Graduate School of Science and Technology Master's Thesis Abstract

Laboratory name (Supervisor)	Cyber Resilience (Youki Kadobayashi (Professor))		
Student ID	2311421	Submission date	2025 / 7 / 23
Name	OYEBODE OLUWATOBI OYEWALE		
Thesis title	Adapting a Cell Phone with Voice Biometric Verification		

Abstract

Secure and reliable user authentication remains a critical concern in mobile banking, particularly in developing regions where low-end devices, noisy environments, and inconsistent audio inputs pose significant challenges. Traditional voice biometric systems often prioritize accuracy while overlooking vulnerabilities such as spoofing, usability under variable conditions, and feasibility on constrained hardware.

This thesis presents a robust and inclusive speaker identification and verification system designed to overcome these limitations through a layered architecture. The system employs a dual-embedding approach that fuses Whisper-based contextual representations with Emphasized Channel Attention, Propagation, and Aggregation - Time Delay Neural Network (ECAPA-TDNN) derived speaker-specific features, enhancing voice discrimination. Cosine similarity and angular distance metrics are used to improve verification reliability. To defend against spoofed audio inputs, a liveness detection module is introduced early in the pipeline, utilizing fused spectral and embedding features for real-time classification. An adaptive enhancement module further ensures consistent feature quality across diverse acoustic conditions.

The proposed system achieves 99.59% speaker identification accuracy and a mean Equal Error Rate (EER) of 0.0017, outperforming existing models in both precision and security. It demonstrates generalizability across languages, achieving a 96.99% F1-score on a tonal African language (Yoruba), and is optimized for real-time deployment on low-end devices through Open Neural Network Exchange (ONNX) model compression.

These results validate the feasibility of deploying secure, accurate, and resource-efficient voice authentication systems for real-world financial applications, offering a practical foundation for biometric verification in environments where traditional methods fall short.