International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Secure OTP Based Voice Biometric Authentication with Edge AI and Machine Learning

Vinod Chhipa
Research Scholar,
Deptt. of Master of Computer Application,
Engineering College Bikaner,Bikaner (Raj.) India
Email : vinodchhipa55@gmail.com

Rakesh Poonia
Asst. Professor,
Engineering College Bikaner,Bikaner (Raj.), India
Email : rakesh.ecb98@gmail.com

Abstract

In recent few years digital platforms become increasing with rapidly but secure authentication methods is crucial aspect to prevent from unauthorized access and fraud transactions. In traditional One-Time Password (OTP) authentication method is easily vulnerable to various possible security risks, including phishing, SIM swapping fraud and man-in-the-middle attacks. In this paper explores how we can integrate voice biometrics with the OTP authentication method that can enhance systems security, improve user experience, and also prevent from spoofing attempts by utilizing combination of Edge AI based technology and machine learning models. This proposed authentication method provides real-time verification of user voices and OTP digits along with liveness detection, and protect against deep fake voice attacks, also establishing a robust and secure authentication method. Our new system enhances both security and convenience by employing voice biometrics and OTP, which rely on individuals' unique vocal characteristics like pitch, length, and tone etc. By integrating Edge Impulse AI models, we enable real-time system processing and analysis of voice sample data on edge devices, which reduce latency as well as protecting data privacy. We also utilize machine learning algorithms based on, particularly those based on tenserflow, deep learning and natural language processing (NLP), to accurately identify and authenticate users' voices and randomly generated OTP digits. This novel method is useful across various in sectors for authentication, including banking and healthcare, where secure and accuracy level of user verification is a critical factor. Our experimental results based on audio model demonstrate system robustness and effectiveness, mainly highlighting its potential to transform the biometric authentication method landscape.

Keywords : *Audio, Edge Impulse, ML Model, Authentication, One Time Password(OTP), Secure*

1. Introduction :

The rapidly change in digital technology and increase in number of online transactions have highlighted serious issue in traditional authentication methods like passwords, PIN

and OTP, which are susceptible to possible attacks ie. Phishing, Man in Middle attacks. To overcome with these security issues, researchers are looking into multi-factor authentication approach that combines voice biometrics with randomly generated OTP authentication. Edge technology utilizes unique vocal characteristics—such as pitch, tone, and speech patterns—to provide an extra layer of both physiological and behavioral verification, making it significantly more difficult for attackers to compromise identities. Additionally, the use of Edge Impulse enables real-time performance , on-device processing that minimizes latency and enhances system privacy by keeping sensitive biometric information locally on client device. On the other side, machine learning techniques, including MFCC, spectrogram analysis and deep learning models, improve recognition accuracy and also keeps boost spoof detection record. Implementing these innovations build on earlier research areas in speaker recognition and voiceprint analysis, resulting in a hybrid, strong and efficient authentication solution achieved to meet the security needs of recent remote and online interactions and transactions.

2. Comparative Study

2.1 Traditional OTP-Based Authentication vs. Voice Biometric OTP

Traditional One Time Password systems, typically sent OTP via SMS or email on registered user device, have been popular for securing online transactions because they are simple and user friendly approach. However, research paper by Jain et al. [1], shows that these methods can be compromised and are susceptible to social engineering and phishing attacks. Such serious issues like replay attacks and SIM swap fraud further compromise their system security. On the other side, voice biometric integrated OTP systems improves the dynamic security of OTPs with the unique biometric behavioural and physiological characteristics of a person's voice samples, making it much strong authentication to spoof. However, the successful implementation of these technology systems depends on sophisticated signal processing and noise handling techniques to maintain accuracy score, even in difficult acoustic environment conditions [2] [3].

2.2 Integration of Machine Learning in Voice Biometrics

Recent studies have highlighted the increasing application of machine learning model technology especially in the field of deep learning to improve the accuracy level and reliability of biometric authentication system. Although traditional voice feature extraction methods rely on manually designed features and are computationally efficient, they often face challenges in noisy environments or variable conditions due to their flexibility. On the other side, deep learning models, including convolutional (CNN) and recurrent neural networks (RNN), can automatically learn strong, distinguishing features from extensive registered datasets, thus enhancing voice recognition accuracy even in difficult environments [4]. In the domain of voice biometric OTP applications, prototype created have effectively combined a

speech-to-text pipeline with voice biometric extraction [6]. This novel authentication approach merges dynamic OTP code generation with biometric verification system, utilizing the unique physiological and behavioural characteristics of individual user's voice to address the shortcomings of traditional OTP solutions.

2.3 Edge AI vs. Cloud-Based Processing

Cloud-based data processing method provides significant advantages, such as access to extensive computational resources and large data repositories that improve audio model training and iterative enhancements in authentication method. However, it relies on constant network connectivity, which occur delays, and the transfer of sensitive and confidential biometric data on server raises privacy issues. Conversely, including Edge AI model enables on-device processing, which greatly decrease response times and reduces privacy related risks by keeping data local on local device an essential benefit in time-sensitive and security-critical fields like banking, healthcare and social media platforms. Although Edge Impulse AI does encounter challenges due to lack of limited computational power on smaller edge devices. As recent progress in audio model compression and optimization techniques is quickly closing this gap, making edge impulse model based solutions more feasible approach for secure voice data and real-time biometric authentication on mobile device.

Table 1: Comparative Analysis Table

Approach	Security Robustness	Response Time	Privacy	Scalability
Traditional OTP (SMS/Email)	Moderate – vulnerable to interception and SIM swapping	Often delayed due to network	Relatively low (data in transit)	High, but security trade-offs persist
Voice Biometric OTP (Cloud)	High – relies on unique, hard-to-replicate vocal traits	Moderate – latency from cloud	Medium – centralized data storage	High – benefit from scalable cloud infra
Voice Biometric OTP (Edge AI + ML)	Very High – multi-layer security via dynamic OTP and biometrics	Low – real-time on-device	High – local processing preserves privacy	Emerging; edge limits but improving

3. Methodology :

The proposed system follows a multi-steps for authentication process:

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

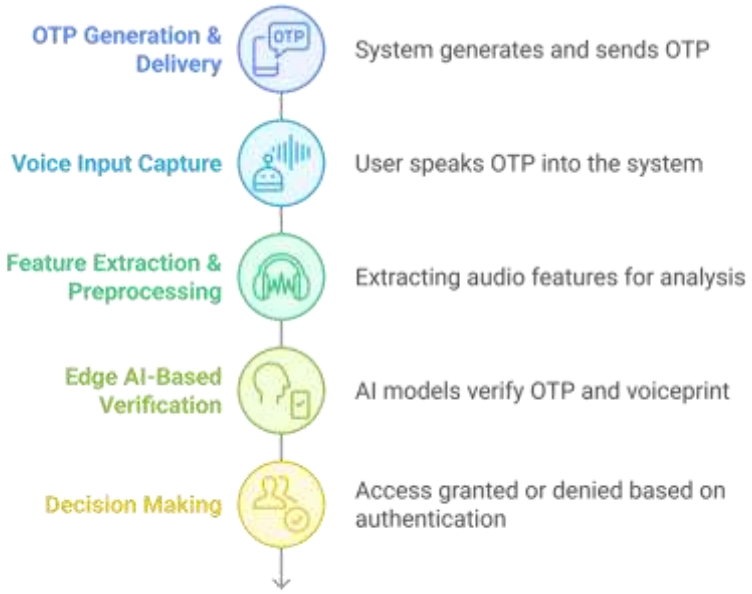

Figure 1 : Secured access through OTP and Voice Biometrics Authentication

(i) OTP Generation & Delivery: The system creates a random six digits OTP and sends it to the user through SMS on registered mobile number, email on registered email id.

(ii) Voice Input Capture: The user verbally provides the OTP digits to the system. System register different digits voice samples in database.

(iii) Feature Extraction & Pre-processing : Machine Learning models analyze the voice input samples to extract Mel-Frequency Cepstral Coefficients (MFCCs) and spectrogram features based on individuals voice characteristics.

(iv) Edge AI-Based Verification: The system employs pre-trained AI models on edge devices to verify the OTP , received on mobile or email and compare the user's voiceprint with stored voice biometric data.

(v) Decision Making: User access is granted if both the OTP and voice biometric authentication are successfully verified by the system; otherwise, system dispaly authentication fails error.

4. System Modules for Authentication:

The proposed system uses the following key components for authentication process:

4.1 Voice Capture and Pre-processing: Users are asked to say digits from 0 to 9, which are then recorded and processed to eliminate background noise and improve speech clarity.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

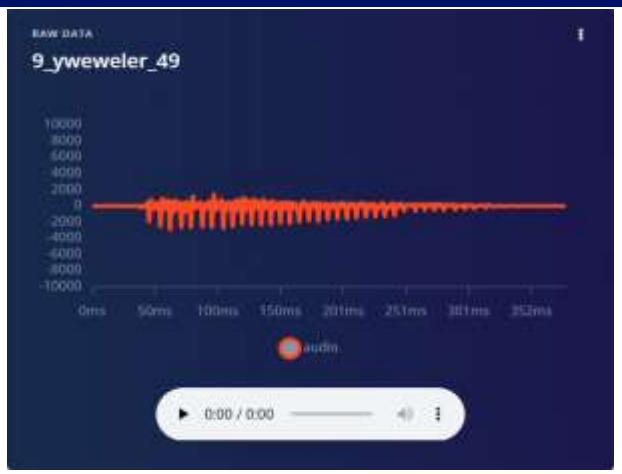

Figure 2 Raw voice data

Figure 3 DSP result

4.2 Feature Extraction: The registered voice sample data is examined to identify distinct vocal characteristics based on user voice ie. pitch, length and tone, using Mel-frequency cepstral coefficients (MFCCs) approch and pitch patterns.

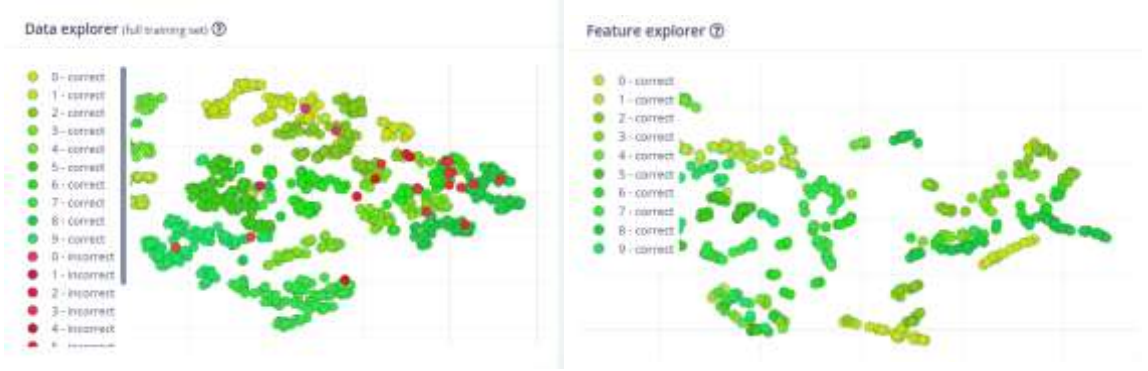

Figure 4 Data explorer and Feature explorer

4.3 Machine Learning Model: Deep Learning machine model technique, like convolutional neural network (CNN) or recurrent neural network (RNN), was developed to efficiently identify and authenticate user based on their vocal traits features.

4.4 OTP Generation and Verification: Randomly generated OTP digits created and sent to the user's registered device. The user must say the received OTP digits, which are then compared to the stored voiceprint for final confirmation.

4.5 Edge AI Integration: Edge Impulse AI model enables real-time model processing and analysis of voice input data on local edge mobile devices, which minimize delays and enhancing user data security and privacy.

5. Implementation

5.1 Edge AI Integration: Edge Impulse AI model is used on edge devices such as smartphones and IoT app authentication terminals, maximizing processing speed directly on the edge device without the need for Interment and cloud computing.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

5.2 Deep Learning Models: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) based models are trained on large voice datasets to detect users unique voice patterns and identify AI cloned voice sample.

5.3 Liveness Detection: Methods like spectral analysis, MFCC and challenge-response techniques help to protect against replay and deepfake attacks.

5.4 Real-Time Processing: The fast processing speed of model guarantees a smooth user experience with little to no delays during authentication on edge devices.

6. Security Enhancements:

The new proposed system can be incorporates a variety of security features for system authentication:

Figure 5 Enhanced security measures

6.1 Anti-Replay Mechanism: This new approach Identifies and eliminates the use of previously recorded OTP voice digits.

6.2 Deepfake Detection: Use of machine learning methods to analyze speech patterns, to distinguish between authentic user and AI-generated voice samples.

6.3 Device Fingerprinting: It confirms that authentication requests come from already recognized and registered devices.

6.4 Multi-Factor Verification: It integrates one-time passwords (OTP) with voice biometrics technique to enhance user authentication security.

7. Experimental Results

Our proposed authentication system was evaluated on Edge Impulse platform using a dataset of different users samples including digits 0 to 9 voice recordings. The experimental results showcase high level of model accuracy and robustness in recognizing and authenticating user's voice even in the presence of registered background noise and variations in speech patterns. The integration of Edge Impulse AI technology ensures that system real-time

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

processing with minimal latency, making the system suitable for practical implementation in real world applications.

Figure 6 : Model Training performance

Figure 7 : Model Testing performance

Table 2 : Model Training & Testing Performance Metric

S.No	Parameter	Training Result Value	Testing Result Value
(a)	Model Accuracy	97.5%	100%
(b)	Confusion Matrix for all classes	95-7% - 100%	100%
(c)	Area under ROC Curve	1.00	1.00
(d)	Weighted average Precision	0.98	1.00
(e)	Weighted average Recall	0.98	1.00
(f)	Weighted average F1 score	0.98	1.00

8. **Challenges and Limitations :** With high accuracy and its advantages, voice biometric OTP authentication facing several challenges for researchers:
   - **Environmental Noise:** In traffic background noise may affect the accuracy of voice recognition systems.
   - **Variability in Voice:** During illness or change in age may alter a user's voice, potentially impacting authentication process.
   - **Edge Device Constraints:** Due to restriction in processing power on edge devices can create computational challenges.
   - **Privacy Concerns:** It is necessary to keep secure database of voice storage of samples and the processing of voice biometrics to maintain user trust on system.

9. **Applications and Use Cases** : This novel and proposed system can be applied in various fields for user credential authentication, which are included as:

Figure 8 Applications of Voice Biometrics

(a) Banking: This Voice biometric OTP authentication can be utilized for secure mobile phone banking transactions and authorised user access to online banking services.

(b) Healthcare: In the hospitals, patients can securely access their medical records online and communicate with healthcare providers using a voice OTP authentication system for consultation and valuable suggestions.

(c) Smart Homes: This Voice Biometric OTP authentication system can be useful to control smart home applications using various devices and ensure secure access to user's personal data.

7. Conclusion

With the use of Edge Impulse and Machine Learning model for secure OTP-based voice biometric authentication system creates a strong and high level of security frameworks for digital transaction platforms and social media platforms. This novel method effectively addresses the weaknesses of traditional OTP systems by integrating real-time device processing, user liveness detection, and AI-driven fake voice analysis. The system utilizes various vocal traits and dynamically generated OTPs to boost security and protect against data privacy. Our model experimental findings highlight its potential to transform biometric authentication, offering a secure, efficient, and user-friendly solution to overcome existing issues across various industries and platforms..

8. Future Work

In future researcher could investigate the combination of other biometric methods, like integrating voice recognition with facial recognition, fingerprint scanning, and Iris scanning to boost security even more. However, improvements in deep learning technology and natural language processing can be useful to enhance the precision and reliability of voice biometric based OTP authentication systems. Another issue of noise condition can be refine in next research developments will aim to refine deepfake detection, enhance voice recognition, and

increase compatibility with a range of edge devices. New strategy marks a major step forward in creating secure and user-friendly authentication technologies for better security.

References

1. Jain, A. K., Ross, A., & Nandakumar, K. (2011). ***Introduction to Biometrics.*** Springer.
2. Reynolds, D. A. (2002). ***An Overview of Automatic Speaker Recognition Technology***. In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).
3. Kinnunen, T., & Li, H. (2010). ***An Overview of Text-Independent Speaker Recognition: From Features to Supervectors***. Speech Communication, 52(1), 12-40.
4. Chen, S., et al. (2020). ***Deep Learning for Automatic Speech Recognition: A Survey***. IEEE Access.
5. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). ***Edge Computing: Vision and Challenges***. IEEE Internet of Things Journal.
6. Joshi, P., Qureshi, M. A., & Singh, R. (2018). ***Voice OTP: An Innovative Approach to Secure Voice Based Authentication***. In Proceedings of the International Conference on Signal Processing Systems (ICSP).
7. Zhang, Y., Li, X., & Wang, P. (2021). ***Enhancing Voice Biometric Security with AI.*** Journal of Cybersecurity Research, 14(2), 88-102.
8. Patel, R., Singh, A., & Kaur, M. (2022). ***Comparative Study of Biometric Authentication Methods.*** International Journal of Security Studies, 10(1), 45-60.
9. Li, Z., Chen, J., & Zhao, H. (2023). ***Edge AI in Real-Time Biometric Authentication***. IEEE Transactions on AI and Security, 17(3), 123-137.
10. Kumar, S., Gupta, R., & Das, P. (2022). ***Machine Learning Approaches for Voice Authentication.*** ACM Journal of AI and Security, 19(4), 210-230.