IMAGINE THE POSSIBILITIES...
Smart Gateways: Active A.I. In Subscriber Networks

Kyle Haefner
Senior Security Engineer
CableLabs
What is A.I.

Artificial Intelligence—Advanced Iteration (A.I.)

Mostly when we talk of A.I. we are really referring to Machine Learning (M.L)

“Machine Learning is the study of computer algorithms that allow computer programs to automatically improve through experience”

– Tom Mitchell, Carnegie Mellon

“Experience” is really iterating over data and updating weights.
**Scale of IoT.**

By end of 2020 there will be 50 Billion IoT Devices

At this scale (50 billion) one compromised device in every 83,000 could create a botnet equal in size to the Mirai Botnet of 2016.
**Device Centric Network Control**

### Academic Examples (Not Comprehensive)

**Princeton IoT Inspector**
https://iot-inspector.princeton.edu/

“An open-source desktop tool with a one-click install process. Automatically discovers IoT devices and analyzes their network traffic.”

**IoT Sentinel: Technische Universität**
https://arxiv.org/abs/1611.04880

“...a system capable of automatically identifying the types of devices being connected to an IoT network and enabling enforcement of rules for constraining the communications of vulnerable devices...”

**AuDI:**
Aalto University
https://tuprints.ulb.tu-darmstadt.de/8511/

“...a probabilistic framework for providing meaningful feedback in device identification, particularly when the device has not been previously observed.”

**DeviceMien:**
Rutgers University

“...a probabilistic framework for providing meaningful feedback in device identification, particularly when the device has not been previously observed.”

---

![Diagram of IoT Sentinel system design](image-url)
Predictability of a device is correlated to the complexity of a device.

Determine where device is on this scale.

**Low Complexity**
- Sensors
- Light Bulbs
- Single Purpose

**High Complexity**
- Laptops
- Smartphone
- General Purpose
Noise to Signal Ratio (NSR)

DBSCAN
Uses density based spatial distance
- Signal are clusters
- Noise non-clusters
NSR Complexity of Devices

Netflow Data

- Connection Based
  - Destination IP addresses
  - Destination Ports
- Traffic Based
  - Packets/s
  - Bytes/s
Behavior

Gaussian Mixture

Set of Gaussians make up behavior

- All points that are more than 2 standard deviations are outliers
- In two-dimensional space these form ellipses.
  - Not limited to two dimensions
- Behavior is the set of all of these Gaussians for a device.
Low Complexity

- Sensors
- Light Bulbs
- Single Purpose

High Complexity

- Laptops
- Smartphone
- General Purpose
Model Effectiveness

F1 Scores for Home devices

- F1 score is the harmonic mean between precision and recall.
- Perfect F1 score is 1.0
- Non-normalized shows correlation between accuracy and NSR complexity
Thank You!

Kyle Haefner
Senior Security Engineer
CableLabs
k.Haefner@cablelabs.com