Deep Dive into the IoT Backend Ecosystem

Said Jawad Saidi, Srdjan Matic, Oliver Gasser, Georgios Smaragdakis, Anja Feldmann

max planck institut informatik

UNIVERSITÄT DES SAARLANDES

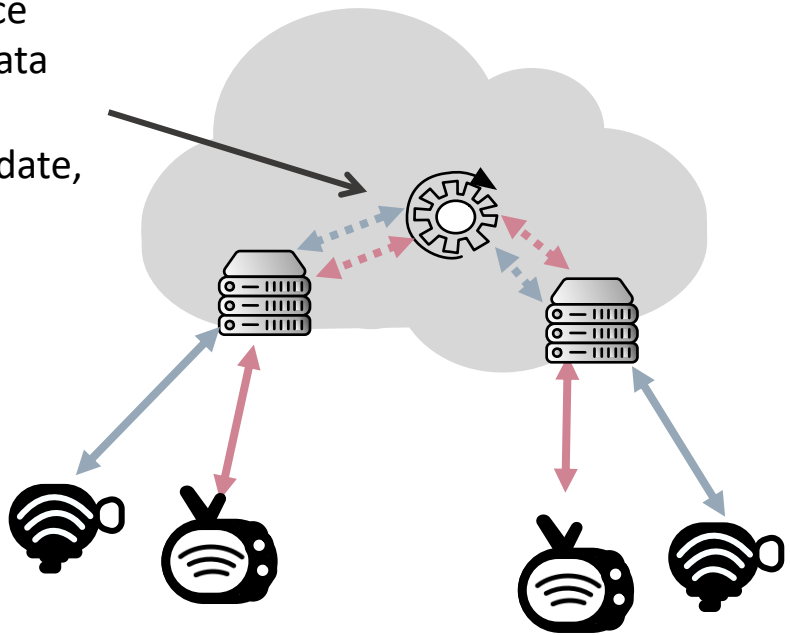
institute IMDEA software

TU Delft

IMC 2022, Nice, France

IoT Backend Server Infrastructure: support IoT devices

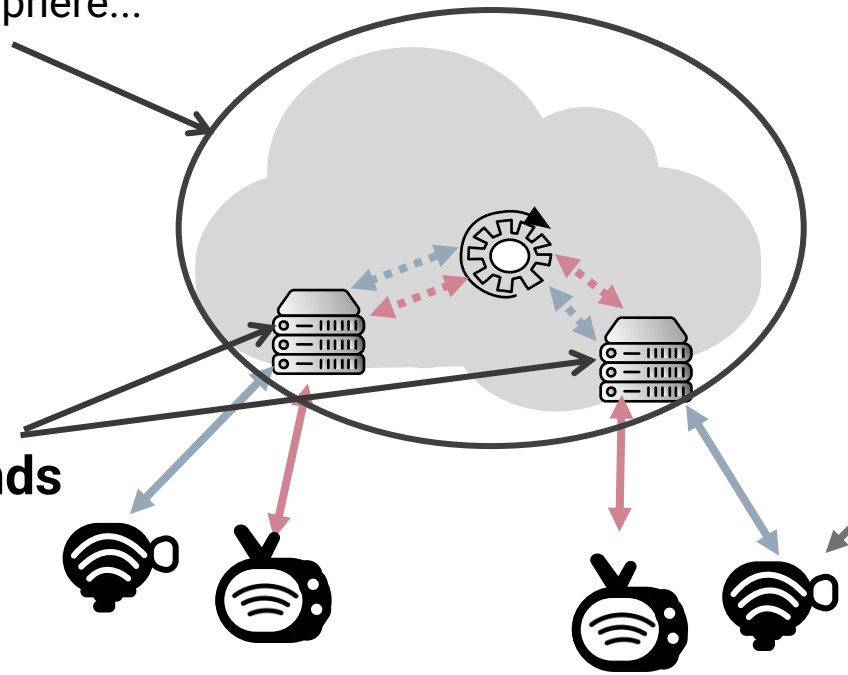
Off-loaded device functionalities, Data Ingestion, Authentication, Update, Monitoring, ...



IoT Backend Providers:

Companies providing specialized IoT backend services

IoT Backend Provider:
Amazon IoT, Google IoT Core
Siemens Mindsphere...



IoT Backends

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

Open for hire: attack trends and misconfiguration pitfalls of IoT devices

Information Exposure From Consumer IoT Devices

A Multidimensional, Network-Informed Measurement Approach

Jingjing Ren
Northeastern University

Daniel J. Dubois
Northeastern University

David Choffne
Northeastern Univer

Anna Maria Mandalari
Imperial College London

Roman Kolcun
Imperial College London

Hamed Haddad
Imperial College Lon

Research Goals:

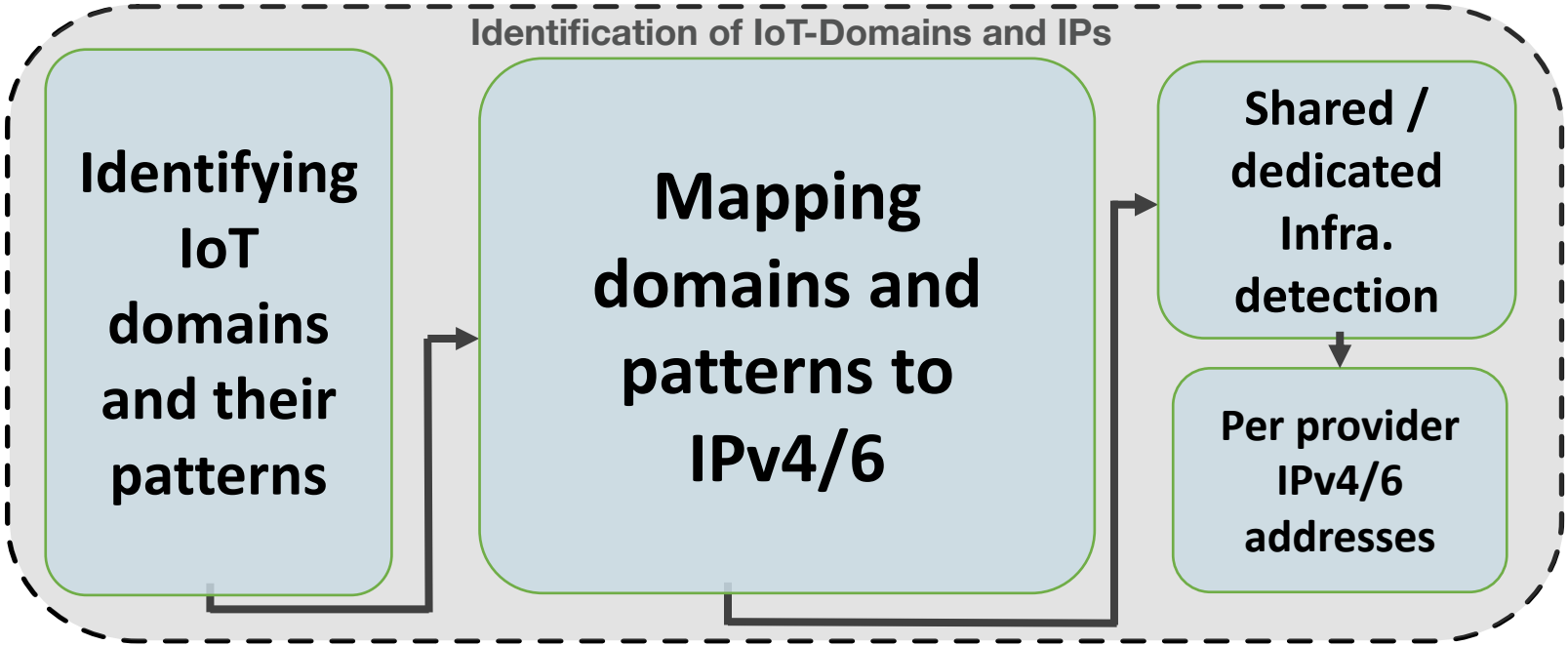
RG1: Inferring and characterizing the IoT Backends

RG2: Characterization of the IoT backend traffic (as seen in a large European ISP)

Outline

- Methodology
- RG1: Inferring and Characterizing the IoT Backends
- RG2: Characterization of the IoT Backend Traffic
- Limitations
- Summary

Methodology Overview



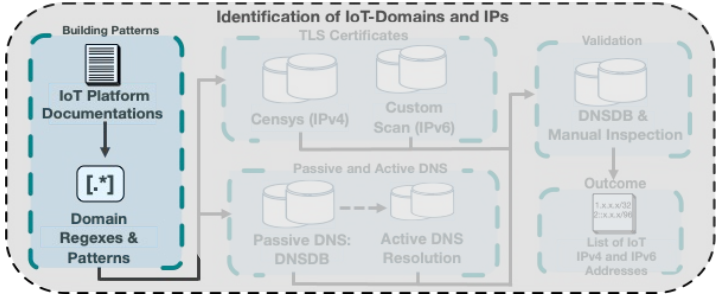
Studied Backend Providers

- Alibaba IoT
- Amazon IoT
- Baidu IoT
- Bosch IoT Hub
- Cisco Kinetic
- Fujitsu IoT
- Google IoT core
- Huawei IoT
- IBM IoT
- Microsoft Azure IoT Hub
- Oracle IoT
- PTC ThingWorx
- SAP IoT
- Siemens Mindsphere
- Sierra Wireless
- Tencent IoT

Account for 90+% of IoT backend revenue *

* IoT Analytics. 2022 List of IoT Platforms Companies. https://iot-analytics.com/product/list-of-iot-platform-companies.

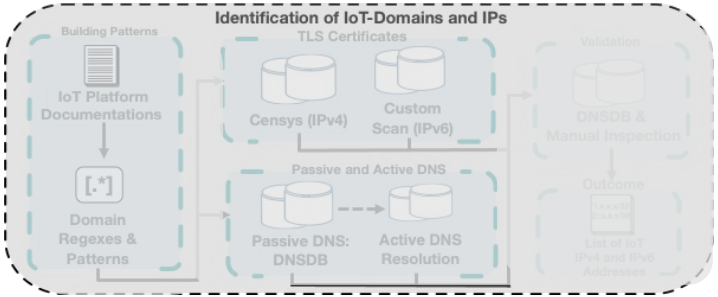
Methodology: Identifying IoT domains and their patterns



- IoT backend domains:
  - specified in the provider's documentation
  - follow well-defined form
  - <**subdomain**>.<**region**>.<**secondlevel-domain(SLD)**>
  - xyszd23d.iot.us-east2.amazonaws.com.

- Regular expression:
  - (.+)(\.iot\.)([[:alnum:]]+(-[[:alnum:]]+)+)?(\.amazonaws\.com\.$)
  - .\.(iot-(coaps|mqtts|https|amqps|api|da)\.).+\.myhuaweicloud\.com\.$

Methodology: Mapping domains and patterns to IPv4/6

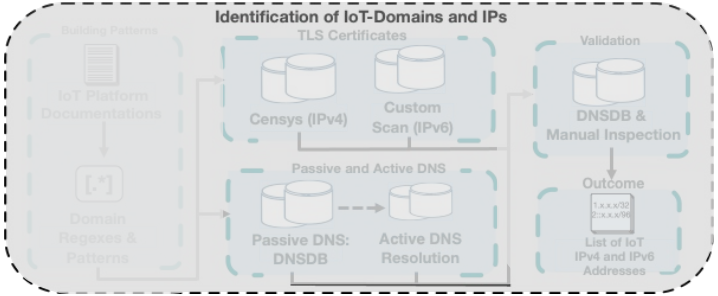


- Find hosts with matching TLS certificates:
  - **Censys**, **IPv4**
  - Custom **IPv6** scan

- DNS lookup to find other hosts(e.g., Non-TLS)
  - Matching A/AAAA/CNAME records in passive DNS datasets like DNSDB
  - Resolving the DNSDB records from multiple locations

Methodology: Shared / dedicated Infra. detection

- Check whether other (non-IoT) services are hosted on same IPs
  - Lookup IPv4/6 addresses in DNSDB to find other domains hosted on same IPs



RG1: Inferring and Characterizing the IoT Backends

- Deployment strategies
- Location
- Supported protocols

Deployment Strategies:

- Diverse deployment methodologies & # IPs
- 6 companies rely on public clouds
- One backend provider relies on another one

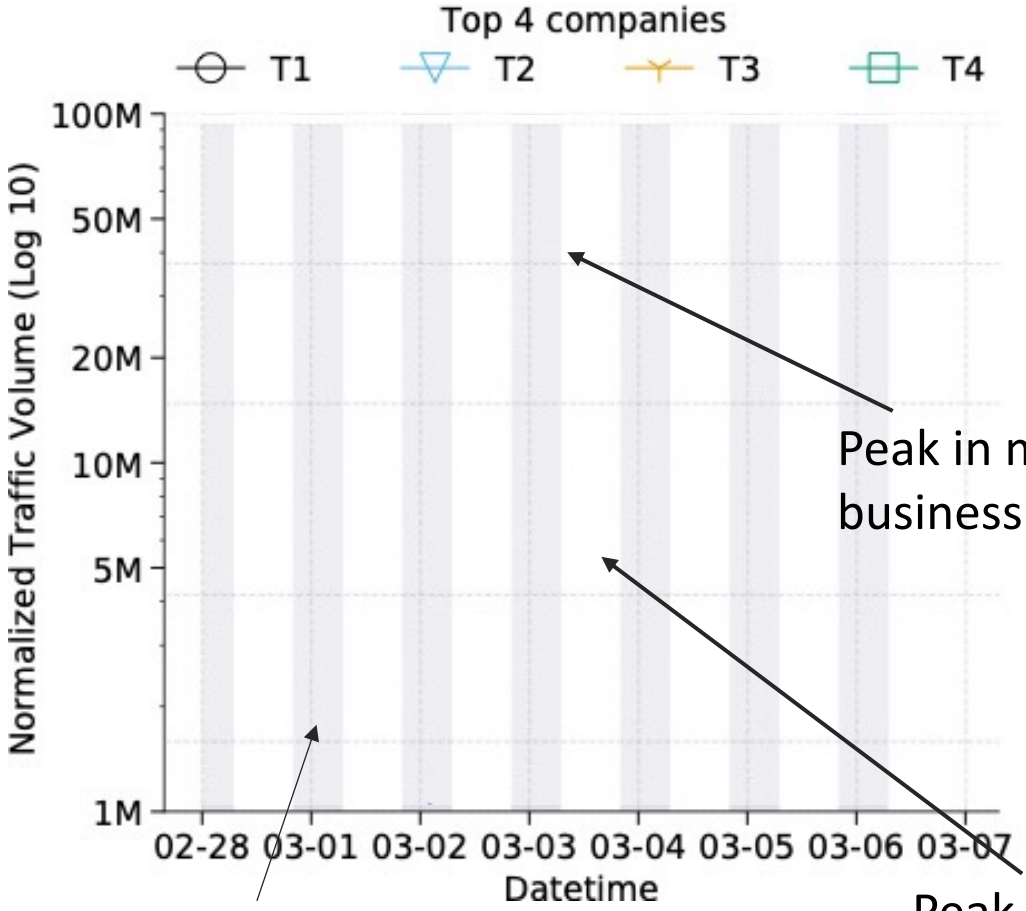
Backend Provider Name [Source]	# AS	# IPv4 /24 (IPv6 /56)	# Locations	# Countries	Protocols (Ports)	Strategy
Alibaba IoT [3–5]	2	73 (2)	27	13	MQTT(1883), HTTPS(443), CoAP(5682)	DI
Amazon IoT [8, 17, 19]	4	9,000 (20)	18	15 +Anycast	MQTT(8883, 443), HTTPS(443, 8443)	DI
Baidu IoT [21–23]	2	26 (1)	2	1	MQTT(1883, 1884, 443), HTTP(80, 443), CoAP(5682, 5683)	DI
Bosch IoT Hub [24]	1	290 (0)	1	1	MQTT(8883), HTTPS(443), AMQP(5671), CoAP(5684)	PR
Cisco Kinetic [28, 29]	2	14 (0)	4	2	MQTT(8883, 443), TCP(9123, 9124)	PR
Fujitsu IoT [45]	1	2 (0)	2	1	MQTT(8883), HTTPS(443)	DI
Google IoT core [49, 51]	1	114 (11)	77	14	MQTT(8883,443), HTTPS(443)	DI
Huawei IoT [58]	1	26 (0)	2	1	MQTT(8883, 443), HTTPS(8943), CoAP(NA)	DI
IBM IoT [59, 60]	2	116 (0)	12	8	MQTT(8883, 1883), HTTP(S)(80,443)	DI
Microsoft Azure IoT Hub [20, 67]	1	282 (0)	39	16	MQTT(8883), HTTPS(443), AQMP(5671),	DI
Oracle IoT [68, 69]	3	67 (0)	10	8	MQTT(8883), HTTPS(443)	DI+PR
PTC ThingWorx [74]	3	881 (0)	10	8	Protocol Agnostic	PR
SAP IoT [81, 82]	6	2.929 (0)	7	5	MQTT(8883), HTTPS(443)	PR
Siemens Mindsphere [84, 85]	4	126 (1)	3	3 +Anycast	MQTT(8883), HTTPS(443), OPC-UA	PR
Sierra Wireless [86–88]	4	7 (2)	4	4	MQTT(8883,1883), HTTP(S)(80,443), CoAP(5682,5686)	PR
Tencent IoT [94, 95]	5	47 (2)	5	4	MQTT(8883,1883), HTTP(S)(80,443), CoAP(5684)	DI

RG2: Characterization of the IoT Backend Traffic (as seen at a large European ISP)

- Traffic patterns
- Traffic Localization:
  How much IoT traffic leaves Europe? (privacy ....)
- Outage:
  Effect of an outage on the IoT traffic (case study)

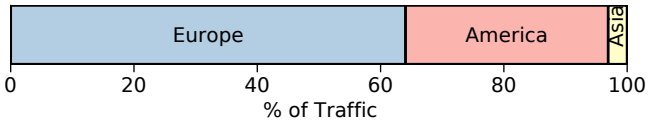
Traffic Patterns

- Diverse traffic patterns



Traffic Localization: GDPR Implications?

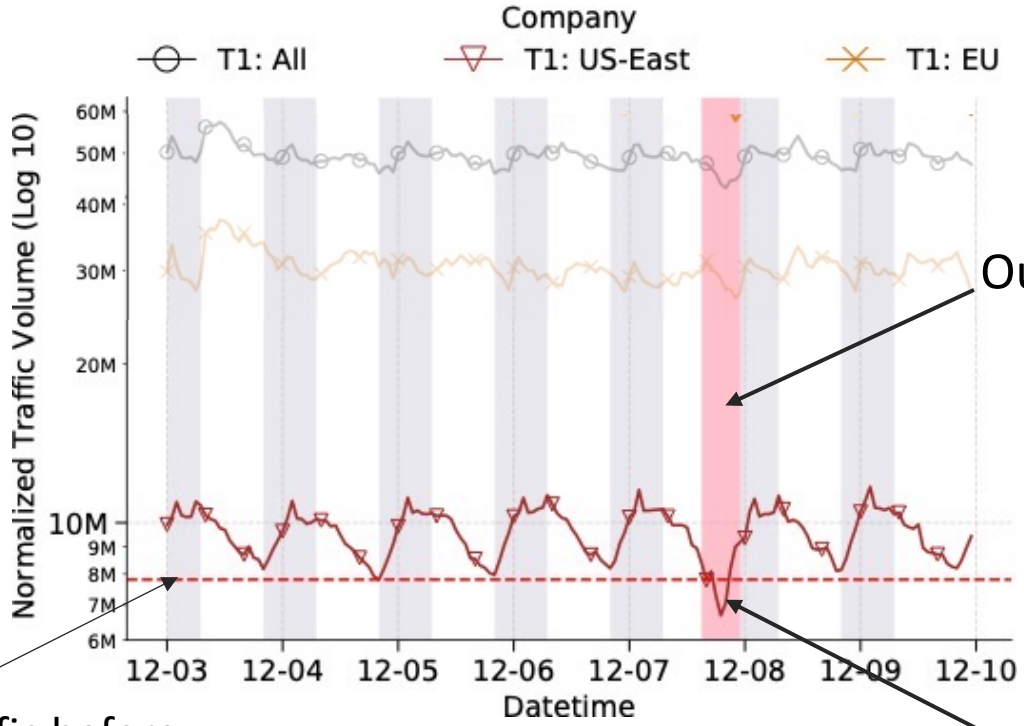
>70% of observed IoT backend server IPs are outside of EU (65% in US)



~>50% of observed IoT subscriber lines contact Non-EU servers

Majority of traffic stays within Europe

AWS Outage: December 7th, 2021



Outage period

Lowest traffic before the outage

14.5+% traffic drop

Limitations

- Reliance on public documentation
- Limited view on IPv6
- Shared infrastructure detection
- Single vantage point

Summary

For Domains
& Regexes
Scan this

- **RG1**: Inferring and characterizing IoT Backends:
  - First study on Infrastructure of IoT Backend Providers
  - Methodology to infer the IoT backends
  - Can be repurposed to identify candidate IoT traffic in passive data
- **RG2**: Characterization of the IoT backend traffic:
  - Diverse IoT traffic patterns
  - ~>50% of observed IoT subscriber lines contact Non-EU servers

Backup Slides

IoT Backend Providers: Important, yet understudied

Amazon Outage Shuts Down IoT Vacuums, Doorbells, Fridges, Even Home Locks

Talk about a single point of failure.

/ Robots & Machines / Alexa / Amazon / Internet Of Things

source: https://futurism.com/amazon-outage-iot

Sierra Wireless partially restores network following ransomware attack

Production lines are operating again at the IoT device manufacturer, but internal IT systems remain down following a cyberattack on March 20.

source: https://www.zdnet.com/article/sierra-wireless-partially-restores-network-following-ransomware-attack/

Recent research focuses on IoT devices

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

Open for hire: attack trends and misconfiguration pitfalls of IoT devices

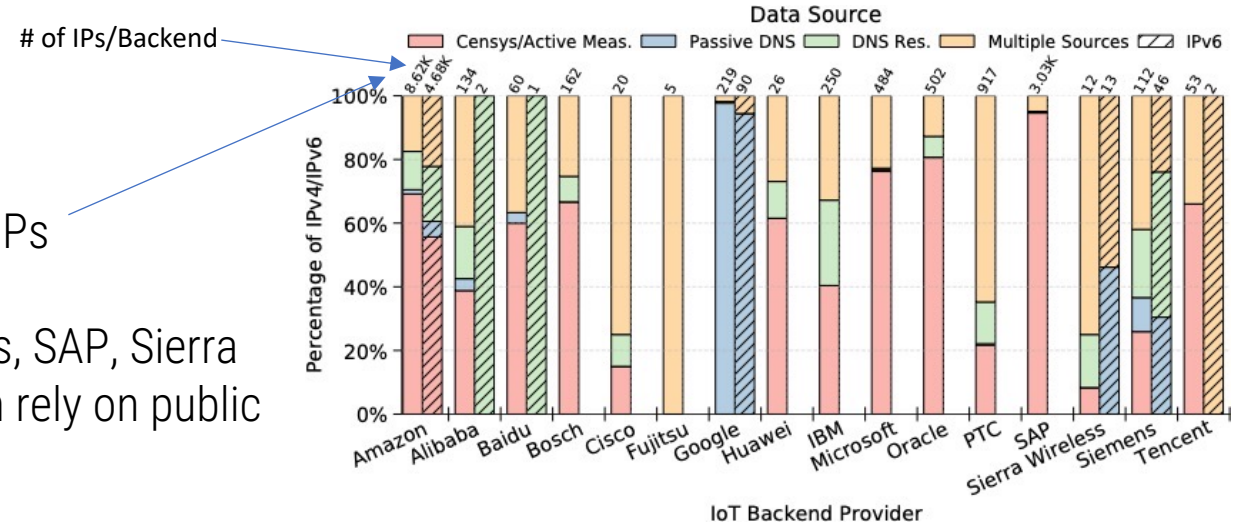
Information Exposure From Consumer IoT Devices:

A Multidimensional, Network-Informed Measurement Approach

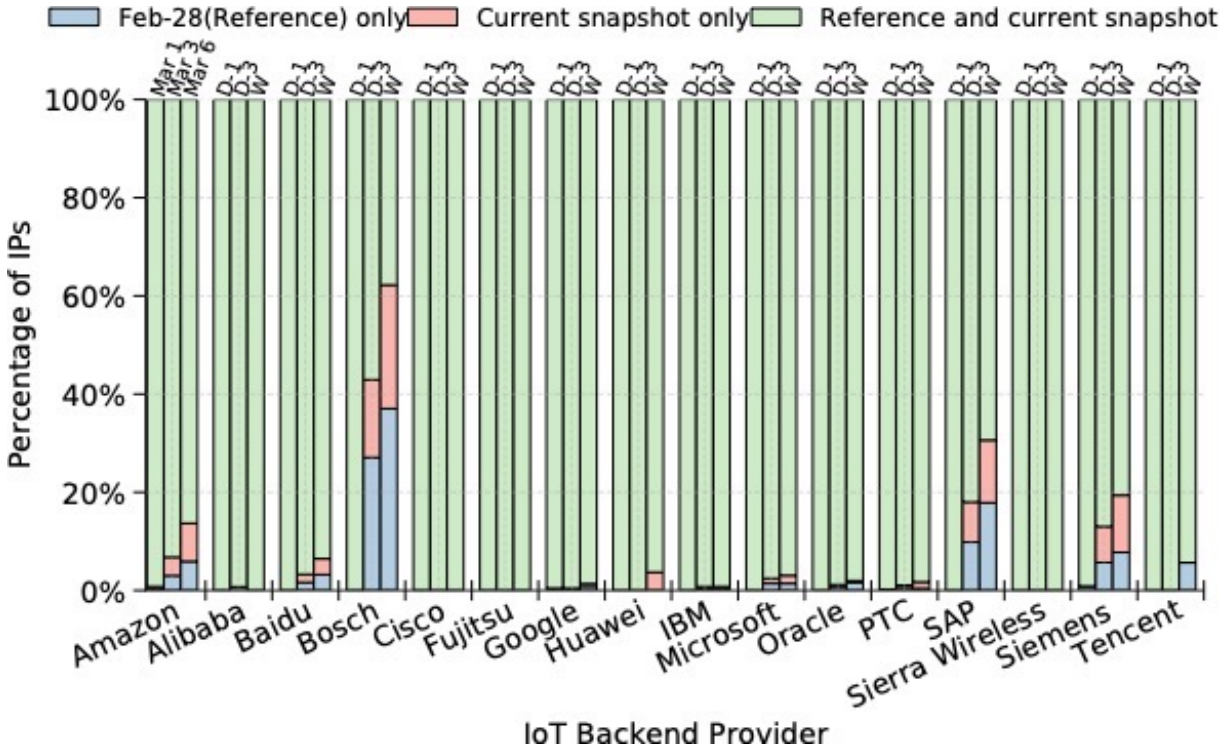
Jingjing Ren, Northeastern University
Daniel J. Dubois, Northeastern University
David Choffnes, Northeastern University
Anna Maria Mandalari, Imperial College London
Roman Kolcun, Imperial College London
Hamed Haddadi, Imperial College London

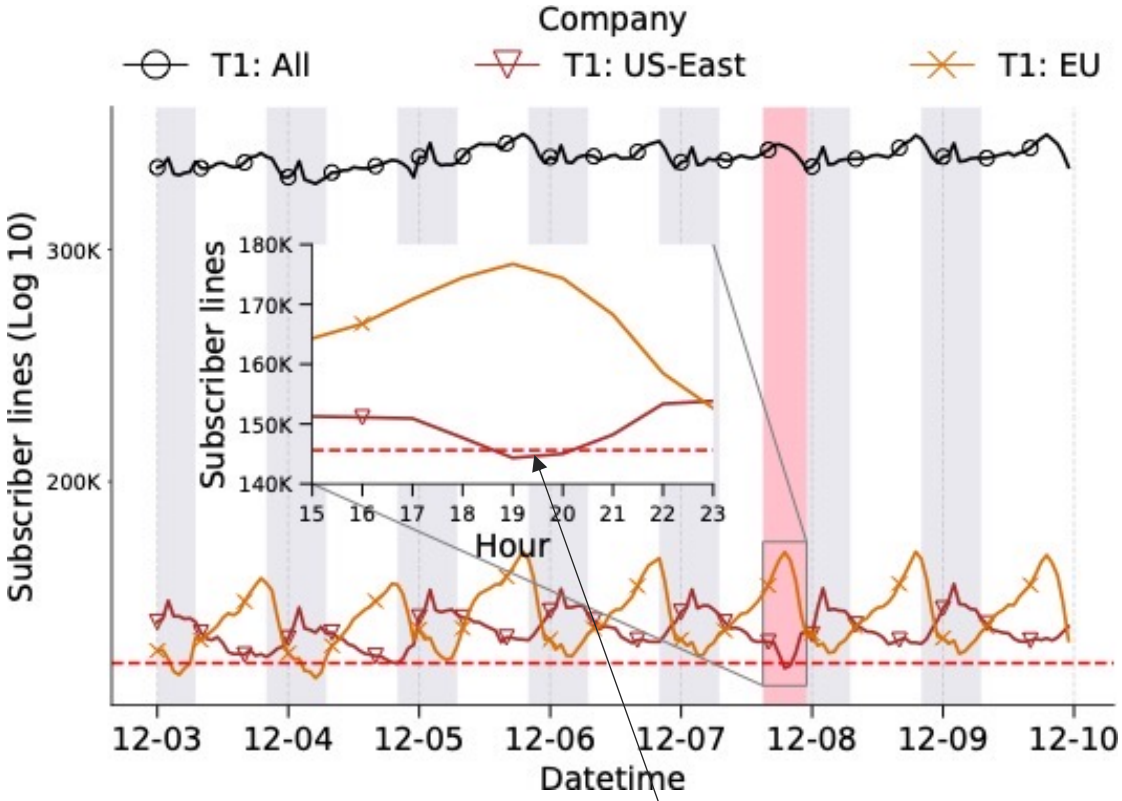
Characterization:

- Diverse deployment methodologies & # IPs
- Cisco, PTC, Siemens, SAP, Sierra Wireless, and Bosch rely on public clouds
- Siemens partially relies on Amazon IoT



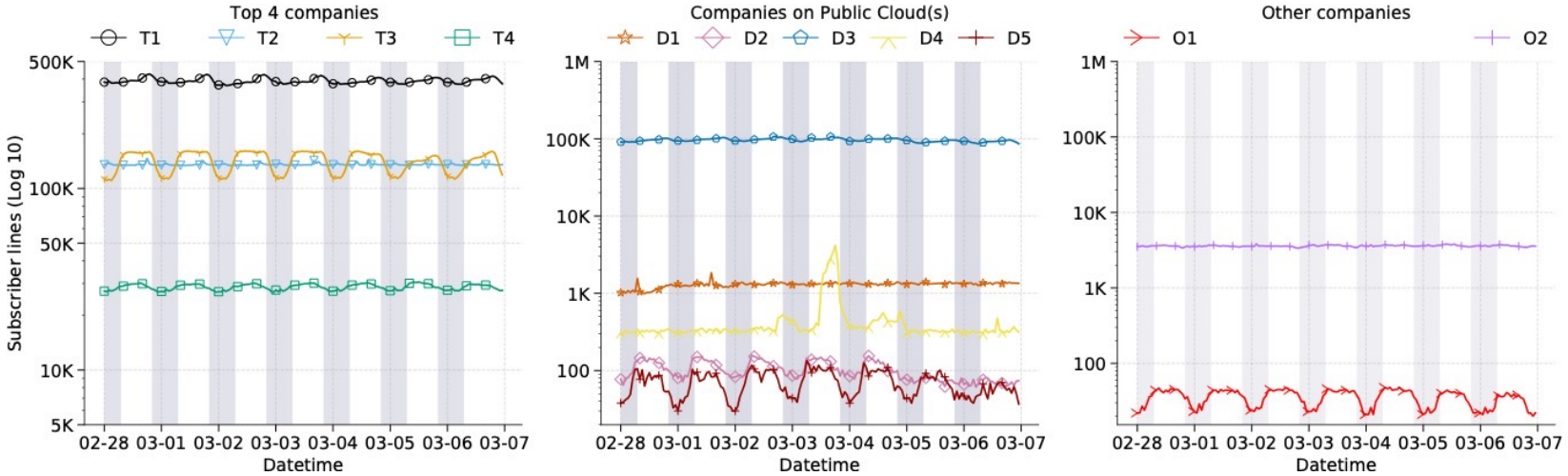
IP Stability



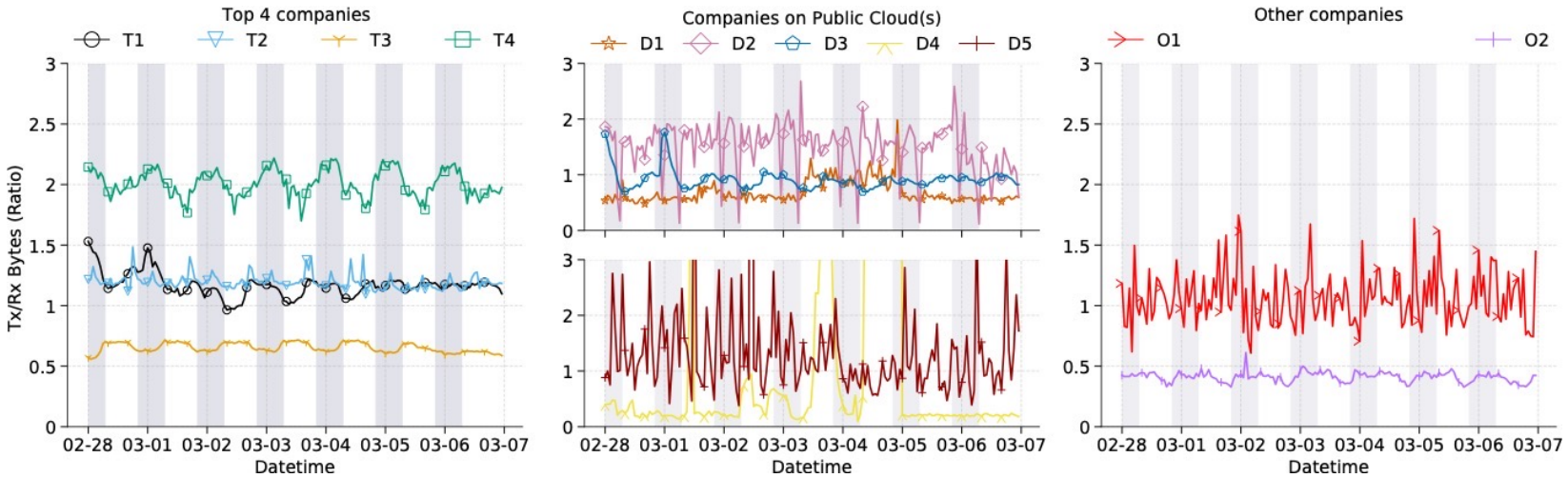


Slight #of European customers were affected

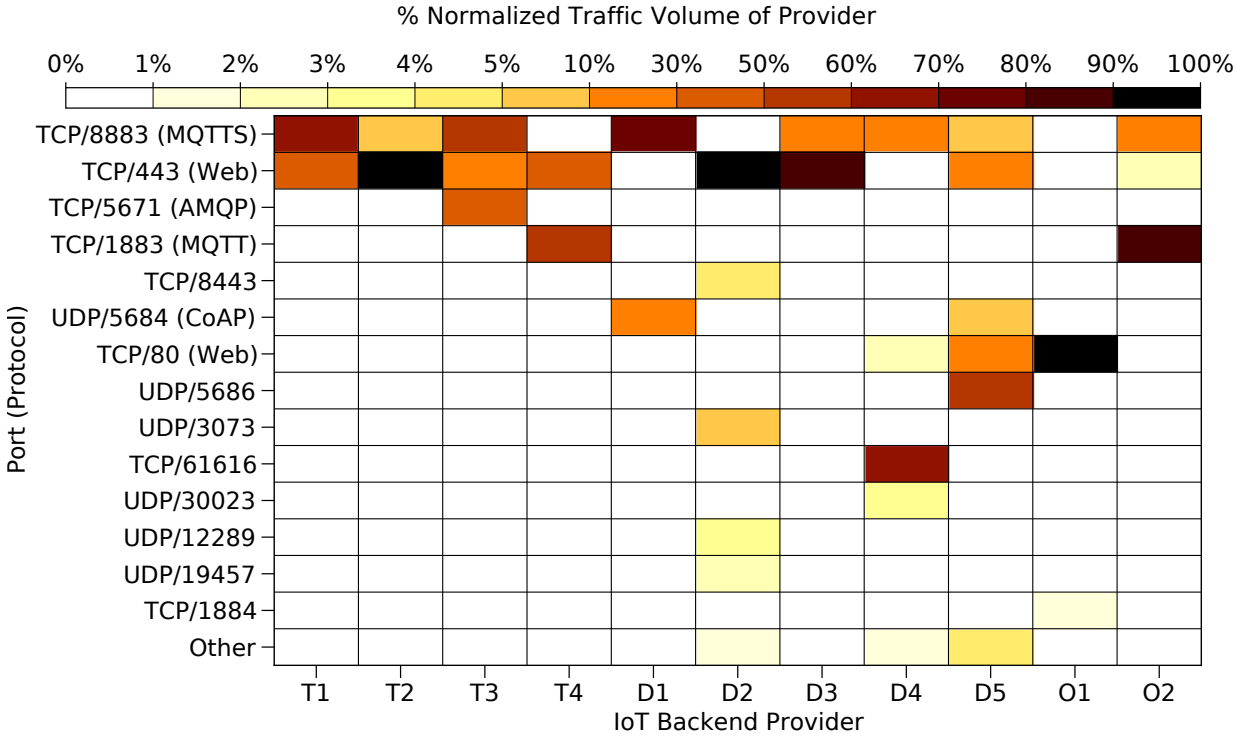
of Subscriber lines per provider



of Downstream vs Upstream Traffic Ratio



Application Mix



Application Mix