IMPLEMENTATION OF CLOUD BASED INTELLIGENT WIRELESS NETWORKS USING DEEP LEARNING

ANISH K
Department of TCE
Dayananda Sagar College of Engineering Bangalore, India

SAGAR S BHAT
Department of TCE
Dayananda Sagar College of Engineering Bangalore, India

ARVIND RAO
Department of TCE
Dayananda Sagar College of Engineering Bangalore, India

VINOD B DURDI
Department of TCE
Dayananda Sagar College of Engineering Bangalore, India

Abstract - Communication can take many different forms, including gestural, written, vocal, and visual. These types can frequently be converted into one another's formats. For someone who cannot understand or does not have the ability to recognise the type of communication being employed, this step becomes extremely crucial. This project's main goal is to convert text formats into speech formats and vice versa using the MFCC Algorithm, which is more accurate and efficient than conventional systems. The project's target audience includes people who cannot read i.e illiterates, and those who have visual impairments i.e blind people.To make the system more user-friendly, a web application is created specifically for this purpose. All messages are uploaded to the cloud, i.e., the firebase console, via a wireless channel.The project makes use of an Arduino UNO and an ESP8266 NodeMCU Wi-Fi module for this purpose. In today's world, any communication involving the internet and network applications necessitates security
And protection. For this purpose, a powerful data hiding RBA algorithm has been developed.

Keywords: **Wireless Communication, Speech Recognition, Deep Learning, Cryptography, Cloud Computing, Web Development**

INTRODUCTION

Communication is a necessary life skill. Better understanding results from effective communication [1]. In our daily lives, we use two types of communication: verbal and electronic [2]. They may appear to be similar, but they are vastly different.
Electronic communication is the use of electronic devicesand systems for acquiring, processing, storing, displaying, analysing, protecting, and transferring information [3]. Because of the invention of TTL (transistor-transistor logic), whichled to the development of IC, communication electronics has seen a rapid growth in research and development (integrated circuits). The communication system is classified into two types based on how signals are transmitted: wired communication systems and wireless communication systems [4]. Wired communication occurs when there is a physical medium between the transmitter and receiver through which the information signal is transferred. They are typically employedfor short-distance communication. They do not necessitate signal modulation. A wireless communication system, on the other hand, is one in which there is no physical medium between the transmitter and the receiver. An antenna is used for both transmitting and receiving the signal. Wireless systems are well-suited for long-distance communication. Its nature is more complicated than that of a wired system. In wireless systems, information is transmitted through an electromagnetic wave of high frequency. The benefits of wireless system include convenience, mobility, productivity, easy setup, expandability, security and reduced cost. Hence it is more suitable for communication. The type of information or data transmitted via a wireless system can be divided into five categories: text, image, audio, video, and animation [5]. They can be represented in two ways: analogue and digital. Digital signals are discrete and finite, whereas analogue signals are continuous [6]. Information, whether analogue or digital, requires protection from cyber-attacks and theft [7]. Most encryption techniques are regarded as 'digital' problems that necessitate the manipulation of bits via the execution of algorithms. These algorithms aid in the conversion of sensitive and confidential data, known as plain text or original message, into cypher text or encrypted message, which is unreadable and difficult to comprehend [8][9]. After encryption, the data is broadcasted into a wireless channel, which uses air or free space as a medium of communication. Antennas are to communicate over free space [10]. An antenna is a device that converts an RF signal travelling on a conductor into an electromagnetic wave. Electromagnetic waves are produced byvibrations between an electric field and a magnetic field that oscillate perpendicularly to each other [11]. Computers that serve as internet sources generate and receive EM waves [10]. The Internet allows us to communicate in real time by sharing and seeking information. It is a global network of billions of computers

that are linked together. Cloud computing is a result of the internet, which provides many resources as services such as software, platforms, and infrastructure [12][13]. Because of virtualization technology, it is possible. It generates a simulated and digital virtual computer that behaves like a physical computer, complete with its own set of hardware devices. The ESP8266 Wi-Fi module is used to connect to this network [11]. It includes TCP/IP networking software as well as the capabilities of a microcontroller. It is a WIFI-enabled system on chip (SoC) that is primarily used for the development of Internet of Things and embedded applications. A web application has been created using the Wi-Fi network, and the information has been made application specific, portable, and secure. The web app was created with MongoDB as the back-end server [14][15].

PARAMETERS	LPCC ALGORITHM	MFCC ALGORITHM
SUCCESS RATE	82.22 %	86.67 %
POWER CONSUMPTION	858 micro watts	560 micro watts
TIME CONSUMPTION	5.16 seconds	0.91 seconds
ERROR RATE	17.78 %	13.33 %

Table 2.1

The second section of this paper compares existing and proposed algorithms.

COMPARISON BETWEEN EXISTING AND PROPOSED ALGORITHM

LPCC is linear prediction system more efficient that predict future values of data based on previous sample whereas MFCC is Mel frequency Cepstrum coefficient that considers nature of speech while it extracts the features [4][5][6]. The success rate or performance percentage is the highest in MFCC algorithm as seen in the table [citation] whereas the power consumed time consumed and also the error rate is highest in LPCC hence, this project has made use of MFCC algorithm instead of LPCC

The following table 2.1 compares the above-mentioned algorithms.The following table 2.2 compares various cryptographic algorithm based on different parameters namely computational power, time period, memory consumption and system efficiency

Table 2.2

Parameters	DNA Based Algorithm	CBR Algorithm	RBA
COMPUTATIONAL POWER	30 %	15 %	7.1 %
MEMORY CONSUMPTION	490 KB	13 MB	17 MB
TIME PERIOD	3.832 seconds	4.547 seconds	Depends on length of the text input. Max(3.165 seconds)
SYSTEM EFFICIENCY	85.2 %	76.36 %	93.3 %

Based on the discussions made above it is evident that MFCC is more efficient and consumes lesser time when compared to existing algorithm. In addition to this it can also be noted that the proposed cryptographic algorithm is more secured and consumes less power. Thethirdsectioninthispaperanalyzesthearchitectureoftheproposedsystem.

METHODOLOGY

In the previous section analysis of different algorithm has been made and it can deduce that MFCC and RBA algorithm are better than their counterparts in terms of performance and security. Hence these two algorithms have been employed in this project. The following figure fig 3.1 shows implementation of the proposed system.

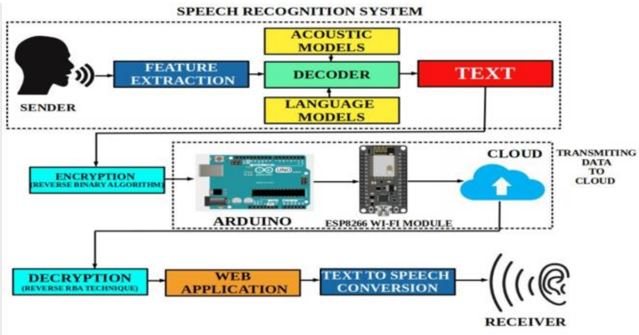
3.1 SPEECH RECOGNITION SYSTEM

This process is a capability that enables a program to process a human speech into a written format. It is mainly divided into two consecutive steps namely feature extraction and decoding.

3.1.1 FEATURE EXTRACTION: In this process the speech signal is transformed to a logical representation that is more discriminative and reliable than the actual signal.Here the initial set of data is reduced by identifying key features of the data for machine learning. This process mainly helps in improving overall quality of audio signal by extracting key features like pitch, power and resonance that makes it much more identifying.Feature extraction identifies the most discriminating characteristics in signals, which a machine learning or a deep learning algorithm can more easily consume.

Fig 3.1: Block Diagram of the Proposed System

The fig 3.1 shows the overview of the proposed system. It has 6 methods, the first method is



In this project MFCC algorithm has been implemented for feature extraction due to it's good error detection, better performance and efficiency.
Here the formula 3.1 is used for calculating Mel frequency Cepstrum Coefficient

$$\boldsymbol{M(f) = 1125\ln(1+\frac{f}{700})}$$

$$Where\ f = frequency\ of\ speech\ signal$$

Equation 3.1 Formula for calculating Mel - frequency

3.1.2 DECODING:

The next step is decoding which helps in generation of text from the audible speech. This leverages the functionalities of two models namely Acoustic and Language models. An acoustic model helps in generation of words by mapping a phoneme i.e., a basic unit of sound to every word. The model is learned from a set of audio recordings and their corresponding transcripts. On the other hand, a language model helps in identifying any errors in syntax, semantics of the given generated words with the help of various statistical and probabilistic techniques.

3.2 ENCRYPTION:

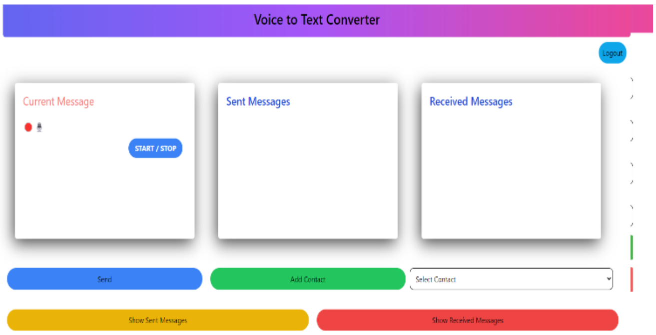
The first step in the backend is to encrypt the data provided. After sending the text generated by Speech Recognition to the backend, an encryption module that employs the RBA algorithm and converts the plain text to cypher text is activated. The above-mentioned generated text is fed into the encryption module. The encryption module is made up of pre-defined keys that are used to generate different character sets for the input string using the RAA technique, which employs the reverse string concept. The generated text is now partially encrypted, but an intruder can easily crack it. As a result, additional parameters are included to increase security and maintain the confidentiality of the text.

3.3TRANSMITTING DATA TO CLOUD:

This is the second backend process. After sending the data to the backend, it is encrypted using the RBA Algorithm and then sent to the cloud using an Arduino Uno and an ESP8266 Wi-Fi module. The data or message generated at the front end is sent to the backend and serially transmitted to Arduino via a python script that runs in an infinite loop on the sender's device. As a result, whenever a message is detected in the backend, it is sent to the Arduino uno via this python script. Because the Arduino cannot be connected to the internet, a nodemcuWi-Fi module was used to connect the Arduino and the ESP8266 NodeMCUWi-Fi-module serially via the Tx and Rx pins. Then after transmitting data to the Wi-Fi-module, the data is sent to mongo dB cloud where the data are stored in NoSQL Format. Hence the data which was device specific or restricted to a device can be accessed from the cloud to any device

3.4 CLOUD TO WEB APPLICATION:

Following the transmission of encrypted data to the cloud, this data can be accessed via the frontend website through the two dedicated buttons. One displays the sent messages, while the other displays the received messages, completing the full duplex communication cycle. This data retrieval process is accomplished by connecting an asynchronous listener to the cloud server. Because the data received from the cloud is encrypted, a function on the receiving end that performs decryption using the Reverse RBA algorithm has been implemented to convert the data back to its original format. Thus, the desired text is encrypted, stored, displayed, and updated in real time on the project's front-end website.



3.5 DECRYPTION:

After retrieving encrypted data from the cloud, it must be converted back to its original format. As a result, the Reverse RBA technique is used to decrypt the data. Because encryption employs a complex algorithm, the decryption algorithm employed is equally complex and consists of four steps: obtaining the equivalent ASCII value of the binary number, converting the ASCII value to their respective character, Reverse Character Rotation, Reverse RAA Technique.

3.6 TEXT TO SPEECH:

After displaying the text message this project deploys a TTS synthesis translator that converts this text message to audio output. TTS, or text-to-speech synthesis, is the automatic transformation of a text into speech that as closely as possible imitates the sound of a native speaker of the language reading that text. A Python library and command-line programme called gTTS (Google Text-to-Speech) is used to interact with the Google Translate text-to-speech API. When using TTS, the text is received as the input, followed by the TTS engine computer algorithm analyses, pre-processes, and synthesises the text talk that includes some mathematical models. TTS's engine typically produces sound data with an audio output.Thefourthsectionexaminestheoutcomeoftheabove-mentionedproposedmodel.

RESULTS

The following images depict the output of the proposed system.

Fig 4.1: Login Page

The above fig 4.1 shows the website's login page. The website has three required fields: the user's mobile number, email address,

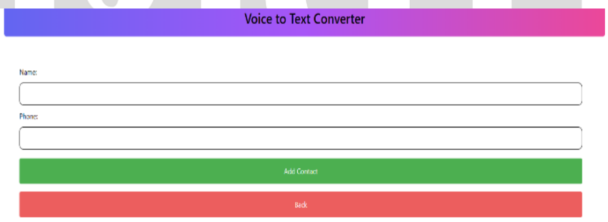
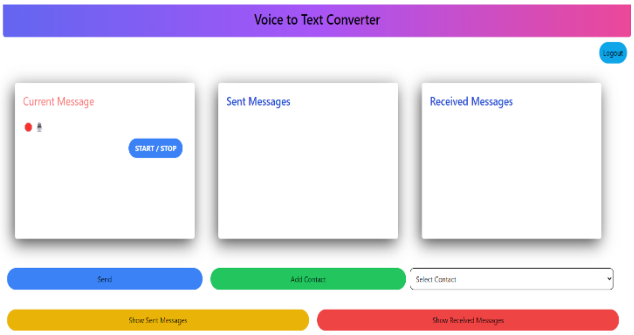


Fig 4.2: Sign up page

and password, as well as two buttons: sign in and sign up. Sign up buttons lead to a sign-up page where new users can register. The sign in button instructs the website to check all required fields and grants access to the homepage only if all three required fields are appropriate.

The fig 4.2 shows the sign-up page. he signs up page is where new users can register for the application. Users can register by providing basic information such as their name, phone number, password, and email address. In addition to this there is also a dedicated button that returns the user back to the login page.



Dedicated button that returns the user back to the login page.

The above fig 4.3 illustrates the home page of the website. The home page consists of 3 main sections which are voice to text converter section, sent message section and received message section. The first section is used to convert voice to text, there is a dedicated button to record the audio input from the user.There is also a send button which transmits user data to the cloud. The second and third sections have dedicated buttons for displaying the text output of sent and received messages for a specific contact as well as converting them to audio output. The contact can be chosen from the drop-down box. There is also an add contact button where the user can add additional contacts.

Fig 4.4: Add Contact page

The figure 4.4 shows the add contact page, where the user can add new contacts to his contact list.

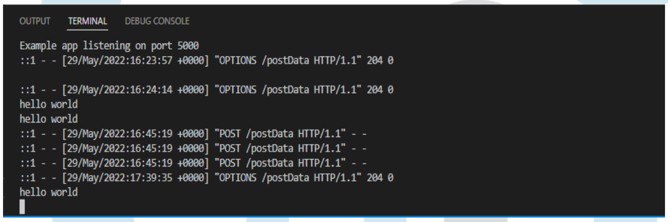


Fig 4.5: Data processing at Backend using Node.js

The figure above 4.5 depicts how by using the SEND button in the website we can initiate the process of transmitting the data to backend

Fig 4.6: Encryption and Decryption

The above figure 4.6 illustrates the encryption module written in Python and implemented to secure the data provided by the user. The encryption module makes used of RBA algorithm which utilizes both symmetric substitution and transposition methods. The reverse process of encryption which is decryption happens at the receivers end to obtain original information.

Fig 4.7: Hardware Implementation

The above figure 4.7 describes the hardware components used for the working of thisproject. It has two components basically-the first is an Arduino Uno ATMEGA 328P which is an microcontroller and an esp8266 Wi-Fi module connected serially with it The Wi-Fi module is used so that we can connect to the cloud via the internet

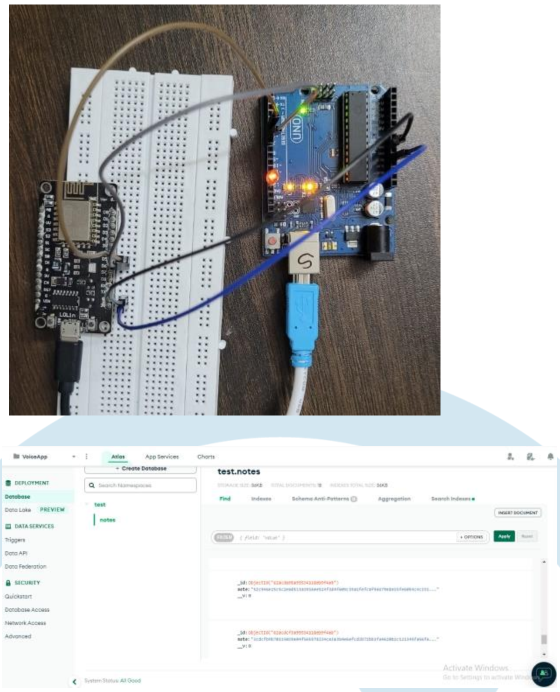


Fig 4.8: Transmitted real-time encrypted data in the Cloud

The cloud platform used for this project is depicted in the figure above 4.8. The data for this project is stored in the mongo atlas cloud. The cloud is a software platform used in this project to load encrypted versions of real-time user input data. All of the data is saved in json format.

Fig 4.9: Computational Consumption of CPU by different algorithms

Key:

- LPCC ALGORITHM
- MFCC ALGORITHM
- DNA BASED - ALGORITHM
- RBA ALGORITHM

Based on the above fig 4.9, following table 4.1 depicts the performance analysis of different algorithms

Table 4.1

Parameters	MFCC	LPC	RBA
Power	560 micro wats	858 micro wats	262 micro wats
Time	0.91 seconds	5.16 seconds	3.165 seconds
CPU Consumption	12.2%	15%	7.1 %
Efficiency	86.67 %	82.2 %	93.3 %

The above table 4.1 concludes that though power and CPU consumption of MFCC Algorithm is less and it also provides faster and efficient results.

Note: The following analysis have been deduced after integrating different speech algorithms with Reverse Binary Algorithm.

CONCLUSION

A front-end plus a back-end make up a text-to-speech system (or "engine") Two key duties lie in the front-end. It does this by first turning raw text that includes symbols like numbers and abbreviations into the equivalent of written words. Common names for this procedure include text normalisation, pre-processing, and tokenization. Each word is subsequently given a phonetic transcription by the front-end, which also separates and labels the text into prosodic units like clauses, phrases, and sentences. Conversion from text to phoneme or grapheme to phoneme is the process of assigning phonetic transcriptions to words. The symbolic language representation produced by the front-end is made up of prosody data and phonetic transcriptions. The symbolic representation of language is next transformed into sound via the back end, also known as the synthesiser.

ACKNOWLEDGEMENT

The researchers would like to acknowledge their colleague for the excellent contributions and discussions on the topics mentioned above.

REFERENCES

[1]. Simon DobriSek, et al., (2019), "HOMER: a Voice-Driven Text-to-Speech System for the Blind IEEE International Symposium on Industrial Electronics (ISIE), July 2019.
[2]. ItunuoluwaIsewon, et al., (2018), Design and implementation of text to speech Conversion for Visually Impaired People " International Journal of Applied Information Systems (IJAIS)
[3]. Rev," All You Need to Know About Automatic Speech Recognition Transcription Models" from https://www.rev.com/blog/guide-to-speechrecognitiontranscriptionmodels#:~:text=They%20provide%20a%20way%20of,Random%20Fields%2C%20and%20Neural%20Networks
[4]. Hyung-Suk Kim ,"Linear Predictive Coding is All-Pole Resonance Modeling", Center for Computer Research in Music and Acoustics, Stanford University
[5]. H Hermansky , "," Perceptual linear predictive (PLP) analysis of speech", from https://pubmed.ncbi.nlm.nih.gov/2341679/
[6]. uday200, "MFCC Technique for Speech Recognition", from https://www.analyticsvidhya.com/blog/2021/06/mfcc-technique-for-speech-recognition/
[7]. Deepraj Pradhan, (2020), et al., " Cryptography Encryption Technique Using Circular Bit Rotation in Binary Field " 8th International Conference on Reliability, Infocom Technologies and Optimization, 2020
[8] AllaLevina, (2021), et al., "Implementation White-Box Cryptography for Elliptic Curve Cryptography" 10th Mediterranean Conference on Embedded Computing (MECO), 2021
[9]. Bahubali Akiwate (2018), et al., "A Dynamic DNA for Key-based Cryptography", International Conference on Computational Techniques, Electronics and Mechanical Systems, 2018
[10]. Sasikumar Gurumurthy, (2019), et al., " Recent Trends In Wireless Technology " National conference on "Network Technologies" [NCNT-2019 January]
[11]. R. Tsafack Pierre, (2019), et al., "INVESTIGATION OF WI-FI (ESP8266) MODULE AND APPLICATION TO AN AUDIOSIGNAL TRANSMISSION". DISSERTATION/ BATCHANOU TATAP PERETS ARNAUD/ FE15A031/LEVEL 500/ UB/FET/2018-2019
[12]. Mohamed Abdalla Mokar, (2019), et al., "Using Firebase Cloud Messaging to Control Mobile Applications" International Conference on Computer, Control, Electrical and Electronics Engineering (ICCCEEE19), 2019
[13]. Chenhao Huang, (2020), et al. "Deciding When to Trade Data Freshness for Performance in MongoDB-as-a-Service" IEEE 36th International Conference on Data Engineering (ICDE), 2020
[14]. Stephan Boersma, (2021), et al, "React-bratus: Visualising React Component Hierarchies", Working Conference on Software Visualization (VISSOFT), 2021
[15]. Maha A. Sayal ,(2016), et al., ., " Building web applications using Cloud Computing" International Joural of Software and Web Sciences-(IJSWS)