



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 Verificatio n 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

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

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 bona fide 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 assume large query sets, when attacked, do not hold. Our analysis shows that failure to detect spoofing in real-time, necessitates the need to integrate key message distinguishing features that are easily distinguishable but are subtle enough to be spoofed. We can still bypass such textual content-based attacks by performing a CMs bypass attack that yields success.

We perform

CMs, bypass

of potential

call centers.

VA systems

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®**

aculab



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



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 the world leader in speech recognition software.¹ It's one more way we're

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 AccessSM](#)

[Third-party Applications](#)

HOW TO PROTECT YOURSELF

[Look Out for Suspicious Emails](#)

[Create a Strong Password](#)

[Monitor Your Accounts](#)

[Take Precautions at Home](#)

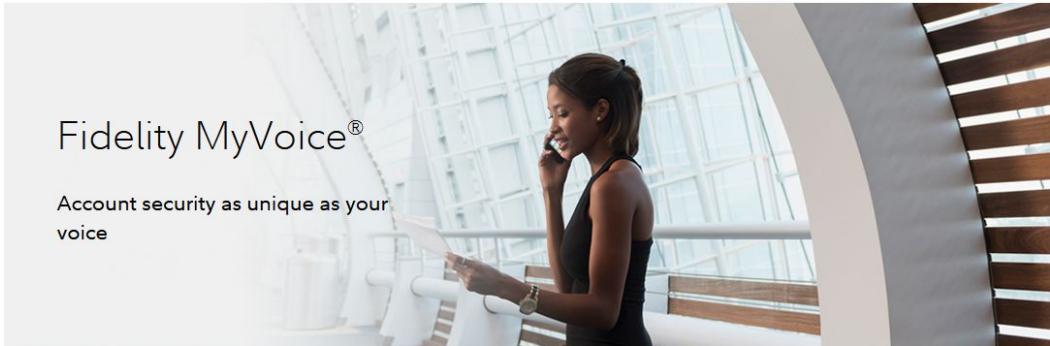
[Protect Your Loved Ones From Financial Exploitation](#)

SUSPICIOUS ACTIVITY

[Report an Online Security Issue](#)

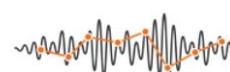
Fidelity MyVoice®

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



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.

*A voiceprint is a combination of your physical and behavioral voice patterns. Like a fingerprint, it's

Demonstration Overview

Content:
Defeated my bank's VVS
using a voice clone

Method:
Created using deep learning techniques

Tools:
Open-source text-to-speech utilized in real-time

Example:
Demo Discussion

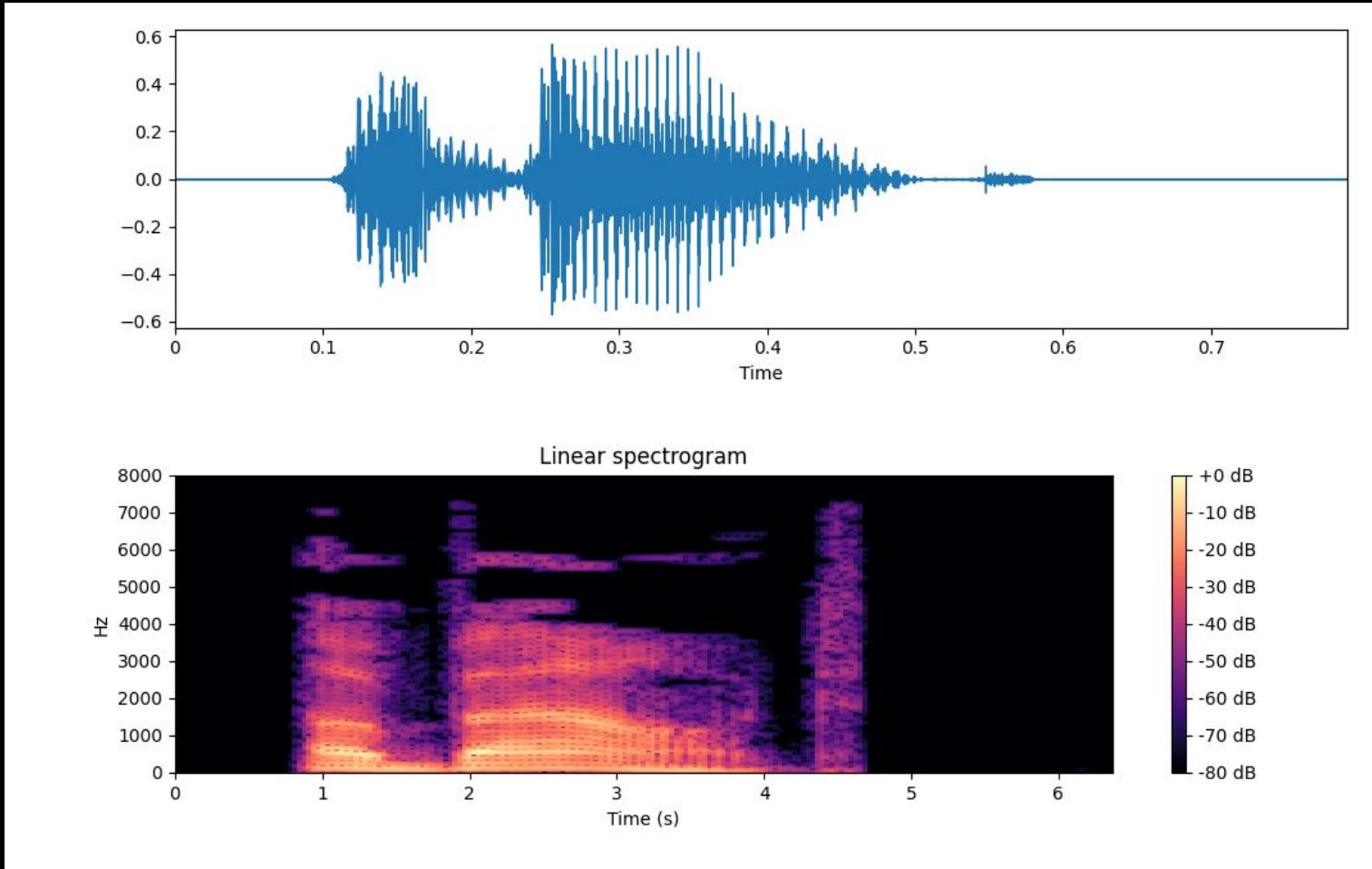
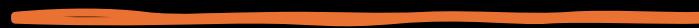
How I Cloned My Voice

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 Spectrogra m



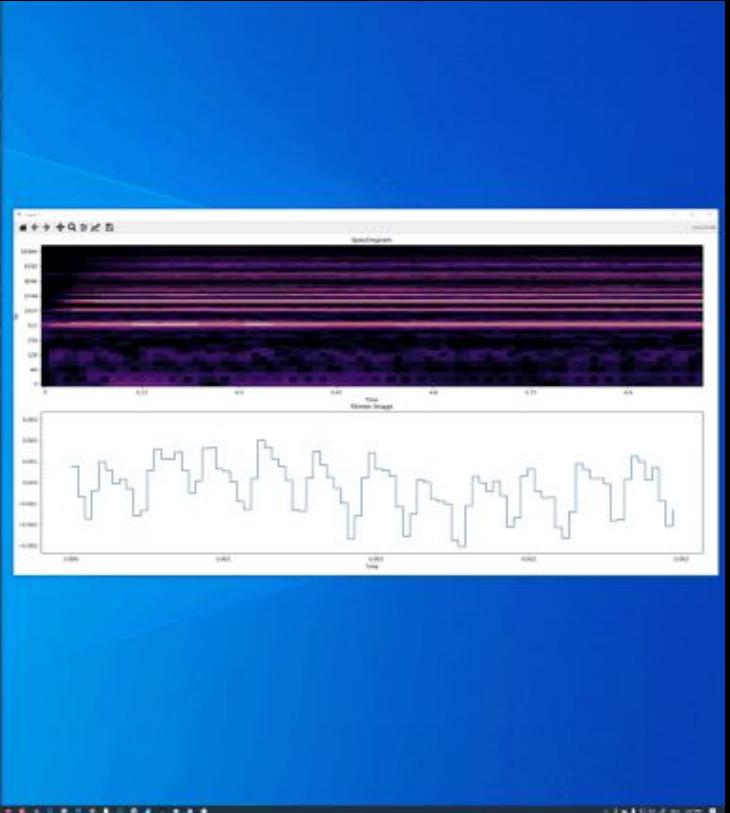
Python

- 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

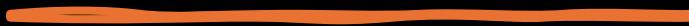
The screenshot shows a Jupyter Notebook interface with the following details:

- EXPLORER**: Shows open editors: `FECKRAMP.ipynb`, `int_py ColabEnvironment.ipynb`, `auth_py ClientEnvironment.ipynb`, `AWS authentification test.ipynb`, `3D sound room plot.ipynb`, `static audio wave frequency representation.ipynb`, and `Import librosa Unlocked.ipynb`. It also lists "NO FOLDER OPENED" and "You have not yet opened a folder." with an "Open Folder" button.
- TERMINAL**: Displays a command-line session:

```
python -u
"C:\Users\vench\MyData\Local\Temp\VisualCode\environment\lisa.py"
[2]
```
- CODE CELLS**: The main area contains two code cells. The first cell is titled "static audio wave frequency representation.py" and imports `librosa`, `librosa.display`, `matplotlib.pyplot as plt`, `numpy as np`, `ipython.display as ipd`, and `from matplotlib.animation import FuncAnimation`. It defines `load_audio` to load audio from a file path at sample rate 44100, and `calculate_spectrogram` to convert mono audio to a spectrogram using STFT and amplitude_to_db. The second cell is titled "import librosa Unlocked.ipynb" and imports `librosa`, `librosa.display`, `matplotlib.pyplot as plt`, and `FuncAnimation`. It defines `animate` to clear axes and update spectrograms, and `animate_stereo_image` to clear axes and update stereo images. It then uses `FuncAnimation` to play a spectrogram and a stereo image. Finally, it checks if the name is __main__ and runs `load_audio` on a file path, calculates a spectrogram, creates a stereo image, and plots them in a figure with frames and interval.



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



Collecting sample
voice data

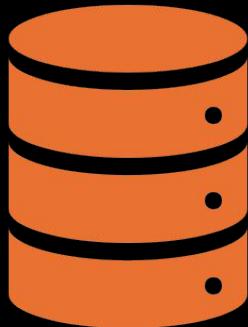


Training AI models
for voice cloning



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 Takeaway

S

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

Unsupervised Learning Categories Newsletter Podcast About Become a Member

Unsupervised Learning > Authors > Daniel Miessler



Daniel Miessler
SECURITY | AI | MEANING :: danielmiessler.com/about
X in YouTube Twitter Instagram

Jul 14, 2024  

Dynamic Content Summaries (DSC)
Dynamic Content Summaries are AI-generated summaries of source content customized for individual principals

 Daniel Miessler

Training Corporate Training Consulting Media

ARCANUM About Us Work with Us

Arcanum x JHaddix



Jason has had a distinguished 20-year career in cybersecurity, previously serving as CISO of Buddobot, CISO of Ubisoft, Head of Trust/Security/Operations at Bugcrowd, Director of Penetration Testing at HP, and Lead Penetration Tester at Redspin.

He has also held positions doing mobile penetration testing, network/infrastructure security assessments, and static analysis.

Jason is a hacker, bug hunter, and is currently ranked 57th all-time on Bugcrowd's bug bounty leaderboards. Currently, he specializes in recon, web application analysis, and emerging technologies.

Jason has also authored many talks for world-renowned conferences like DEF CON, Bsides, Black Hat, RSA, OWASP, Nullcon, SANS, IANS, BruCon, ToorCon, and many more.

Work with Us



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