Physical (PHY) Layer Features for Device Authentication

Bryan Manzi[+], Gian Claude[*], Peter Ekpewoh[*], Timothy J. Pierson[+], Michel Reece-Kornegay[*] and David Kotz[+]

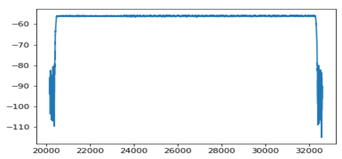
[+] Dartmouth College, [*] Morgan State University

Authenticating devices is essential to maintain IoT data integrity

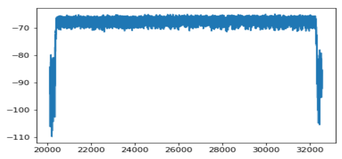
- Many researchers imagine networks of billions of deployed Internet of Things (IoT) devices collecting data about their local environment
- Data from these deployed devices is envisioned to be aggregated and analyzed to yield insight into the behavior of complex systems
- Adversarial devices, operating amongst legitimate devices, may try to masquerade as legitimate devices
- These adversaries may then try to inject inaccurate or misleading data into the network, potentially resulting in an erroneous characterization of the environment
- Reliable device authentication is needed to prevent such impersonation attacks

PHY-layer radio signal features could help authenticate wireless devices

- Currently, authentication is typically done at upper levels of the network stack using cryptography
- Key management and computational complexity can make cryptography difficult to implement and maintain on a large network of small, possibly mobile, computationally constrained IoT devices
- We evaluate radio signal features at the physical (PHY) layer as a complimentary approach to cryptography
- Difficult-to-clone manufacturing imperfections create device-unique PHY-layer signal ‘fingerprints’
- These unique fingerprints can be used to identify and authenticate devices, potentially stopping impersonation attacks



Legitimate device’s signal amplitude



Software Defined Radio impersonating a legitimate device

DARTMOUTH

MORGAN STATE UNIVERSITY

Device authentication process

1) Develop device signal profile

- When devices are first introduced (e.g., paired), devices note the signal ‘fingerprint’ of the new device
- Device creates profile of the new device based on signal fingerprint features

Machine learning

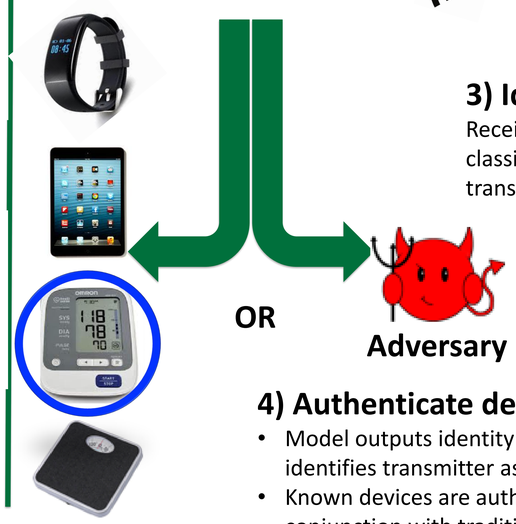
2) Create classification models

- Machine learning produces classification models based on signal profile
- Models attempts to identify individual devices (e.g., device number 123)
- Models may identify device types (e.g., iPhone X)

Classification models

3) Identify transmitter

Received signals features are input to classification models to identify the transmitting device



OR

Adversary

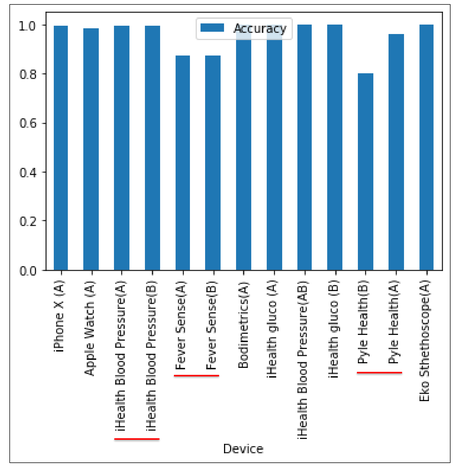
Legitimate devices

4) Authenticate device

- Model outputs identity of transmitting device or identifies transmitter as potential adversary
- Known devices are authenticated (perhaps in conjunction with traditional cryptography)
- User alerted to potential adversaries and device refuses to further communicate with adversary

Preliminary results are promising

We can distinguish between devices from the same manufacturer and between individual devices



Process

- Signals features input into trained models
- Device identity predicted

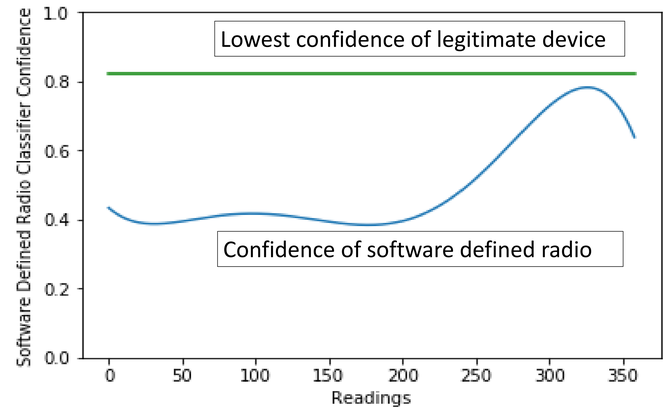
Performance

- Average Precision: 96.9%
- Average Accuracy: 97.7%
- Average Recall: 95.8%
- Average F1: 96.3%
- Average Confidence: 92.3%

Takeaways

- Some devices are easier to classify than others
- Classifier accuracy decreases if there are multiple instances of the same type of device

We can distinguish between legitimate devices and SDR impersonation attacks



Software Defined Radio (SDR) replay attacks are always detected

- HackRF[1] SDR replaying legitimate signals results in model predictions with low confidence
- Replayed signals always result in lower confidence than lowest confidence of legitimate devices

National Science Foundation Award CNS-1329686

[1] HackRF One https://greatscottgadgets.com/hackrf/one/