CSE227 – Graduate Computer Security

Side Channels Pt. 2 + IoT Devices

UC San Diego

Housekeeping

General course things to know

- Course projects
  - I will provide some initial thoughts and feedback on each of your project ideas by **this Friday** via email — really enjoyed them so far
  - Start meeting with your teams, ideating, and reaching out to me if you have things you want to chat about!

Getting started

Week 3 advice on projects

- For this week, if you're looking for something to "do…"
- Consume a bunch of papers / blog posts / videos / etc. related to what you're looking at. Spend a few hours investing yourself in the research world you're planning on entering
- Meet with your group, at least twice, and discuss the ideas you're finding. How do they help you think more clearly about your own project?
- Start to outline what you need for the project to be successful.
  - Data? Code? An algorithm?
  - Start collecting what you need — look for anything that can help you

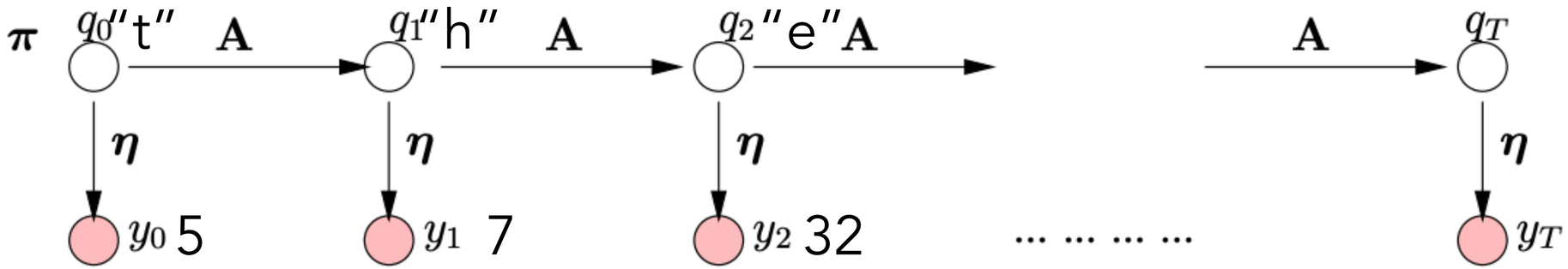
Today's lecture

Learning Objectives

- Finish up keyboard emanations discussion
- Discuss the "Cold boot" attack on DRAM
- Learn about IoT devices, voice interfaces, and how those interfaces fall apart in practice
- Discuss "Skill Squatting Attacks" paper

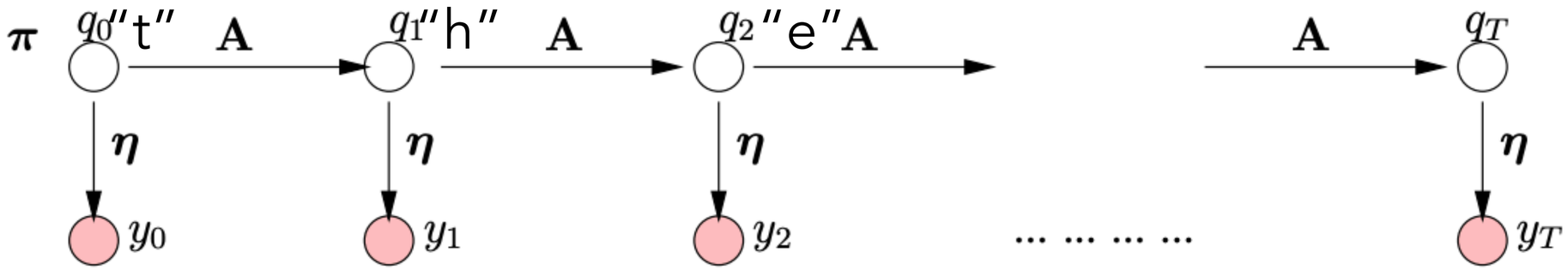
Keyboard Acoustic Emanations

Bi-grams of characters



- What do the circles represent in the model? What are shaded and unshaded?
- What is *q*? ***Characters pressed***
- What is *y*? ***Cluster labels***

Bi-grams of characters



- What is **A**, otherwise known as the *transition matrix?*
- How do we populate **A?**
- What is *eta*?
- How do the authors populate *eta*?

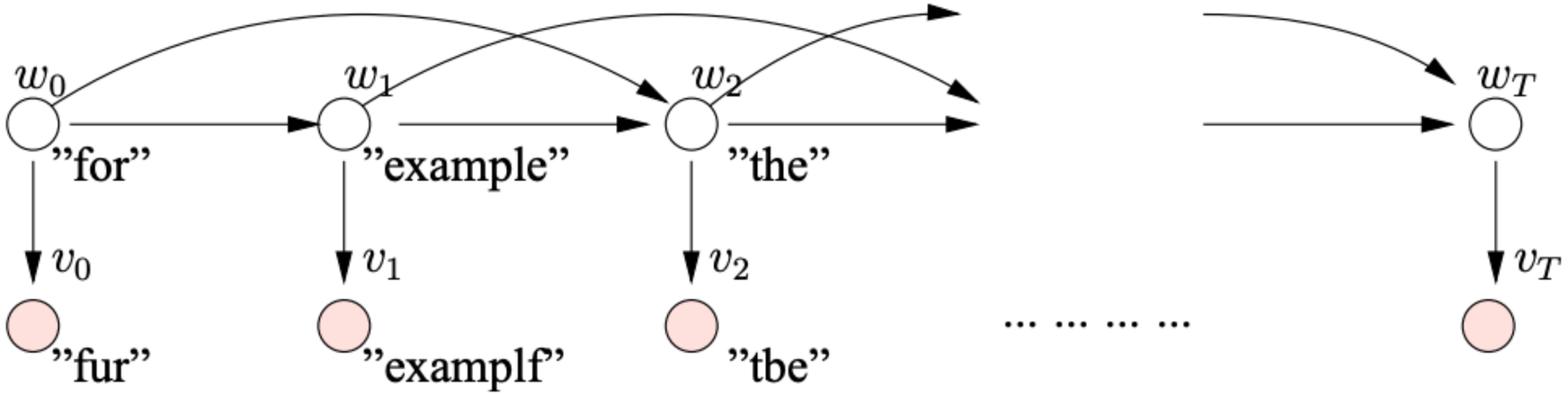
So now we have characters…. are they right?

- Authors used *spellcheck* to try and make the text more readable, but it still made mistakes
  - e.g., “fur example” vs. “for example”
- Can get more readable text using an n-gram language model
  - What’s an n-gram language model?

So now we have characters…. are they right?

- Authors used *spellcheck* to try and make the text more readable, but it still made mistakes
  - e.g., “fur example” vs. “for example”
- Can get more readable text using an n-gram language model
  - What’s an n-gram language model?
- n-gram: sequence of *n* adjacent items in text, speech, genomes, etc.

Tri-grams of words (HMMs to the rescue, again)



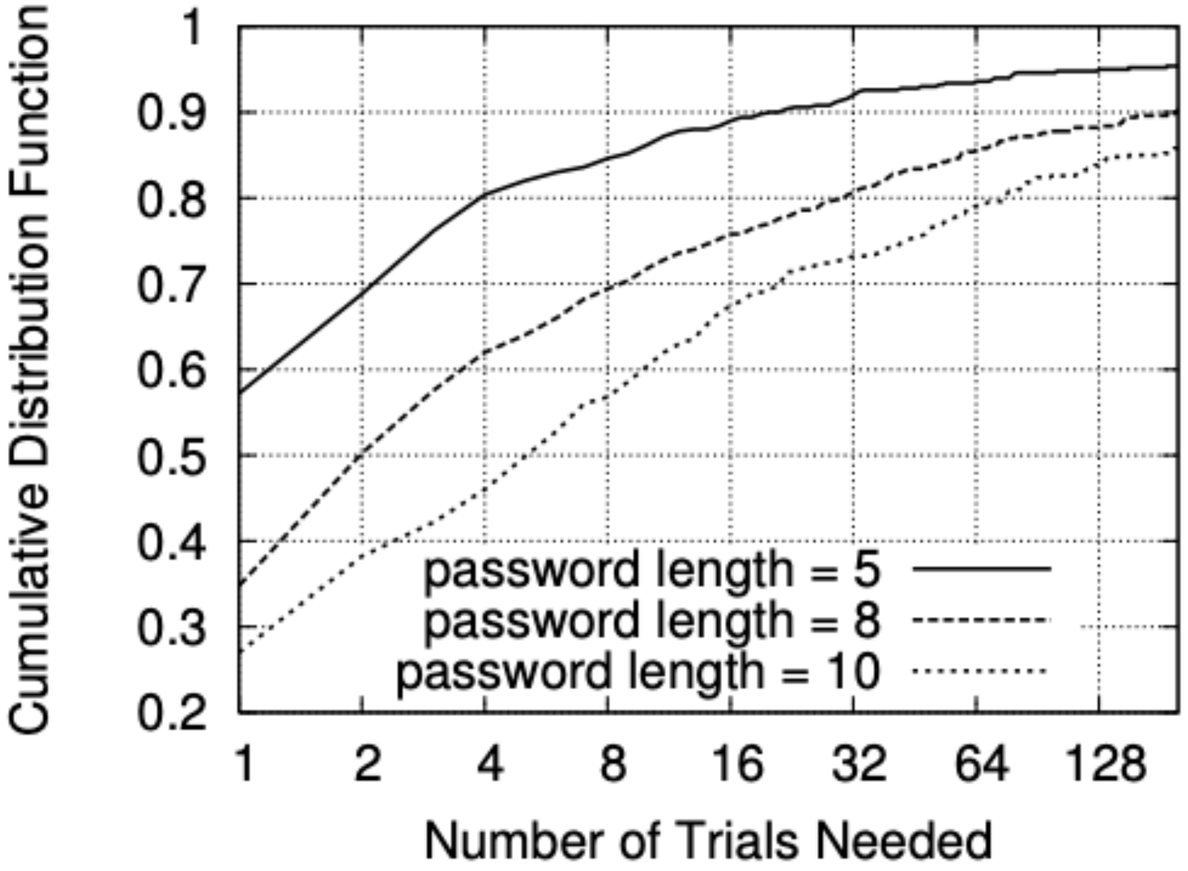
- Hidden variables are the original words
- This HMM depends on *two* layers (previous word **and** previous previous word)
- This is a great strategy if you have **unknowns** you want to predict from a known distribution!

Evaluation

- What was the evaluation setup for the attack?
- How many environments did the authors test their attack in?
- Are all keyboards vulnerable to this kind of attack? How did the authors evaluate this?

So it works on english, does it work on passwords?

- Yes!
- Authors found that they could recover 90% of 5-character passwords, 77% of 8-character passwords, and 69% of 10-character passwords
- Probabilities form a “hit list” of potential passwords to try



5-minute discussion: Meta points

- How feasible is this attack?
- Do you believe this attack will work in practice? Why or why not?
- What do we think about side channel research?

Side channels can be even crazier….

RSA Key Extraction via Low-Bandwidth Acoustic Cryptanalysis

Side channels can be even crazier….

LEVERAGING MACHINE LEARNING TO ENHANCE ACOUSTIC EAVESDROPPING ATTACKS (PART 1 OF 4)

October 14th, 2025 by Brian

This multi-part series explores how machine learning can enhance eavesdropping on cellular audio using gyroscopes and accelerometers — inertial sensors commonly built into mobile devices to measure motion through Micro-Electro-Mechanical Systems (MEMS) technology. The research was conducted over the summer by one of our interns, Alec K., and a newly hired full-time engineer, August H.

https://cc-sw.com/leveraging-machine-learning-to-enhance-acoustic-eavesdropping-attacks-part-1-of-4/

Lest We Remember: Cold-Boot Attacks on Encryption Keys

What is DRAM?

What is DRAM?

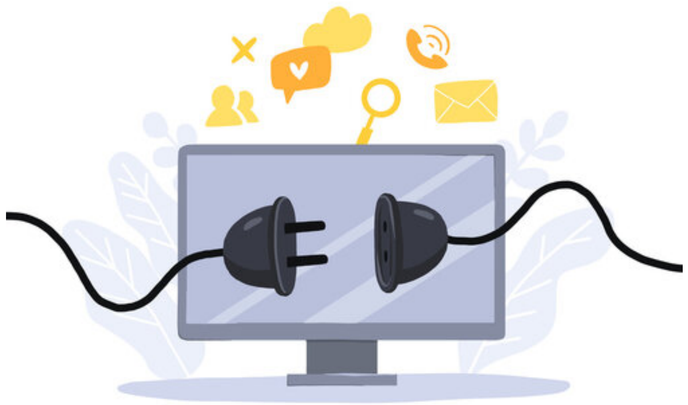
DRAM: Dynamic random-access memory – a type of computer memory (hardware)

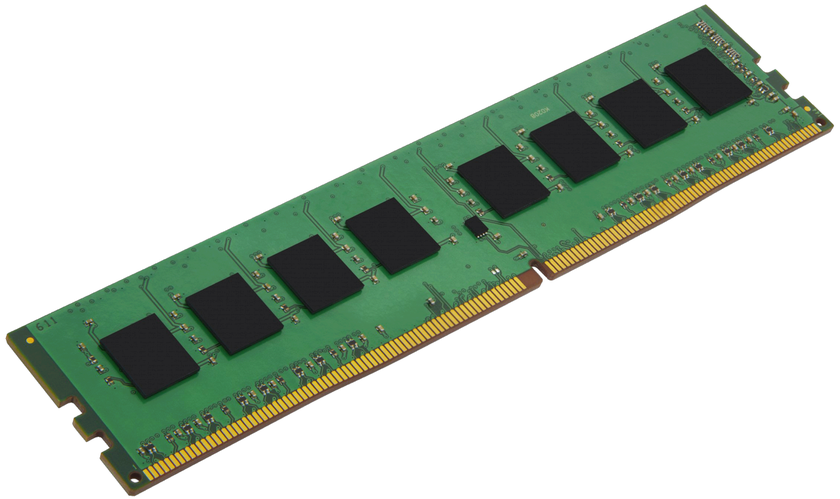
Cold Boot Attack

What is the attack the authors want to conduct?

Cold Boot Attack

What is the attack the authors want to conduct?





Cold Boot Attack

What is the attack the authors want to conduct?

What makes the attack possible?

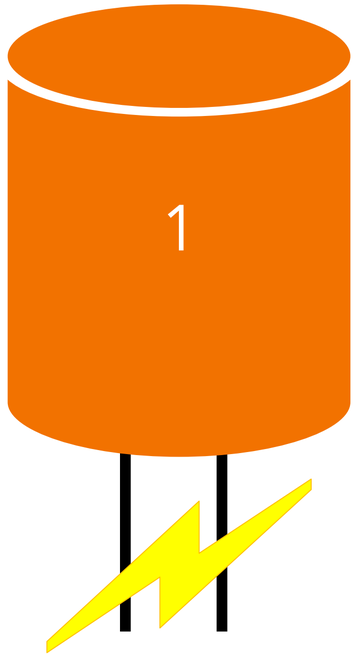
Memory remanance: Most DRAM lose contents *gradually* over a period of seconds not all at once. This creates an opportunity to inspect what's in DRAM!

How does DRAM work?

DRAM cells are essentially just capacitors. What is a capacitor?

How does DRAM work?

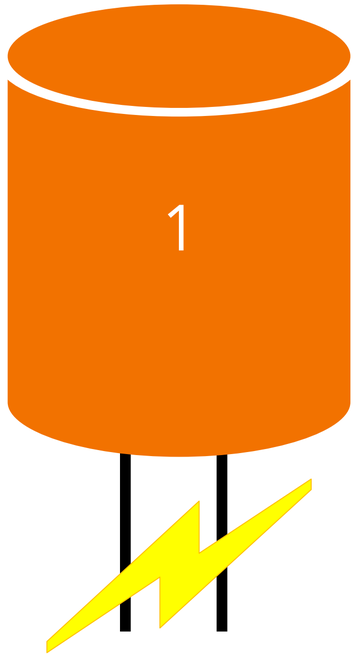
DRAM cells are essentially just capacitors. What is a capacitor?



Write “1”

How does DRAM work?

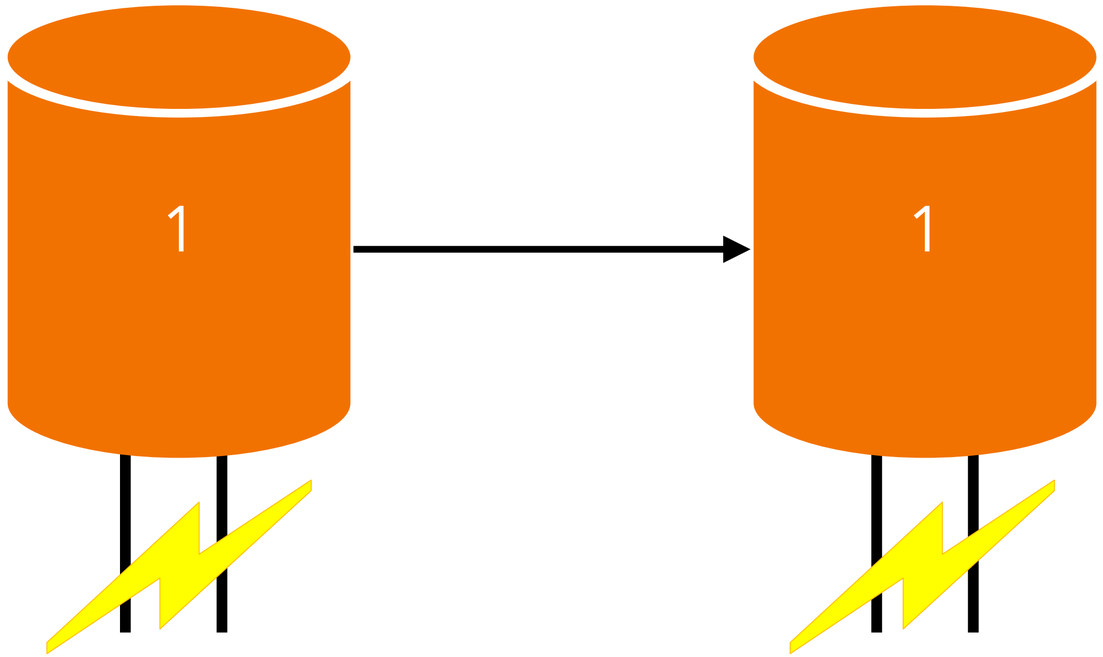
What is a *refresh* in DRAM, and how does it work?



Write “1”

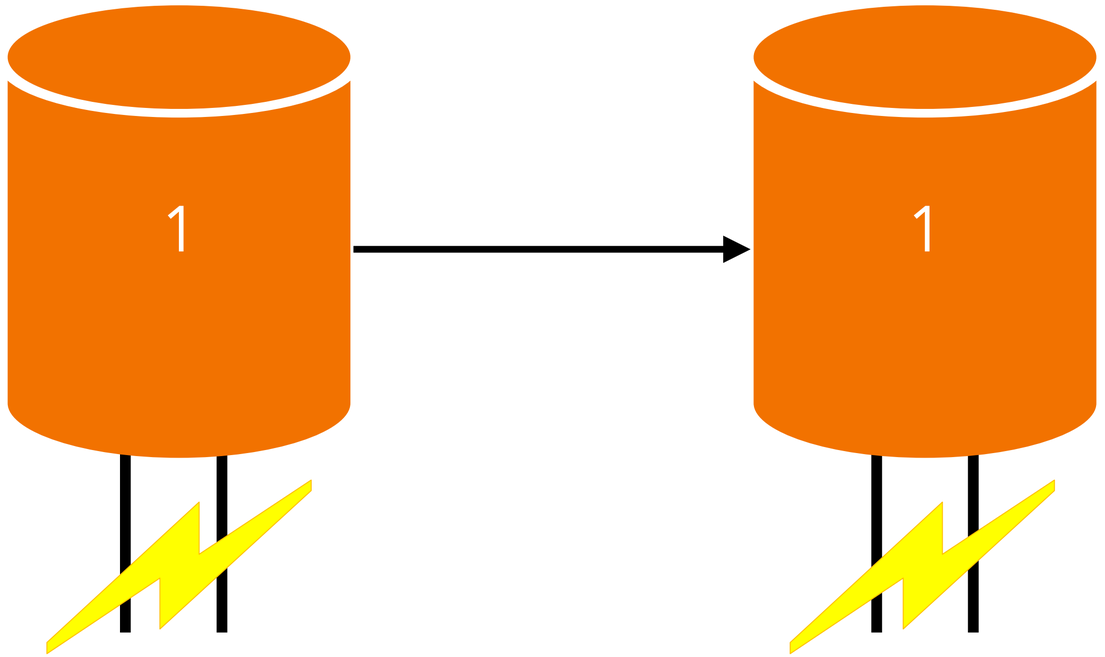
How does DRAM work?

What is a *refresh* in DRAM, and how does it work?



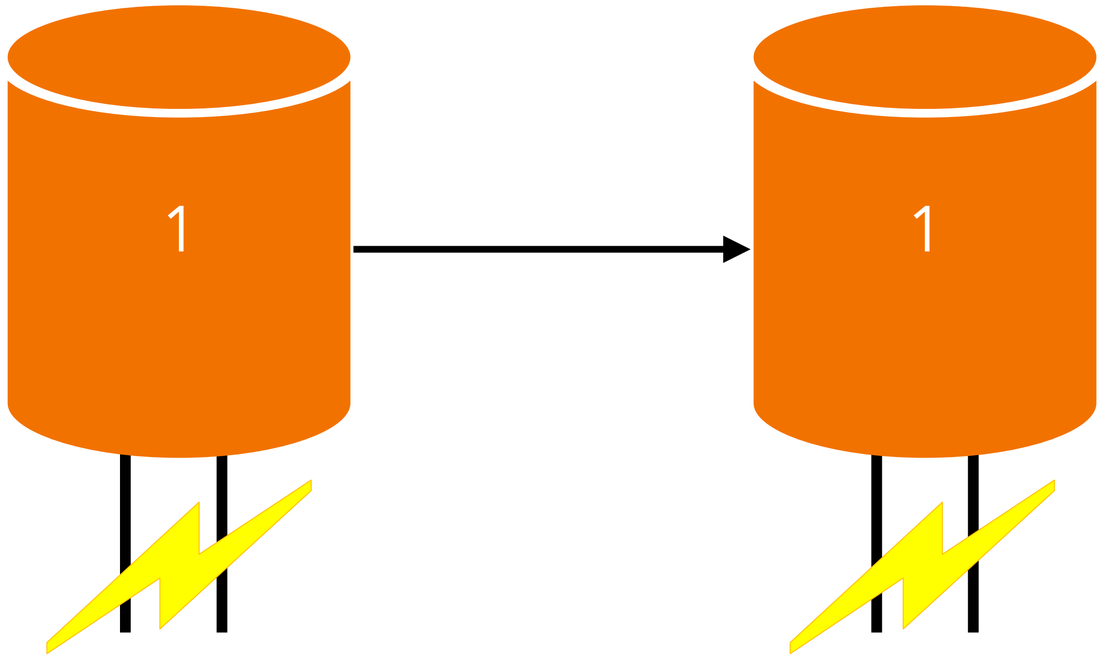
How does DRAM work?

What is a refresh interval?



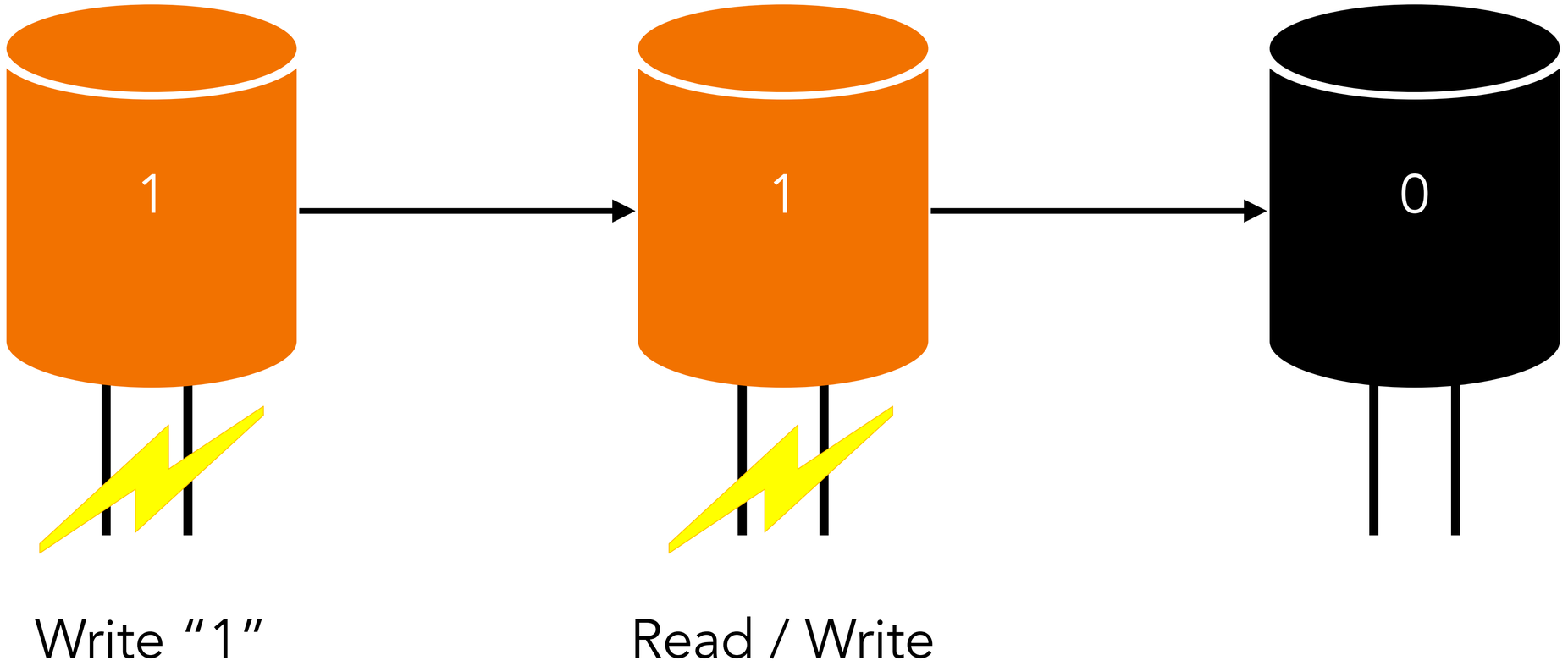
How does DRAM work?

What happens if we don’t refresh the cell?



How does DRAM work?

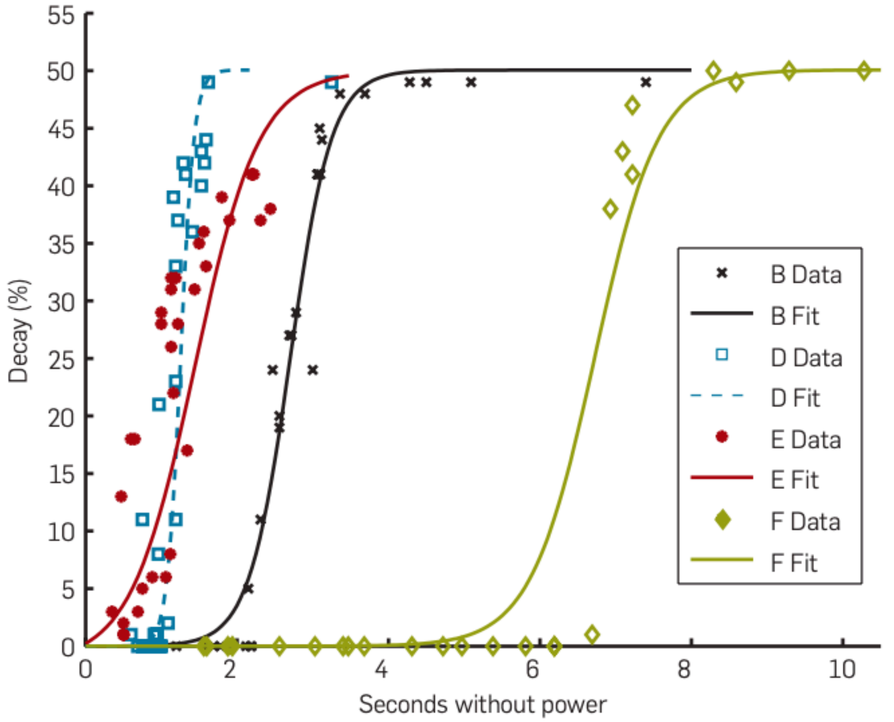
What happens if we don’t refresh the cell?



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?

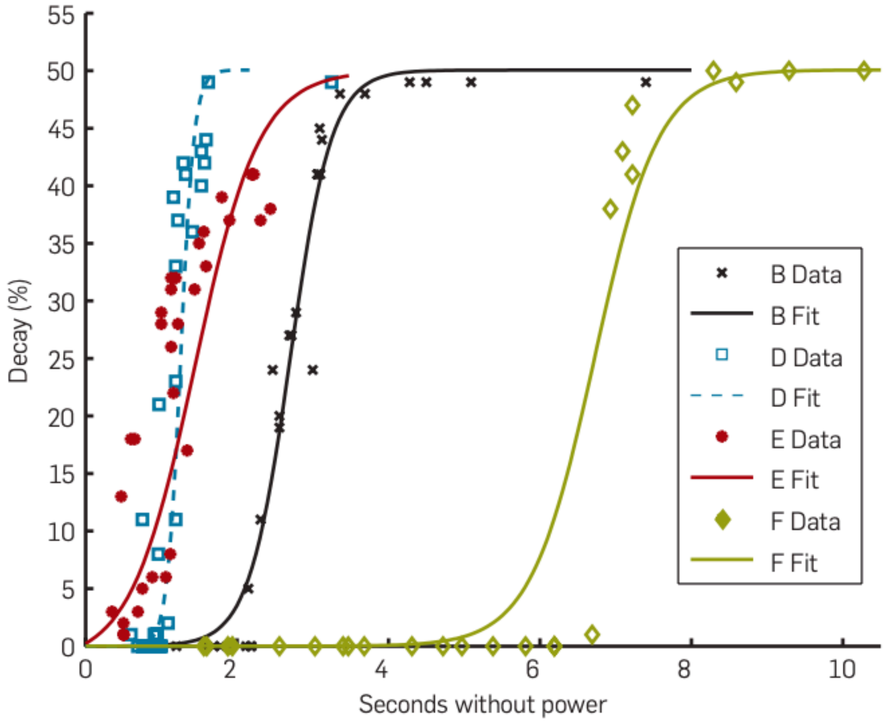
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?
- How did the authors measure errors?

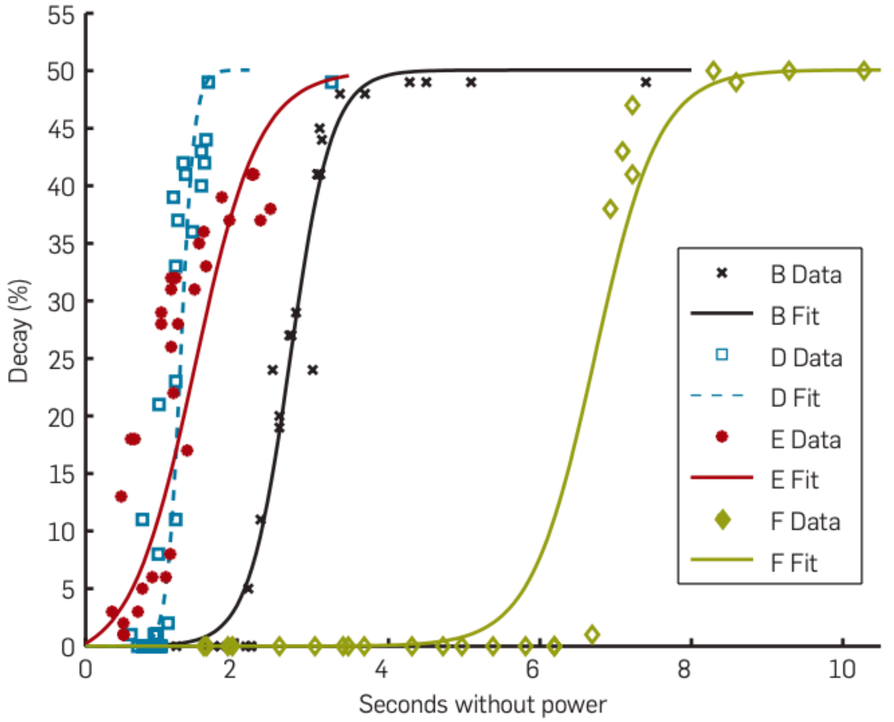
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?
- How did the authors measure errors?
  - **Hamming distance:** number of bit errors divided by the total number of bits

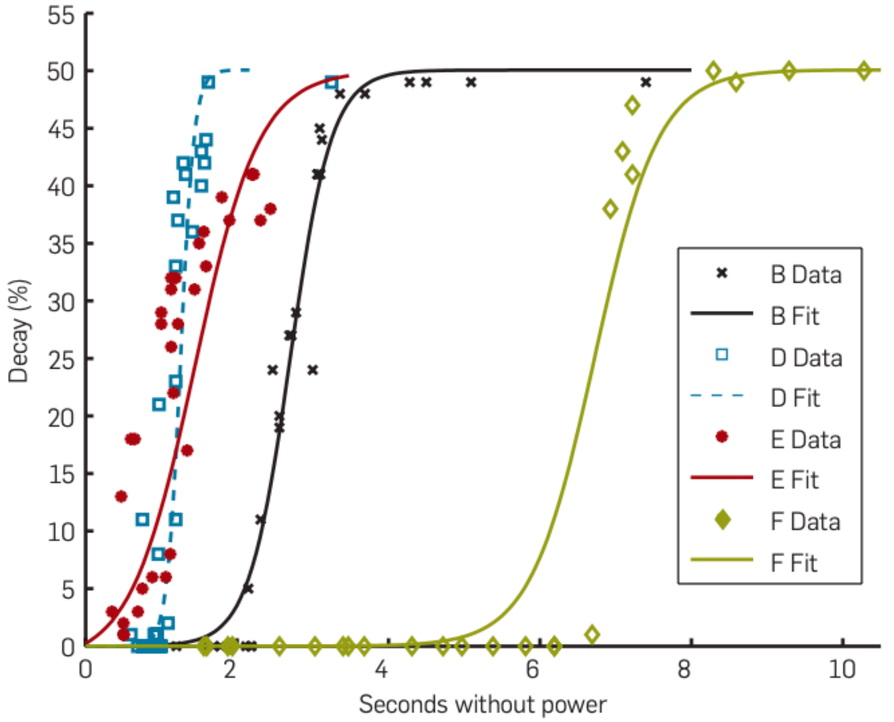
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?
- How did the authors measure errors?
  - **Hamming distance:** number of bit errors divided by the total number of bits
- What would the error rate be if memory had fully decayed?

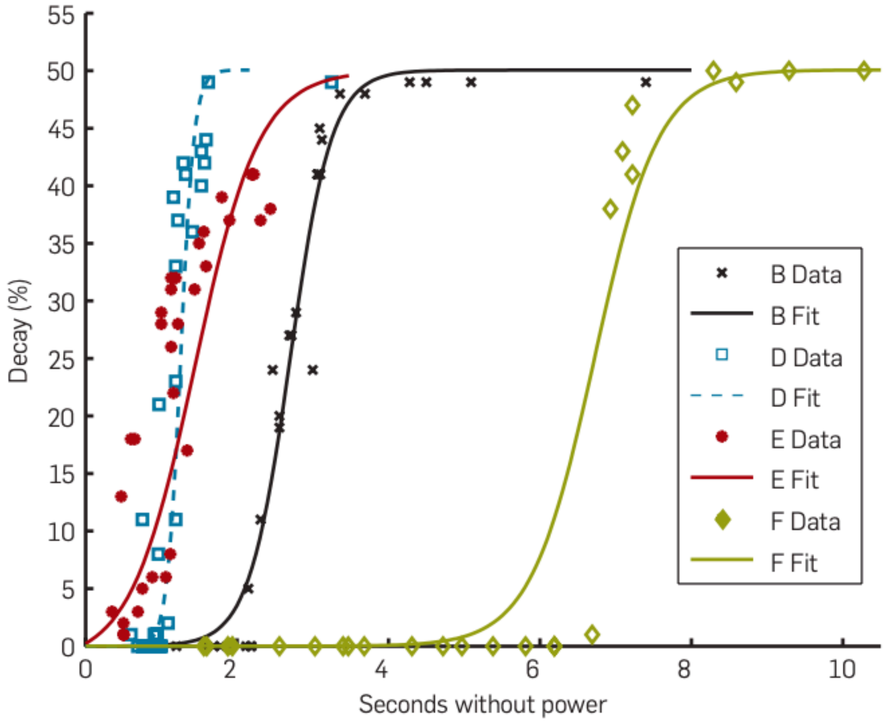
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



Reduced Temperature Experiments

- How did changing the temperature affect decay times?

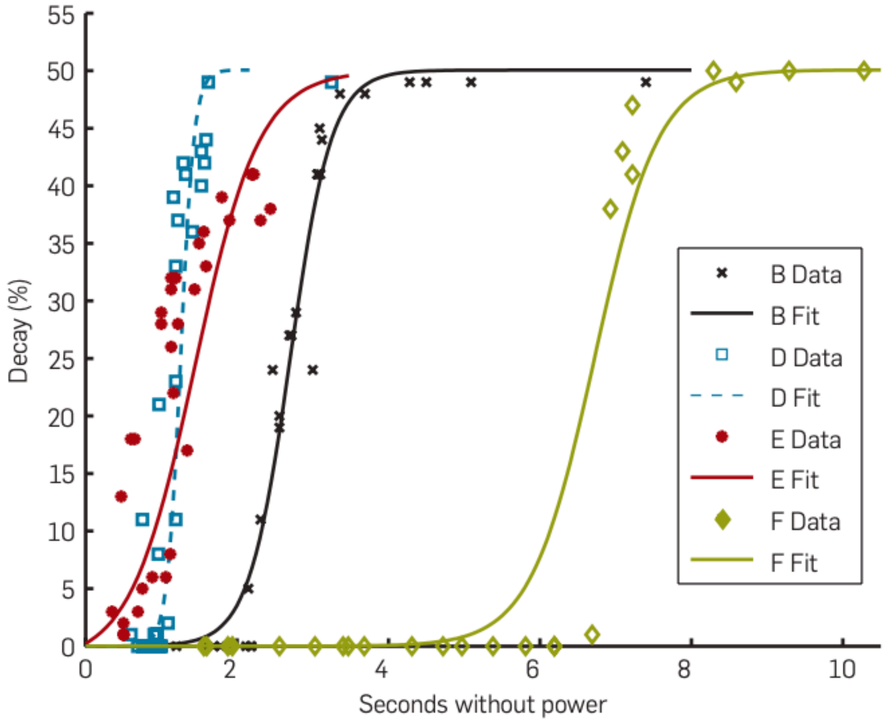
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



Reduced Temperature Experiments

- How did changing the temperature affect decay times?
  - Reduced temperature to -50 degrees celsius (-58 degrees F) – attacker could cut power for 1 minute and recover at least 99.9% of bits correctly

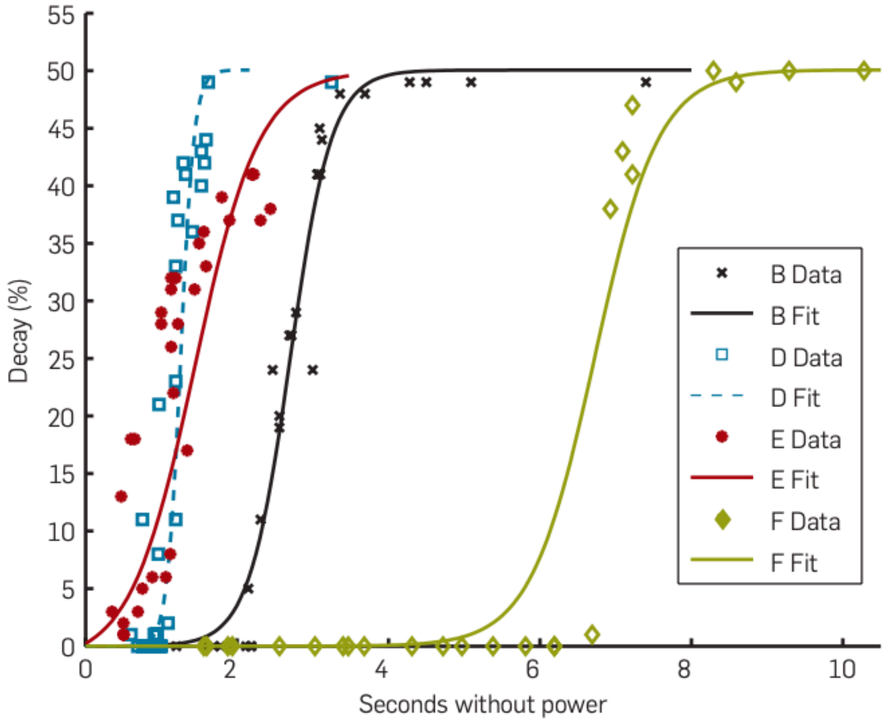
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.

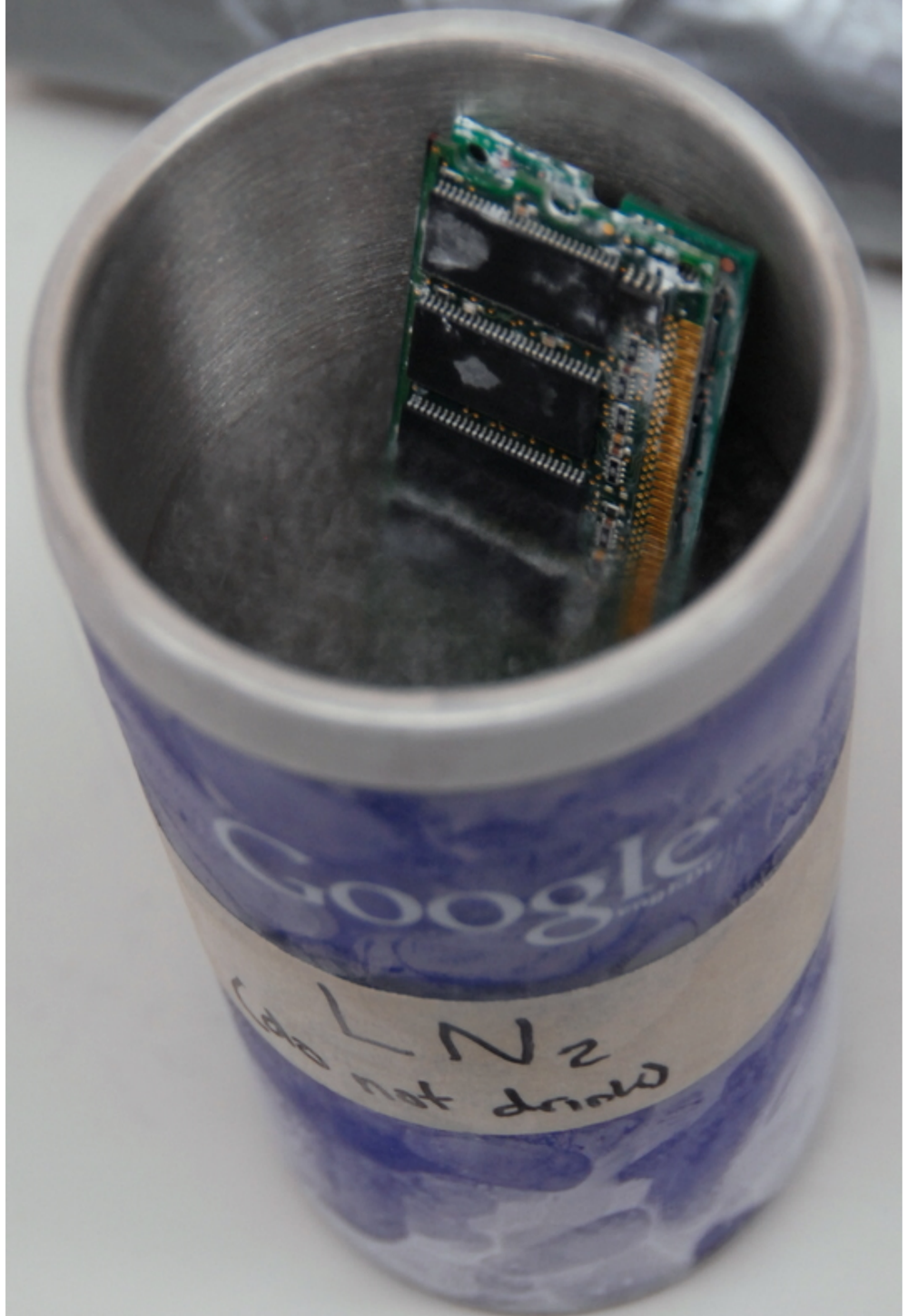


Reduced Temperature Experiments

- How did changing the temperature affect decay times?
  - Reduced temperature to -50 degrees celsius (-58 degrees F) – attacker could cut power for 1 minute and recover at least 99.9% of bits correctly
- How did the authors make it even colder?

Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.




Google
LN2

Getting data out of DRAM

- Turns out, this is nontrivial
  - Rebooting immediately starts to refresh DRAM cells, which could **erase** memory that was previously written in those cells
- You could write a small program that can copy memory to another medium
  - Authors essentially created a few extremely tiny programs to do this to show proof of concept – we won't discuss this but it's pretty cool

What can we do with this?

- Paper goes into lots of fun detail about reconstructing cryptographic keys
  - DES, AES, RSA —> all of these can be reconstructed
- What is the fundamental reason why keys can be reconstructed?

What can we do with this?

- Paper goes into lots of fun detail about reconstructing cryptographic keys
  - DES, AES, RSA —> all of these can be reconstructed
- What is the fundamental reason why keys can be reconstructed?
  - Keys have unique signatures: AES has a **key schedule** with a repeatable pattern that you can exploit as a side channel to find an AES key :)

Attacking Encrypted Disks

- What is an encrypted disk?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?
- How does BitLocker encrypt data on the disk?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?
- How does BitLocker encrypt data on the disk?
  - **Authors defeated BitLocker, FileVault, TrueCrypt, dm-crypt, Loop-AES, and could probably have done a lot more!**

Feasibility

- How feasible is this attack?
- Do you believe this attack will work in practice? Why or why not?

What do these attacks teach us about *trust?*

What can we do about side channels?

Break Time + Attendance

https://tinyurl.com/cse227-attend

Codeword:
Chann3L

Skill Squatting Attacks

What's an IoT Device?

What's an IoT Device?

IoT devices are devices with sensors, processing, software, and other tech that allow them to exchange data with other devices and actuate in the real world

What’s a voice user interface?

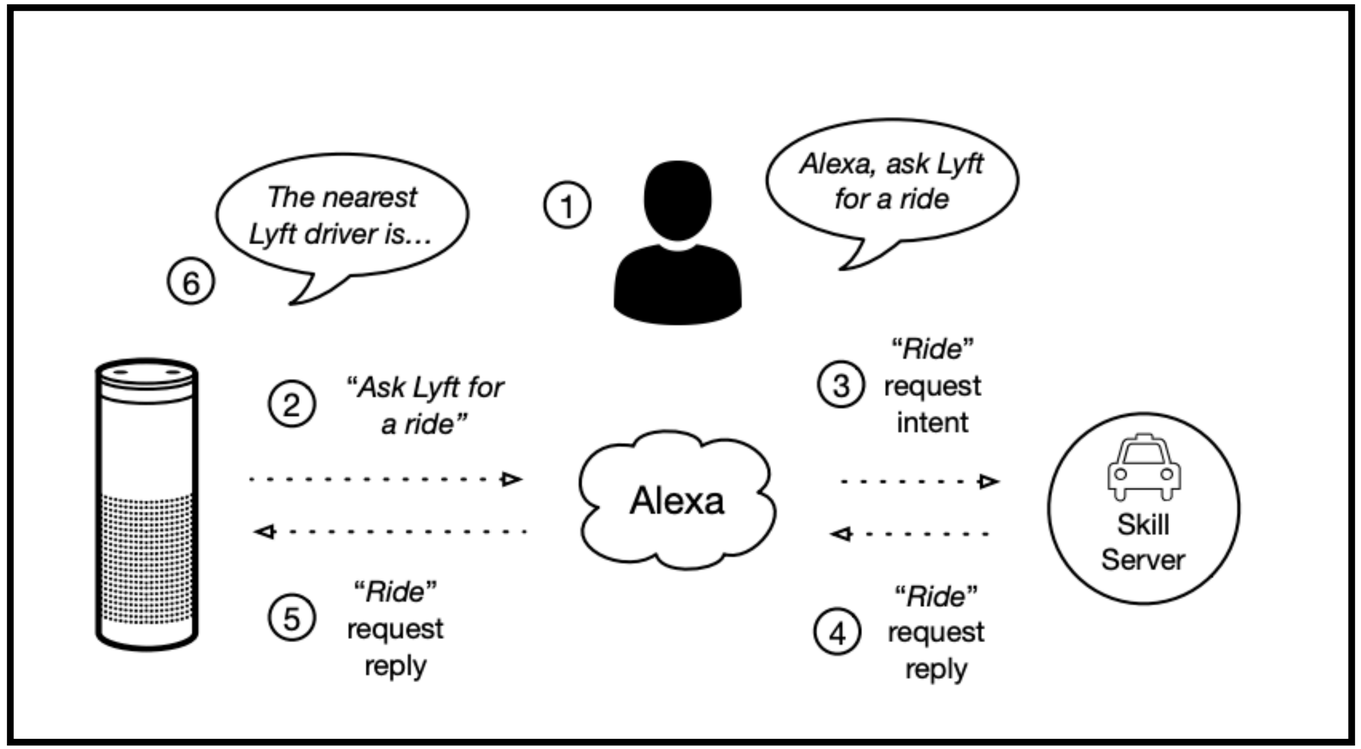
What’s a voice user interface?

Technology that lets you engage with a device or application using speech recognition.

What's a "skill?" (Amazon Alexa)

What's a "skill?" (Amazon Alexa)

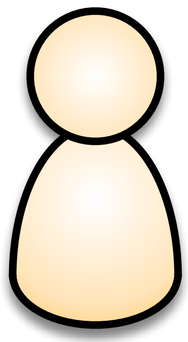
Essentially, an "app" controlled by the user's voice.

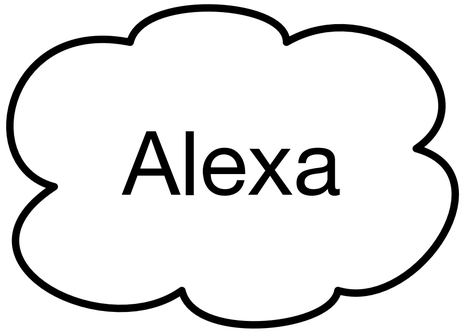


Skill Squatting Attack

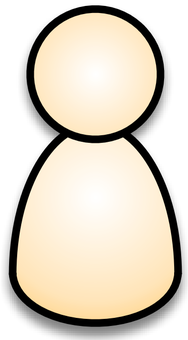
What is the attack the authors want to conduct?

“Alexa, tell me some cat facts!”





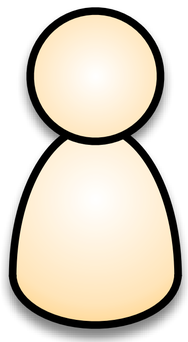
“Alexa, tell me some cat facts!”



Alexa

Skills
…
…
cat forks
cat fast
cat facts
…
…
…
…

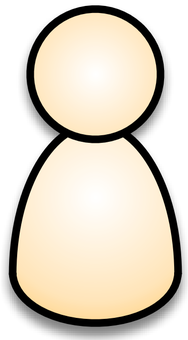
“Alexa, tell me some cat facts!”



Alexa

Skills
…
…
cat forks
cat fast
cat facts
…
…
…
…

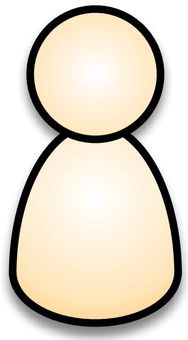
“Alexa, tell me some cat facts!”



Alexa

Skills
…
…
cat forks
cat fast
cat facts
…
…
…
…

“Alexa, tell me some cat facts!”



Alexa

Skills

…

…

cat forks

cat fast

cat facts

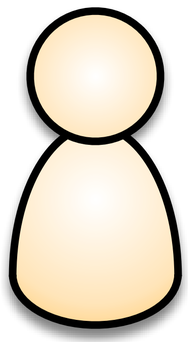
…

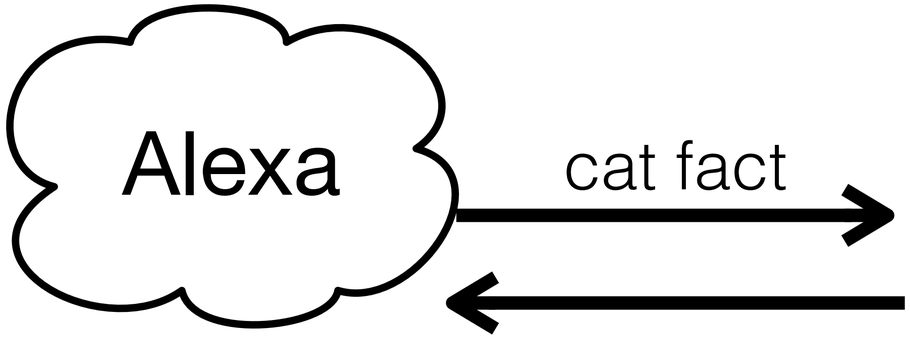
…

…

…

“Alexa, tell me some cat facts!”





Skills

…

…

cat forks

cat fast

cat facts

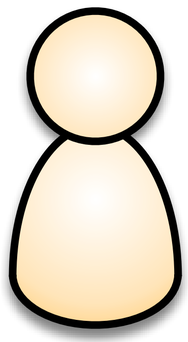
…

…

…

…

“Alexa, tell me some cat facts!”



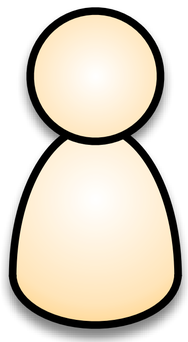
Alexa

cat fact

Skills
…
…
cat forks
cat fast
cat facts
…
…
…
…

“A group of cats is called a clowder!”

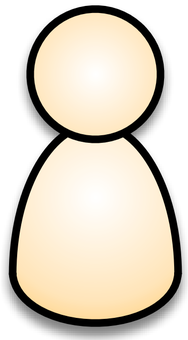
“Alexa, tell me some cat facts!”



Alexa

Skills
…
…
cat forks
cat fast
cat facts
cat fax
…
…
…

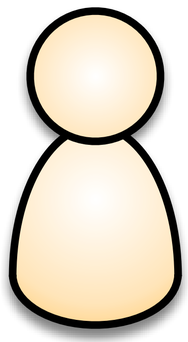
“Alexa, tell me some cat facts!”

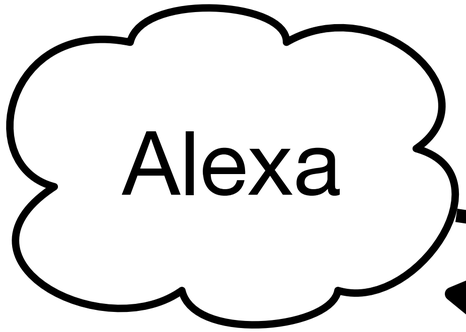


Alexa

Skills
…
…
cat forks
cat fast
cat facts
cat fax
…
…
…

“Alexa, tell me some cat facts!”





cat fact

Skills
…
…
cat forks
cat fast
cat facts
cat fax
…
…
…

“A group of cats is called a murder!”

Skill Squatting Attack

What is the attack the authors want to conduct?

What makes the attack possible?

Skill Squatting Attack

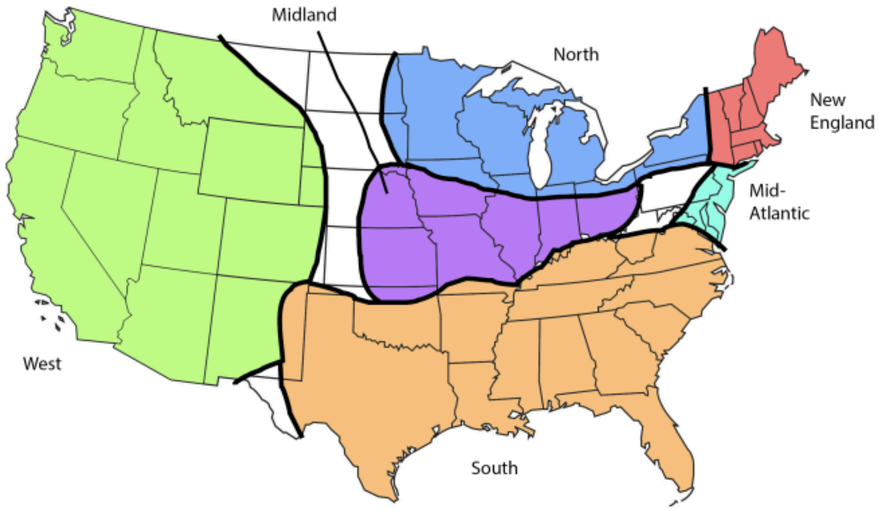
What is the attack the authors want to conduct?

What makes the attack possible?

> Authors found that Amazon's speech-to-text algorithm makes **predictable errors**

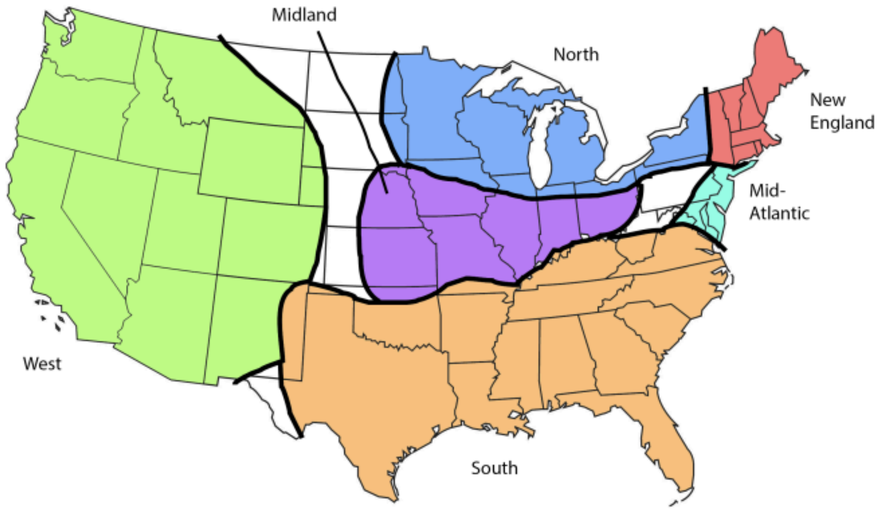
Identifying Predictable Errors

- What data did the authors use to identify predictable errors?
- What was the strategy the authors used to *record* what Alexa interpreted on the other end of speech recognition?
- How well did Alexa do?



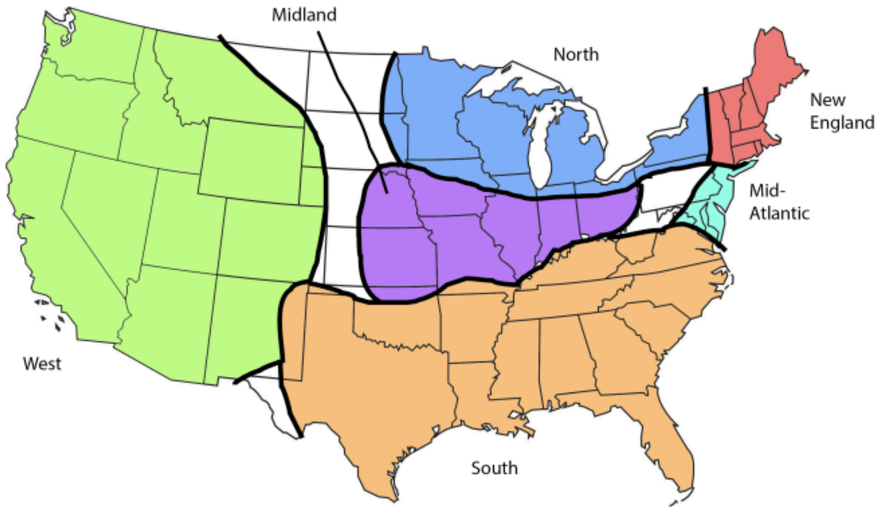
Identifying Predictable Errors

- What data did the authors use to identify predictable errors?
- What was the strategy the authors used to *record* what Alexa interpreted on the other end of speech recognition?
- How well did Alexa do?
  - **68.9% success rate**



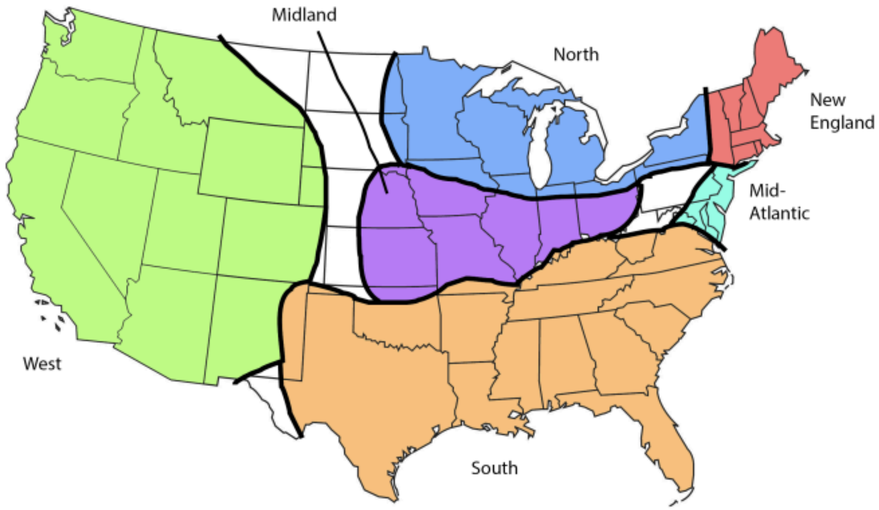
Reasons for Errors

- What's a homophone?
- What's a compound word?
- What's a phoneme?



Reasons for Errors

- What's a homophone?
- What's a compound word?
- What's a phoneme?
  - A single basic sound – the smallest possible phonetic unit – that helps distinguish one word from another in a language



Predictable Errors

Word	Prediction
Sail	Sale
Rip	Rap
Outshine	Outshyne
Lung	Lang
Accelerate	Xcelerate
Mill	No
Preferably	Preferrably
Earthy	Fi
Calm	Com
Coal	Call
Outdoors	Out Doors
Loud	Louder

Word	Prediction
Superhighway	Super Highway
Wet	What
Main	Maine
Boil	Boyle
Sell	Cell
Full	Four
Dime	Time
Bean	Been
Dull	Doll
Sweeten	Sweden
Luck	Lock
Con	Khan

Predictable Errors

Word	Prediction
Sail	Sale
Rip	Rap
Outshine	Outshyne
Lung	Lang
Accelerate	Xcelerate
Mill	No
Preferably	Preferrably
Earthy	Fi
Calm	Com
Coal	Call
Outdoors	Out Doors
Loud	Louder

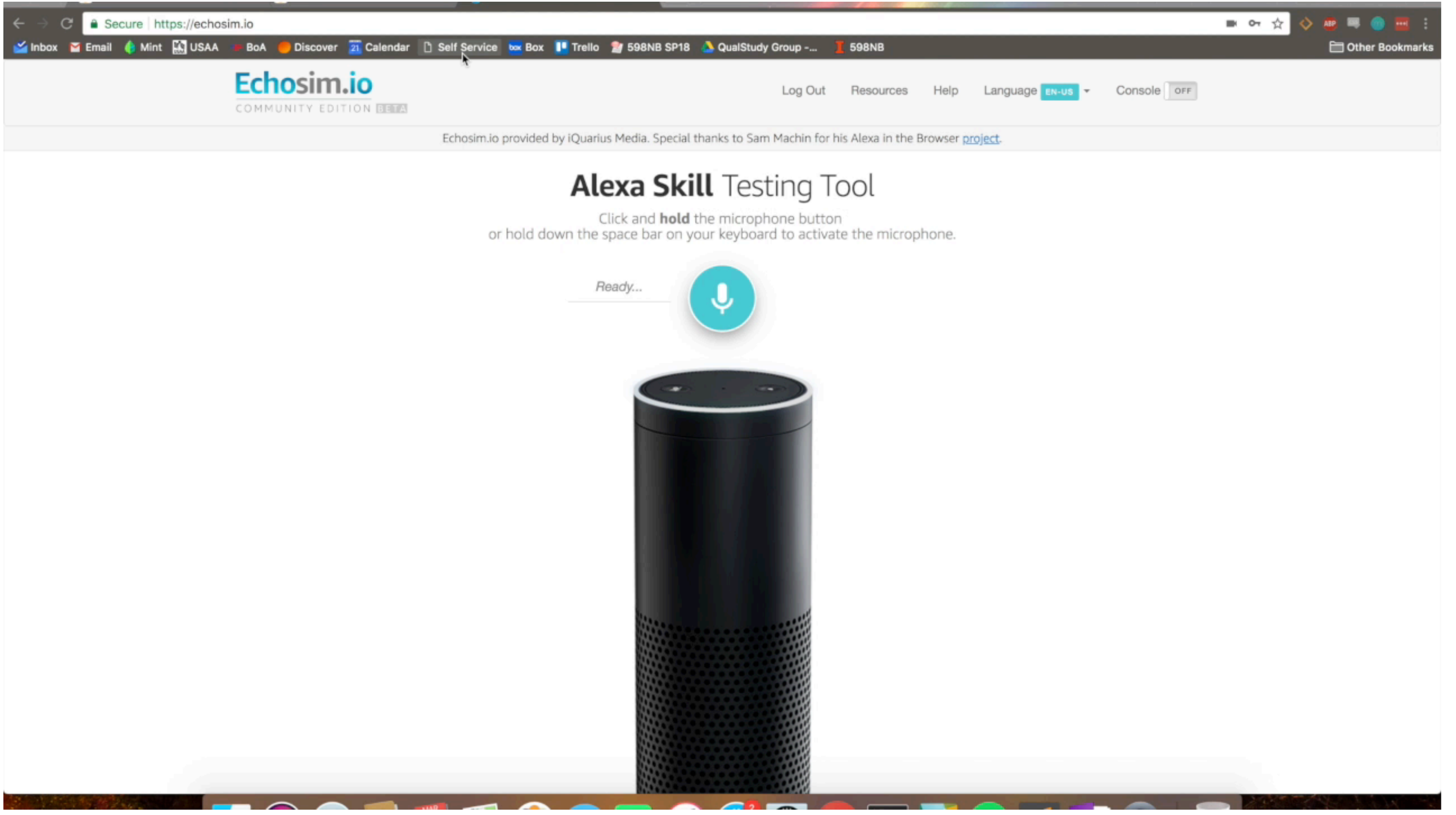
Word	Prediction
Superhighway	Super Highway
Wet	What
Main	Maine
Boil	Boyle
Sell	Cell
Full	Four
Dime	Time
Bean	Been
Dull	Doll
Sweeten	Sweden
Luck	Lock
Con	Khan

Extending the attack with phonemes

- The authors extended their attack beyond words using phonemes
  - What was their basic attack strategy?
  - How effective was their attack?

Extending the attack with phonemes

- The authors extended their attack beyond words using phonemes
  - What was their basic attack strategy?
  - How effective was their attack?
    - Increased possible word errors from 188 —> **3606** (17.5x increase)


Secure https://echosim.io
Inbox Email Mint USAA BoA Discover Calendar Self Service Box Trello 598NB SP18 QualStudy Group -... 598NB Other Bookmarks
Echosim.io
COMMUNITY EDITION BETA
Log Out Resources Help Language EN-US Console OFF
Echosim.io provided by iQuarius Media. Special thanks to Sam Machin for his Alexa in the Browser project.
Alexa Skill Testing Tool
Click and hold the microphone button
or hold down the space bar on your keyboard to activate the microphone.
Ready...

Meta-points

- What are the capabilities of the attacker?
- What's the threat model in this attack?
- Do you believe this attack will work in practice? Why or why not?
- How would you defend against this attack?

Next time…

- Moving away from software and devices, moving towards the **web**
  - Next two weeks on web, next *four* weeks on “web + networking”
- Keep working on your projects!

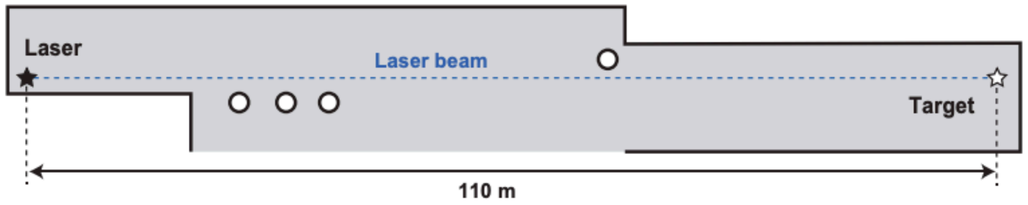
Light Commands

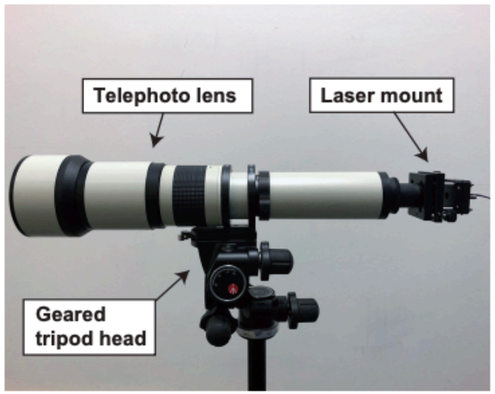
Light Commands Attack

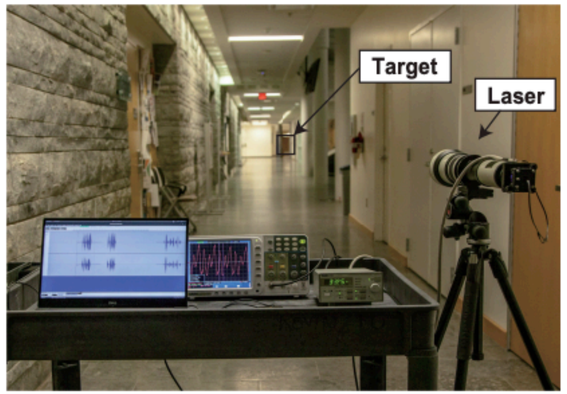
What is the attack the authors want to conduct?

Light Commands Attack

What is the attack the authors want to conduct?







Light Commands Attack

What is the attack the authors want to conduct?

What makes the attack possible?

Light Commands Attack

What is the attack the authors want to conduct?

What makes the attack possible?

> “A semantic gap between the physics and specifications of microphones, where microphones often unintentionally respond to light as if it was sound.”

Light Commands – Threat Model

- What are the capabilities of the attacker?

Light Commands – Threat Model

- What are the capabilities of the attacker?
  - No physical access
  - Line of sight
  - Device feedback
  - Device characteristics (they can test at home)

Light Commands Attack

What is the attack the authors want to conduct?

What makes the attack possible?

> “A semantic gap between the physics and specifications of microphones, where microphones often unintentionally respond to light as if it was sound.”

Injecting Sounds as Light

- How did the authors convert a sound (e.g., an audio file saying “OK Google”) into light?

Injecting Sounds as Light

- How did the authors convert a sound (e.g., an audio file saying “OK Google”) into light?
  - Essentially, modulated laser intensity as a function of the audio waveform

Injecting Sounds as Light

- How did the authors convert a sound (e.g., an audio file saying “OK Google”) into light?
  - Essentially, modulated laser intensity as a function of the audio waveform
- Why are microphones sensitive to light in this way?

Injecting Sounds as Light

- How did the authors convert a sound (e.g., an audio file saying “OK Google”) into light?
  - Essentially, modulated laser intensity as a function of the audio waveform
- Why are microphones sensitive to light in this way?
  - Photoelectric effects – emission of electrons due to light hitting the circuit
  - Photoacoustic effect – the diaphragm **moves** because the high intensity light hits it – and moves with intensity thus recording the sound

Authors tried this on a number of devices – and it worked!

Device	Backend	Category	Authen-tication	Minimum Power [mW]*	Max Distance at 60 mW [m]**	Max Distance at 5 mW [m]***
Google Home	Google Assistant	Speaker	No	0.5	50+	110+
Google Home Mini	Google Assistant	Speaker	No	16	20	—
Google Nest Cam IQ	Google Assistant	Camera	No	9	50+	—
Echo Plus 1st Generation	Alexa	Speaker	No	2.4	50+	110+
Echo Plus 2nd Generation	Alexa	Speaker	No	2.9	50+	50
Echo	Alexa	Speaker	No	25	50+	—
Echo Dot 2nd Generation	Alexa	Speaker	No	7	50+	—
Echo Dot 3rd Generation	Alexa	Speaker	No	9	50+	—
Echo Show 5	Alexa	Speaker	No	17	50+	—
Echo Spot	Alexa	Speaker	No	29	50+	—
Facebook Portal Mini (Front Mic)	Alexa	Speaker	No	1	50+	40
Facebook Portal Mini (Front Mic)§	Portal	Speaker	No	6	40	—
Fire Cube TV	Alexa	Streamer	No	13	20	—
EcoBee 4	Alexa	Thermostat	No	1.7	50+	70
iPhone XR (Front Mic)	Siri	Phone	Yes	21	10	—
iPad 6th Gen	Siri	Tablet	Yes	27	20	—
Samsung Galaxy S9 (Bottom Mic)	Google Assistant	Phone	Yes	60	5	—
Google Pixel 2 (Bottom Mic)	Google Assistant	Phone	Yes	46	5	—

*at 30 cm distance, **Data limited to a 50 m long corridor, ***Data limited to a 110 m long corridor, §Data generated using only the first 3 commands.

What were the authors able to do?

What were the authors able to do?

- Ask for the time
- Set the volume to zero
- Purchase items
- Open connected devices (e.g., doors)
- Generate commands + even trigger skills!

What were the authors able to do?

What were the authors able to do?

- Ask for the time
- Set the volume to zero
- Purchase items
- Open connected devices (e.g., doors)
- Generate commands + even trigger skills!

Feasibility

- Do you believe this attack will work in practice? Why or why not?
- How would you defend against this