EasyChair Preprint

№ 3493

Study on Voice Password Based Secured Communication Using RSA and Elgamal Algorithm

Prashnatita Pal, Bikash Chandra Sahana, Sagnik Ghosh, Jayanta Poray and Professor Amiya Kumar Mallick

May 28, 2020

Study on voice password based secured communication using RSA and Elgamal algorithm

Prashnatita Pal[1], Dr.Bikash Chandra Sahana[2], Sagnik Ghosh[3] Dr. Jayanta Poray[4], Dr. Amiya Kumar Mallick[5]

[1,2]Electronics & Communication Engineering
National Institute Of Technology
Patna, India

[3]Electronics & Communication Engineering
St Thomas College of Engineering Technology
Kolkata, India

[4]Computer Science & Engineering
Techno India University
Kolkata, India

[5] Formerly Professor, Electrical Electronics & Communication Engineering
Indian Institute Of Technology
Kharagpur, India

Communication Email: prashnatitp@gmail.com

Abstract- Secured voice authentication based communication is the main aim of this study. Here eight speech keywords have been recorded and stored in computer memory. One speaker recognition model was used for voice password authentication. Then the speech keywords were encrypted using private key. The message was encrypted using RSA or Elgamal algorithm. The message was modulated using FSK digital modulation technique using and sent through the communication channel. The speech samples were demodulated and decrypted at the receiver. The received speech samples matched with the original transmitted voice samples. The equality ratio for this study is 0.6 and above. In this study secured authentication technique has been adopted. After voice authentication, secured communication has been done successfully.

Keywords- **Speech Recognition, Cryptography, RSA algorithm, ElGamal algorithm, voice password**

I. INTRODUCTION

Speech communication is an important aspect of our life. Security of speech to maintain its confidentiality, proper access control integrity and availability has been a major issue in speech communication. Therefore, protection of speech password or data from misuse is essential. Today in the generation of electronic gadgets, the necessity to prevent data from miscreants is increasing day by day. Cryptography is the process of utilization of codes to prevent anyone from violating speech security. Speech protection can be accomplished by changing the original speech by any means to some other speech codes, so that if someone gets that speech by means of hacking then also it must remain in useless bits of speech for that person. This process can be achieved by encrypting that speech by some means of algorithms which are known to the sender and on the other side the similar decryption algorithms must be known to only the desired receiver such that it can convert that encrypted speech back to the user understandable data or signal. To improve the protection mechanism, RSA (named after its authors - Rivest, Shamir and Adleman) is one of the most popular public key cryptographic algorithm that is used to ensure speech

communication security. It consists of two main cryptographic processes. Firstly, a public key is used to convert an input speech into an unrecognizable encrypted output called cipher speech (encryption process), which makes it practically infeasible to recover the original speech without the encryption key. Secondly, a private key is used which converts the unrecognizable speech back to its original form (decryption process). Another popular algorithm is the ElGamal encryption system which is an asymmetric key encryption algorithm described by Taher ElGamal in 1985 [17]. Its security depends upon the ability of a hacker to compute discrete logarithms. Encrypted data can be transmitted using FSK digital modulation technique. Here FSK is generated using reflex klystron. Frequency Shift Keying (FSK) is one of the popular digital modulation technique in which the frequency of the carrier signal varies according to the digital signal changes. The frequency of output of a FSK modulated wave is high for a binary high input and is low for a binary low input. It plays an important role in long distance communication. Reflex klystron is basically a microwave generator where velocity modulation technique has been utilized to form a high energy density bunch of electrons which suitably reflect to generate high frequency RF oscillation in a re-entered cavity; It was used as a local oscillator in some radar receivers and a modulator in microwave transmitters the 1950s and 1960s.Demodulation and decryption is done at receiver.

The conventional MFCC used in [1], [6], [9] and [15] has the disadvantage in removing the specific band noise, has inherently low recognition rates. We have used Noise reduction using spectral gating filtering. This algorithm is based on the one outlined by Audacity for the noise reduction effect. The disadvantage of [2] and [4] is that it is able to recognize voice signals of very short time span containing at the most one word. Our study deals not only with voice clips containing one word but is also able to deal with voice clips containing several words. This is an enormous advantage over the above mentioned literatures. Since most of the voice recognition systems required in the present world needs to handle voice clips containing multiple words spanned over a long interval. In [5] the accuracy of voice recognition was not sufficient; we have improved the recognition accuracy by employing a digital filter. In [10] and [12], the degree of effect of voice disguise on the recognition rate varies with different disguising types. So it is not easy to understand if a voice is disguised. Our system can easily recognize a disguised voice because it compares the input speech signal and therefore the reference speech signal and allows the communication to happen when equality ratio is over 0.6. Experimental results in our system have shown that a disguised voice could not achieve an equality ratio of a minimum of 0.6. [11] has used LabView Programming Model which has several disadvantages like lot of memory is needed and also time consuming. Developer edition is very costly. It also has debugging issues. We have used Python which is open source language and resource efficient. We used Jupyter Notebooks embedded in Anaconda which is well-known software for executing Python programs. Several other IDEs are available which open source.

II. METHODOLOGY

Speech samples are recorded using mobile recorder in .mp3 format. But .wav format is desirable for working on the speech sample. Therefore, speech samples in .mp3 format are converted to .wav format using one converter application. Ploted the amplitude vs. time graph for each of the speech samples in MATLAB, which are uploaded into python based voice identification system. Jupyter notebooks embedded in Anaconda which is very well-known software for executing Python programs. Now, when a person speaks his speech is compared with the recorder speech samples. If it is matching with one of the speech samples, then he is an authenticated speaker and his speech is processed for transmission to the receiver. Otherwise, he is an unauthenticated speaker and transmission to the receiver will not take place. After a match is found, the spectrogram of the corresponding speech sample is plotted after eliminating the noise. RSA algorithm or Elgamal algorithm is applied on noise eliminated signal for encryption of speech signal. Digital signal was then passed through Reflex Klystron to convert this digital signal into modulated signal using FSK (Frequency Shift Keying) and transmitted to the receiver as shown in Fig.1. Reflex Klystron will assign two frequencies where high frequency (f_1)

is assigned for a binary high input and low frequency (f_2) is assigned for a binary low input. At the receiver, in order to identify f_1 and f_2, FSK signal is passed through a coupler which divides the corresponding signal into two parts. These two parts contain both f_1 and f_2 frequencies in same phase. The resulting two signals are passed through two different resonating cavities of frequencies of f_1 and f_2 to identify them. The resulting two signals are then summed up using adder circuit to get the original speech signal. This signal is amplified and applied to DAC (Digital to Analog Converter) to get back analog signal. RSA decryption algorithm or Elgamal decryption algorithm (whichever is applied) is applied on analog signal for decryption and get back original speech signal which is spoken as shown in Fig.2.

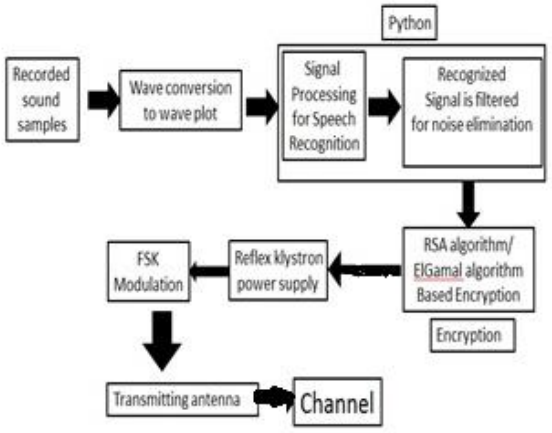

Fig.1. voice authentication, Encryption and modulation Block at transmitter

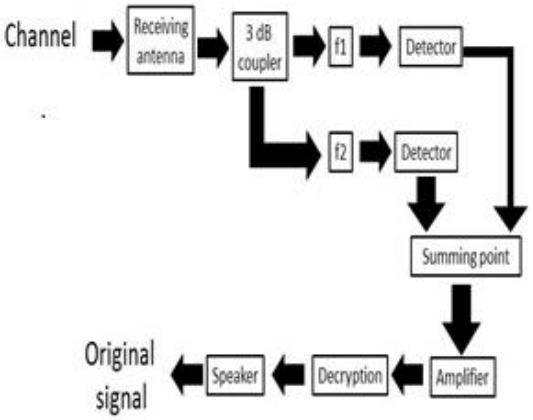

Fig.2.Demodulation and decryption at receiver

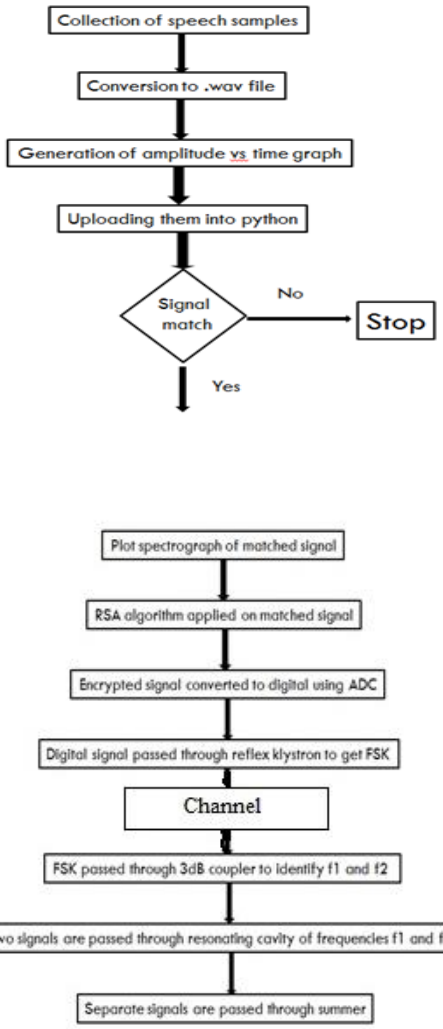

Fig.3 Flow chart of voice password authentication and Modulation and demodulation of speech signal

The characteristics curve for reflex klystron is shown below in Fig.4.

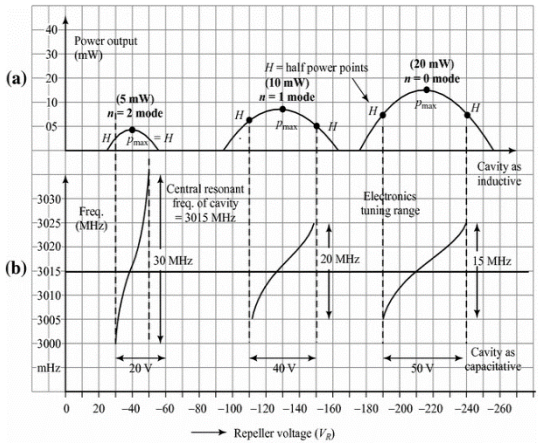

Fig.4. Characteristic curve of reflex klystron

Frequency Shift Keying is the digital modulation technique in which the frequency of the carrier signal varies according to the digital signal changes. FSK is a scheme of frequency modulation. The output of a FSK modulated wave is high in frequency for a binary High input and is low in frequency for a binary Low input. The binary 1s and 0s are called Mark and Space frequencies. Fig.5 is the diagrammatic representation of FSK modulated waveform along with its input.

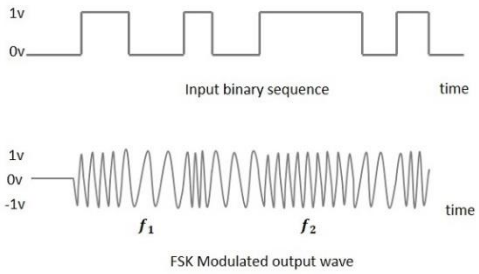

Fig.5. Frequency Shift keying

III. ENCRYPTION AND DECRYPTION TECHNIQUE USING RSA AND ELGAMAL ALGORITHM :

The RSA Algorithm[16] was named after Ronald Rivest, Adi Shamir and Leonard Adelman, who first published the algorithm in April, 1977. Since that time, the algorithm has been employed in the most widely-used Internet electronic communications encryption program. It is also employed in both the Netscape Navigator and Microsoft Explorer web browsing programs in their implementations of the Secure Sockets Layer (SSL), and by MasterCard and VISA in the Secure Electronic Transactions (SET) protocol for credit card transactions. The RSA Algorithm is only one implementation of the more general concept of public key cryptography. Typical encryption techniques use mathematical operations to transform a message (represented as a number or a series of numbers) into a cipher text. RSA Algorithm is used for public –key cryptography. RSA Algorithm involves three steps namely

Key generation, Encryption, Decryption.

Key Generation
Select p, q where p & q both prime, and p ≠ q
Calculate n=p ×q
Calculate Ø(n) = (p-1) × (q-1)
Select integer 'e' such that gcd (Ø(n),e) =1;1<e<Ø(n)
Calculate d, d≡e ^ (-1) (mod Ø(n)) or d.e ≡1 (mod Ø(n))
Public Key: PU = {e, n} and
Private Key: PR = {d, n}
Encryption
Plaintext: M<n
Cipher text: C= M ^ e mod n
Decryption
Cipher text: C
Plaintext: M= C ^ d mod n

ElGamal Algorithm [17]: **ElGamal encryption system** which is an asymmetric key encryption algorithm for public-key cryptography is based on the Diffie–Hellman key exchange. It was described by Taher ElGamal in 1985. ElGamal encryption is used in the free GNU Privacy Guard software, recent versions of PGP, and other cryptosystems. The Digital Signature Algorithm (DSA) is a variant of the ElGamal signature scheme, which should not be confused with ElGamal encryption. ElGamal encryption can be defined over any cyclic group G, like multiplicative group of integers modulo *n*. Its security depends upon the difficulty of a certain problem in related to computing discrete logarithms. **ElGamal** Algorithm involves three steps,
Key generation, Encryption, Decryption as in Fig.7 :

- **Key Generation**
  - Select a large prime as a q
  - Select x *to* be a member of the group **G** = < Zq^*, X >, x must be "$1 \le x \le q-1$"
  - Select g *to* be a primitive root (generator) in the group **G** = < Zq^*, X >
  - $y = g^x \bmod q$
  - Public key ← (g, y, q)
  - *Private key* ← x
- **Encryption**
  - Select a random integer r in the group **G** = < Zq^*, X >, r must be "$1 \le r \le q-1$"
  - $C_1 = g^r \bmod q$
  - $C_2 = (p \,.\, y^r) \bmod q$ // p is the plaintext
- **Decryption**
  - $P = [\, C_2 (C_1{}^x)^{-1} \,] \bmod q$

Fig.7 ElGamal algorithm

Proposed Methodology :
1. Start
2. Speech samples are recorded using mobile recorder.
3. Conversion into.wav format for ease.
4. Plotting the amplitude vs. time graph for each of the speech samples in MATLAB/Python.
5. Uploading it into Python for further analysis

6. Voice recognition is done by matching with a reference stored speech keyword data base.
7. After a match is found, the spectrograph of the corresponding speech sample is plotted in Python after eliminating the noise.
8. RSA algorithm or Elgamal algorithm is applied on noise eliminated signal for encryption of speech signal.

9. Digital signal was then passed through Reflex Klystron to convert this digital signal into FSK (Frequency Shift Keying) and transmitted to the receiver.
10. At the receiver, in order to identify f1 and f2, FSK signal is passed through a coupler which divides the corresponding signal into two parts. These two parts contain frequencies both frequencies f1 and f2 in same phase.
11. The resulting two signals are passed through two different resonating cavities of frequencies of f1 and f2 to identify them and then summed up using circuit to get the original speech signal.

12. RSA decryption algorithm or Elgamal decryption algorithm (whichever is applied in Step 8) is applied on analog signal for decryption and get back original speech signal which is authenticated in Step 6.
13. Stop

IV. RESULTS AND DISCUSSIONS

The relevant spectrograms obtained from the voice signal preceded by the plotting of wave plots and recognition and implementation of RSA algorithm on this sound signal are shown respectively in Fig 8 and Fig.9. It shows two such samples used for recognition.

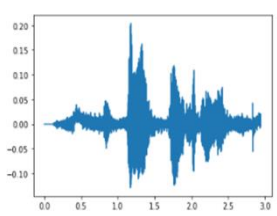

Fig8. Sample-1

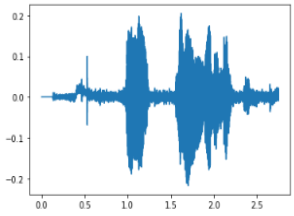

Fig9. Sample-2

The spectrograms after recognition of the correct voice signal and distinguishing into its components is shown below in Fig.10

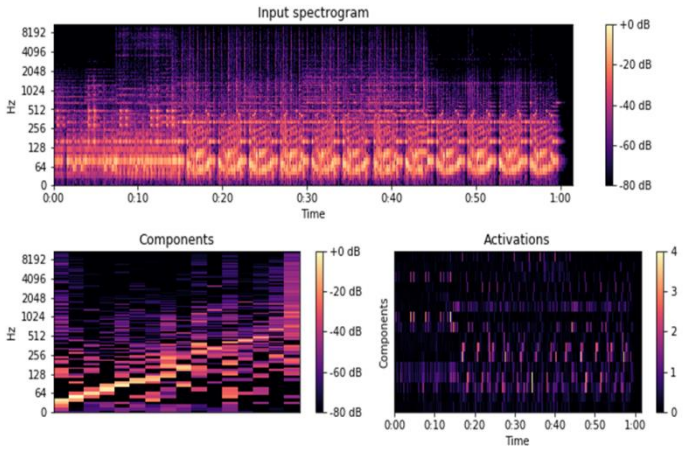

Fig.10.Spectrograms of the recognized voice signal and its components

The sound wave is further processed to plot its harmonic, percussive and full power spectrogram. This is depicted in Fig.11. These breakdowns are suitable when sound analysis is done at higher levels of processing.

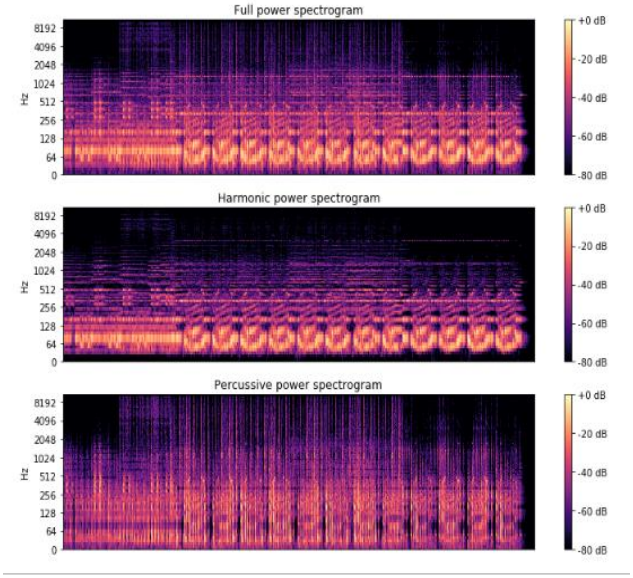

Fig.11. The three corresponding spectrograms

The ultimate stage in the encryption stage involves implementing RSA algorithm. The waveforms are shown below in Fig.12.

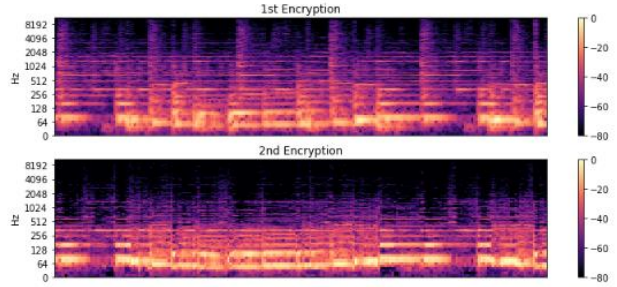

Fig.12. Encrypted Waves

The recovered waveform after applying the decryption algorithm is shown below in Fig.13.

This waveform is obtained based on software simulation.

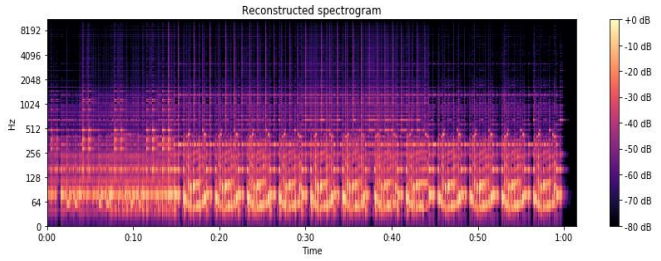

Fig.13. Reconstructed Spectrogram

The graph for klystron characteristics using external modulation is shown in Fig.14[22]

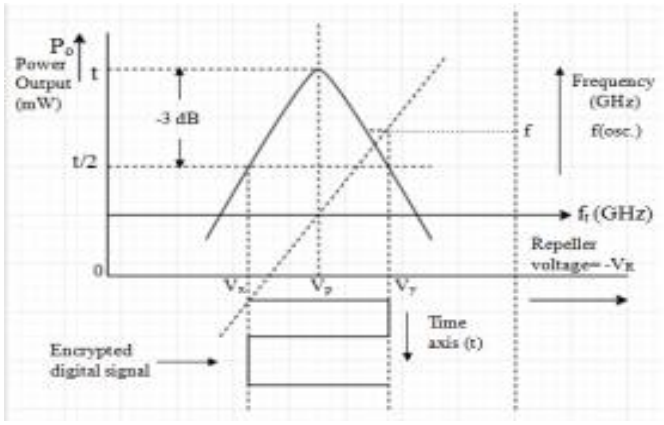

Fig.14. Graph of klystron characteristics using external modulation

The setup for obtaining this graph as a result is shown below in Fig.15.

Fig.15. Setup of klystron characteristics using external modulation

Comments on results: The results were obtained using python with the help of librosa, matplolib, numpy, PIL and glob. As the results are based on software simulation hence the original and reconstructed spectrograms shows almost match. The DAC was implemented and the output was brought into effect in the form of a LED signal.

CONCLUSION

In addition to RSA or ElGamal encryption techniques, voice based authentication of authentic person's gives additional security. Therefore, only authentic persons data would be taken for encryption and finally for transmission. The methodology can be used in highly secured environment like in defense applications. High power microwave devices like reflex klystron is used for generation of FSK modulated signals. The signals are reconstructed after FSK demodulation and decryption process. Similarity index of the reconstructed signal is very high with respect to transmitted signal. Here The encryption and decryption time of the RSA algorithm is better than the ElGamal algorithm. Ciphertext RSA has fewer numbers than ElGamal algorithm. The ElGamal algorithm has a ciphertext pair. Each encrypted plaintext will generate two ciphertext values. RSA algorithm and ElGamal algorithm are asymmetric algorithms which have different formulas for encryption and decryption. RSA algorithm is faster than ElGamal algorithm. Regarding security, the ElGamal algorithm will be more challenging to solve than the RSA algorithm because ElGamal has a complicated calculation to solve discrete logarithms.

REFERENCES

[1] H. Bae, H. Lee and S. Lee, "Voice recognition based on adaptive MFCC and deep learning," 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, 2016, pp. 1542-1546.

[2] N. C. Bui, J. J. Monbaron and J. Michel, "An Integrated Voice Recognition System," ESSCIRC '82: Eighth European Solid-State Circuits Conference, Brussels, 1982, pp. 158-161.

[3] S. J. Wenndt and R. L. Mitchell, "Machine recognition vs. human recognition of voices," 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, 2012, pp. 4245-4248.

[4] Yiu-Kei Lau and Chok-Ki Chan, "Speech recognition based on zero crossing rate and energy," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 33, no. 1, pp. 320-323, February 1985.

[5] Y. Yamazaki, M. Tamaki, C. Premachandra, C. J. Perera, S. Sumathipala and B. H. Sudantha, "Victim Detection Using UAV with On-board Voice Recognition System," 2019 Third IEEE International Conference on Robotic Computing (IRC), Naples, Italy, 2019, pp. 555-559.

[6] J. Pak and M. Kim, "Convolutional Neural Network Approach for Aircraft Noise Detection," 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Okinawa, Japan, 2019, pp. 430-434.

[7] H. AlShu'eili, G. S. Gupta and S. Mukhopadhyay, "Voice recognition based wireless home automation system," 2011 4th International Conference on Mechatronics (ICOM), Kuala Lumpur, 2011, pp. 1-6.

[8] N. Aktar, I. Jaharr and B. Lala, "Voice Recognition based intelligent Wheelchair and GPS Tracking System," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, 2019, pp. 1-6.

[9] M. S. I. Sharifuddin, S. Nordin and A. M. Ali, "Voice Control Intelligent Wheelchair Movement Using CNNs," 2019 1st International Conference on Artificial Intelligence and Data Sciences (AiDAS), Ipoh, Perak, Malaysia, 2019, pp. 40-43.

[10] T. Tan, "The effect of voice disguise on Automatic Speaker Recognition," 2010 3rd International Congress on Image and Signal Processing, Yantai, 2010, pp. 3538-3541.

[11] S. Pleshkova, Z. Zahariev and A. Bekiarski, "Development of Speech Recognition Algorithm and LabView Model for Voice Command Control of Mobille Robot Motio," 2018 International Conference on High Technology for Sustainable Development (HiTech), Sofia, 2018, pp. 1-4.

[12] N. Obin and A. Roebel, "Similarity Search of Acted Voices for Automatic Voice Casting," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 9, pp. 1642-1651, Sept. 2016.

[13] R. A. Rashid, N. H. Mahalin, M. A. Sarijari and A. A. Abdul Aziz, "Security system using biometric technology: Design and implementation of Voice Recognition System (VRS)," 2008 International Conference on Computer and Communication Engineering, Kuala Lumpur, 2008, pp. 898-902.

[14] A. Prodeus and K. Kukharicheva, "Automatic speech recognition performance for training on noised speech," 2017 2nd International Conference on Advanced Information and Communication Technologies (AICT), Lviv, 2017, pp. 71-74.

[15] Che Yong Yeo, S. A. R. Al-Haddad and C. K. Ng, "Animal voice recognition for identification (ID) detection system," 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, Penang, 2011, pp. 198-201.

[16] R.L. Rivest, A. Shamir, and L. Adleman , A Method for Obtaining Digital Signatures and Public-Key Cryptosystems , Communications of the ACM ,February 1978

[17] T. Elgamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in IEEE Transactions on Information Theory, vol. 31, no. 4, pp. 469-472, July 1985.

[18] Chakraborty Mohuya & Mallick, Amiya. (2010). AES Encrypted FSK Generation at X-Band Frequency using a Single Reflex Klystron. Wireless Communication over ZigBee for Automotive Inclination Measurement. China Communications. 7. 1-9.

[19] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. Commun. of the ACM, 21:120 - 126, 1978.

[20] Tahir, Ari. . Design and Implementation of the RSA Algorithm using FPGA. International Journal of Computers & Technology. Vol 14. 6361-6367.(2015)

[21] Richard W. Middlestead, "Frequency shift keying (fsk) modulation, demodulation, and performance," in Digital Communications with Emphasis on Data Modems: Theory, Analysis, Design,Simulation,Testing, and Applications, Wiley, 2017, pp.207-225

[22] Prashnatita Pal,Bikash Chandra Sahana,Jayanta Poray, and Amiya Kumar Mallick “Generation of Encrypted FSK RF Signals for Secured Communication Inspired with High Frequency Technique “International conference on Recent Trends in Artificial Intelligence, IOT, Smart Cities & Applications (ICAISC-2020)