# IP-Based IoT Device Detection

**Hang Guo** and John Heidemann
USC/Information Sciences Institute
August 20, 2018  IoT S&P'18

# IoT Devices Cause Record-Breaking DDoS Attacks

2016-09-20: **620 Gb/s** attack against cybersecurity blog KrebsOnSecurity.com

2016-09-23: Est **1 Tb/s** attack on French cloud-computing provider OVH

2016-10-21: Est **1 Tb/s** attack against DNS provider Dyn

- Last about **5 and a half hours** in total, affect **69 services**

Airbnb
**Amazon.com**
Ancestry.com
The A.V. Club
BBC
The Boston Globe
Box
Business Insider
CNN
Comcast
CrunchBase
DirecTV
The Elder Scrolls Online
Electronic Arts
Etsy
FiveThirtyEight
Fox News
The Guardian
GitHub
HostGator
iHeartRadio
Imgur
Indiegogo
Mashable
NHL
**Netflix**
The New York Times
Overstock.com
PayPal
Pinterest
Pixlr
PlayStation Network
Qualtrics
Ruby Lane
RuneScape
SaneBox
Seamless
Second Life
Shopify
Slack
SoundCloud
Squarespace
Spotify
Starbucks
Storify
Swedish Government
Swedish Civil Contingencies Agency
Visa
Vox Media
Walgreens
The Wall Street Journal
Wikia
Wired
Wix.com
WWE Network
Xbox Live
Yammer
Yelp
Zillow

Vulnerable IoT Devices Broke the Internet !

USC Viterbi School of Engineering Information Sciences Institute ant.isi.edu

# Motivation

- Vulnerable IoT devices threaten the security of Internet ecosystem
- To understand these threats requires knowledge IoT devices.
  - Such as locations, distribution and growth
- These knowledge help guide the design and deployment of future IoT security solutions
  - Reveal the scale of IoT security problem, the problem's growth and distribution

# Contributions

- A new method to detect IoT devices from observations of Internet traffic.

- Apply our method to real-world network traffic
  - Find at least 35 IoT devices on a college campus and 60 in customers of an IXP.

# Talk Structure

**Method**

Detection Results

Future Work and More

# Insight for Detection

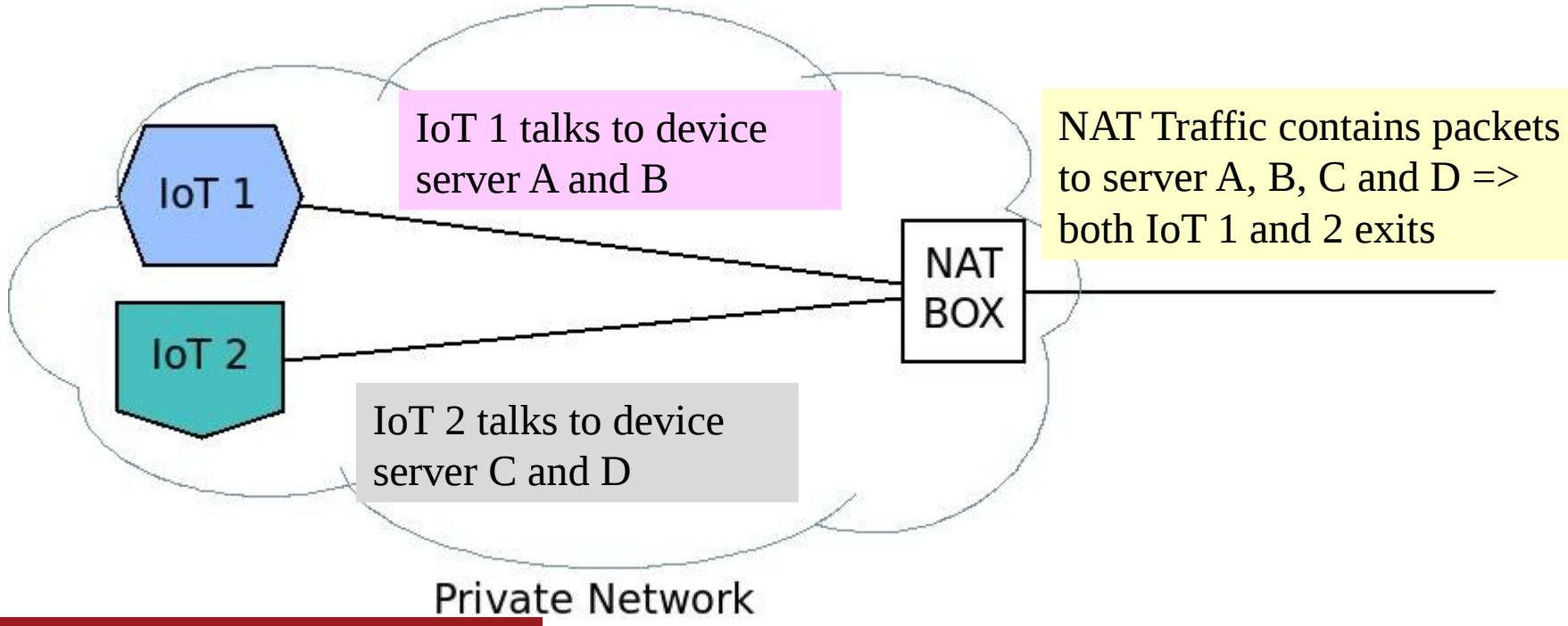
- Most IoT devices exchange packets with servers run by their manufacturers (called device server).

- We can **identify IoT devices** by watching for these packet exchanges.

- We can **identify the types of devices** since servers are usually unique for device type

# Packet Exchanges Robust to NAT

- IoT devices often behind NAT:
  - Can still identify packet exchanges when NATs mix traffic from multiple devices.



# Detection Method: Learning

- **Determine device servers to look for**

- Look for device servers in Traffic

# Determine Device Servers: Find Candidate

**Find Server Candidate**

Filter Candidates

Handle Shared Servers

Track Server IP

Server names from DNS A requests made by **sample IoT devices** during operation.

- devices we own
- devices from others with public traffic measurements.

# Determine Device Servers: 2-Step Filtering

Find Server Candidate

**Filter Candidates**

Handle Shared Servers

Track Server IP

Risk: confuse among devices that use public services (e.g. time/NTP) provided by 3rd party servers

*Test: server domains contain no manufacturer info**

filter **Third-Party Servers**

filter **Human-Facing Servers**

*Test: server respond HTTP(s) requests with contents**

Risk: confuse laptops or cellphones (operated by human) as IoT devices.

*Exact test used here in our paper

# Determine Device Servers: Filtering Result

Find Server Candidate

**Filter Candidates**

Handle Shared Servers

Track Server IP

What remains: **device-facing manufacturer server**, or just **device servers**.

Run by IoT manufacturers and serve devices only.

We only use device servers for detection.

# Determine Device Servers: Shared Servers

Find Server Candidate

Filter Candidates

**Handle Shared Servers**

Track Server IP

different device types share the **exact set** of device server names

→ treat devices as same type

different device types share **partial** device server names

→ can't guarantee they're distinguishable

→ when detect devices sharing servers, conservatively report detect at least one of them

# Determine Device Servers: Track Server IPs

- Find Server Candidate
- Filter Candidates
- Handle Shared Servers
- **Track Server IP**

Need: discover device servers by domain names
but search for device server by IPs in traffic
===> need to track server name to IP resolution

→ by DNS querying server names every hour

Risk: server names are long-lived
server IP sometimes change/depend on location.

→ track server IPs at roughly the same time and from the same location as network traces

# Detection Method: Running

- Determine device servers to look for

- **Look for device servers in Traffic**

# Look for device servers in Traffic

- Track a list of device server names that each type of device talks to.

- Define a **threshold**: typically a majority, but not all servers we track for a device*
  - Traffic to that number of server names from a given IP address ==> presence of that IoT device.



*more details about this threshold in our paper

# Talk Structure

Method

**Detection Results**

Future Work and More

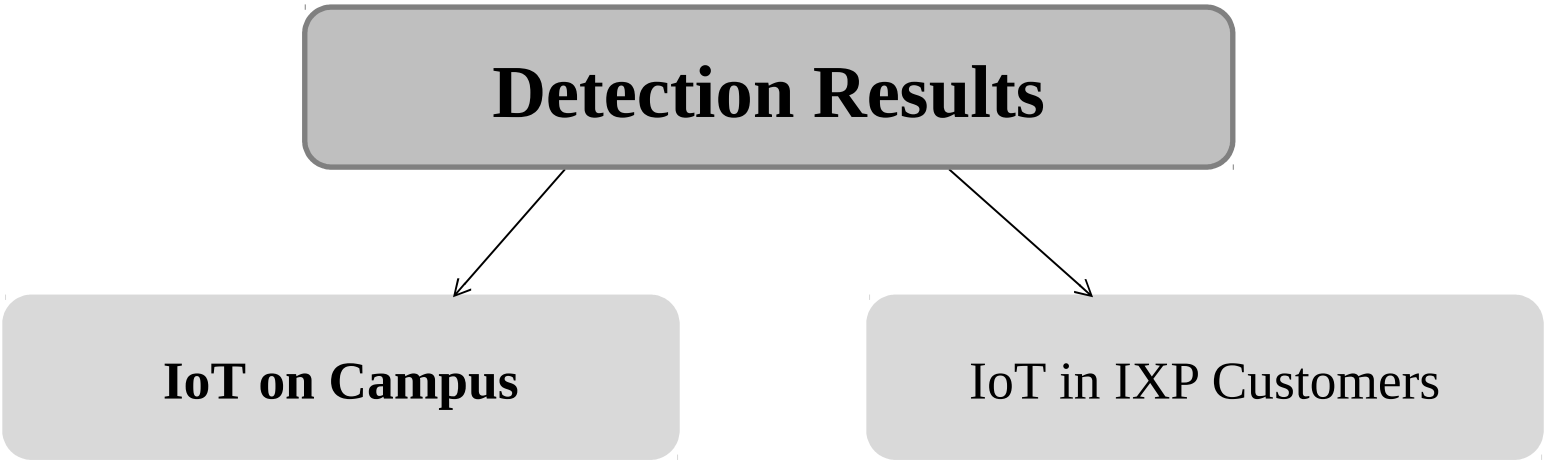
# Devices and Server Names for Detection

- We use 26 types of IoT devices
  - 10 devices we own
  - 16 devices from data provided by the University of New South Wales*
- We extract 99 distinct device server names from these 26 devices for detection

USC Viterbi School of Engineering Information Sciences Institute ant.isi.edu

*Sivanathan, A., Sherratt, D., Gharakheili, H. H., Radford, A., Wijenayake, C., Vishwanath, A., and Sivaraman, V. Characterizing and classifying IoT traffic in smart cities and campuses. In Proceedings of the IEEE Infocom Workshop on Smart Cities and Urban Computing (May 2017), pp. 559–564.

# Talk Structure



# IoT on Campus: Inputs

- **Network Traces:** USC campus guest Wi-Fi traffic
  - 6-day measurement (2017-10-06 to 2017-10-11)
  - For user privacy: drop payload; anonymize IP.

- **Server IPs:** IPs for our 99 IoT device servers , tracked across the same 6-day period, at USC.

# IoT on Campus: Detection Results

- 35 detections from 8 user IPs

| IP A & B & C | IP D | IP E | IP F | IP G & H |
|---|---|---|---|---|
| Belkin_SmartPlug | Samsung_IPCam | Samsung_IPCam | HP_Printer | Amazon_* |
| Samsung_IPCam | HP_Printer | HP_Printer | Withings_Scale | |
| HP_Printer | LiFX_SmartLightBulb | LiFX_SmartLightBulb | Amazon_* | |
| Netatmo_WeatherStation | Amazon_* | Amazon_* | | |
| LiFX_SmartLightBulb | Withings_* | | | |
| Amazon_* | | | | |
| Withings_* | | | | |

at least one of Amazon_FireTV and Amazon_Echo

at least one of Withings_Scale and Withings_SleepSensor

IPs with IoT devices often have several devices ==> use of NAT

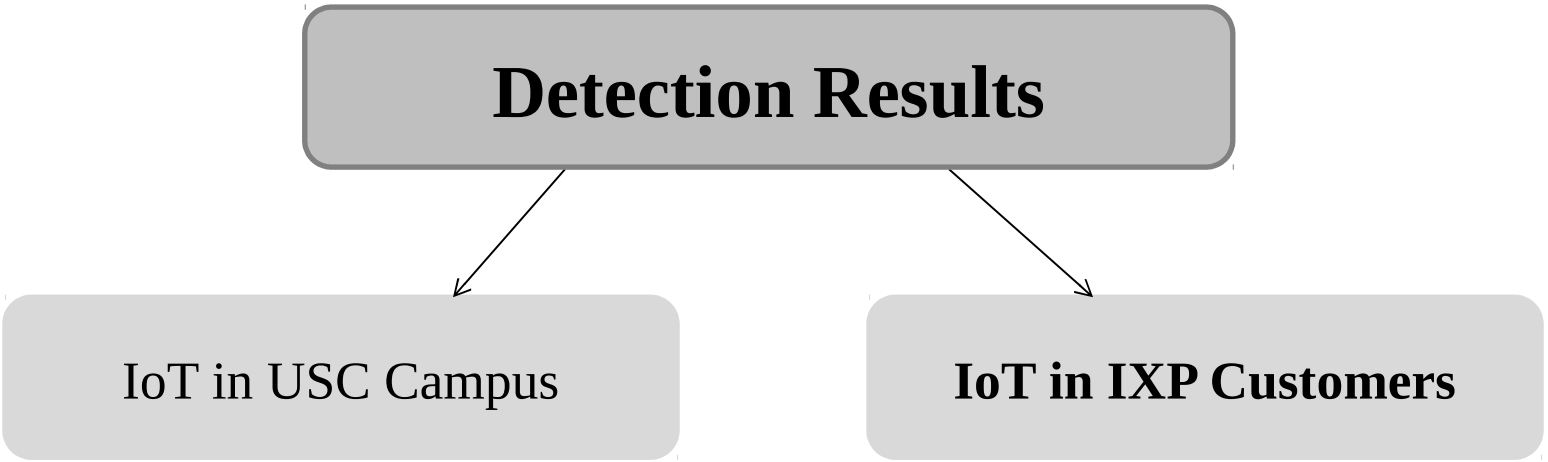
Look like one user whose IP changes 3 times during our measurement ===> Can't confirm due to lack of access to raw traffic capture and DHCP data

# IoT Devices in Campus: Conclusion

- Our approach works at a real campus.

- Our result in an under estimation of actual IoT deployment on USC Campus.
  - Guest Wifi is only a small fraction of all campus traffic.

# Talk Structure



# IoT at an IXP: Inputs

- **Network Traces:** Flow-level traffic from IXP FRGP (Front-Range Gigapop)
  - 10-day measurement (2015-05-10 to 2015-05-19 )
  - shared by Colorado State University.

- **Server IPs:** Historical Passive DNS data from Farisight Security*
  - Since we don't have IPs for our device servers in 2015.



*Security, F. Passive DNS historical Internet database: Farsight DNSDB. https://www.farsightsecurity.com/solutions/dnsdb/

# IoT at an IXP: Detection Results

- 60 detections of only two types of IoT devices
  - Withings_SmartScale (58) and PIXSTAR_PhotoFrame (2).

- We believe incomplete server IPs cause this under-detection*
  - Farsight data only covers IPs for 51 of our 99 device servers
  - Even for the covered server names, the Farsight data potentially miss IPs:
    - For server names that rotate IPs within pool, Farsight's passive measurement may miss part of the pool



*More discussions in paper

# IoT at an IXP: Redo Detection

- Redo detection with a more complete list of server IPs:
  - **Network Traces:** first half day's FRGP data
  - **Server IPs:** IP history we track for our 99 server names
    - From 2017-10-12 to 2018-02-23, at USC
    - We show these IPs are likely applicable, by proving IP for our 99 server names are either stable or from stable pools*.

- Detection result shows at least 4 more types of device detected
  - Despite we only use first half day's FRGP traffic.



*More details in paper

# IoT at an IXP: Conclusion

- It’s hard to detect IoT devices in the past with our IP-based detection method
  - Server IPs change over time
  - Commercial DNS data has limited coverage.

# Talk Structure

Method

Detection Results

**Future Work and More**

# Future Work

- IP-based method require historical DNS data for detection in archived traces
- We are working on a new method that directly work with server's DNS domain names
- We are also looking at approaches to learn new server names as part of the detection process

# Conclusion

- We develop an IP-based IoT detection method.
- We apply it to traffic from a college campus and an IXP and find real-world IoT devices.

Check our paper for validation and more about our IP-based method and apply it to you dataset!

Check our tech report* for our new **DNS-based** and **certificate-based** IoT detection methods!

USC Viterbi School of Engineering Information Sciences Institute ant.isi.edu

*Tech report at:
https://www.isi.edu/~hangguo/papers/Guo18c.pdf