Recurring Verification of Interaction Authenticity Within Bluetooth Networks

@ The 14th ACM Conference on
Security and Privacy in Wireless and Mobile Networks (WiSec'21)
June 28th – July 2nd, 2021

Travis Peters[1], Timothy J. Pierson[2], Sougata Sen[3], José Camacho[4], David Kotz[2]

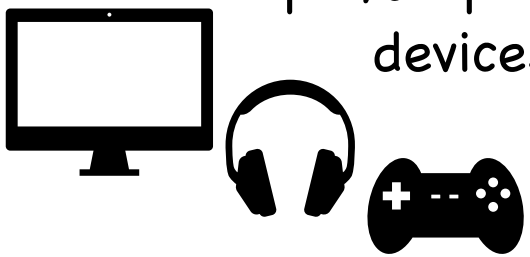
[1] Montana State University, USA (now @ Include Security)

[2] Dartmouth College, USA

[3] BITS Pilani, Goa Campus, India

[4] University of Granada, Spain

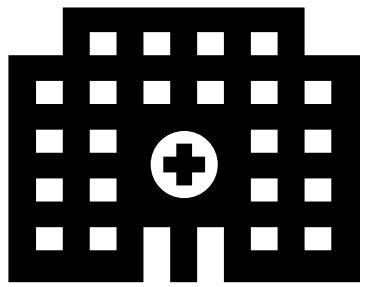
Traditional user input/output (I/O) devices



Smart home devices

Smart fitness & wellness devices

Smart medical devices



Problem: *Have I been compromised? Am I vulnerable?*

LILY HAY NEWMAN SECURITY 07.29.19 11:04 AM

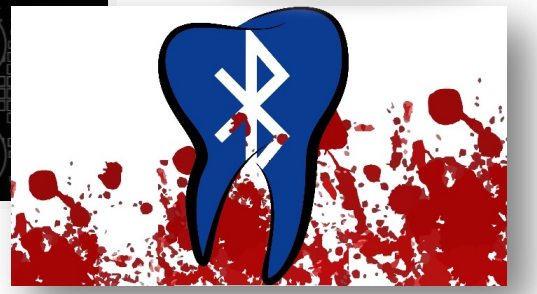
AN OPERATING SYSTEM BUG EXPOSES 200 MILLION CRITICAL DEVICES

Think of how the WannaCry ransomware used the Eternal Blue Windows vulnerability to spread across networks and around the world. It's like that, but with firewalls, industrial equipment, and medical devices instead of Windows machines. The result could be anything from device malfunctions to full system takedowns.

Forescout and JSOF Disclose New DNS Vulnerabilities, Impacting Millions of Enterprise and Consumer Devices

FORESCOUT RESEARCH LABS | APRIL 12, 2021

BlueBorne



Problem: *Unsatisfactory "solutions"...*

LILY HAY NEWMAN SECURITY 07.29.19 11:04 AM

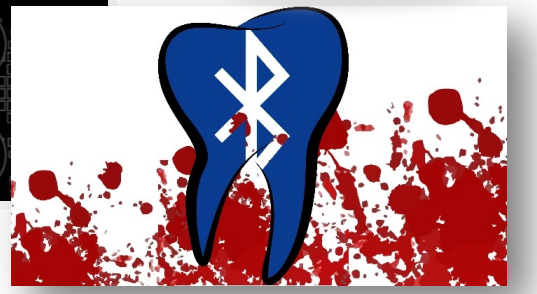
AN OPERATING SYSTEM BUG EXPOSES 200 MILLION CRITICAL DEVICES

Think of how the WannaCry ransomware used the Eternal Blue Windows vulnerability to spread across networks and around the world. It's like that, but with firewalls, industrial equipment, and medical devices instead of Windows machines. The result could be anything from device malfunctions to full system takedowns.

Forescout and JSOF Disclose New DNS Vulnerabilities, Impacting Millions of Enterprise and Consumer Devices

FORESCOUT RESEARCH LABS | APRIL 12, 2021

BlueBorne



The "solutions"...

TC

*"How do you stay safe? **Keep all of your devices updated** regularly and **be wary of older IoT devices**."*

THE VERGE

*"The simplest protection is to **leave Bluetooth off**, but since phones are still vulnerable when they're connected to a Bluetooth device, **the only recommendation is not to use Bluetooth at all**."*

There must be a better way...

Problem: *The challenges are plentiful…*

Challenges

- Complex ecosystem
- Inability to modify peripheral devices
- Lack of transparency
- Resource-constrained devices
- Myriads of vulnerable apps/devices
- Non-technical users / risky behaviors
- Unmanaged devices & networks
- “Always on” requirement

Pair Once,
Trust Indefinitely

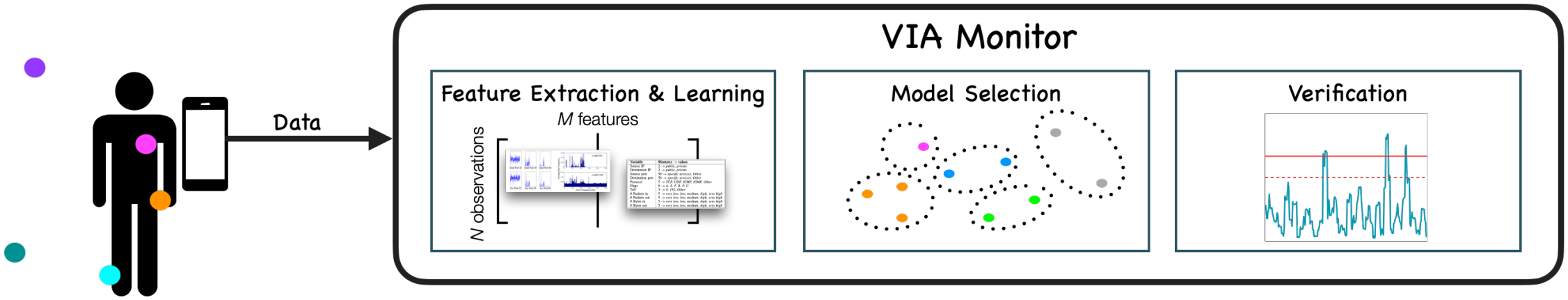
Towards a Solution...

Recurring Verification

- *What is connected right now?*
- *Are interactions consistent w/ expectations?*

Remediation

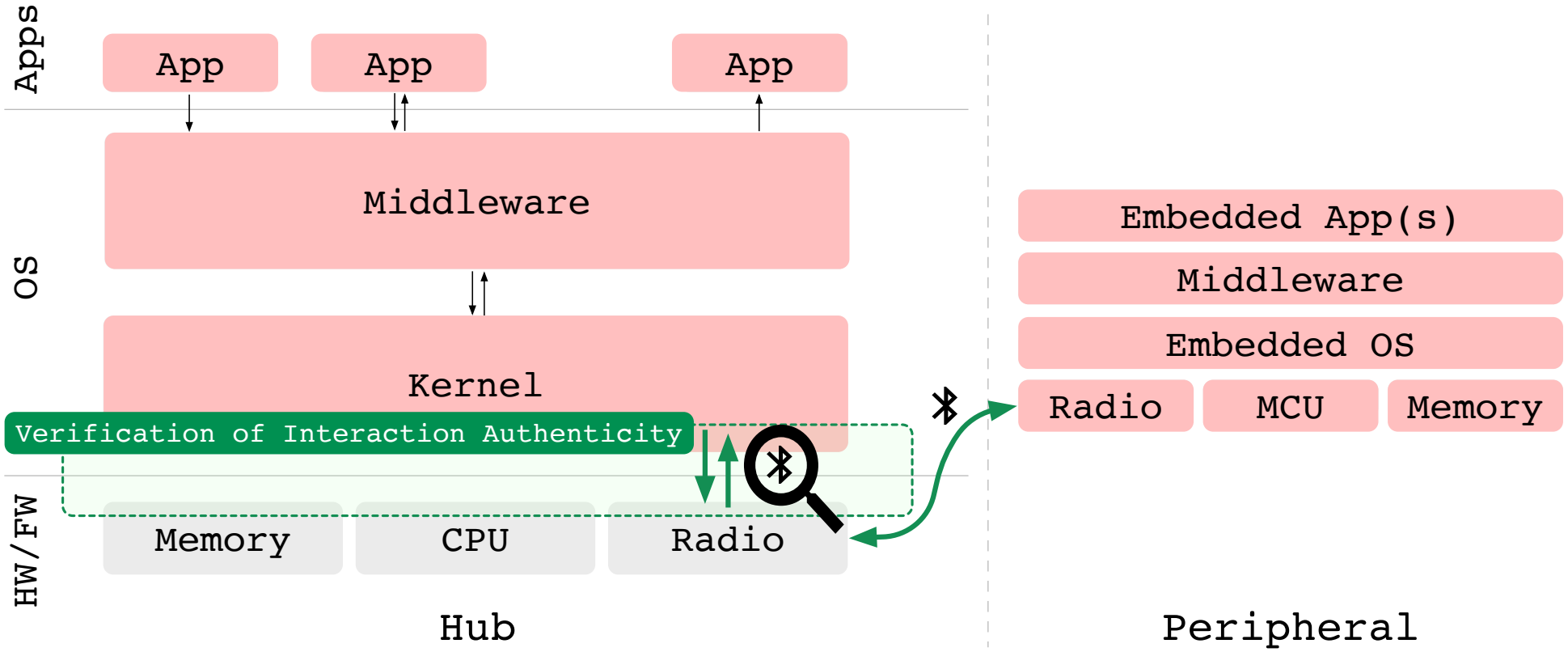
- *Is it safe to keep using device X?*
- *What can be done?*
  (e.g., notify user, terminate connection)



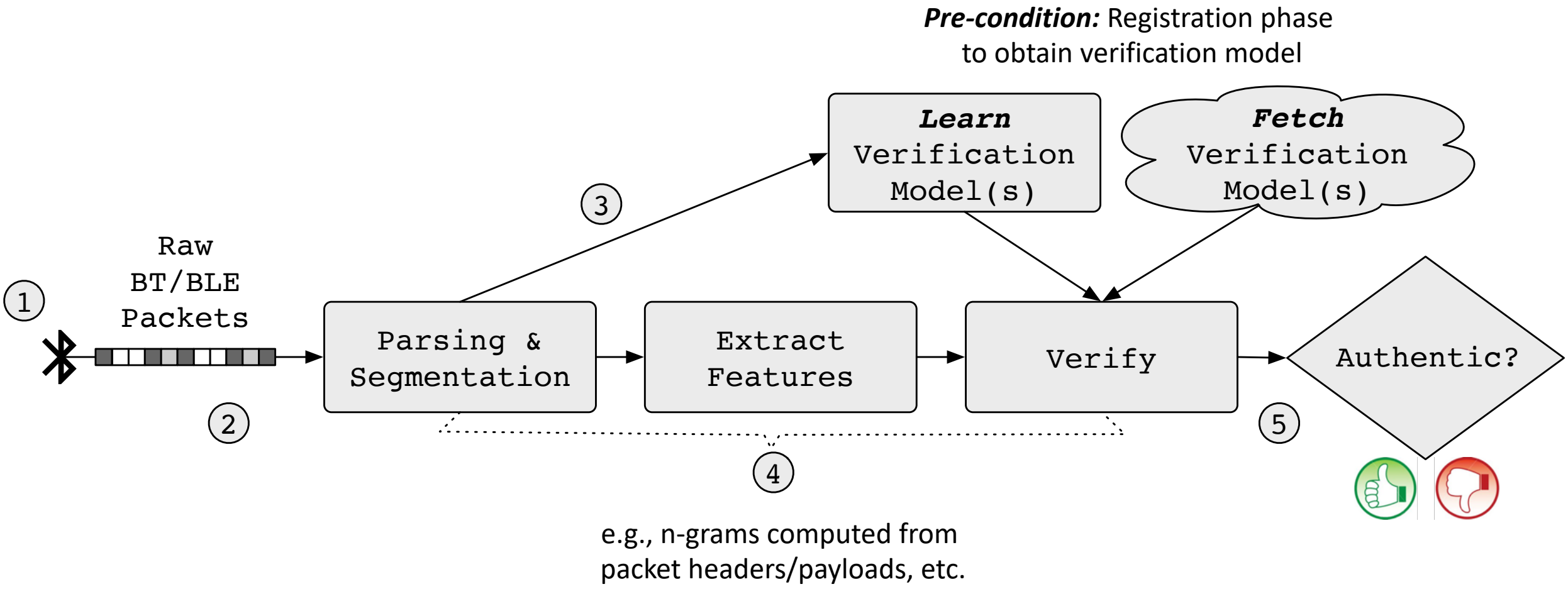
Goal: Realize a new, noninvasive solution to regularly re-evaluate trust decisions *(e.g., long-term pairings, short-term connections/sessions)*

➔ Monitor ongoing network traffic
➔ Detect deviations from accepted verification models
➔ Take appropriate action

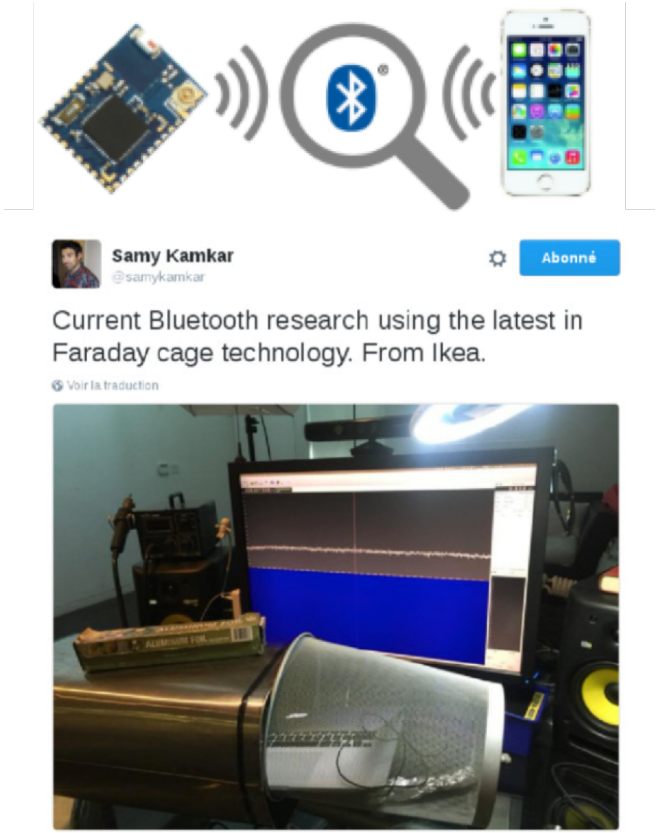
System Overview



System Overview (cont.)

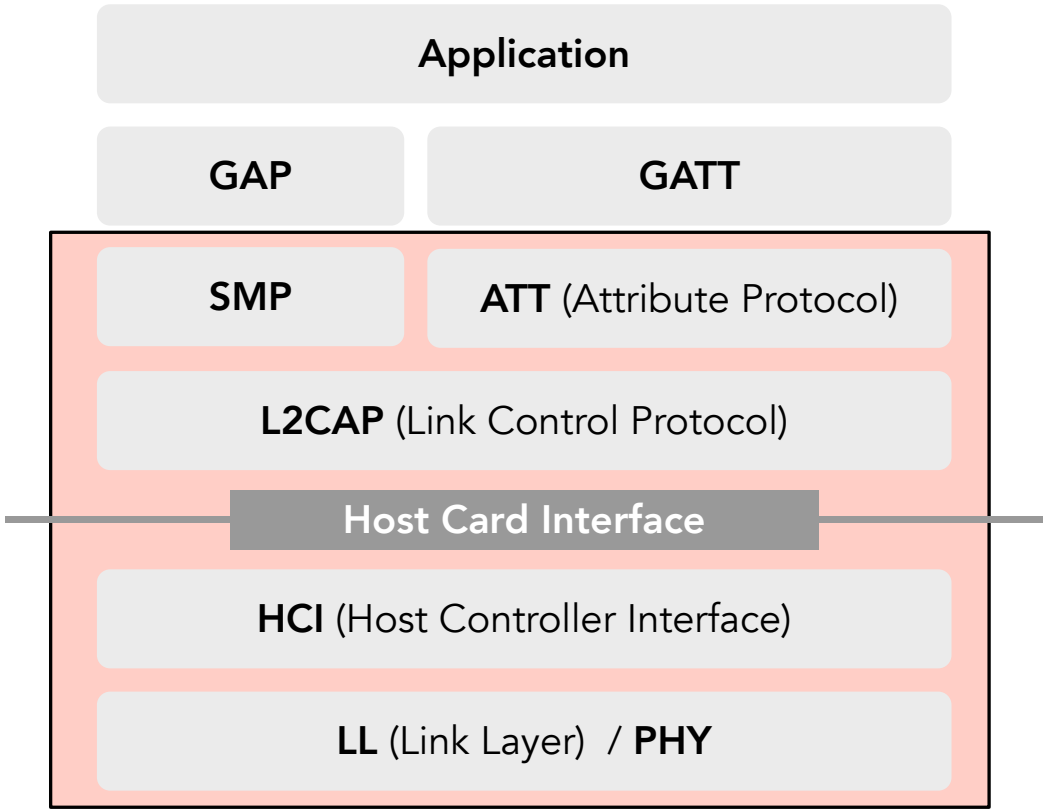


Bluetooth & Data Capture



Sniff OTA traffic (unreliable)…

vs.

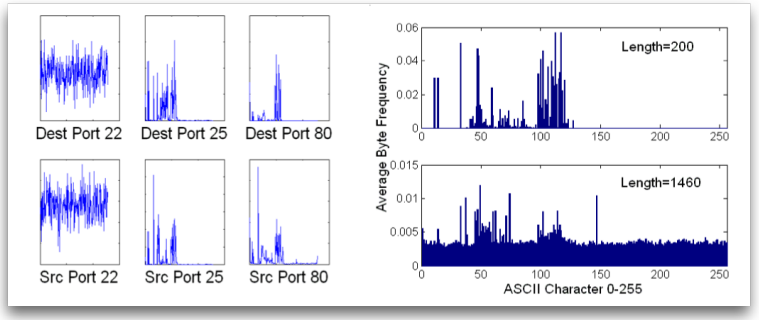


…sniff ***within*** device,
between HW and SW (reliable)

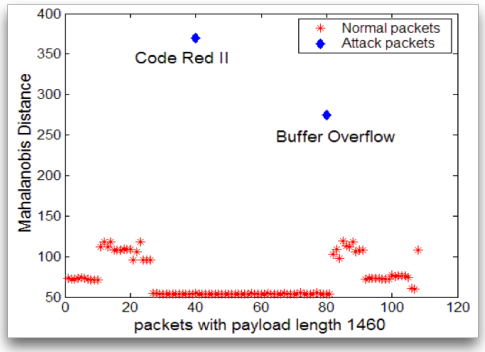
W. Albazrqaoe, et al. ***Practical Bluetooth Traffic Sniffing***. Proceedings of MobiSys, 2016.
J. Gutierrez del Arroyo, et al. ***Enabling Bluetooth Low Energy auditing through synchronized tracking of multiple connections***. *International Journal of Critical Infrastructure Protection*, vol. 18, pp. 58–70, 2017.

Traffic Modeling

- Fine-Grained Modeling & Analysis
  - Deep packet inspection; **n-grams** based on aspects of packet headers/payload
  - models based on packet length, src/dst identifiers, ports, **byte distributions**, etc.
  - n-grams ➔ choose n=1 ➔ simple, efficient, effective, resilient to mimicry attacks



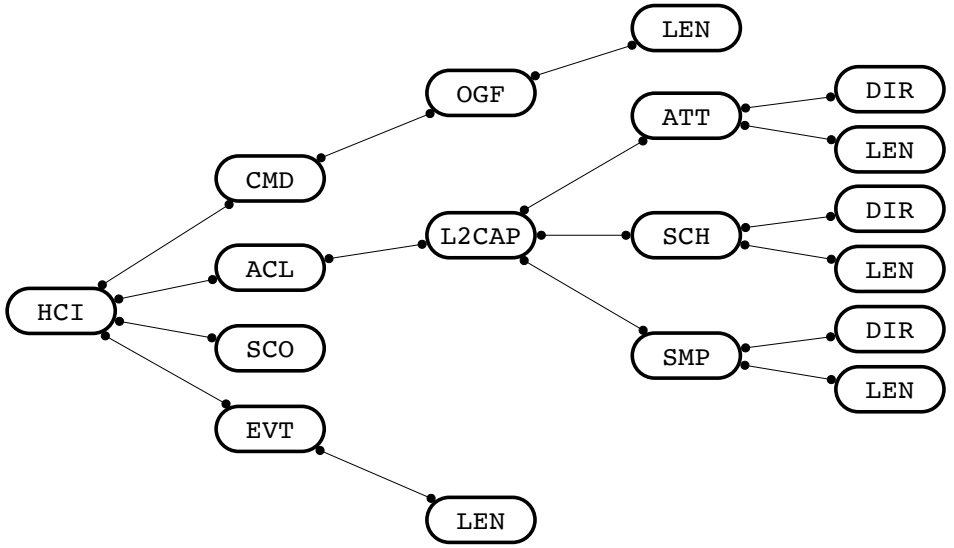
$$\phi : x \rightarrow \left(\phi_s(x)\right)_{s \in S} \quad \text{with} \quad \phi_s(x) = \text{occ}(s, x)$$



$$d^2(x, \overline{y}) = (x - \overline{y})^T C^{-1} (x - \overline{y})$$

PAYL (K. Wang and S. Stolfo, 2004)

- Hierarchical Segmentation
  - Different model for each protocol based on length & direction
  - Target specific protocols and underlying semantics
  - Isolate device-specific characteristics

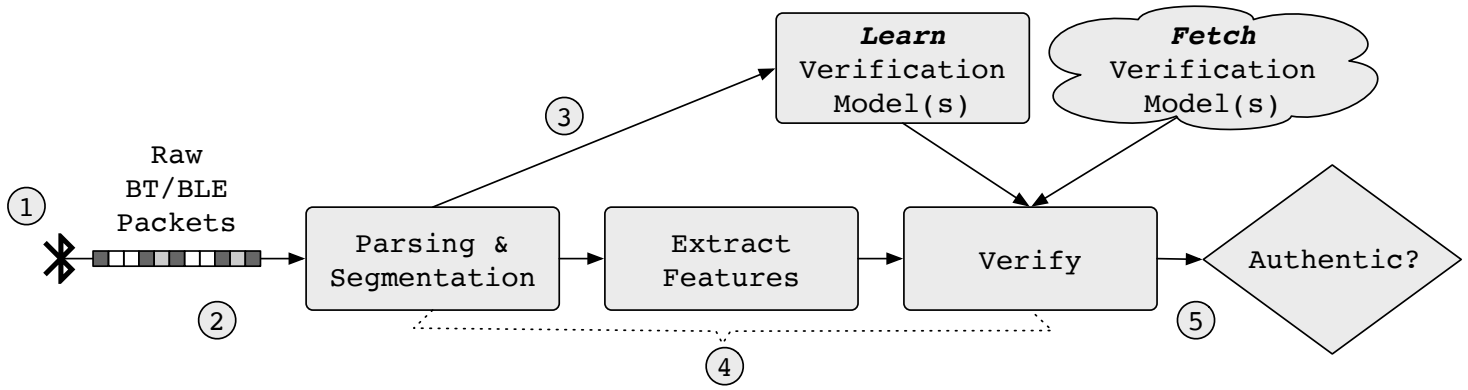


K. Wang and S. Stolfo. **Anomalous Payload-Based Network Intrusion Detection**. Recent Advances in Intrusion Detection, 2004.

C. Wressnegger et al. **A Close Look on N-grams in Intrusion Detection: Anomaly Detection vs. Classification**. Proceedings of the ACM Workshop on Artificial Intelligence and Security, 2013.

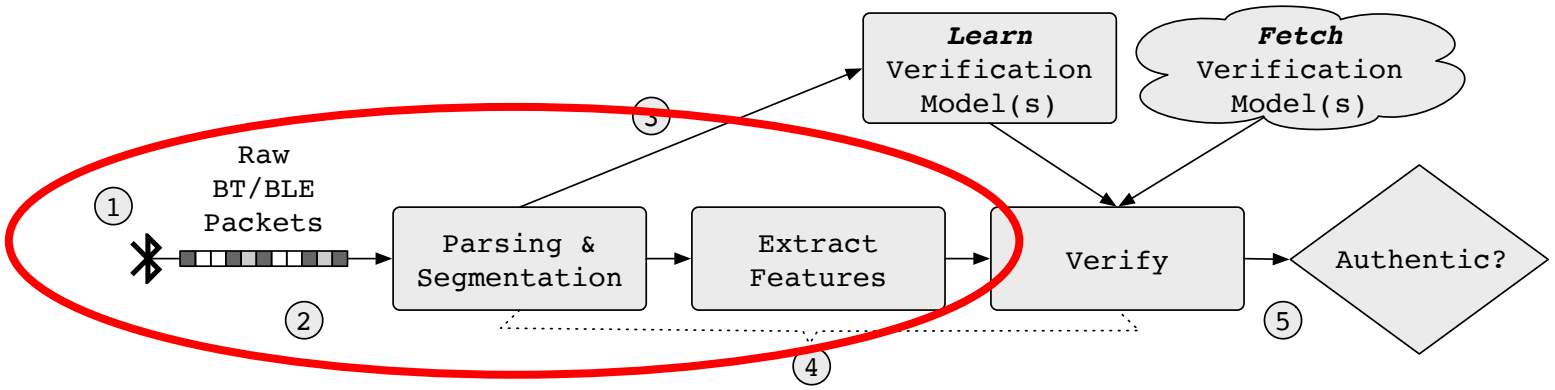
Tools & Dataset

An overview of our analysis tools & testbed of Bluetooth-enabled consumer devices



Tools & Dataset

An overview of our analysis tools & testbed of Bluetooth-enabled consumer devices



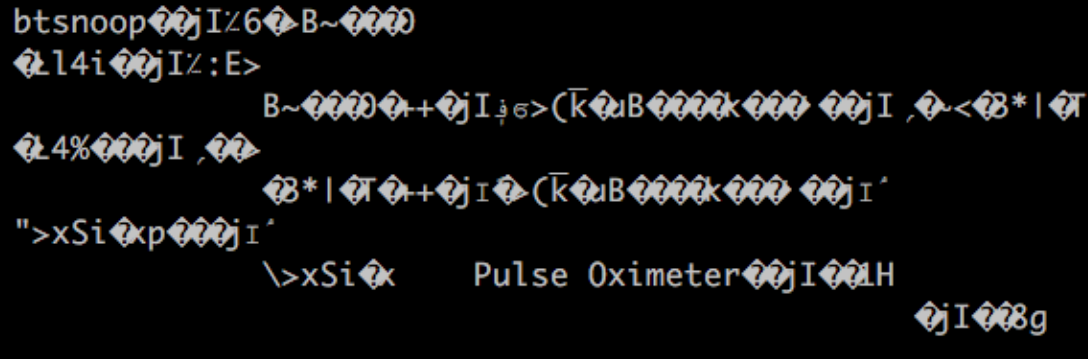
Bluetooth Dataset

- **>300 Bluetooth app-device HCI traces**
  Events, Commands, and ACL traffic
- **20* distinct devices**
  10+ traces each (3-10min each)
- **2 distinct device categories**
  smart home, smart health

Identifier	Device Model
Smart Health	
BP Monitor iHealth [wrist] (1)	iHealth View Bluetooth Wrist Blood Pressure Monitor
BP Monitor iHealth [upperarm] (1)	iHealth Feel Bluetooth Upper Arm Blood Pressure Monitor
BP Monitor Omron [wrist] (1)	OMRON 10 Series Wireless Wrist Blood Pressure Monitor
BP Monitor Omron [upperarm] (1)	OMRON Evolv Wireless Upper Arm Blood Pressure Monitor
BP Monitor Choice [upperarm] (1)	Choice Wireless Blood Pressure Monitor, Upper Arm
Glucosemonitor iHealth [na] (1)	iHealth Wireless Smart Blood Sugar Test Kit
Glucosemonitor Choice [na] (1)	Choice Wireless Blood Glucose Monitor
HR Monitor PolarH7 [chest] (1)	Polar H7 Wearable Heart Rate Monitor (Chest)
HR Monitor PolarH7 [chest] (2)	Polar H7 Wearable Heart Rate Monitor (Chest)
HR Monitor Zephyr [chest] (1)	Zephyr Wearable Heart Rate Monitor (Chest)
Pulse Oximeter iHealth [finger] (1)	iHealth Air Wireless Fingertip Pulse Oximeter
Scale Gurus [floor] (1)	Bluetooth Smart Body Fat Scale by Weight Gurus
Scale Renpho [floor] (1)	RENPHO Smart Bluetooth Body Fat Scale
TENS Unit Omron [na] (1)	OMRON Avail Dual Channel TENS unit
Thermometer Kinsa [ear] (1)	KINSA Smart Ear (in-ear smart thermometer)
Thermometer Kinsa [oral] (1)	KINSA QuickCare (oral smart thermometer)
Smart Home	
Env Sensor Inkbird [na] (1)	Inkbird combo mini Bluetooth (temp/hum) sensor
Env Sensor Inkbird [na] (2)	Inkbird combo mini Bluetooth (temp/hum) sensor
Smart Lock August [door] (1)	August Smart Lock Pro + Connect (3^{rd} Gen.)
Smart Lock Schlage [door] (1)	Schlage Sense Smart Deadbolt

Format: Device & Manufacturer [Location] (Device ID)

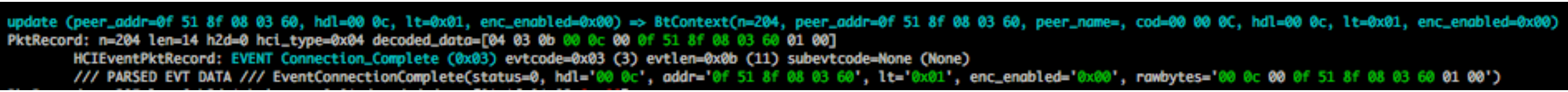
App	Corresponding Device(s)
Smart Health	
RENPHO	RENPHO scale
Weight Gurus	Weight Gurus scale
iHealth MyVitals	iHealth blood-pressure monitors, pulse oximeter
OMRON Connect	OMRON blood-pressure monitors
Choice Blood Pressure	Choice blood-pressure monitor
Polar Beat	Polar and Zephyr heart-rate monitors
OMRON TENS	OMRON TENS unit
iHealth Gluco-Smart	iHealth blood-glucose meter
AgaMatrix Diabetes Manager	Choice blood-glucose meter
Kinsa	Kinsa oral and ear thermometers
Smart Home	
Schlage Home	Schlage smart deadbolt
August Home	August smart lock
Engbird	Environment sensors



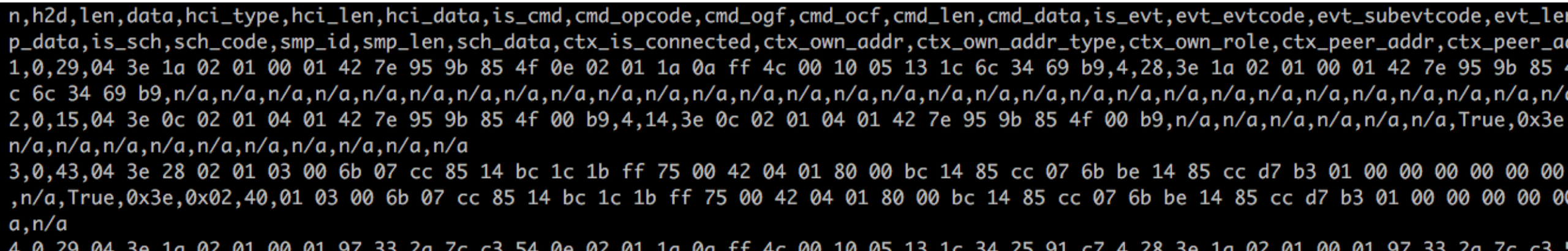
Raw, Binary Data

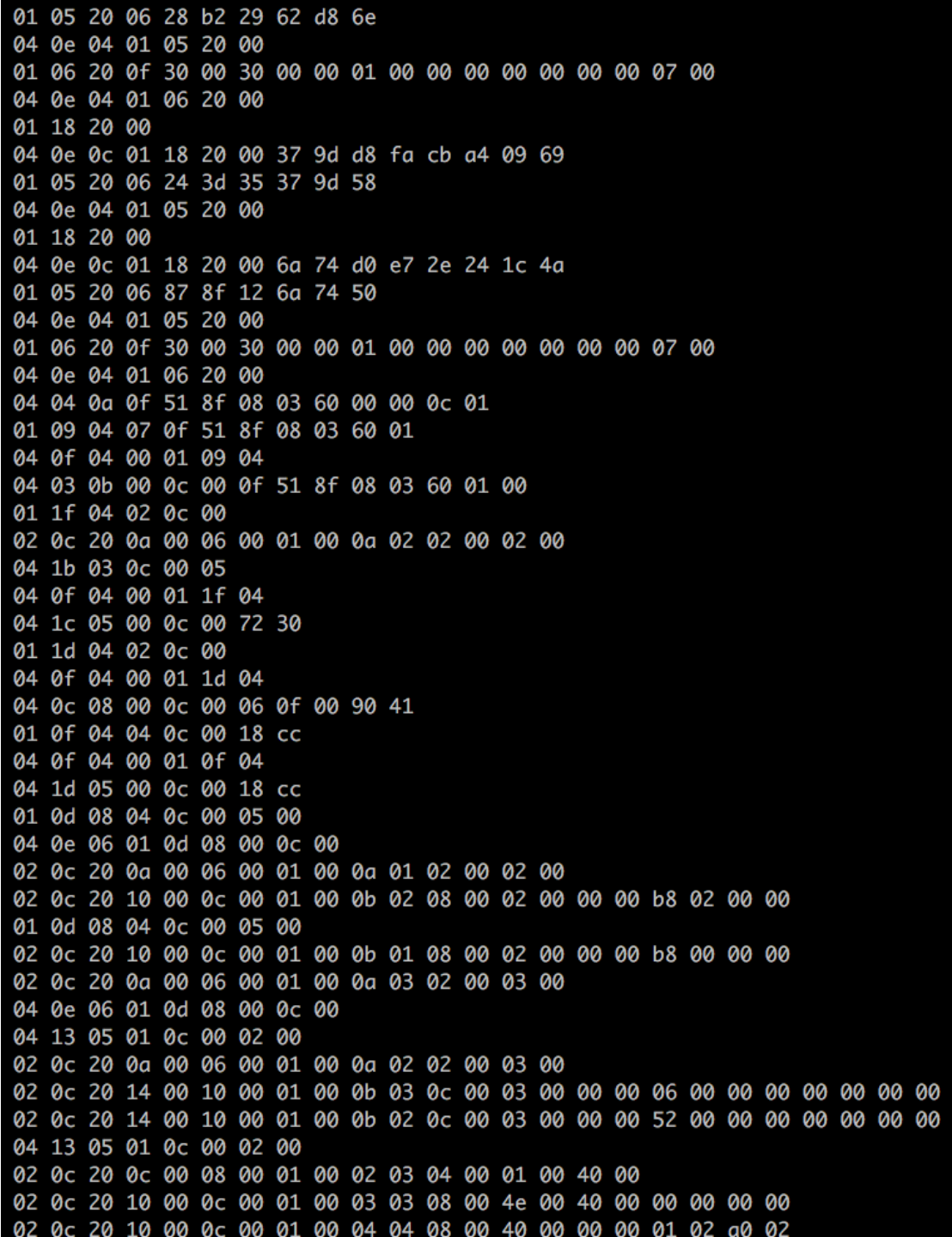
(BTSnoop v1, HCI UART)

Parse & extract features (e.g., pkt types, lengths, direction, conn. context, byte-frequency dist.)

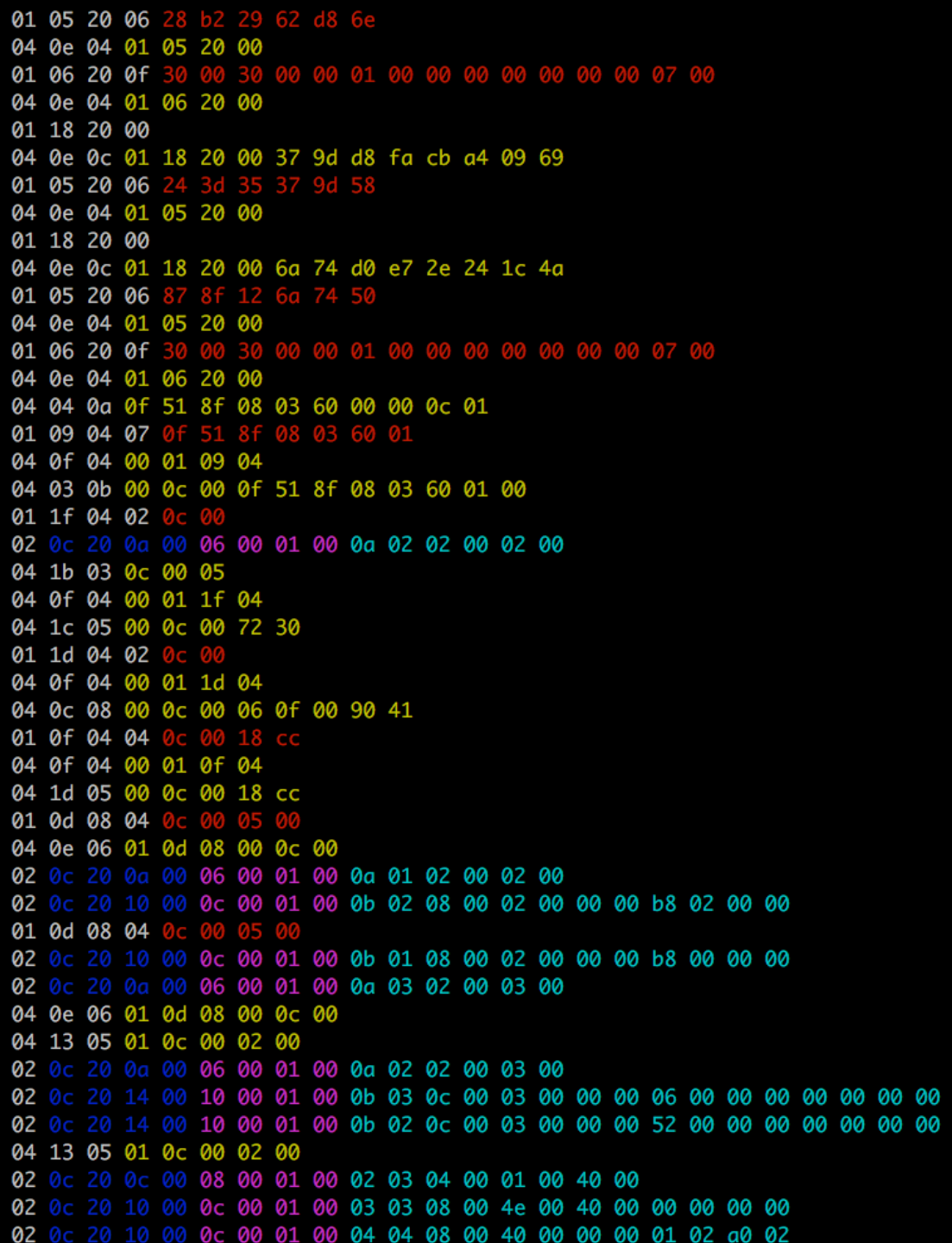


Write per-packet features to subsequent file for analysis



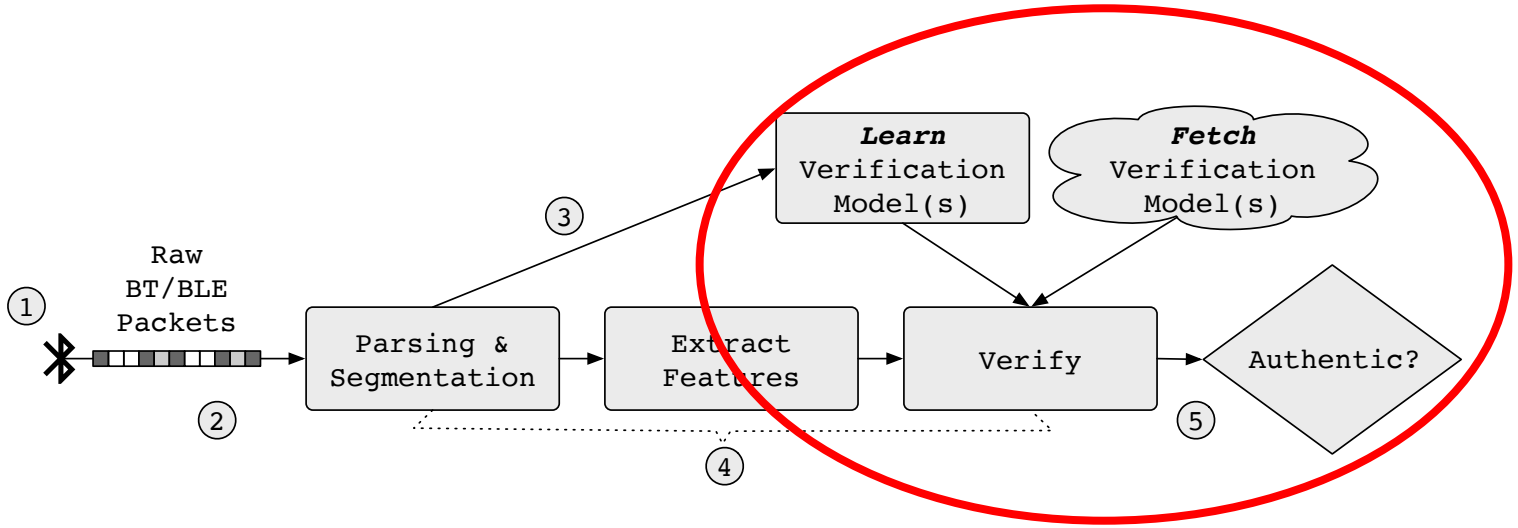


Raw packet bytes used to compute byte-frequency distributions (i.e., n-grams)



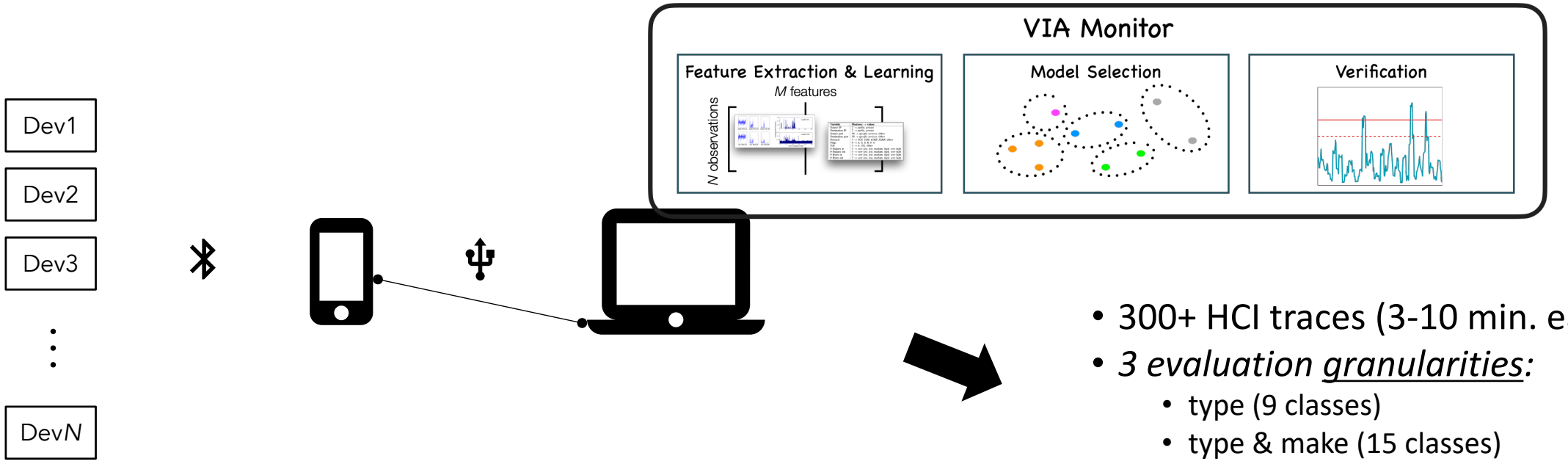
Raw, segmented packet bytes used to compute byte-frequency distributions (i.e., n-grams)

Experimental Results



Experimental Procedure

RQ: *Can we reliably verify whether a device's behavior is consistent with a previously-learned, authentic behavioral model?*



testbed
(20 smart devices)

smartphone
(trace collection)

laptop VM/Container
(parsing+learning+analysis)

- 300+ HCI traces (3-10 min. each)
- *3 evaluation granularities:*
  - type (9 classes)
  - type & make (15 classes)
  - unique device (20 classes)
- One-vs-all (re)labeling strategy
- *Off-the-shelf classifiers (sklearn):* RF, SVM, kNN, Dummy/baseline
- Stratified 10-fold cross validation

Verification by Device Type

- **Q**: Are ***functionally-different*** devices distinct from one another?
- **Observation(s)**: clear separation between the different ***types*** of devices
- **Takeaway(s)**: Yes!
  F1-Score ≥ 0.90 in all cases

Device Type	Precision	Recall	F1-Score
BP Monitor	1.0	0.96	0.98
Env Sensor	1.0	0.81	0.90
Glucosemonitor	1.0	0.96	0.98
HR Monitor	1.0	0.91	0.95
Pulse Oximeter	1.0	1.00	1.00
Scale	1.0	0.90	0.95
Smart Lock	1.0	1.00	1.00
TENS Unit	1.0	0.92	0.96
Thermometer	1.0	0.97	0.98

Verification by Device Instance

- **Q**: Are ***individual devices*** distinct from one another?

- **Observation(s)**:
  - potential confusion within same device ***type*** & ***manufacturer***
  - potential confusion between ***identical devices***

- **Takeaway(s)**: Yes!
  *F1-Score ≥ 0.86 in most cases**

Device Instance	Precision	Recall	F1-Score
BP Monitor Choice [upperarm] (1)	1.00	1.00	1.00
BP Monitor iHealth [upperarm] (1)	1.00	0.92	0.96
BP Monitor iHealth [wrist] (1)	1.00	0.94	0.97
BP Monitor Omron [upperarm] (1)	1.00	1.00	1.00
BP Monitor Omron [wrist] (1)	1.00	1.00	1.00
Env Sensor Inkbird [na] (1)	0.50	0.43	0.46
Env Sensor Inkbird [na] (2)	0.50	0.22	0.31
Glucosemonitor Choice [na] (1)	1.00	1.00	1.00
Glucosemonitor iHealth [na] (1)	1.00	0.91	0.95
HR Monitor PolarH7 [chest] (1)	1.00	0.45	0.62
HR Monitor PolarH7 [chest] (2)	1.00	0.18	0.31
HR Monitor Zephyr [chest] (1)	1.00	0.75	0.86
Pulse Oximeter iHealth [finger] (1)	1.00	1.00	1.00
Scale Gurus [floor] (1)	1.00	0.90	0.95
Scale Renpho [floor] (1)	1.00	0.80	0.89
Smart Lock August [door] (1)	0.91	1.00	0.95
Smart Lock Schlage [door] (1)	1.00	0.86	0.92
TENS Unit Omron [na] (1)	1.00	0.92	0.96
Thermometer Kinsa [ear] (1)	1.00	1.00	1.00
Thermometer Kinsa [oral] (1)	1.00	0.94	0.97

**The paper presents a detailed evaluation of various special cases*

Conclusion

Summary

- A new recurring verification scheme for Bluetooth
- A new ***dataset*** (>300 HCI traces) and ***tools*** to aid future work in Bluetooth research
- Evaluation off-the-shelf ML models to perform verification procedure
  - f1-score >=0.86 in most cases
  - adept at verification of ***functionally-similar devices***
  - adept at verification of ***unique device instances****

Future Opportunities

- Expand dataset to include more devices, adversarial examples, etc.
- Real-world deployment (e.g., in mobile hub) to evaluate practical issues *(power consumption, verification latency, etc.)*

Please see our paper for more details...

- *Discussion and experiments conducted for special cases
- Features for verification models
- Case study: Bluetooth-enabled blood-pressure monitors

Thank You! Questions? Comments?

Recurring Verification of Interaction Authenticity Within Bluetooth Networks

@ The 14th ACM Conference on
Security and Privacy in Wireless and Mobile Networks (WiSec'21)
June 28th – July 2nd, 2021

Travis Peters[1], Timothy J. Pierson[2], Sougata Sen[3], José Camacho[4], David Kotz[2]

[1] Montana State University, USA (now @ Include Security)

[2] Dartmouth College, USA

[3] BITS Pilani, Goa Campus, India

[4] University of Granada, Spain