Tracking 802.11 stations without relying on the link layer identifier

Mathy Vanhoef[†], Célestin Matte[‡], Mathieu Cunche[‡], Leonardo S. Cardoso[‡], Frank Piessens[†]

[†] iMinds-Distrinet, KU Leuven , [‡] Univ Lyon, INSA Lyon, Inria, CITI, France

IEEE P802E - 14th April 2016

Introduction

- The paper
  - *Why MAC Address Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms*. Accepted at AsiaCCS 2016.

Introduction

- MAC address randomization proposed to prevent tracking
  - Idea of a disposable link-layer identifier
  - Being deployed in major OSes
    - iOS 8, Android 6, Windows 10, Linux kernel 3.18
- Is it enough to prevent tracking ?
  - Probe requests contains a lot of other information
  - Can we track devices despite the lack of a stable link-layer identifier ?
  - Can we link together probes from the same device based on their content ?
  - Can we force a device to reveal its real MAC address ?

Introduction

- Attacker capabilities
  - Monitoring wireless channels
  - Injecting 802.11 frames
- Attacker objectives
  - Group together frames belonging to the same device
- Link-Layer identifier is assumed to change periodically

Introduction

Datasets

Table : Details of the probe requests datasets.

Dataset	Lab	Train-station	Sapienza[1]
#MAC addr.	500	10 000	160 000
#Probe Req.	120 000	110 000	8 million
Time frame	Oct '15	Oct/Nov '15	Feb/May '13
Location	Lab	Train Station	Rome

[1] Marco V. Barbera et al. *CRAWDAD dataset sapienza/probe-requests (v. 2013-09-10)*. Downloaded from http://crawdad.org/sapienza/probe-requests/20130910. Sept. 2013. DOI: 10.15783/C76C7Z.

Fingerprinting using Information Elements

- Information elements (a.k.a. tagged parameters, or tags)
  - Indicates the support of capabilities
  - Ex. `Supported Rates`, `High Throughput capabilities` and `Interworking Capabilities`
- High diversity in term of values and in term of information elements present in probe requests
  - Idea: Exploit this diversity to fingerprint devices

Fingerprinting using Information Elements

```
▼Tag: HT Capabilities (802.11n D1.10)
  Tag Number: HT Capabilities (802.11n D1.10) (45)
  Tag length: 26
 ▼HT Capabilities Info: 0x100c
    .... .... .... ...0 = HT LDPC coding capability: Transmitter does not support receiving LDPC coded packets
    .... .... .... ..0. = HT Support channel width: Transmitter only supports 20MHz operation
    .... .... .... 11.. = HT SM Power Save: SM Power Save disabled (0x0003)
    .... .... ...0 .... = HT Green Field: Transmitter is not able to receive PPDUs with Green Field (GF) preamble
    .... .... ..0. .... = HT Short GI for 20MHz: Not supported
    .... .... .0.. .... = HT Short GI for 40MHz: Not supported
    .... .... 0... .... = HT Tx STBC: Not supported
    .... ..00 .... .... = HT Rx STBC: No Rx STBC support (0x0000)
    .... .0.. .... .... = HT Delayed Block ACK: Transmitter does not support HT-Delayed BlockAck
    .... 0... .... .... = HT Max A-MSDU length: 3839 bytes
    ...1 .... .... .... = HT DSSS/CCK mode in 40MHz: Will/Can use DSSS/CCK in 40 MHz
    ..0. .... .... .... = HT PSMP Support: Won't/Can't support PSMP operation
    .0.. .... .... .... = HT Forty MHz Intolerant: Use of 40 MHz transmissions unrestricted/allowed
    0... .... .... .... = HT L-SIG TXOP Protection support: Not supported
 ▼A-MPDU Parameters: 0x19
    .... ..01 = Maximum Rx A-MPDU Length: 0x01 (16383[Bytes])
    ...1 10.. = MPDU Density: 8 [usec] (0x06)
    000. .... = Reserved: 0x00
 ▶Rx Supported Modulation and Coding Scheme Set: MCS Set
 ▶HT Extended Capabilities: 0x0000
 ▶Transmit Beam Forming (TxBF) Capabilities: 0x0000
 ▶Antenna Selection (ASEL) Capabilities: 0x00
```

Figure : Example of the `HT_Extended_capabilities` Information Element

Fingerprinting using Information Elements

Empirical evaluation using the datasets

- Considered metrics
  - Fraction of affected devices
  - Entropy: amount of identifying information
- Single Information Elements
  - Can provide up to 5.24 bits of entropy
  - Some IE are found in almost all device (`Supported rates`)
  - Ex. `HT capabilities info` (`Train-station`) : 4.74 bits of entropy, 90% of devices affected, stable for 95.9% devices
- Global fingerprint based on most common IE
  - Entropy : 7.03 bits (`Train-station`)
  - Enough to uniquely identify 1 device among 128 (in average)

Element	Entropy (bits)			Stability			Affected devices		
	Lab	Station	Sapienza	Lab	Station	Sapienza	Lab	Station	Sapienza
HT capabilities info	3.94	4.74	3.35	96.0%	95.9%	99.6%	90.9%	90.0%	81.1%
Ordered list of tags numbers	4.23	5.24	4.10	93.6%	94.2%	91.2%	100%	100%	100%
Extended capabilities	2.59	2.57	0.064	98.5%	99.4%	99.9%	55.4%	51.3%	0.6%
HT A-MPDU parameters	2.59	2.67	2.54	97.8%	99.1%	99.7%	90.9%	90.0%	81.1%
HT MCS set bitmask	1.49	1.43	1.16	97.6%	99.0%	99.9%	90.9%	90.0%	81.1%
Supported rates	1.18	2.10	1.36	98.2%	95.9%	99.8%	100%	99.9%	100%
Interworking - access net. type	1.08	1.11	0.006	99.6%	99.6%	100.0%	47.5%	46.1%	0.04%
Extended supported rates	1.00	1.77	0.886	98.0%	96.3%	99.4%	99.1%	72.6%	99.7%
WPS UUID	0.878	0.788	0.658	98.2%	99.2%	99.6%	8.4%	5.5%	3.6%
HT extended capabilities	0.654	0.623	0.779	97.8%	98.9%	99.9%	90.9%	90.0%	81.1%
HT TxBeam Forming Cap.	0.598	0.587	0.712	97.8%	98.9%	99.9%	90.9%	90.0%	81.1%
HT Antenna Selection Cap.	0.579	0.576	0.711	98.0%	98.9%	99.9%	90.9%	90.0%	81.1%
Overall	5.48	7.03	5.65	92.5%	90.7%	88.8%	-	-	-

Fingerprinting using Information Elements

Wi-Fi Protected Setup (WPS)

- Information element dedicated to WPS
  - Includes a UUID field
- Universally Unique Identifier UUID
  - A unique identifier *by definition*
  - Generally derived from the MAC address[2]
  - Could be reversed to reveal the original MAC
- Re-identification attack on the datasets
  - UUID derived from the real Wi-Fi MAC address in 75% of the cases

[2] P. Leach, M. Mealling, and R. Salz. *A Universally Unique IDentifier (UUID) URN Namespace*. RFC 4122 (Proposed Standard). Internet Engineering Task Force, July 2005. URL: http://www.ietf.org/rfc/rfc4122.txt.

Predictable scrambler seed

Predictable fields

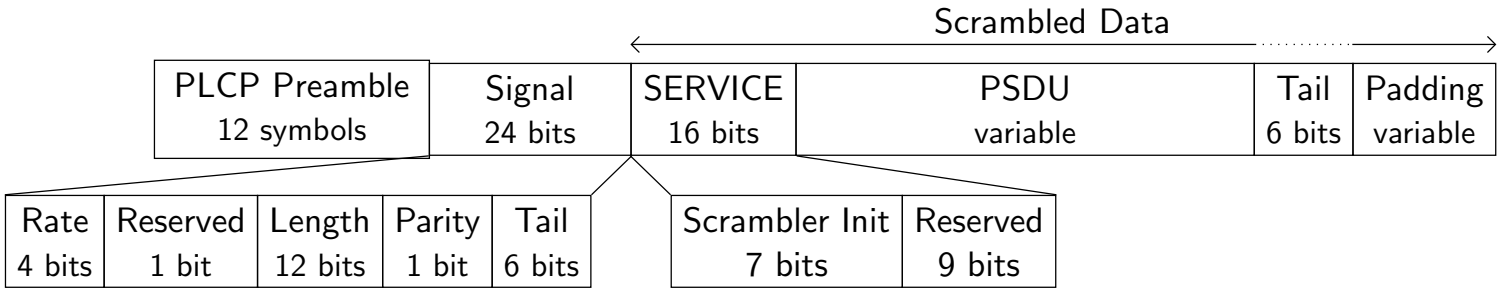
- Predictable fields in 802.11 frames
  - Fields with a content that can change over time
  - Value in a given frame can be predicted from the previous frames
  - Example: `Sequence Number` field
    - Incremented for each frame
    - Not reset when MAC address is changed in iOS[3]
    - Can be used to trivially defeat MAC Randomization

[3] Julien Freudiger. "How Talkative is Your Mobile Device? An Experimental Study of Wi-Fi Probe Requests". In: *WiSec*. 2015.

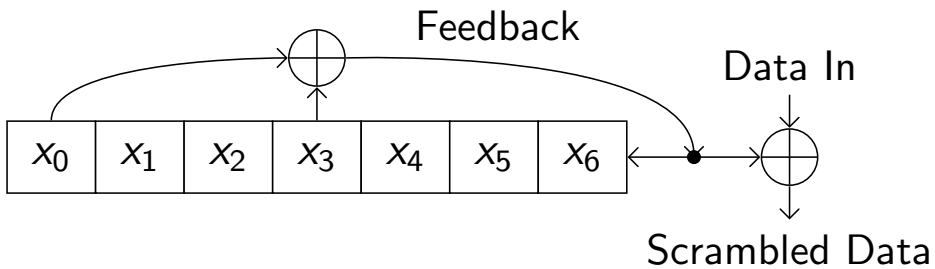
Predictable scrambler seed

Scrambler seed

- Scrambler in OFDM frames of 802.11 PHY
  - Scrambler used from the `SERVICE` field to the end
  - Seed contained in the 7 first bits of `SERVICE` field



- Scrambling sequence generated by a Linear Feedback Shift Register (LFSR)
  - Seed set the initial state of LFSR



Predictable scrambler seed

- Scrambler seeds can be predictable
  - Bloessl. et al. showed that it is the case for two prototype implementation of 802.11p[4]
  - No specification in the standard on how to generate the seeds
  - Implementation choice taken by the vendor
- What about commodity 802.11 implementations ?

[4] B. Bloessl et al. "The scrambler attack: A robust physical layer attack on location privacy in vehicular networks". In: *ICNC*. 2015.

Predictable scrambler seed

- Study of scrambler seeds in 802.11 commodity hardware
  - Experimental setup
    - 11 Wi-Fi commodity hardware
    - GNU-Radio implementation of 802.11 based on `gr-ieee802-11`[5]
    - USRP N210
    - Faraday room from FIT CortexLab[6]

[5] Bastian Bloessl et al. "An IEEE 802.11 a/g/p OFDM Receiver for GNU Radio". In: *SRIF Workshop*. 2013.

[6] `http://www.cortexlab.fr/`

Predictable scrambler seed

- Observed behaviors
  - Freewheeling: State of the LFSR at the end of a frame is reused for the next frame
    - Sometime with a constant number of shift of the LFSR
  - Constant seed, or limited to a small set (bug ?)
  - Incremental: seed value is incremented by one at each frame

Active attacks

Active attacks

- Karma attack
  - Fake AP with popular SSID
  - Trigger authentication/association from STA
  - STA switch back to their real MAC when connecting to AP
- Exploiting Hotspot 2.0
  - Enable Wi-Fi roaming
  - STA send ANQP query to AP to retrieve list of available services
  - STA switch back to their real MAC addr. when querying
  - Query also contain predictable counter that could help tracking

Countermeasures

- Information elements in probe requests
  - Are they really needed ?
  - Remove them or restrict to a bare minimum
- Scrambler seed and counters
  - Reset to a random value upon MAC addr change
  - Unpredictable scrambler seeds
    - Use a crypto PRNG to generate seeds
    - Chipsets allowing a reset of the seed
- Active attacks
  - Keep random MAC addr. when sending ANQP queries