

Exploiting Voice Cloning in Adversarial Simulation

Mark Foudy

Boston Hacker

@0xM4rk7homas

DEFCON 32

Adversary Village

Who am I?

- Father, Husband
- AI Offensive Security Researcher
- Founder of Neurodiverse Hackers
- DEFCON 508
- Passion for pushing technological boundaries, and a commitment to improving the safety and functionality of digital systems



Voice Verification Systems

Thesis:

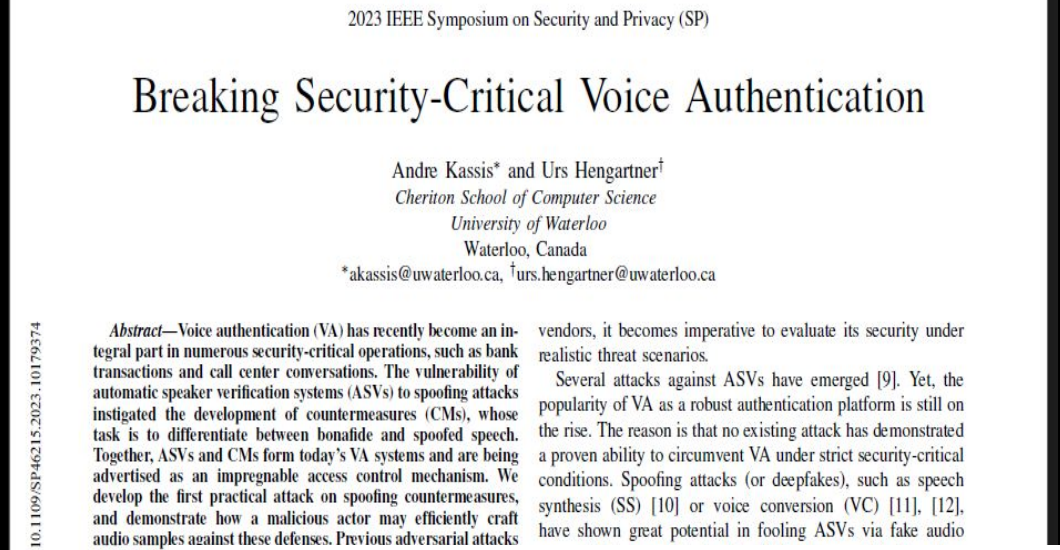
Voice Verification Systems (VVS) are susceptible to spoofing attacks with modest effort using victim's audio

Objective:

Demonstration spoofing my own cloned voice to defeat my bank's VVS

Research & Recent Headlines

Latest news and studies show that criminals increasingly use deepfake voices to bypass previously secure voice authentication systems



2023 IEEE Symposium on Security and Privacy (SP)

Breaking Security-Critical Voice Authentication

Andre Kassis* and Urs Hengartner†
Cheriton School of Computer Science
University of Waterloo
Waterloo, Canada
*akassis@uwaterloo.ca, †urs.hengartner@uwaterloo.ca

***Abstract*—Voice authentication (VA) has recently become an integral part in numerous security-critical operations, such as bank transactions and call center conversations. The vulnerability of automatic speaker verification systems (ASVs) to spoofing attacks instigated the development of countermeasures (CMs), whose task is to differentiate between bonafide and spoofed speech. Together, ASVs and CMs form today's VA systems and are being advertised as an impregnable access control mechanism. We develop the first practical attack on spoofing countermeasures, and demonstrate how a malicious actor may efficiently craft audio samples against these defenses. Previous adversarial attacks against VA**
which assu
large query
When attac
not hold. O
of failure t
real-time, n
need to inte
key messag
distinguish
that are eas
are subtle
can still by
textual cont
attack that
yielding suc
we perform
CMs, bypas
of potential
call centers.
VA systems

ivacy (SP) | 978-1-6654-9336-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/SP46215.2023.10179374

vendors, it becomes imperative to evaluate its security under realistic threat scenarios.

Several attacks against ASVs have emerged [9]. Yet, the popularity of VA as a robust authentication platform is still on the rise. The reason is that no existing attack has demonstrated a proven ability to circumvent VA under strict security-critical conditions. Spoofing attacks (or deepfakes), such as speech synthesis (SS) [10] or voice conversion (VC) [11], [12], have shown great potential in fooling ASVs via fake audio

MOTHERBOARD
TECH BY VICE

How I Broke Into a Bank Account With an AI-Generated Voice

Banks in the U.S. and Europe tout voice ID as a secure way to log into your account. I proved it's possible to trick such systems with free or cheap AI-generated voices.

By Joseph Cox

Overview of Research

EFFECTIVENESS: VOICE CLONING CAN BE A POWERFUL METHOD TO GAIN INFORMATION AND ADVANTAGES

FOCUS: DEFEATING VOICE VERIFICATION SYSTEMS

MECHANISM: UNDERSTANDING HOW VVS DETECT HUMAN SPEECH AND IDENTIFY SPOOFED AUDIO

Companies Producing VVS



AURAYA

World leaders in voice biometrics



VERINT®

Powering
Actionable Intelligence®









Financial Institutions as Key Customers

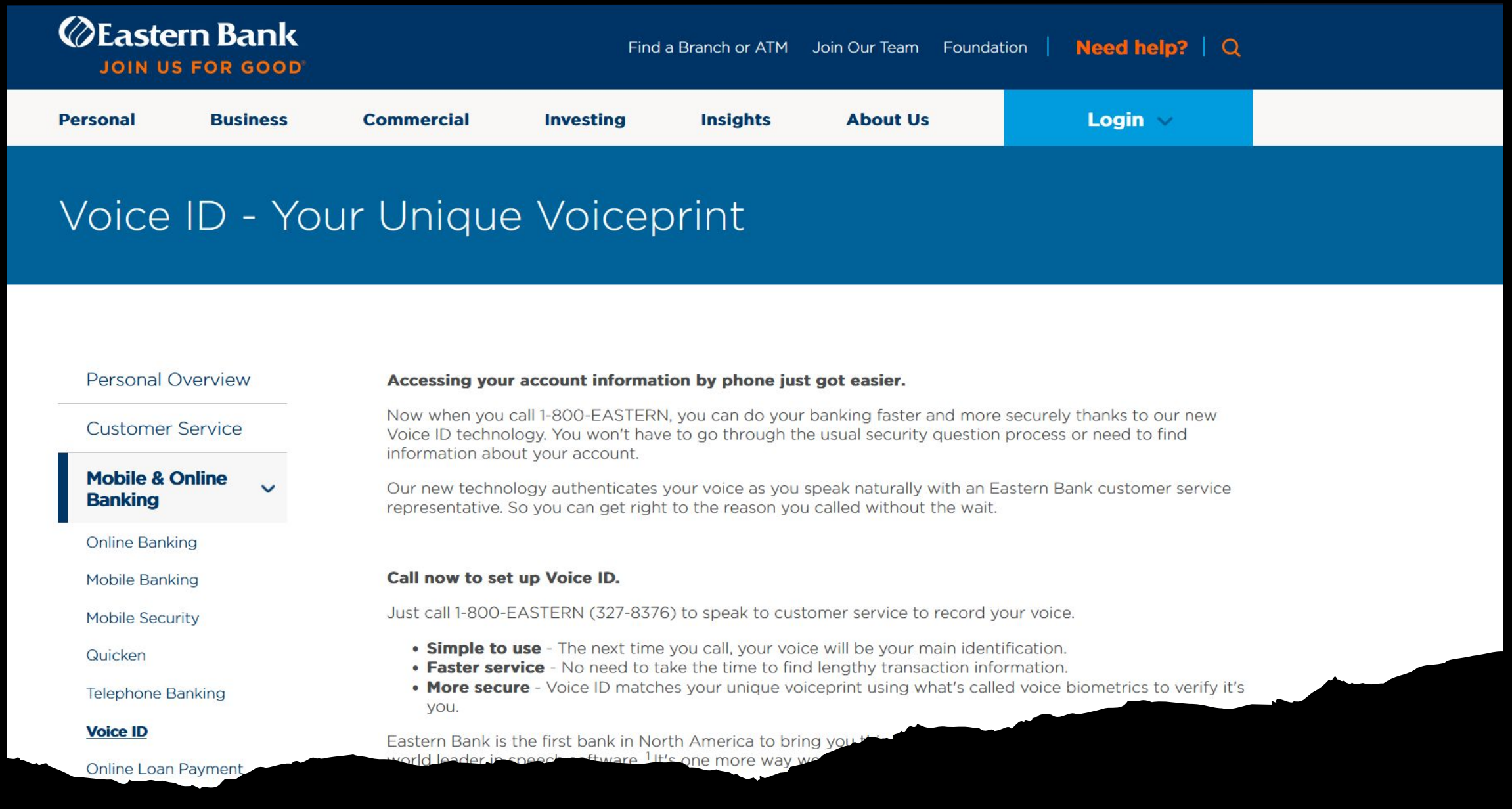
- Importance: Financial institutions rely heavily on VVS
- Security Benefits: Voice verification systems add a layer of biometric security
- Example: My own bank and my financial services provider



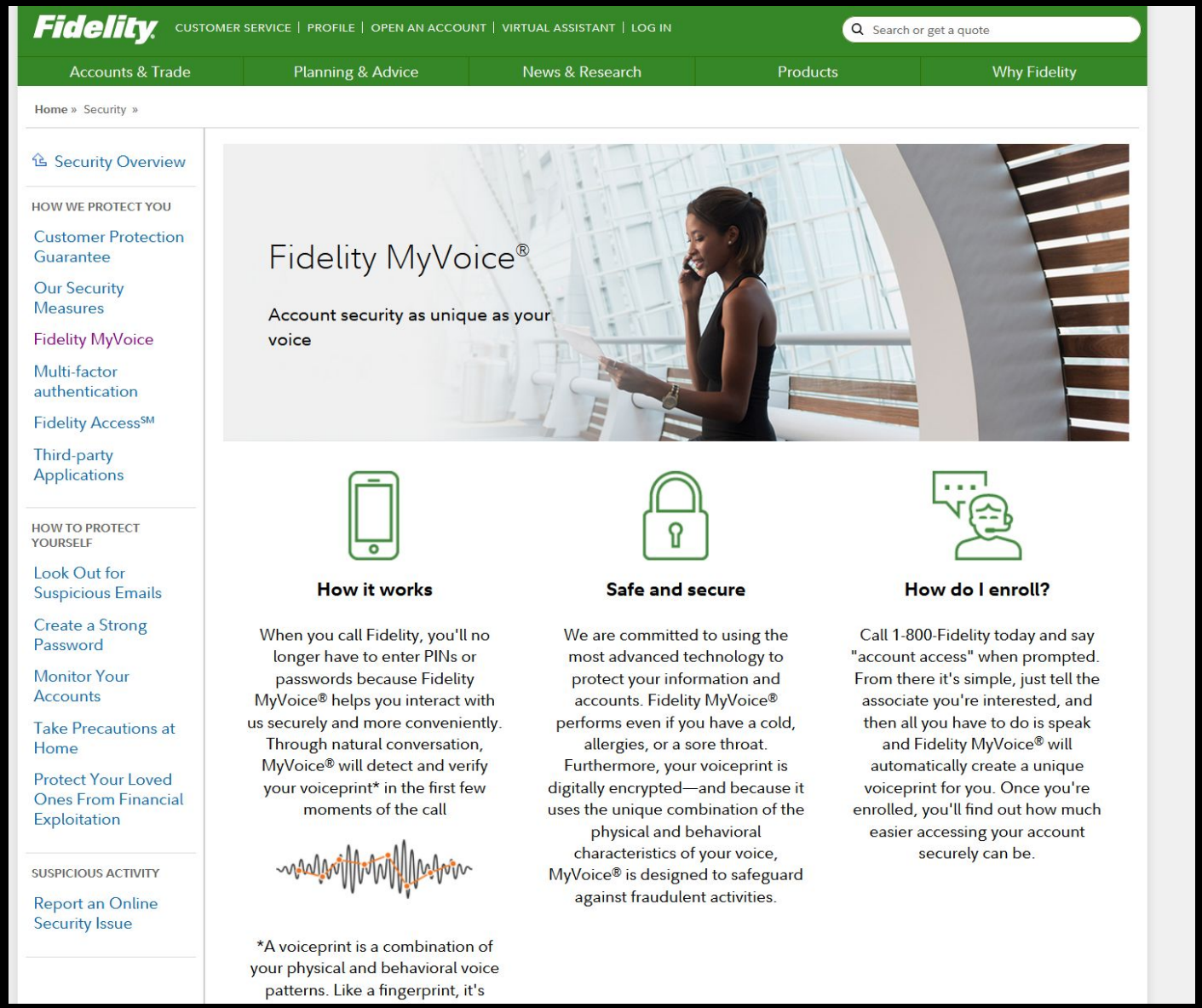
Demonstration Discussion

The institution’s name and specific details are anonymized to maintain confidentiality and adhere to ethical standards in security research.

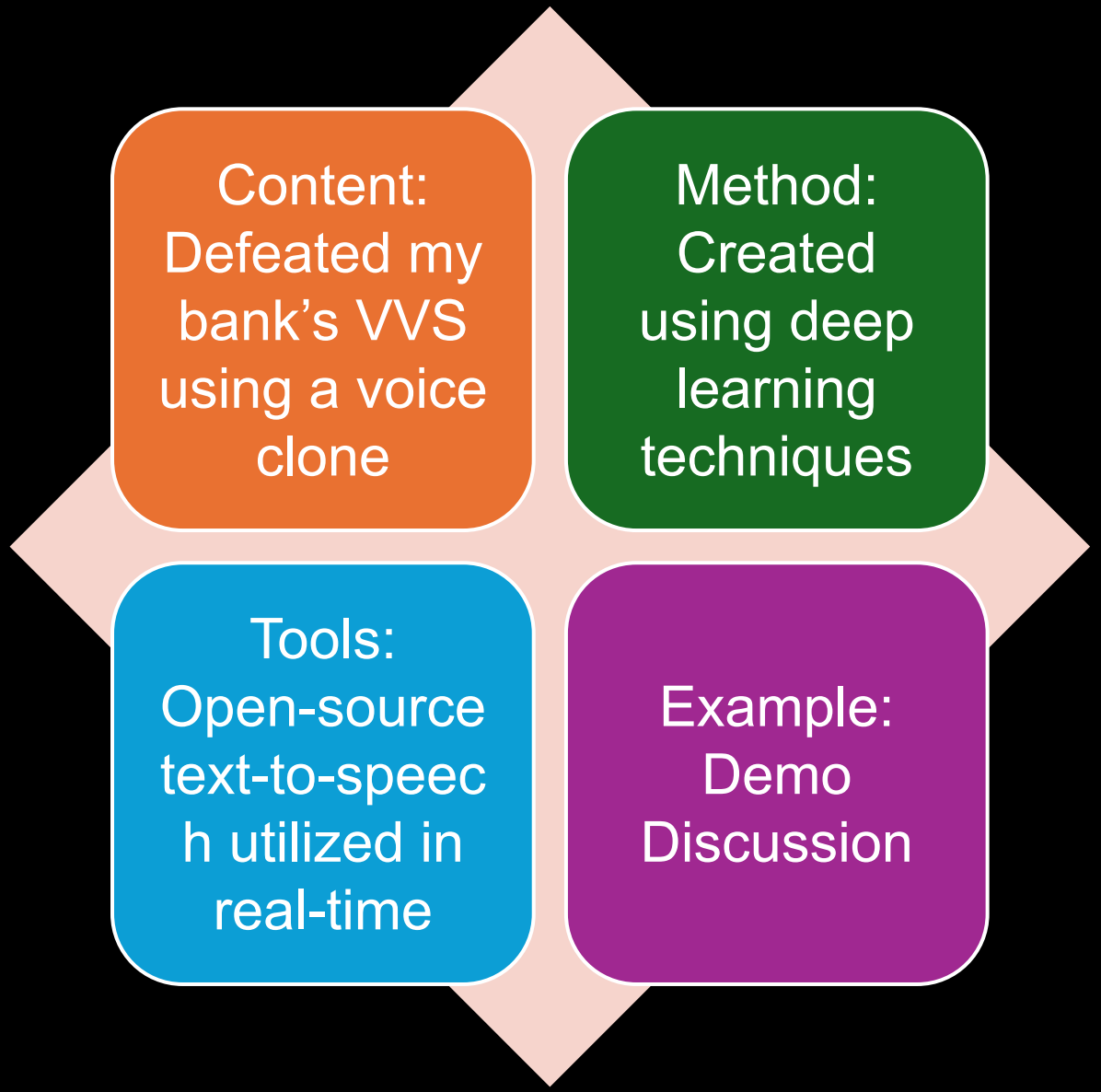
Examples


Eastern Bank
JOIN US FOR GOOD
Find a Branch or ATM
Join Our Team
Foundation
Need help?
Personal
Business
Commercial
Investing
Insights
About Us
Login
Voice ID - Your Unique Voiceprint
Personal Overview
Customer Service
Mobile & Online Banking
Online Banking
Mobile Banking
Mobile Security
Quicken
Telephone Banking
Voice ID
Online Loan Payment
Accessing your account information by phone just got easier.
Now when you call 1-800-EASTERN, you can do your banking faster and more securely thanks to our new Voice ID technology. You won't have to go through the usual security question process or need to find information about your account.
Our new technology authenticates your voice as you speak naturally with an Eastern Bank customer service representative. So you can get right to the reason you called without the wait.
Call now to set up Voice ID.
Just call 1-800-EASTERN (327-8376) to speak to customer service to record your voice.
Simple to use - The next time you call, your voice will be your main identification.
Faster service - No need to take the time to find lengthy transaction information.
More secure - Voice ID matches your unique voiceprint using what's called voice biometrics to verify it's you.
Eastern Bank is the first bank in North America to bring you
world leader in speech
ftware.[1] It's one more way we

Examples


Fidelity
CUSTOMER SERVICE | PROFILE | OPEN AN ACCOUNT | VIRTUAL ASSISTANT | LOG IN
Search or get a quote
Accounts & Trade
Planning & Advice
News & Research
Products
Why Fidelity
Home » Security »
Security Overview
HOW WE PROTECT YOU
Customer Protection Guarantee
Our Security Measures
Fidelity MyVoice
Multi-factor authentication
Fidelity Access℠
Third-party Applications
HOW TO PROTECT YOURSELF
Look Out for Suspicious Emails
Create a Strong Password
Monitor Your Accounts
Take Precautions at Home
Protect Your Loved Ones From Financial Exploitation
SUSPICIOUS ACTIVITY
Report an Online Security Issue
Fidelity MyVoice®
Account security as unique as your voice
How it works
When you call Fidelity, you'll no longer have to enter PINs or passwords because Fidelity MyVoice® helps you interact with us securely and more conveniently. Through natural conversation, MyVoice® will detect and verify your voiceprint* in the first few moments of the call
*A voiceprint is a combination of your physical and behavioral voice patterns. Like a fingerprint, it's
Safe and secure
We are committed to using the most advanced technology to protect your information and accounts. Fidelity MyVoice® performs even if you have a cold, allergies, or a sore throat. Furthermore, your voiceprint is digitally encrypted—and because it uses the unique combination of the physical and behavioral characteristics of your voice, MyVoice® is designed to safeguard against fraudulent activities.
How do I enroll?
Call 1-800-Fidelity today and say "account access" when prompted. From there it's simple, just tell the associate you're interested, and then all you have to do is speak and Fidelity MyVoice® will automatically create a unique voiceprint for you. Once you're enrolled, you'll find out how much easier accessing your account securely can be.

Demonstration Overview



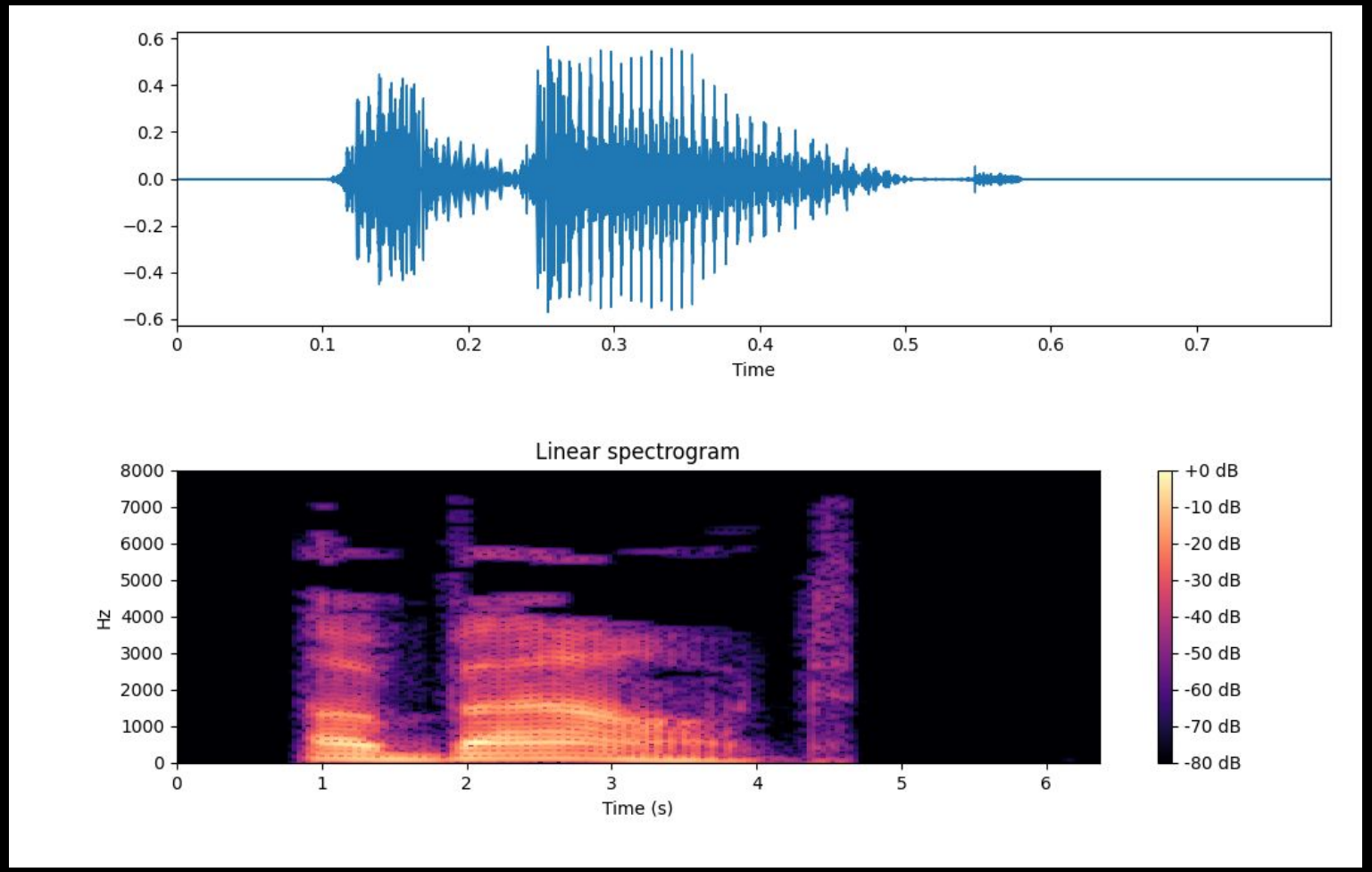
How I Cloned My Voice

Spectrogram Analysis: Visual representation of the spectrum of frequencies generated by AI

Setup: Overview of the hardware and software setup – Lambda Labs AI Desktop Vector

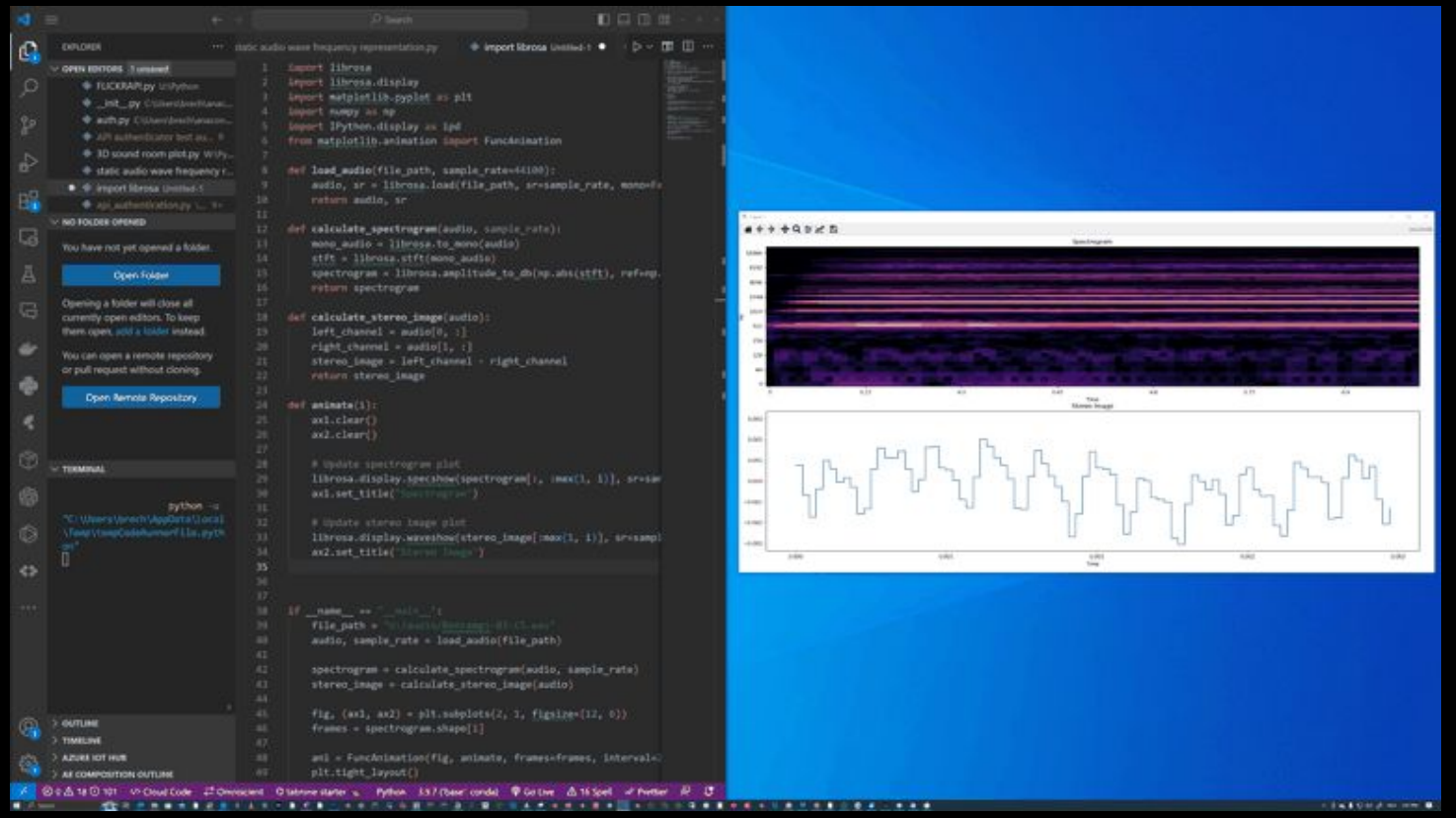
Frameworks used: Tacotron-2, WaveGlow, Mozilla TTS

From a Wave File to a Spectrogram



Python

- Python can be used to convert a wave file to a Mel spectrogram using various libraries like librosa, matplotlib, and numpy.
- This code snippet will convert a wave file to a Mel spectrogram and display it using a color-coded heat map, where the intensity represents the amplitude of the frequency



Lambda Labs AI Vector



Methodology

Experiment Design:

1. Selection of VVS to be tested (Nuance)

2. Tools & Technology used for Voice Cloning (Tacotron 2, WaveGlow, Mozilla TTS, AI Vector)

3. Procedure for creating and testing cloned voices

Steps

1

Collecting sample voice data

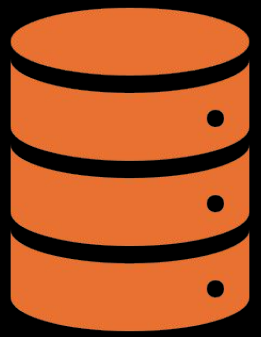
2

Training AI models for voice cloning

3

Attempting to bypass VVS using cloned voices

Voice Cloning Process



Data Collection – Several hours



My voice samples were collected (Blue Yeti, Audacity) ensuring high-quality recordings for accurate cloning

AI Models



TACOTRON 2 &
WAVEGLOW



TRAINING PROCESS &
PARAMETERS

Experiment Execution

Testing Phase:

- How my cloned voice was tested against the VVS
- Parameters and criteria for success/failure

Results:

- Success rate of my cloned voice in bypassing VVS
- Specific observations and anomalies

Results Analysis

How & Why the VVS was bypassed

Patterns or weaknesses identified in the VVS

Countermeasures (CMs)

Preventative Measures: How to strengthen VVS against spoofing attacks

MFA, Liveness Detection, Advanced ML Algorithms, Speech Signal Analysis, Continuous Authentication, Behavioral Biometrics

Training Systems to Detect Differences in Real v. Synthetic Voices

- Enhanced Security Against Fraud
- Improving Detection Algorithms
- Preserving User Trust
- Adaptation to Evolving Threats



4 Key Takeaways

1. VVS can be compromised using advanced AI techniques

2. The cloned voice successfully bypassed the VVS, showcasing the sophistication and risks of modern voice synthesis technologies.

3. The experiment underscores a significant security flaw in relying solely on voice verification for sensitive operations like banking transactions.

4. Financial institutions should integrate additional security layers such as behavioral biometrics, context-based authentication, and continuous monitoring.

Real- World Implications

If I can do it, cyber criminals can do it.

Adversary emulation is a critical practice for cybersecurity professionals because it allows them to simulate real-world cyber attacks in a controlled environment, closely mirroring the tactics, techniques, and procedures used by actual threat actors. By understanding the methods and tools cybercriminals employ, security teams can better anticipate potential vulnerabilities and strengthen their defenses accordingly. Demonstrating the ease with which cybercriminals can achieve their malicious goals serves as a stark wake-up call, highlighting the urgent need for robust security measures. This approach not only educates and trains cybersecurity personnel but also underscores the importance of proactive security strategies, fostering a culture of vigilance and continuous improvement in defense mechanisms.



References

- https://brechtcorbeelsaudioengineering.quora.com/https-www-quora-com-Can-you-provide-a-step-by-step-guide-on-creating-visual-representations-of-audio-data-such-as-spec?
- https://eeghacker.blogspot.com/2014/05/eeg-as-wav-files-go-spectrograms.html
- https://www.kaggle.com/code/msripooja/steps-to-convert-audio-clip-to-spectrogram
- https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53
- https://github.com/andrekassis/Breaking-Security-Critical-Voice-Authentication



Special Thanks to:

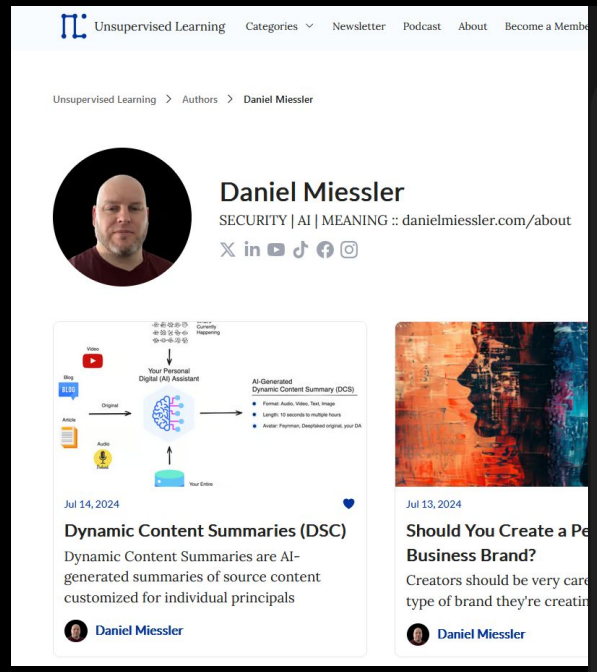


Andre Kassis - Machine Learning Security and Privacy Researcher - University of Waterloo / Ph.D., Computer Science.

Lambda Labs - https://lambdalabs.com/

Adversary Village and in particular, Abhijith B R

Thank you also to all the “homies” who helped inspire this presentation





Also, thank you to the Immaculata, Charlotte, Michael, Therese, and Gianna for always rooting me on.