# Lecture #5: IoT botnet Measurements 2

Antonia Affinito, Etienne Khan, Ting-Han Chen, and Cristian Hesselman

University of Twente | June 5, 2024

UNIVERSITY OF TWENTE.

SIDN LABS

ITS NOT MALWARE
WH
.GOV
ITS ALTERNATIVE SOFTWARE
memegenerator.net

# Admin

# Important dates

- Lab report (PDF) and required files: **Wed Jun 19, 9 AM CEST**

- Written exam: **Wed July 3, 13:45-15:45** ⚠

- Alle summaries and lab reports to be submitted through CANVAS

# Where are you with your lab assignment?

- Still trying to find the instructions on the SSI site
- Designing measurement setup
- Analyzing measurements
- Writing lab report
- Just need to click “submit” in Canvas

# Interactive lectures, so please speak up!

# Schedule

| Lecture | Date | Contents |
|---|---|---|
| R1 | May 1 | Course introduction |
| R2 | May 8 | IoT and Internet Core Protocols |
| G1 | May 14 | How the core of the Internet works |
| R3 | May 15 | IoT Edge Security Systems |
|  | May 22 | No lecture (as several of your teachers will be in Dresden :) |
| R4 | May 29 | IoT Botnet Measurements 1 |
| R5 | Jun 5 | IoT Botnet Measurements 2 |
| R6 | Jun 12 | IoT Security in Non-Carpeted Areas |
| G2 | Jun 14 | Maarten Bodlaender, Nokia, title TBP |
| R7 | Jun 19 | IoT Device Security |
|  | Jun 26 | No lecture (so you can study for the exam :) |

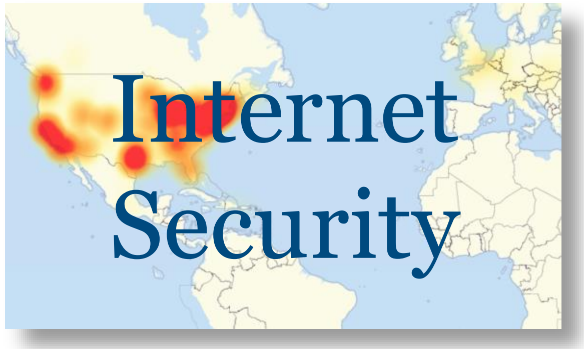
# Introduction to today’s lecture

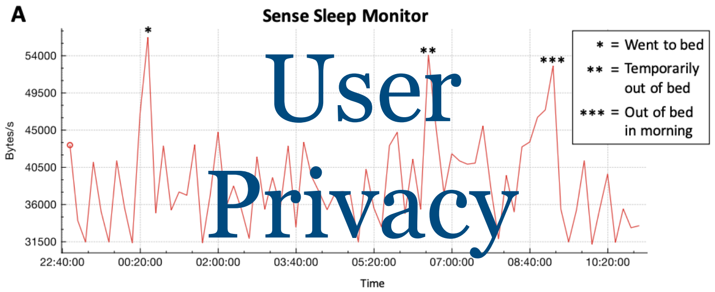
# Motivation: mitigation of IoT botnets

- Requires **scalable** mechanisms to understand **IoT bot behavior** as well **where IoT devices are**

- Challenging because of wide variety of IoT devices and their increasing number and distribution across multiple network operators

- Example mechanisms:
  - Post-mortem analysis [Mirai, Hajime]
  - Automated malware analysis [RIoTMAN]
  - Identification of IoT devices “in the wild” [Haystack]


Internet Security



User Privacy

User Safety

UNIVERSITY OF TWENTE.

SIDN LABS

# So that’s why we selected today’s papers for you

[RIoTMAN] A. Darki, and M. Faloutsos, “RIoTMAN: a systematic analysis of IoT malware behavior”, CoNEXT ’20: Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies, November 2020

[Haystack] S.J. Saidi, A.M. Mandalari, R. Kolcun, H. Haddadi, D.J. Dubois, D. Choffnes, G. Smaragdakis, and A. Feldmann, “A Haystack Full of Needles: Scalable Detection of IoT Devices in the Wild”, 20st ACM Internet Measurement Conference (IMC 2020), October 2020

# Today's learning objective

- After the lecture, you will be able to discuss scalable mechanisms to identify IoT endpoints and the behavior of devices that have been infected with a bot/malware

- Contributes to SSI learning goal #1: "Understand IoT concepts and applications, security threats, technical solutions, and a few relevant standardization efforts in the IETF"

# But first: group discussion for a broader perspective

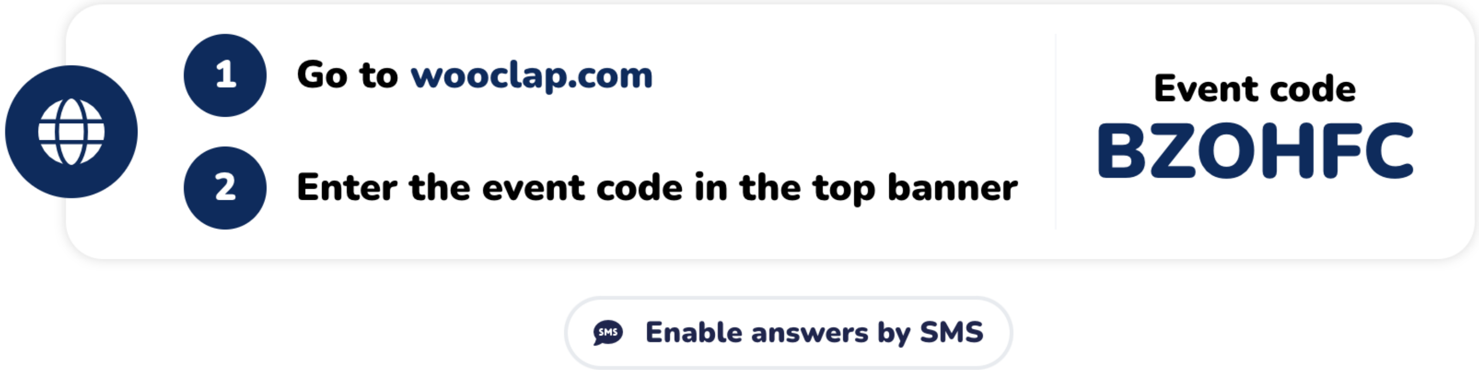
- What **other** mechanisms would players in and outside the IoT ecosystem need to identify IoT endpoints and clean those infected with a bot?
- Think device manufacturers, operators of back-end services, software and hardware engineers, regulators, and so forth
- Split up in groups of around 5 and discuss!
- Take 5 minutes ☺

UNIVERSITY OF TWENTE. SIDN LABS

# “RIoTMAN: a systematic analysis of IoT malware behavior”

16th International Conference on emerging Networking EXperiments and Technologies (CoNEXT), November 2020

# Get your phones ready!



1 Go to wooclap.com

2 Enter the event code in the top banner

Event code

**BZOHFC**

Enable answers by SMS

# What struck you about the paper?

# Challenge: profiling IoT malware

- What needs to be profiled?

- Why is profiling a challenge?

- Why do we need to solve it?

# RIoTMAN: profiling IoT malware binaries

- What's their overall approach?

- What's the advantage of their approach?

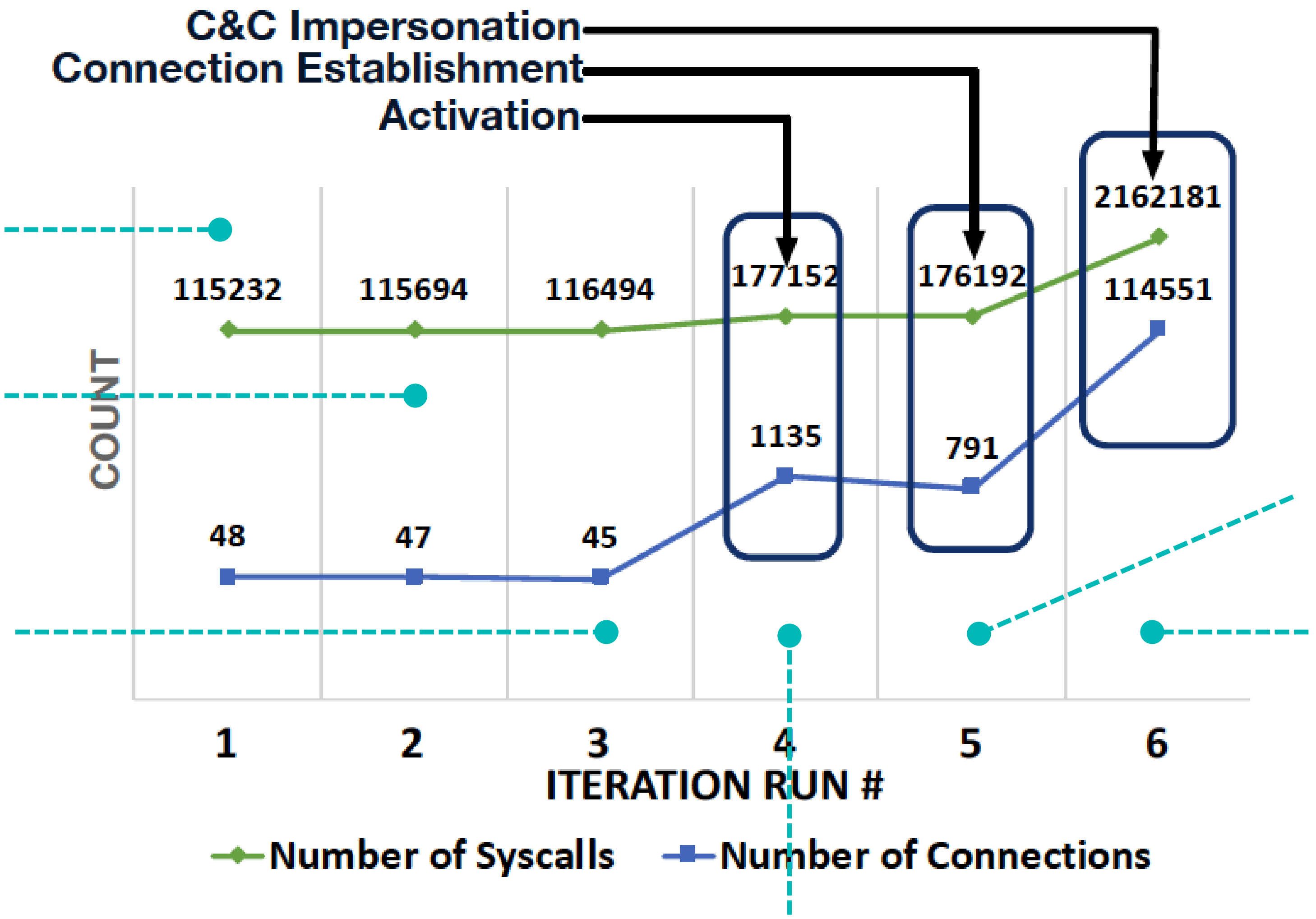
- What malware states does RIoTMAN distinguish?

# Example: Linux.Tsunami

Run 1: uClibc and libnvram libs missing

Run 2: Malware fails to access NVRAM memory

Run 3: Missing source configuration information (a device name in this case) from the simulated NVRAM.

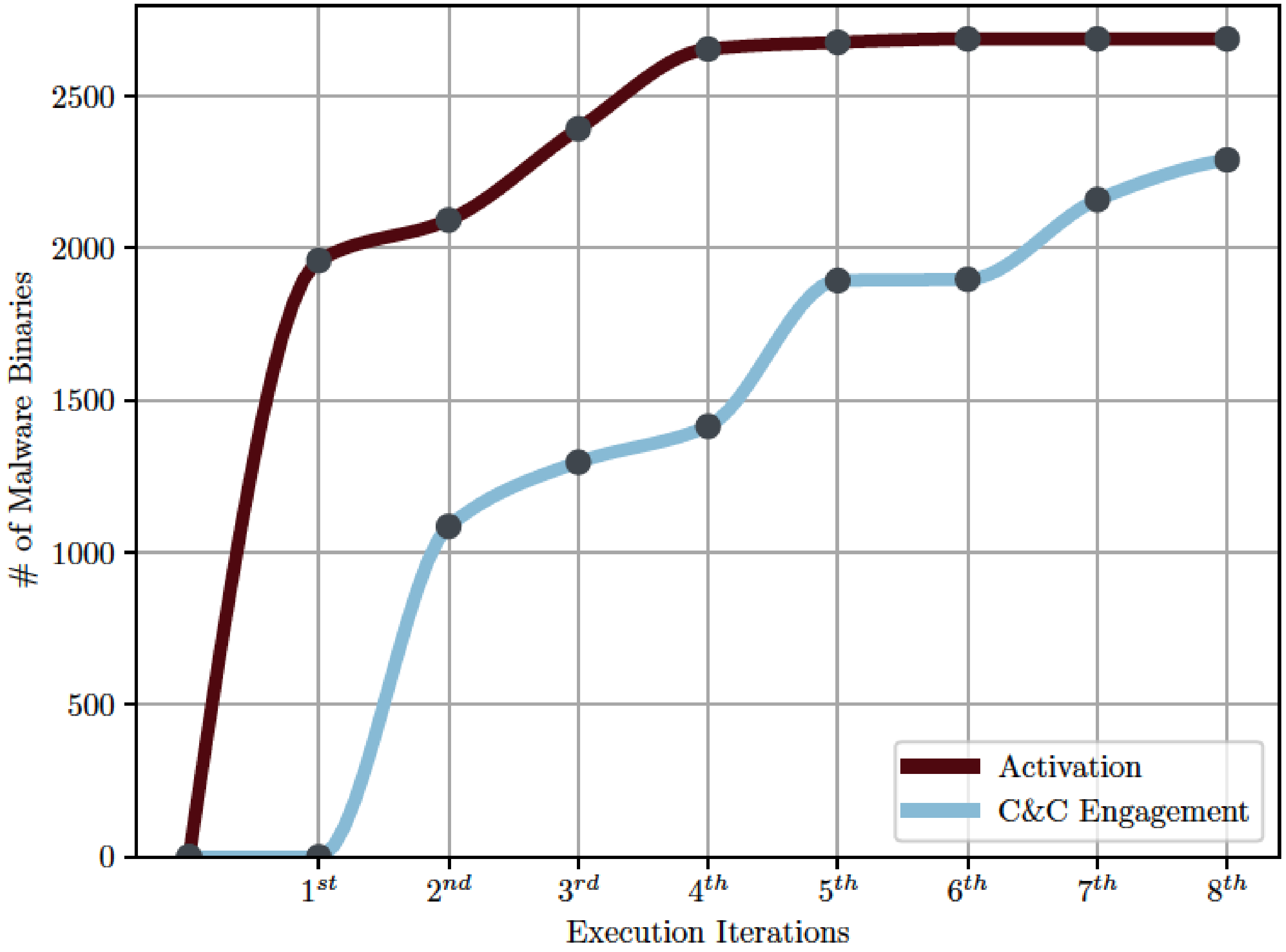


Run 5: malware awaiting instructions

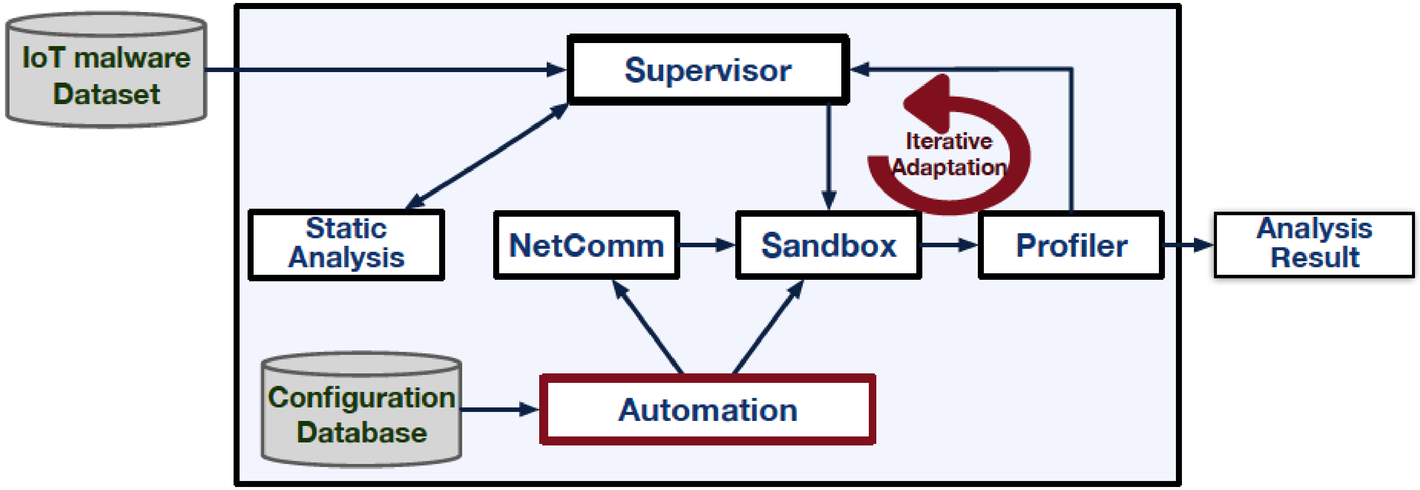
Run 6: successfully sent scan command

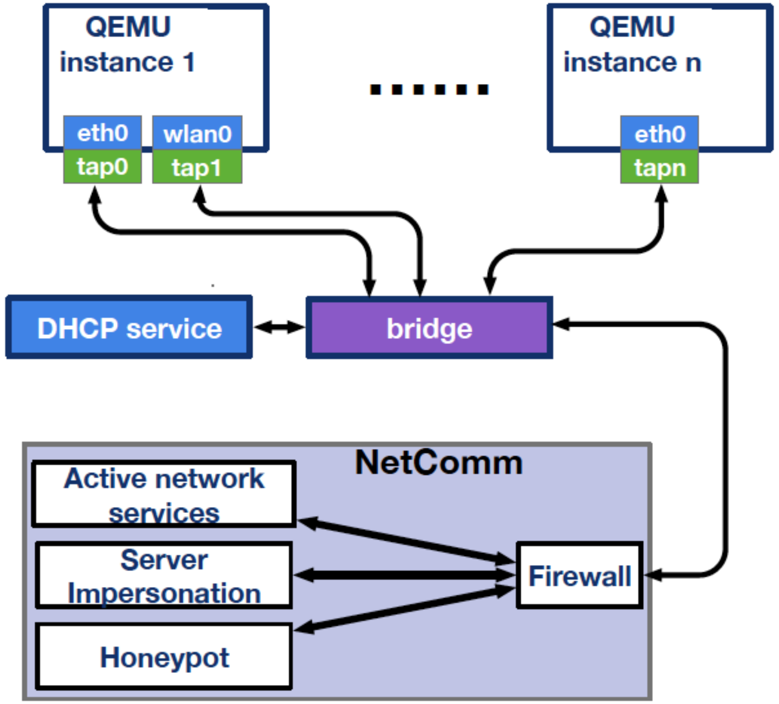
Run 4: malware activated

# Key measurement result – what are we looking at?



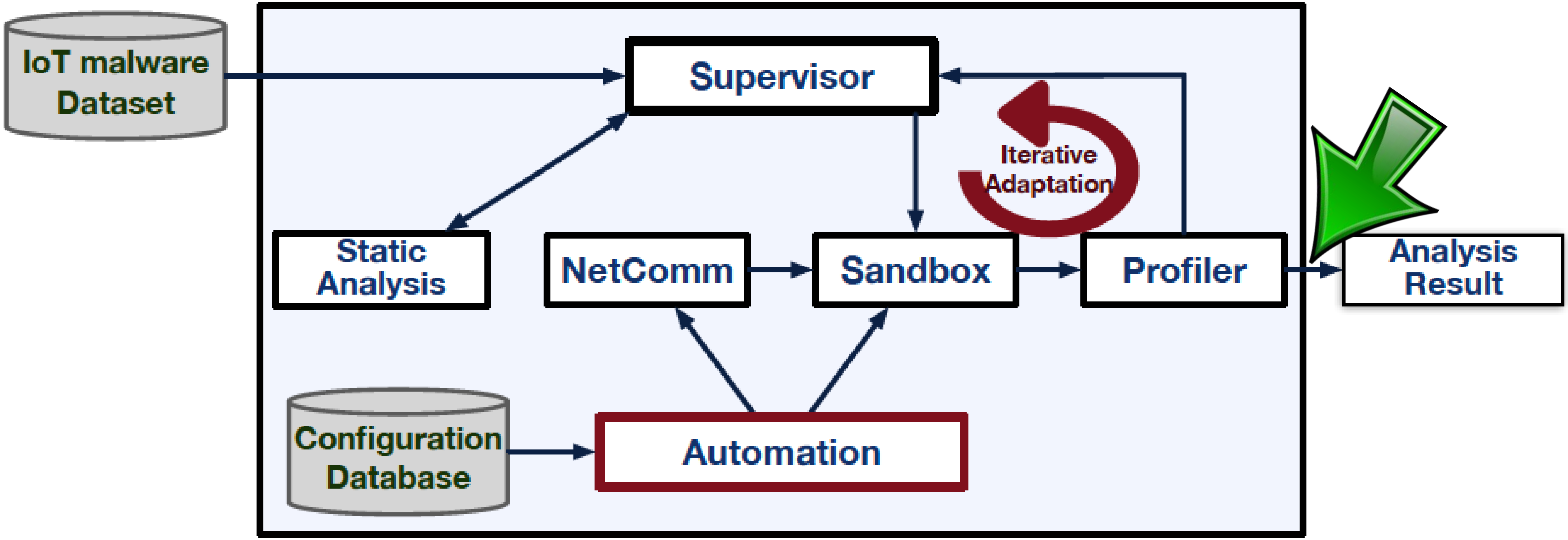
# RIoTMAN measurement architecture



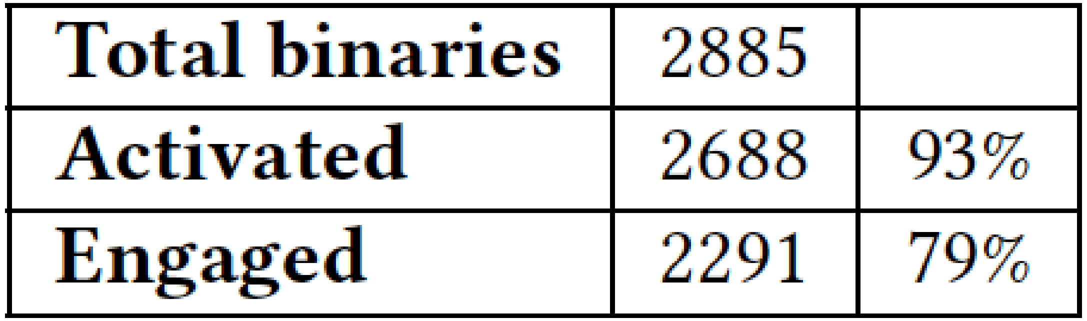


What are the responsibilities of the components?

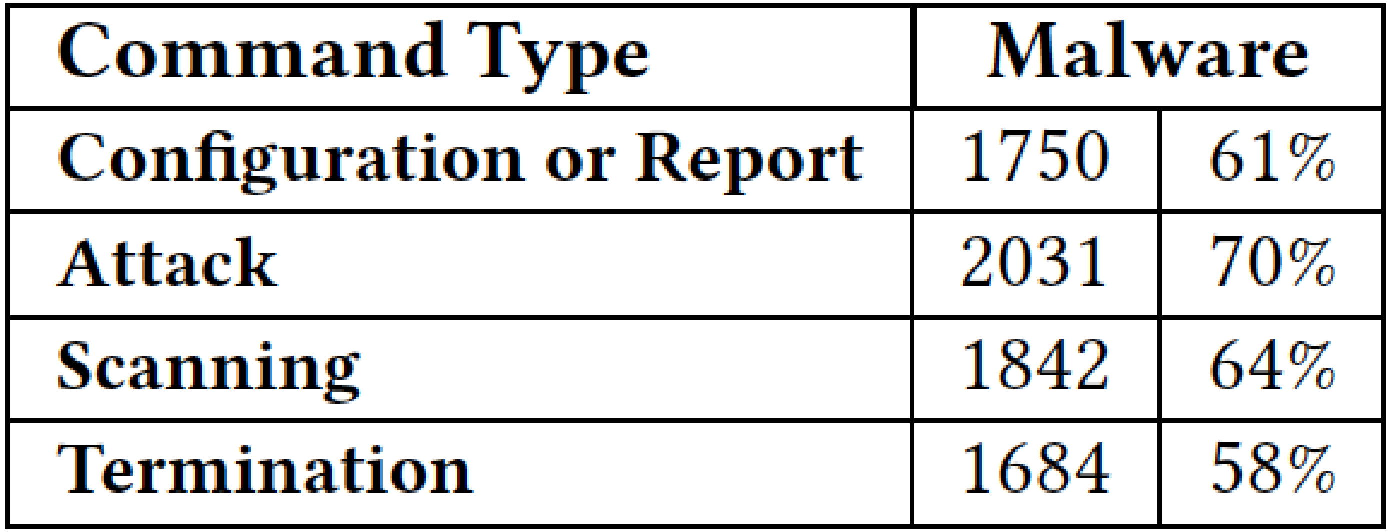
# RIoTMAN profiles



# Measurement results



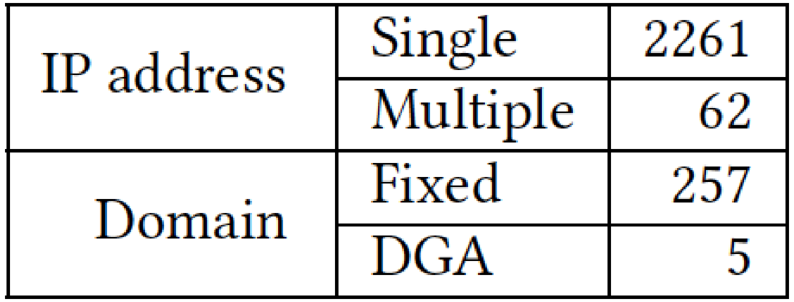
| | | |
|---|---|---|
| Total binaries | 2885 | |
| Activated | 2688 | 93% |
| Engaged | 2291 | 79% |



| Command Type | Malware | |
|---|---|---|
| Configuration or Report | 1750 | 61% |
| Attack | 2031 | 70% |
| Scanning | 1842 | 64% |
| Termination | 1684 | 58% |

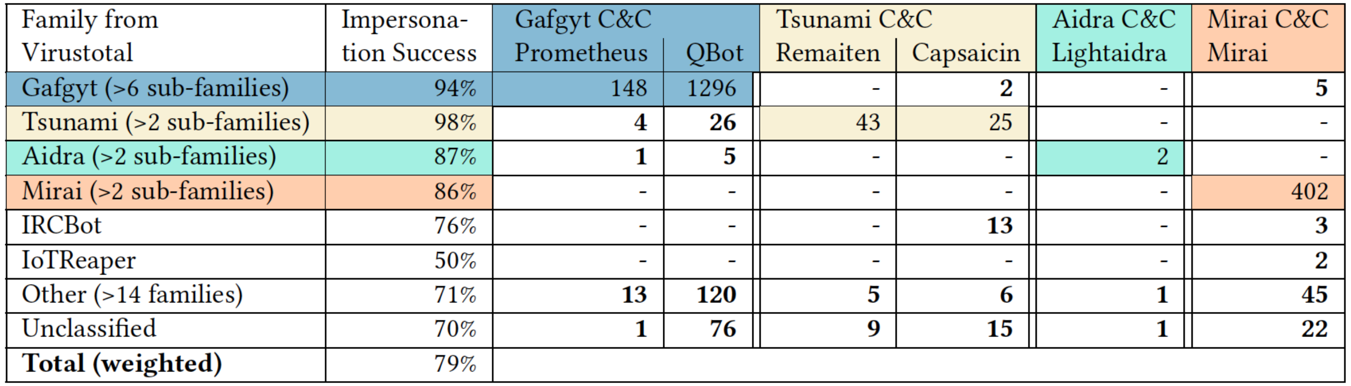
# IoT malware behaviors – how can we leverage that?

C&C discovery

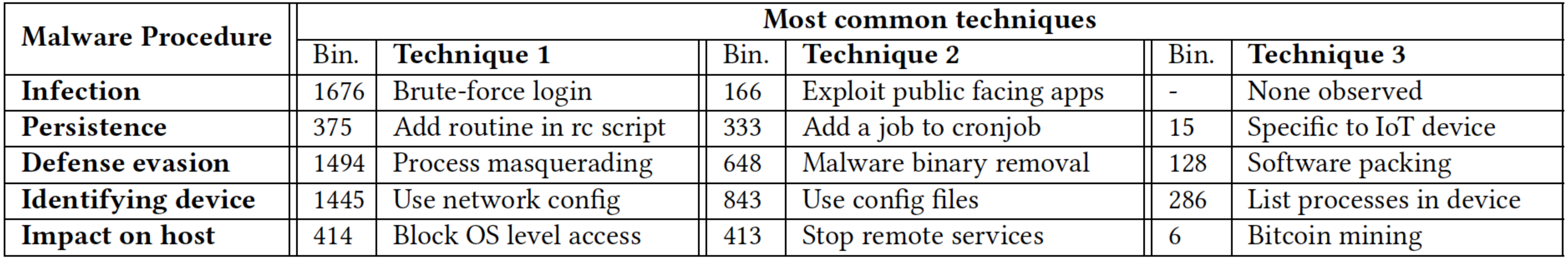


| | | |
|---|---|---|
| IP address | Single | 2261 |
| | Multiple | 62 |
| Domain | Fixed | 257 |
| | DGA | 5 |

Cross-talk in binaries



| Family from Virustotal | Impersonation Success | Gafgyt C&C | | Tsunami C&C | | Aidra C&C | Mirai C&C |
|---|---|---|---|---|---|---|---|
| | | Prometheus | QBot | Remaiten | Capsaicin | Lightaidra | Mirai |
| Gafgyt (>6 sub-families) | 94% | 148 | 1296 | - | 2 | - | 5 |
| Tsunami (>2 sub-families) | 98% | 4 | 26 | 43 | 25 | - | - |
| Aidra (>2 sub-families) | 87% | 1 | 5 | - | - | 2 | - |
| Mirai (>2 sub-families) | 86% | - | - | - | - | - | 402 |
| IRCBot | 76% | - | - | - | 13 | - | 3 |
| IoTReaper | 50% | - | - | - | - | - | 2 |
| Other (>14 families) | 71% | 13 | 120 | 5 | 6 | 1 | 45 |
| Unclassified | 70% | 1 | 76 | 9 | 15 | 1 | 22 |
| **Total (weighted)** | 79% | | | | | | |



| **Malware Procedure** | **Most common techniques** | | | | | |
|---|---|---|---|---|---|---|
| | Bin. | **Technique 1** | Bin. | **Technique 2** | Bin. | **Technique 3** |
| **Infection** | 1676 | Brute-force login | 166 | Exploit public facing apps | - | None observed |
| **Persistence** | 375 | Add routine in rc script | 333 | Add a job to cronjob | 15 | Specific to IoT device |
| **Defense evasion** | 1494 | Process masquerading | 648 | Malware binary removal | 128 | Software packing |
| **Identifying device** | 1445 | Use network config | 843 | Use config files | 286 | List processes in device |
| **Impact on host** | 414 | Block OS level access | 413 | Stop remote services | 6 | Bitcoin mining |

Advanced behaviors

# Limitations

- Linux-based IoT devices only

- They exclude botnets that use encryption, P2P botnets, and IPv6 communications

# Key takeaways

- Dynamic analysis of IoT malware, limited manual effort
- Important to understand, detect, and mitigate IoT botnets at scale
- One piece of the "IoT botnet mitigation puzzle"
- Significant amount of work in terms of engineering, finding datasets, and analysis
- Next challenge: how will RIoTMAN-like systems work in practice (higher TRLs)?

# Coffee break

# “A Haystack Full of Needles: Scalable Detection of IoT Devices in the Wild”

Internet Measurement Conference (IMC 2020)

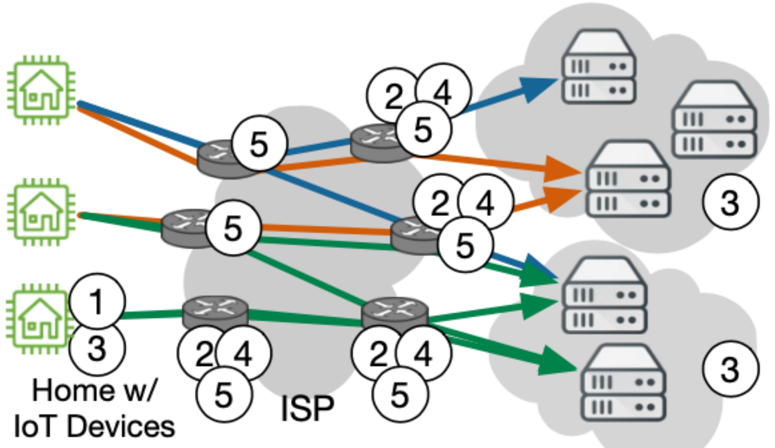
# What struck you about the paper?

# Your opinion

1. What is the paper about?
2. Why is it important to identify IoT devices?
3. How might the takeaways of this study influence future research or industry practices?

# Group Discussion - The Three Parts

How can you replicate this methodology? Why is it scalable?



1 Generate and capture ground truth (GT) IoT traffic in the labs and household - Section 2

2 Capture GT traffic in ISP Vantage point
Evaluate visibility of GT IoT traffic in the ISP-VP - Section 3

3 Identify IoT domains, service IPs , and port numbers, generate detection rules - Section 4

4 Cross check detection rules by inferring devices on GT data - Section 5

5 Detect IoT devices in the wild - Section 6

**Figure 2: General methodology overview.**

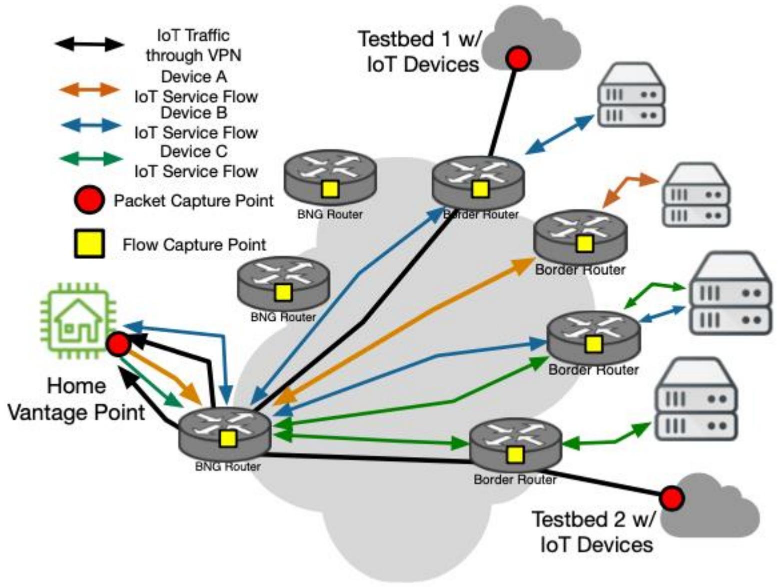


**Figure 3: ISP setup & flow collection points.**

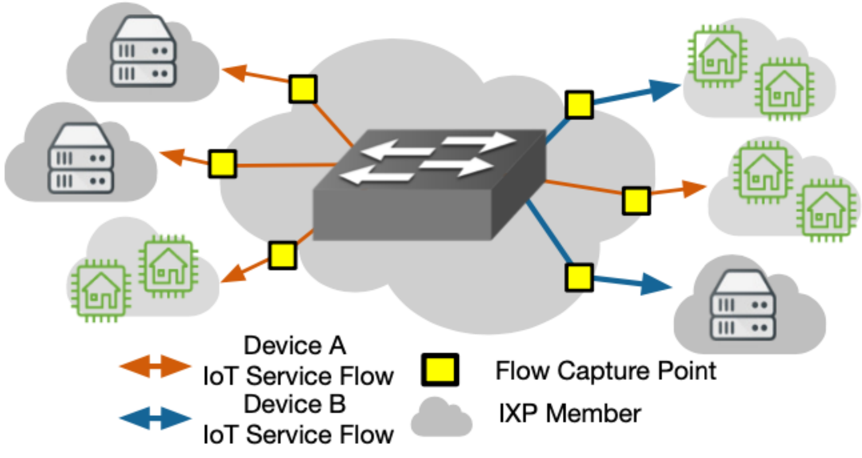
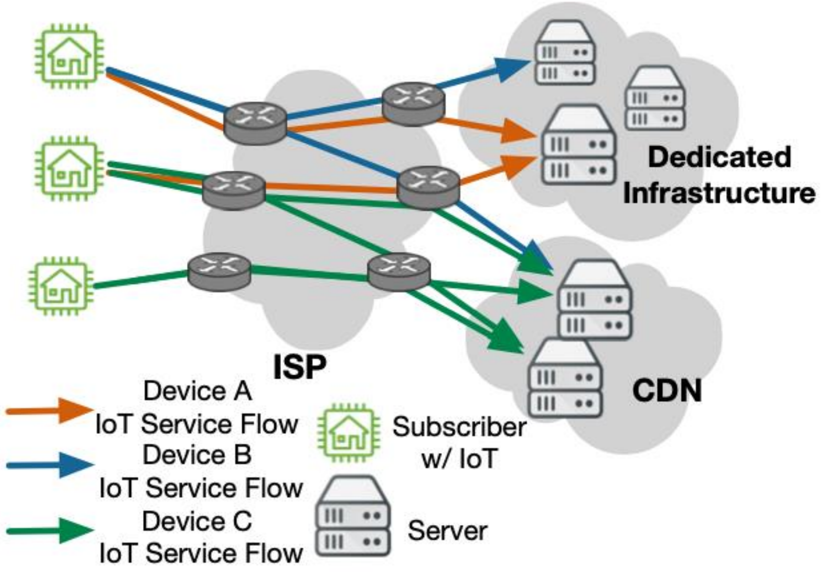


**Figure 4: IXP setup & flow collection points.**

# Scalable detection of IoT devices

The main method of IoT device detection

1. Platform-level
2. Manufacturer-level
3. Product-level



Dedicated Infrastructure
ISP
CDN
Device A IoT Service Flow
Device B IoT Service Flow
Device C IoT Service Flow
Subscriber w/ IoT
Server

Figure 1: Simplified IoT communication patterns.

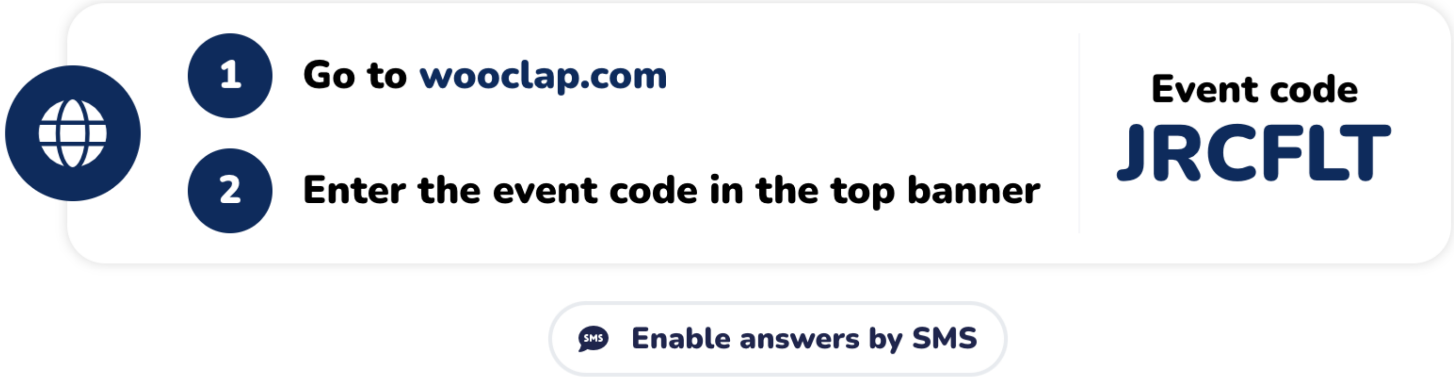
# Controlled experiments

Tunnel traffic to an ISP to establish ground truth.

Why do this? And why exactly like this?

# Get your phones ready!

## How to participate?



1 Go to wooclap.com

2 Enter the event code in the top banner

Event code
JRCFLT

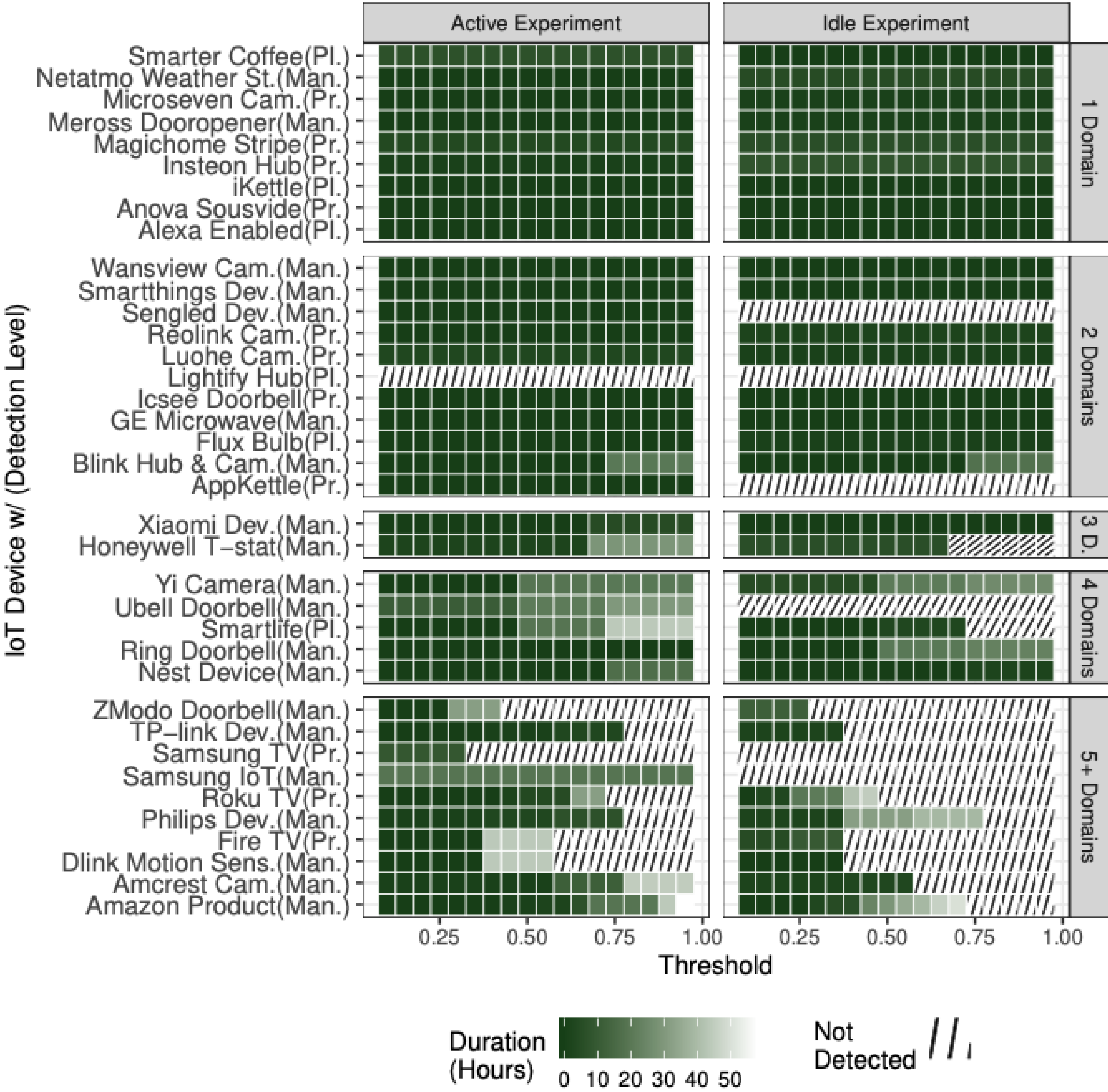
Enable answers by SMS

# Home-VP

Time to detect IoT

Domains per IoT device

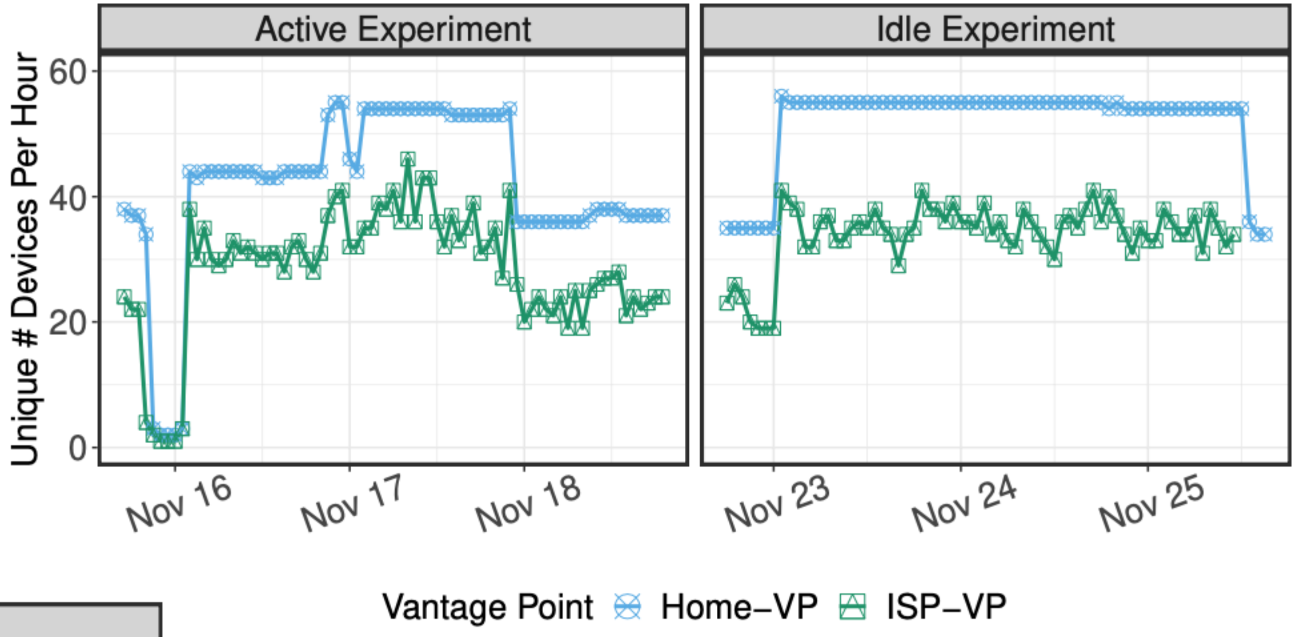
Threshold for detection



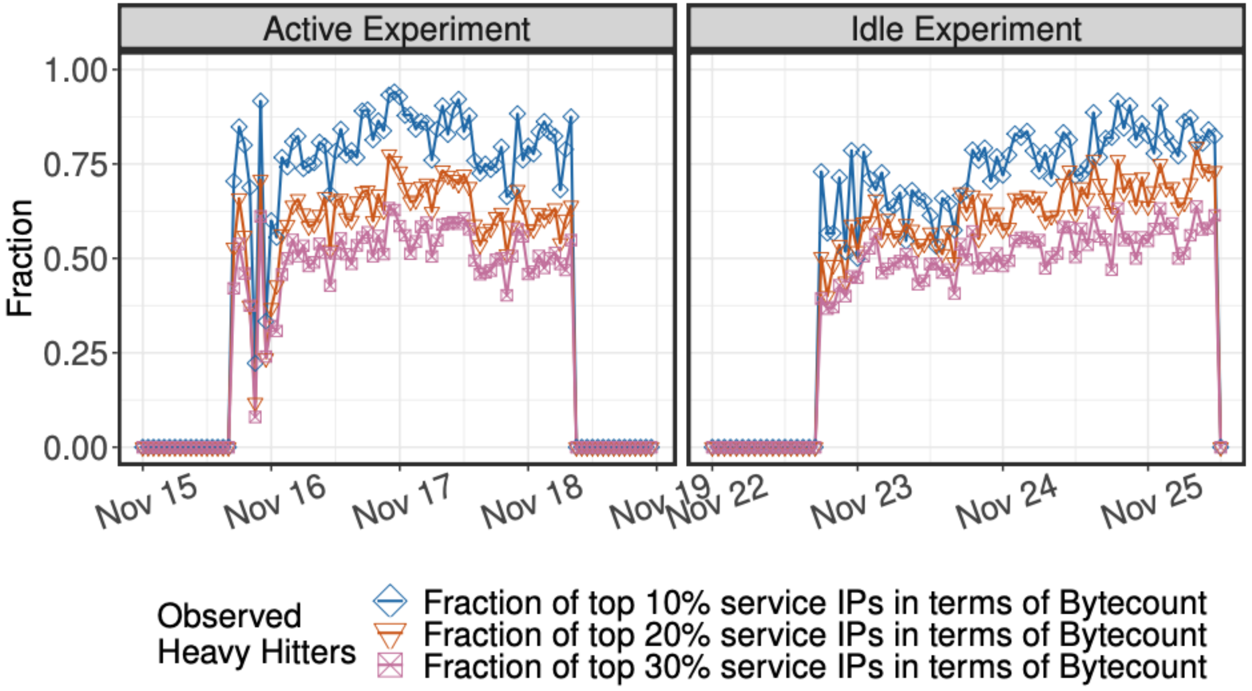
# ISP vantage point

12M subscribers

What can they see?



(d) # Unique IoT devices per hour.



# IXP vantage point

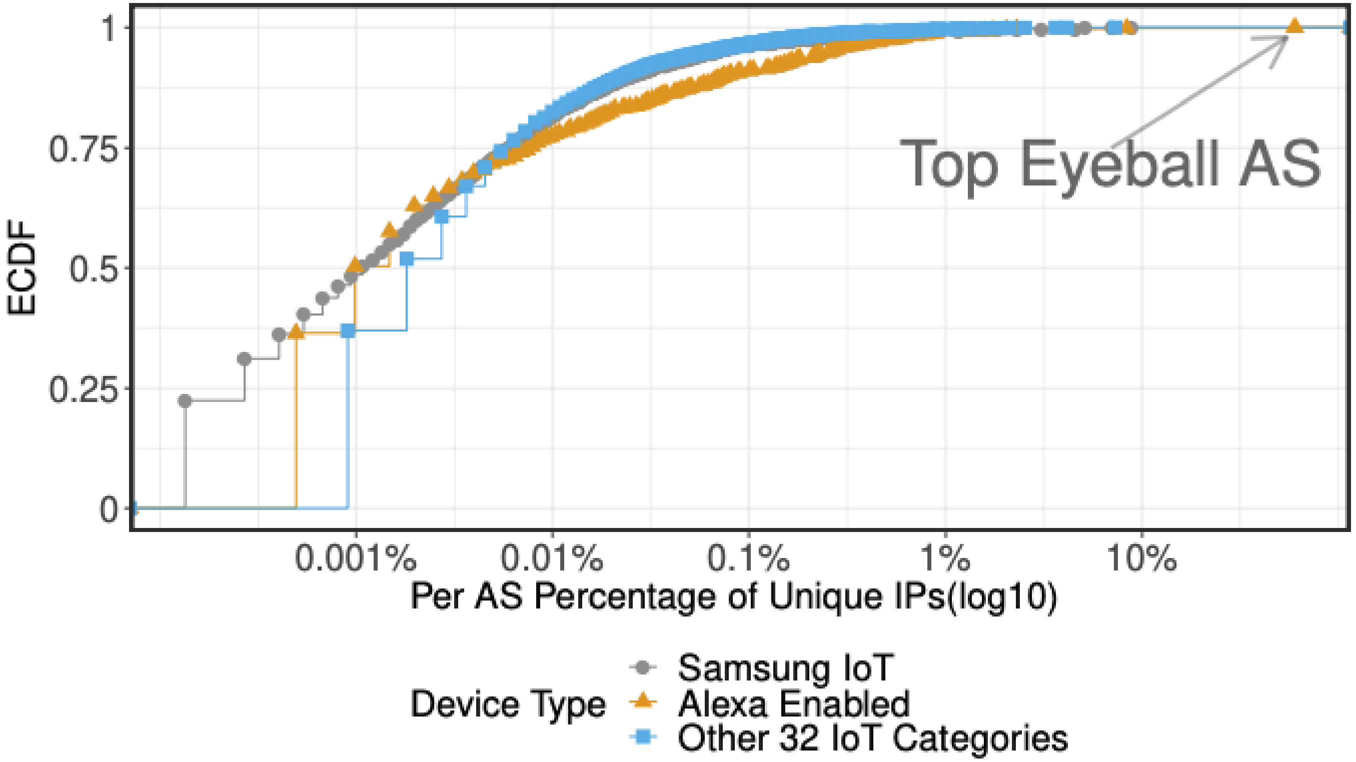


**Figure 16: IXP: ECDF of Per-ASN Percentage (# Unique IPs) - Day 15-11-2020.**

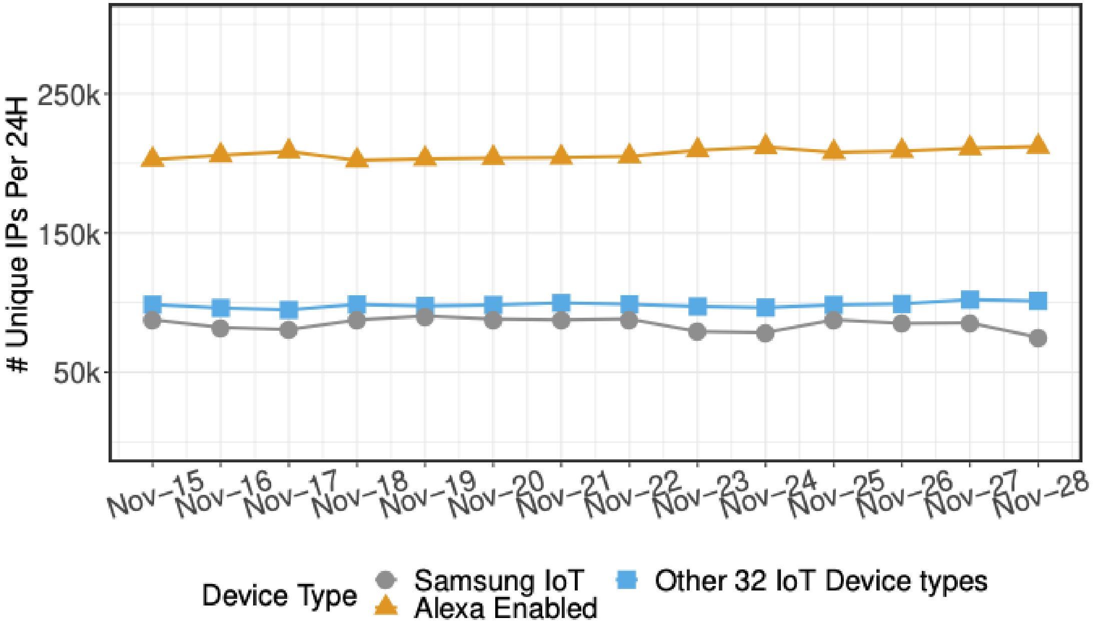


**Figure 15: IXP: Number of Samsung IoT, Alexa Enabled, and Other 32 IoT device types IPs observed/day.**

# Discussion on Security Benefits

“For example, an ISP can use our methodology for redirecting the IoT devices traffic to a new backend infrastructure that offers privacy notices or security patches for devices that are no longer supported by their manufacturers.”

“Moreover, if an IoT device is misbehaving, e.g., if it is involved in network attacks or part of a botnet [31], our methodology can help the ISP/IXP in identifying what devices are common among the subscriber lines with suspicious traffic.”

# Discussion

Our analysis could be simplified if an ISP/IXP had access to all DNS queries and responses. Even having a partial list, e.g., from the local DNS resolver of the ISP, could improve our methodology.

# Key takeaways

- Combining passive and active monitoring techniques to comprehensively detect IoT devices
- 20% of 15 million subscriber lines used at least one of the 56 different IoT products
- Important to understand, detect, and mitigate IoT botnets at scale

**Next regular lecture:**

Wed June 12, 10:45-12:30

Topic: IoT Security in Non-Carpeted Areas

**Dr. Antonia Affinito** | a.affinito@utwente.nl
**Prof. Cristian Hesselman** | c.e.w.hesselman@utwente.nl