浙江大学
ZHEJIANG UNIVERSITY

DolphinAttack: Inaudible Voice Commands

Guoming Zhang Chen Yan Xiaoyu Ji

Tianchen Zhang Taimin Zhang Wenyuan Xu

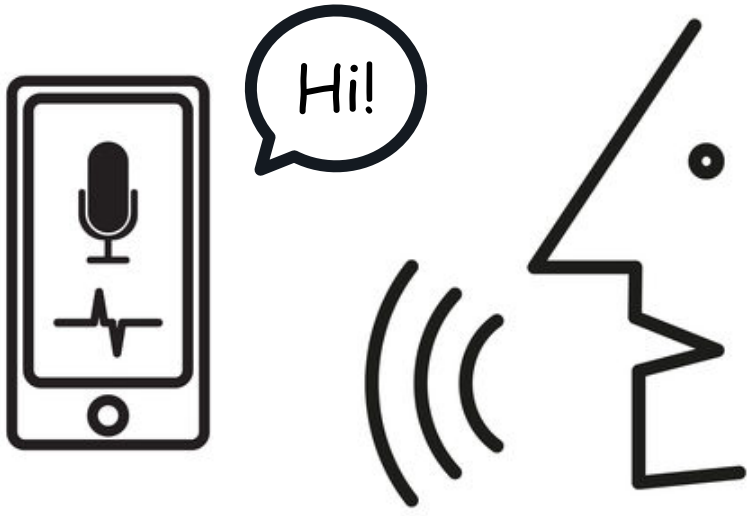
USSLAB, Zhejiang University

智能系统安全实验室
UBIQUITOUS SYSTEM SECURITY LAB.

Outline

1. Background of Voice Assistants
2. Design of DolphinAttack
3. Attack Scenarios
4. Evaluation
5. Defense & Responsible Disclosure

Voice becomes an increasingly important interface

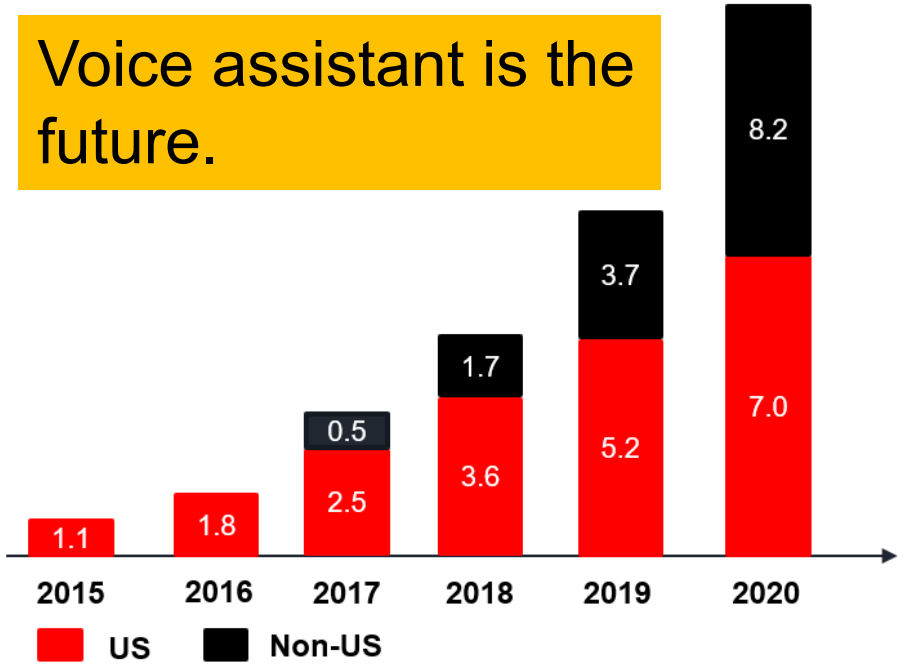


Siri Google Now Alexa Cortana S Voice Hi Voice

Digital Voice-activated Assistant Device Shipments Worldwide, US vs. Non-us, 2015-2020

millions

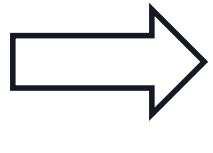
Voice assistant is the future.

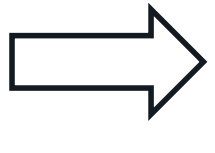


Source: Strategy Analytics as cited in press release, Aug,26,2016

How do voice assistants work?

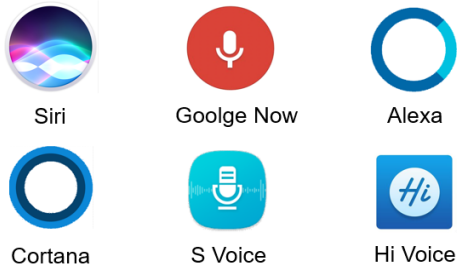
Sound *Electrical signal* *Command*





Voice Command Input

Speech Recognition



Command Execution

What can voice assistants do?

What can a malicious user achieve?

What's on my calendar today?

Sensitive information

Open evil.com

Malicious website

Tell my wife I love her

Fake message

Send an email to my boss

Social engineering

Open the front door

Break-in

Buy something on Amazon

Lose money

Transfer $100 to Eve

Steal money

Call 1234567890

Eavesdrop

Facetime Eve

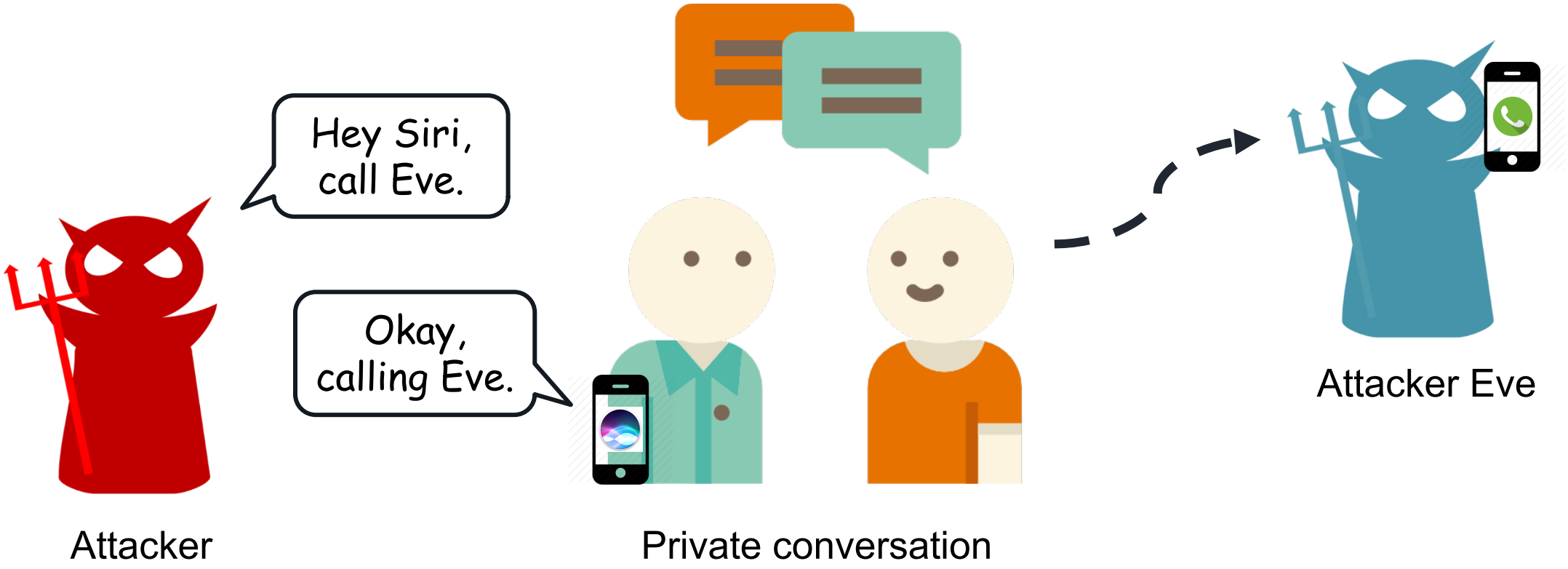
Spy

Drive me to Austin

Mislead

Attack Scenario 1: fake online orders

Attack Scenario 2: spying phone/video calls



Attack Scenario 3: exposing user privacy

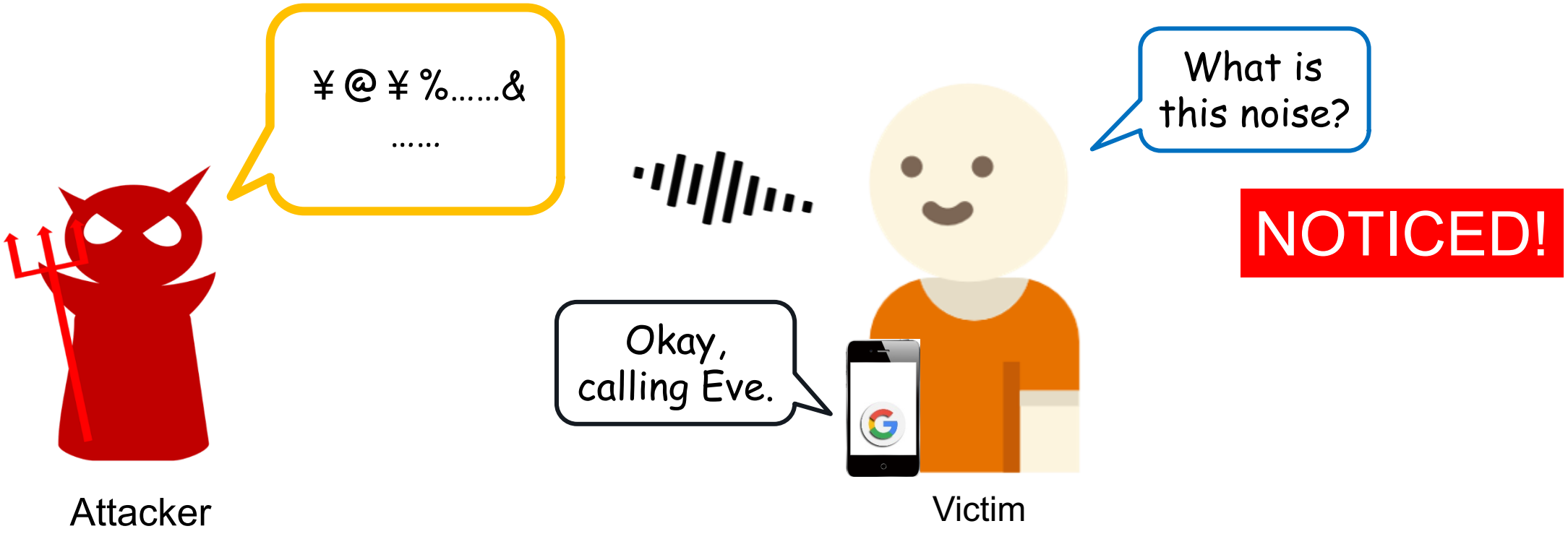
Related Work

The attacking commands are **audible**, and can be **noticed**!

Related Work

Vaidya et al., **Cocaine Noodles** (WOOT 2015)
Carlin et al., **Hidden Voice Commands** (Usenix Security 2016)

The attacking commands are still **audible.**



DolphinAttack

ATTACKED DEVICE : AMAZON ECHO

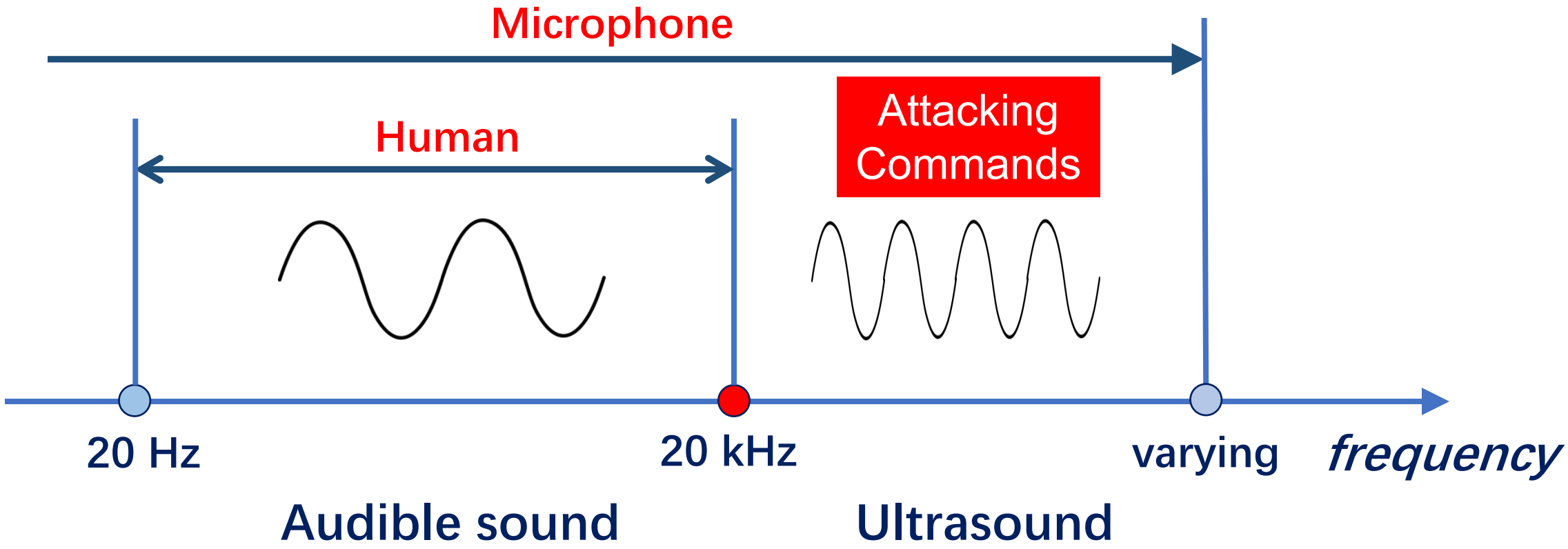
Attack Scenario

- Order stuff
- Make a call
- Read to-do list
- Open the door


Scenario 1: Shopping

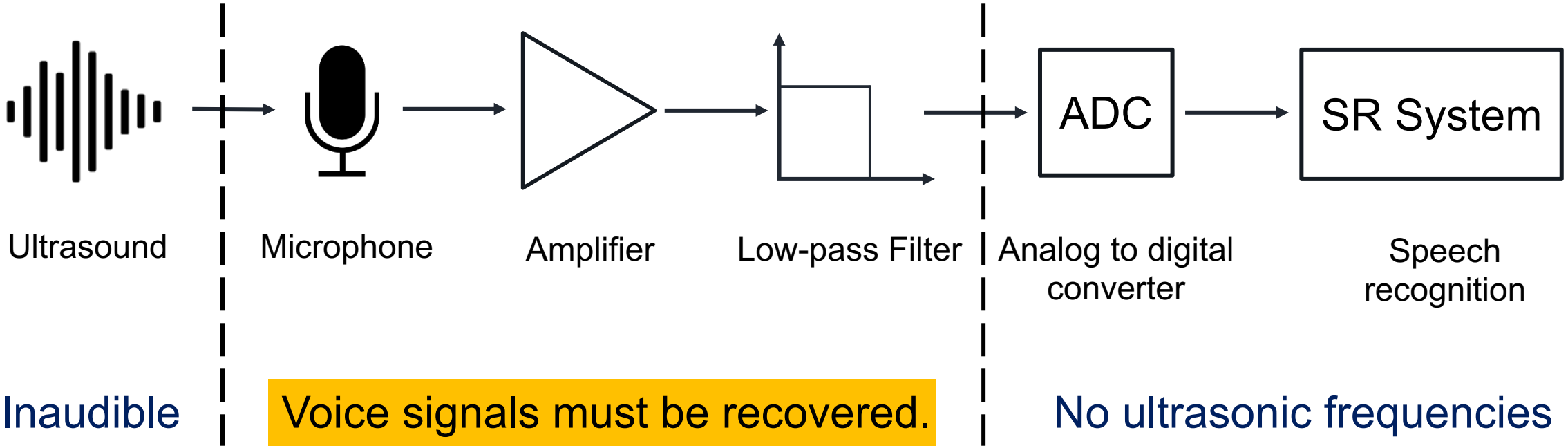
Hearing Range of Human and Microphone

Speech recognition systems only accept signals of **audible sound**.

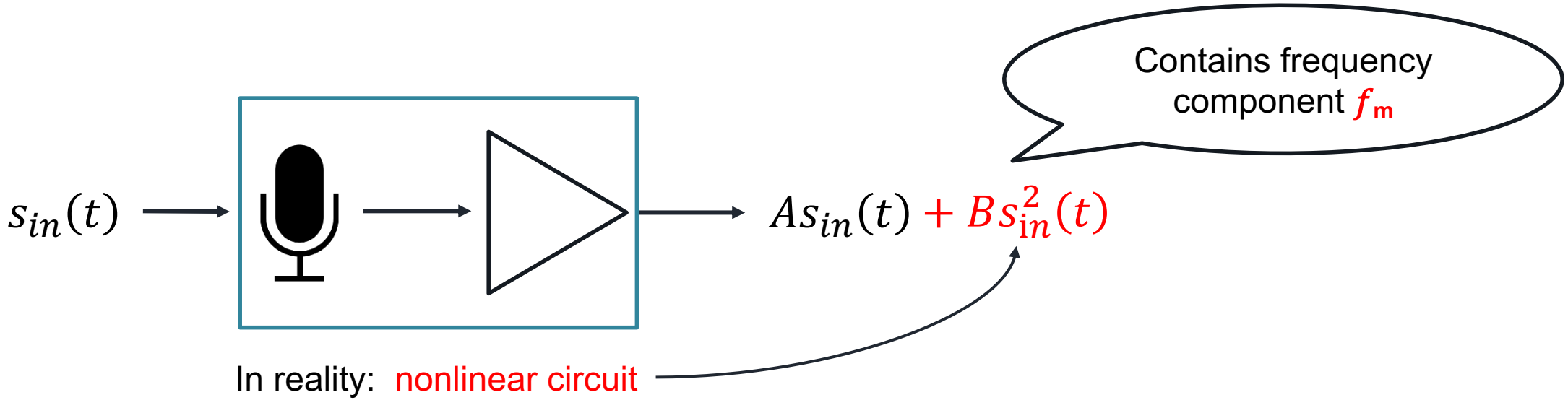


How can voice assistants accept ultrasound?

- The low-pass filter will remove ultrasonic frequencies to avoid aliasing.



Exploiting the Nonlinearity of Microphone

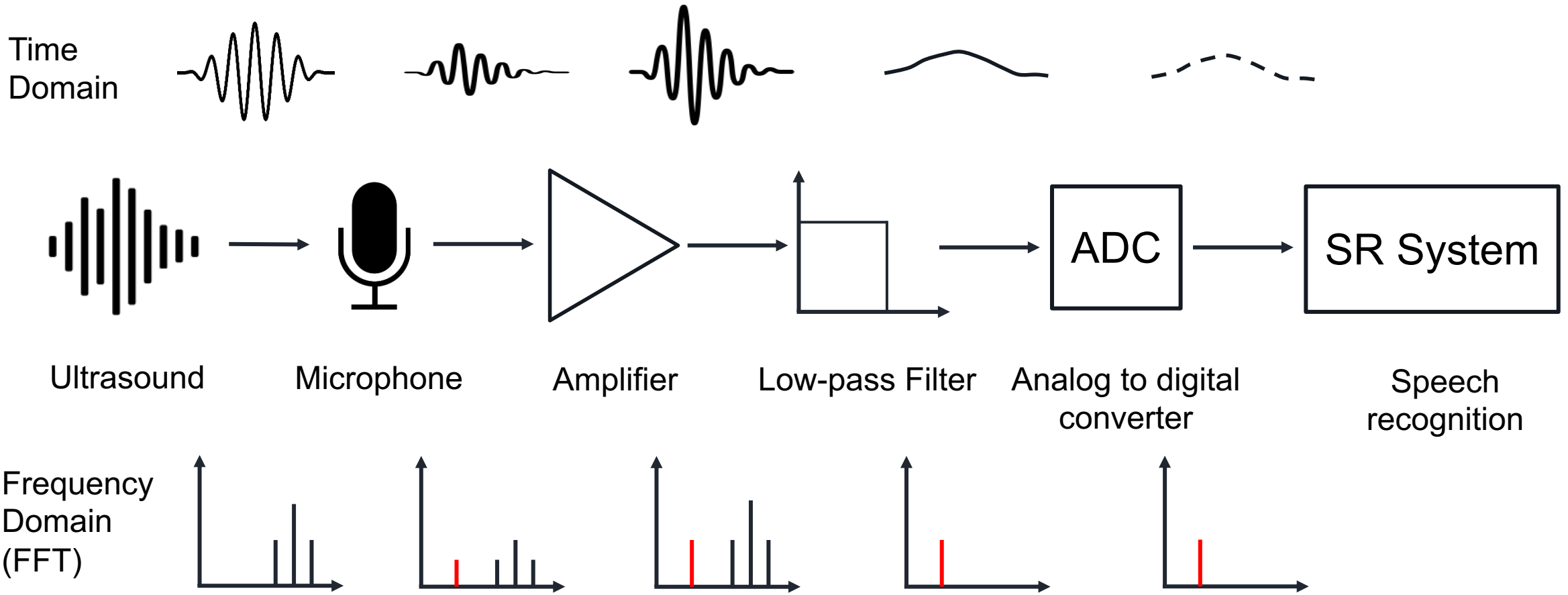


Let input be $s_{in}(t) = m(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$

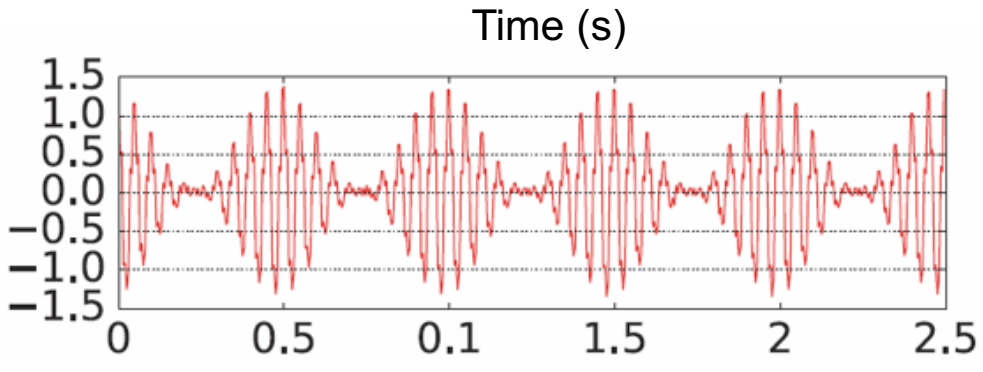
Where $m(t)$ is a baseband voice signal, $m(t) = \cos(2\pi f_m t)$

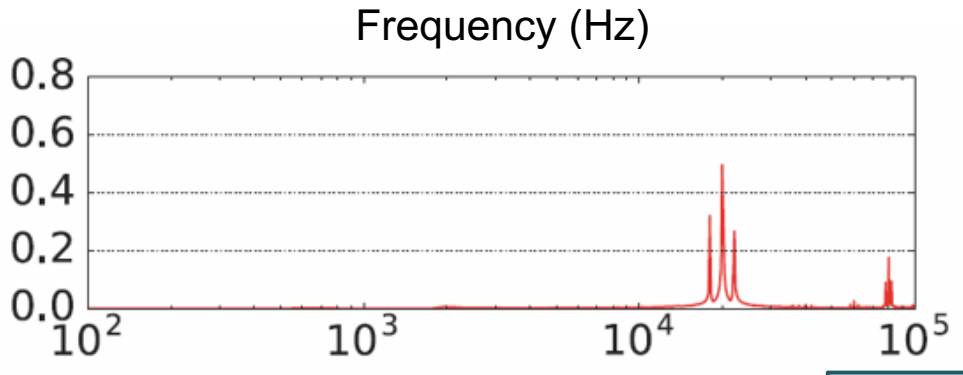
The baseband voice signals can be demodulated by microphones.

Signal Flow of DolphinAttack



Nonlinearity Effect Validation

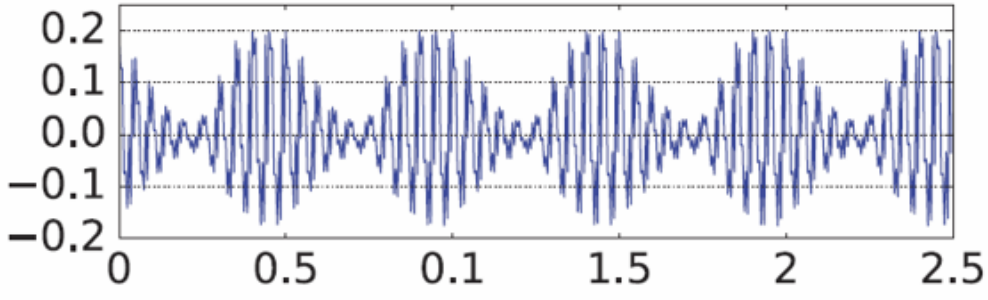


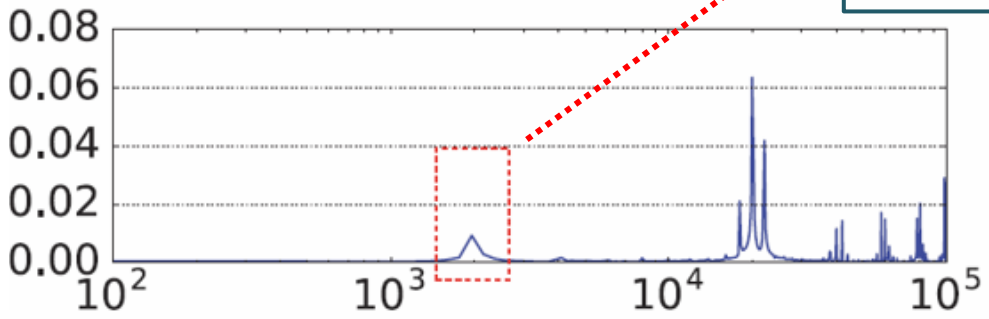


$f_c = 22$ kHz, $f_m = 2$ kHz

Signals of DolphinAttack

Nonlinearity Effect

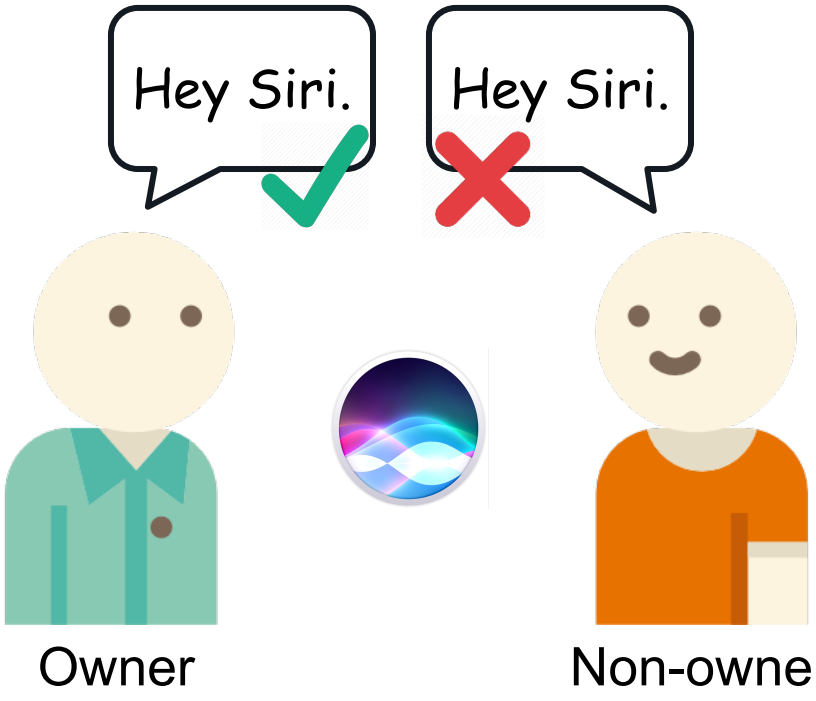




Signals received by a MEMS microphone

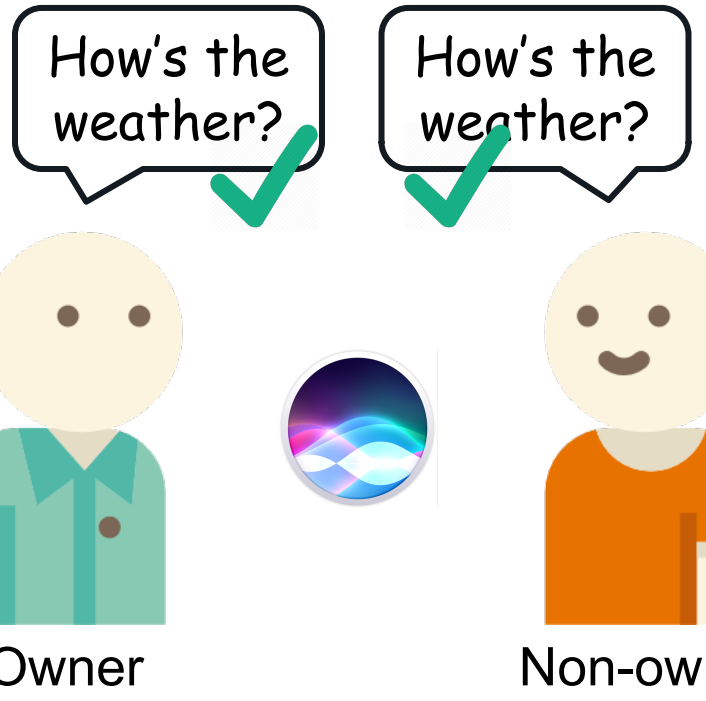
Speaker Dependent vs Speaker Independent

Activation Commands



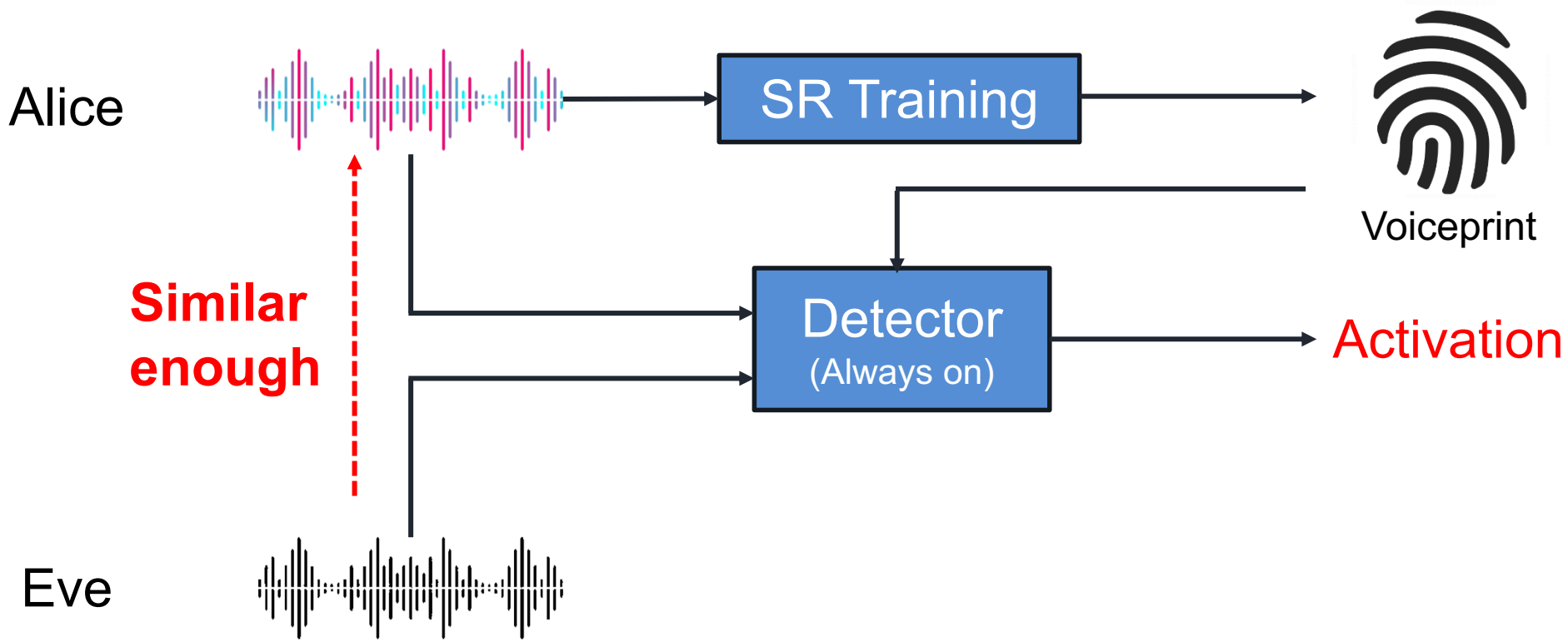
+

Control Commands

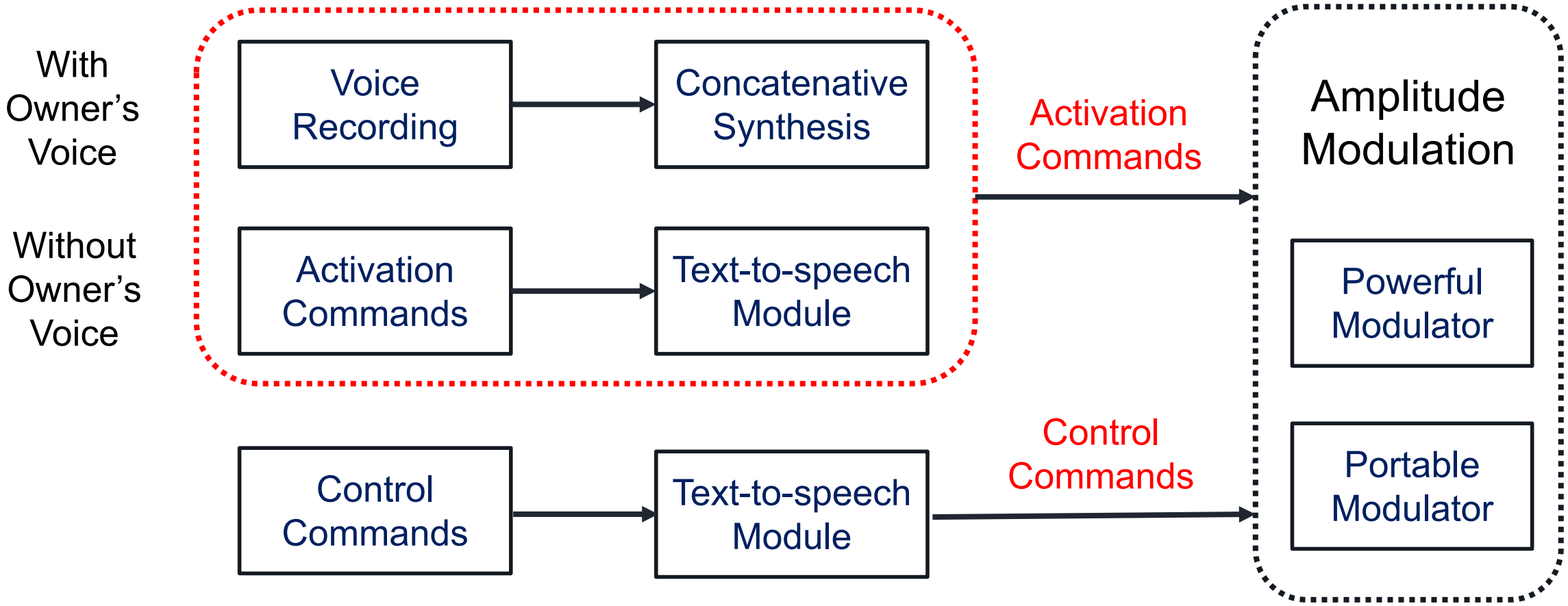


Both **activation** and **control** commands are required for DolphinAttack.

Speaker Dependent SR – Activation



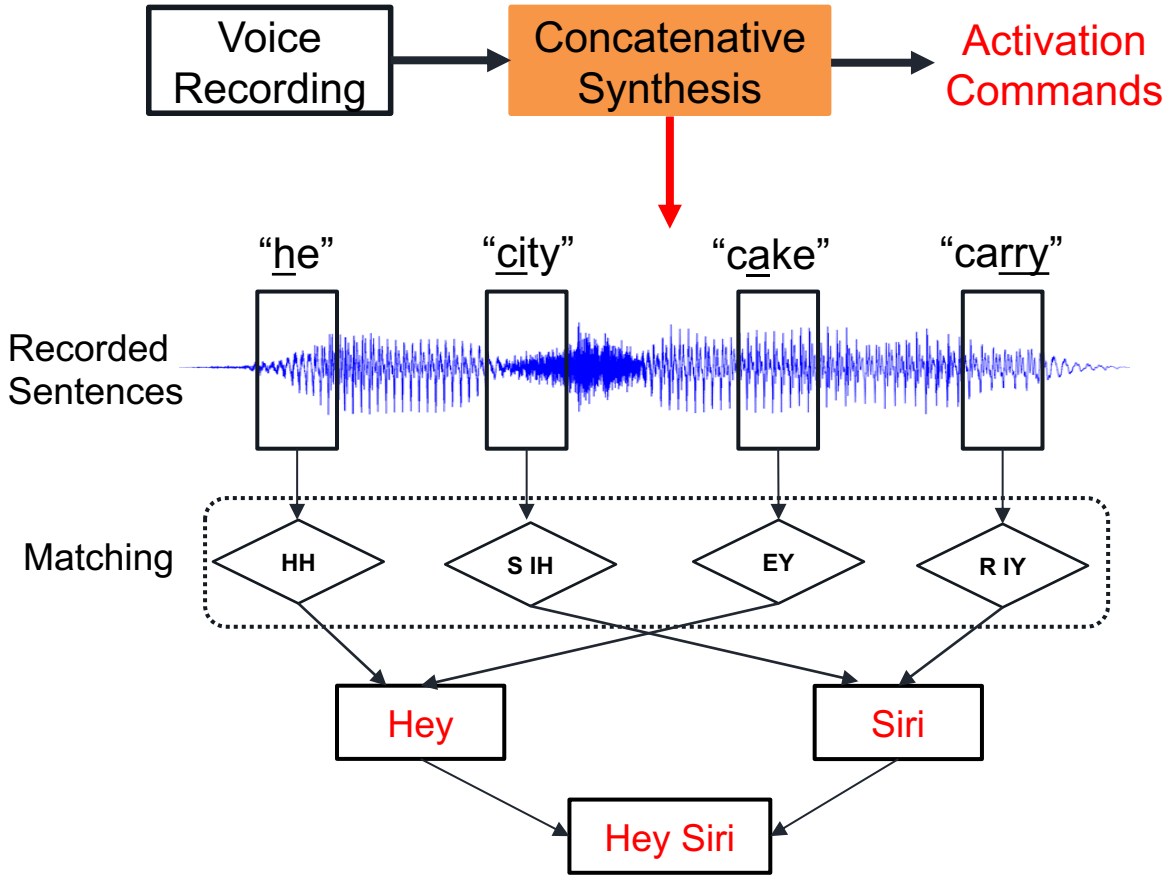
Design of DolphinAttack



1. Concatenative Synthesis – with owner's voice

- 44 phonemes in English.
- "Hey Siri" includes 6 of them

  (i.e., HH, EY, S, IH, R, IY).
- Synthesize a desired activation command by searching for relevant phonemes from other words in available recordings.



Concatenative synthesis of "Hey Siri"

2. TTS-based Approach – without owner's voice

TTS: Text to Speech

Observation

- Two users with similar vocal tones can activate the other's Siri.

Method

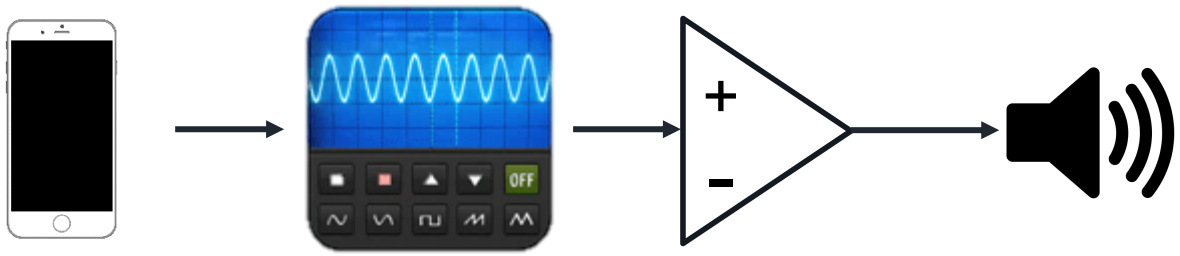
- 35 out of 89 TTS systems can successfully activate a trained Siri.

TTS Systems	voice type #	# of successful types	
		Call 12..90	Hey Siri
Selvy Speech [51]	4	4	2
Baidu [8]	1	1	0
Sestek [45]	7	7	2
NeoSpeech [39]	8	8	2
Innoetics [59]	12	12	7
Vocalware [63]	15	15	8
CereProc [12]	22	22	9
Acapela [22]	13	13	1
Fromtexttospeech [58]	7	7	4

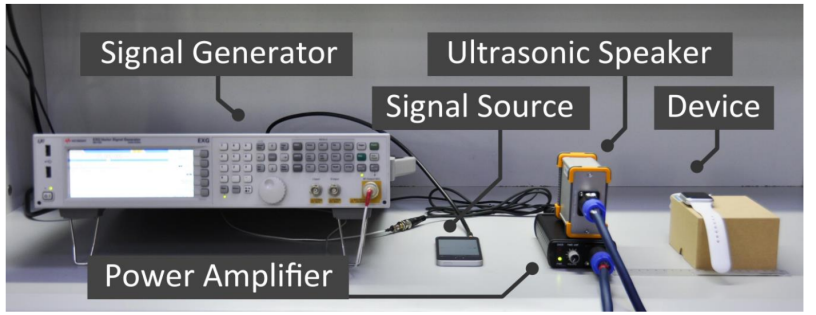
The list of TTS systems used for attacking the Siri trained by the Google TTS system, and the evaluation results on activation and control commands.

Inaudible Voice Commands Transmitter

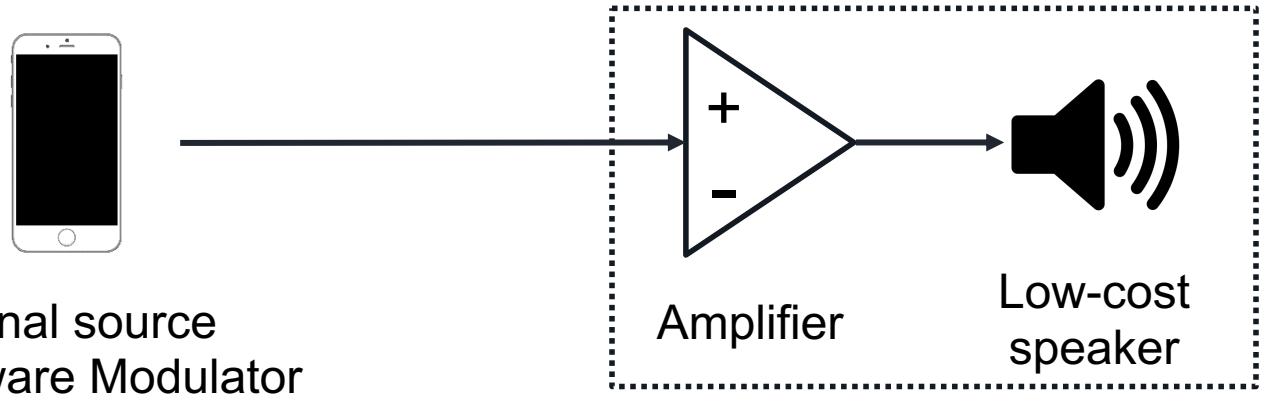
Rich man solution



Signal source | Signal generator (Hardware Modulator) | Amplifier | Professional speaker

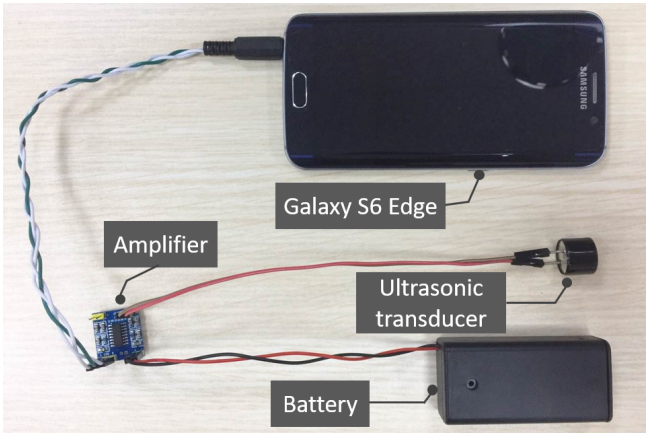


Poor man solution



Signal source & Software Modulator

Amplifier | Low-cost speaker

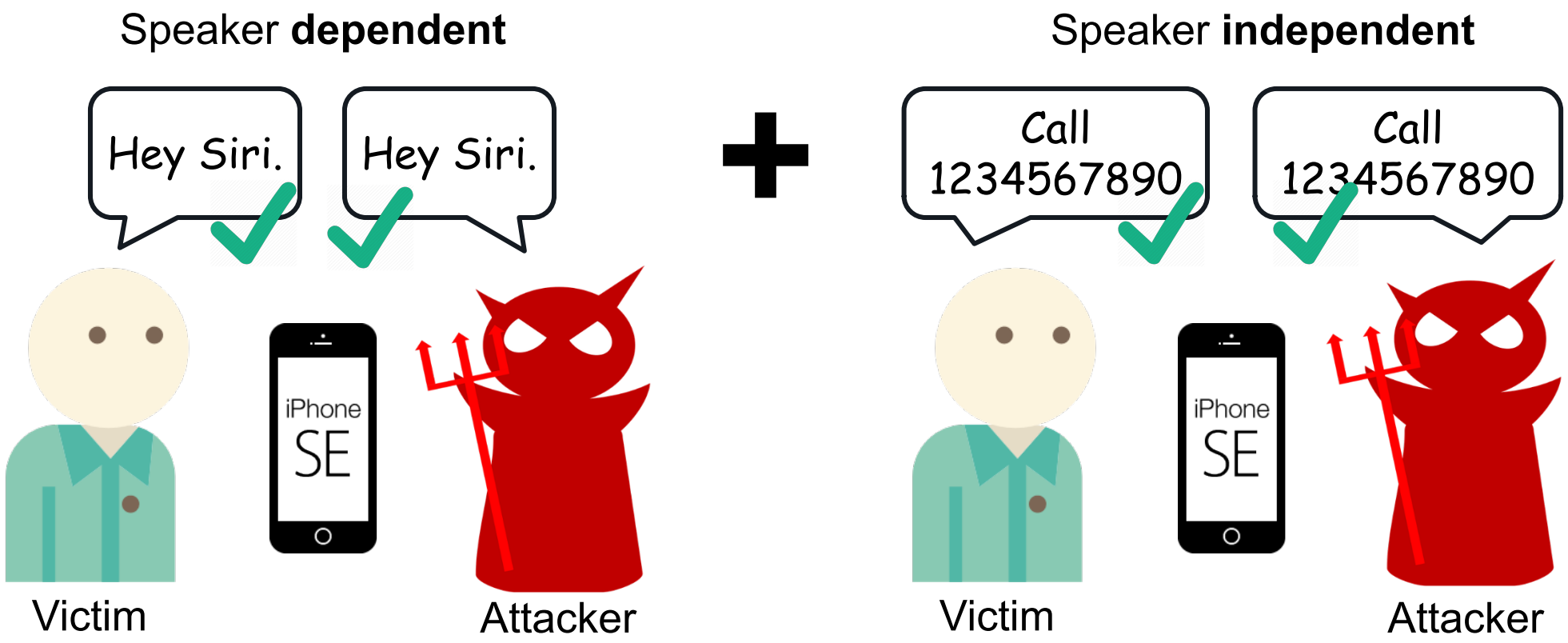


Less than $3

DolphinAttack

ATTACKED DEVICE : IPHONE SE

Attack Scenario: Make Spying Phone Call

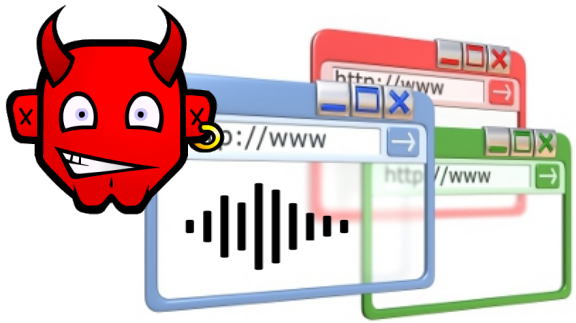


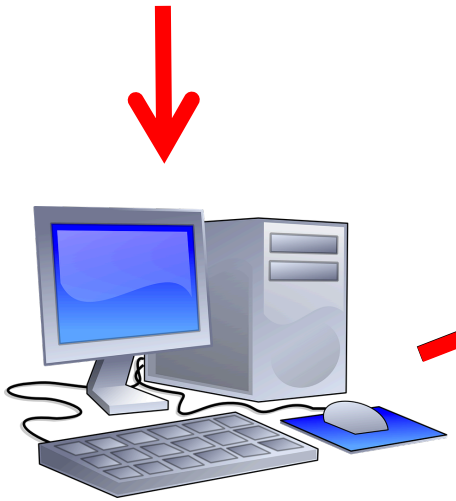

555
Activate Siri and make a phone call with a normal voice.

DolphinAttack

ATTACKED DEVICE: APPLE WATCH

Attack Scenario: Remote Attack





Computer

Commodity Speaker

Smart devices


"Facetime 1551072xxxx"
Under attack

DolphinAttack

COMPROMISED DEVICES

Manuf.	Model	OS/Ver.	SR System	Attacks		Modulation Parameters		Max Dist. (cm)	
				Recog.	Activ.	f_c (kHz) & [Prime f_c] ‡	Depth	Recog.	Activ.
Apple	iPhone 4s	iOS 9.3.5	Siri	√	√	20–42 [27.9]	≥ 9%	175	110
Apple	iPhone 5s	iOS 10.0.2	Siri	√	√	24.1 26.2 27 29.3 [24.1]	100%	7.5	10
Apple	iPhone SE	iOS 10.3.1	Siri	√	√	22–28 33 [22.6]	≥ 47%	30	25
			Chrome	√	N/A	22–26 28 [22.6]	≥ 37%	16	N/A
Apple	iPhone SE †	iOS 10.3.2	Siri	√	√	21–29 31 33 [22.4]	≥ 43%	21	24
Apple	iPhone 6s *	iOS 10.2.1	Siri	√	√	26 [26]	100%	4	12
Apple	iPhone 6 Plus *	iOS 10.3.1	Siri	×	√	— [24]	—	—	2
Apple	iPhone 7 Plus *	iOS 10.3.1	Siri	√	√	21 24-29 [25.3]	≥ 50%	18	12
Apple	watch	watchOS 3.1	Siri	√	√	20–37 [22.3]	≥ 5%	111	164
Apple	iPad mini 4	iOS 10.2.1	Siri	√	√	22–40 [28.8]	≥ 25%	91.6	50.5
Apple	MacBook	macOS Sierra	Siri	√	N/A	20-22 24-25 27-37 39 [22.8]	≥ 76%	31	N/A
LG	Nexus 5X	Android 7.1.1	Google Now	√	√	30.7 [30.7]	100%	6	11
Asus	Nexus 7	Android 6.0.1	Google Now	√	√	24–39 [24.1]	≥ 5%	88	87
Samsung	Galaxy S6 edge	Android 6.0.1	S Voice	√	√	20–38 [28.4]	≥ 17%	36.1	56.2
Huawei	Honor 7	Android 6.0	HiVoice	√	√	29–37 [29.5]	≥ 17%	13	14
Lenovo	ThinkPad T440p	Windows 10	Cortana	√	√	23.4–29 [23.6]	≥ 35%	58	8
Amazon	Echo *	5589	Alexa	√	√	20-21 23-31 33-34 [24]	≥ 20%	165	165
Audi	Q3	N/A	N/A	√	N/A	21–23 [22]	100%	10	N/A

‡ Prime f_c is the carrier wave frequency that exhibits highest baseband amplitude after demodulation. — No result

† Another iPhone SE with identical technical spec.

* Experimented with the front/top microphones on devices.

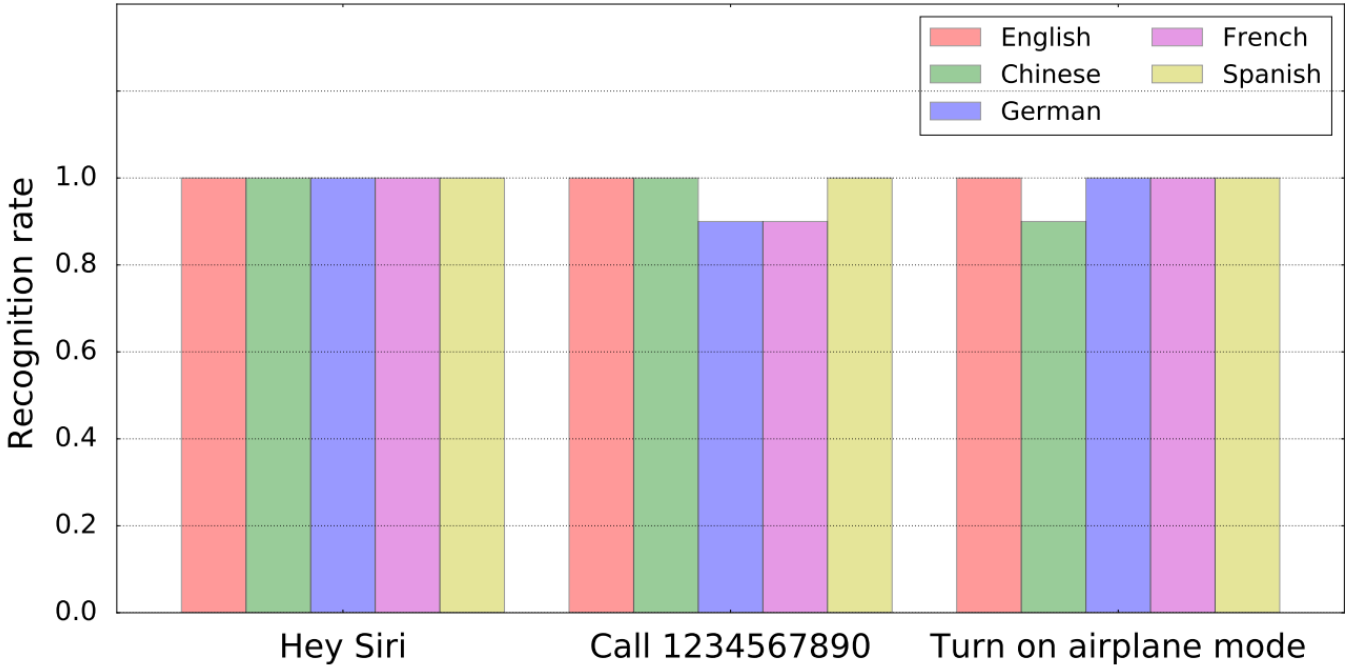
Evaluation

- Impact of languages
- Impact of attack distance
- Impact of background noise
- Impact of sound pressure level

Evaluation: Impact of Languages

DolphinAttack is effective for various languages and voice commands.

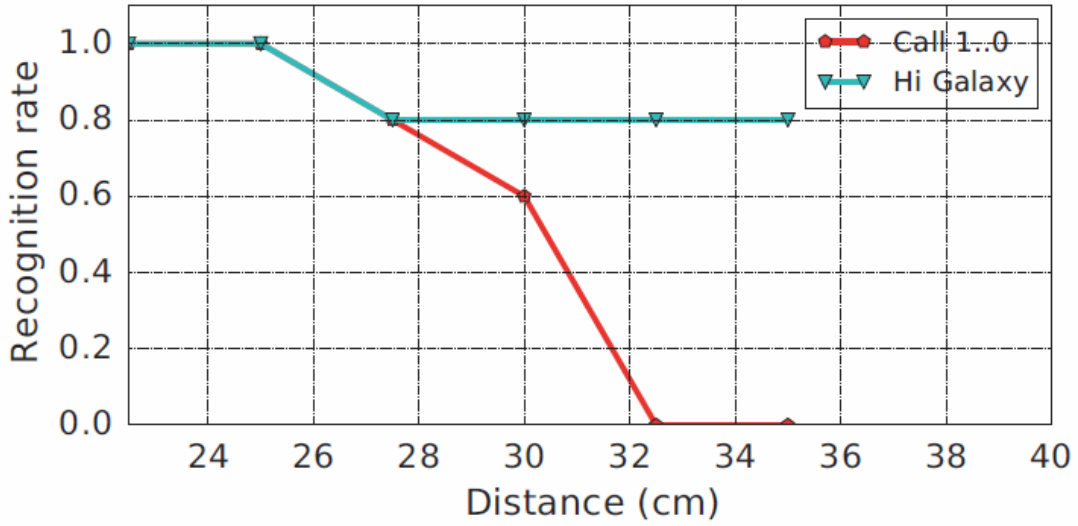
- English
- Chinese
- French
- German
- Spanish



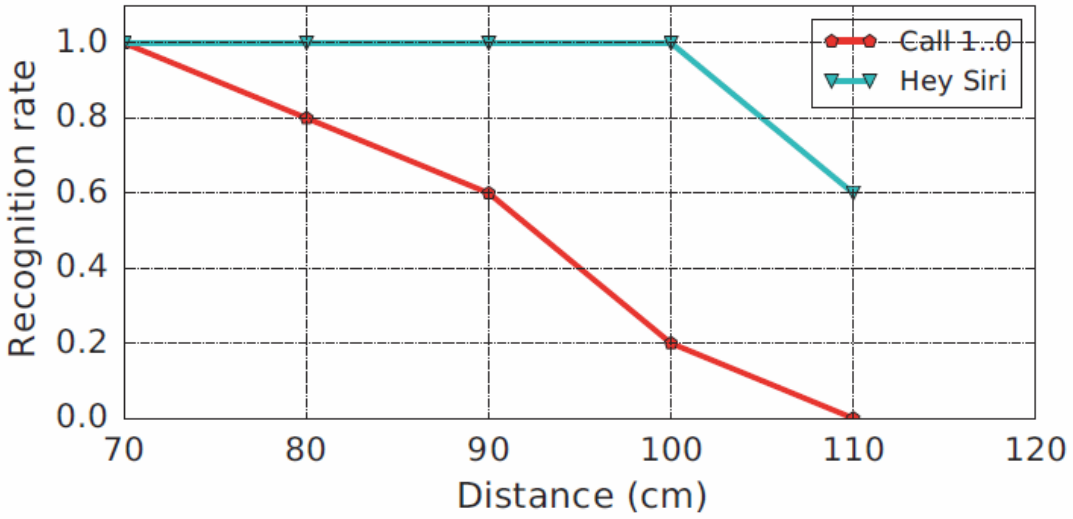
The recognition rates of voice commands in five languages

Evaluation: Impact of Attack Distance

The attack distance has fundamental impact on the effectiveness of DolphinAttack and is device dependent.



(a) The recognition rates of the Galaxy S6 Edge

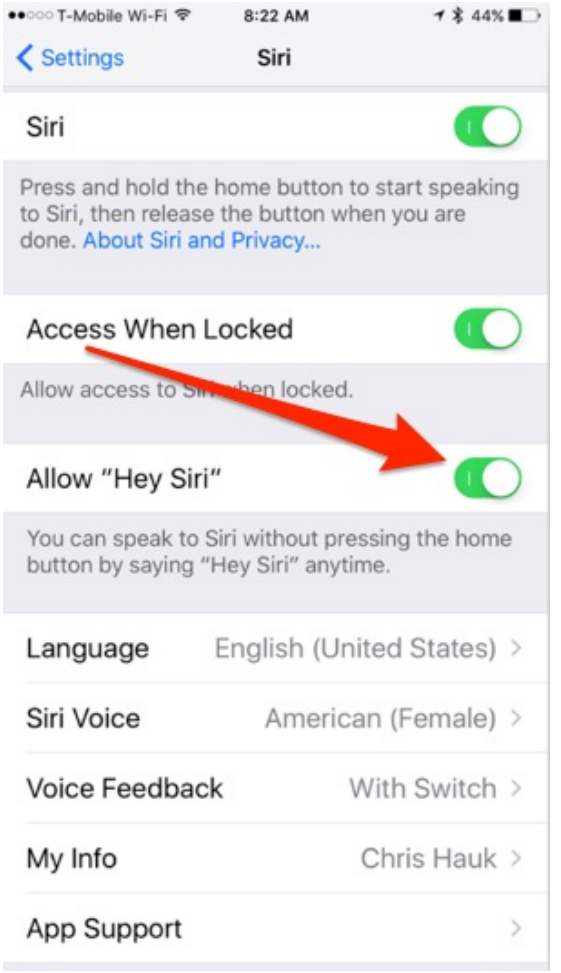


(b) The recognition rates of the Apple watch

The impact of attack distances on the recognition rates for S6 Edge and Apple watch.

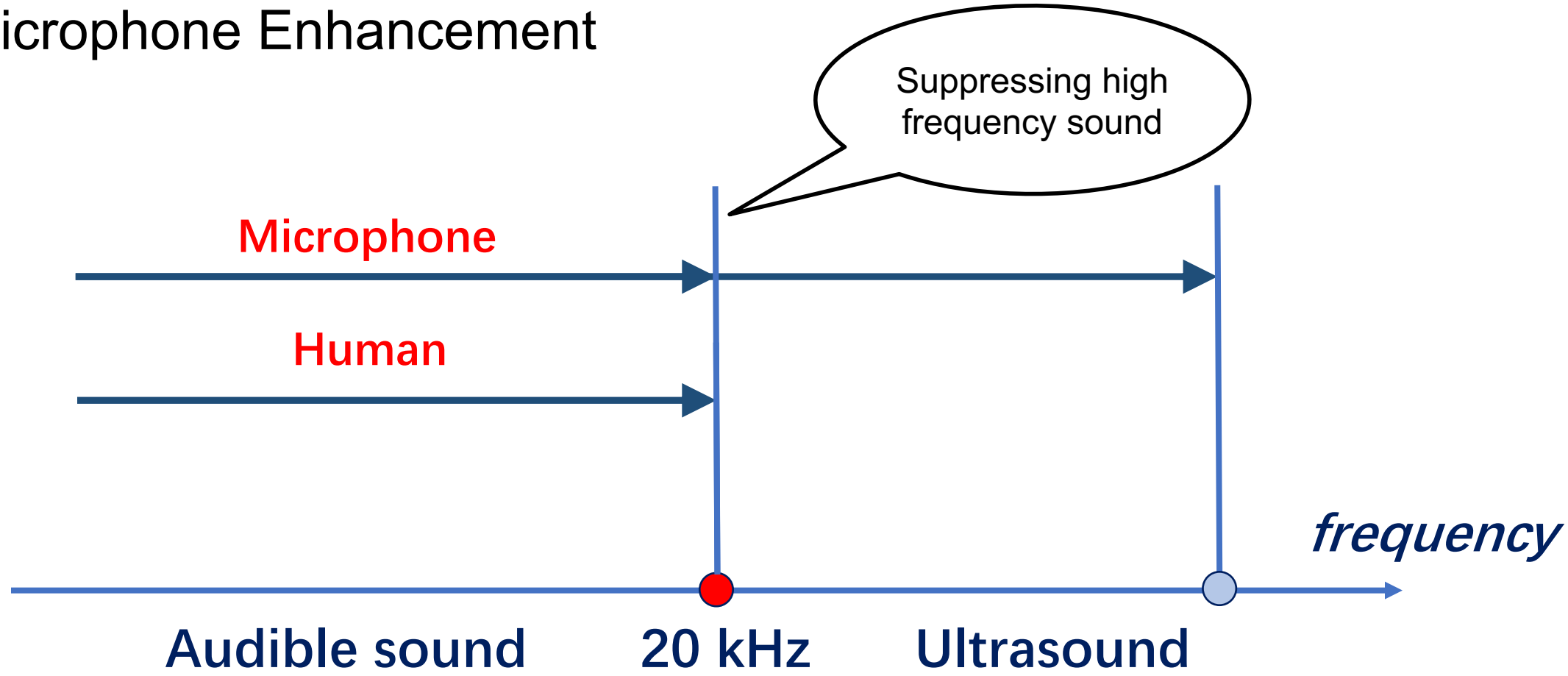
Defense


T-Mobile Wi-Fi 8:21 AM 44%
Settings
Do Not Disturb
General
Display & Brightness
Wallpaper
Sounds
Siri
Touch ID & Passcode
Battery
Privacy
iTunes & App Store
Wallet & Apple Pay


T-Mobile Wi-Fi 8:22 AM 44%
Settings
Siri
Siri
Press and hold the home button to start speaking to Siri, then release the button when you are done. About Siri and Privacy...
Access When Locked
Allow access to Siri when locked.
Allow "Hey Siri"
You can speak to Siri without pressing the home button by saying "Hey Siri" anytime.
Language English (United States)
Siri Voice American (Female)
Voice Feedback With Switch
My Info Chris Hauk
App Support

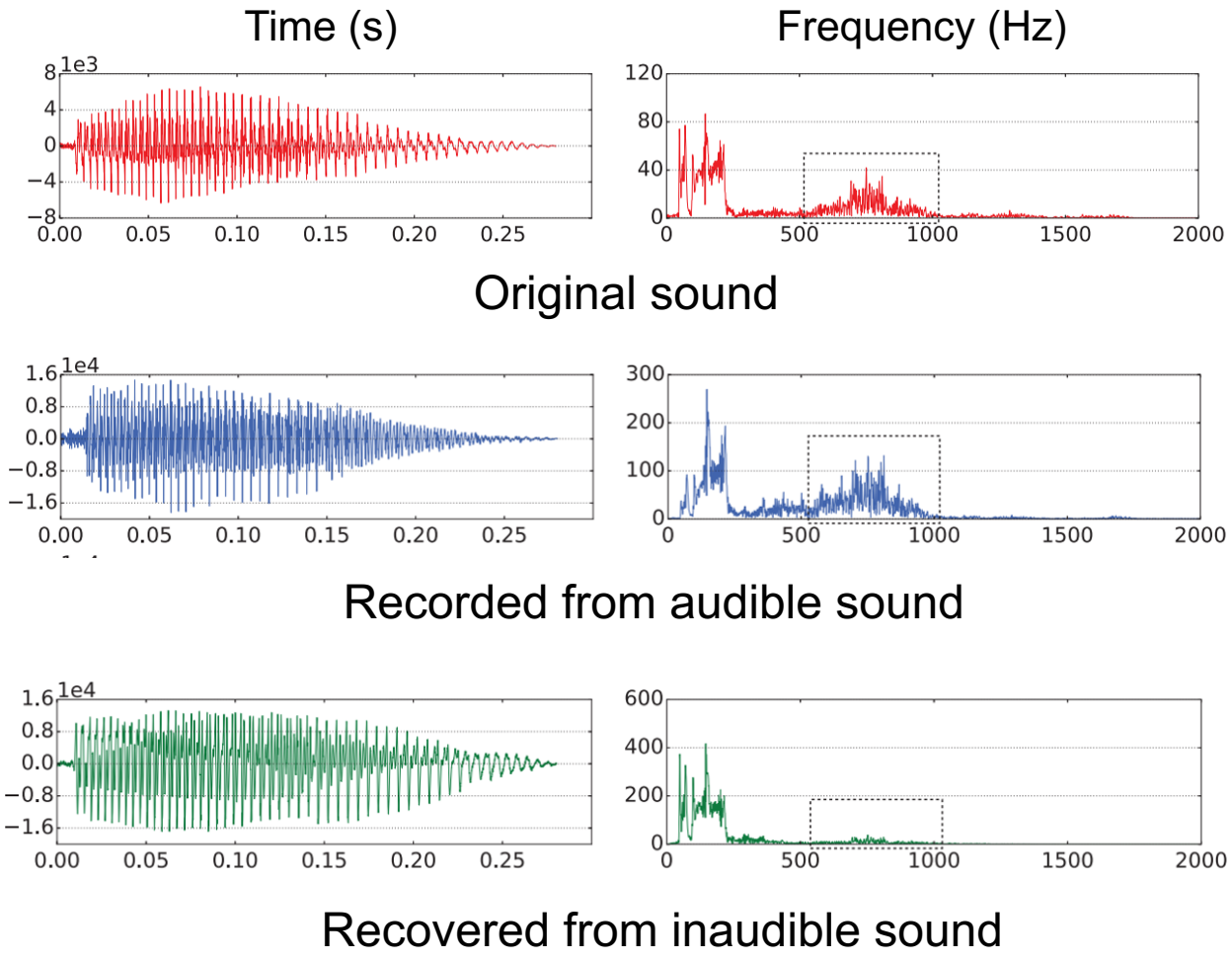
Hardware-Based Defense

- Microphone Enhancement



Software-Based Defense

- Modulated voice commands are distinctive from genuine ones.
- Supported vector machine (SVM) as the classifier to detect the malicious command from the normal command.
- Result: 100% true positive rate (7/7) and 100% true negative rate (7/7).



Responsible Disclosure

- We have contacted the product security team at Apple, Amazon, Google, Android, Huawei, and Samsung, and received their feedback.

Hello Chen,

Thank you for sharing an advance copy of your paper. We are reviewi ng it and will provide you with our f eedback.

Best regards,
Deven
Apple Product Security

Chen,
Thank you for your report. We are investigating your finding and have assigned you case ID RM029916357. Please be sure to reference that number if you have any follow up questions or want to provide additional information related to your finding. Amazon takes security very seriously and we always appreciate it when researchers work with us to improve our product security.

Kind Regards,
Ryan

sh...@google.com added comment #10:
Thanks again for this report. After investigation by the Android Security team and the feature team we believe that this is something best addressed through hardware changes on the microphone in future devices.

Thanks,
Android Security Team

Hey,

thanks for reaching out. This sounds (or, rather, doesn't?) like a cool attack!

I recommend to report this as an Android bug instead. I'm honestly not sure if hardware issues are covered by this team, but they should at least know the correct point of contact. If they're not responsible either, feel free to circle back to us.

Daniel
Google Security Team

Dear Chen Yan,

We highly appreciate your concern about the security problems of Huawei products.

we have analyzed the information sent by you regarding potential security issues in Huawei products. In order to verify and address the mentioned problem, could you please provide in addition the following information to help us with verification:
1)the snapshot of software version in your Honor7, you can find it in the following menus: Settings --> About phone.
2)Detailed steps about DolphinAttack.
3)As you mentioned in your email, it is a universal issue. Did you tested other products, and have you reported it to other vendors, like: Google.
4)The paper you reported to ACM CCS 2017.
5)If CERT assign you a track id, could you plerase share it to us?

Thank you and best regards,
Huawei PSIRT

Dear Chen Yan,

We would like to thank you for sharing a potential security issue for Samsung mobile device.

We are looking into the issue you shared, and we want to get any sample sound or Proof of Concept to verify and analyze.

We also want to ask you to share your slides for the ACM CCS conference prior to submittal, so that we can enhance and secure our product.

Thank you.

NOTE: Please note that we may ask you to report it a different channel if our analysis concludes that there is no security impact.

Very Respectfully,
Samsung Mobile Security Technologies

Summary

- Voice assistant has become an increasingly popular human-computer interaction mechanism, but they are vulnerable to attacks.
- DolphinAttack is a totally inaudible attack from a new perspective, could attack Siri, Alexa, Google Now, Cortana, Samsung S Voice, Huawei Hi Voice.
- To avoid the abuse of DolphinAttack in reality, we propose two defense solutions from the aspects of both hardware and software.

Questions

DolphinAttack Homepage: http://dolphinattack.com/

USS Lab Homepage: http://usslab.org/