Lecture #7: IoT Device Behavior

Cristian Hesselman, Elmer Lastdrager, Ramin Yazdani, and Etienne Khan

University of Twente | June 3, 2020

Lab assignment

- MUD descriptions: you'll need to generate them yourselves, tools are available
- IoT devices: you'll need to work with the actual hardware, no emulations (unless as an extra)
- Use IoT devices without a browser-like interface, such as light bulbs, audio speakers, doorbells
- Do not use multi-purpose devices like tablets, phones, laptops
- At least 2 IoT devices per group of 3 and at least 3 devices per group of 4
- Etienne Khan available for assistance

UNIVERSITY OF TWENTE.

SIDN LABS

Lab assignment (2)

- If you have not yet started, start today!

- Don't overcomplicate recording: a laptop suffices.

UNIVERSITY OF TWENTE.

SIDN LABS

Paper summaries

- You must have handed in your two summaries BEFORE this lecture
- You can use the summaries during the oral exam (“open book”)
- You **cannot** complete SSI without submitting 12 paper summaries!

Interactive Lecture

- Goal: enable you to learn from each other and further increase your understanding of the papers (contributes to preparing yourself for the oral exam)

- Format:
  1. We'll ask someone to provide **their opinion** of the paper
  2. A summary by teachers (put any questions in the chat)
  3. Questions: discussion starters and fact questions
  4. Discussion (use your mic)
  5. We may ask someone specific to start the discussion

- Experimental format resulting from Corona pandemic, please provide feedback!

Today's papers

Are about measuring IoT botnets

- [**AuDI**] Marchal, S., Miettinen, M., Nguyen, T. D., Sadeghi, A-R., & Asokan, N. (Accepted/In press). AuDI: Towards Autonomous IoT Device-Type Identification using Periodic Communication. IEEE Journal on Selected Areas in Communications
- [**IMC**] J. Ren, D. J. Dubois, D. Choffnes, A. M. Mandalari, R. Kolcun, and H. Haddadi, "Information Exposure from Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach", Internet Measurement Conference (IMC2019), Amsterdam, Netherlands, Oct 2019

“AUDI: Towards Autonomous IoT Device-Type Identification using Periodic Communication”

Passive monitoring

Encryption-agnostic

See also:

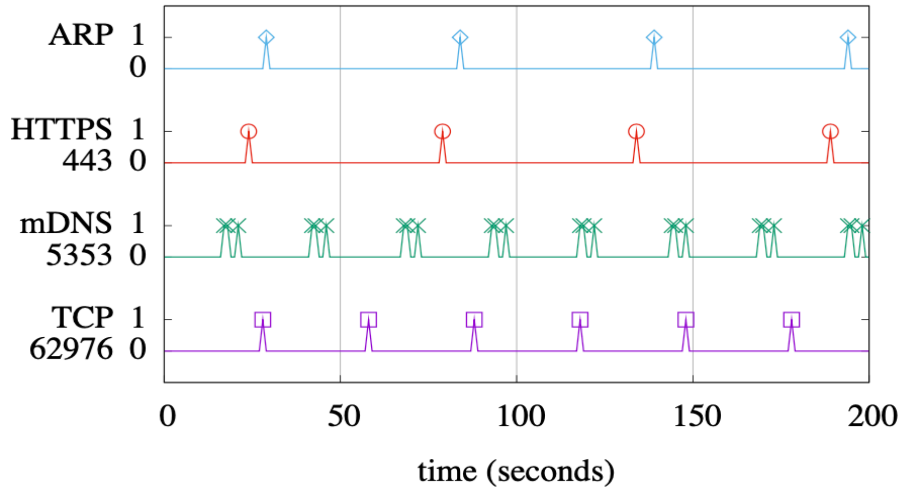
Noah Apthorpe, Dillon Reisman, Nick Feamster, "**A Smart Home is No Castle: Privacy Vulnerabilities of Encrypted IoT Traffic**", Workshop on Data and Algorithmic Transparency (DAT '16), New York University Law School, November 2016

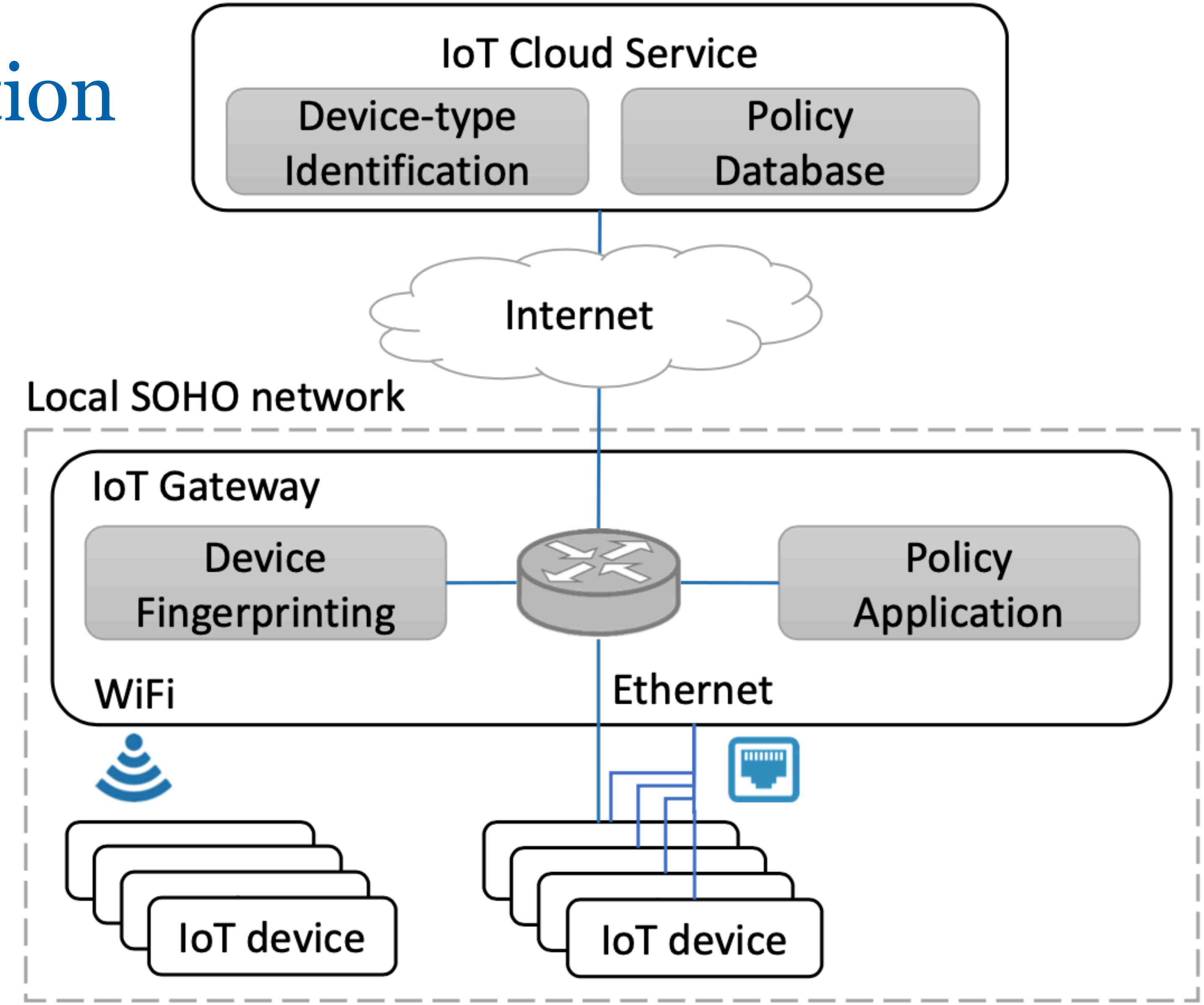
UNIVERSITY OF TWENTE.

SIDN LABS

Device Type identification

- Goal: “quickly, accurately and *autonomously* identifying the type of IoT devices”
- QoS or security policies
- Passive fingerprinting of periodic network traffic
- 98.2% accuracy in tests





UNIVERSITY OF TWENTE.

SIDN LABS

How do they do it?

- Periodic background network traffic
- Analyse per flow
- Time series: traffic 1/0 every second
- Compute periods Fourier transform
- Autocorrelation to find periodicity
- Fingerprinting periods

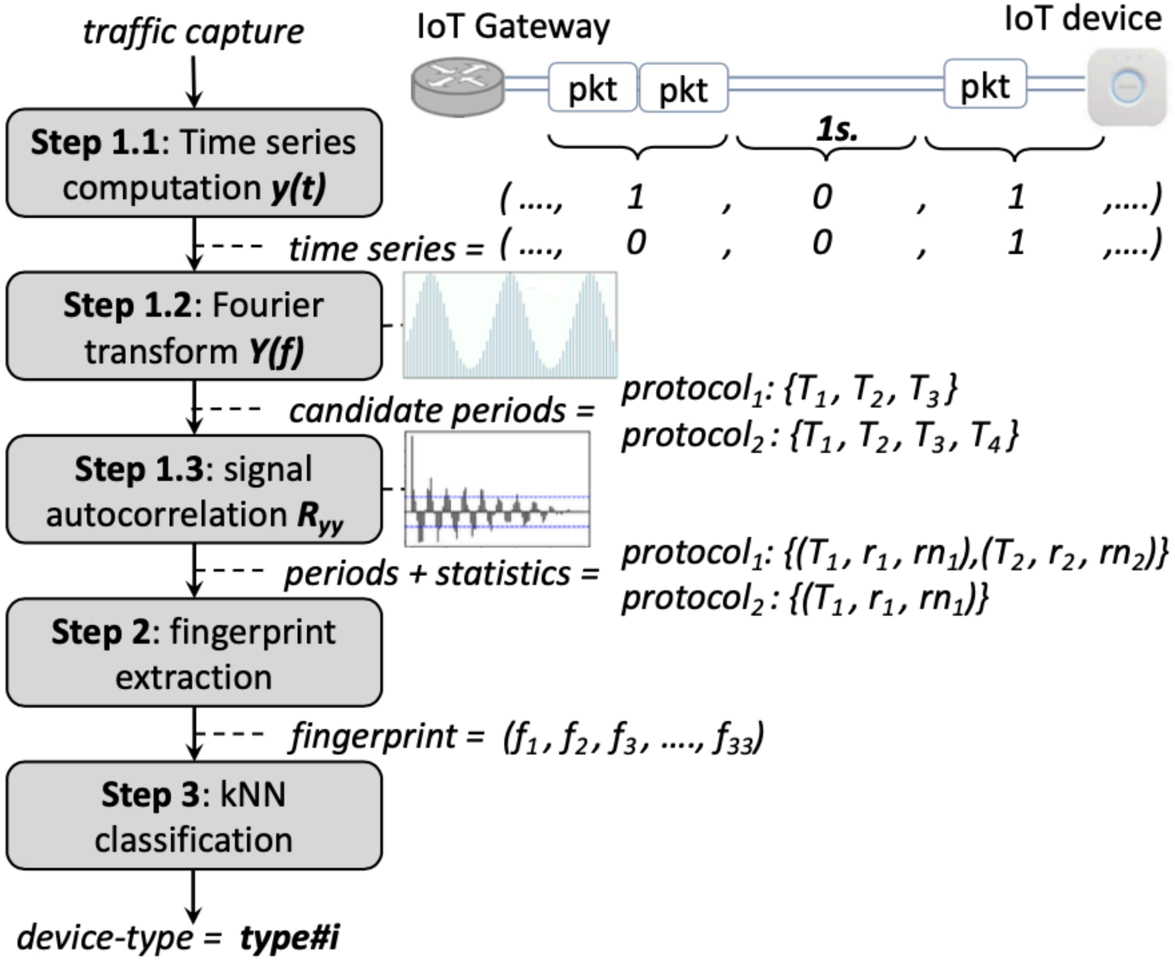


Fig. 2: Overview of device-type identification.

UNIVERSITY OF TWENTE. SDN LABS

Quiz: countering detection

How can you avoid getting fingerprinted?

A. Generate a constant stream of traffic
B. Encrypt the network traffic
C. Open connections to random hosts
D. Disable the ICMP finger protocol
E. You can’t

Fingerprints

33 features in 4 categories

Manually designed

Category	f	Description	Importance
periodic flows	1	# periodic flows	0.440
	2	# periodic flows (protocol $\leq$ layer 4)	0.465
	3	Mean periods per flow	0.068
	4	SD periods per flow	0.037
	5	# flows having only one period	0.429
	6	# flows having multiple periods	0.176
	7	# flows with static source port	0.533
	8	Mean frequency source port change	0.310
	9	SD frequency source port change	0.137
period accuracy	10	# periods inferred in all sub-captures	0.329
	11	Mean period inference success	0.037
	12	SD period inference success	0.022
period duration	13	# periods $\in [5s.; 29s.]$	0.409
	14	# periods $\in [30s.; 59s.]$	0.408
	15	# periods $\in [60s.; 119s.]$	0.467
	16	# periods $\in [120s.; 600s.]$	0.419
period stability	17	# Mean(r) $\in [0.2; 0.7[$	0.386
	18	# Mean(r) $\in [0.7; 1[$	0.436
	19	# Mean(r) $\in [1; 2[$	0.239
	20	# Mean(r) $\in [2; +\infty[$	0.124
	21	# SD(r) $\in [0; 0.02[$	0.185
	22	# SD(r) $\in [0.02; 0.1[$	0.151
	23	# SD(r) $\in [0.1; +\infty[$	0.185
	24	# Mean(rn) $\in [0.2; 0.7[$	0.288
	25	# Mean(rn) $\in [0.7; 1[$	0.307
	26	# Mean(rn) $\in [1; 2[$	0.313
	27	# Mean(rn) $\in [2; +\infty[$	0.246
	28	# SD(rn) $\in [0; 0.02[$	0.217
	29	# SD(rn) $\in [0.02; 0.1[$	0.217
	30	# SD(rn) $\in [0.1; +\infty[$	0.220
	31	# Mean(rn)$-$ Mean(r) $\in [0; 0.02[$	0.408
	32	# Mean(rn)$-$ Mean(r) $\in [0.02; 0.1[$	0.248
	33	# Mean(rn)$-$ Mean(r) $\in [0.1; +\infty[$	0.482

IoT Cloud service

- Fingerprints are sent to IoT Cloud service
- Cloud services uses fingerprints to learn (and find) device types (i.e., step 3)
- Fingerprints per 30 minutes.
- Unsupervised (?) clustering algorithm: autonomously group these fingerprints into clusters and create an abstract label for each cluster

Evaluation

33 devices

Background + activity

6224 fingerprints

ID in +- 30 minutes

Device-type	Identifier	Device model	WiFi	Ethernet	Other	Background	Activity
type#01	ApexisCam	Apexis IP Camera APM-J011	●	●	○	●	●
type#02	CamHi	Cooau Megapixel IP Camera	●	●	○	●	●
type#03	D-LinkCamDCH935L	D-Link HD IP Camera DCH-935L	●	○	○	●	●
type#04	D-LinkCamDCS930L	D-Link WiFi Day Camera DCS-930L	●	●	○	●	○
	D-LinkCamDCS932L	D-Link WiFi Camera DCS-932L	●	●	○	●	○
type#05	D-LinkDoorSensor	D-Link Door & Window sensor	○	○	●	●	●
	D-LinkSensor	D-Link WiFi Motion sensor DCH-S150	●	○	○	●	●
	D-LinkSiren	D-Link Siren DCH-S220	●	○	○	●	●
	D-LinkSwitch	D-Link Smart plug DSP-W215	●	○	○	●	●
	D-LinkWaterSensor	D-Link Water sensor DCH-S160	●	○	○	●	●
type#06	EdimaxCamIC3115	Edimax IC-3115W Smart HD WiFi Network Camera	●	●	○	●	●
	EdimaxCamIC3115(2)	Edimax IC-3115W Smart HD WiFi Network Camera	●	●	○	●	●
type#07	EdimaxPlug1101W	Edimax SP-1101W Smart Plug Switch	●	○	○	●	●
	EdimaxPlug2101W	Edimax SP-2101W Smart Plug Switch	●	○	○	●	●
type#08	EdnetCam	Ednet Wireless indoor IP camera Cube	●	●	○	●	●
type#09	EdnetGateway	Ednet.living Starter kit power Gateway	●	○	●	●	●
type#10	HomeMaticPlug	Homematic pluggable switch HMIP-PS	○	○	●	●	●
type#11	Lightify	Osram Lightify Gateway	●	○	●	●	●
type#12	SmcRouter	SMC router SMCWBR14S-N4 EU	●	●	○	●	●
type#13	TP-LinkPlugHS100	TP-Link WiFi Smart plug HS100	●	○	○	●	●
	TP-LinkPlugHS110	TP-Link WiFi Smart plug HS110	●	○	○	●	●
type#14	UbnTAirRouter	Ubnt airRouter HP	●	●	○	●	●
type#15	WansviewCam	Wansview 720p HD Wireless IP Camera K2	●	○	○	●	●
type#16	WeMoLink	WeMo Link Lighting Bridge model F7C031vf	●	○	●	●	●
type#17	WeMoInsightSwitch	WeMo Insight Switch model F7C029de	●	○	○	●	●
	WeMoSwitch	WeMo Switch model F7C027de	●	○	○	●	●
type#18	HueSwitch	Philips Hue Light Switch PTM 215Z	○	○	●	●	●
type#19	AmazonEcho	Amazon Echo	●	○	○	○	●
type#20	AmazonEchoDot	Amazon Echo Dot	●	○	○	○	●
type#21	GoogleHome	Google Home	●	○	○	●	○
type#22	Netatmo	Netatmo weather station with wind gauge	●	○	●	●	○
type#23	iKettle2	Smarter iKettle 2.0 water kettle SMK20-EU	●	○	○	●	●
	SmarterCoffee	Smarter SmarterCoffee coffee machine SMC10-EU	●	○	○	●	●

Quiz: attack!

Devices can spoof their fingerprint. How do the authors counter this?

A. The gateway will detect this thanks to the ReliefF feature selection

B. They propose to add active scanning as future work

C. Add the device’s MAC address as a feature

D. They assume that the device is not infected during the first 30 minutes

Discussion

- Privacy implications?
- Sharing policies with central cloud service
- Fingerprinting attack traffic?

UNIVERSITY OF TWENTE.

SIDN LABS

Information Exposure From Consumer IoT Devices

A Multidimensional, Network-Informed Measurement Approach

UNIVERSITY OF TWENTE.

SIDN LABS

Motivation

IoT devices are the new normal (+7.000.000.000 devices around us)

- But don't just take my word for it, take Bosch's
  - https://www.youtube.com/watch?v=v2kV6pgJxu0

But time and time again we have seen that:

- IoT cameras might record you in unexpected scenarios
- IoT assistants might activate/record unexpectedly
- IoT TVs show you ads in your launcher/menu
  - https://www.thedrum.com/news/2019/09/10/the-first-thing-you-see-lg-smart-tv-now-ad

UNIVERSITY OF TWENTE. SIDN LABS

Expectations

- My IoT device only connects to the server of the manufacturer
- My IoT device only transmits its data in an encrypted fashion
- My IoT device only transmits relevant data to the manufacturer
- My IoT device only does its IoT task when I ask it to do so
- My IoT device purchased in my region, won't connect to any other jurisdiction

- Quick question: Do you have any additional expectations?

Data Collection Methodology

- Well to see if our expectations hold true, lets put them to the test
- 81 different IoT devices in two different jurisdictions: UK and US
- All traffic is captured at a central server before egressing into the Internet

- But how do we test? As we've seen before, there is no standard IoT testbed.
- How do we test smart assistants?

UNIVERSITY OF TWENTE. SIDN LABS

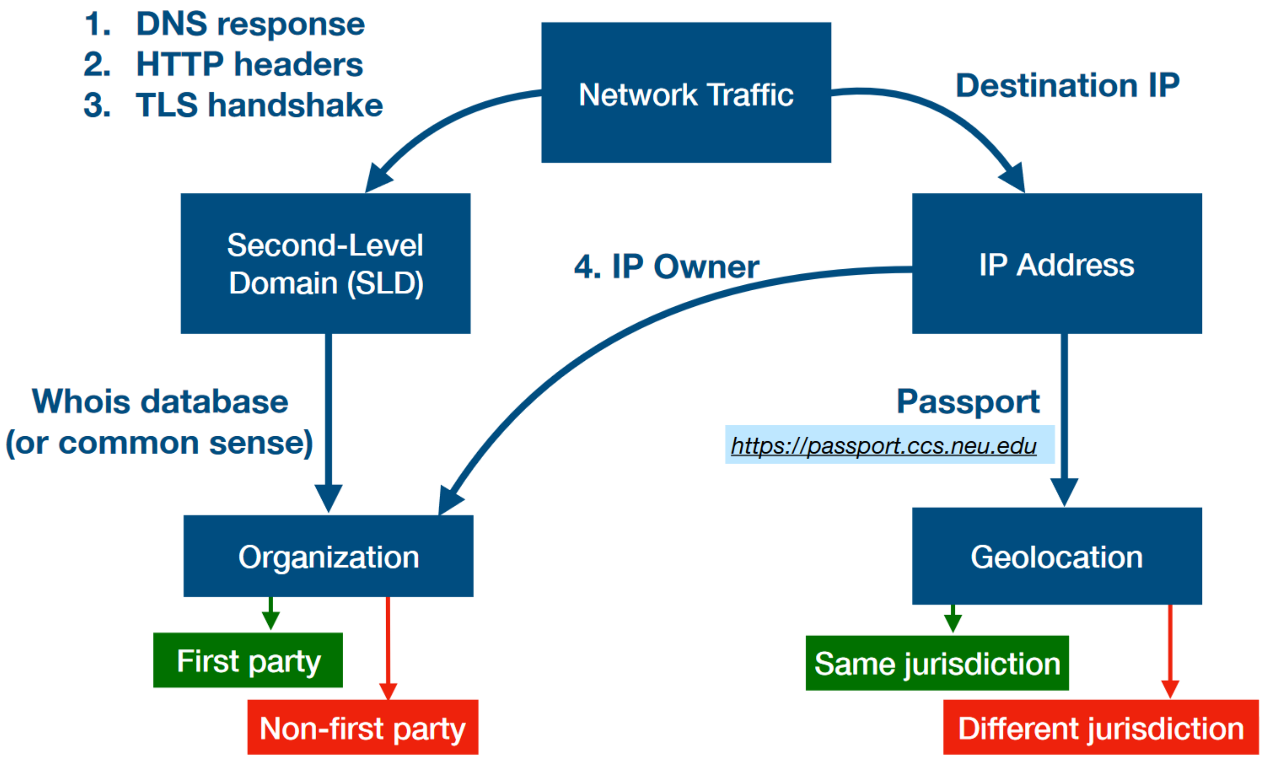
Data Collection Methodology

34,586 experiments (92.6% automated)

- **Controlled interactions**
  - Manual (repeated 3 times)
  - Automated (repeated 30 times)
    - Text-to-speech to smart assistants (Alexa/Google/Cortana/Bixby)
    - Monkey instrumented control from Android companion apps

Activity	Description
Power	power on/off the device
Voice	voice commands for speakers
Video	record/watch video
On/Off	turn on/off bulbs/plugs
Motion	move in front of device
Others	change volume, browse menu

Destination Analysis



- Number of devices contacting non-first party organizations

High reliance on cloud and CDN providers

Nearly all TVs contact Netflix w/o it being logged in or used

Chinese cloud providers

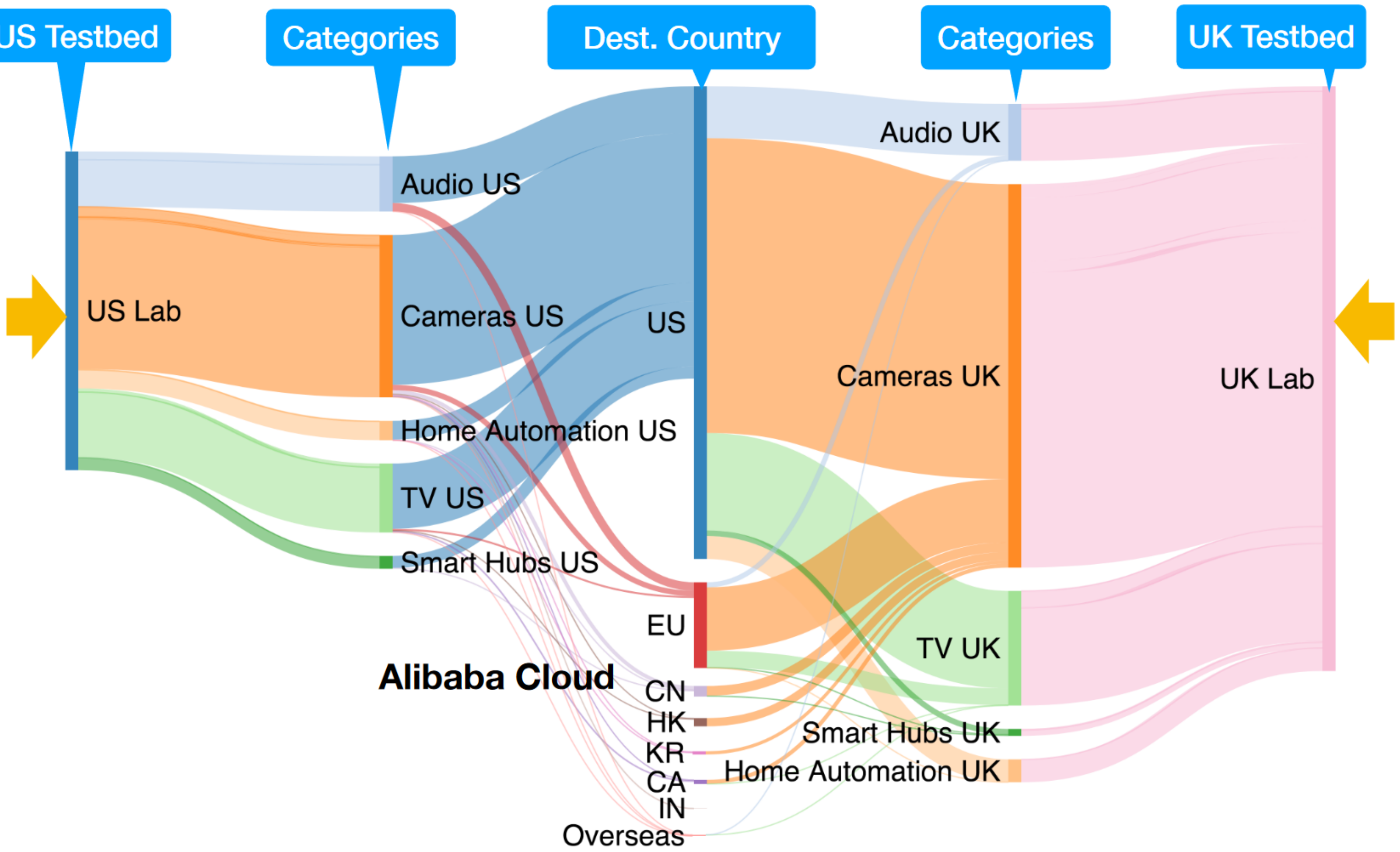
Organization	US 46	UK 35	US Common 24	UK Common 24
Amazon	31	24	16	17
Google	14	9	10	8
Akamai	10	6	6	5
Microsoft	6	4	1	1
Netflix	4	2	3	2
Kingsoft	3	3	1	1
21Vianet	3	3	1	1
Alibaba	3	4	2	2
Beijing Huaxiay	3	3	1	1
AT&T	2	0	1	1

Regional differences

UNIVERSITY OF TWENTE.

SIDN LABS

Destination Analysis



UNIVERSITY OF TWENTE.

SIDN LABS

Encryption Analysis

- Remove everything which is not detected by Wireshark as TLS or QUIC
- Get a baseline entropy for HTTP (0.25) and HTTPS (TLS) (0.85) traffic
- But depending on the content (IMC 2019 websites) you might get different results:
    - HTTP (0.55, max = 0.62) / fernet (0.73, min = 0
- Suddenly the picture isn't so clear anymore?
- Open discussion: What do you think the unidentified traffic might be?
- Open discussion: Shouldn't MITM analysis be deployed as well?

Enc	Range	US	UK	US∩	UK∩	VPN US→UK	VPN UK→US	VPN US∩	VPN UK∩
✗	>75	0	0	0	0	0	0	0	0
	50-75	1	1	0	0	2	0	1	0
	25-50	4	1	1	1	3	2	0	1
	<25	41	31	24	24	41	31	24	24
✓	>75	7	7	5	5	4	5	3	3
	50-75	5	7	4	6	7	8	5	7
	25-50	10	5	5	4	12	5	7	5
	<25	24	14	11	10	23	15	10	10
?	>75	16	10	8	7	17	11	8	7
	50-75	11	6	5	5	11	5	5	4
	25-50	11	7	6	5	13	10	8	9
	<25	8	10	6	8	5	7	4	5

Table 5: Number of devices by encryption percentage in quartile groups across lab and network.

UNIVERSITY OF TWENTE.

SIDN LABS

Quiz: *Which of these items was <u>not</u> sent unencrypted?*

A. MAC address (incl. timestamp)
B. Firmware
C. On/Off signals
D. Video
E. Voice data

UNIVERSITY OF TWENTE.
SIDN LABS

Unexpected Behavior

Popular doorbells

Video recording on detected motion (cannot be disabled)

Popular smart TVs

Contact **Netflix**, **Google**, and **Facebook** unexpectedly

Alexa-enabled devices

Frequently falsely triggered (e.g. "**I like S**tar Trek")

- Other notable cases of activities detected when idle
  - Cameras reporting **motion** in absence of movement
  - Devices spontaneously **restarting** or reconnecting

UNIVERSITY OF TWENTE.

SIDN LABS

Conclusion

- **First step towards more large-scale IoT measurements:**
  - 81 devices, 2 countries, 34K experiments
- **Main results:**
  - 57% (50%) of destinations of the US (UK) devices are not first-party
  - 56% (84%) of the US (UK) devices have at least one destination abroad
  - 89% (86%) of the US (UK) devices are vulnerable to at least one activity inference
  - Activity inference can be used to identify *unexpected* activities
- **Impact:**
  - Press coverage
  - Working with manufacturers to understand information exposure
  - **Testbed/analysis framework and data are publicly available**

https://moniotrlab.ccis.neu.edu/imc19/

Discussion (if time permits)

- We heard one opinion at the beginning, maybe some more?
- How would you improve this study?
- Can we say anything about the long-term feasibility of projects like these?

Volg ons

SIDN.nl

@SIDN

SIDN

Other Questions or Feedback?

Next lecture: **Wed June 10, 10:45-12:30**
Topic: IoT Network Security

UNIVERSITY OF TWENTE.

SIDN LABS