UNIVERSITÄT BONN

IT Security

Device identification in wireless networks

Daniel Vogel
Friedrich-Hirzebruch-Allee 8
53115 Bonn
vogel@cs.uni-bonn.de

December 05, 2024

UNIVERSITÄT BONN

Goal of this lecture

- What is Device Identification (DI) and what role does DI fill?
- Why is DI relevant to us and you?
- What techniques can be used to do DI?
- Improve awareness on what wireless devices can and will tell the environment about itselves and the users

UNIVERSITÄT BONN

Device Identification

An introduction

Device Identification

- Definition Device Identification:
  - Identification is the ability to uniquely identify a user or device based on a unique ID (such as MAC address, IMEI or MEID for phones)

Device Identification

- Definition Device Identification:
  - Identification is the ability to uniquely identify a user or device based on a unique ID (such as MAC address, IMEI or MEID for phones)
- Coverage: Mostly wireless environments. Some presented concepts can be applied to wired connections as well.

Device Identification Goals

- Goals: to identify / differentiate

Device types	TV, phone, notebook, lamp, ...
Device manufacturers	Samsung, Google, Cisco, AVM, …
Device models	Galaxy S22, Galaxy A13, ...
Device properties	Support for different protocols or features, screen size, ...
Unique devices in a network	My phone vs your phone

Using Device Identification

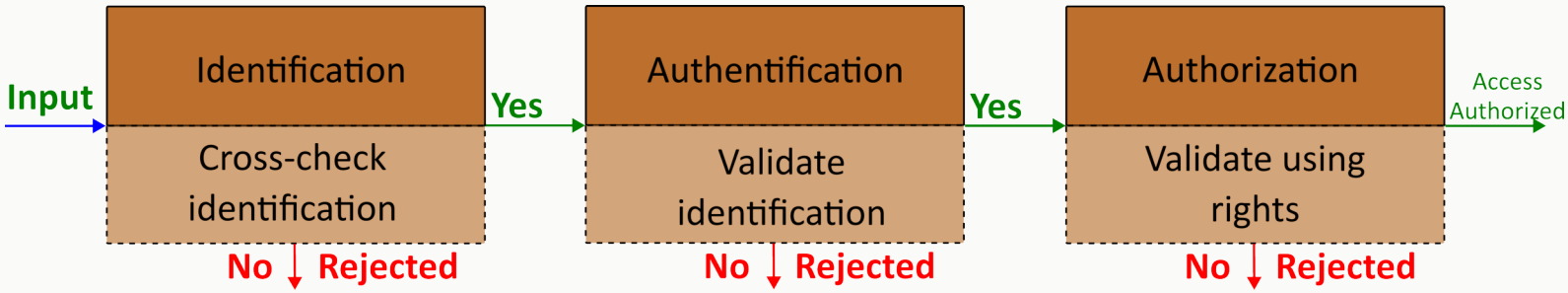
- RF Device Identification and Authentication
- Localization, Navigation, Tracking
- Intrusion detection
- New application domains

UNIVERSITÄT BONN

On the example RF Device Identification

- RF Fingerprinting
  - Identify devices based on the RF characteristics of their transmissions
  - Most often hardware-specific, where the wireless chip adds some form of individual effect
  - Imperfections in manufacturing lead to slight deviations in frequency, transmission power, clock skew, etc.
  - Very hard to forge a specific RF fingerprint

Using Device Identification for Authentication



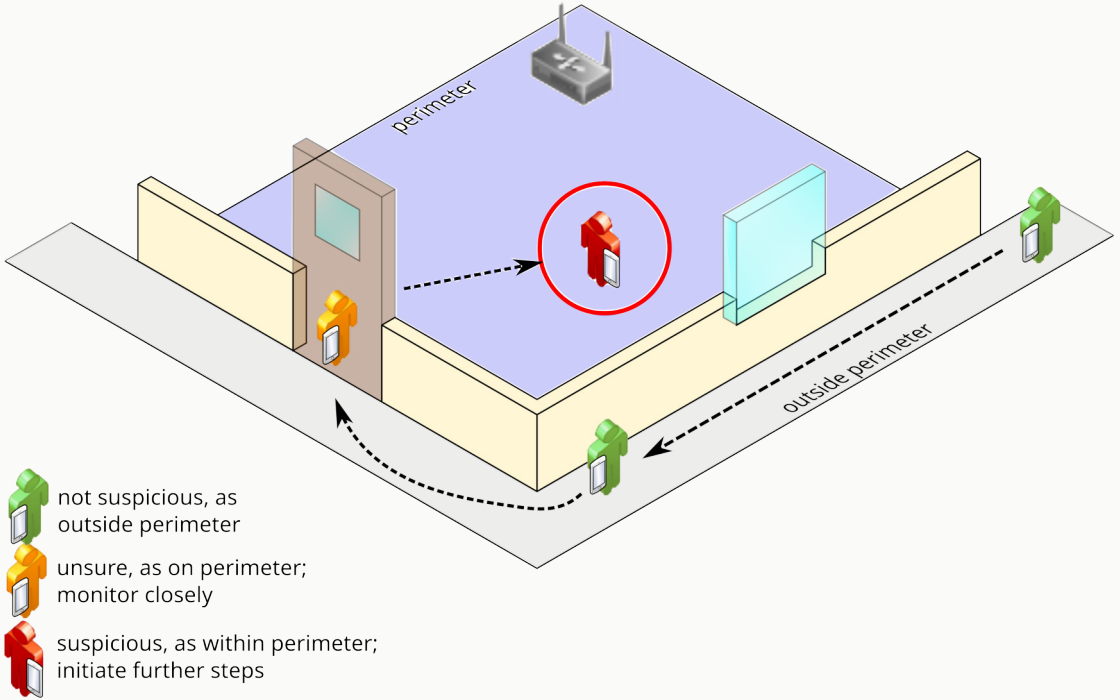
- Identification using unique IDs
- Authentication: proving that a user/device is genuinely who that user/device claims to be
- Authorization: checking legitimate permission

See: Jagannath: A Comprehensive Survey on RF Fingerprinting: Traditional Approaches, Deep Learning, and Open Challenges (2015)

Using Device Identification for Localization, Navigation, Tracking

- Usable for search & rescue operations
  - Victims or rescue operators can be identified and tracked based on their RF fingerprint
- Law enforcement
  - Identify and differentiate unique devices in a certain scenario
  - Then use positioning/localization techniques to detect intruders or track fleeing criminals
  - Identify contraband wireless devices in jails based on RF fingerprints
- (Indoor) Navigation
- Techniques potentially independent of specific packets sent by devices

Using Device Identification for Localization, Tracking, Navigation



UNIVERSITÄT BONN

Using Device Identification for Intrusion detection

- In wireless environments: Devices can be used while mobile
  - Physical intrusion or in networks with limited spacial coverage
- Identify replay/relay attacks (authenticity)
- RF fingerprint database
  - Identify outsider

Using Device Identification for new Application Domains

- Intelligent TeleHealth
  - Such as Smart Health Care
- Autonomous UAV or V2X
- Smart Grid 2.0
- and many more to come

UNIVERSITÄT BONN

Different perspectives

- The **network authority perspective (defensive)**
  - Law enforcement agencies can maintain some measure of control and regulatory power (e.g. illegal transmitters)
  - Mobile operators can identify cloned cell phones
  - Administrator can identify and track problematic hosts
  - Prevention/protection against some types of attacks on the network operation (e.g. identity spoofing)

UNIVERSITÄT BONN

Different perspectives

- The **attacker perspective (offensive)**
  - Identify valuable targets (e.g. hosts) to break into a network
  - Privacy violation (e.g. unauthorized tracking or self defense)
  - Protocol compromise (e.g. “Shake them up”, where it matters that a third party cannot distinguish which device sent a certain packet)

UNIVERSITÄT BONN

Device Identification Basics

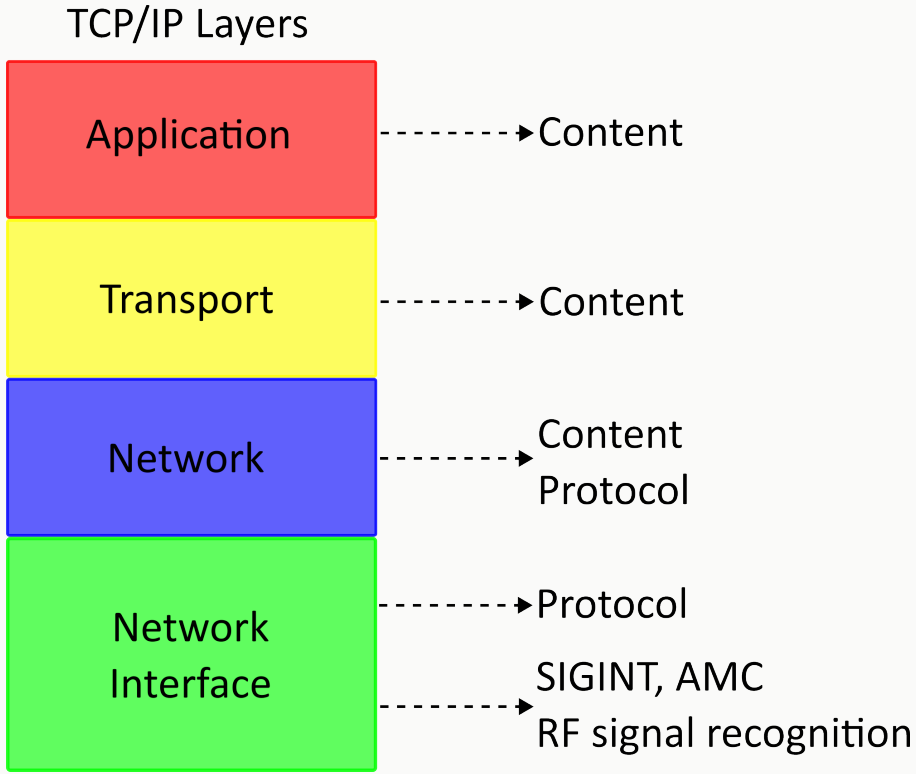
scenarios, layers, characteristics

UNIVERSITÄT BONN

Scenario

- In a typical scenario
  - The **fingerprinter** observes traffic *to and from a targeted device* (**fingerprintee**) in order to find characteristics that (uniquely) distinguish the device or its components
- Fingerprinting looks for characteristics in all layers
- Observables depend on communication (technology)
  - i.e. Wi-Fi provides other observable characteristics than cellular networks do

FP characteristics per layer



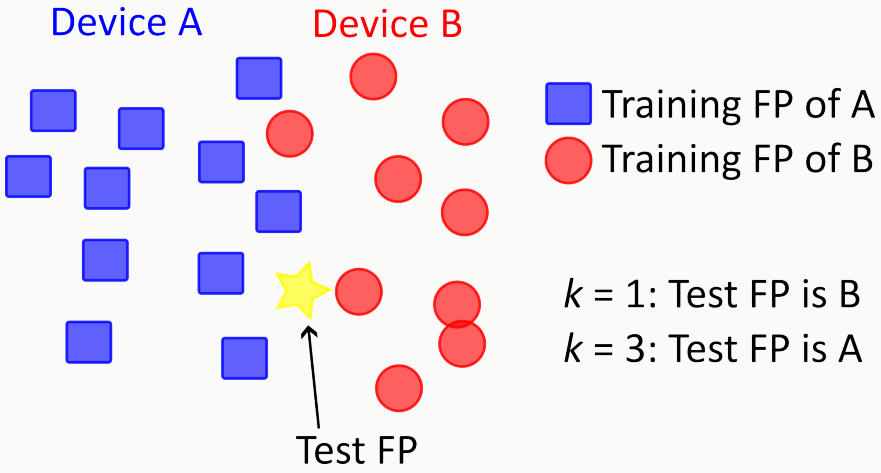
UNIVERSITÄT BONN

Characteristics

- Collect characteristics to be used as **Device Identificating Data (DID)** or **fingerprint**
  - Device identifiers such as MAC, IMEI, etc.
  - Differences in software implementations of specs (i.e. 802.11, features, behavior, etc.)
  - Hardware imperfections (see RF fingerprinting)
- **Direct DID**:
  - DID that qualify for device identification by themselves
- **Indirect DID**:
  - DID that by themselves are not enough but can help gathering additional info on the targeted device

Classification of Characteristics

- Classification of the characteristics (fingerprints)
  - Typically by means of some standard *classifier*
  - When the number of devices is known in advance: Nearest Neighbor Classifier
- Classification Error Rate (ER)
  - % of misclassified test FP over total of test FP



More on Characteristics

- Some characteristics are *unique* to certain scenarios
- Some characteristics are *always* observable in a certain scenario
- Some characteristics *may be absent or can be triggered*

More on Characteristics

- Some characteristics are *unique* to certain scenarios
  - i.e. protocol dependent. Wi-Fi packets between connected devices
- Some characteristics are *always* observable in a certain scenario

- Some characteristics *may be absent or can be triggered*

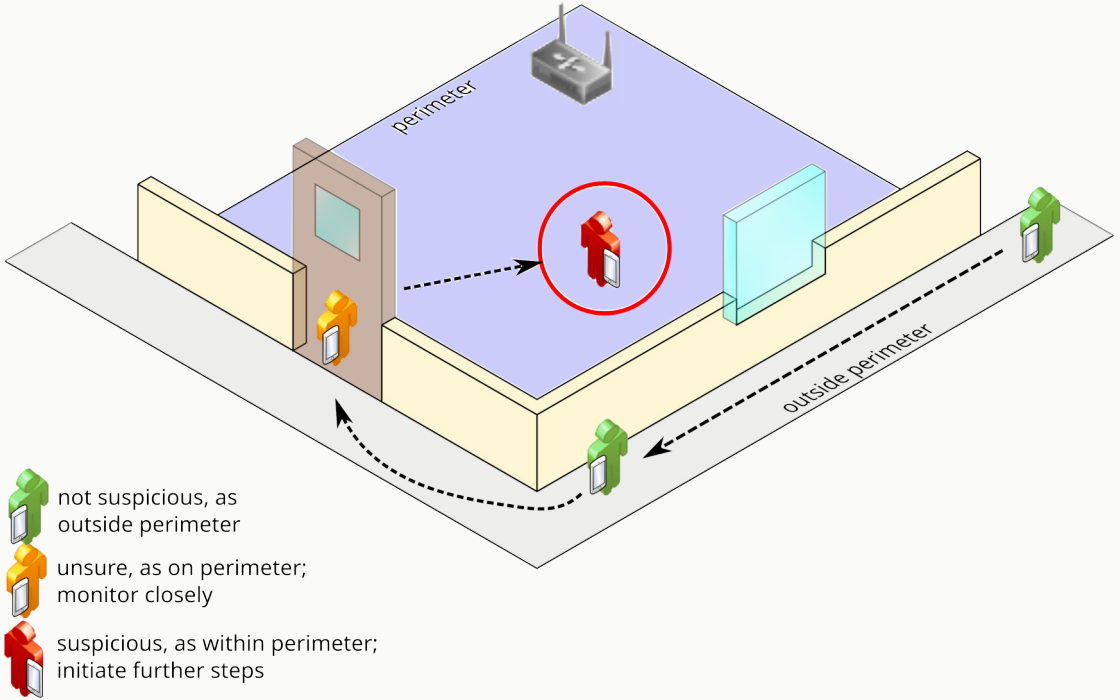
More on Characteristics

- Some characteristics are *unique* to certain scenarios
  - i.e. protocol dependent. Wi-Fi packets between connected devices
- Some characteristics are *always* observable in a certain scenario
  - i.e. frequency, modulation details
- Some characteristics *may be absent or can be triggered*

More on Characteristics

- Some characteristics are *unique* to certain scenarios
  - i.e. protocol dependent. Wi-Fi packets between connected devices
- Some characteristics are *always* observable in a certain scenario
  - i.e. frequency, modulation details
- Some characteristics *may be absent or can be triggered*
  - i.e. optional header fields, Wi-Fi features, behavioral

Example Use Cases



UNIVERSITÄT BONN

Attacker Models

- Passive attackers of different power levels
  - With access to their own hardware for their specific use case (relatively weak)
  - With access to a multitude of sensors distributed on a larger area, i.e. an airport, a malicious cellular service provider, etc.
  - With access to sensor data from around the globe (google sensorvault, ...)
- Active attackers of different power levels
  - Depends on how sophisticated the attack / how many attackers / how many devices

Real Life Examples

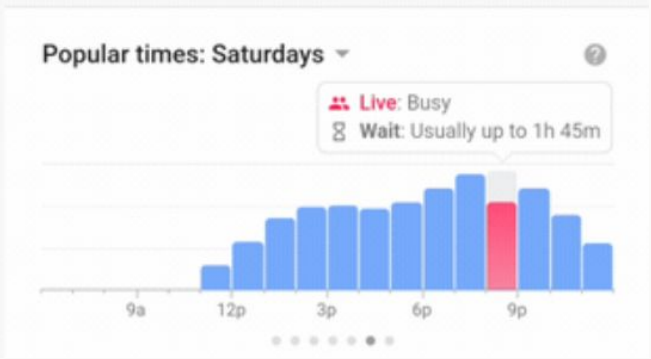
- Wardriving (warbiking, warcycling, warwalking, warflying, etc.)
  - Identifying Routers and their location, uploading to services like WiGLE.net

https://farm3.static.flickr.com/2010/2657835607_f95accdf1f.jpg

http://www.focushacks.com/photo/warcycle.jpg

Real Life Examples

- Counting visitors or load factor



https://support.google.com/business/answer/6263531?hl=de

ANALYSIS

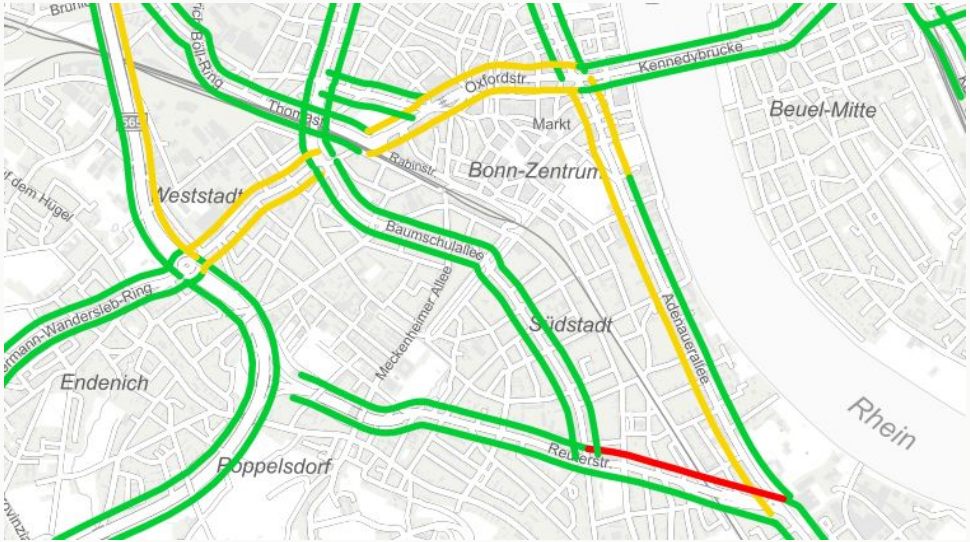
Attention shoppers: Retailers can now track you across the mall

Your favorite big box retailer or discount warehouse will soon be able to track your movements via your smartphone. Meet the next big thing in analytics: You.

https://www.computerworld.com/article/2832804/attention-shoppers--retailers-can-now-track-you-across-the-mall.html

Real Life Examples

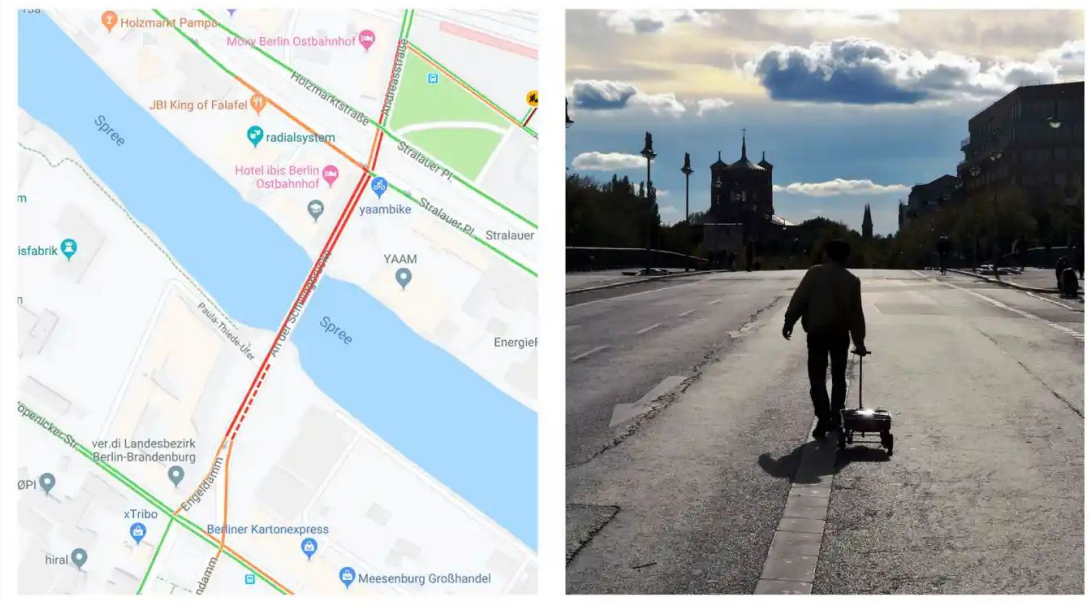
- Traffic load: The city of Bonn uses Bluetooth to measure the time a devices needs to travel between two points



https://www.bonn.de/themen-entdecken/verkehr-mobilitaet/aktuelle-verkehrslage.php

Real Life Examples

- Virtual traffic jam caused by 99 mobile phones being dragged over a bridge in Berlin



https://www.heise.de/newsticker/meldung/Virtueller-Stau-auf-Google-Maps-als-Kunstwerk-4651651.html

Real Life Examples

- UAV Detection on Airports
  - Detection and tracking of drones that might impede aerial traffic and bring airports to a halt
  - UAVs permanently send their position and aeronautical data
  - i.e. aerialarmor.com, droneshield.com

Real Life Examples

- Device tracking as a general concern when thinking about powerful actors, i.e. NSA, google sensorvault

Tech > Tech Industry

NSA tracks hundreds of millions of cell phones worldwide

You can add location tracking to the surveillance activities carried out by the secretive US agency, The Washington Post reports. And though the NSA says Americans aren't targets, data on some does get sucked up.

Edward Moyer
Dec. 4, 2013 2:51 p.m. PT

3 min read

https://www.cnet.com/tech/tech-industry/nsa-tracks-hundreds-of-millions-of-cell-phones-worldwide/

UNIVERSITÄT BONN

Device Identification Techniques

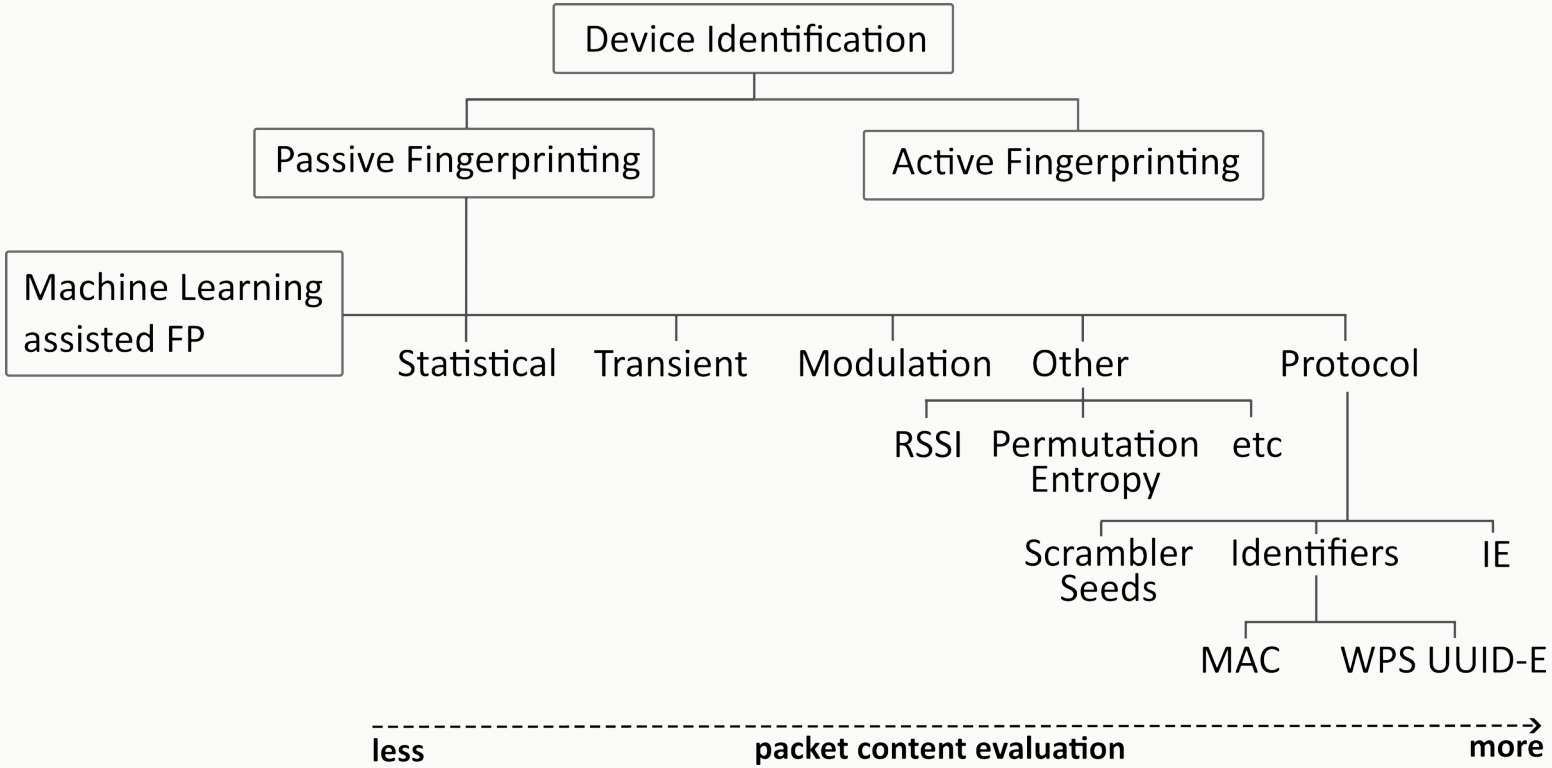
passive and active identification

UNIVERSITÄT BONN

Passive vs active identification

- Passive identification methods:
  - Only observing the communication traffic of the TD
  - Dependent on communication state the TD is in *(devices connected to a/your network may be more talkative and provide additional information)*
- Active identification methods:
  - Generating purpose-built traffic with the TD and then observe their behavior
  - Also dependent on communication state the TD is in
  - Sometimes TD can be tricked into connecting to the attackers network, thus leaking additional info

Passive identification techniques by packet evaluation



Signal intelligence vs packet evaluation

- RF Signal Intelligence (SIGINT) is its own research field
  - Extracting signal characteristics such as modulation, bandwidth, frequency, protocols etc. from unknown RF signals in the spectrum of interest
  - SIGINT methods can be utilized in device identification
  - We're basically looking at the signal, not the content
- Packet evaluation looks at the content transmitted by the signal
  - Symbols, bit entropy, encoding, header info, actual payload, etc.

DI methods: A selection

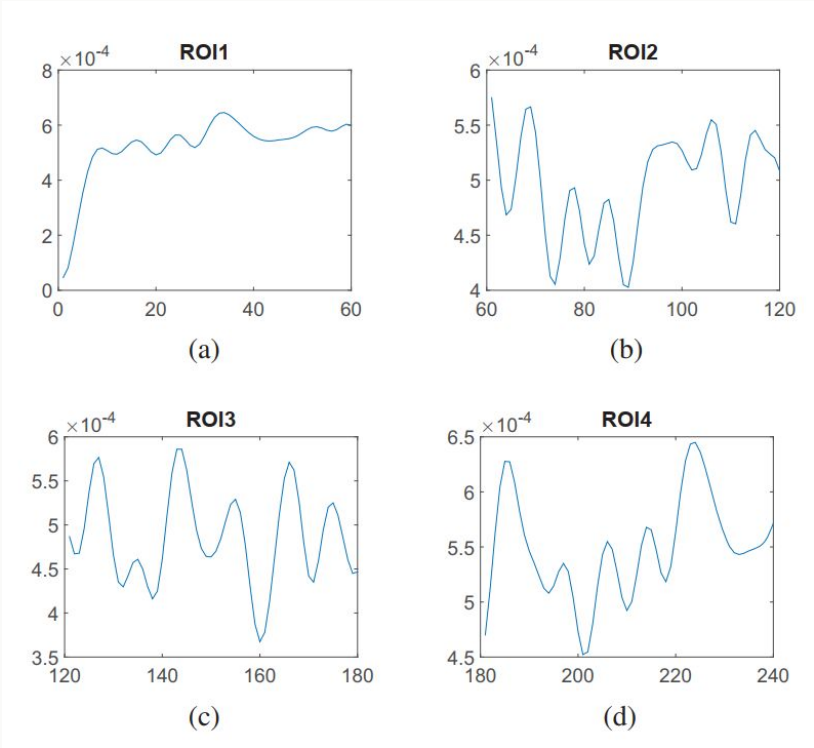
- A selection of device identification methods will be presented next
- Only summarizations. Sources cited on page. Citations can be found on the very last slides.
  - There are a lot of different approaches and papers. This selection aims to show a broad spectrum of methods that illustrate different approaches to identify devices

Passive DI: Statistical approaches

- Using statistical classification and regression models on RF features
- Non-parametric features of complex IQ Signals of ZigBee devices [1]
  - Record thousands of ZigBee preambles of different ZigBee devices
  - Divide those into 32 equal size Regions of Interest (ROI)

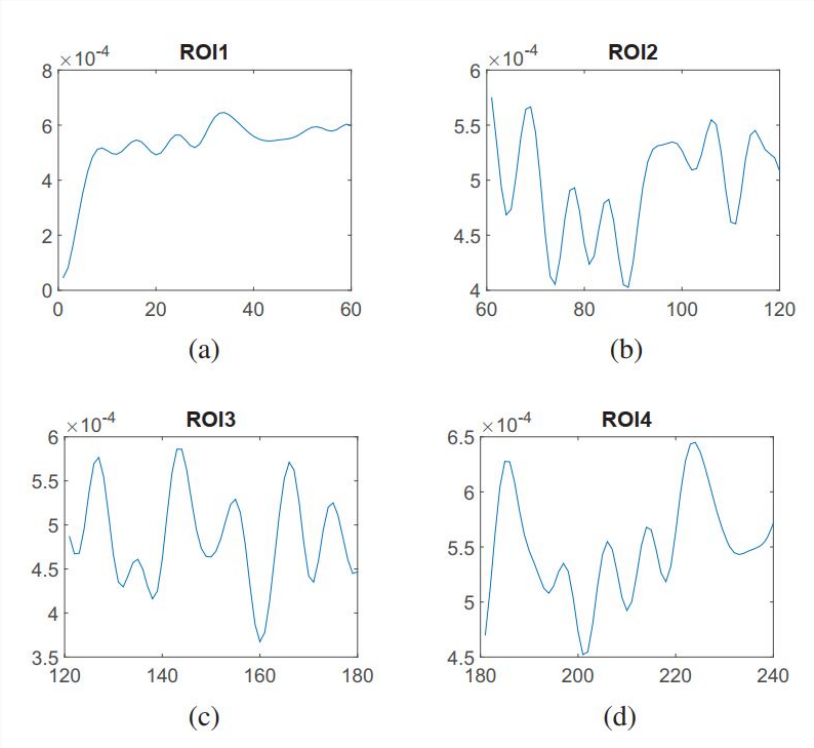
[1] H.Patel: Non-Parametric Feature Generation for RF-Fingerprinting on ZigBee Devices (2015)

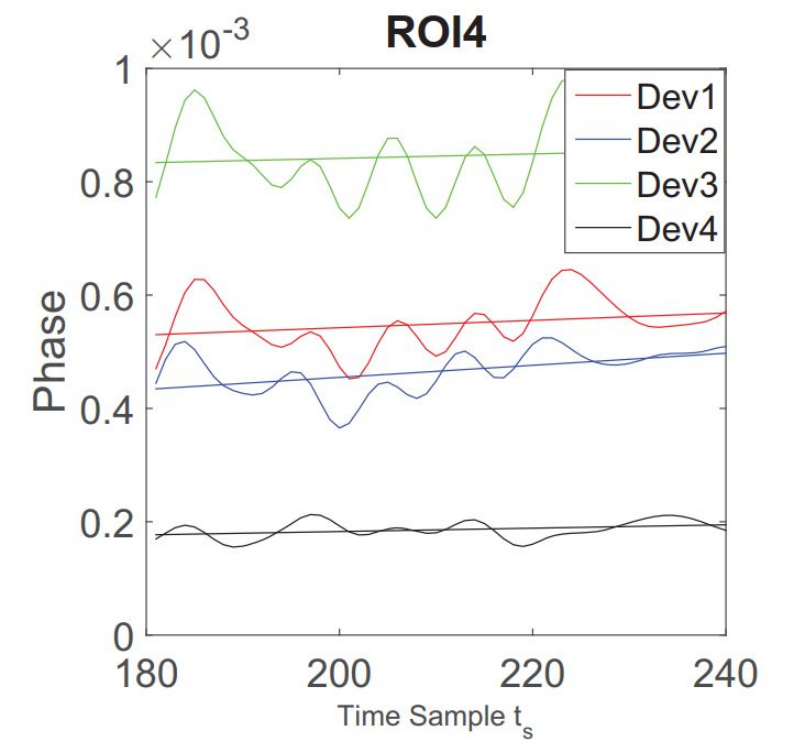
Passive DI: Statistical approaches



[1] H.Patel: Non-Parametric Feature Generation for RF-Fingerprinting on ZigBee Devices (2015)

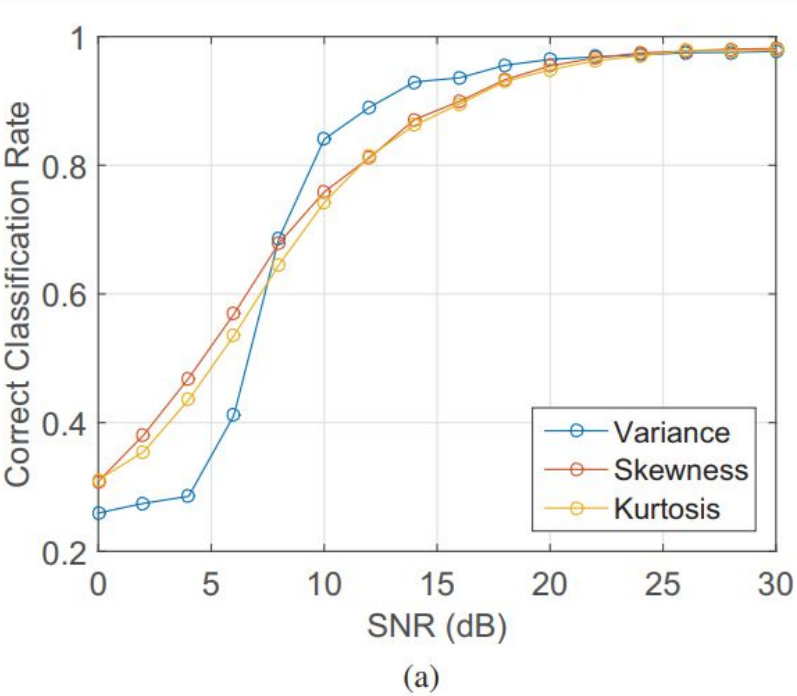
Passive DI: Statistical approaches

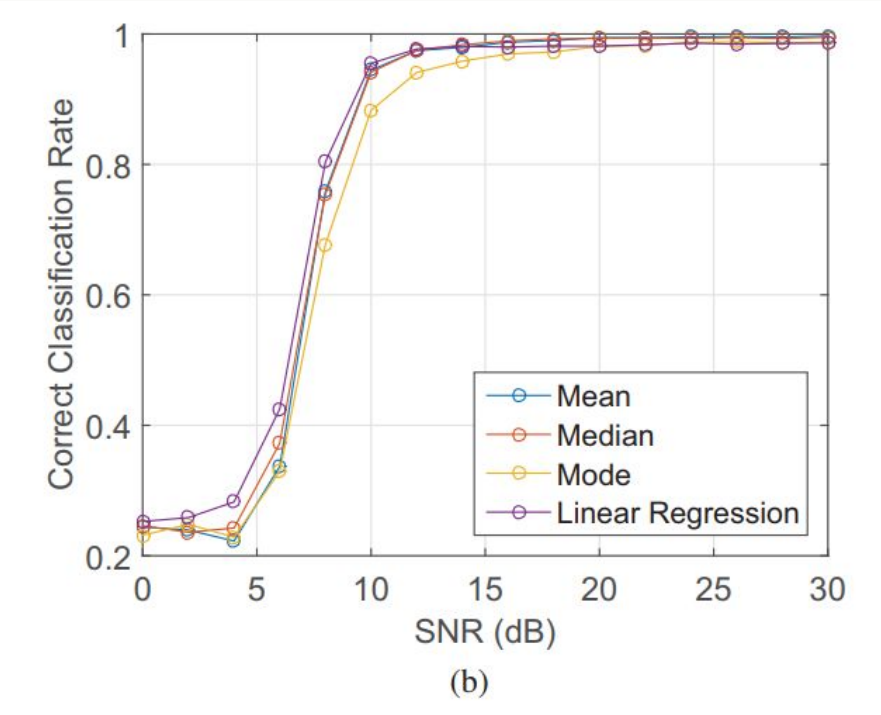




[1] H.Patel: Non-Parametric Feature Generation for RF-Fingerprinting on ZigBee Devices (2015)

Passive DI: Statistical approaches

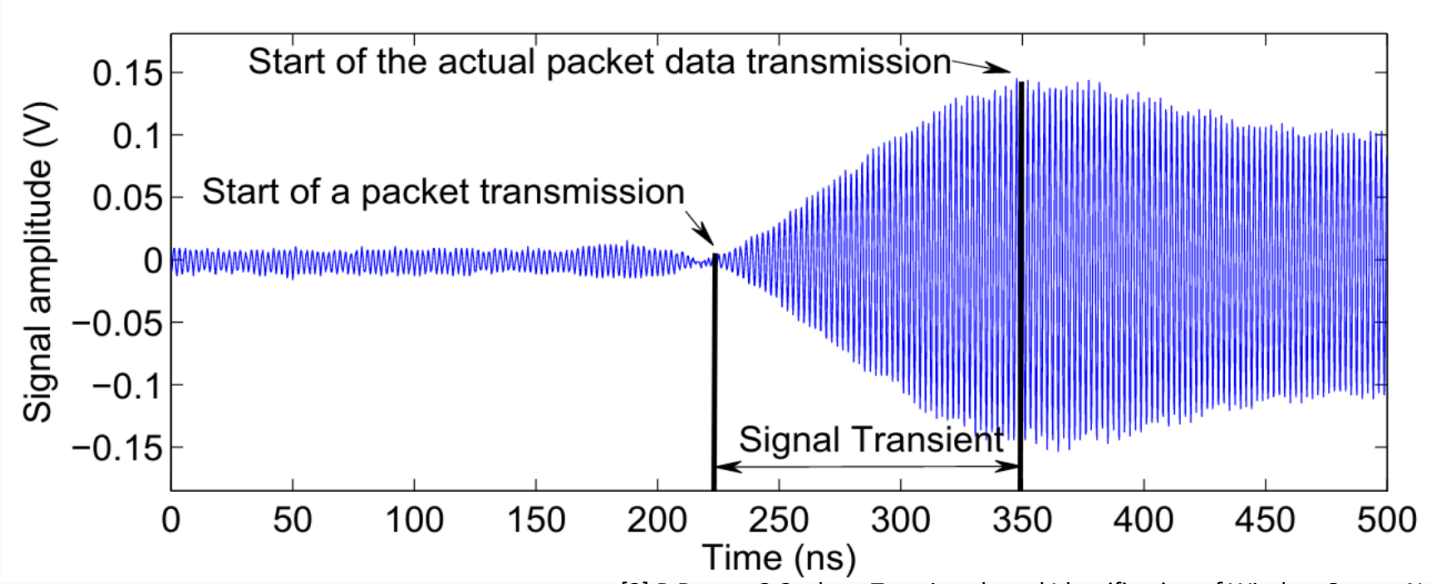




[1] H.Patel: Non-Parametric Feature Generation for RF-Fingerprinting on ZigBee Devices (2015)

Passive DI: Transient-based approaches

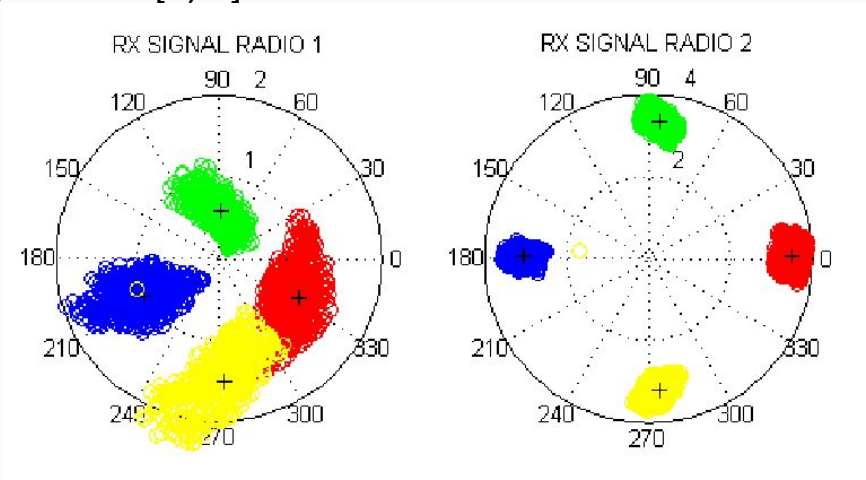
- Signal ramps up from channel noise to full power before a new transmission
- The time between the start of ramping up to full power is called the *transient signal*



[2] B.Danev, S.Capkun: Transient-based Identification of Wireless Sensor Nodes (2009)

Passive DI: Modulation-based approaches

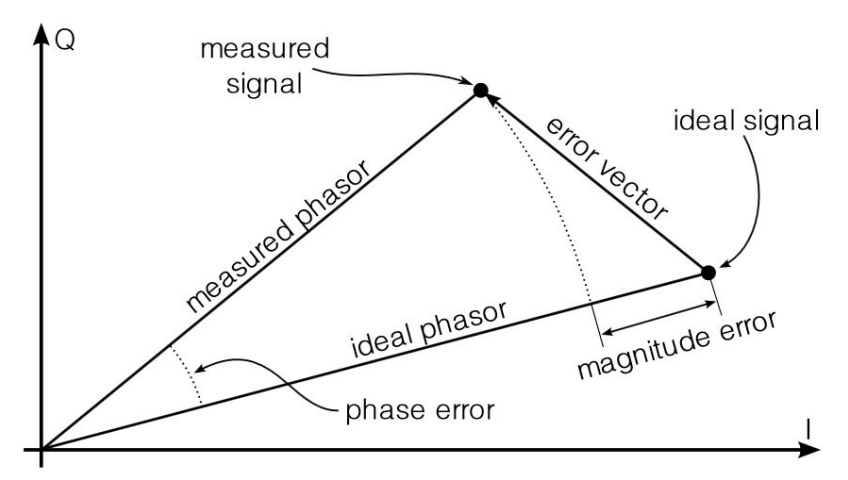
- Using modulation-based errors to identify devices
- Devices using “identical” chips should have very similar modulation characteristics
- Classifiers can be trained for specific (families of) devices [3, 4]



[3] Candore: Robust stable radiometric fingerprinting for wireless devices (2009)

Passive DI: Modulation-based approaches

- Distinguishing between identical devices
- Record and extract errors over a set of QPSK symbols
  - Frequency Offset
  - SYNC correlation
  - I/Q origin offset
  - Magniture/Phase offset



- Compute Classification error rate using error vectors representing above errors
- Based on classifier less than 5% error rate

[4] Brik: Wireless Device Identification with Radiometric Signatures (2009)

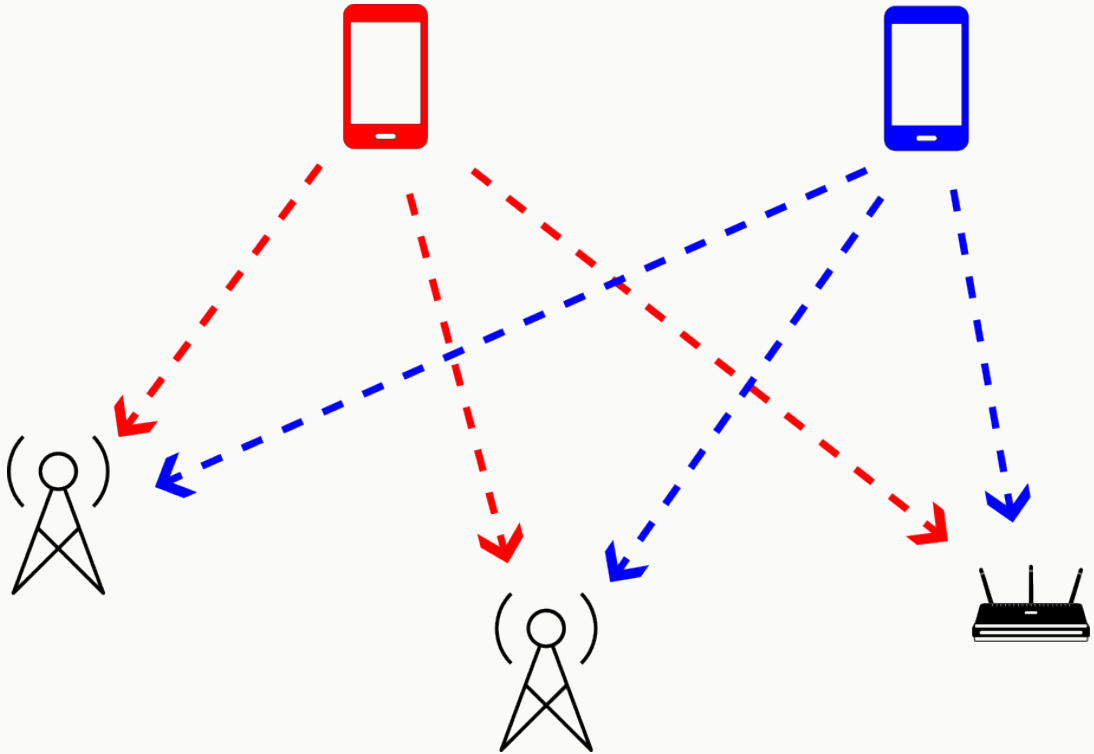
Passive DI: Other signal-based approaches

- RSS-based DI (also: location/positioning based)
  - Differentiate transmitting devices via RSS (and other metrics) of signals sent by them [5]
- Permutation-Entropy-based DI [6]
  - PE is the measure of complexity for a (chaotic) time series

[5] Yuan: MFMCF: A novel indoor location method combining multiple fingerprints and multiple classifiers (2019)

[6] Deng: Radio Frequency Fingerprint Extraction based on Multidimensional Permutation Entropy (2017)

Passive DI: Positioning-based DI

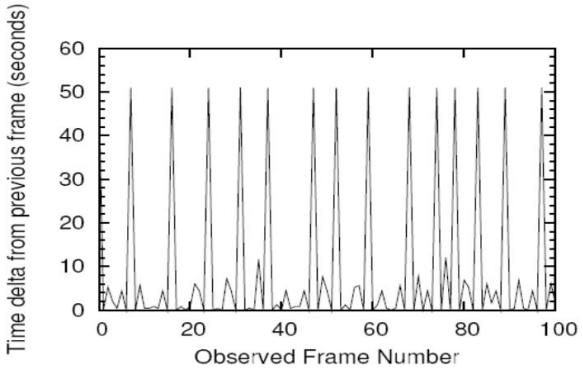


Passive DI: Approaches not using packet content

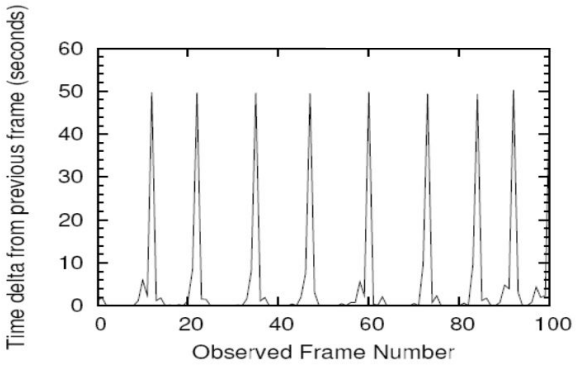
- Approaches so far are not using the packet's contents for DI
- We can use different parts of signals and derive identifying metrics
  - From preambles, transient, steady-state signals
  - From modulation errors
  - From relative signal power and Channel State Information (CSI)
- Usually best suited to identify specific devices with specific chips that classifiers have been trained for

Passive DI: Approaches using Behavior

- Behavior referring to observable behavior of devices, drivers etc.
- [7]: statistical analysis of the rate at which 802.11 data link layer frames are transmitted by a wireless device dependent on the specific driver



D-Link driver
D-Link DWL-G520 PCI Wireless NIC



Cisco driver
AIR-CB21AG-A-K9 PCI Wireless NIC

[7] Franklin: Passive Data Link 802.11 Wireless Device Driver Fingerprinting (2006)

Passive DI: Approaches using Scrambler-seeds

- *Scrambling* encodes a packet before transmitting, reverted by a *descrambler* at the receivers side
- Ensure that the sent signal has certain desired properties for transmitting and receiving
  - High entropy, reduce packet loss, make identical packets be transmitted differently, makes signal power spectrum independent of transmitted data
- Usually done by generating a pseudo-random bitstring added on top of the actual data stream (LFSR)
- Not to be confused with encryption
  - Encryption typically happens digitally, scrambling not (always) with the intention to provide confidentiality

Passive DI: Approaches using Scrambler-seeds

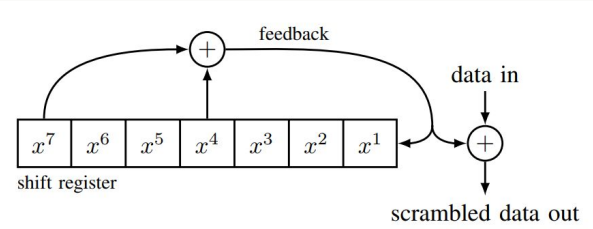


Figure 1. Schematic overview of the IEEE 802.11 scrambling algorithm.

[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)

Passive DI: Approaches using Scrambler-seeds

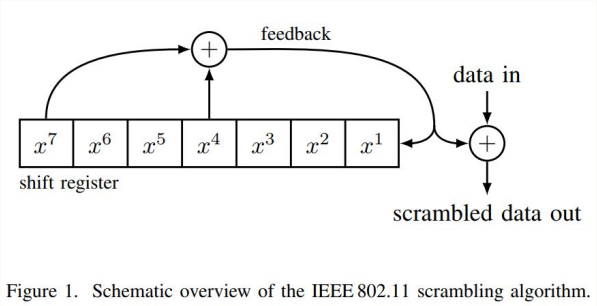
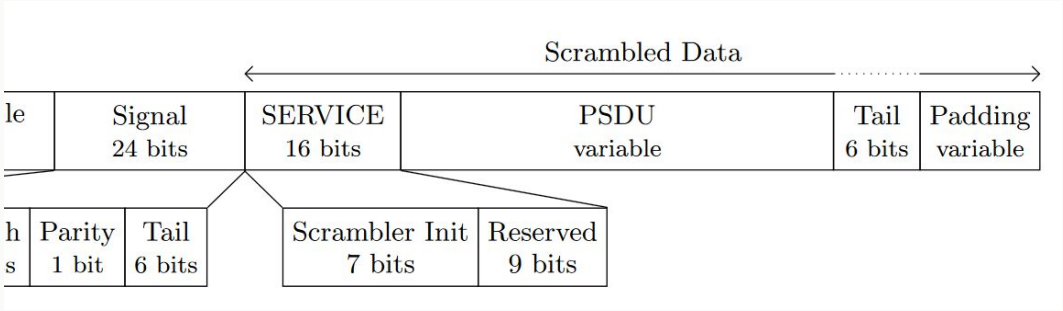


Figure 1. Schematic overview of the IEEE 802.11 scrambling algorithm.



[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)

Passive DI: Approaches using Scrambler-seeds

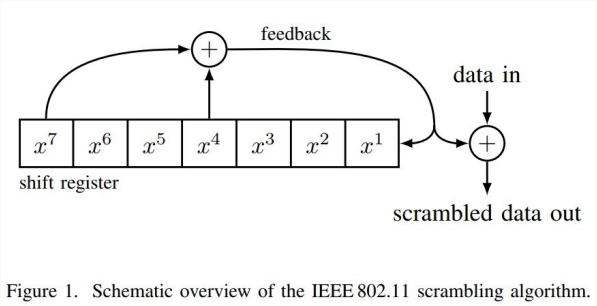
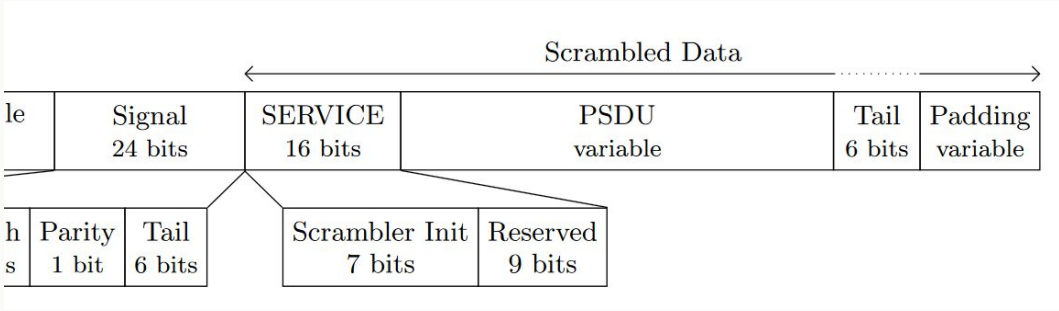


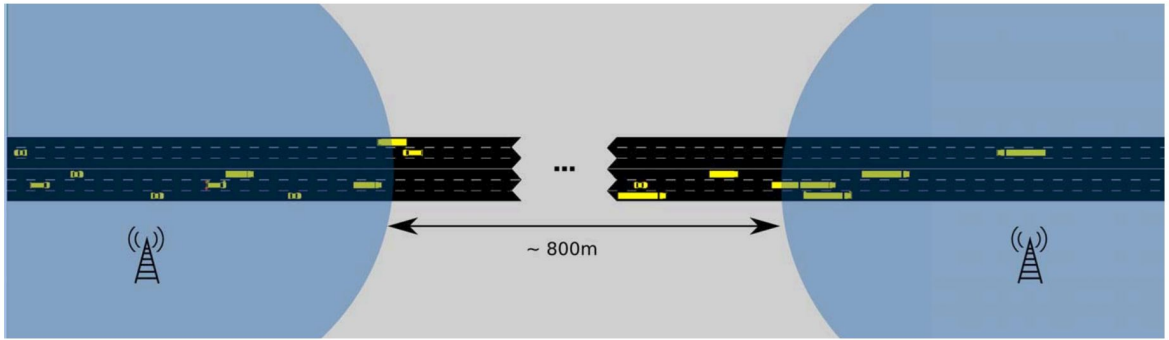
Figure 1. Schematic overview of the IEEE 802.11 scrambling algorithm.

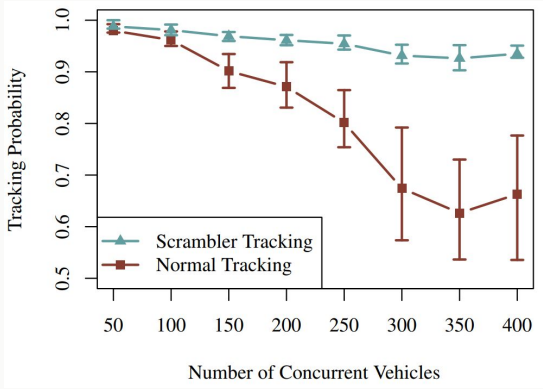


- Self-synchronizing Scrambler: *Scrambler Init* field initially all-zeros: after scrambling contains first 7 feedback bits
  - LFSR-state after 7 shifts.
- Since we *know* the scrambler state of a frame (Scrambler Init field) we can calculate the following states of the LFSR → predict scrambling sequence for next packet from this TD [8, 9]

[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)
[9] Bloessl: The Scrambler Attack: A Robust Physical Layer Attack on Location Privacy in Vehicular Networks (2015)

Passive DI: Approaches using Scrambler-seeds





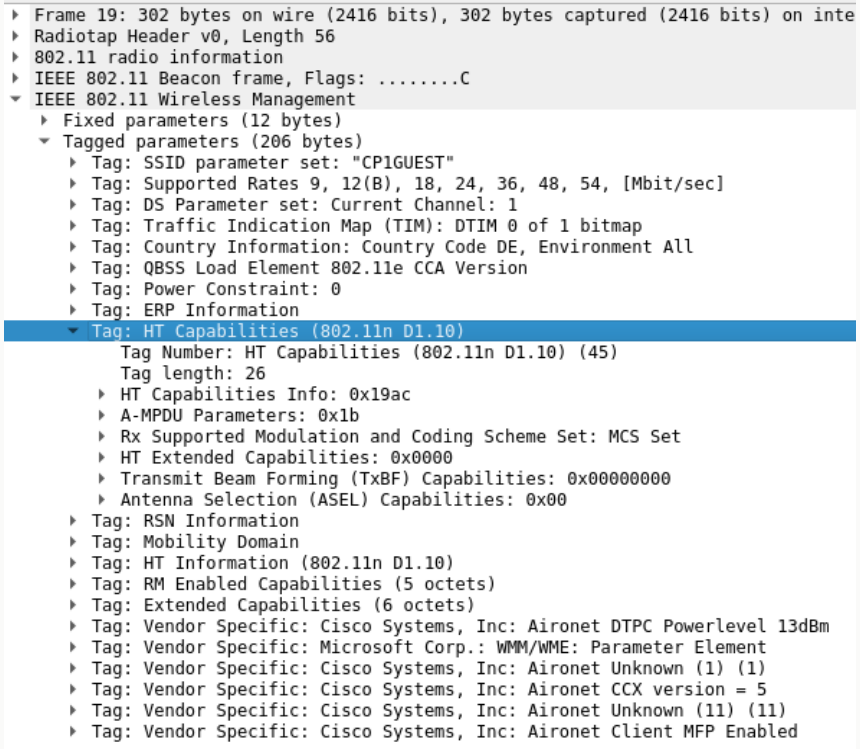
[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)
[9] Bloessl: The Scrambler Attack: A Robust Physical Layer Attack on Location Privacy in Vehicular Networks (2015)

Passive DI: Approaches using Identifiers

- The most obvious and trivial way to identify devices: **use device identifiers**
- Protocols use device identifiers to enable device addressability, e.g. **MAC address, IMEI, WPS UUID**
- Assuming the content of frame or packet headers are understandable by an eavesdropper
  - We can look into the header format and extract direct DIDs straight from there
  - If these identifiers are not changed from frame to frame per device, reidentification is possible
  - MAC headers may be sent unencryptedly, so eavesdropping may be possible
- Obvious problem for device privacy, so there are some countermeasures known
  - e.g. MAC randomization

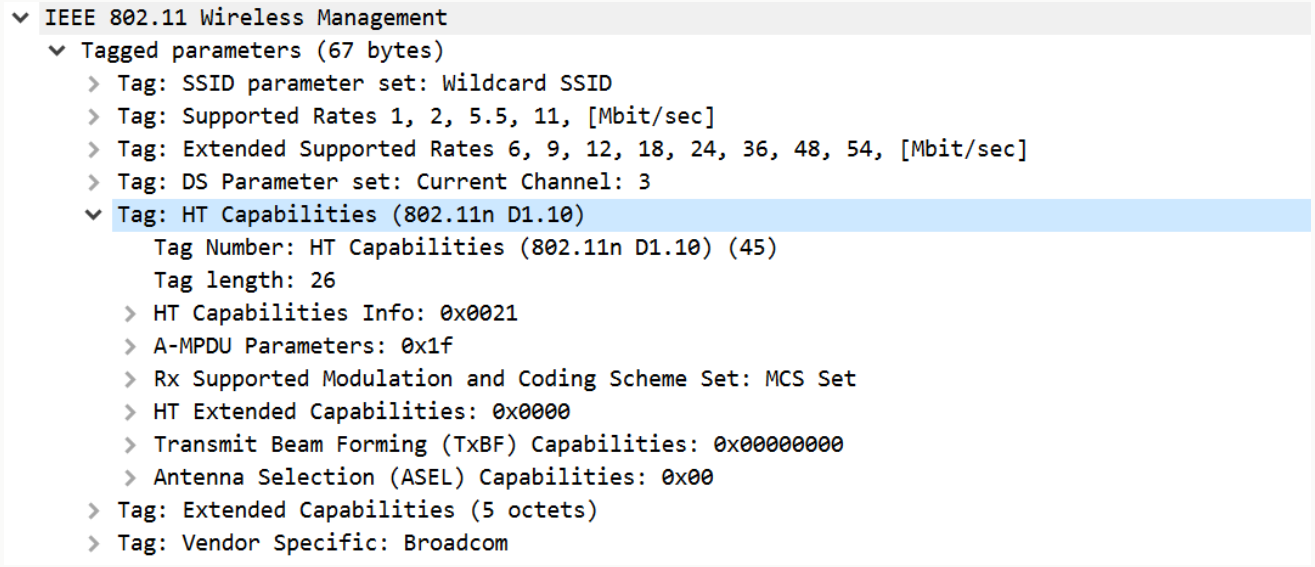
Passive DI: Approaches using Information Elements (IEs)

- IEs are a Wi-Fi specific part of the protocol
- Some IEs are mandatory for the protocol to function
- Many are optional, hinting at capabilities for other participating device to allow certain features to be used
- Certain combinations of IEs will de-anonymize devices even when the MAC-address is randomized
  → exercise
- e.g. Probe Request FP, SSID FP [8]



[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)

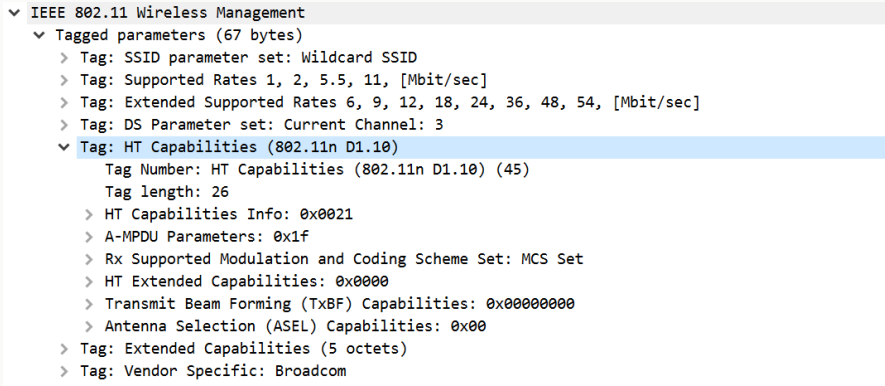
Passive DI: Approaches using Information Elements (IEs)



- Fingerprint using Vanhoef's technique:
  0,1,50,3,45,127,extcap:0000008001,extrates:0c1218243048606c,htagg:1f,htampdu:1f,htasel:00,htcap:0021,htext:0000,htmcs:000000000000000000000000000000ff,httx:00000000,supra:02040b16

Passive DI: Vanhoef Fingerprinting

- Vanhoef fingerprinting using IEs [8]
- Extract tags of existing IEs in PR to note existence and order of used IEs
- For each IE: note name and value as string
  → concatenate into a single FP string
- String compare allows to filter for packets using the exact same IE tags and IE values

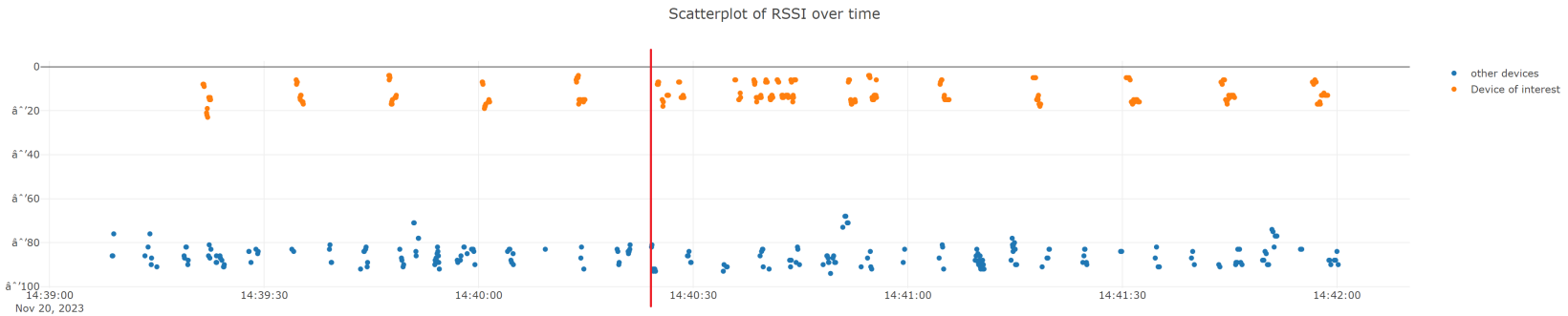


[8] Vanhoef: Why MAC Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms (2017)

Passive DI: Approaches using Information Elements (IEs)

CapturePi Web Interface START

Threshold: -25dBm



Time	Delta	MAC	Fingerprints	SSIDs
14:39:21	0.00s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:39:34	11.97s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:39:47	11.96s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:00	12.01s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:13	12.10s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:24	10.12s	e0:[redacted]:47:7f	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:27	1.42s	22:[redacted]:1a:76	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:35	7.13s	22:[redacted]:1a:76	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:43	0.37s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:40:51	7.38s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:41:04	8.90s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:41:17	11.75s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:41:30	11.97s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:41:43	11.43s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	
14:41:56	10.89s	ae:[redacted]:6e:6c	0,1,50,3,45,127,191,221(0x50f2,8),supra:02040b0c12161824,extrates:3048606c,htcap:016f,htagg:17,htmcs:000000000000000000000001000000ff,extcp:0000000000000040	

UNIVERSITÄT BONN

Passive DI recap

- Passive device identification means *listening in* on packets sent by devices over the shared medium and using these packets to *extract DID*
- DID can either be extracted by examining the *received signal and its properties* or by examining the *packet and its contents*
- Some approaches theoretically work on any RF transmission, if a large enough knowledge base was there to distinguish between all possible devices
- Some approaches are protocol specific and require devices to behave in a certain way, e.g. send their MAC address in clear, use QPSK, use predictable scrambler seeds
- Very few approaches provide direct DIDs and are generally easy to use
- In practice: Devices want to be identified to use services, so they provide identifying information

Active DI

- Active identification methods:
  - Generating purpose-built traffic with the TD and then observe their behavior
  - Also dependent on communication state the TD is in
  - Sometimes TD can be tricked into connecting to the attackers network, thus leaking additional info
- We actively engage the TD in order to have them send us/the network packets we can use for identification/tracking

MAC Randomization and why it's a problem

- Communication in heterogeneous networks like wireless networks requires addressability
- Communication protocols use device identifiers to make devices addressable
- MAC addresses are used for many different protocols, i.e. Wi-Fi, Bluetooth
- Sending the device MAC in clear all the time allows for addressability but also easy tracking

MAC Randomization and why it's a problem

- A commonly used countermeasure is MAC randomization
  - Instead of the global MAC address that identifies the RF chip a random MAC address is sent with the packets
  - Packets with a randomized MAC are not (easily) matched to individual devices anymore
- Doesn't this also make these devices not addressable anymore?

MAC Randomization and why it's a problem

- A commonly used countermeasure is MAC randomization
  - Instead of the global MAC address that identifies the RF chip a random MAC address is sent with the packets
  - Packets with a randomized MAC are not (easily) matched to individual devices anymore
- Doesn't this also make these devices not addressable anymore?
  - ***YES and NO***, not all sent packets need to be matched to an individual device to serve the purpose
  - i.e. Probe requests are used for network discovery, not for network connections
  - In general: When connection is not the purpose, any MAC suffices

MAC Randomization: global vs local MAC

- MAC randomization is not part of any specification
  - There is no clear instruction when and where and how to use it or if at all and for what packets etc
- We define the MAC address that the RF chip is identified by as the *global MAC address*
- Since randomized MACs can be used if the purpose is not to connect, how do devices get addressable when a connection is desired?
  - They can fall back to the global MAC address and use it for any connection (attempts)
  - They can also use a randomized address for a connection, if it doesn't change during the connection
  - We call this a *local MAC address* or a session MAC address

MAC Randomization identification

- How do we identify an observed MAC address as a randomized MAC?

MAC Randomization identification

- How do we identify an observed MAC address as a randomized MAC?
- MAC randomization is not part of any specification

MAC Randomization identification

- How do we identify an observed MAC address as a randomized MAC?
- MAC randomization is not part of any specification
  - In Wi-Fi: next to last Bit of the first MAC Byte is called *universal/local Bit*
  - If MAC global: set to 0
  - If MAC local: set to 1

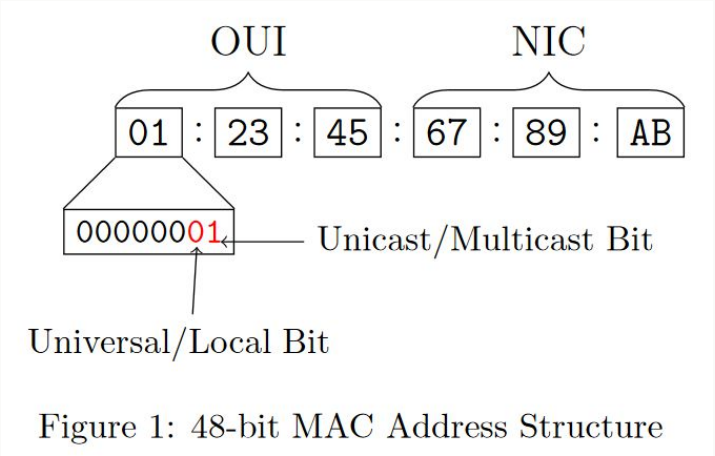


Figure 1: 48-bit MAC Address Structure

[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

MAC Randomization identification

- Different vendors implement MAC randomization differently
  - Not all implementations respect the universal/local Bit
  - Apple devices randomize all 47 Bits
  - Android devices vary in their implementation
  - Many devices only randomize the NIC, leaving the OUI intact
  - We can often still determine the vendor

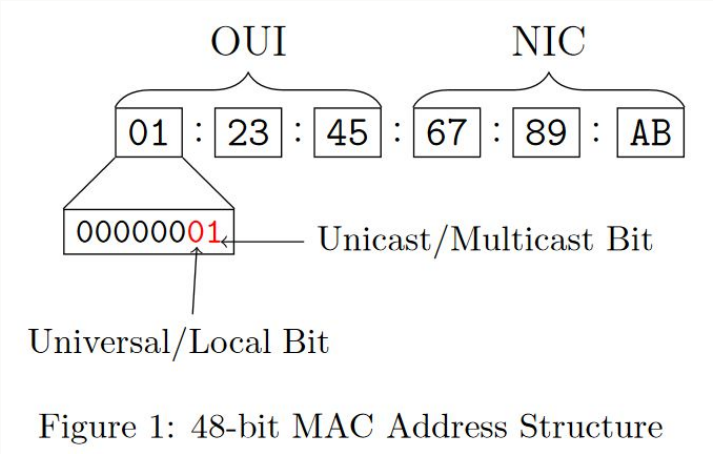


Figure 1: 48-bit MAC Address Structure

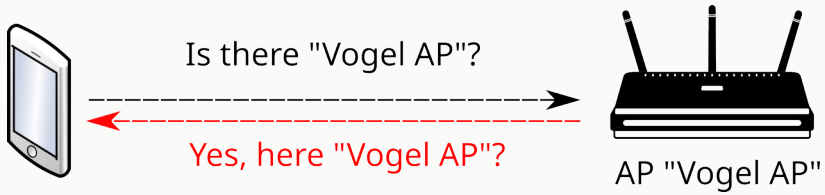
[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

UNIVERSITÄT BONN

Active DI vs MAC Randomization

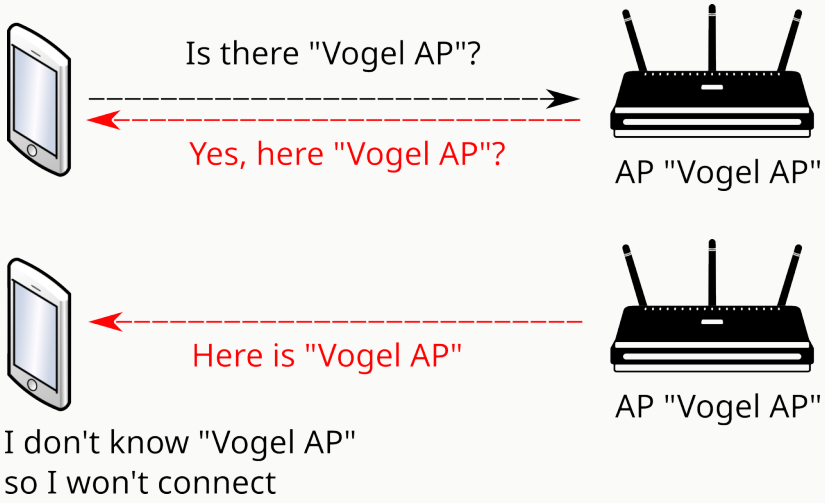
- We don’t really need active DI when we can observe direct DID passively
  - Most active DI techniques are basically attacks on MAC Randomization or other protection mechanisms
  - Some aim to increase the amount of traffic generated
- General idea is mostly the same: send a ”stimulus” packet and observe response or the lack of a response

Active DI: AP Impersonation attacks



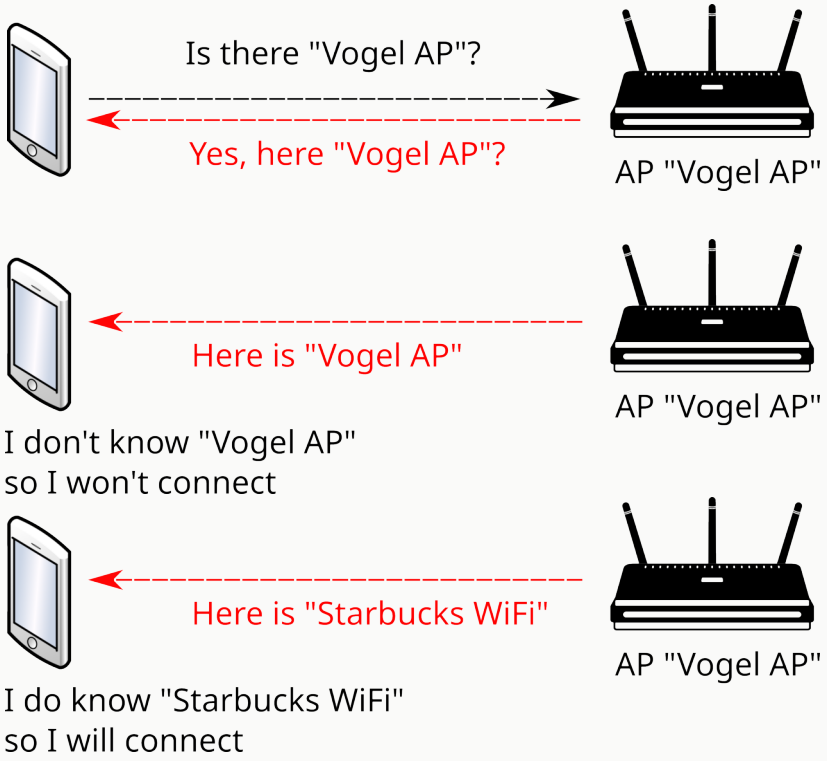
[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

Active DI: AP Impersonation attacks



[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

Active DI: AP Impersonation attacks



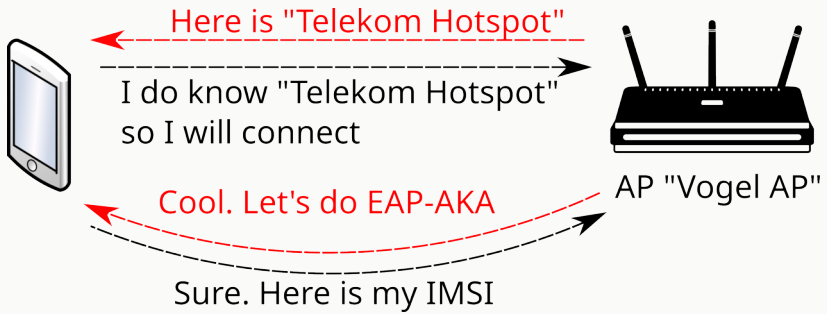
[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

Active DI: AP Impersonation attacks

- The *Karma Attack* listens to directed probe requests by the TD and impersonates an AP using a SSID that was observed
- The *Mana Attack* listens in to reconstruct the Preferred Network Lists (PNLs) of nearby devices before engaging in active communication
- The *Known Beacon Attack* tries to brute force the PNL of a TD trying commonly used SSIDs first
- WiFi-based IMSI Collector
- Hotspot 2.0 rogue AP

[9] Martin: A Study of MAC Address Randomization in Mobile Devices and When it Fails (2017)

Active DI: AP Impersonation attacks



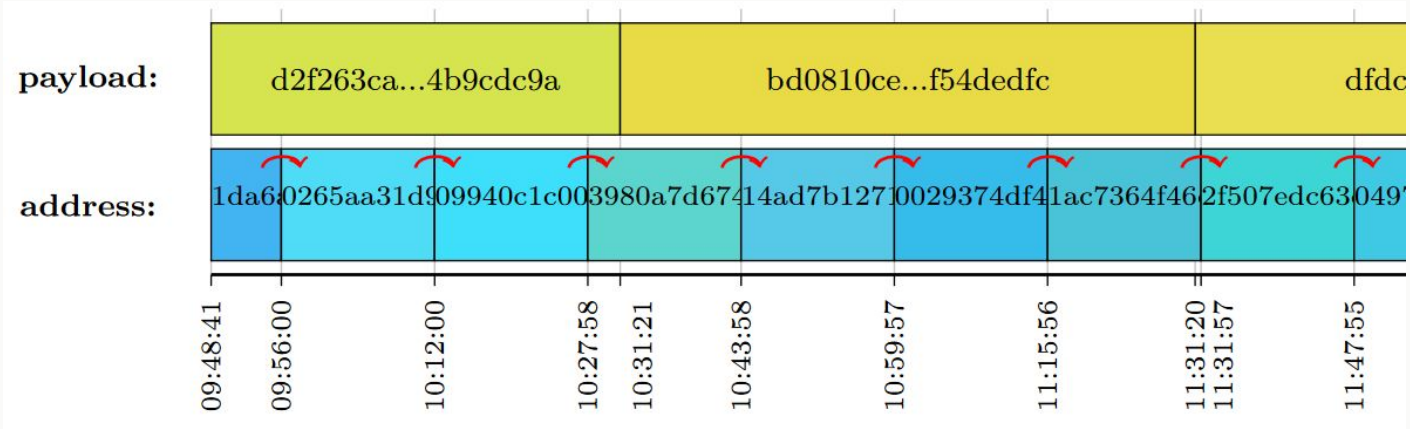
[10] O'Hanlon: Mobile Subscriber WiFi Privacy (2017)

Active DI: Bluetooth and Wi-Fi

- Bluetooth also uses MAC addresses and also uses MAC randomization
- In Apple iPhones Wi-Fi and Bluetooth MAC addresses were correlating starting with the iPhone 3G
  - Basically we can derive one MAC from the other using [9]

Active DI: Bluetooth Randomization

- Randomized Bluetooth MACs update at a different interval than the content of the sent packets
  - This allows to track bluetooth devices through MAC randomization as long as the observation is somewhat gapless



[11] Becker: Tracking Anonymized Bluetooth Devices (2019)

Passive DI vs Active DI recap

- Passive DI
  - Application Layer FP can be circumvented by changing the application parameters
  - Physical Layer FP is more difficult to compromise due to inherent physical properties
  - Good features are very difficult to find and collect
- Active DI
  - More flexible approaches and allows to explore different scenarios
  - Easier to detect from the authority

UNIVERSITÄT BONN

Summary Device Identification

- Techniques for device identification span all network layers for a variety of objectives (defensive or offensive)
- The accuracy of identification depends on many factors (i.e. hardware, experimental conditions, features and classification procedures)
- There is comparably little work on the resilience of identification with respect to the above and other attacks
- Major research interest as wireless devices get more and more ubiquitous
  - Identifying a device often also identifies the person carrying it

UNIVERSITÄT BONN

Some notes on the exercise

- Exercise sheet contains a practical task, where you examine a Wi-Fi capture of a fictional burglary, aim to identify the intruding device, and extract information about it
  - You will need to implement a fingerprinting method and use it to identify the target device
  - Please note the derived fingerprint for the target device and answer the given questions using that fingerprint
  - You have 1 week for this exercise

UNIVERSITÄT BONN

IT Security

Device identification in wireless networks

Tasks via mailing list

Daniel Vogel
Friedrich-Hirzebruch-Allee 8
53115 Bonn
vogel@cs.uni-bonn.de

December 05, 2024

Lightning Surveys ⚡