INTELLEGENT VAULT WITH VOICE BASED APPROACH

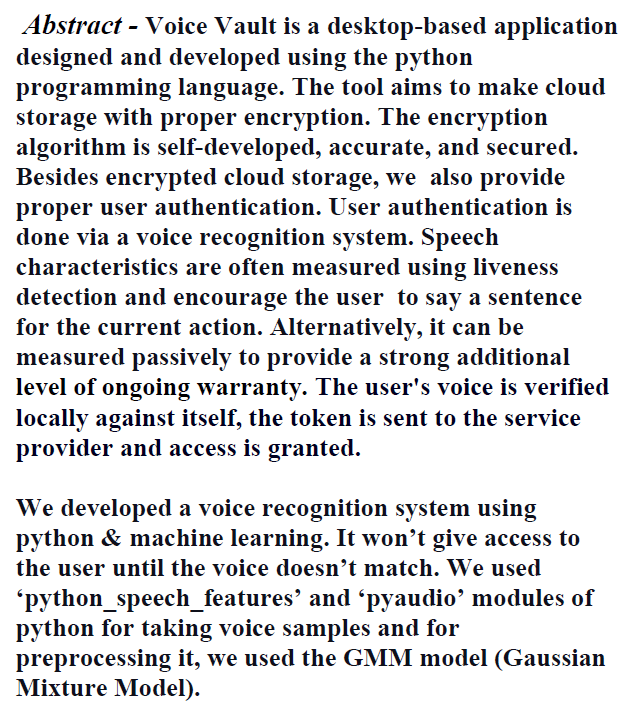
Mr. BODDU NITISH KUMAR RAO, Mr. SINGULURI SAI MITHILESH,

Miss. KAVALAPATI CHANDRA PRIYA, Mr. RAJAMUDILI AJAY

Mrs. Ch. Swapna Priya Associate Professor

Department of Computer Science and Engineering

Vignan’s Institute of Information Technology, Visakhapatnam - 530046



# **INTRODUCTION**

The study of manipulating, managing, transforming, and encoding information is known as computer science. It covers the concept and strategies of records processing in virtual computers, in addition to the layout of pc hardware and software program and pc applications. This branch of study

uses computer science theories to design, test, and analyze concepts. It is possible for computers to solve a variety of problems in business, engineering, healthcare, and other areas, but the steps required to find a solution depend on the domain of the problem and the Computer Science skills of the problem-solver. Knowledge of both are required. Engineering teaches people how to make hardware and software.

**II. LITERATURE REVIEW**

The speaker recognition model using GMM was introduced using text-independent speaker recognition with the new method. The feature vector was estimated using the MelfrequencyCepstral Coefficients (MFCC) of the preprocessed audio signal. The audio signal is extracted in the form of 10 to 15 feature vectors and converted to frames. GMM is used to train the total voice features. Estimate the speaker model using the maximum likelihood method. The maximum likelihood detector algorithm is used in the decision-making process. Recent advances in speech processing technology have improved the performance of speech recognition systems and require further improvements.

This study compares several parameters used in an automated speech recognition system to improve system performance. Because it employs biometric features derived from a natural action, speaking, voice recognition is one of the simplest kinds of automated identification. Speech Identification (SI) & Speech Verification (SV) are the two forms of voice recognition. The purpose of voice identification is to figure out which known group of sounds the input sample best fits. Voice confirmation is twofold choice that entails determining if a voice signal belongs to a specific individual or not. Text-dependent and text-independent voice identification are the two tasks. The spoken sentence is recognized by the system with a text-dependent identifier, but not with the text-independent identifier. As part of this task, we evaluated the model parameters that were taken into account in an automatic speech recognition system. Speech recognition is a technology that automatically recognizes who is speaking based on the individuality of the speech wave.

The main goal of this study is to use the provided approach to improve the computational speed, approximation quality, and accuracy of language identification systems. Usefulness of feature extraction strategies for more robust voice recognition will be investigated in future research. To improve accuracy, the hybrid classifier is also examined with better correction functions.

**Audit on Voice Recognition Technique**

The human voice is the most common and primary means of communication. The use of voice is a common method of computer interaction. Voice is the primary means of communication between people and the most natural and efficient way to exchange information. The most common way of changing over a sound sign into an arrangement of words utilizing a calculation executed as a PC program is called acknowledgment. One of the most captivating parts of sign handling is voice handling. Digitalizing voice has been a huge difficulty in the recent years. People naturally expect voice interfaces with computers because spoken language dominates human communication. This paper discusses different types of voices. Voice recognition is a subset of pattern recognition. It's possible to think of the voice recognition system as having four stages of operation 1) analysis 2) feature extraction 3) modelling 4) testing. For the extraction of voice features, MFCC is used, with GMM and HMM being the best choices.

**The identification of voices using Mel frequency cepstral coefficients**

In order to reduce the amount of data processed, we used the Mel Frequency Cepstrum Coefficient (MFCC) and vector quantization techniques. The task of voice identification is to use speech samples or speech from a speech population to determine the identity of the person who generated the speech. A sample of voice is used in voice verification to verify whether the voice was actually generated by the person who claims it.

For voice identification, MFCC technique has been used. VQ is used to reduce the extracted feature's data. According to the findings, the system's identification rate increases as the number of centroids increases. It has been determined that the combination of mel frequency and humming window results in the best results. In this study, we demonstrate that even a linear scale can have a reasonable detection rate when using more centroids, even if the number of votes increases. Therefore, you will need to increase the number of centers of gravity when the number of votes increases. However, as the number of voices increases, the recognition rate drops significantly on a linear scale. Mel scale is also less susceptible to vocal cord changes over time. By incorporating HMM, you can improve segmentation efficiency and accuracy while dealing with crosstalk, laughter, and a typical voice. The extracted parametric representation of the acoustic signal can be used to further improve the identification rate using a more effective normalization algorithm.

**GMM Training**

MFCC features are always trained on GMM models. The main key steps in GMM training are clustering and expected value. Here, the cluster number of each vector is obtained by the general K-Mean algorithm .

This method was useful for setting the centroids from the vectors observation. The centers are returned by the cluster model and for each clusters it refers to the nearest member. As a result of the k means algorithm, distortion is minimised by taking the square of the distance between every observation and the centroid. The distance is found using the euclidean distance between the two points.

**Algorithm**

To estimate the parameters of the GMM model from the training data, maximum likelihood (ML) estimation is used. The expectation-maximization approach is used to repeatedly acquire ML parameter estimations. Mean , Variance , Weight , and log Likelihood are all returned.

**GMM Recognition**

The speaker is identified via GMM recognition using log probability. It recalculates the speech vector's log likelihood and compares it to the value previously saved. Access to the full speaker is granted by a log likelihood equal to the stored value.

**III. System Requirements**

**3.1. Windows / Linux / Mac OS**

The proposed application is system independent.

**3.2. Microsoft Visual Studio 2010 Or Higher**

This is an IDE that is used to develop this windows based desktop application.

**3.3. SQL**

To query the data My SQL is used which is an open source relational database management system.

**3.4. Tkinter**

Used to create Graphical User Interface for the application.

**3.5. Cryptography And Steganography**

Algorithms associated with these techniques would help develop this application.

**IV. Algorithm**

AES plain text must be 128 bits long. The

key size can be 128, 192, or 56 bits. 128-bit AES encryption pseudocode:

A. Performs a one-time initialization process:

a. Expand the key of 16-byte

b. Performs initialization of a 16-byte plain text block named State

B. In each round, continue with the next operation.

a. Apply the S-Box to state

b. Rotate plaintext line K by K bytes

c.Performing a column merge operation

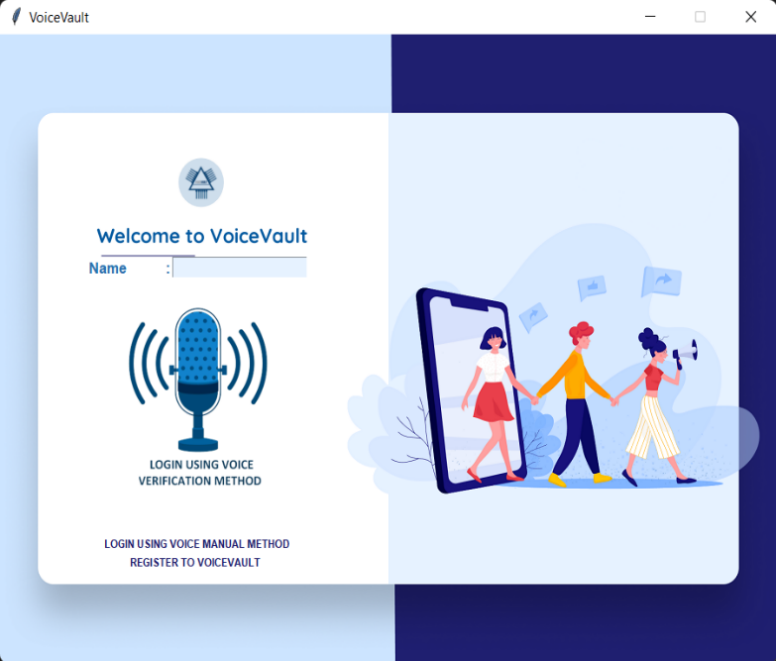
d. XOR state with the button.

**V. System** **Working**

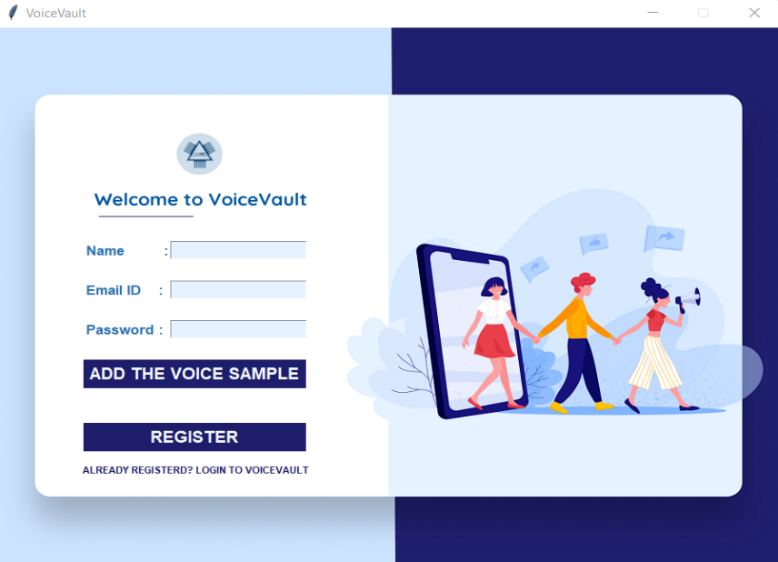
A user must first create an account with a username and master password. The user will be authorised and logged in using voice passwords after the initial login. All logins will be based on voice recognition methods and will be confirmed and logged into the interface as a result of this. After that, the user will be directed to a home page. The file will be locked using steganography technique - AES(Advanced Encryption Standard) algorithm and uploaded to the cloud storage while being uploaded. To unlock the text file, the user must first locate and download the encrypted file from the database. It'll be saved in the Downloads folder. This application is extremely user-friendly. It decreases the amount of work required to remember all of the passwords.

**VI. Diagramatic Representation of Application**

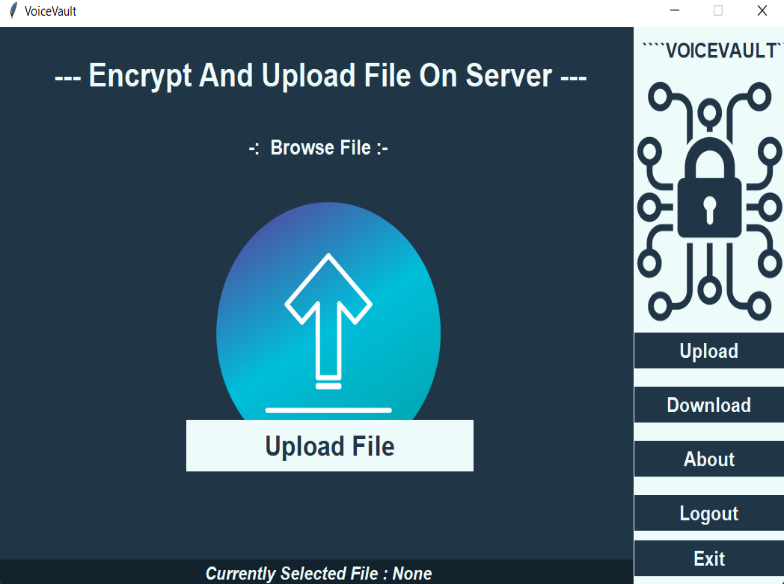
**Login Page**



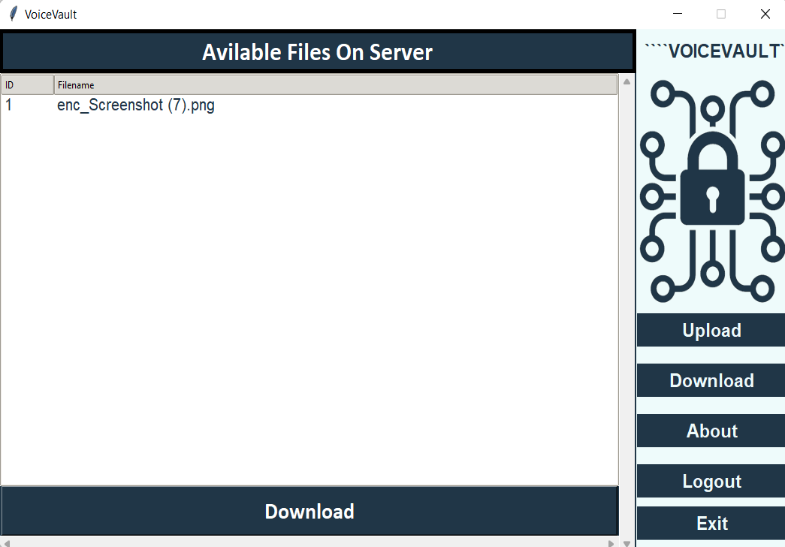
**Registration Page**



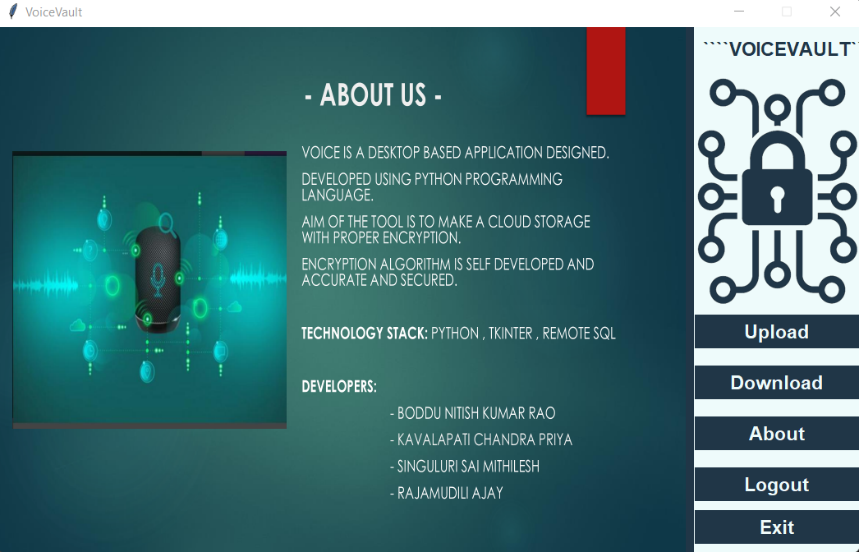
**Encryption/Uploading**



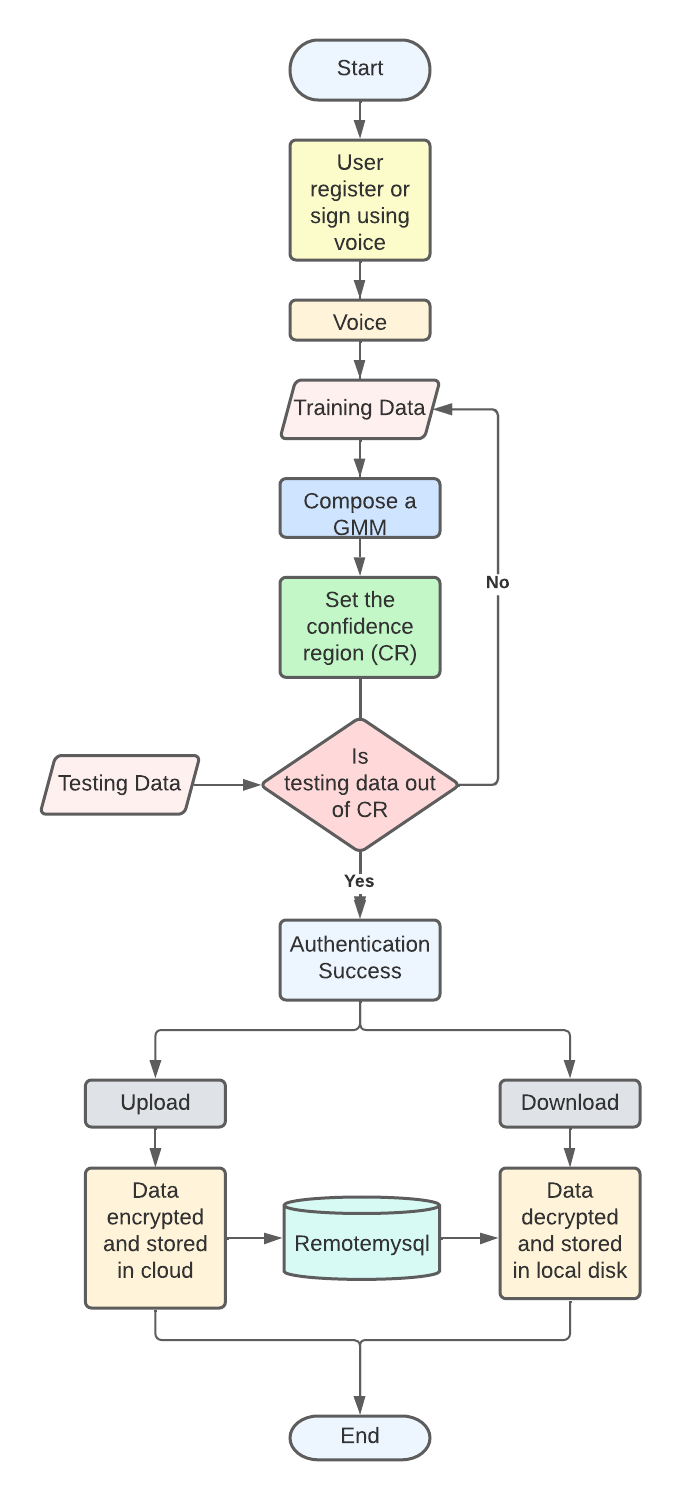
**Decryption/Downloading**



**About Page**



**Flowchart**



**VII. Conclusion**

This chapter concludes the proposed system. It is intended to develop a real-time voice identification system. For feature extraction, uses MFCC, and for training, it uses GMM.   The voice is first captured using a microphone, and then voice features are retrieved. The hamming window is used to reduce dis-continuities at the frame's edge, resulting in smooth frequency transmission in speech signals. Mel Frequency Cepstral Coefficients produces 15 MFCC coefficients using 40 Mel filters. These coefficients are subsequently sent to GMM to be used in the training phase. Users are identified by comparing the logarithmic probability with the system specified threshold. It decreases the amount of work required to remember all of the passwords. The application aids in the security of all text files, photos, videos, and other media. It prevents unauthorised users from gaining access. The application is made even more secure by using a voice security technique for authentication and then encrypting the files. Cryptography technology (AES Encryption) increases the security system by virtually eliminating the possibility of a compromise. As a result, a new security solution has been developed that allows you to lock and unlock your files while also storing them in the cloud. This programme is also extremely user friendly and cost-effective.

**VIII. References**

**[1].** Do, Minh N. "An Automatic Speaker Recognition System", Digital Signal Processing Mini-Project.

June 2005.

**[2].** Md. Rashidul Hasan, Mustafa Jamil, Md. Golam Rabbani and Md. Saifur Rahman, "Speaker identification using Mel frequency Cepstral coefficients", 3rd International Conference on Electrical & Computer Engineering ICECE 2004, December 2004.

**[3].** Hagai Aronowitz, David Burshtein, " Efficient Speaker Identification and Retrieval ". http://en.wikipedia.org/wiki/MATLAB. http://www.mathworks.com/help/pdf\_doc/matlab/ getstart.pdf .

**[4].** Archana Shende, Subhash Mishra and Shiv Kumar, " Comparison of different parameters used in GMM based automatic voice recognition ", International Journal of Soft Computing and Engineering (IJSCE), July 2011.