CLOUD-BASED VOICE RECOGNITION SECURITY SYSTEM

STUDENTS: MEGAN BUI, DEVYANSH GUPTA, KELLY HO

OBJECTIVE

- Develop a cloud-based voice recognition system that identifies BECU members as they speak with an employee at a Neighborhood Financial Center in the United States
- Real-time voice recognition from captured audio streams
- Provide a second layer of authentication (after PIN verification) that strengthen security, reduce caller fraud, and maintains operation efficiency

SYSTEM OVERVIEW & WEB DEVELOPMENT

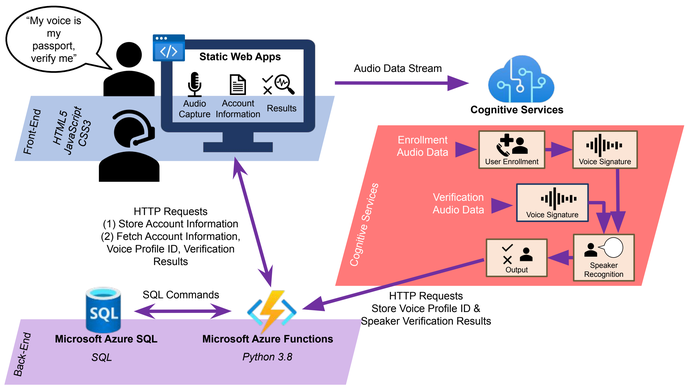


Figure 1: System Diagram

USE CASES

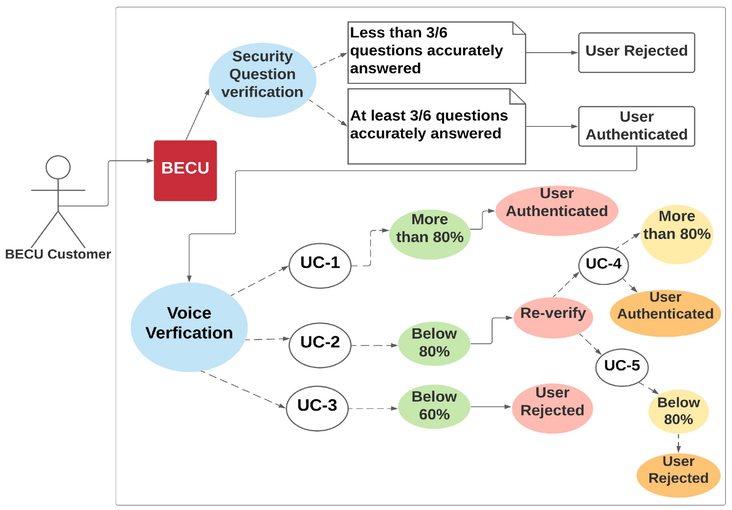


Figure 2: Use Cases

FRONT-END

- Implemented with Azure Static Web Apps [1]
- Built the website with HTML5, CSS3, and JavaScript
- Wireframe Design Software: Adobe and Lucidchart
- Audio Capture System: MediaStream API [2] for recording and streaming audio data
- Azure Cognitive Services Speaker Recognition [3][4] in JavaScript:
  - Text-Dependent
  - Pass Phrase: "My voice is my passport, verify me"
  - Enrollment: 3 Audio Recordings
  - Verification: 1 Audio Recording

ADMIN PORTAL

- Admin Login Page: Allow BECU employees to sign into admin portal
- Add New Customer Page: Register customers and enroll voice signatures
- Customer Database Page: Verify customers with security questions and voice recognition
- System Dashboard Page: Show voice recognition system performance

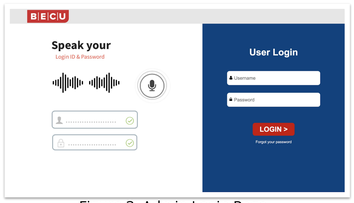
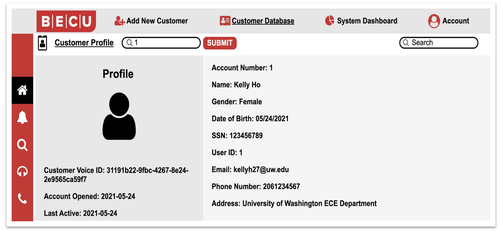


Figure 3: Admin Login Page



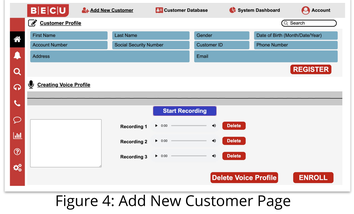
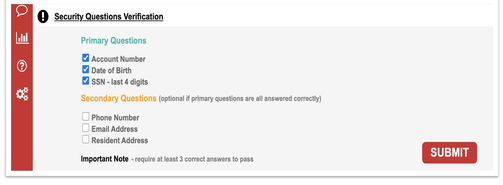


Figure 4: Add New Customer Page



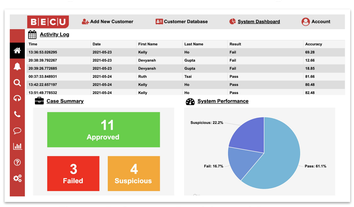


Figure 5: System Dashboard Page

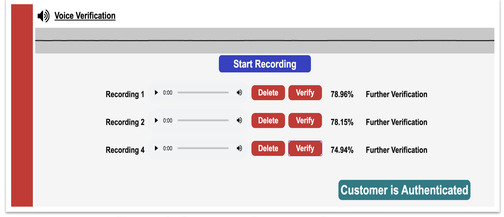


Figure 6: Customer Database Page

BACK-END: API

- Implemented with Azure Functions [5] in Python
- Builds the web API with a RESTful and serverless architecture
- Handles HTTP requests using HTTP trigger function
- Responds to database transactions with pyodbc package
- Creates and fetches customer information and voice profile ID for voice recognition
- Updates and retrieves voice recognition system performance statistics

BACK-END: DATABASE

- Implemented with Azure SQL Database [6]
- Stores customers information when they are enrolled for the first time
- Stores employee credentials and system performance metrics
- Connects with Azure Cognitive Services to store the Voice Profile ID generated to verify the customer

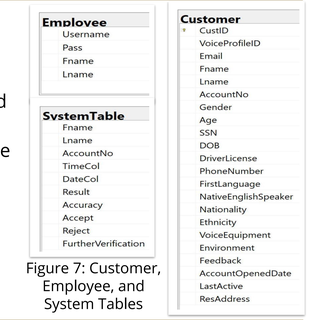


Figure 7: Customer, Employee, and System Tables

MACHINE LEARNING RESULTS

- The voice signature of 17 users were enrolled with recordings in quiet environments
- Thresholds: Reject = <70%; Further Verification = 70-80%; Accept = >80%

Summary	Quiet	Noisy	Slow	Fast
Rejected	0.0000	0.1176	0.0000	0.0000
Needed Further Verification	0.1176	0.5294	0.8750	1.0000
Accepted	0.8824	0.3529	0.1250	0.0000
Average Confidence %	0.9204	0.7847	0.7550	0.7052

Table 1 : Testing Scenarios
(1) quiet environments, (2) noisy environments, (3) users spoke slowly, & (4) users spoke quickly

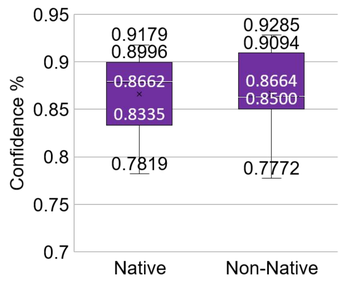


Figure 8 : Native vs. Non-Native English Speakers Tested in Quiet Environments

CONCLUSION & FUTURE WORK

Our team successfully:

- Developed a web application to achieve a cloud-based voice recognition solution
  - Authenticate new and existing members within 30 sec with >80% confidence level
- Used accessible microphones to capture live audio streams and perform real-time voice recognition
- Deployed Microsoft Azure resources for effective and secure cloud-based functionalities that can be easily integrated with BECU's existing infrastructure and flexible to scale-up

Future Goals:

- Explore Blockchain as an alternative to SQL database
- Further test the MS Speaker Recognition API against a more diverse, large dataset

REFERENCES

[1] Azure Static Web Apps: https://azure.microsoft.com/en-us/services/app-service/static/
[2] MediaStreams API: https://developer.mozilla.org/en-US/docs/Web/API/MediaStream
[3] Azure Cognitive Services: https://azure.microsoft.com/en-us/services/cognitive-services/speaker-recognition/
[4] Xiao, X., Kanda, N., Chen, Z., Zhou, T., Yoshioka, T., Chen, S., . . . Gong, Y. (2021). Microsoft Speaker Diarization System for the Voxceleb Speaker Recognition Challenge 2020. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). doi:10.1109/icassp39728.2021.9413832
[5] Azure Functions: https://docs.microsoft.com/en-us/azure/azure-functions/functions-create-serverless-api
[6] Azure SQL Database: https://azure.microsoft.com/en-us/products/azure-sql/database/

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

INDUSTRY MENTOR: JIM JENKINS, RITU GOEL, SIEKHUOY HOUT
FACULTY ADVISOR: PAYMAN ARABSHAHI, ELI SHLIZERMAN
GRADUATE STUDENTS: XIULONG LIU, KUN SU, HAOBO ZHANG

SPONSOR: BOEING EMPLOYEES' CREDIT UNION (BECU)