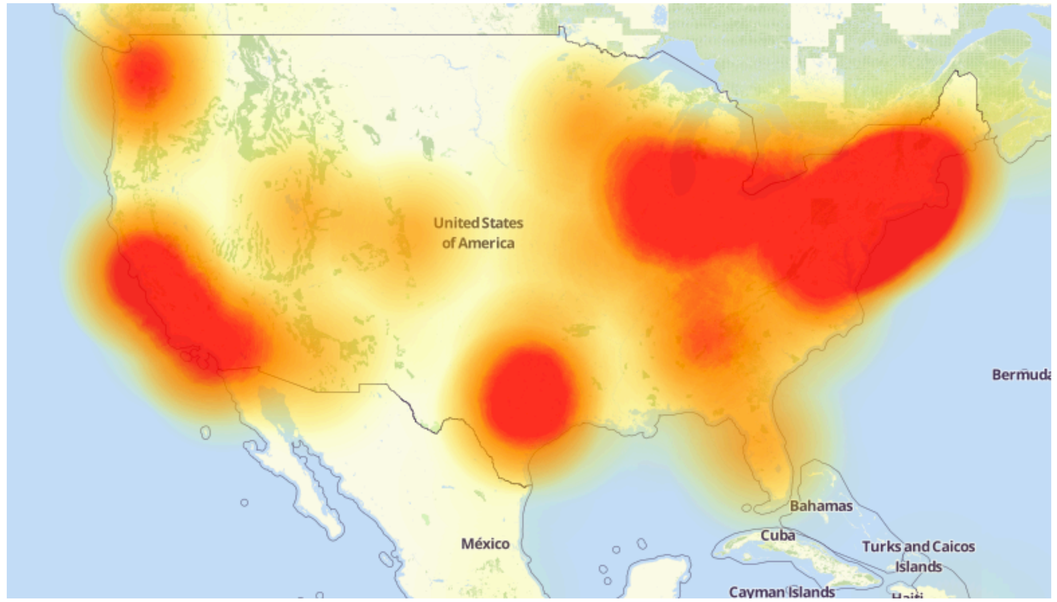
A Principled Approach to Measuring the IoT Ecosystem

Deepak Kumar
University of Illinois

Cyberattack Knocks Out Access to Websites

Popular sites such as Twitter, Netflix and PayPal were unreachable for part of the day



Measuring the Mirai Botnet

Data Source	Size
Network Telescope	4.7M unused IPs
Active Scanning	136 IPv4 scans
Telnet Honeypots	434 binaries
Malware Repository	594 binaries
Active/Passive DNS	499M Daily RRs
C2 Milkers	64K issued attacks
Krebs DDoS Attack	170K attacker IPs
Dyn DDoS Attack	108K attacker IPs

Understanding the Mirai Botnet – USENIX 2017

What can we learn about the IoT ecosystem by using varied measurement perspectives and techniques?

Outline

- Relevant background/motivation
- Brief discussion of completed work
- Proposed future projects
- Discussion and future directions

Measurement Perspectives

Measurement Perspectives

- Internal
  - Many IoT devices are behind NATs, requiring a local network perspective to study devices

Measurement Perspectives

- Internal
  - Many IoT devices are behind NATs, requiring a local network perspective to study devices
- External
  - Public fingerprint of a device is often the only perspective researchers have for security analysis

Measurement Perspectives – Limitations

- Internal scanning is an effective method to learning what IoT devices inside homes really look like, but threat model is stricter
- External scanning can give us a sense of devices that are *immediately* vulnerable

Measurement Techniques

Measurement Techniques

- Active
  - Probe devices (e.g., send TCP SYN) to learn of their *server* capabilities

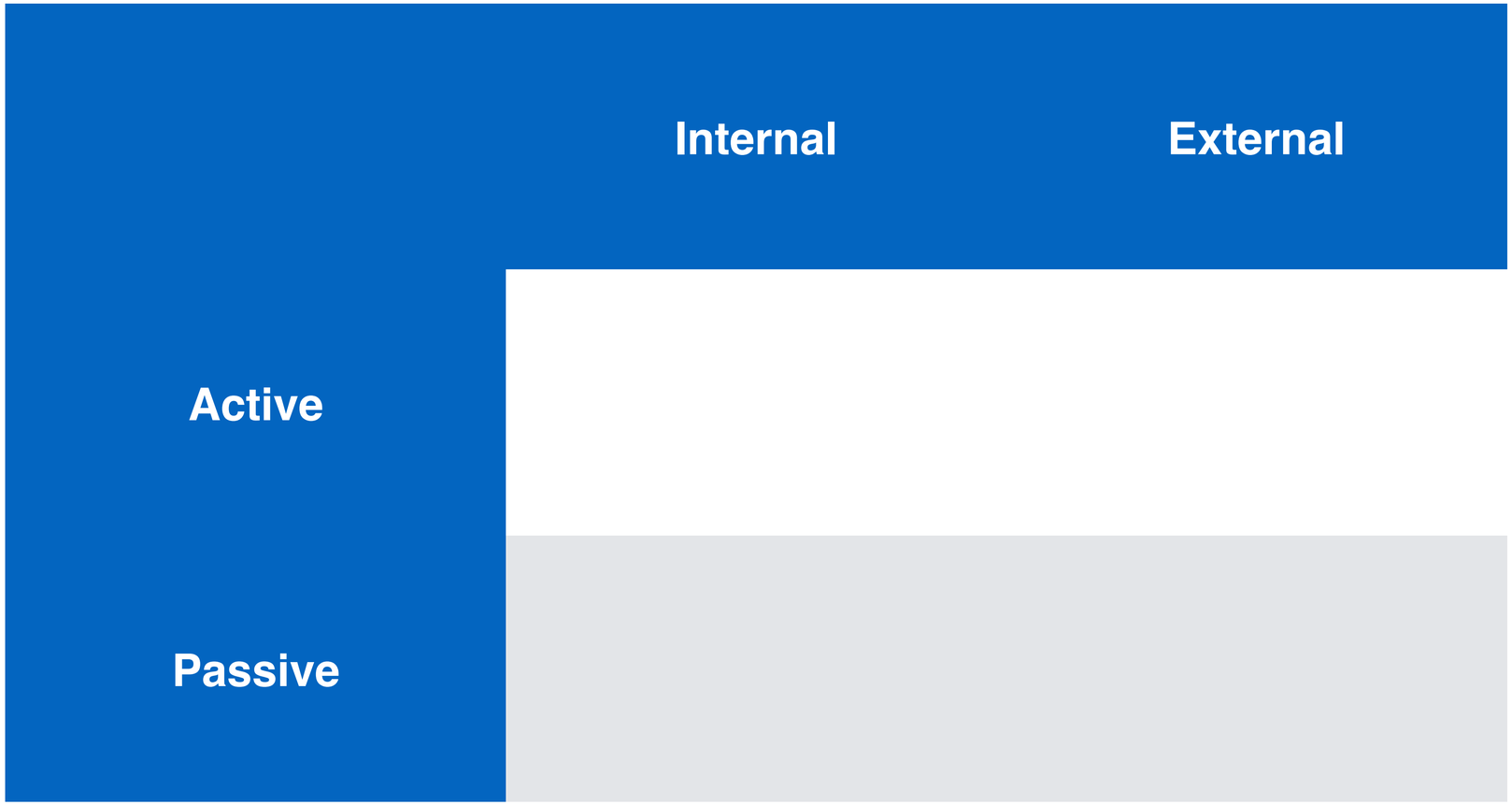
Measurement Techniques

- Active
  - Probe devices (e.g., send TCP SYN) to learn of their *server* capabilities
- Passive
  - Observe devices (e.g., network tap) to learn of their *client* behavior

Measurement Techniques – Limitations

- Active probing enumerates *all* the server capabilities of a device, but can't tell you how the device is used
- Passive observation tells you the network behavior of devices, but doesn't enumerate capabilities

Thesis Plan



Thesis Plan

	Internal	External
Active	“All Things Considered: An Analysis of IoT Devices on Home Networks”	
Passive		

All Things Considered: An Analysis of IoT Devices on Home Networks

Deepak Kumar
University of Illinois

Kelly Shen
Stanford University

Benton Case
Stanford University

Deepali Garg
Avast Software

Galina Alperovich
Avast Software

Dmitry Kuznetsov
Avast Software

Rajarshi Gupta
Avast Software

Zakir Durumeric
Stanford University

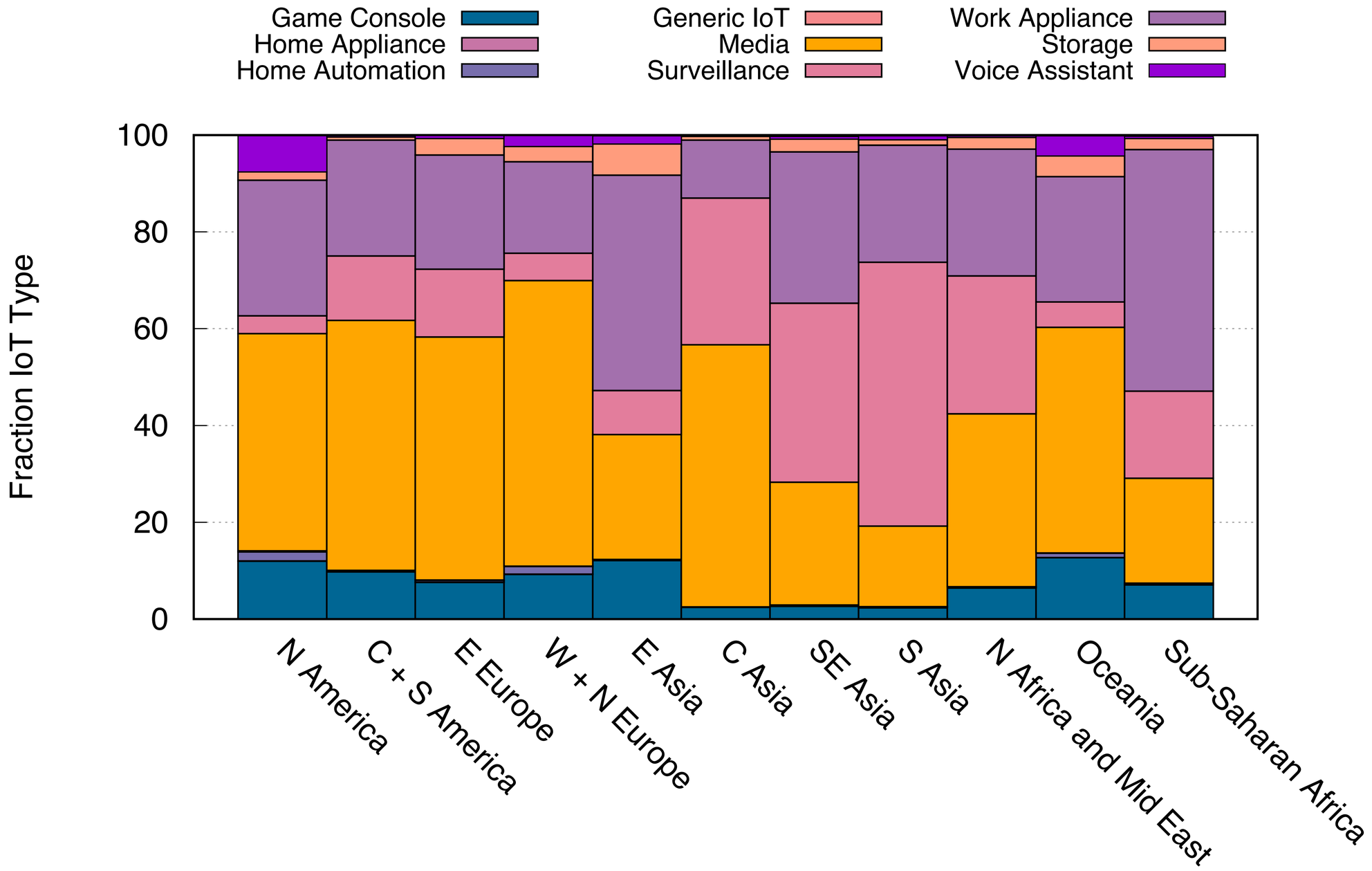
USENIX Security 2019

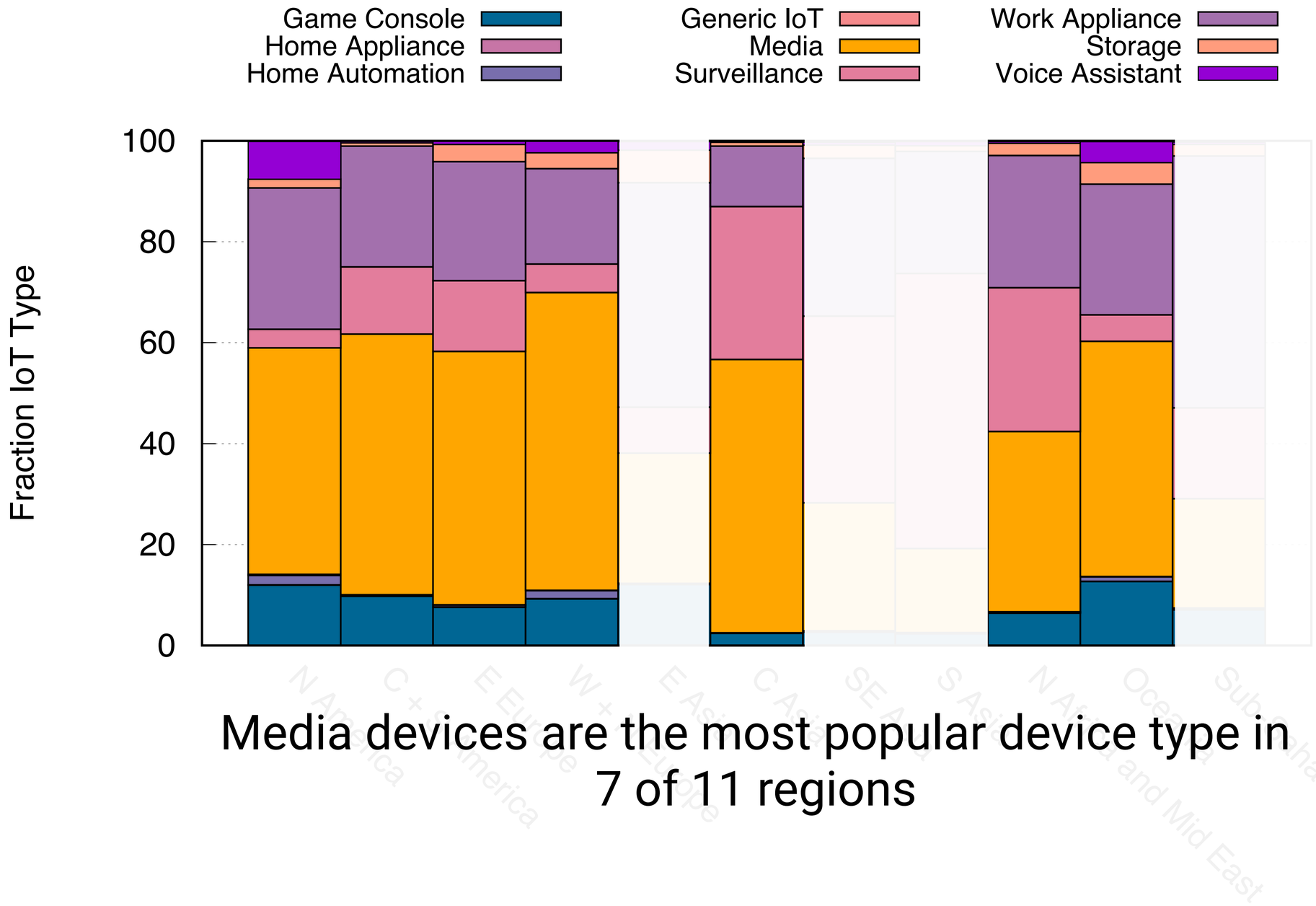
Avast Wi-Fi Inspector

- Performs active internal network scans and checks devices for weak security
  - Device identification
  - Weak default credentials
  - Vulnerability to known recent CVEs

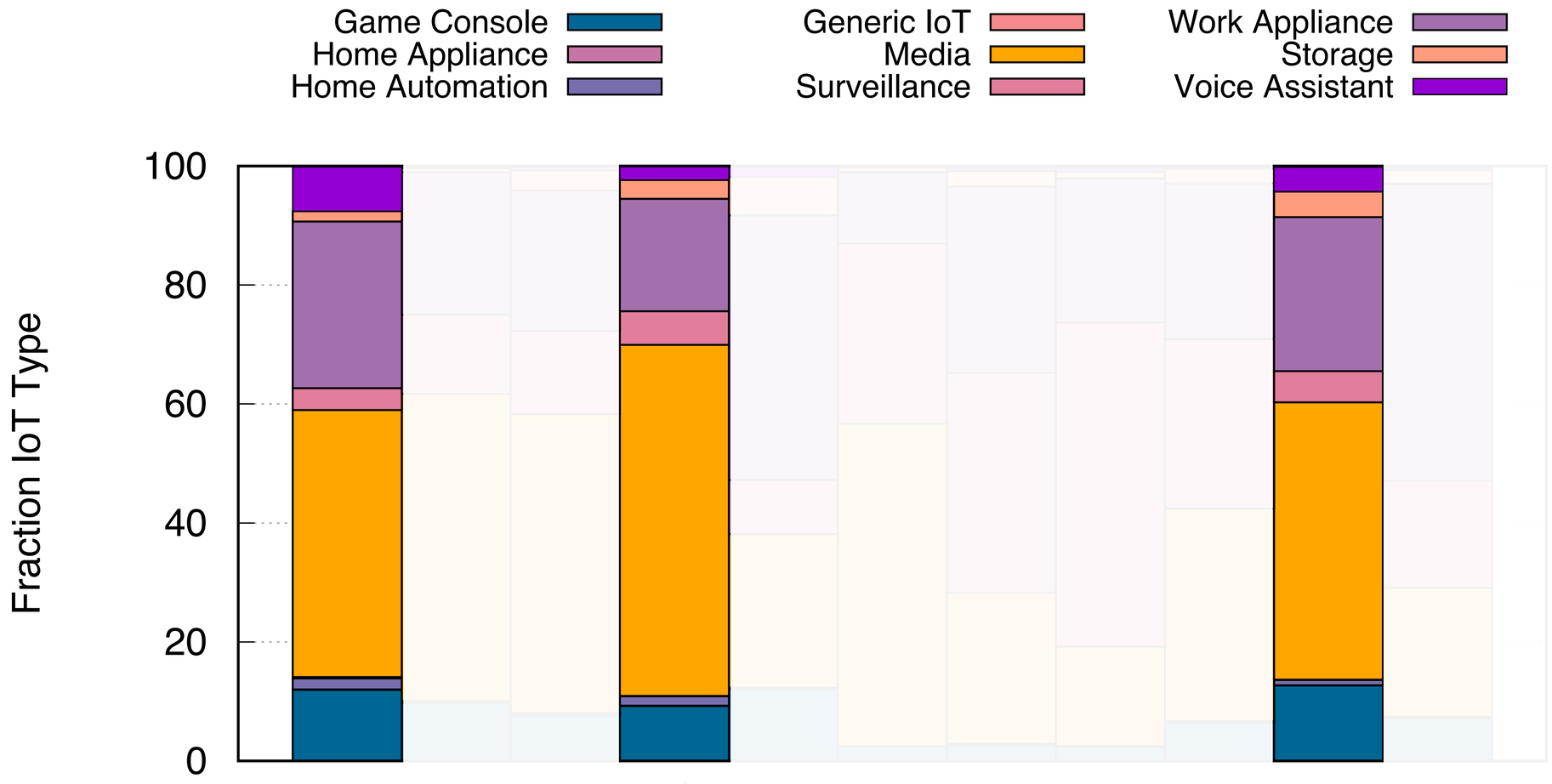
Dataset

Network scans collected from
15.5 million homes, spanning
83 million devices across
11 geographic regions

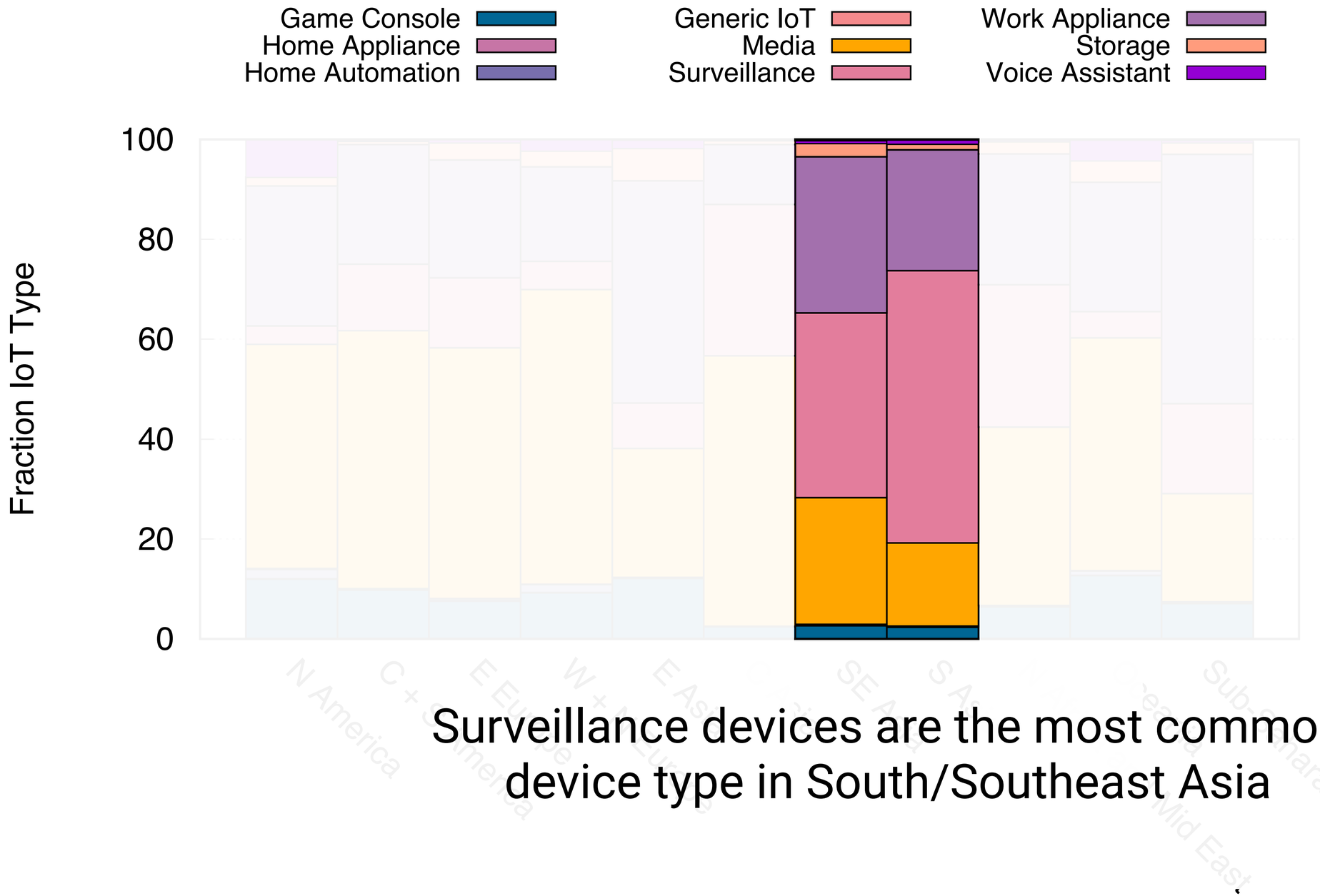

Game Console
Home Appliance
Home Automation
Generic IoT
Media
Surveillance
Work Appliance
Storage
Voice Assistant
Fraction IoT Type
100
80
60
40
20
0
N America
C + S America
E Europe
W + N Europe
E Asia
C Asia
SE Asia
S Asia
N Africa and Mid East
Oceania
Sub-Saharan Africa



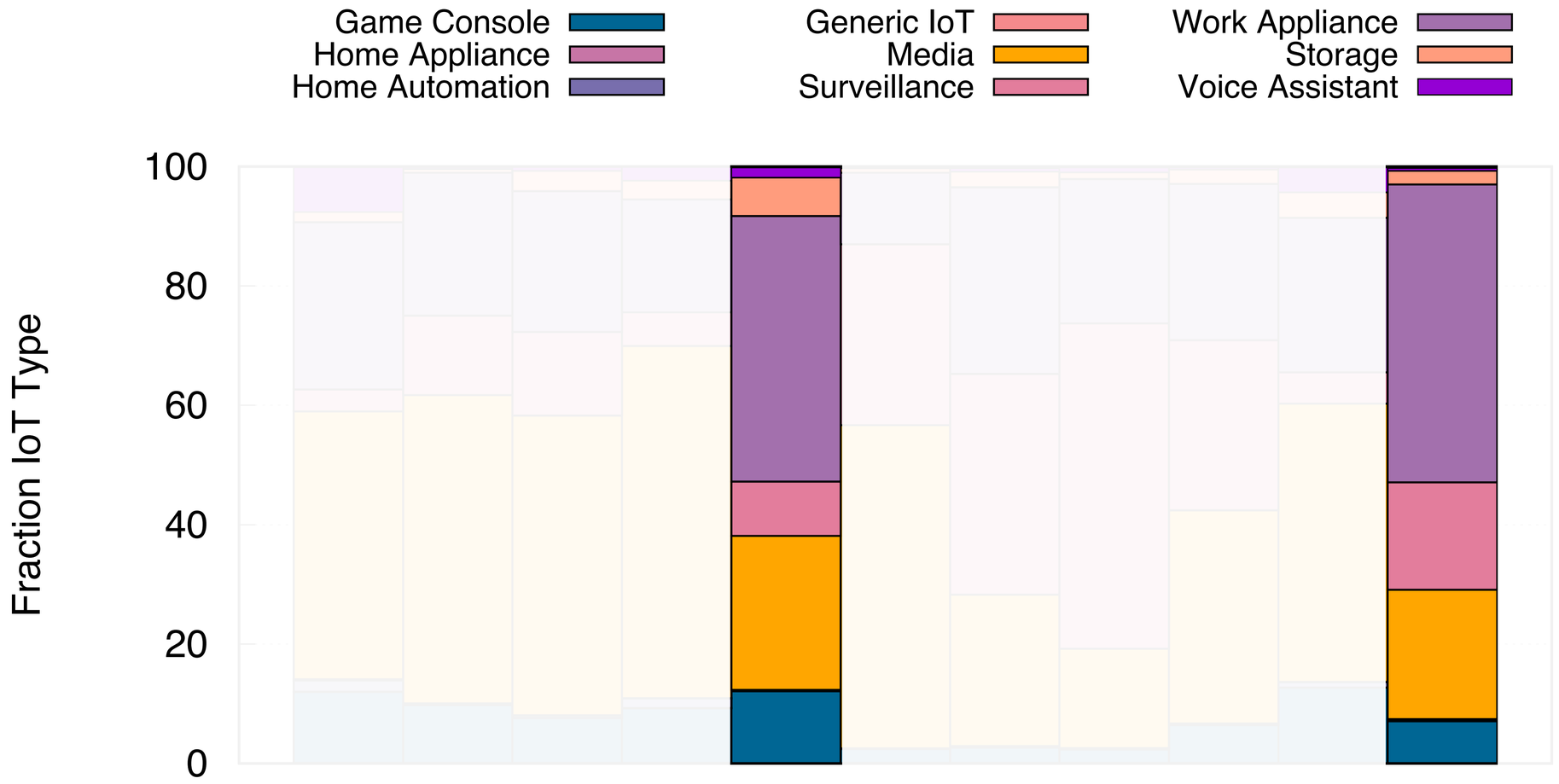
Media devices are the most popular device type in 7 of 11 regions



Home automation and voice assistants are only prevalent (>1% of homes) in North America, Western Europe, Oceania



Surveillance devices are the most common device type in South/Southeast Asia



Work appliances are the most common device type in East Asia/Sub-Saharan Africa

Each region has unique device type "preferences"

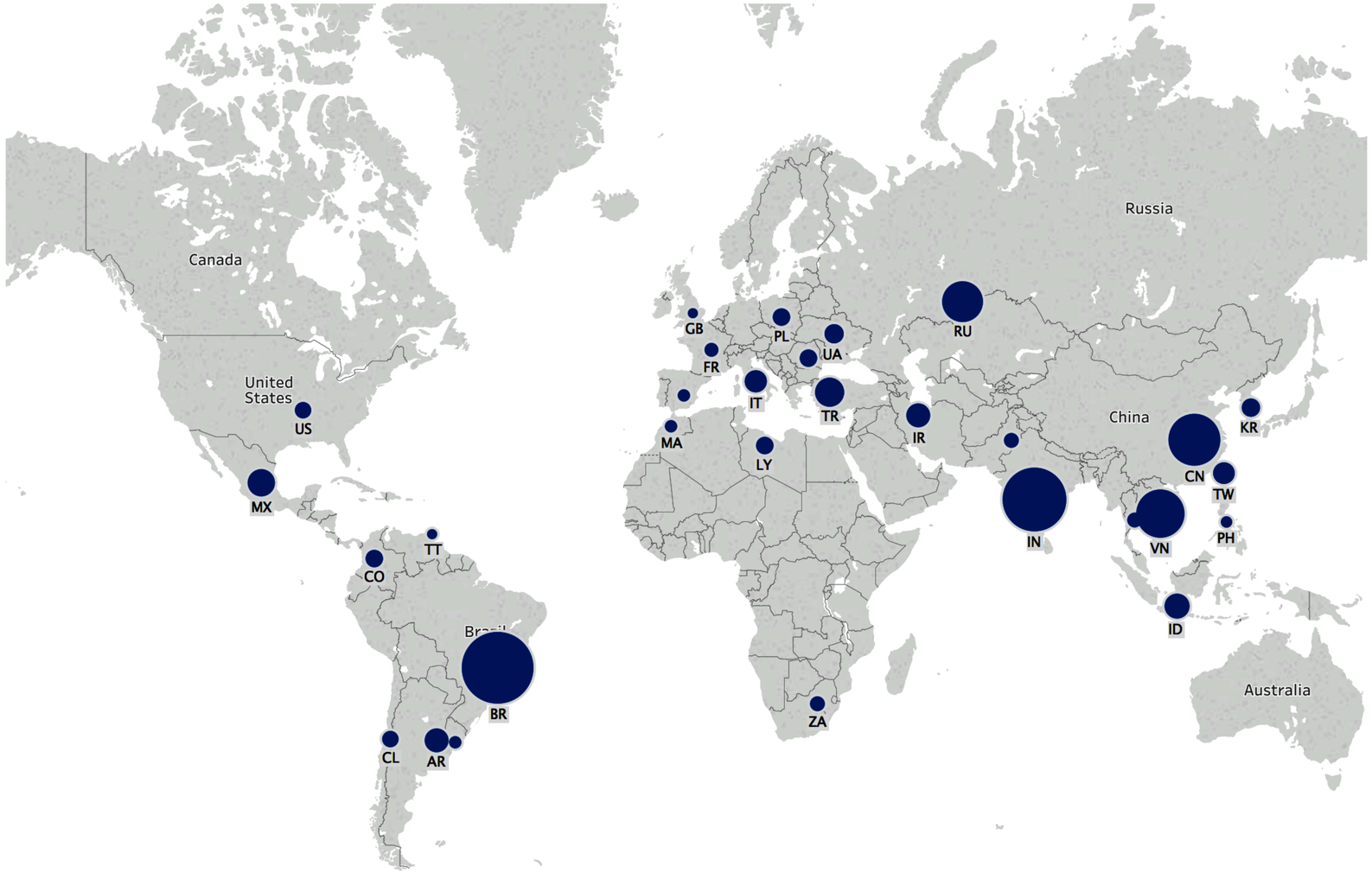
Case Study: Weak Telnet Credentials

Device Type	% Support Telnet	% Weak Telnet
Surveillance	**14.6%**	**10.7%**
Router	**14.6%**	**1.9%**
Home Appliance	3.2%	1.6%
Media	1.4%	0.9%

Case Study: Weak Telnet Credentials

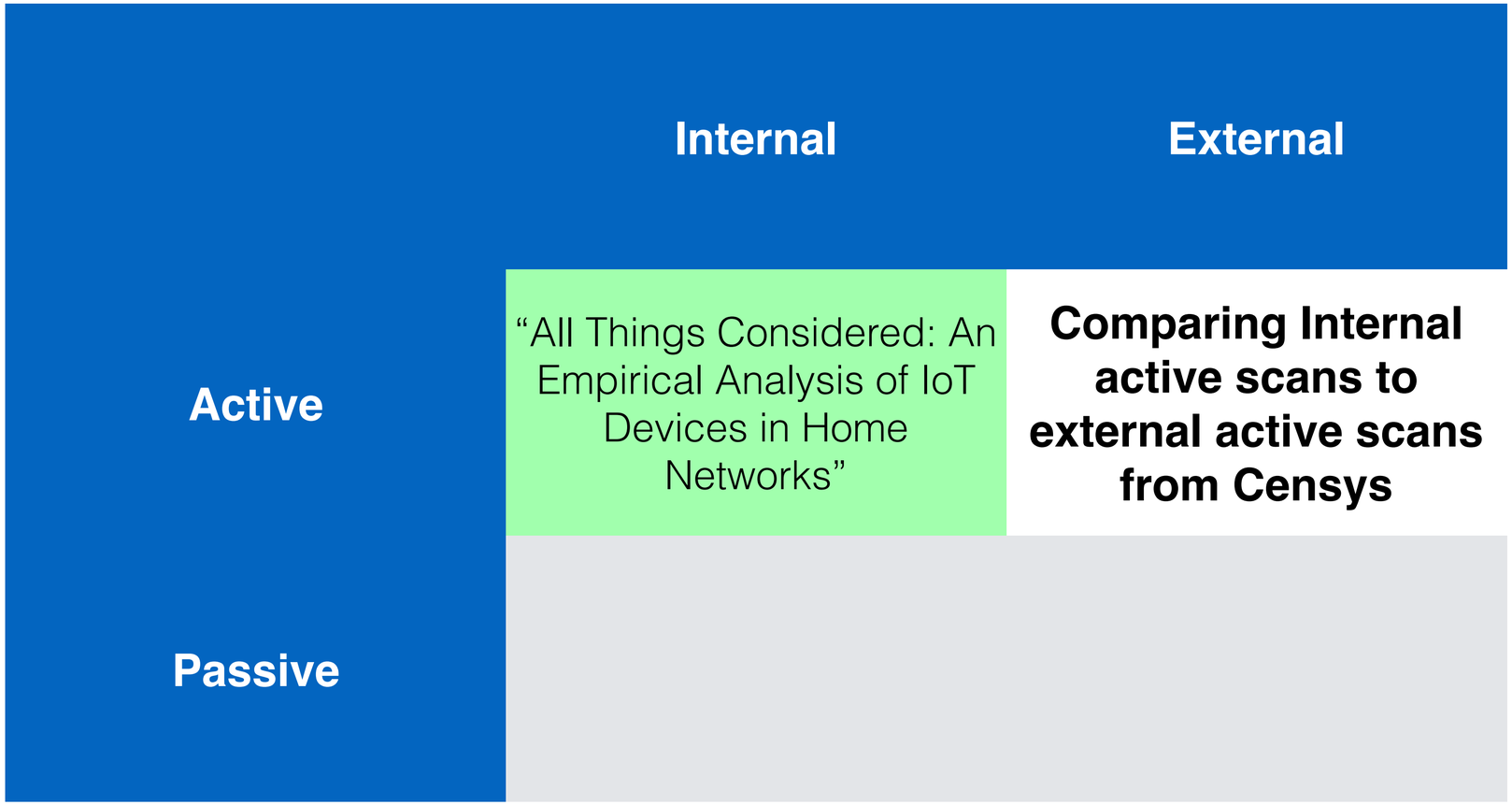
Region	% IoT Weak Telnet	% Surveillance
North America	0.5%	3.7%
South America	**4.9%**	**13.3%**
Eastern Europe	**3.0%**	**14.0%**
Western Europe	1.0%	5.6%
East Asia	0.4%	9.1%
Central Asia	**4.9%**	**30.3%**
SE Asia	**3.6%**	**37.0%**
South Asia	**2.9%**	**54.5%**
Oceania	0.7%	4.3%
N. Africa + Middle East	**4.8%**	**28.5%**
Sub-Saharan Africa	**1.1%**	**18%**

Mirai Infections



What can other perspectives and techniques tell us?

Proposed Work



Active, External Scans

- Active, external scans form the foundation of much research in the measurement community

Active, External Scans

- Active, external scans form the foundation of much research in the measurement community
  - ZMap, Censys, Shodan, Massscan have changed our *access* to data

Project 1: Proposal

- In this project, I propose comparing the external, active measurement perspective (Censys) to the internal, active measurement perspective (Avast)

Project 1: Proposal

- In this project, I propose comparing the external, active measurement perspective (Censys) to the internal, active measurement perspective (Avast)
  - *What are the measurement biases introduced by only studying the public Internet?*

Research Plan

- Collect raw data from Censys
  - Censys regularly scans IPv4 space on a fixed set of ports and collects application layer data
- Investigate and explain network services, device type distributions differences between two vantage points
- Tie back into published measurement research

Proposed Work

	Internal	External
Active	“All Things Considered: An Empirical Analysis of IoT Devices in Home Networks”	**Comparing Internal active scans to external active scans from Censys**
Passive		

Proposed Work

	Internal	External
Active	“All Things Considered: An Empirical Analysis of IoT Devices in Home Networks”	Comparing Internal active scans to external active scans from Censys
Passive	**Comparing Internal active scans to Internal passive scans**	

Internal Scans

- Internal scans are harder to come by in the research community, but a historically desired perspective

Internal Scans

- Internal scans are harder to come by in the research community, but a historically desired perspective
  - Netalyzer, Bismark were deployed at smaller scale to investigate network bandwith, misconfigurations, security problems

Project 2: Proposal

- In this project, I propose comparing an active, internal perspective with a passive, internal perspective

Research Plan

- Partnered with IoT-Inspector team to instrument their tool to perform passive *and* active scanning inside a home
  - Currently, the tool works by ARP-spoofing and serving as a MiTM, logging aggregate statistics and some flow data
- Deploy the tool to ~10K users currently on the waiting list
- Enumerate the differences between the two perspectives
  - *What don't you see by studying client behavior alone?*

Proposed Work

	Internal	External
Active	“All Things Considered: An Empirical Analysis of IoT Devices in Home Networks”	Comparing Internal active scans to external active scans from Censys
Passive	Comparing Internal active and Internal passive scans	

Future Directions

- How do users actually *configure* their IoT devices?
  - Partnered with IFTTT, a trigger-action platform that enables users to configure “network rules” for their homes
  - Starting work with Prof. Bates
- Exploring the passive external perspective for device fingerprinting and device enumerations
  - DNS can be a way to fingerprint devices (Alrawi et. al, IEEE S&P 2019)