RAPID: ROBUST AND ADAPTIVE DETECTION OF DISTRIBUTED DENIAL-OF-SERVICE TRAFFIC FROM THE INTERNET OF THINGS

Samuel Mergendahl and Jun Li

June 30th, 2020

University of Oregon

Center for Cyber Security and Privacy (CCSP)

IEEE Conference on Communications and Network Security

UNIVERSITY OF OREGON

Center for Cyber Security and Privacy

INTRODUCTION

- **The Internet of Things (IoT)**
  - Influx of novel applications with nearly 7 billion Internet connected devices in 2018
- IoT networks often exhibit poor security practices
  - E.g., default passwords

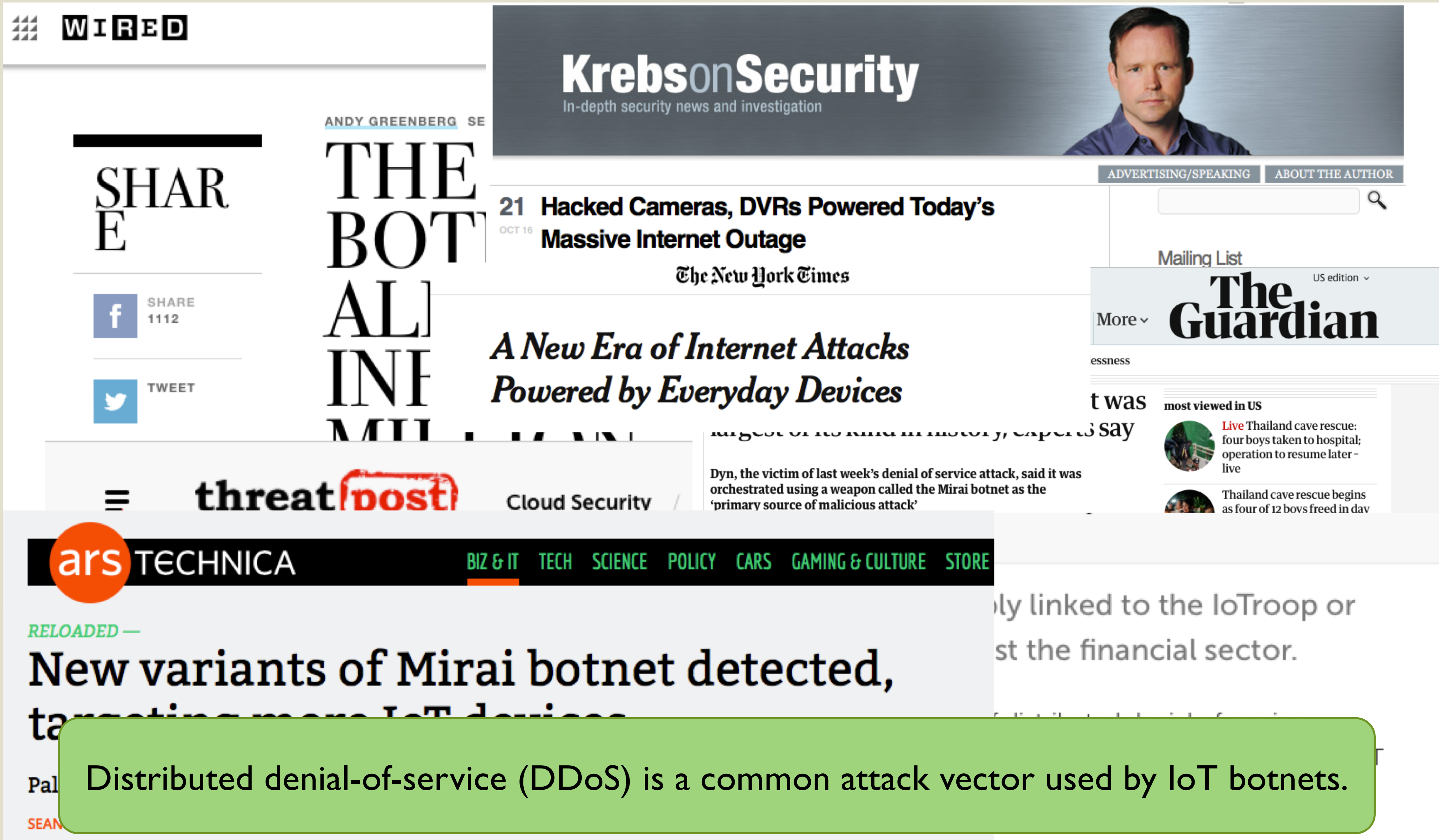
Password	Device Type	Password	Device Type	Password	Device Type
123456	ACTi IP Camera	klv1234	HiSilicon IP Camera	1111	Xerox Printer
anko	ANKO Products DVR	jvbzd	HiSilicon IP Camera	Zte521	ZTE Router
pass	Axis IP Camera	admin	IPX-DDK Network Camera	1234	Unknown
888888	Dahua DVR	system	IQinVision Cameras	12345	Unknown
666666	Dahua DVR	meinsm	Mobotix Network Camera	admin1234	Unknown
vizxv	Dahua IP Camera	54321	Packet8 VOIP Phone	default	Unknown
7ujMko0vizxv	Dahua IP Camera	00000000	Panasonic Printer	fucker	Unknown
7ujMko0admin	Dahua IP Camera	realtek	RealTek Routers	guest	Unknown
666666	Dahua IP Camera	1111111	Samsung IP Camera	password	Unknown
dreambox	Dreambox TV Receiver	xmhdipc	Shenzhen Anran Camera	root	Unknown
juantech	Guangzhou Juan Optical	smcadmin	SMC Routers	service	Unknown
xc3511	H.264 Chinese DVR	ikwb	Toshiba Network Camera	support	Unknown
OxhlwSG8	HiSilicon IP Camera	ubnt	Ubiquiti AirOS Router	tech	Unknown
cat1029	HiSilicon IP Camera	supervisor	VideoIQ	user	Unknown
hi3518	HiSilicon IP Camera	<none>	Vivotek IP Camera	zlxx.	Unknown
klv123	HiSilicon IP Camera				

Default passwords leveraged by Mirai to create a large-scale IoT botnet.

Antonakakis et al. Usenix Security, 2017

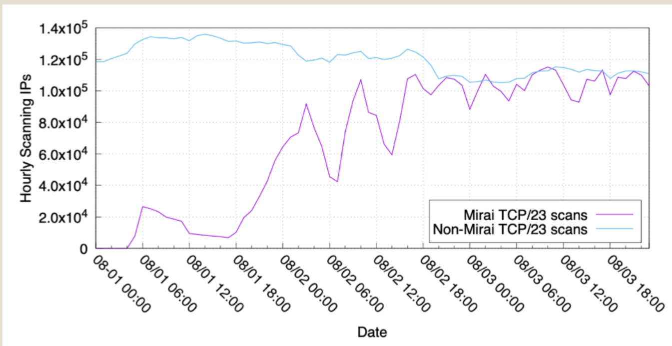
IoT devices often become compromised and recruited into large-scale botnets.

IOT-ENABLED DDOS



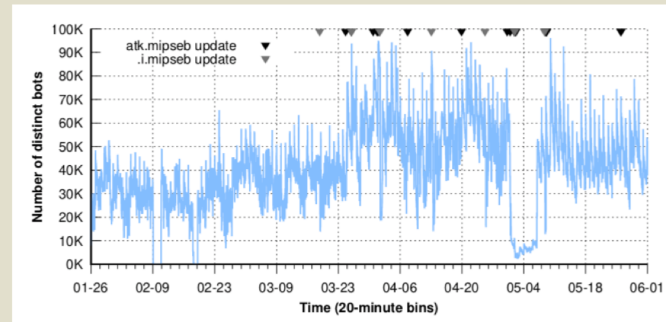
Distributed denial-of-service (DDoS) is a common attack vector used by IoT botnets.

IOT-ENABLED DDOS



Mirai infection rate.

Antonakakis et al. Usenix Security, 2017



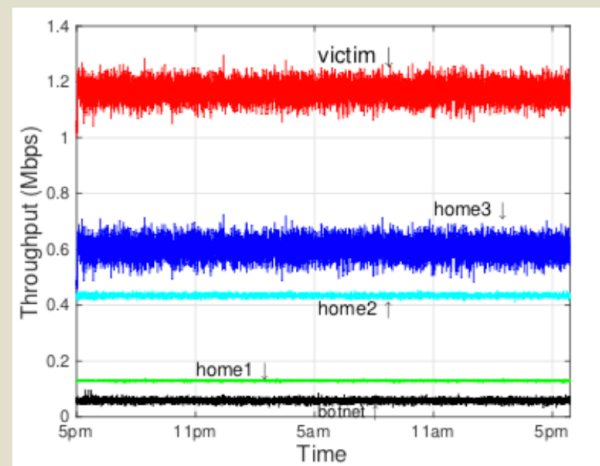
Number of unique Hajime bots over time.

Herwig et al. NDSS, 2019

Attack Type	Attacks	Targets	Class
HTTP flood	2,736	1,035	A
UDP-PLAIN flood	2,542	1,278	V
UDP flood	2,440	1,479	V
ACK flood	2,173	875	S
SYN flood	1,935	764	S
GRE-IP flood	994	587	A
ACK-STOMP flood	830	359	S
VSE flood	809	550	A
DNS flood	417	173	A
GRE-ETH flood	318	210	A

Mirai attacks between Sep. 2016-Feb. 2017.

Antonakakis et al. Usenix Security, 2017



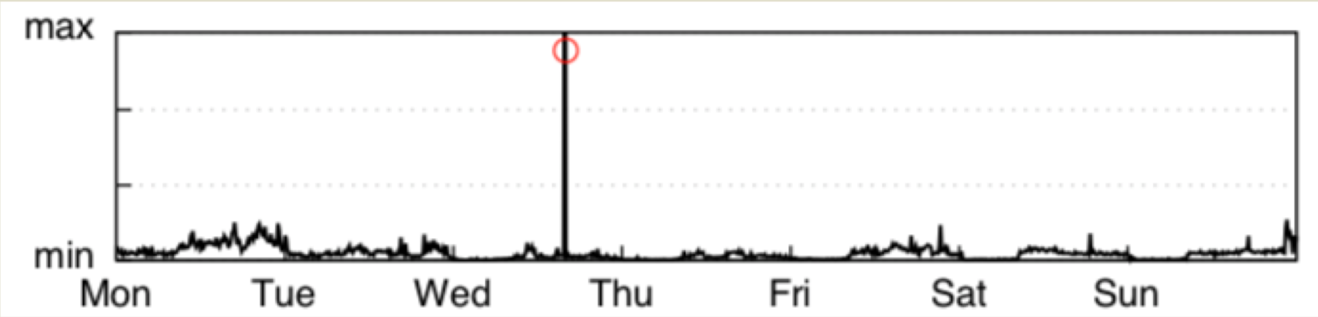
IoT smart-home reflection capability.

Sivaraman et al. WiSec, 2017

Recent quantitative studies suggest IoT-enabled DDoS is a massive threat.

ANOMALY DETECTION

- Key idea:
  - Malicious DDoS traffic exhibits statistically different behavior than normal, benign traffic
  - Detect statistical outliers ⇔ Detect malicious traffic
- Derivation techniques for classification boundaries:
  - Manual statistical investigation of previous traffic to select **static thresholds**
  - Automatic derivation through **machine learning algorithms** that train on past traffic
    - Black-box approach with **neural networks**



Liu et al. Internet Measurement Conference, 2015

Anomaly detection plays a pivotal role in the detection and mitigation of IoT-enabled attacks.

IOT ANOMALY DETECTION: REAL-WORLD DEPLOYMENT CHALLENGES

1. **Sufficient accuracy**
   - False positives in detection lead mitigation to drop benign traffic
   - Causes increased retransmission and energy consumption for constrained IoT devices
2. **Easy deployment**
   - Many IoT networks deploy through non-security professionals
   - Cannot rely on manual parameter tuning to achieve sufficient accuracy
3. **Domain shift**
   - Heterogeneity of IoT leads to the failure of pre-trained models
4. **Explainable classifications**
   - IoT often interacts with the physical world
   - Must allow a human-in-the-loop to make structured changes if needed

These are **conflicting challenges**:

One specific IoT deployment challenge often ***fundamentally neglects or contradicts*** *a different IoT deployment challenge.*

IOT ANOMALY DETECTION: DESIGN GOALS

Design Goals of Rapid

1. **Achieve impressive accuracy**
   - Currently through neural networks
2. **Provide a pre-trained model**
   - Ready to deploy in any IoT network
3. **Operate in real-time**
   - Extract computationally efficient features
4. **Provide diagnostic insight**
   - Special design of neural network
5. **Automatically adapt to domain shift**
   - Leverage a novel active learning technique

Rapid: Robust and Adaptive Detection

Past anomaly detection focus:

Opprentice
Liu et al.
IMC, 2015

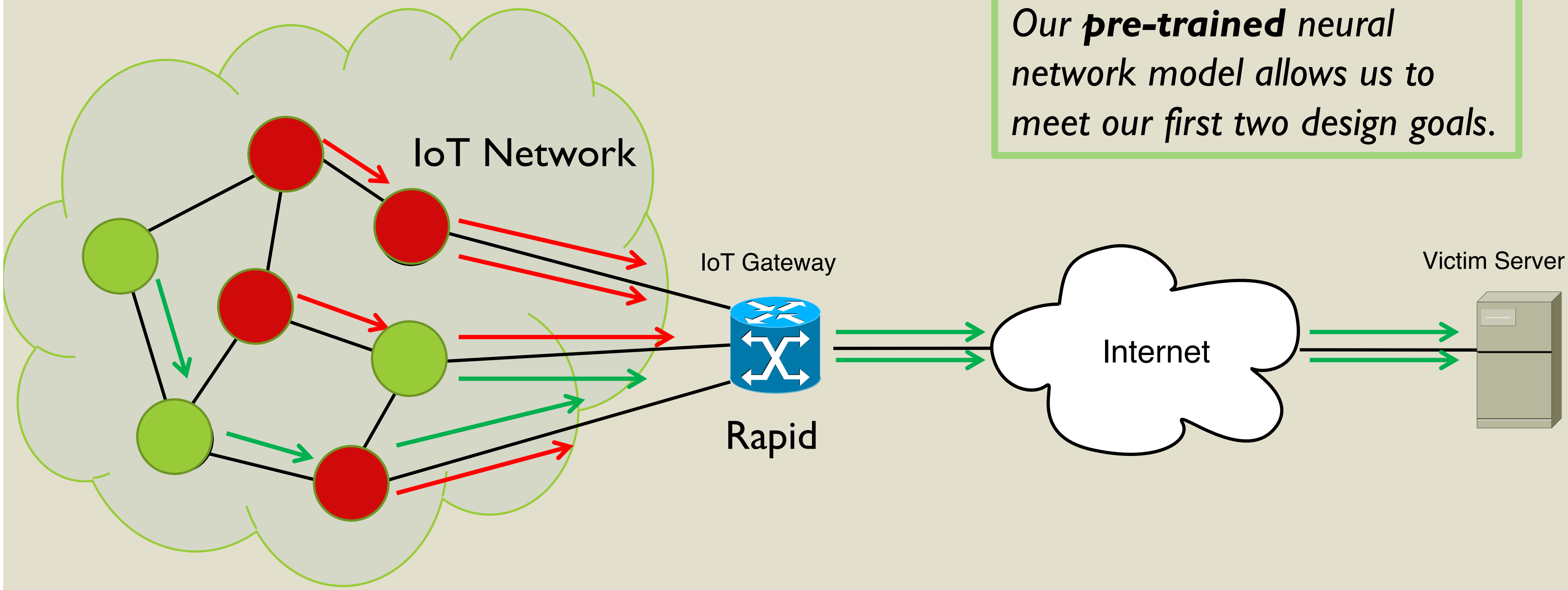
DeepLog
Du et al.
CCS, 2017

IDS-NNM
Linda et al.
IJCNN, 2009

RAPID: OVERVIEW

- Rapid resides at the gateway of a **generic** IoT network
  - E.g., Rapid can defend a smart-home, healthcare facility, large-scale factory, etc.

*Our **pre-trained** neural network model allows us to meet our first two design goals.*



Rapid deploys a neural network to detect any DDoS traffic that leaves the IoT network.

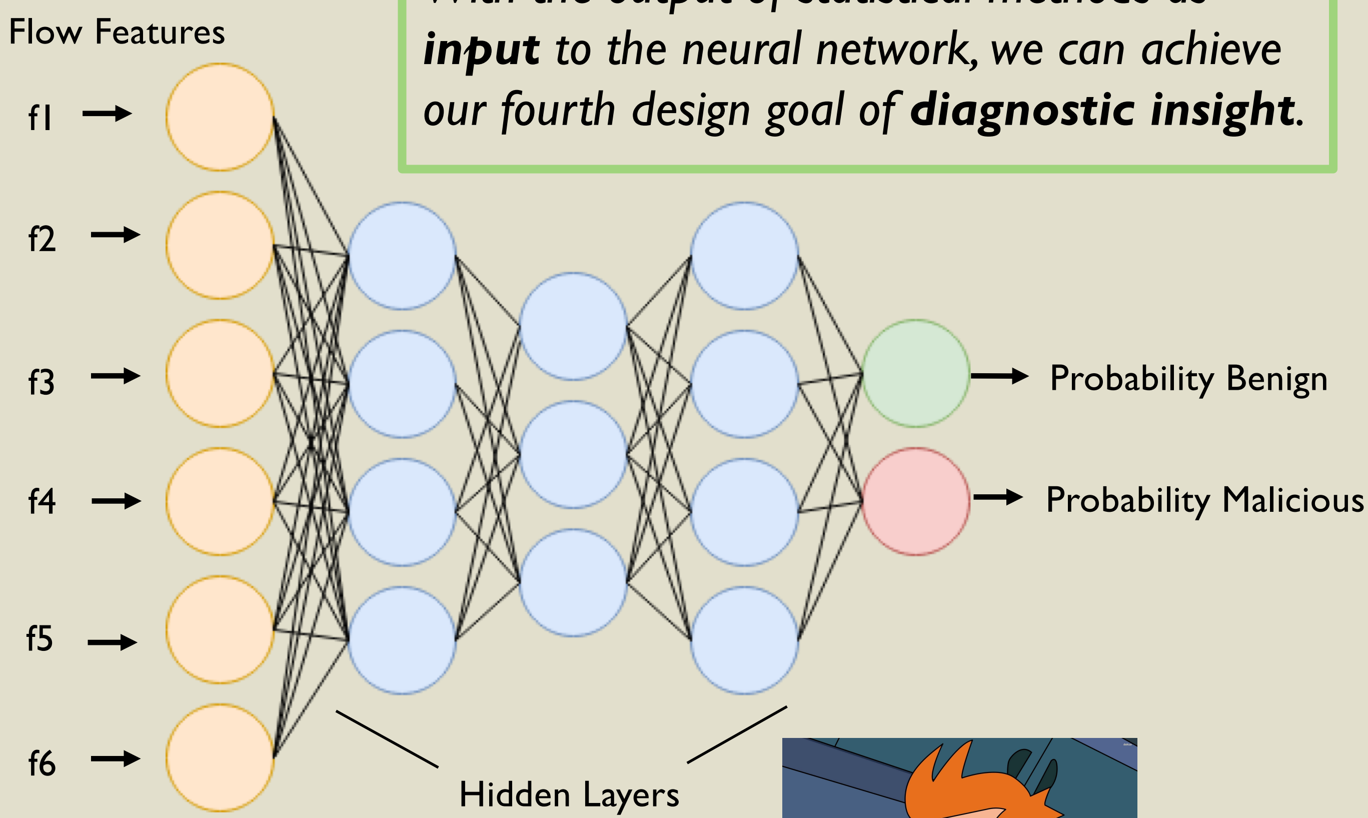
FLOW PRE-PROCESSING: REAL-TIME OPERATION

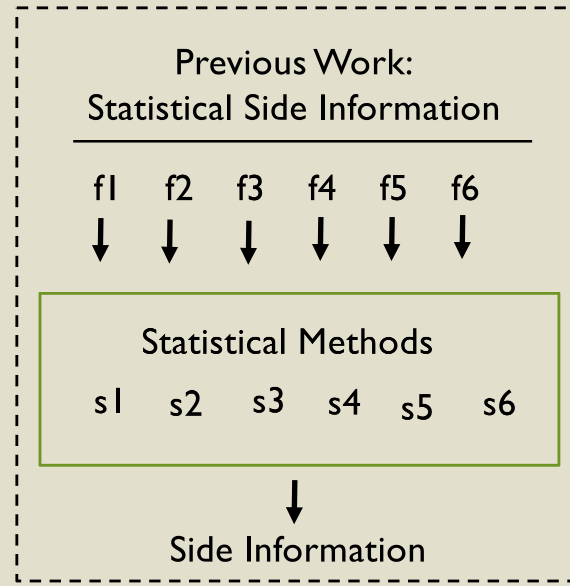
- Collect **sFlow** streams at gateway and separate into:
  - **Aggregate Flows**
    - Each flow has the same external IP address
    - Used for Attack Detection (not discussed in this presentation)
  - **Granular Flows**
    - Each flow has the same internal and external IP address
    - Used for Attack Classification
- Extract **four features** for each flow during each time window:
  1. Total outgoing bytes
  2. Ratio of incoming/outgoing bytes
  3. Total outgoing packets
  4. Ratio of incoming/outgoing packets
- We call these features ***basic detectors***
  - Early DDoS detection solutions directly used these for detection (with thresholds)

*Our **computationally efficient** and well-studied DDoS features allow us to meet our third design goal.*

DIAGNOSTIC-AWARE CLASSIFICATION

With the output of statistical methods as ***input*** *to the neural network, we can achieve our fourth design goal of* ***diagnostic insight****.*



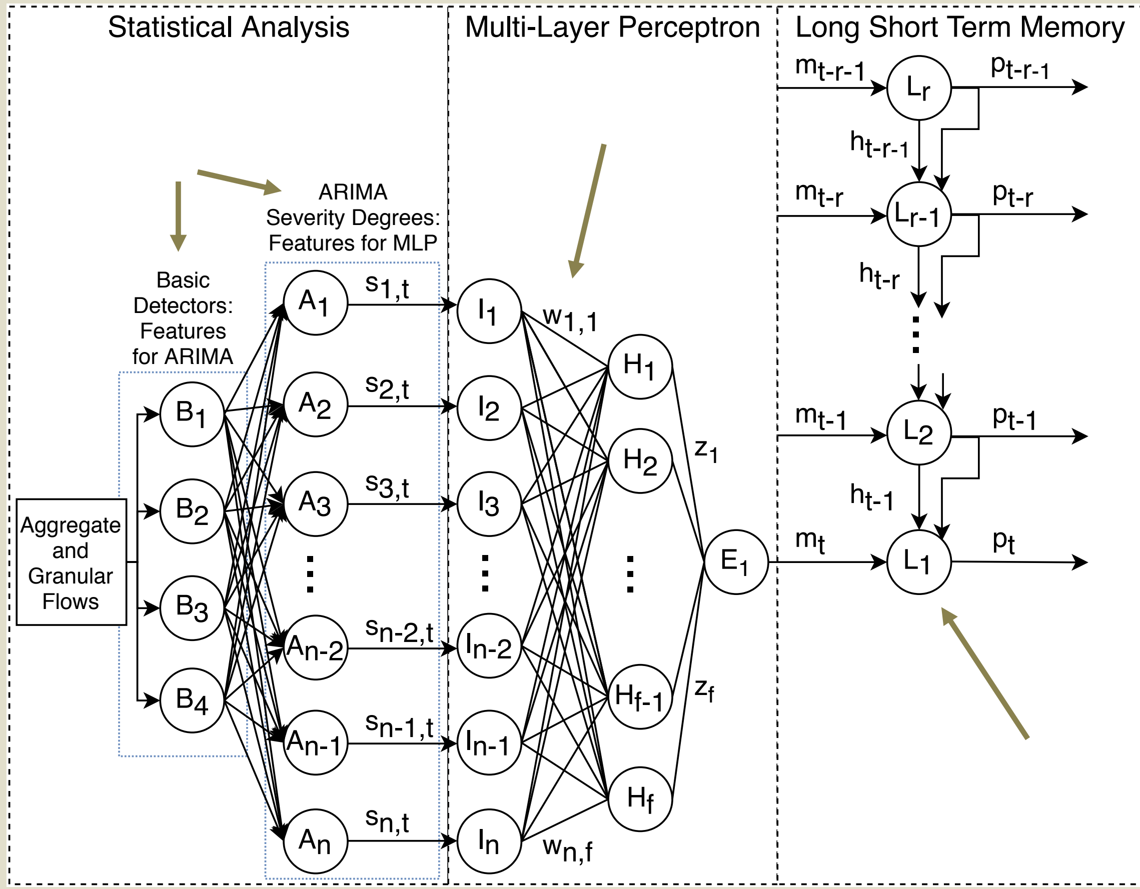


Fundamentally can't know if the side information is truly related to the classification

Rapid employs a special neural network design to preserve diagnostic insight.

ENSEMBLED CLASSIFICATION WITH DEEP LEARNING

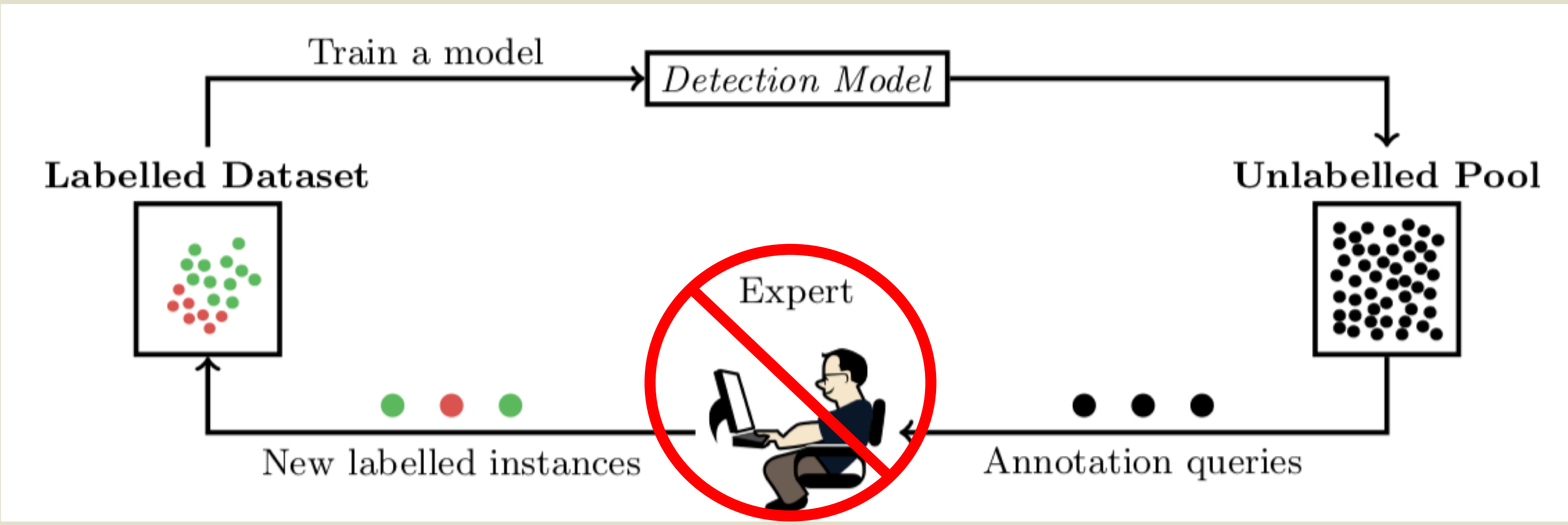
- We use Auto-regressive Integrated Moving Average (**ARIMA**) as our statistical analysis
  - ARIMA forecasts the next value in each basic detector time-series
  - Each ARIMA algorithm outputs a ***severity degree***
- A Multi-Layer Perceptron (**MLP**) ensembles the severity degrees
- Long Short Term Memory (**LSTM**) analyzes the output of the MLP
  - Over many time windows
  - Outputs a single severity degree



Rapid ensembles ARIMA severity degrees with an MLP and LSTM.

DOMAIN ADAPTATION WITH ACTIVE LEARNING

- Unfortunately, a pre-trained model will fail when ported to a new environment
  - I.e., domain shift causes trained systems to fail
  - Can use ***active learning*** to collect new labeled data and re-train under domain shift
- Current active learning solutions are not sufficient for IoT networks
  - The network operator of many IoT networks is not a security professional

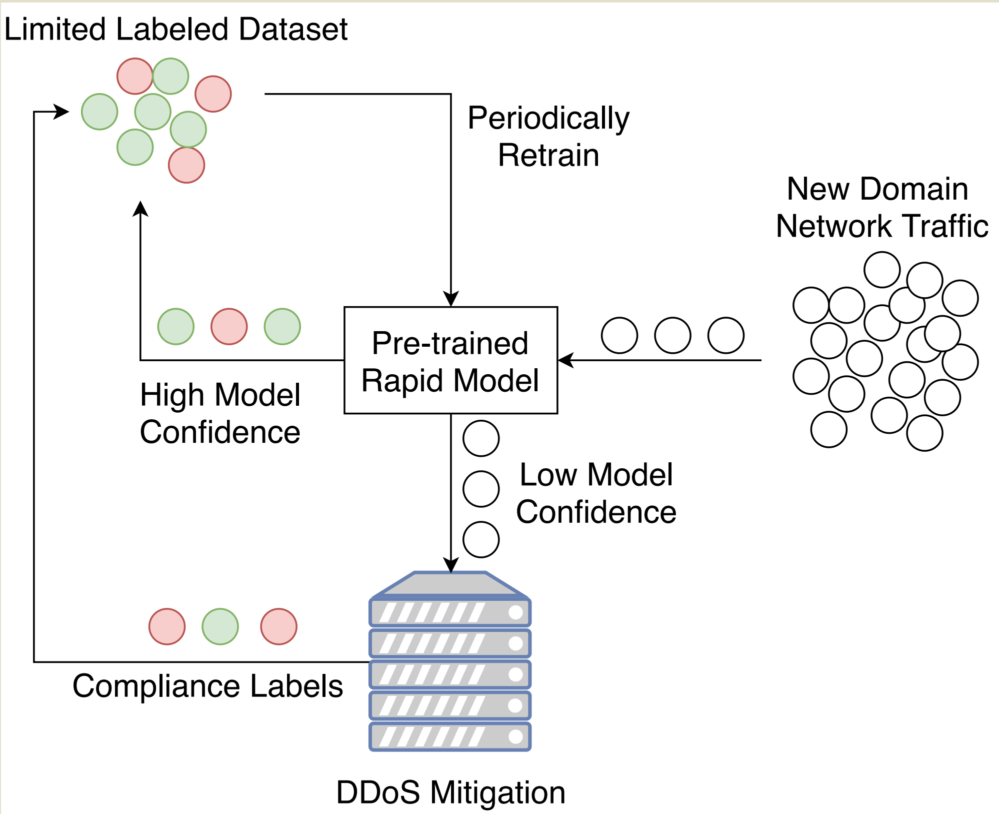


Active learning example.

Beaugnon et al. RAID, 2017

Rapid must perform active learning to combat domain shift.

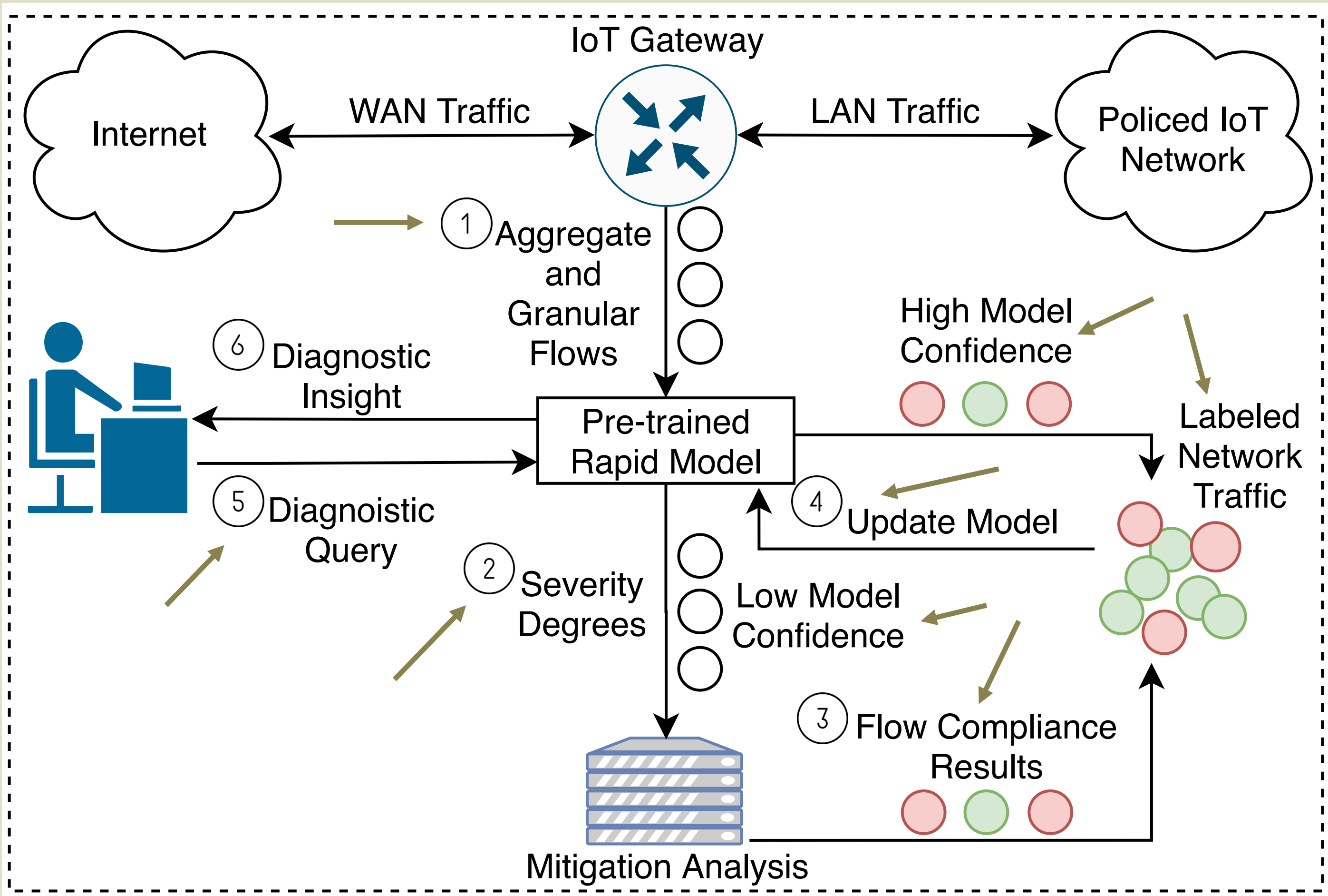
DDOS MITIGATION AS SECURITY EXPERT



- Replace the security expert with comprehensive DDoS mitigation
  - Automates the process for this particular domain
- Recent DDoS attack mitigation:
  - A connection's response to traffic engineering techniques can further identify malice
    - E.g., Dropping a TCP connection
      - Should result in reduced send rate
- Treat mitigation compliance as the labels for low model confidence

Rapid interweaves with attack mitigation to adapt to new domains without a security expert.

RAPID SYSTEM REVIEW



Rapid detects IoT-enabled DDoS with high accuracy, domain adaptability, and diagnostic insight.

EVALUATION OVERVIEW

- Evaluation goals:
  - Accuracy of Rapid compared to state of the art anomaly detection systems
    - Opprentice (Random Forest), IDS-NNM (MLP), DeepLog (LSTM)
  - Test Rapid under domain shift
    - Sensitivity and specificity
    - Model calibration and reliability
  - Attack detection flexibility

*Our evaluation demonstrates the real-world **deployability** of Rapid.*

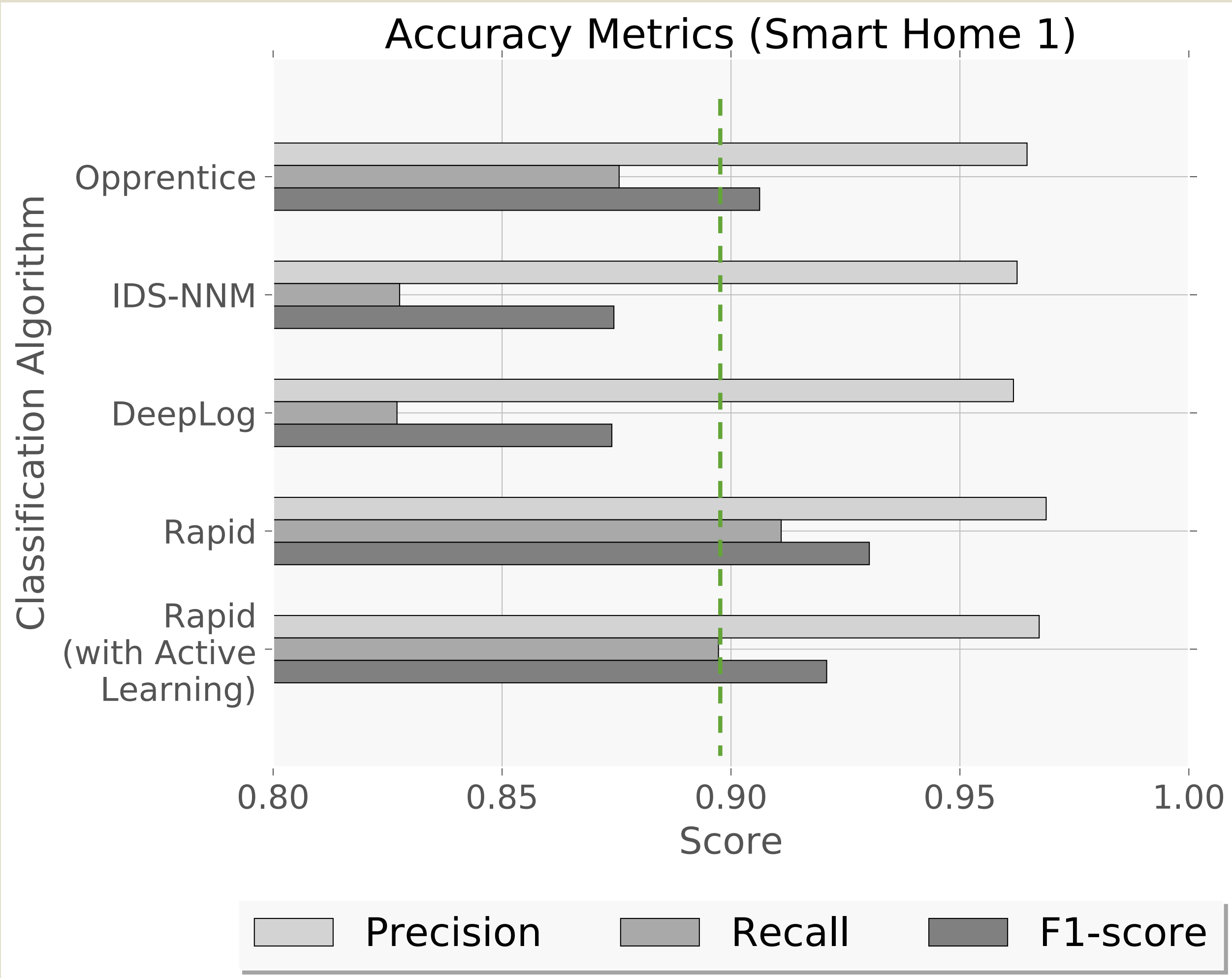
- Datasets
  - Test Rapid under multiple types of IoT traffic
  - Test Rapid under multiple types of DDoS attacks

Typical evaluation of train-test split

Dataset	Benign IoT Traffic	Benign Non-Iot Traffic	UDP Flood	TCP SYN Flood	HTTP Flood	DNS Flood
Smart Home 1 [32]	✓	✓	✓	✓	✓	✗
Smart Home 2 [33]	✓	✗	✗	✗	✗	✗
Smart Hospital [34]	✓	✓	✗	✗	✗	✗
CAIDA DDoS Attack [35]	✗	✗	✗	✓	✗	✗
Booter DDoS Attack [36]	✗	✗	✗	✗	✗	✓
DARPA DDoS Attack [37]	✗	✓	✗	✓	✗	✗

Never seen during training

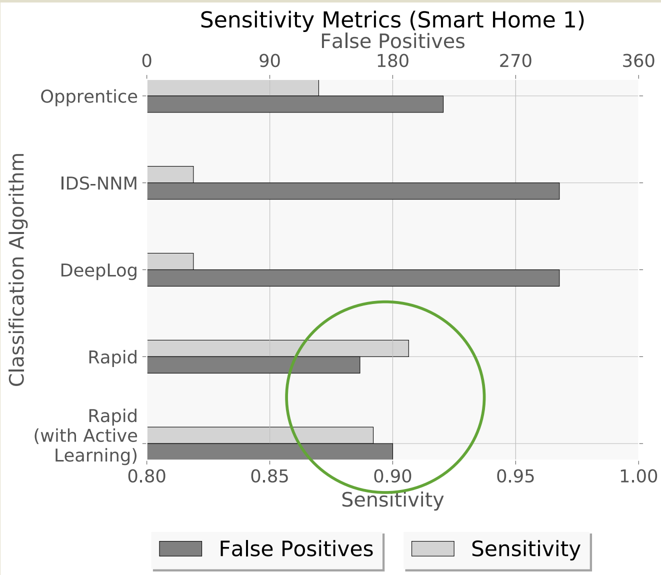
ACCURACY

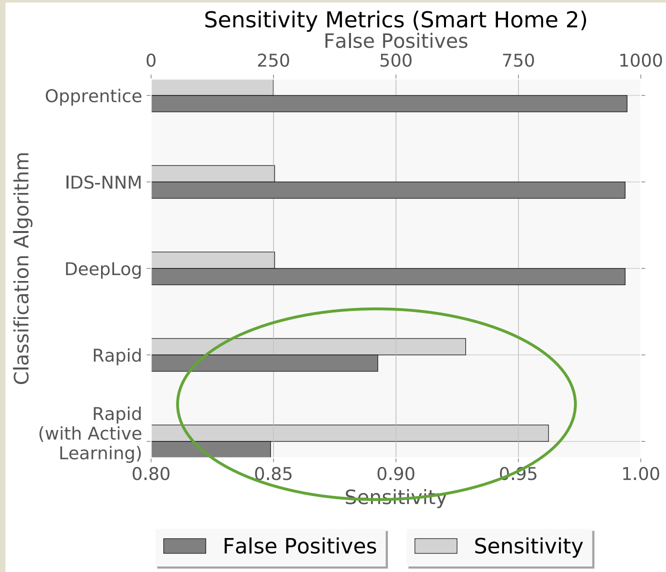


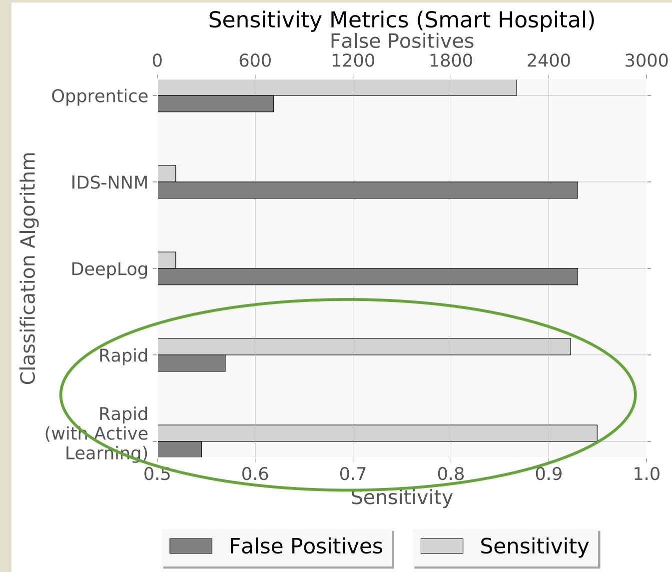
- Precision
  - TP / (TP + FP)
- Recall
  - TP / (TP + FN)
- F1-score
  - 2TP / (2TP + FP + FN)

Rapid achieves state of the art accuracy.

SENSITIVITY AND FALSE POSITIVES



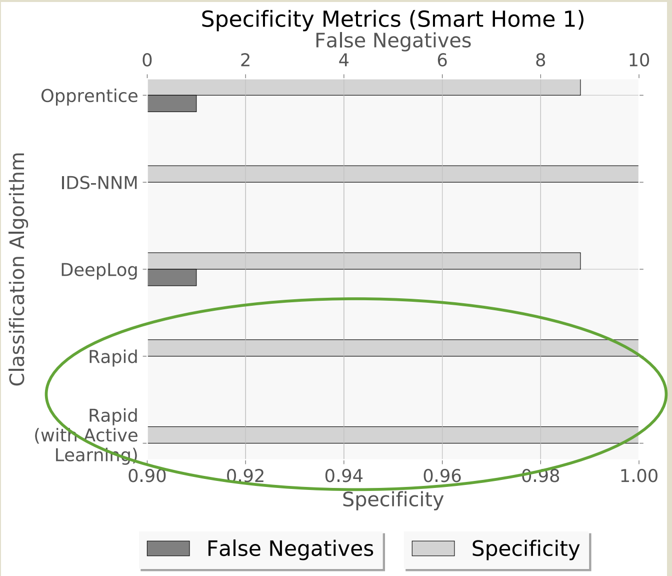


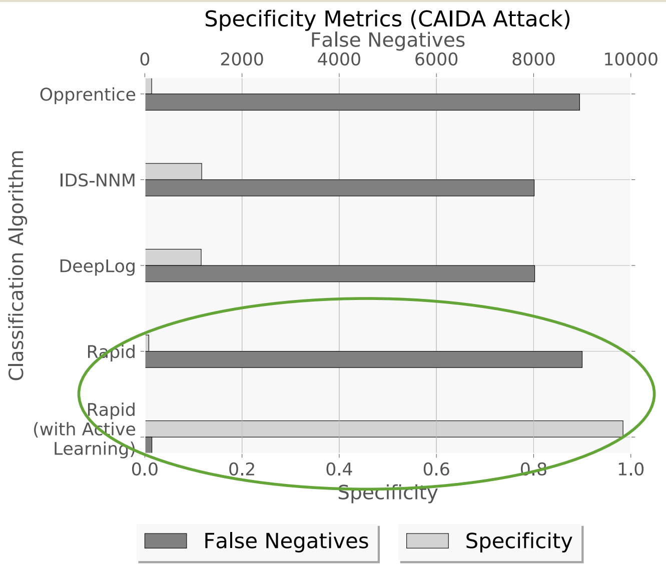


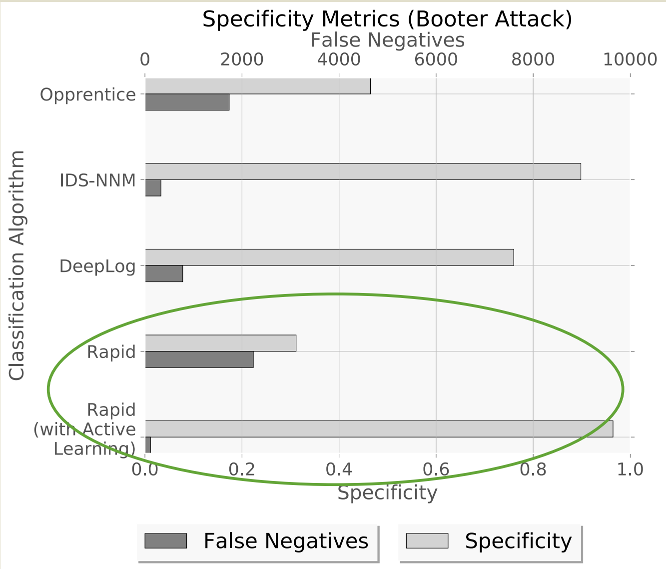
- Sensitivity
  - TP / (TP + FN)
- False positives
  - Cannot show false positive rate since TN = 0

Rapid reduces false positives and improves sensitivity under domain shift.

SPECIFICITY AND FALSE NEGATIVES







- Specificity
  - TN / (TN + FP)
- False negatives
  - Cannot show false negative rate since TP = 0

Rapid reduces false negatives and improves specificity.

CONCLUSION

- We presented a new anomaly detection system, Rapid
  - Detects IoT-enabled DDoS attacks at the gateway of an IoT network
  - Specifically designed for real-world deployment

- Key features of Rapid:
  - Leverages neural network techniques for **state of the art accuracy**
  - **Automatically adapts** to domain shift with novel active learning techniques
  - Provides **diagnostic insight** into classifications
  - **Comprehensive evaluation** of multiple real-world IoT and DDoS datasets

ACKNOWLEDGEMENTS

Special thanks to
my collaborator Dr. Jun Li
and the rest of
the Center for Cyber Security and Privacy
at the University of Oregon

Thanks for listening!