iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.80779

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Voiceprint Authentication for File System

Vaishnavi Kshirsagar [1], Gayatri Kharat[2], Vishakha Jadhav[3], Malti Jadon[4], Prof. Dr. Ketaki Amit Malgi[5]
Department of Information Technology, Bharati Vidyapeeth's College Of Engineering For Women, Pune, Maharashtra, India

Abstract: ***The rise in online attacks has made many traditional ways of confirming someone's identity, like passwords and pins, easier for criminals to steal or guess as well as allowing them access to people's stored passwords, making these forms of identity verifications easy to compromise. The use of biological or behavioural traits associated with a person (biometric traits) can replace traditional passwords and pins as secure forms of identity verification. This research paper describes a method of accessing secure files through verifying an individual's voice and subsequently encrypting that access using a modern encryption algorithm (AES-GCM). The method of accessing secure files through this process is achieved by extracting Mel Frequency Cepstral Coefficients (MFCCs) from an individual's voice when the individual uses their voice to verify their identity. An experimental analysis was completed whereby voice samples were obtained from 20 users to evaluate the proposed secure voice-based authentication system. The experimental results indicate that the proposed secure voice-based authentication system achieved an authentication success rate of approximately 92%. The proposed system also achieved an average False Acceptance (FAR) of 4.1% and a False Rejection (FRR) rate of 3.9% which illustrates that the proposed method of using an individual's voice to access secure files will provide a reliable solution while having very little impact on the computational overhead required to verify the identity of a user. The results suggest that voice-based biometric authentication systems represent a viable method for improving data security and for this area to be researched further.***

Keywords: voice authentication, biometrics, MFCC, AES-GCM encryption, speaker recognition, file security, client-side processing

I. INTRODUCTION

As a result of rapidly changing technology, ensuring that individuals' personal data remains safe and secure from malicious actors has become the biggest hurdle to successful identity and access management (IAM) when using traditional methods (username/password) and tokens (PINs found on ATM cards). Biometric authentication (fingerprints, voice, etc.) are viewed as a more secure method than traditional password-based methods; voice-based systems are especially effective because it is virtually impossible to replicate a human voice, it is natural to use, and every person has a unique voice. Numerous studies have compared biometric methods of authentication, like voice recognition, with traditional password-based systems to assess their viability for secure user authentication [1]. The "Voice-Based Authentication System for File Sharing" (VBA-System) introduces a type of security mechanism that eliminates the use of passwords by validating user identity based on the sound of their voice. The system employs Mel Frequency Cepstral Coefficients (MFCCs), spectrograms, machine learning (e.g., TensorFlow using Keras and Librosa) and AES encryption for secure access to files. The VBA-System solutions utilize both artificial intelligence (AI) and authentication "anti-spoofing" mechanisms to prevent the exploitation of an individual's voice for malicious purposes, e.g. replayed recordings and/or artificially-generated voices. The cost of implementing a microphone and laptop (or other computer) needed to deploy VBA-System in a secure manner across multiple industries including healthcare, financial, educational, and government sectors makes this solution attractive, user-friendly and a cost-saver when accounting for the many inherent weaknesses common to previous or existing/authentication technologies.

The principal contributions of this research include the following:

1) To develop a secure biometric authentication system based on voice print technology that will provide secure access to files.
2) Combining MFCC feature extraction and AES-GCM encryption for file security and increased security against hacking.
3) Provide full client-side architectures to prevent the loss of personal biometric data from the Internet.
4) Evaluate the performance of the system under different environmental conditions.

II. RELATED WORK

Voice biometrics and speaker verification methods have been studied in numerous evaluations of secure authentication systems [1], [4]. Use of individual vocal characteristics for authenticating the user in a secure environment, without requiring the user to provide a physical object, such as a token or password has brought a notable surge of interest to voice based authenticating methods. Kinnunen and Li [5] provide a detailed overview of speaker recognition techniques, describing the effectiveness of feature

extraction methods (e.g., Mel Frequency Cepstral Coefficients (MFCC) and Spectrogram analysis) for establishing individual speaker differentiation. The techniques described by Kinnunen and Li are widely utilized in contemporary voice biometric systems due to their ability to extract essential acoustical properties of speech signals.

Recent research has begun employing Deep Learning Models (DLMs), such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), for enhancing the performance of speaker recognition systems [3], [7], particularly in difficult listening environments. These studies have also highlighted the importance of including anti-spoofing methods to protect against replay attacks and synthetic voice attacks in all biometric authentication systems [6].

In addition, methods for enhancing the performance of voice biometric systems, including analysis of signal variation and liveness detection, have been proposed as potential solutions to improve the effectiveness of these systems.

There are numerous studies that have looked at enhancing data security by combining the use of biometric authentication and cryptography. By combining biometric verification with encryption algorithms, such as AES, you can help secure sensitive files even if someone tries to get to them without permission. Many of the existing systems use server-side processing, which creates privacy and storage concerns for biometric data. The Voiceprint Authentication for File System proposal integrates MFCC based voice recognition and AES-GCM encryption in a client-side architecture to provide secure and user friendly file protection while respecting individual privacy.

A. Novelty of the Proposed work

This novel improvement in voice based biometric authentication lies in the incorporation of digital encryption of AES-GCM into a completely client-side implementation, whereas existing approaches do not use the server to process the biometric data. By eliminating the storage of biometrics on third party servers this implementation provides greater privacy for users. The implementation is able to provide real-time and efficient authentication by utilizing lightweight, web based APIs to perform the authentication process without the need for any specific hardware.

III. SYSTEM ARCHITECTURE

This novel improvement in voice based biometric authentication lies in the incorporation of digital encryption of AES-GCM into a completely client-side implementation, whereas existing approaches do not use the server to process the biometric data. By eliminating the storage of biometrics on third party servers this implementation provides greater privacy for users. The implementation is able to provide real-time and efficient authentication by utilizing lightweight, web based APIs to perform the authentication process without the need for any specific hardware.

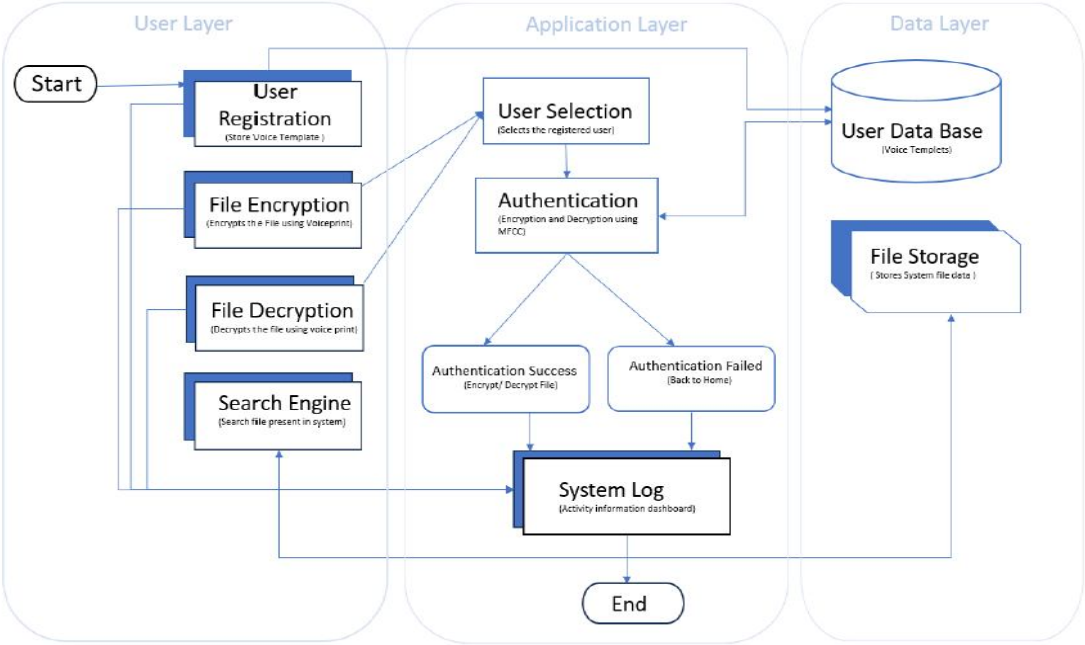


Fig. 1. System Architecture of the Voiceprint Authentication System

The architecture consists of five layers:

1) UI Layer - The use of React components for registering and logging in users, as well as for both encrypting and decrypting files.
2) Application State Layer - Manages all user-related information and provides coordination between the various components.
3) Core Logic Modules - The biometric engine will generate voiceprints, while the cryptography engine will encrypt and decrypt files.
4) Browser API Layer - Allows local biometric processing while encrypting data without using any back-end services; only front-end is used for security reasons.
5) External Services Layer - Uses the Google Gemini API for providing AI-driven explanations about how to better understand the content being created.

A. *Authentication Process*

The proposed System's authentication consists of seven sequential processes that ensure the privacy and verification of users through authentication:

1) Voice Input Acquisition - To create a voice via the Media Devices API, the user uses the microphone of their device will be used to make an audio recording. This audio is captured in real-time for later usage.
2) Audio Preprocessing - The audio that was recorded will be processed by applying noise-reduction techniques to lessen unwanted noise in the audio so that the extracted features will be as accurate as possible.
3) Feature Extraction - Mel Frequency Cepstral Coefficients (MFCC) will be extracted from the audio signal using the Web Audio API, which represent the features of the user's voice that would only exist because they heard that voice.
4) Creation of Voiceprint - The MFCC features that were previously extracted will be combined together to create the voice of the user, which is the user's biometric identification template.
5) Voice Matching - At the time when the user is attempting to authenticate, their created voiceprint will be compared to the stored version of their voiceprint via cosine similarity when checking if they are the same person.
6) Access Decision - Once the cosine similarity score has been calculated, if it is greater than the predetermined threshold, the user has been successfully authenticated; otherwise, the authentication will fail, and access will be denied.
7) Secure File Access - When the user has been authenticated and has passed all possible tests, the encrypted and decrypted files will be done with AES-GCM encryption using a key generated via PBKDF2 (Password-Based Key Derivation Function 2)

IV. IMPLEMENTATION AND RESULTS

Voiceprint Authentication for the file system has been built with an entirely client-side software architecture to preserve user privacy. All biometric and cryptography functions occur on the user's browser rather than on a third-party server outside of your control. The implementation is done via React, TypeScript and Tailwind CSS for an interactive and responsive user interface. Voice recording and feature extraction are done through the Media Devices API and the Web Audio API. The voice recording is processed to obtain Mel Frequency Cepstral Coefficients (MFCC), which can then be used to create a unique voiceprint representation of the user. The file will be encrypted securely using AES-GCM, with key creation being accomplished through PBKDF2 via the Web Crypto API.

A. *Experimental Setup*

The evaluation of the performance of the proposed system was accomplished through the collection of voice samples from 20 individual users during the enrolment process. Each user contributed five different voice recordings for use in generating their voiceprint templates. Three separate authentication tests were carried out on each user to determine the effectiveness of the proposed approach by testing both genuine user authentication attempts as well as simulated unauthorized access attempts. To evaluate the reliability and robustness of the proposed method, experiments were conducted under a variety of different environmental conditions, such as quiet, indoor areas and areas with moderate background noises (e.g. coffee shops). It is important to note that due to the relatively small number of users in this study, the resulting product of this work has limited potential for generalization. Future studies should utilize a larger number of users to replicate these findings and assess generalizability, thereby increasing their validity.

B. *Performance Evaluation*

An analysis of the functioning of the proposed system was undertaken. The parameters used to assess the evaluation included: authentication accuracy; response time; and reliability (within acceptable limits).

The overall accuracy of authentication for the proposed system was found to be approximately 92%, while the False Acceptance Rate (FAR) and the False Rejection Rate (FRR) were found to be 4.1% and 3.9%, respectively. Although these amply demonstrated FAR levels compared to high security biometric metrics from literature, they were consistent with the literature for browser-based MFCC systems using general purpose hardware and also had good performance on a comparative basis to similar client-side implementations of biometric technologies. In terms of the average time needed to complete the authentication process, the proposed system was anticipated to take approximately 2-3 seconds, thus providing a very good user experience. The average processing time for both the encryption and decryption operations using AES-GCM was approximately 1 second. See Table 1 for a summary of these results.

TABLE I
PERFORMANCE EVALUATION OF THE PROPOSED SYSTEM

Metric	Result
Authentication Accuracy	92%
False Acceptance Rate (FAR)	4.1%
False Rejection Rate (FRR)	3.9%
Average Authentication Time	2.3 sec
Encryption/Decryption Time	1.1 sec

C. *Discussion*

The findings from the research study suggest that voice authentication is an effective replacement solution to the standard passwords used for authenticating users. The proposed approach combines the security of biometric verification with the security of AES-GCM encryption, resulting in higher levels of security than traditional password authentication methods. The ability to execute voice authentication using client devices provides additional privacy when compared to current methods which use server-side storage of biometric identifiers. Limitations of the research include less accurate authentication on noisy backgrounds, changes in the sound of a user's voice due to illness and emotional state. Future implementations will incorporate noise-resilient models and adaptive voice profiling in order to assist in resolving these limitations.

V. APPLICATIONS AND LIMITATIONS

A. *Applications*

1) File Access Protection: A means of protecting sensitive files (e.g. Fleet records, Fleet budgets) from unauthorized access using voice recognition technology.
2) Healthcare Systems: Voice recognition technology can be used to protect patients' medical records from unauthorized access using fingerprint or other biometric access control methods.
3) Finance and Banking: Banks will use voice recognition technology to secure sensitive customer/financial information (e.g. Bank account information) as well as provide secure verifications for customers of their financial institutions.
4) Educational institutions & Research: Educational institutions will use voice recognition technology to secure student records, research results and examination documents, without the use of hardware tokens.
5) Government and Corporate: Voice recognition technology will be used to enhance the security of classified documents and classified corporate information.

B. *Limitations*

1) Noise Sensitivity: If there is noise in the environment or there are multiple competing voices, the authentication accuracy will decrease.
2) Voice Variability: Anytime the user is stressed, ill, or emotional, their voice can change and this will impact how well the system can recognize them.
3) Spoofing Challenges: Anti-spoofing protection has been applied, but highly sophisticated AI generated voice clones are still a serious potential risk to the system.
4) Device Dependency: The quality of the microphone and recording hardware impacts the reliability of voice features extracted for authentication.
5) Limited Dataset: As processing is client-side, the system will not be able to use any scalable cloud-based datasets for model improvement and the 20-user evaluation will limit the ability of the system to generalize.

VI. FUTURE WORK

Combining several kinds of biometric methods (like using fingerprints and/or facial recognition) helps increase accuracy when authenticating someone or even creating a more secure method of authentication through the use of multiple biometrics.

Deep Learning techniques provide a way for more advanced anti-spoofing methods to help identify signs of an artificial or replayed voice.

Noise-resilient models provide improved performance in a variety of environments (e.g., outdoors and/or noise) through the use of advanced signal processing and machine learning techniques.

Cloud-based secure access to your system can enhance the functionality of your systems with access to encrypted files stored securely in the cloud up to the point where you are able to access these secure files from any device you own.

User behaviour learning is a technique that uses an adaptive model which tracks voice characteristics over time to help reduce the risk of users being incorrectly rejected due to events such as illness or changes in mood.

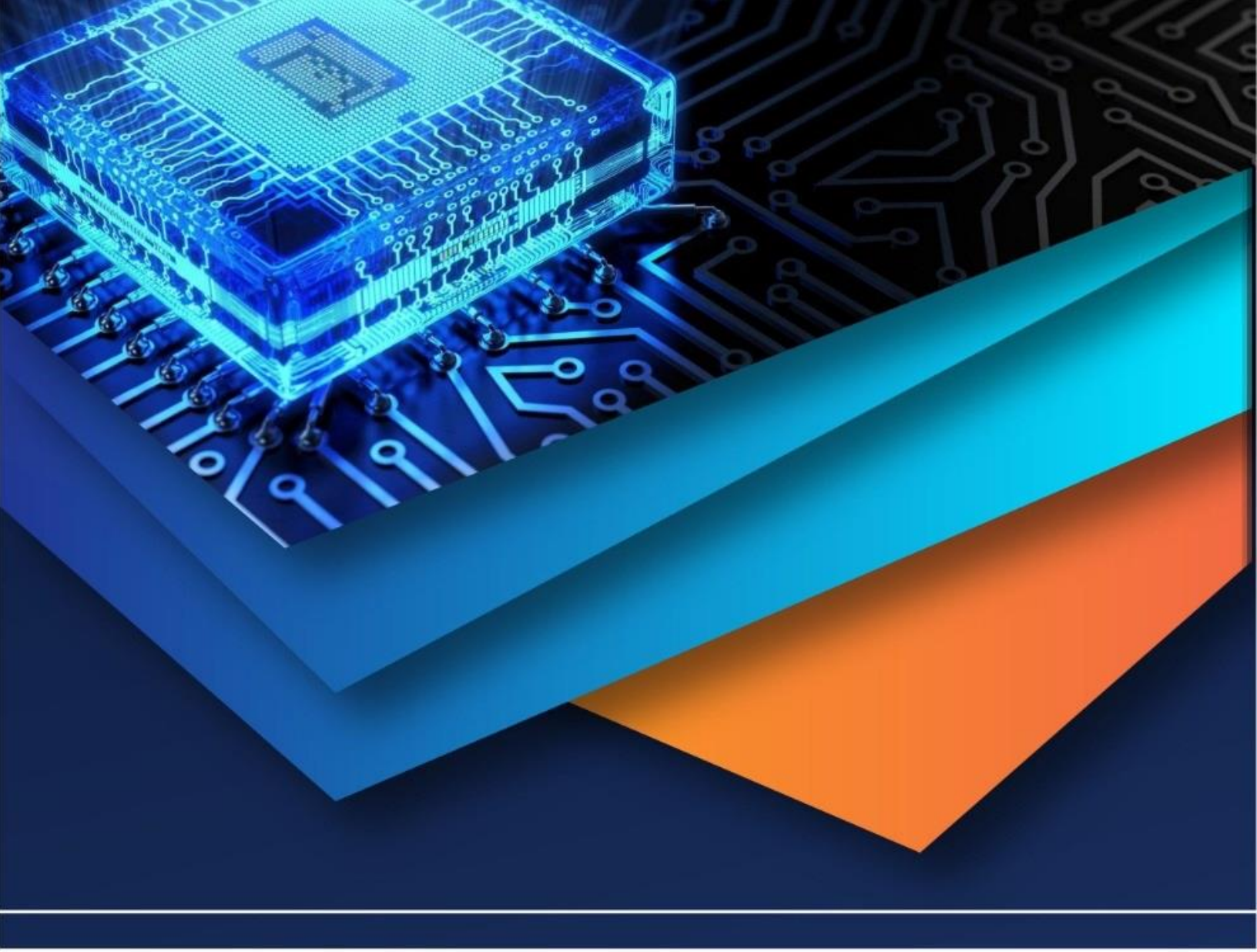
VII. CONCLUSIONS

The Voiceprint Authentication for File System exemplifies a new, user-friendly, safe, and modern method of securing files through the combination of voice biometrics and AES-GCM encryption. There is no third-party server processing within the system because all processes occur in the client's browser. According to experimental data, using voiceprints rather than passwords provides an approximately 92% accurate authentication, suggesting that voiceprints can serve as viable substitutes for traditional passwords while providing a more natural and intuitive experience. There are still challenges associated with this method, such as noise, variability of

the user's voice, and the risk of being spoofed; however, it lays the groundwork for future improvements. Ultimately, this system offers a strong, expandable, and future-proof solution for protecting sensitive electronic data in numerous real-world settings.

REFERENCES

[1] D. O'Shaughnessy, "Review of methods for automatic speaker verification," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 32, pp. 1776–1789, 2023.
[2] R. Sharma, D. Govind, J. Mishra, A. Dubey, K. Deepak, and S. Prasanna, "Milestones in speaker recognition," Artificial Intelligence Review, vol. 57, no. 3, p. 58, 2024.
[3] M. Jakubec, R. Jarina, E. Lieskovska, and P. Kasak, "Deep speaker embeddings for speaker verification: Review and experimental comparison," Engineering Applications of Artificial Intelligence, vol. 127, p. 107232, 2024.
[4] R. M. Hanifa, K. Isa, and S. Mohamad, "A review on speaker recognition: Technology and challenges," Computers & Electrical Engineering, vol. 90, p. 107005, 2021.
[5] T. Kinnunen and H. Li, "An overview of text-independent speaker recognition: From features to supervectors," Speech Communication, vol. 52, no. 1, pp. 12–40, 2010.
[6] H. Delgado et al., "ASV spoof 2021: Automatic speaker verification spoofing and countermeasures challenge evaluation plan," arXiv:2109.00535, 2021.
[7] S. Wang, Z. Wu, and H. Meng, "Deep learning approaches for speaker verification: A survey," IEEE Access, 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY