AI Based Online Verification of Scheme Beneficiary

Vibhuti Velgekar[1]*, Shreya Sawant[2], Samiksha Gawade [3], Vishvesh shirgaonkar[4], Basil Jose[5], Snehal Bhogan[6], Gaurav Naik[7]

[1-6]Department of Computer Engineering, Agnel Institute of Technology and Design, Goa University, Mapusa, India
[7]Infotech Corporation of Goa, Panaji Goa

Corresponding Author: 17co60@aitdgoa.edu.in

Abstract—In order to benefit from different social schemes, the beneficiaries need to submit his/her Life Certificate in offline or online modes which are not affordable by the common man. This certificate is used as a proof to ensure that the person mentioned in the 'Life Certificate' is alive. The current prevailing methods for this are time-consuming as well as tedious for senior citizens as the have to stand in long queues. This paper presents a computationally simple and efficient enhancement technique that uses voice-verification algorithms to distinguish the beneficiaries and perform verification. The proposed method uses online mode along with feature extraction and pattern matching techniques.

Keywords—MFCC(Mel Frequency Cepstral Coefficient), DTW(Dynamic Time Warping), Speaker Identification, Spekaer recognition, IVRS(Interactive Voice Response System), Identification, Verification, DCT(Discrete Fourier Transform)

I. INTRODUCTION

Authentication is very important in today's generation. It helps to recognize the user of the system and verify that the device actually belongs to the said owner. Addition to granting access to the system, authentication also helps to secure the system from unauthorized access and hackers. A number of tools and devices are already available to perform authentication such as biometric scanners, using user credentials, multifactor authentication etc. But when it comes to actually verify the liveliness of the user, some of these existing methods impose problems.

One of the newly emerging method of user authentication is voice-based authentication. It is one of the cost-effective solution and time saving technique. This technique can be used in the process of generating the 'Life Certificate'. The term 'Life Certificate' refers to a proof that the person is alive. To take the benefits of such social welfare schemes every year user has to renew the certificates in the one of the possible ways.

Verification done through voice based eliminates the use of prevailing expensive biometric scanning devices as well as the time consuming manual procedure. The proposed project targets making the whole cycle web-based utilizing web-based framework as the primary concern in correspondence. By using the web-based frameworks along with voice biometrics for the verification, a customer will feel that the person is in effect better took care of his/her solicitation. The client may have numerous inquiries and utilizing online and cost-effective mode a brisk reaction is consistently accessible. One needs to adhere to the guidelines from the arrangement of alternatives given by the framework cautiously. The reaction is educational and proficient regarding time.

Speaker Identification process is generally implemented by means of two processes, one after the other. The first process (or) the front-end module is called feature extraction, which involves extraction of the speech parameters from the clean speech signals [1]. This phase is also called as the training phase. Following the training phase is the second process (or) the back-end module, called as the feature matching which tries to match the extracted parameters of speech samples in both test and the train phase [2].

II. RELATED WORK

Syed Akhtar Imam1, Priyanka Bansal2, Vimlesh Singh proposed Speaker Recognition using Automated Systems which provides massive security. With the change in technology speaker recognition and validation of speech gained visibility and importance in the public as audio content, speech technology, and e-commerce. With an increase in demand for audio materials, speaker recognition creates growing interest. This paper explains various steps of speaker recognition such as speaker verification, speaker identification, feature extraction technique like LPC, LPCC, RCC, MFCC & PLPC, speaker modelling and concluding remark of this technique.[1]

Tahira Mahboob1 , Memoona Khanum2 , Malik Sikandar 3 Hayat Khiyal4 , Ruqia Bibi5 presented Text dependent recognition that identifies user against a phrase and text independent recognition identifies the user irrespective of what he is saying. The success in both cases depends upon

the various speaker characteristics which differentiate the one speaker from other. This paper targets the implementation of MFCC with GMM techniques in order to identify a speaker.[2]

S. Malini1 ,R. Kousalya2, presented Speaker Identification using MFCC and DTW Technique on the Enhanced Speech Signal in a Noisy Environment. The speech signals were collected from the NOIZEUS database and the feature extraction technique used is Dynamic Time Warping (DTW). The entire project results were simulated and verified using MATLAB 2010a. The results of the proposed technique are found to be more efficient than the results produced by the existing technique which uses MFCC alone for feature extraction. [3]

1Kaustubh B. Joshi, 2Vijay V. Patil presented the paper where the signal processing techniques, MFCC and DTW are explained and discussed in detail along with a Matlab program where these techniques have been implemented. The choices made in signal processing, feature extraction and pattern matching are determined by discussions of available studies on these topics. The results indicate that it is possible to program text-dependent speaker verification systems that are functional in clean conditions with tools like Matlab [4].

III. METHODOLOGY.

The proposed method consists of a web-based application that have the features to carry on the various functions related to the users who are beneficiaries to such social welfare schemes. The beneficiaries will be able to verify themselves after their successful registration towards the scheme.

As we are working with speaker verification as the main verification purpose for the system the entire system itself consists of many sub components.

On analyzing the existing methods, we have seen that MFCC technique performs exceptional speaker identification in a noiseless environment [3]. Coming to the feature matching phase, we are using DTW for the purpose. This method, combined with MFCC feature extraction, improves single digit recognition rate from 85.3% to 99% [4]. We have considered text dependent speaker verification instead of digits. The beneficiaries are asked to prompt particular words and their records are captured for further analysis.

A speaker recognition system consists of feature extraction of the speech signal, training for the particular user and recognition in the later phase of verification.
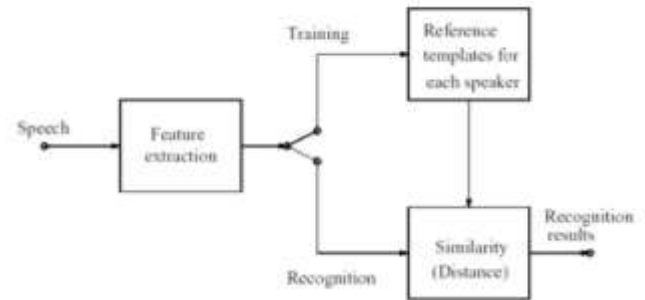

Fig 1: Structure of speaker recognition system

The very first step is to register the users who are the beneficiaries who will use the system to verify themselves. They can get themselves registered by visiting the office and providing their necessary details. An admin will access them in the process of registration. The admin here is the authenticated user who have the access to the database generally the employees of the government office. Admin can add the new beneficiary, delete existing one, update the details and also send reminders to the user based on their renewal dates. After successfully registering themselves a user gets a unique identification number.

The second step is to extract feature of the users and save them to the database. The feature extraction is done at the time of registration process. After the details of the user are stored they will be asked to prompt certain number of words. This analog voice signal is recorded by a microphone and gets converted to a digital signal through the use of an A/D-converter [3]. The features of all these words will be extracted and stored in the database. This is called the training phase.

The last phase is the verification phase in which the beneficiaries are verified based on their voice biometrics. The result of this phase is acceptance or rejection of the voice biometric features. Verification of each user is carried out by allowing the person utter the word belonging to the identity he is claiming into the microphone. The analog voice signal is then converted to a digital signal and when the features get extracted they are compared with the voice-print from the database with the help of a pattern matching algorithm. The comparison will give a match score that will be used for checking that the user is accepted or rejected.

A. Mel Frequency Cepstral Coefficients
Mel Frequency Cepstral constant features an immense accomplishment in speaker recognition system. The MFCC is best acknowledged and most extensively used for each speech and speaker recognition. Once the frequency bands are placed logarithmically in MFCC, it estimates the human system response a lot of fastidiously than the other system. The tactic of process MFCC is predicated on the short-term analysis, and therefore from every frame a MFCC vector is computed. So as to get the coefficients, the speech samples is taken because the input and playing window is applied to cut back the disruption of a signal.

Then Discrete Fourier Transform (DFT) will be used to produce the Mel filter bank. MFCC is calculated by victimization the formula [17]- Mel (f) = 2595*log10 (1+f/700) [2]. Most reliable verification techniques are text-dependent [4]. So words are considered in the proposed system for extraction.

Fig 2: Block diagram for obtaining MFC coefficients

B. Dynamic Time Warping (DTW)

This notable algorithmic principle is utilized in a few territories. It's as of now used in Speech acknowledgment, language acknowledgment and motions acknowledgment, penmanship and on-line signature coordinating, information mining and measurement bunching, observation, super particle grouping arrangement and substance designing, music and sign handling. Dynamic Time twisting calculation is projected by Sadaoki Furui in 1981.This calculation quantifies the similitude between two arrangement which can shift as expected and speed. This calculation finds an ideal match between two given arrangements. the regular of the 2 examples is taken to make a shiny new format. This interaction is consistent till all the instructing expressions have been joined into one format. this framework coordinates a check contribution from a multi-dimensional component vector T= [t1, t2… ti] with a reference layout R= [r1, r2… rj]. It finds the capacity w(i) as demonstrated inside the beneath figure. In Speaker Recognition framework, each info discourse is contrasted and the vocalization in the data. For each examination, the space measure is determined inside the estimations. Lower distance demonstrates higher comparability [3].

Fig 3: Dynamic Time Warping

Fig 4: Flowchart of Verification

IV. RESULTS AND DISCUSSION

The system implemented consists of the database designed in PostgreSQL along with the Django framework and python. The database consists of various tables as can be seen in Fig 5.

Fig 5 :Schema of the database

The authenticated admin needs to login in order to perform various operations.The basic login/logout functionalities is also given to the admin.
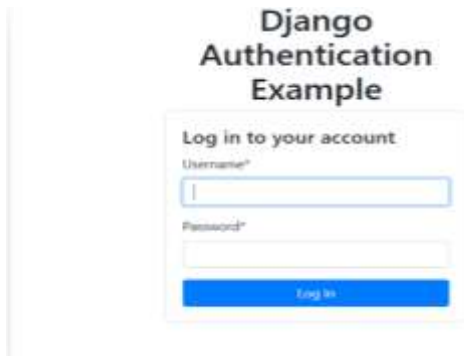
Fig 6: Login page

After the user login he can perform various functionalities.
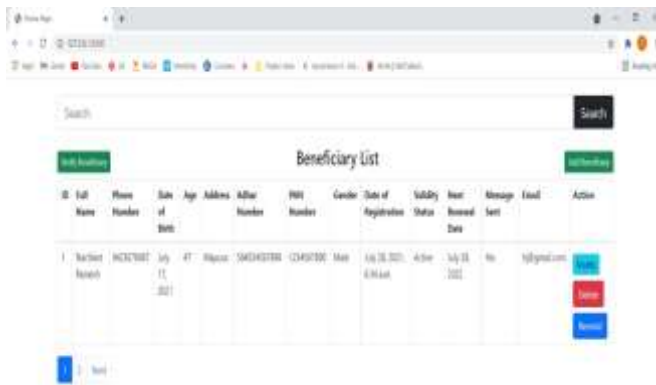

Fig 7: Beneficiary Details Page

Add Beneficiary: Add new beneficiaries to the system.
Modify: Change the existing details of the beneficiary.
Delete: Remove existing beneficiaries.
Remind: Send reminder to the beneficiaries regarding their renewal dates for the scheme.
Verify Beneficiary: Verify the beneficiaries based on their voice biometric.
Search: Search field allows admin to search for beneficiaries with certain name.
After a beneficiary is added using Add Beneficiary the features of his voice sample are extracted and stored in the database for further verification process. In extraction phase, for each word generated two voice samples are stored and threshold is set accordingly.

Fig 8: Feature Extraction Page

For the verification purpose the beneficiary needs to enter his/her identification number and record voice sample for random generated word. The voice sample is then analyzed and the result is updated.

Fig 9: User Verification page

V. CONCLUSION AND FUTURE SCOPE

Authentication systems are used almost in every areas. Authentication helps in increased security. It is very crucial that each and every individual is familiar with the technologies. Considering the aged people, they face difficulty in using the advanced technologies. Hence there is a need of more simple and familiarized authentication method. AI based online verification can help in multiple ways as discussed above.

The test was conducted on 8 people (for both male and female). It was observed that, for each user the threshold varies. If we keep a fixed threshold, common for all the users, then for two different users the pattern was found not to vary much. So, the system was showing a success message even when the intended users were different. Hence, the threshold is dependent on the user. Increasing the number of voice sample stored for each word increases the efficiency of the user getting verified accurately.

Voice sample were captured in pleasant as well as noisy environment. The degree of percentage of matching the two samples were observed to be independent of noise. However, it is recommended that the voice be captured in a cool noise free condition for better accuracy.

The entire system can be also modified to verifying the user in offline mode that is using IVRS. This proposed addition can make the entire system more cheaper and easy to use. Even a mobile phone without internet will be able to use the system.

REFERENCES

[1] *Researchgate.net.* [Online]. Available: https://www.researchgate.net/publication/329595695_Some_Commonly_Used_Speech_Feature_Extraction_Algorithms/citation/download. **[Accessed: 27-Aug-2021].**

[2] F. Bimbot *et al.*, "A tutorial on text-independent speaker verification," *EURASIP J. Adv. Signal Process.*, **vol. 2004, no. 4, 2004.**

[3] S. Malini and R. Kousalya, "Speaker Identification using MFCC and DTW Technique on the Enhanced Speech Signal in a Noisy Environment," *Int. J. Eng. Res. Technol. (Ahmedabad)*, **vol. 4, no. 14, 2018.**

[4] A. H.Mansour, G. Zen Alabdeen Salh, and K. A. Mohammed, "Voice recognition using dynamic time warping and Mel-frequency cepstral coefficients algorithms," *Int. J. Comput. Appl.*, **vol. 116, no. 2, pp. 34–41, 2015.**

AUTHORS PROFILE

VIBHUTI VELGEKAR, Final year student at Agnel Institute of Techology and Design, Goa

SAMIKSHA GAWADE, Final year student at Agnel Institute of Techology and Design,Goa

SHREYA SAWANT, Final year student at Agnel Institute of Techology and Design,Goa

VISHVESH SHIRGAONKAR, Final year student at Agnel Institute of Techology and Design,Goa

BASIL JOSE, Professor at Department of Computer Engineering, Agnel Institute of Techology and Design,Goa

SNEHAL BHOGAN, Professor at Department of Computer Engineering, Agnel Institute of Techology and Design,Goa

GAURAV NAIK, Assistant Manager (Software) Infotech Corporation of Goa, Panaji-Goa