A Principled Approach to Measuring the IoT Ecosystem

Deepak Kumar

Advised By: Michael Bailey

Committee: Michael Bailey, Nikita Borisov, Adam Bates, Gang Wang, Zakir Durumeric

Thesis Statement: Network measurements drawn from a single vantage point or technique can lead to a biased view of the IoT ecosystem

Thesis Outline

	Active Measurement	**Passive Measurement**
Internal Vantage Point	Avast, IoT Inspector	IoT Inspector
External Vantage Point	Avast, Mirai	Avast, IoT Inspector

Comparing Active Internal and Active External

- Collected IoT devices from Avast and Censys active scans on a single day, May 2, 2020
- Performed external device type identification using five protocols: HTTPS, CWMP, Telnet, SSH, FTP
- CWMP, Telnet most descriptive (by label), while HTTPS, CWMP offered on most devices
- Labeled 6.1M external IoT devices, 1.8M internal IoT devices

Protocol	% IPv4 Responded	% IPs Labeled
443/HTTPS	38.3%	3.7%
7547/CWMP	19.3%	16.3%
22/SSH	16%	2.4%
21/FTP	9.1%	6.8%
23/Telnet	2.4%	20.6%

Comparing Active Internal and Active External

Service	Protocol	Avast Rank	Censys Rank
80	**HTTP**	1 (72%)	3 (15.9%)
53	**DNS**	2 (37%)	8 (3%)
443	**HTTPS**	3 (30%)	2 (27%)
8080	**HTTP**	4 (14.5%)	7 (5.8%)
445	**SMB**	5 (13%)	11 (0.4%)
22	**SSH**	6 (12%)	6 (7%)
631	**IPP**	7 (10%)	12 (0.35%)

Service	Protocol	Avast Rank	Censys Rank
23	**Telnet**	8 (8.3%)	5 (9.2%)
21	**FTP**	9 (7.4%)	4 (11.6%)
7547	**CWMP**	10 (3.9%)	1 (59%)
8888	**HTTP**	11 (2.4%)	14 (0.3%)
8883	**MQTT**	12 (1.6%)	37 (0%)

Comparing Active Internal and Active External

Service	Protocol	Avast Rank	Censys Rank
80	HTTP	1 (72%)	3 (15.9%)
[illegible]	DNS	[illegible]	[illegible]
443	[illegible]	[illegible]	[illegible]
8080	HTTP	4 (14.5%)	7 (5.8%)
445	SMB	5 (13%)	11 (0.4%)
22	SSH	6 (12%)	6 (7%)
631	IPP	7 (10%)	12 (0.35%)

Service	Protocol	Avast Rank	Censys Rank
23	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	4 (11.6%)
7547	CWMP	10 (3.9%)	1 (59%)
8888	HTTP	11 (2.4%)	14 (0.3%)
8883	MQTT	12 (1.6%)	37 (0%)

External scanning underreports IoT specific protocols

Comparing Active Internal and Active External

Device Type	% Internal Devices	% External Devices
Router	61.9%	92.3%
Media	20.7%	1.5%
Work Appliances	6.7%	0.7%
Camera	3.4%	0.6%
Generic IoT	1.1%	0.4%
Storage	0.9%	4.3%

Comparing Active Internal and Active External

Device Type	% Internal Devices	% External Devices
Router	**61.9%**	**92.3%**
Media	**20.7%**	**1.5%**
Work Appliances	**6.7%**	0.7%
Camera	**3.4%**	0.6%
Generic IoT	1.1%	0.4%
Storage	0.9%	**4.3%**

Comparing Active Internal and Active External

Devices that support SSH

SSH	
Internal	**External**
Router (88.3%)	Router (97.9%)
Storage (4.2%)	Generic IoT (0.2%)
Work (3.2%)	
Camera (2%)	
Media (1.4%)	

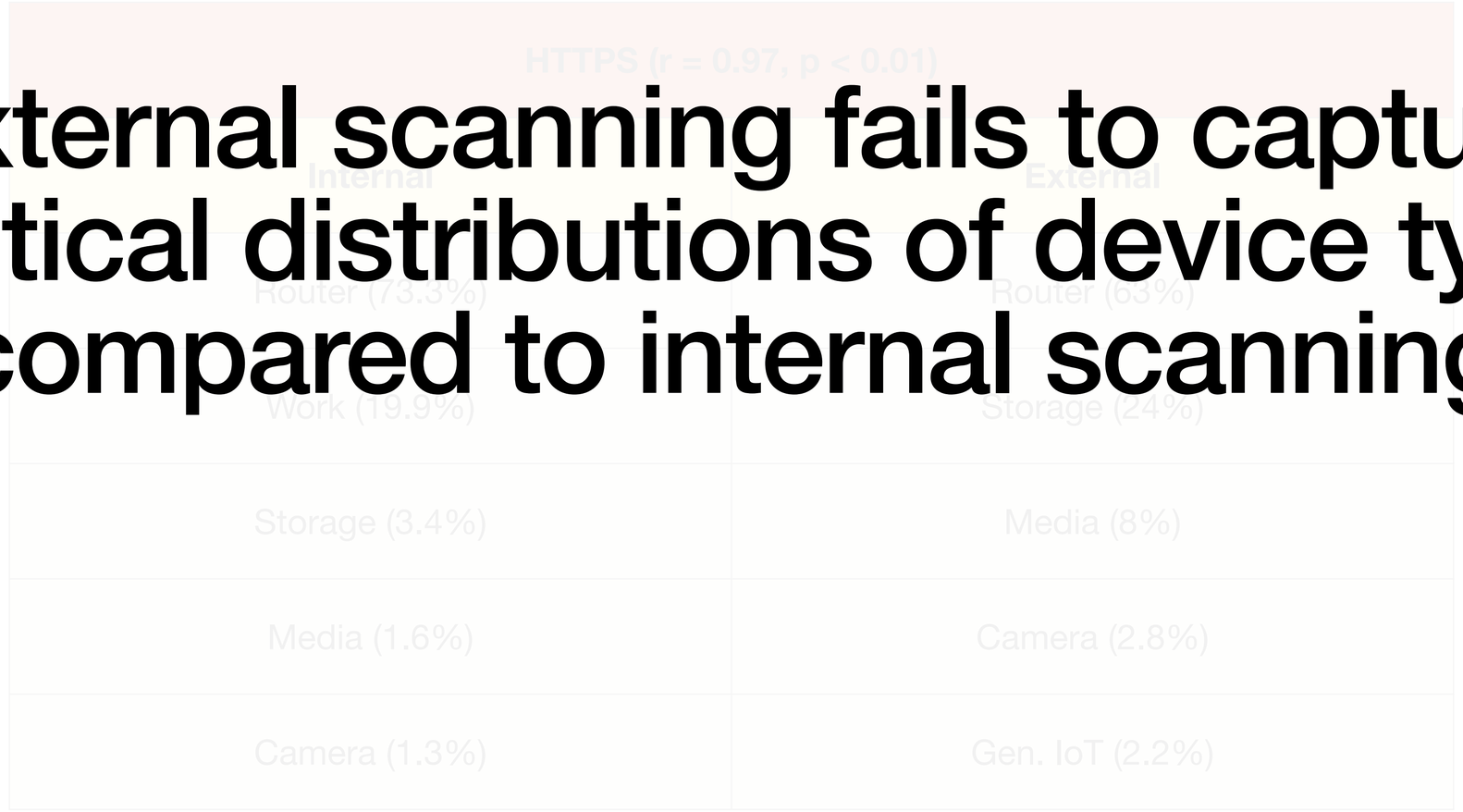
Comparing Active Internal and Active External

Devices that support HTTPS

HTTPS ($r = 0.97$, $p < 0.01$)	
Internal	**External**
Router (73.3%)	Router (63%)
Work (19.9%)	Storage (24%)
Storage (3.4%)	Media (8%)
Media (1.6%)	Camera (2.8%)
Camera (1.3%)	Gen. IoT (2.2%)

Comparing Active Internal and Active External

Devices that support HTTPS



External scanning fails to capture identical distributions of device types compared to internal scanning

Comparing Passive Internal to Passive External

- Devices only use a small fraction of ports when communicating externally
- Primarily features communication over ports 53, 443, 80, all standard protocols to support DNS, HTTPS, and HTTP

Protocol	% Devices Communicated
53/UDP	69%
443/TCP	67%
80/TCP	43%
123/UDP	31%
443/UDP	12%
8883/TCP	6.4%

Comparing Passive Internal to Passive External

Device Type	Fraction Observed Externally
Smart Home IoT	35%
Media	6%
Voice Assistant	27%
Camera	75%
Television	22%
Work Appliance	12.5%
Storage	42%

Comparing Passive Internal to Passive External

Device Type	Fraction Observed Externally
Smart Home IoT	35%
Voice Assistant	27%
Camera	75%
Television	22%
Work Appliance	12.5%
Storage	42%

External device behavior does not match internal device behavior

Comparing Active Internal and Passive Internal

- Compared the protocols that were offered (through SYN scan) to the protocols we observed used by devices
- Devices use only a median 50% of services offered during our measurement period
- Many of the unused services are security critical (e.g., 23/Telnet, 111/rpcbind)

Comparing Active Internal and Passive Internal

- Compared the protocols that were offered (through SYN scan) to the protocols we observed used by devices
- Devices use only a median 50% of services offered during our measurement period
- Many of the unused services are security critical (e.g., 23/Telnet, 111/rpcbind)

Protocol	% Devices Unused	Can Explain?
22/SSH	100%	✘
9100/CUPS	100%	✔
8081/HTTP	100%	✘
111/rpcbind	100%	✘
8443/MQTT	97%	✔
23/Telnet	96%	✘

Comparing Active Internal and Passive Internal

- Compared the protocols that were offered (through SYN scan) to the protocols we observed used by devices
- Devices use only a median 50% of services offered during the measurement period
- Many of these services are security critical (e.g., 23/Telnet, 111/rpcbind)

Protocol	Devices Used	Can Explain?
22/SSH	100%	✗
80/HTTP	100%	✓
111/rpcbind	100%	✗
8443/MQTT	97%	✓
23/Telnet	96%	✗

Device capabilities through active probing are often much larger than device behavior through passive observation

Comparing Active Internal to Passive External

- Devices inside home networks support a host of active services
- We *expect* the behavior of devices to the outside world to be different from the services offered inside networks
- We find no correlation between the services offered from active internal and passive external

Protocol	% Devices Support
8008/HTTP	36%
8443/HTTPS	36%
80/HTTP	31%
443/HTTPS	17%
8080/HTTP	12%
1843/—	11%
1443/—	11%
22/SSH	8%
8060/—	0%

Active services on IoT devices

Comparing Active External to Passive Behavior

- Internal and external device behavior cannot be observed from active probing on the outside
  - At best, can attribute behavior to an externally facing router, but unlikely that the router is producing the traffic

Principles for Measuring IoT

- An internal vantage point provides access to finer-grained data about device types, vendors, and behaviors compared to external vantage points
- Active, external probing can provide sets of IoT devices, but lacks the diversity of IoT devices inside networks and underreports import IoT protocols
- IoT Devices often exhibit different behaviors and capabilities, requiring both passive and active measurement to properly quantify