the client program runs on android based mobile device.

A user (sender) is connected to other users by registering, once registered, the application provides the list of other active clients that can be chosen as designated receivers, receivers must have been registered to the application and saved to the contact phonebook. Before the message gets to the database server (during the process of recording, the message is encrypted in real-time as the transmission start using the system entropy in AES-256) and when part of the message gets to the server side (database), in a progressive mode, the server pushes a response to the receiver side (designated user) almost immediately as in form of notification for incoming message and the message is decrypted. On the receivers' application user interface, the location of the sender is tracked in longitude and latitude and can be mapped using the global positioning system online for further information. In the case where the sender is in danger, may be his/her mobile phone is seized or even destroyed by the abductors while he or she is being kidnapped, the last recorded voice data and exact position is already saved in the memory of paired devices and cloud database, which could aid emergency rescue team or law enforcement agents in knowing the exact time and location the incident happened for quick response.

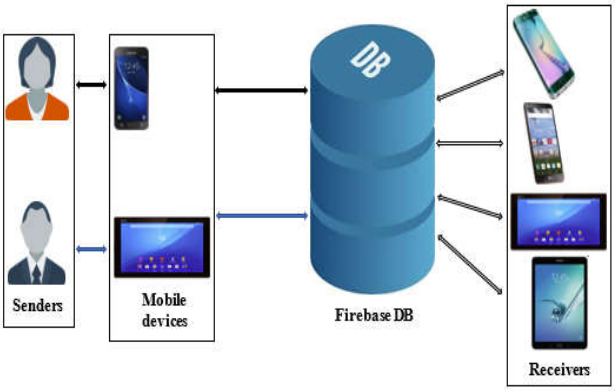

Figure 2. System Architecture

The figure below shows a typical chain of how a data is transmitted. The architecture follows the following steps:

a. Data is digitalized as recording begins taking its source from the MIC (smartphone speaker) in **.3gpp** format.
b. The recording data is encoded using an audio compression scheme optimized for speech coding - AMR-NB (Adaptive Multi-Rate Narrow Band) speech codec to generate a compressed bit rate and data to be transmitted. This codec provides voice activity detection, reduces noise generation, handles frames erasures and packet loss.
c. The recording audio is encrypted in real-time using AES-256 encryption as transmission start.
d. A web socket server orchestrates the communication. This is carried out by Firebase server.
e. At the receivers' end, the audio is decrypted and the received data is decoded using an audio decoder.

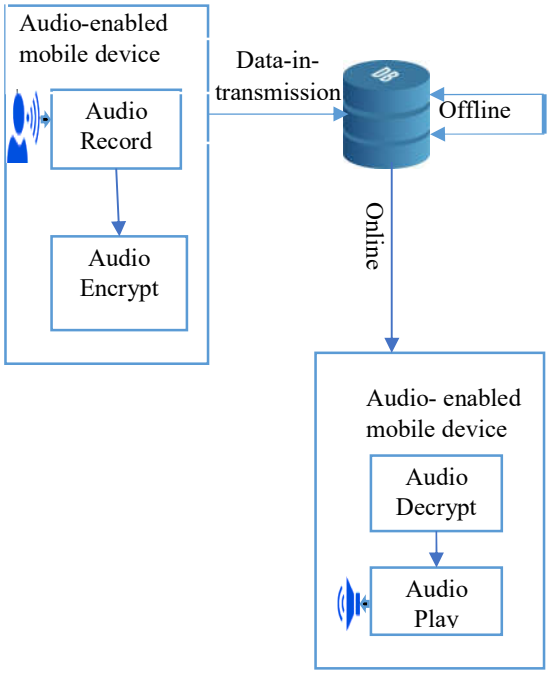

Figure 3. Audio Transmission Application Architecture

3.1. How Encryption is achieved

The Javax Crypto Security Key library was used to handle the encryption of each created recorded audio file before real-time uploading to the cloud through network connection happens. This allow the secured sharing of each recorded audio and it decryption by selected contacts. AES Algorithm was used and a 256-bit encryption was used for the audio encryption, to generate the random passphrase keys for encryption and decryption, the system's entropy is used in line with block size of audio file.

Advanced Encryption Standard (AES) algorithm is a block cipher algorithm, it encrypts each message received as a block one at a time producing its output one at a time. It uses a symmetric key encryption. In AES, each round consists of 4 layers; byte substitution using a substitution table, rows shifting of the state array by different offsets, column mixing of the data within each column of the state array and key addition but in it last round, mixed column is absent.

AES supports three-key length size where the number of rounds depends on the key length as illustrated in the table below.

Table 1. Key Block Round Combination for AES Algorithm

Length of Cypher Key	Key Length	Block Size	Number of Rounds
128	4	4	10
192	6	4	12
256	8	4	14

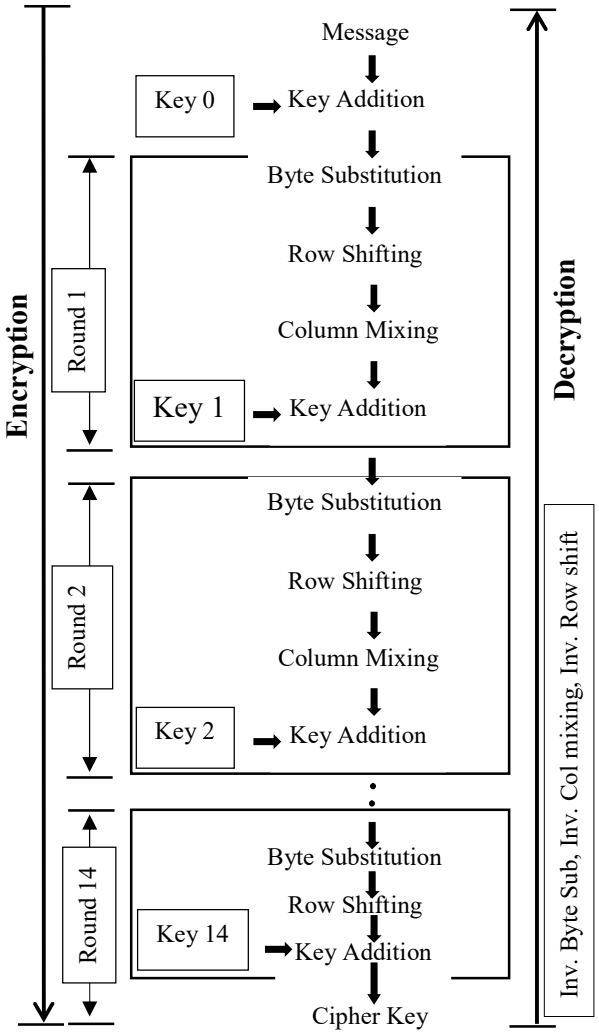

Figure 4. Structure of an AES Algorithm

3.2. How Bit Manipulation Occurs

According to [16], all internal operations of AES are based on finite fields and a new number system is generated. Finite can be observed to be countable elements or elements having limit and a field can be deduced as a mathematical set or entities under the operation addition, subtraction, multiplication and inverse in its region. From one of the conditions that must hold for a finite field to exist which is if there are ($\mathbf{P^m}$) elements where P is a prime number and M is a positive integer, if a finite field with 256 elements is to be considered, its representation will be given as $\mathbf{FF(2^8)}$ which falls under the AES Field.

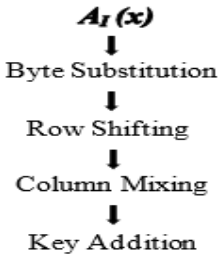

Figure 5. Round Transformation

At the start of the cipher, the incoming audio (A_I) is been broken down in bytes to fit into a state array (a two dimensional array of bytes ($S_{R,C}$) which consist of 4 rows of bytes containing block length divided by 32). The four bytes in each column of the state array forms a 32-bit words, where all the bytes are interpreted as finite field elements following the modulus of an AES Irreducible Polynomial to ensure that the result can be thus represented as a byte not greater than the byte in the column array.

The Byte Substitution is used to provide confusion by taking the multiplicative inverse of (A_I) in $\mathbf{GF(2^8)}$ then applying affine mapping matrix transformation to it and the inverse of the byte substitution when decrypting is the byte substitution where the inverse table is applied, obtained by the inverse of the affine mapping, then taking the multiplicative inverse in $\mathbf{GF(2^8)}$.

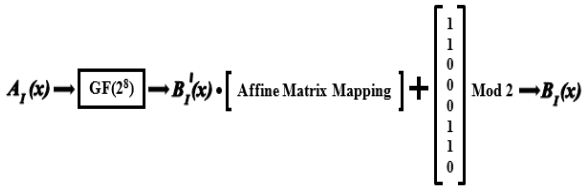

Figure 6. How bit substitution is performed

Row shifting is done to scramble the data obtained from $B_I(x)$. Here byte permutation or re-ordering of the byte is carried out. Bitflops is performed over a different numbers of bytes (offsets). In Column mixing, linear transformation occurs on the state column-by-column instead of the rows, the data is spread across the data path. The columns are considered as a polynomial of $GF(2^8)$, since each columns has 4 rows, Matrix transformation is treated as 4x4 followed by the addition of round key transformation added to the state by a bitwise XOR operation. Row shifting and Column mixing is done to provide diffusion.

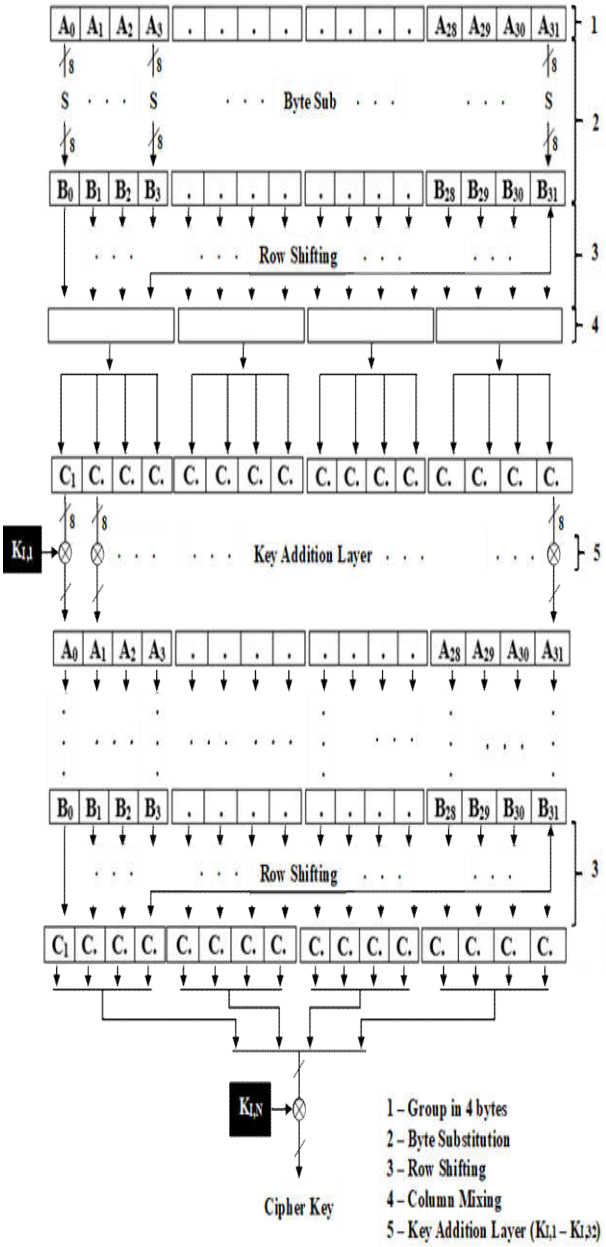

Figure 7. Generation of AES Cypher Key

3.3. Tracking Sender's Location

Integrating geocoder from Google map API service library to track location of audio sender, the Geocoder is localized for a given locales and the result given is a best guess. Google location listener package is used for the location updates and changes as audio recording happens. Real-time location update is possible provided there is internet connectivity in place.

4. System Implementation

The Android application called **"Smartrec"** is a mobile application with fast and helpful integrated feature to allow sharing of recorded audio among selected circle members.
In this study, the following devices were used for testing:

Phone brand: Infinix Hot 4 lite
Operating system: Android
Android version: 6.0
Size of encrypted audio: 1.2mb
Record time in minutes: 5

Phone brand: Tecno Camon CX
Operating system: Android
Android version: 8.1
Size of encrypted audio: 1.1mb
Record time in minutes: 5

4.1. User Application Interface

Figure 8. On lunch Activity

On lunching application, the user interface is displayed with three sections, make application record and send to selected contact during onlunch, choosing designated receivers and user registration to be able to send and receive recordings.

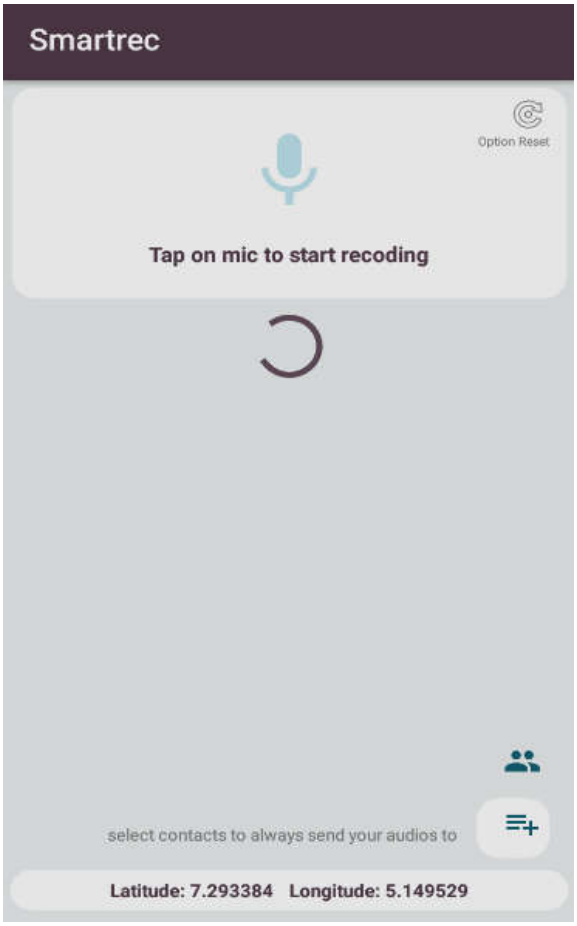

Figure 9. Recording Interface

The MIC symbol is tapped to start recording, it can be stopped to automatically begin upload to its receiver or left alone to begin upload after the given record time has elapsed.

After the elapsed time given for audio recording or explicit tapping on MIC image to stop recording, the file is being processed and an automatic upload to the designated receivers is done.

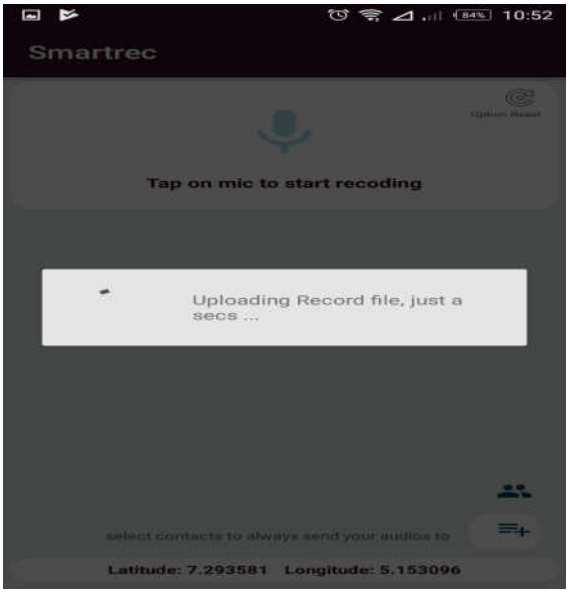

Figure 10. Recorded Audio File Transmission

Figure 11. Received Audio Interface

The mic image is tapped to decode and listen to the audio. The location beneath shows the location of the audio receiver and the location on each audio sent shows the location of the audio sender in longitude and latitude.

5. Conclusion

With rapid spread of smartphones incorporating a lot of features, Mobile devices becomes a very important aspect of every human life since it assists and enable us to access a large variety of services. This study presented an overview of secured smartphone audio transmission among users. It provides a suitable security to the audio by encrypting the audio through AES. Then, transmitting the audio through a cloud-based storage in real-time to designated receivers.

Future research and development should be carried out, so that improvement can be made and users can be rest assured of the security of their data. Developers working on implementation should consider various survey and encryption techniques that can be incorporated with the specifics of the database server needed in order to achieve a complete functional and desired system. The system should be given higher level of enhancement by incorporating other mobile devices products running different operating systems.

References

[1] EL BAKRY, H. M., TAKI, A. E. DEEN E. and EL TENGY, A. H. (2016) Implementation of an Encryption Scheme for Voice Calls. *International Journal of Computer Applications* volume 144 (02): 24-27.

[2] BRANDAU, M. A. (2008) "Implementation of a Real-time Voice Encryption System," *M.S. Thesis* University of Applied Science Cologne, Fachhochschule, Köln.

[3] BELDA, R., ARCE, P., DE FEZ, I., FRAILE, F. and GUERRI. C. (2012) Android Real-time audio communications over Local Wireless. Accessed at https://riunet.upv.es/handle/10251/57677 on 15/6/2017.

[4] KHALIL, M. I. (2017) Quaternion-based Encryption/Decryption of Audio Signal using Digital Images as a variable key. *International Journal of Communication Networks and Information Security (IJCNIS)* volume 9 (2): 216-221.

[5] KAUR, M. and KAUR, S. (2014) Survey of Various Encryption Techniques for Audio data. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)* volume 4 (5): 1314-1317.

[6] MAKWANA, V. and PARMA, N. (2014) Encrypt an Audio file using combine approach of Transformation and Cryptography. *International Journal of Computer Science and Information Technologies (IJCSIT)* volume 5(3): 4473-4476.

[7] CHOO,E. LEE, J., LEE, H. and NAM, G. (2008) "A Lightweight Encryption Scheme for Secure Real-time Multimedia Transmission," ITRC program of the Korea Ministry of Information & Communications and the Basic Research program of the Korea Science & Engineering Foundation.

[8] HALLALE, S., GAIKWAD, N., KACHOLE, A. and KAMBLE, S. (2017) Secured Audio Data transfer over Internet using Steganography. *International Research Journal of Engineering and Technology (IRJET)* volume 4(1): 996-1001.

[9] SHARMA, D. (2012) Five Level Cryptography in Speech Processing using Multi-Hash and Repositioning of Speech Elements. *International Journal of Emerging Technology and Advanced Engineering* volume 2 (3): 21-27.

[10] HINGMIRE, A. TIPARI, M., GOPALAN, R. and CHAVAN, S. (2015) Implementation of Voice, Video and Text Data over Wi-Fi. *International Journal of Engineering Research and General Scienc,* volume 03 (02): 1467-1473.

[11] SINGH, P. K., TRIPATHI, P., KUMAR R. and KUMAR D. (2017) Secured Data Transmission. *International Research Journal of Engineering and Technology (IRJET)* volume 4 (4): 217-222**.**

[12] ALIBRAHEEMI, K. H. and ALREKABY, W. A.(2016) Design and Implementation of Encrypted Call Application on Android System. *International Journal of Computer Science and Mobile Computing (IJCSMC,* volume 4 (9): 261-269.

[13] GANDHI, R. A. and GOSAI, A. M. (2015) A Study on Current Scenario of Audio Encryption. *International Journal of Computer Applications* volume 116 (7): 13-17.

[14] SHARMA, S., KUMAR, L. and SHARMA, H. (2013) Encryption of an Audio file on Lower Frequency Band for Secure Communication. *International Journal of Advanced Research in Computer Science and Software Engineering* volume 3 (7): 79-84.

[15] PRABU, A.V., SRINIVASARAO, S., APPARAO, T., JAGANMOHAN, M. and BABU RAO, K. (2012) Audio Encryption in Handsets. *International Journal of Computer Applications,* volume 40 (6): 40-45.

[16] PAAR, C. and PELZL, J. (2002) *Understanding Cryptography* (New York: Springer).