Tubes Among US: Analog Attack on Automatic Speaker Identification

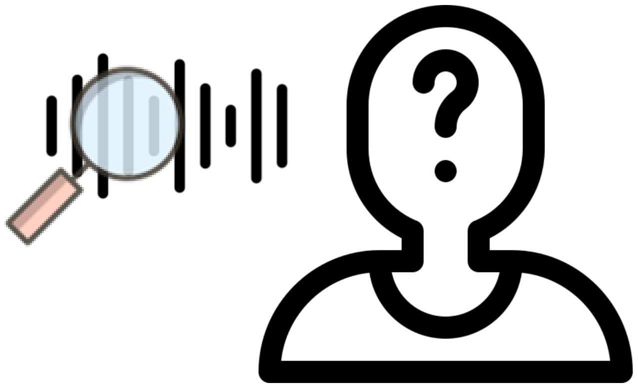
Shimaa Ahmed, Yash Wani, Ali Shahin Shamsabadi, Mohammad Yaghini, Ilia Shumailov, Nicolas Papernot, and Kassem Fawaz

The 32st USENIX Security Symposium, 2023

Anaheim, CA, USA

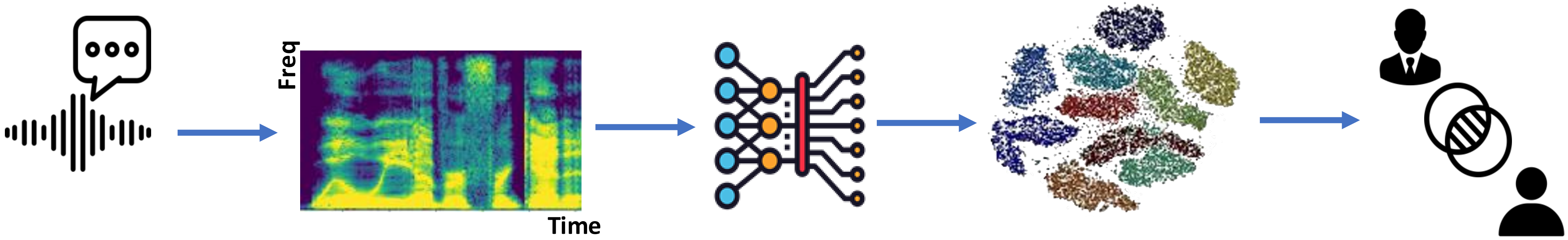
“Speak, so that I may *identify* you.”

- Our voices are distinct
- Voice biometrics can identify people



This Technology is called Speaker Identification

Speaker Identification

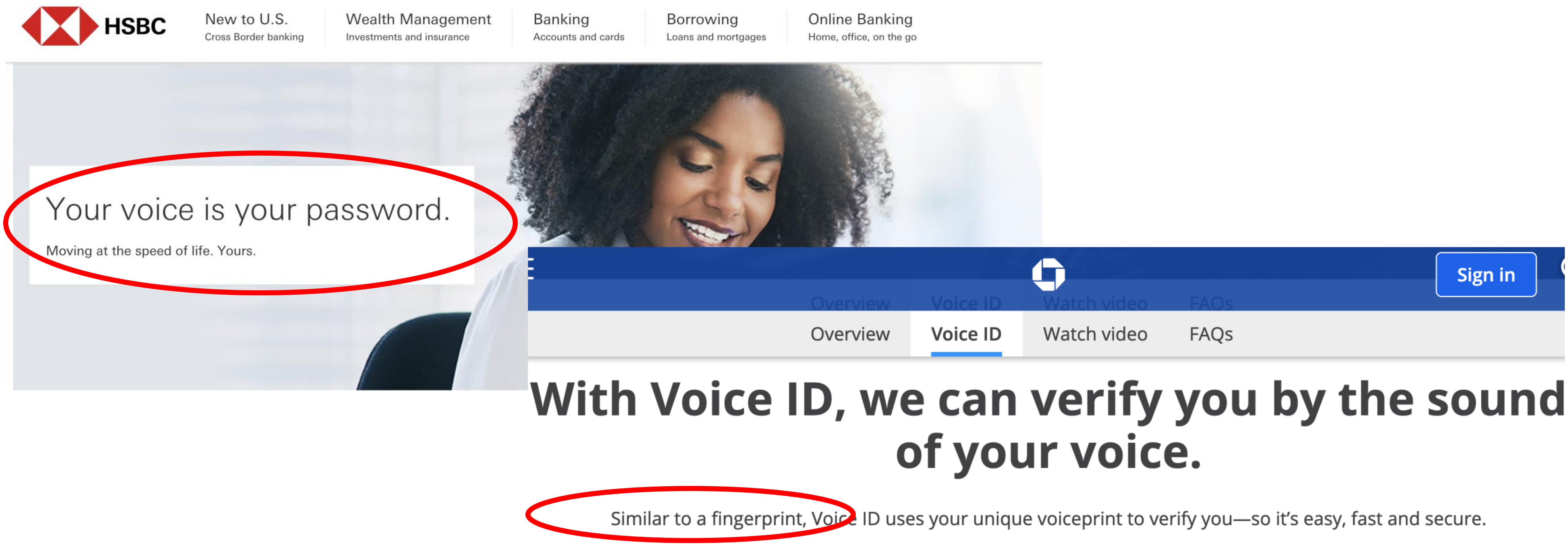


Applications of Speaker Identification

- Voice-enabled devices; Siri
  - Integrity
  - Personalization

- Phone banking
  - Seamless Identification and Authentication

Security-critical applications



Is voice as secure as a password or a fingerprint?

THE WALL STREET JOURNAL.

English Edition ▾ | Print Edition | Video | Podcasts | Latest Headlines

Home World U.S. Politics Economy Business Tech Markets Opinion Books & Arts Real Estate Life & Work WSJ. Magazine Sports

PRO CYBER NEWS

Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case

Scams using artificial intelligence are a new challenge for companies

PHOTO: SIMON DAWSON/BLOOMBERG NEWS

By Catherine Stupp

UPCOMING EVENTS

Nov 9 2021 — 7:00 PM - 8:00 PM EDT — WSJ Opinion: A Mom Term for the Suprem

Nov 17 2021 — 11:00 AM - 3:00 PM EDT — WSJ Pro Sustainable Forum

Nov 22 2021 — 12:00 PM - 1:45 PM EDT — WSJ Women In: Navi New Normal at Work

ADD TO CALENDAR

MOST POPULAR NEWS

1. Link Between Covid-19 Vaccines and Myocarditis Probed

Forbes

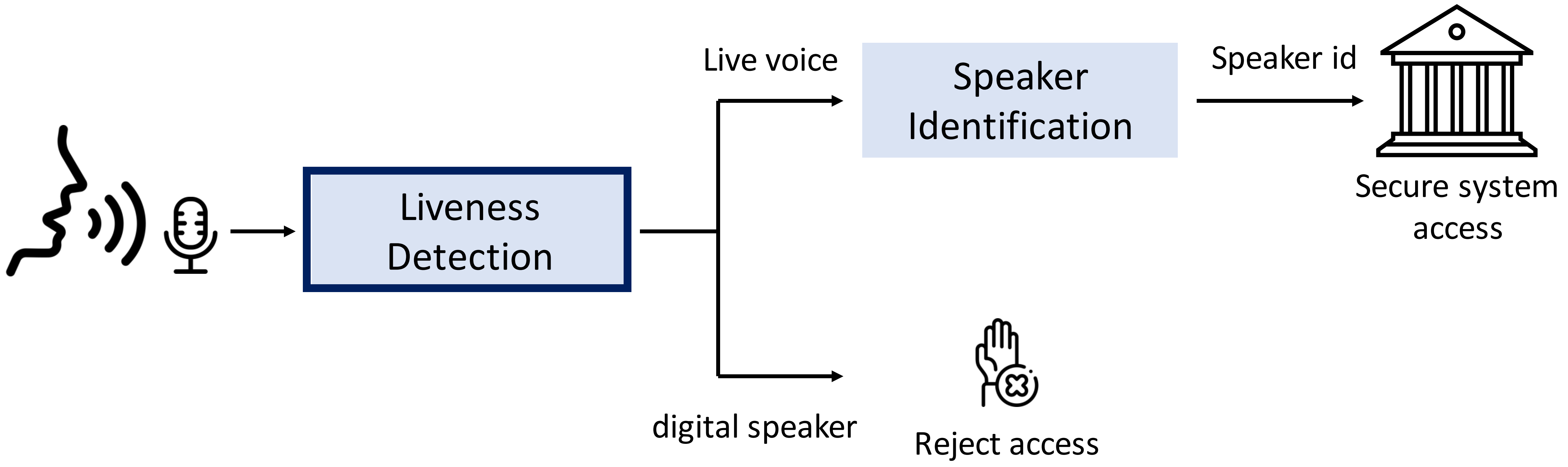
EDITORS' PICK | Oct 14, 2021, 07:01am EDT | 73,591 views

Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find

Thomas Brewster Forbes Staff
Cybersecurity
Associate editor at Forbes, covering cybercrime, privacy, security and surveillance.

Securing Speaker Identification against attacks

- Speaker Identification + Liveness detection



Liveness Detection

- <u>Assumption</u>: Voice is authentic *if* it comes from a human

HSBC

New to U.S.
Cross Border banking

Wealth Management
Investments and insurance

Banking
Accounts and cards

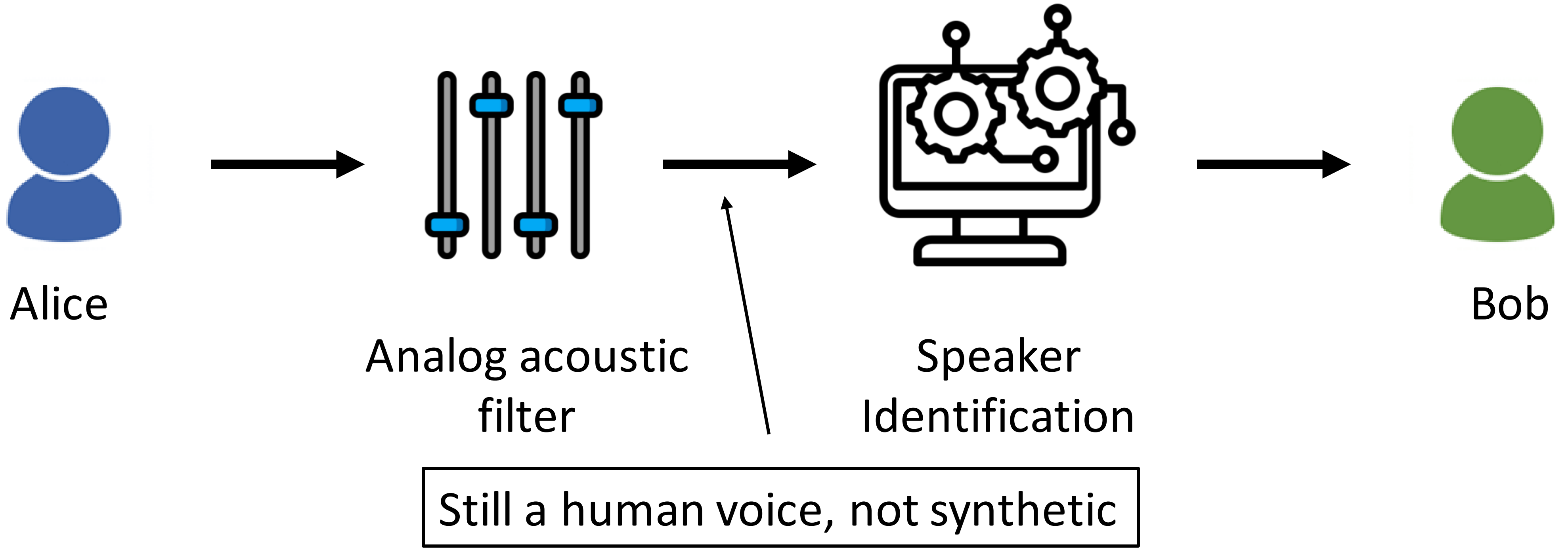
Borrowing
Loans and mortgages

Online Banking
Home, office, on the go

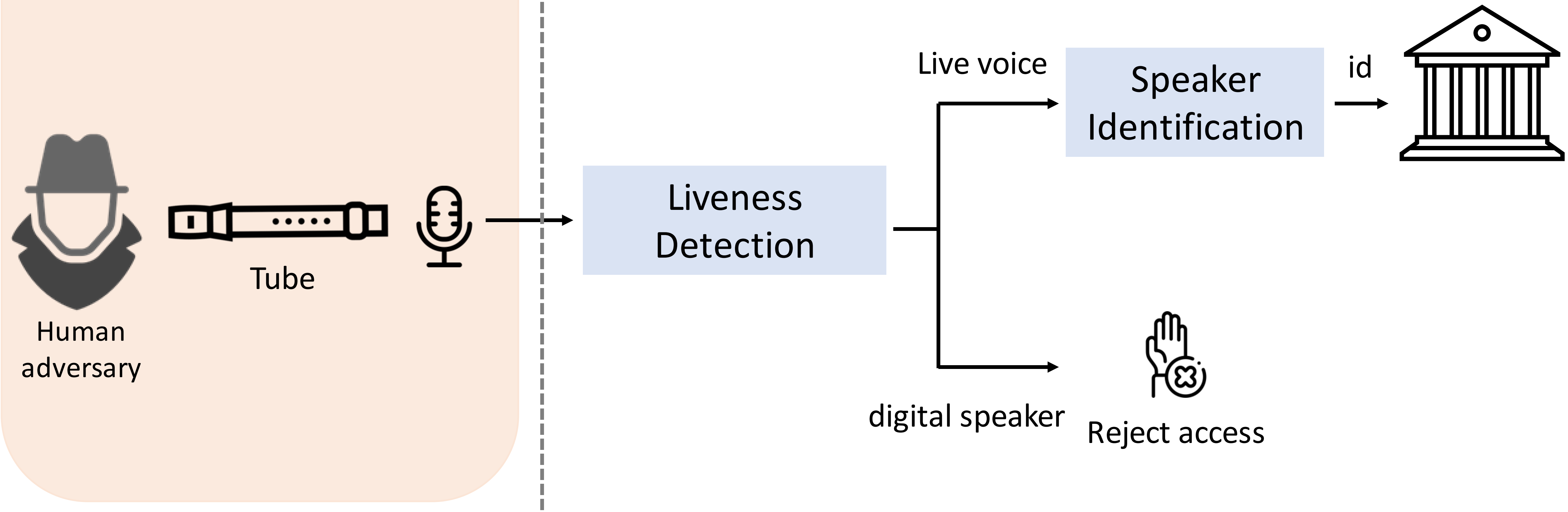
How is it safer?
Fraudsters and hackers may be able to steal or guess your passwords, but they can't replicate your voice. We measure the mechanics of how sounds are produced, rather than the sounds themselves. Voice ID is sensitive and sophisticated enough to detect if someone is impersonating you or playing a recording – and recognize you, even if you have a cold or sore throat.

Breaking the assumption

- What if there is an attack that is *not synthetic*
- An attack that can reshape one's voice to sound like another



Mystique



Acoustic Environment

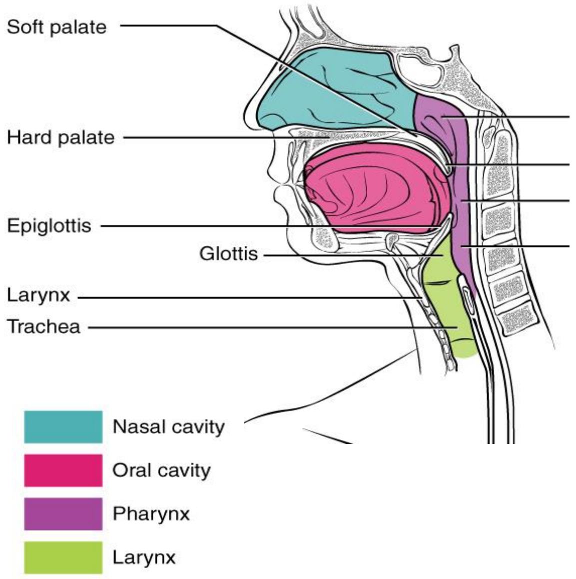
System Under Attack

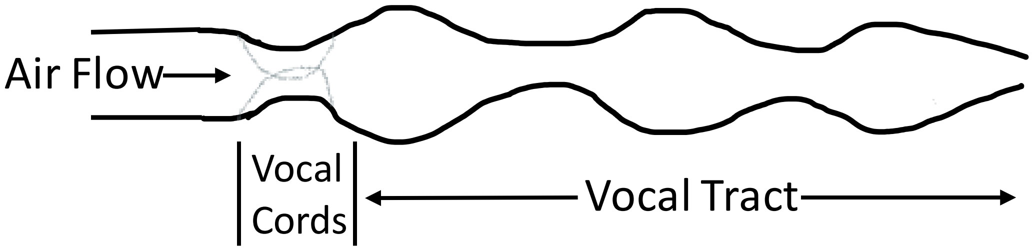
Lea

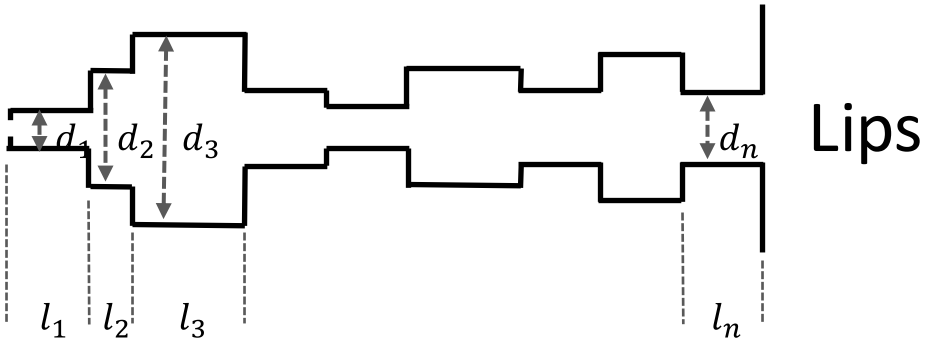
WiPi

Why it works?

The human vocal tract is a resonator



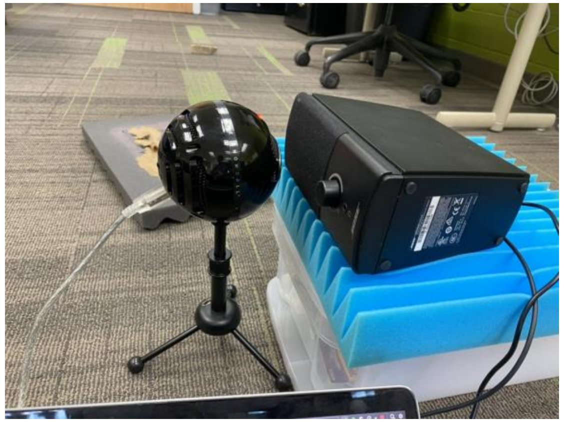




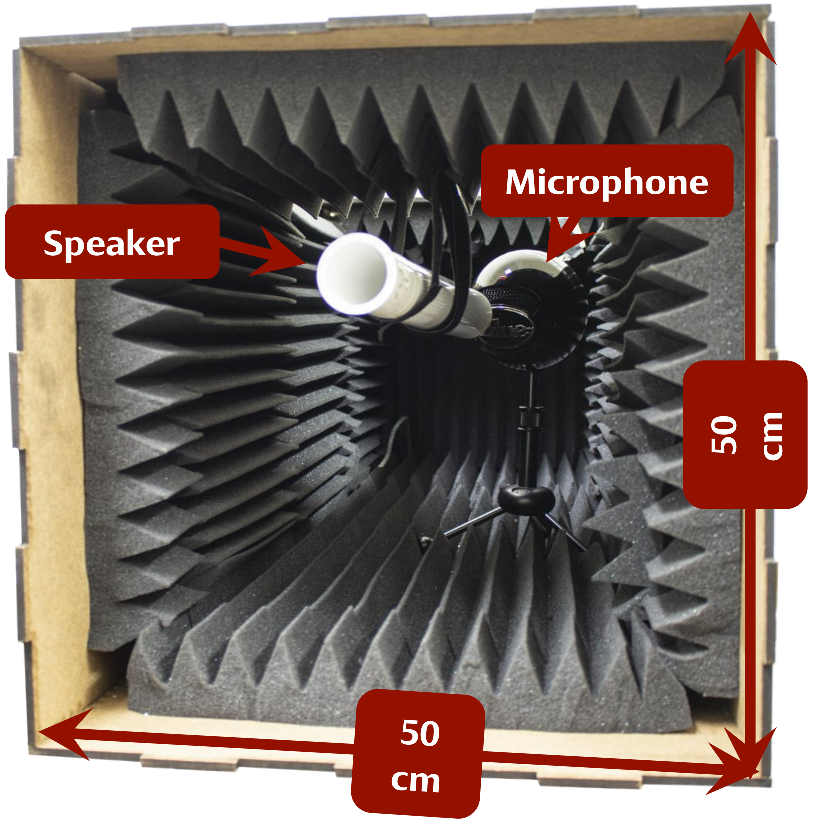
A tube is a controllable extension of the vocal tract

Experiment Setup

- Lots of unsuccessful trials!

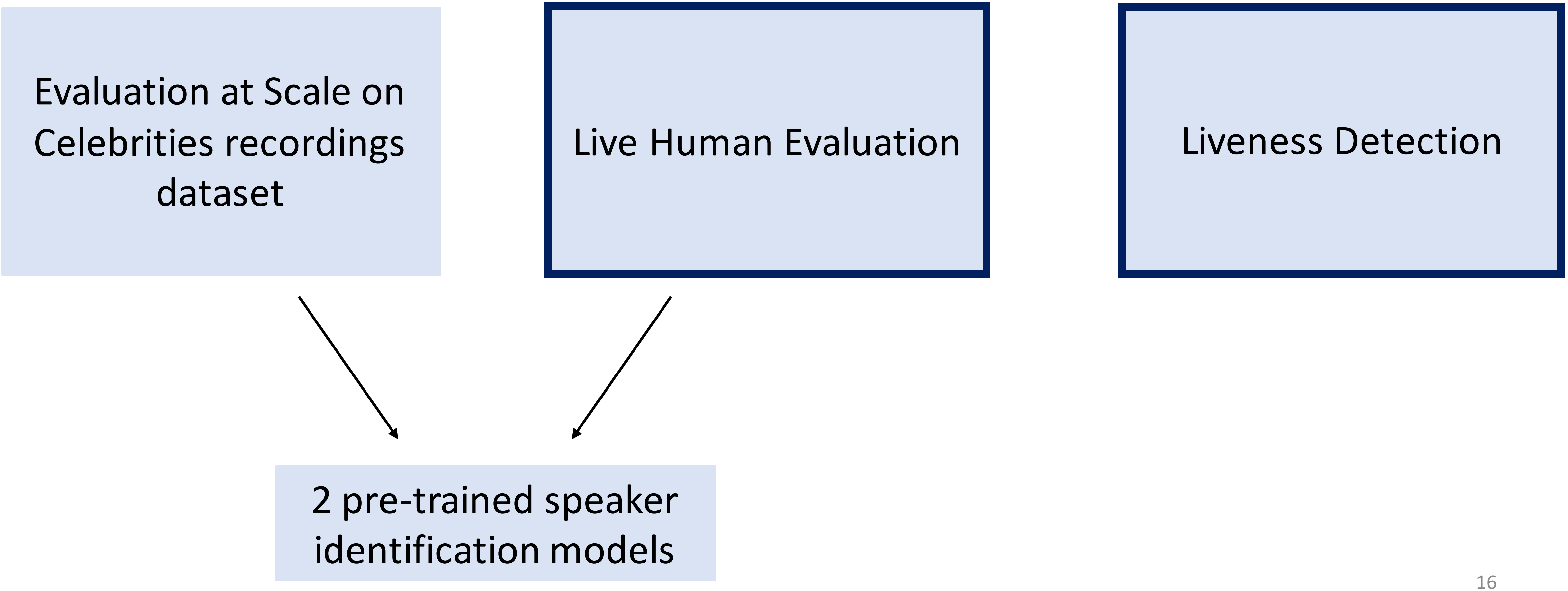



Experiment Setup: The one that works



➤ Ask me what the tube filter look like in the measurements

Evaluation pipeline



Human Evaluation of Mystique

- 14 participants: 8 male and 6 female
- Using Mystique to impersonate celebrities

ID	0	1	2	3	4	5	6	7	8	9	10	11	12	13
sex	F	M	M	F	F	M	M	F	M	M	F	M	M	F
Success Rate (%)	56	66	72	74	65	49	60	75	47	53	70	61	35	78

- Attack success rate ranges from 35% to 75%

➢ Ask me about the baseline impersonation success rate (without Mystique)

Liveness detection

- State-of-the-art liveness detection methods fail to detect Mystique as an attack

Model	EER	FAR @ FRR=0
LA-LCNN	30%	77%
PA-LCNN	31%	99%
Void-SVM	62%	98%
Void-DNN	35%	92%
Void-LCNN	33%	93%

Lower is better

- ML models fit to their training data
  - A tube as an analog acoustic filter falls outside their training distribution

➢ Ask me about our attempts to break Mystique

Conclusion

- Speaker identification models make assumptions about the physical world
  - Expected noise and reverberation
  - Hardware
- These assumptions do not always hold in real-life
  - The analog space is very diverse
- Adversaries can exploit these assumptions to impleme ML services
- Biometric authentication is not as secure as advertised