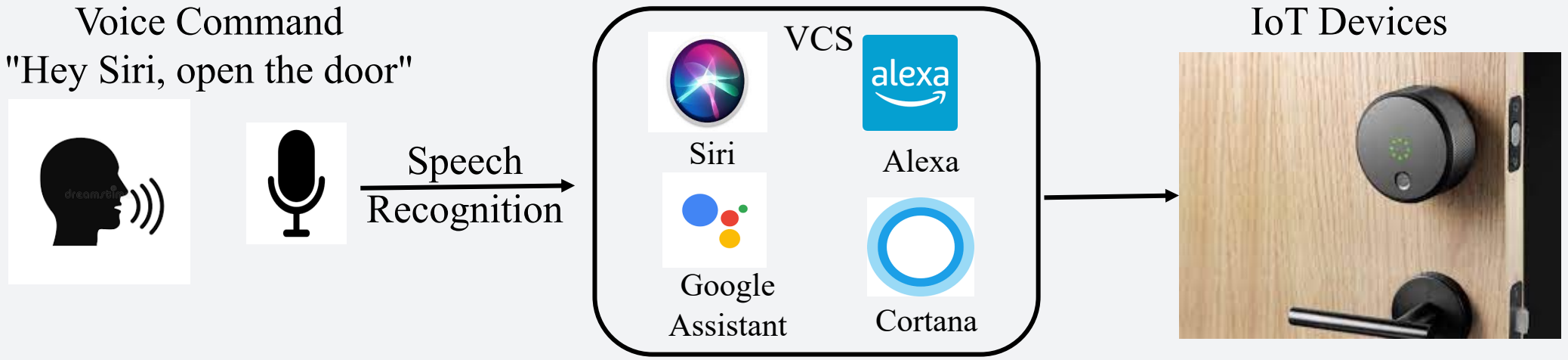
Near-Ultrasound Inaudible Trojan (NUIT): Exploiting Your Speaker to Attack Your Microphone.

Presented by and Qi Xia (UTSA PhD Candidate)
Co-Author: Dr. Guenevere Chen (UTSA), Dr. Shouhuai Xu (UCCS)

UTSA
The University of Texas at San Antonio™
Department of Electrical and Computer Engineering

Voice Controllable System (VCSs)



Motivation

Existing Attack (Dolphin Attack[1]): (Inaudible/Physical attack)

1) Use DSB-AM to modulate voice command to ultrasonic frequency
2) Attack VCS inaudibly by exploiting Microphone's nonlinearity

Ultrasonic Transducer sends out DSB-AM modulated ultrasonic command

Reference:
[1] Zhang, Guoming, et al. "Dolphinattack: Inaudible voice commands." *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. 2017.

Research Question:

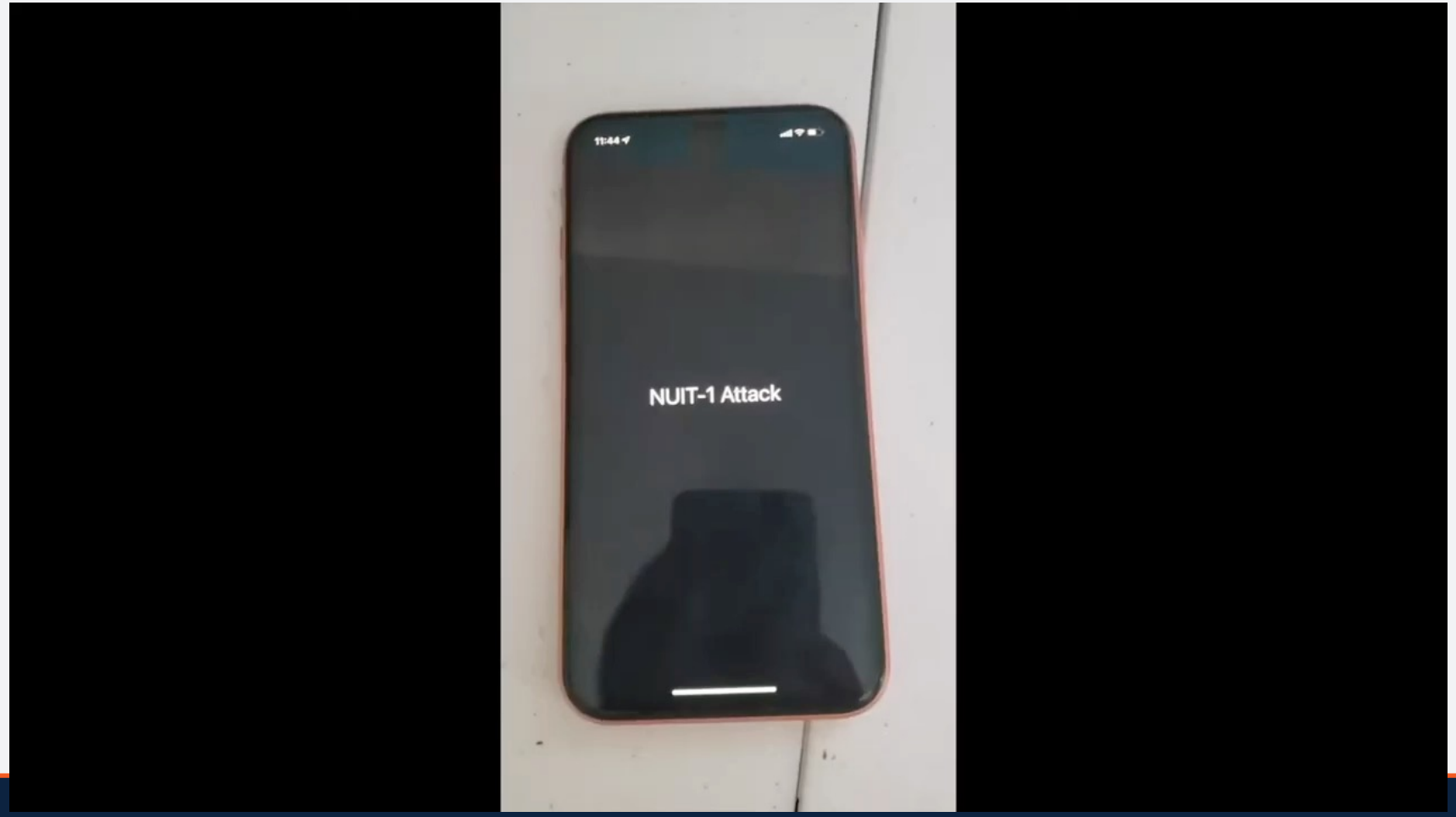
Is it possible to

- remotely wage Dolphin Attack (inaudible attacks) like a Trojan Horse Virus?
- Use victim's own speaker to attack his/her own microphone?

NUIT-1 Attack Demo

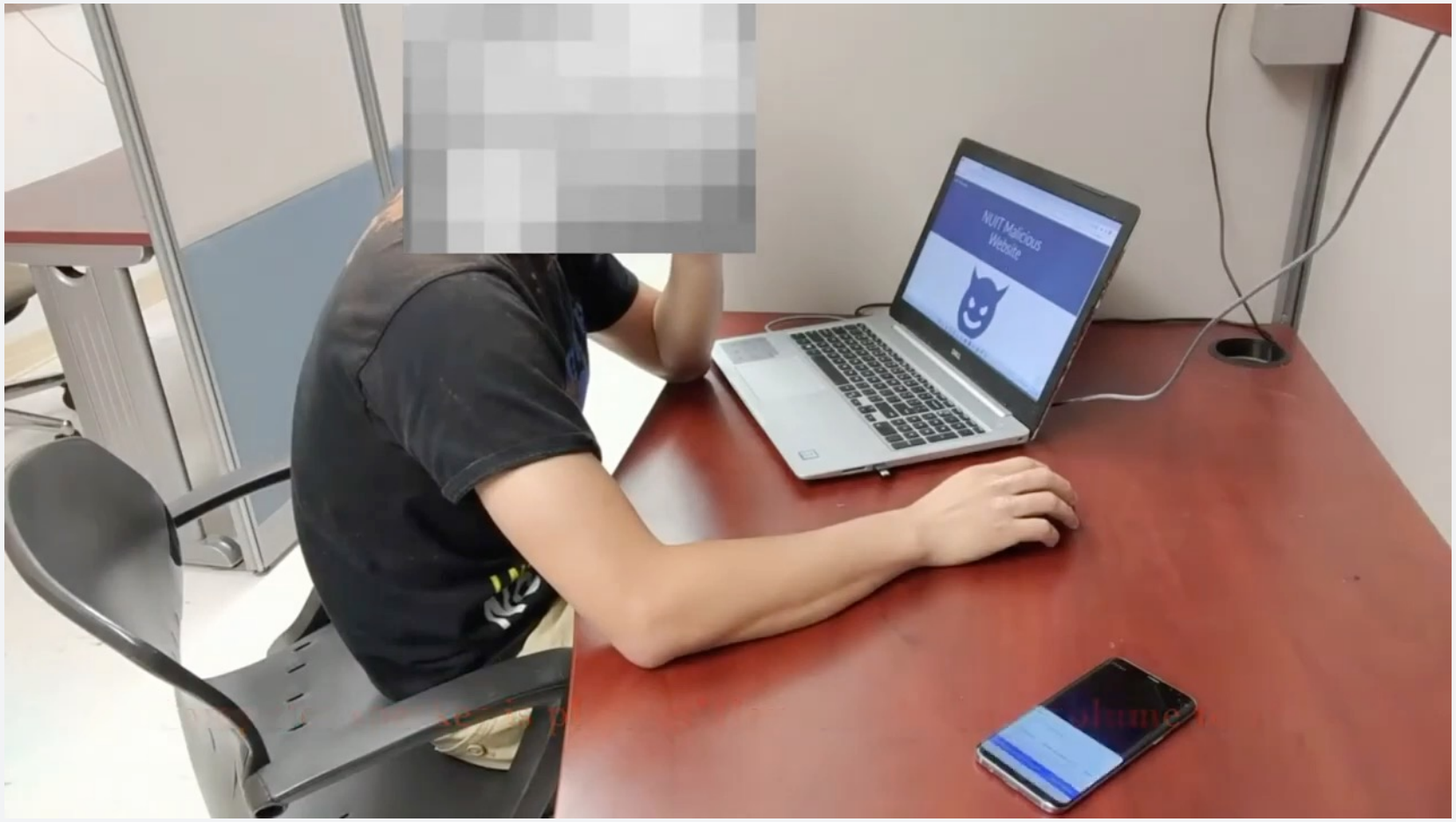
More demos can be seen at https://sites.google.com/view/nuitattack/home

NUIT-1: Attacker Exploits Victim's Speaker to Attack Victim's Microphone on the Same Device



NUIT-2 Attack Demos

NUIT2: Attacker Exploits Victim's Speaker to Attack Victim's Microphone on a Different Device



Contributions

1. NUIT attacks: *inaudibility* , *remote capability* and the *unnoticeable* as devices permit

2. Theoretically innovation: SSB-AM Nonlinear Demodulation

3. *New single-factor software-based* defense: *leverage input NUIT attack signal*

NUIT Attack Implementation

Assumption: Attacker can access victim's Voiceprint for Authentication

Step1: Prepare & record malicious voice commands

Step2: Modulates the input audio files to ensure inaudibility

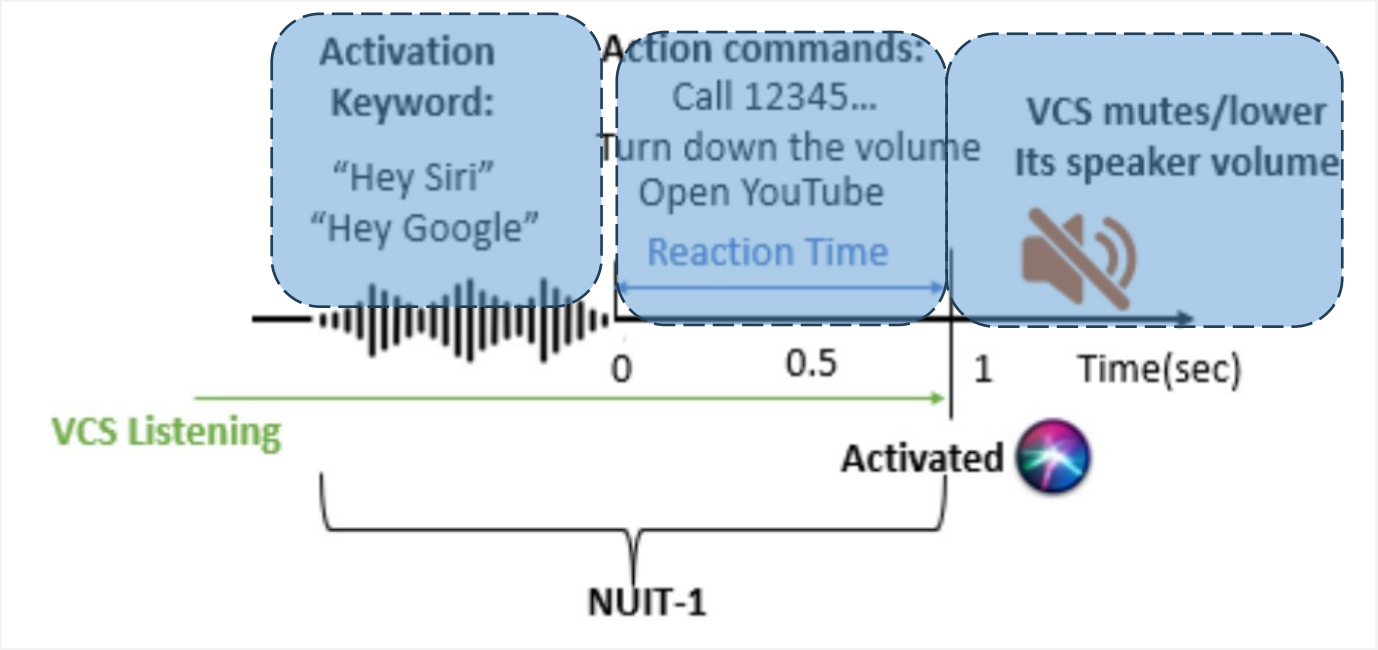
Step3: Remotely deliver the attack signal to the target device

- Embed into online audio/website/ app
- Trick victim to play through social engineer

Step1. Prepare & record malicious voice commands

- Attack command must be at least 6kHz to be recognizable by VCSs.

- For NUIT1: action command length <=0.77 s (reaction time window).



- Silent Response (for iPhones only): "Hey Siri, speak 6%" as the first command

NUIT Attack Implementation

Assumption: Attacker can access Voiceprint/Authentication

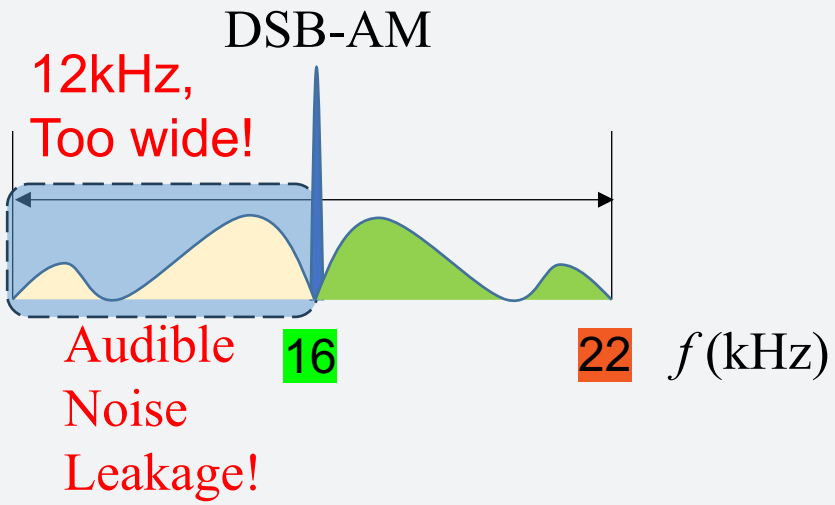
Step1: Prepare & record malicious voice commands

Step2: Modulates the input audio files to ensure inaudibility

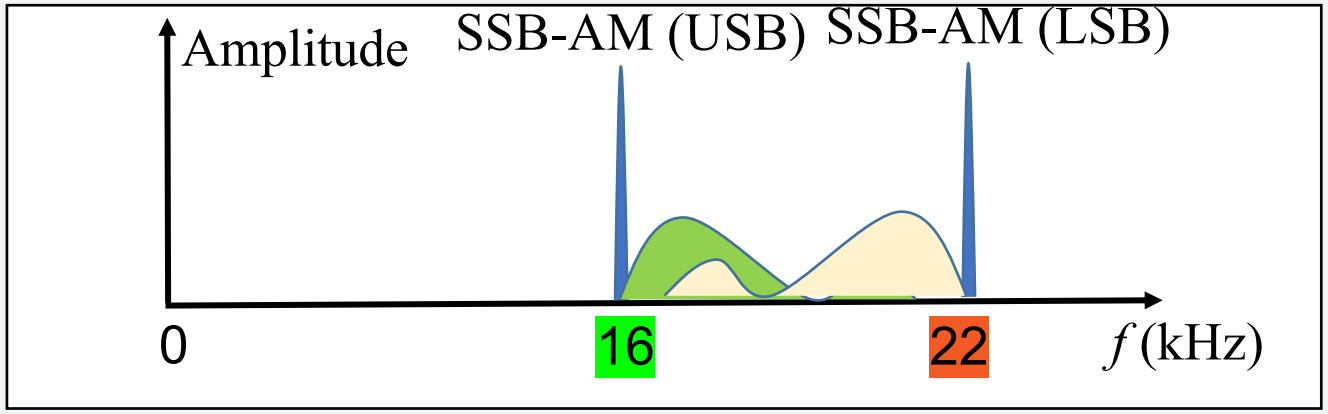
Step3: Remotely deliver the attack signal to the target device

- Embed into online audio/website/ app
- Trick victim to play through social engineer

Step2. Modulates the input audio files to ensure inaudibility



- Goal: modulate voice command into passband between 16kHz-22kHz
- What modulation scheme to use?
- DSB-AM fails, use SSB-AM!
- Two variations of SSB-AM:
  - USB -AM, Carrier Frequency: 16kHz
  - LSB-AM, Carrier Frequency: 22kHz



NUIT Attack Implementation

Assumption: Attacker can access Voiceprint/Authentication

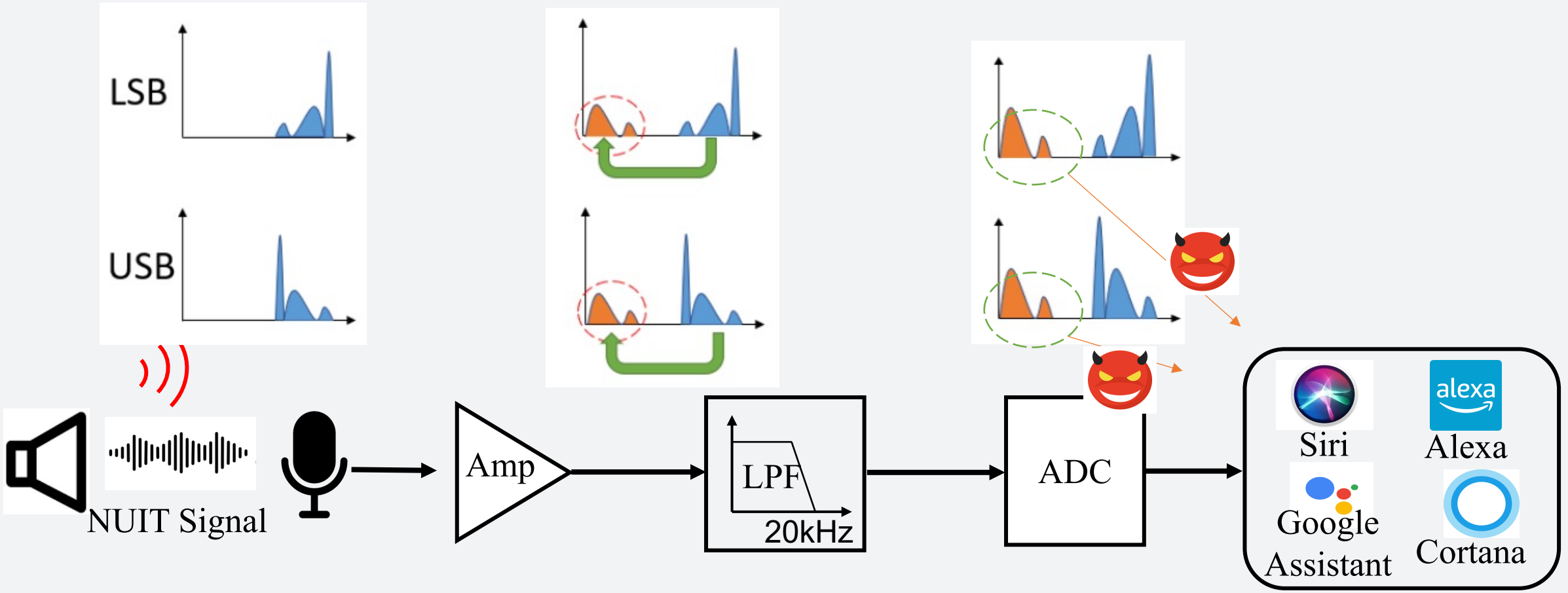
Step1: Prepare & record malicious voice commands

Step2: Modulates the input audio files to ensure inaudibility

Step3: Remotely deliver the attack signal to the target device

- Embed NUIT into online audio/website/ app
- Trick victim to play the attack signal through social engineering

NUIT Exploits the Vulnerability of Microphone Nonlinearly.



Experiments and Results

Devices Vulnerable to NUIT

- NUIT-2 impact more devices than NUIT-1
- Only Apple Siri is vulnerable to silent response NUIT attacks

Table 5: Devices vulnerable to NUIT, where ✓ means an attack succeeds with end-to-end unnoticeability, ✓* means an attack succeeds with inaudible attack signals but not silent response, and × means an attack fails.

Target VCS Device	NUIT-1	NUIT-2
iPhone: X, XR, 8	✓	✓
MacBook: Pro-2021, Air-2017	✓*	✓
Galaxy: S8, S9, A10e	✓*	✓
Echo Dot Gen1	✓*	✓
Dell Inspiron 15	✓*	✓*
Apple Watch 3	×	✓
Google Pixel 3	×	✓
Galaxy Tab S4	×	✓
LG Think Q V35	×	✓
Google Home 1	×	✓
Google Home 2	×	✓
iPhone 6 plus	×	×

NUIT-2 Attack Range

- Devices with powerful speaker (e.g. TV, Laptop, Vehicle Speaker) have longer attack range.
- Devices with low-power (e.g. phone speaker) have shorter attack range.

Table 14: Effectiveness of NUIT-2, where each cell describes the maximum distance (in centimeters) between the victim speaker device and the victim microphone device at which NUIT-2 succeeds with effectiveness $\geq 80\%$, and $\times$ means NUIT-2 fails.

Victim Speaker \ Victim Microphone		Siri			Google Phone Assistant				Alexa	Google Assistant	Cortana	
		iPhone XR	MacBook Pro-2021	Apple Watch 2	Google Pixel 3	Galaxy S9	LG Think Q V35	Galaxy Tab S4	Echo Dot Gen 1	Google Home 2	Dell Inspiron 15	MS Surface
Apple Devices	iPhone XR	3	3	3	4	6	50	5	6	7	6	8
	MacBook Pro	9	8	10	20	25	130	20	30	25	310	320
	iPhone13 mini	3	3	3	4	6	50	5	5	7	6	8
	iMac 27' 2021	13	12	15	13	30	390	20	50	60	370	350
Android Devices	LG Think Q V35	×	×	×	×	×	×	×	×	×	×	×
	Samsung Galaxy S9	4	4	4	6	4	60	6	7	5	7	7
	Samsung Galaxy Tab S4	9	9	10	27	20	150	20	40	50	25	30
Vehicle Audio Sys.	Ford Fusion 2017	30	28	35	102	82	320	70	210	230	160	140
	Nissan Versa S	×	×	×	110	70	300	65	190	220	150	150
Smart Home Devices	Samsung TV	35	32	46	120	80	460	90	350	320	150	100
	Google Home2	3	2	2	15	25	380	27	38	39	58	60
	Echo Dot Gen1	2	1	1	17	29	320	26	42	33	62	69
Windows Laptop	Dell Inspiron15	×	×	×	25	20	300	25	90	100	50	45

Other Experiment

- Impact of Different Languages.
- Impact of Audio Format.
- Impact of Background Noise
- Impact of Directionality

Defense

Basic Idea:
Verify if the baseband component is shadowed from near-ultrasound frequency component (>16kHz)

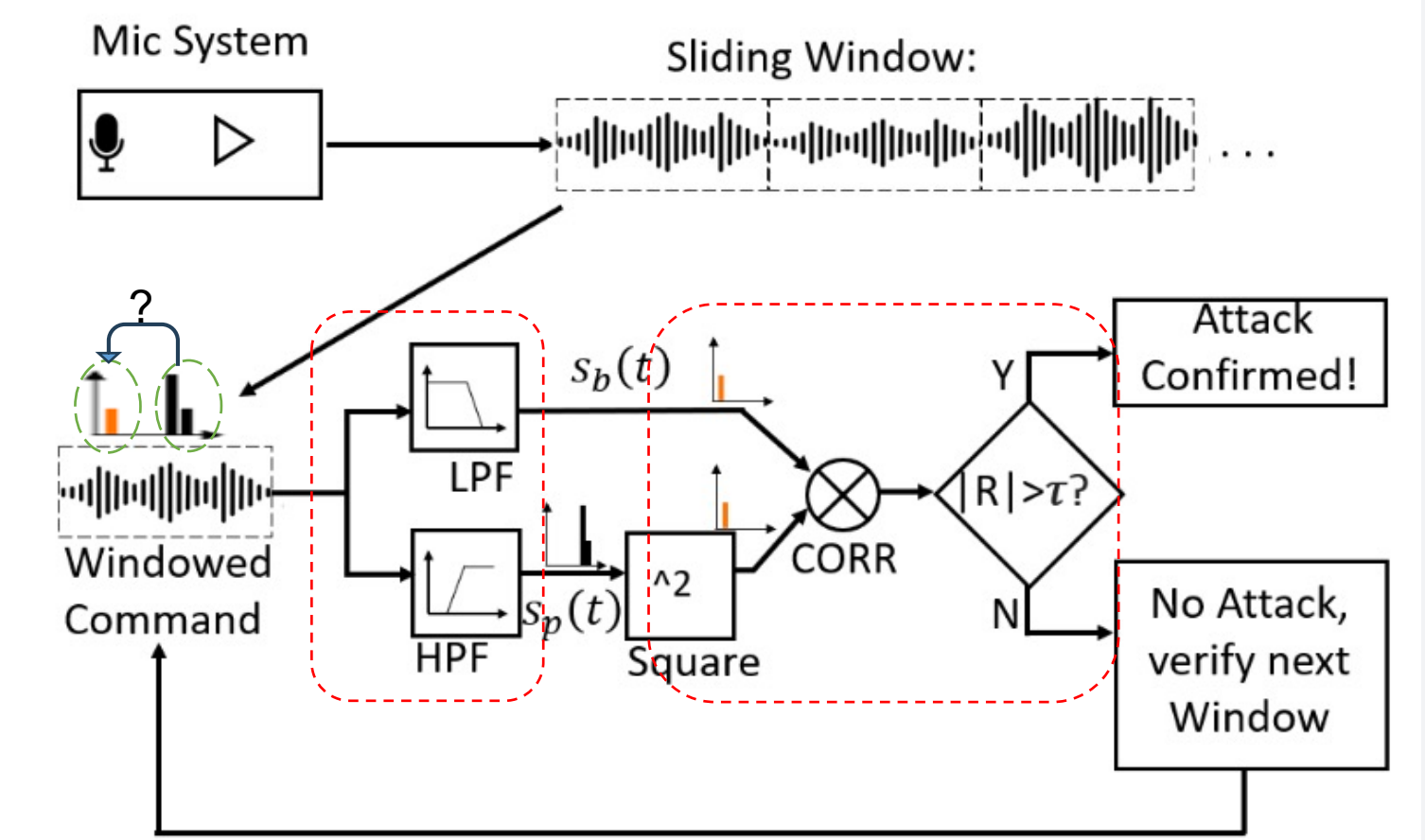


Figure 8: Basic idea for detecting NUIT.

Limitations

- Near Ultrasound is audible to some young kids
- NUIT-1: End-to-end unnoticeability can be achieved by Siri devices only (silent response)
- NUIT-2: The attack distance is short if attacker exploits mobile device's speaker to launch attack.

Conclusion

- NUIT is an attack against VCS, that is both remote and inaudible
- Two instances: NUIT-1 and NUIT-2
  - NUIT-1: Exploit a Speaker to attack the Microphone on same device
  - NUIT-2: Exploit a Speaker to attack the Microphone on a different device
- NUIT is achieved by using SSB-AM modulate to overcome audible leakage

Thank you!

Contact us:
lot.ece@utsa.edu