📱

Link-Layer Device Classification

Research Proposal

Research Statement:

IoT devices are now more prevalent in modern smart home, industrial, government, health, business, and many other impactful societal settings. It has been shown that devices, simply based off of signals alone, can be classified into their respective types using techniques such as machine learning. This may be undesirable from a privacy or security standpoint as knowing the devices present can lead to a larger attack surface: adversaries can exploit known vulnerabilities, cripple a network with DoS attacks, perform further inference-based attacks (peek-a-boo) to gain information about personnel interaction on these networks.

Research Goal:

Perform, and mitigate against, machine learning-based device classification on the link layer (for Bluetooth, Zigbee, and WiFi).

Research Question(s):

1. Is it possible to successfully identify IoT devices using machine learning on the link-layer?
2. Is it viable to supplement the encrypted traffic with unencrypted traffic to achieve high classification results?
3. Can diffusion-based generative modeling be used to mitigate, or defeat, link-layer IoT identification classifiers?

Research Plan:

1. Take existing data we have (or collect more data for testing/validation) and perform basic multi-class classification on it. Performance will answer question.
    a. We can benefit from creating more traffic, and collecting with more robust methods, for longer periods of time. (Bluetooth and Zigbee specifically).
2. We can take other datasets (CIC-IOT, UNSW, etc) **UNENCRYPTED,** and train a model from this, and use our collected sets **(LINK-LAYER ENCRYPTED)** as a testing set. The performance will answer the question, because they are two distinct types of sets.
    a. We can either strive for protocol agnosticism again, or still use the tool but have distinct models for each protocol instead of one for all three. We get more of an idea of which protocols have the highest viability.

[MOTIVATION FOR #3] GANs have been used and are great, but have downfalls to them; diffusion-based models are now state-of-the-art and are more powerful than GANs. This is a unique thing for our space so far.

3. We have a trained model to classify devices from the prior question(s). Based off of what is generated, a flow still needs to be input (tool usable here) to the model to see what it does.
    a. OPTION ONE: Cloaking device - we learn to generate packets to emulate device behavior (conditioning) and inject that into the network. This is more practical, so we would see if the model detects that the generated traffic successfully cloaks by classifying that device
    b. OPTION TWO: Complete mitigation (harder) - Have a type of feedback like a GAN but using diffusion. We can either generate packet captures or flowtables, to still beat out the classifier on the other end. TL;DR: Replicate a GAN but with diffusion.

Related Work:

Related Work:

Background:

- GAN
- Diffusion
- Device Classification

Methods:

Part I: Link-Layer Device Classification

Part II: Data Supplementation

Part III: Diffusion-Based Modeling for Mitigation

Device Abstraction Mapping

DEVICE	ABSTRACT CATEGORY
ASUS Router RT-N12	Router
ASUS RT-AC1200GE	Router
August Smart Lock	Smart Lock
Barnes & Noble Nook	ereader
Withings Blood Pressure monitor	Smart BP Monitor
Bose Home Speaker 300	Smart Speaker
C by GE 3-Wire On/Off Toggle	Smart Switch
Echo W. Hub	Smart Assistant
Fitbit 4 Health & Fitness Tracker	Fitness Tracker

Galaxy A21	Smartphone
Garmin Index S2 Smart Scale	Smart Scale
iRobot Roomba	Smart Vacuum
Kindle	Ereader
Kinsa Quickcare Smart Thermometer	Smart Thermometer
PETKIT WiFi Feeder	Smart Pet Feeder
Phillips Hue Bridge	Smart Bridge
Pixel 4a	Smartphone
Samsung Galaxy Watch Active	Fitness Tracker

ZIGBEE DEVICE ADDRESSES:

Philliphs Hue Bridge	00:17:88:01:05:45:c1:86	0x0001
Phillips Hue Bulb	00:17:88:01:0b:5c:9e:15	0x2ce0
Phillips Hue Bulb	00:17:88:01:0b:61:37:39	0xf70f
Phillips Hue Bulb	00:17:88:01:0b:61:35:fe	0x8b39
Smart Plug	00:12:4b:00:22:eb:07:a7	0x1de6
Smart Plug	00:12:4b:00:22:e9:1c:f2	0x55fb
Smart Plug	00:12:4b:00:22:ea:9d:34	0xddaf
Smart Plug	00:12:4b:00:22:e9:20:c6	0xe645
Smart Plug	00:12:4b:00:22:ea:93:4f	0x5d81
Smart Plug	00:12:4b:00:22:d3:06:c7	0x3a95
Smart Plug	00:12:4b:00:22:ea:97:f9	0x7bd2
Smart Plug	00:12:4b:00:22:eb:07:eb	0x74fa
Smart Plug	00:12:4b:00:22:ea:ec:0d	0x8a91
Smart Plug	00:12:4b:00:22:e9:30:5c	0x237f
Tp-Link Kasa Router	00:12:4b:00:0a:e7:dc:55	0x000
Motion Sensor	00:12:4b:00:15:fe:9d:b0	0xff4f
Motion Sensor	00:12:4b:00:07:fc:f0:17	

Open/Close Sensor	00:12:4b:00:07:fc:cc:f4	0x335b
Open/Close Sensor	00:12:4b:00:07:fc:cc:ac/cc	0xf770
Alexa	00:15:5f:00:40:bc:2f:ee	0x0000

channels: 11, 15, 25