Privacy-Enhancing Technologies Against Physical-Layer and Link-Layer Device Tracking

Trends, Challenges, and Future Directions

Apolline Zehner, Iness Ben Guirat, Jan Tobias Mühlberg

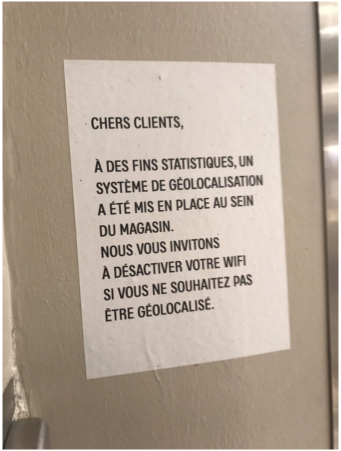
Université libre de Bruxelles

28-02-2025

Let's go shopping!

Translation

Dear customers,
For statistical purposes, a geolocation system has been set up within the store.
We invite you to deactivate your Wi-Fi if you do not wish to be geolocated.

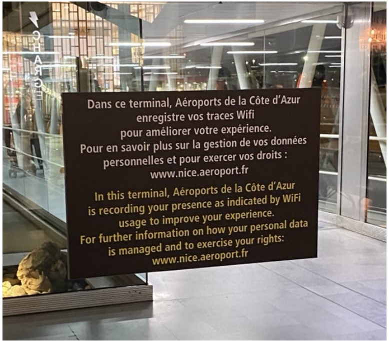


CHERS CLIENTS,

À DES FINS STATISTIQUES, UN SYSTÈME DE GÉOLOCALISATION A ÉTÉ MIS EN PLACE AU SEIN DU MAGASIN.
NOUS VOUS INVITONS À DÉSACTIVER VOTRE WIFI SI VOUS NE SOUHAITEZ PAS ÊTRE GÉOLOCALISÉ.

Source:
`https://xcancel.com/adhavet/status/891693199424729092`

It's not that uncommon...



Dans ce terminal, Aéroports de la Côte d'Azur
enregistre vos traces Wifi
pour améliorer votre expérience.
Pour en savoir plus sur la gestion de vos données
personnelles et pour exercer vos droits :
www.nice.aeroport.fr

In this terminal, Aéroports de la Côte d'Azur
is recording your presence as indicated by WiFi
usage to improve your experience.
For further information on how your personal data
is managed and to exercise your rights:
www.nice.aeroport.fr

Source: a friend of mine

Source: https://www.radiofrance.fr/franceinter/metro-des-ecrans-publicitaires-video-pour-capter-votre-attention-mais-jure-pas-vos-donnees-1176690

What do manufacturers say?

”Key use cases of BLE technology”

BLE beacons are used for location tracking and can provide the most accurate way to track exact location in indoor spaces. (...)

By combining BLE with other tracking technologies such as Wi-Fi and RFID, businesses can create detailed customer profiles and track customer movements throughout the facility in real time.

Source: `https://spaces.cisco.com/key-use-cases-of-ble-technology/`

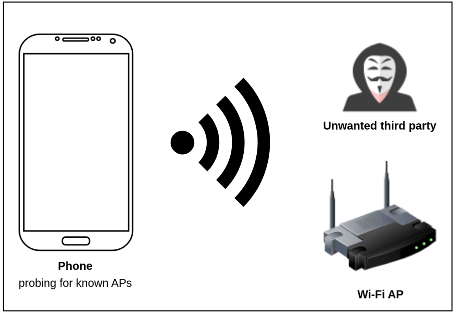
Why is it important?

- You only need a few anonymized location data points to re-identify someone[1];
- Data brokers sell anonymized location data points of people across the globe, allowing re-identification of targeted people[2].

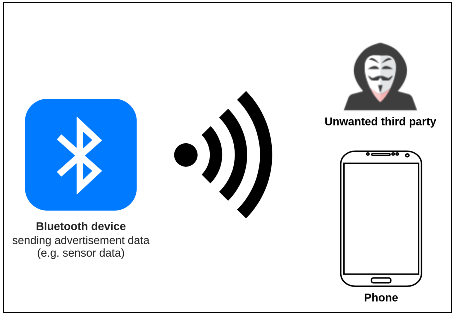
[1] Y.-A. de Montjoye *et al.*, “Unique in the crowd: The privacy bounds of human mobility,” *Scientific Reports*, vol. 3, no. 1, p. 1376, Mar. 25, 2013, Publisher: Nature Publishing Group, ISSN: 2045-2322. DOI: 10.1038/srep01376. [Online]. Available: https://www.nature.com/articles/srep01376 (visited on 15-02-2025)

[2] S. Meineck and I. Dachwitz, “Data Broker Files: How data brokers sell our location data and jeopardise national security,” netzpolitik.org, (Jul. 16, 2024), [Online]. Available: https://netzpolitik.org/2024/data-broker-files-how-data-brokers-sell-our-location-data-and-jeopardise-national-security/ (visited on 15-02-2025)

How does it work?



Schematic representation of probing requests



Schematic representation of advertisements by a Bluetooth device

Solution: randomizing MAC addresses

MAC address randomization: devices regularly randomize their MAC addresses.

Mobile operating systems implements it since at least 2014 (Apple) or 2017 (Google).

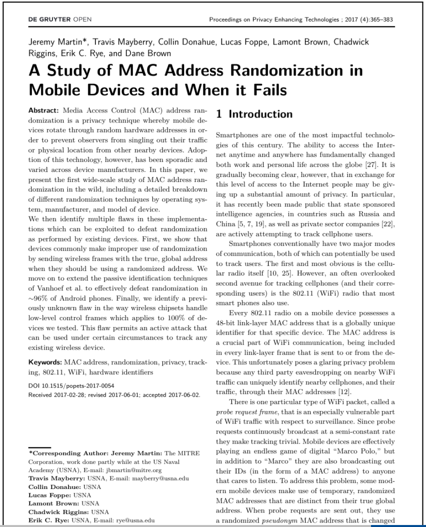
Could my favorite supermarket still track my device (thus me)?

Implementation flaws

Implementation flaws allows device de-anonymization...

- by using sequential sequence numbers;
- by sending data with the real address, not the randomized one;
- through fingerprinting, as devices have different signatures.

J. Martin *et al.*, "A study of MAC address randomization in mobile devices and when it fails," *Proceedings on Privacy Enhancing Technologies*, vol. 2017, no. 4, pp. 365–383, Oct. 1, 2017, ISSN: 2299-0984. DOI: 10.1515/popets-2017-0054. [Online]. Available: `https://petsymposium.org/popets/2017/popets-2017-0054.php` (visited on 11-01-2025)



DE GRUYTER OPEN

Proceedings on Privacy Enhancing Technologies ; 2017 (4):365–383

Jeremy Martin*, Travis Mayberry, Collin Donahue, Lucas Foppe, Lamont Brown, Chadwick Riggins, Erik C. Rye, and Dane Brown

A Study of MAC Address Randomization in Mobile Devices and When it Fails

Abstract: Media Access Control (MAC) address randomization is a privacy technique whereby mobile devices rotate through random hardware addresses in order to prevent observers from singling out their traffic or physical location from other nearby devices. Adoption of this technology, however, has been sporadic and varied across device manufacturers. In this paper, we present the first wide-scale study of MAC address randomization in the wild, including a detailed breakdown of different randomization techniques by operating system, manufacturer, and model of device.

We then identify multiple flaws in these implementations which can be exploited to defeat randomization as performed by existing devices. First, we show that devices commonly make improper use of randomization by sending wireless frames with the true, global address when they should be using a randomized address. We move on to extend the passive identification techniques of Vanhoef et al. to effectively defeat randomization in ~96% of Android phones. Finally, we identify a previously unknown flaw in the way wireless chipsets handle low-level control frames which applies to 100% of devices we tested. This flaw permits an active attack that can be used under certain circumstances to track any existing wireless device.

Keywords: MAC address, randomization, privacy, tracking, 802.11, WiFi, hardware identifiers

DOI 10.1515/popets-2017-0054

Received 2017-02-28; revised 2017-06-01; accepted 2017-06-02.

1 Introduction

Smartphones are one of the most impactful technologies of this century. The ability to access the Internet anytime and anywhere has fundamentally changed both work and personal life across the globe [27]. It is gradually becoming clear, however, that in exchange for this level of access to the Internet people may be giving up a substantial amount of privacy. In particular, it has recently been made public that state sponsored intelligence agencies, in countries such as Russia and China [5, 7, 19], as well as private sector companies [22], are actively attempting to track cellphone users.

Smartphones conventionally have two major modes of communication, both of which can potentially be used to track users. The first and most obvious is the cellular radio itself [10, 25]. However, an often overlooked second avenue for tracking cellphones (and their corresponding users) is the 802.11 (WiFi) radio that most smart phones also use.

Every 802.11 radio on a mobile device possesses a 48-bit link-layer MAC address that is a globally unique identifier for that specific device. The MAC address is a crucial part of WiFi communication, being included in every link-layer frame that is sent to or from the device. This unfortunately poses a glaring privacy problem because any third party eavesdropping on nearby WiFi traffic can uniquely identify nearby cellphones, and their traffic, through their MAC addresses [12].

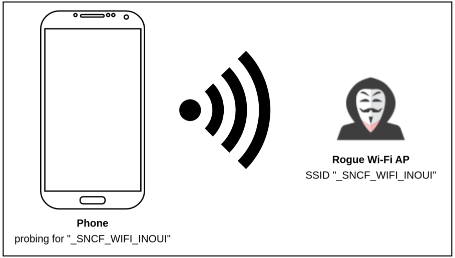
There is one particular type of WiFi packet, called a *probe request frame*, that is an especially vulnerable part of WiFi traffic with respect to surveillance. Since probe requests continuously broadcast at a semi-constant rate they make tracking trivial. Mobile devices are effectively playing an endless game of digital "Marco Polo," but in addition to "Marco" they are also broadcasting out their IDs (in the form of a MAC address) to anyone that cares to listen. To address this problem, some modern mobile devices make use of temporary, randomized MAC addresses that are distinct from their true global address. When probe requests are sent out, they use a randomized *pseudonym* MAC address that is changed

***Corresponding Author: Jeremy Martin:** The MITRE Corporation, work done partly while at the US Naval Academy (USNA), E-mail: jbmartin@mitre.org
Travis Mayberry: USNA, E-mail: mayberry@usna.edu
Collin Donahue: USNA
Lucas Foppe: USNA
Lamont Brown: USNA
Chadwick Riggins: USNA
Erik C. Rye: USNA, E-mail: rye@usna.edu

How to get a device to reveal its MAC address?

Devices switch to their real MAC addresses when associating to an AP.

- All you need is setting up a rogue AP that uses popular SSIDs!



Schematic representation of probing and connecting to a rogue AP.

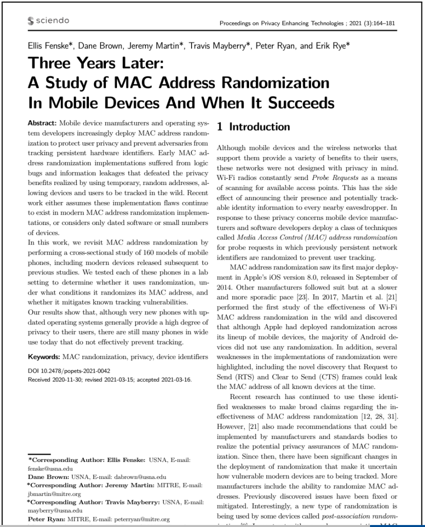
Fingerprinting devices

Most implementation flaws are now patched.

- Devices randomize sequence numbers;
- Devices use randomized MAC addresses even while connected.

What about fingerprinting?

E. Fenske *et al.*, "Three years later: A study of MAC address randomization in mobile devices and when it succeeds," *Proceedings on Privacy Enhancing Technologies*, vol. 2021, no. 3, pp. 164–181, Jul. 1, 2021, ISSN: 2299-0984. DOI: 10.2478/popets-2021-0042. [Online]. Available: https://petsymposium.org/popets/2021/popets-2021-0042.php (visited on 11-01-2025)



sciendo
Proceedings on Privacy Enhancing Technologies ; 2021 (3):164–181

Ellis Fenske*, Dane Brown, Jeremy Martin*, Travis Mayberry*, Peter Ryan, and Erik Rye*

Three Years Later: A Study of MAC Address Randomization In Mobile Devices And When It Succeeds

Abstract: Mobile device manufacturers and operating system developers increasingly deploy MAC address randomization to protect user privacy and prevent adversaries from tracking persistent hardware identifiers. Early MAC address randomization implementations suffered from logic bugs and information leakages that defeated the privacy benefits realized by using temporary, random addresses, allowing devices and users to be tracked in the wild. Recent work either assumes these implementation flaws continue to exist in modern MAC address randomization implementations, or considers only dated software or small numbers of devices.

In this work, we revisit MAC address randomization by performing a cross-sectional study of 160 models of mobile phones, including modern devices released subsequent to previous studies. We tested each of these phones in a lab setting to determine whether it uses randomization, under what conditions it randomizes its MAC address, and whether it mitigates known tracking vulnerabilities.

Our results show that, although very new phones with updated operating systems generally provide a high degree of privacy to their users, there are still many phones in wide use today that do not effectively prevent tracking.

Keywords: MAC randomization, privacy, device identifiers

DOI 10.2478/popets-2021-0042
Received 2020-11-30; revised 2021-03-15; accepted 2021-03-16.

1 Introduction

Although mobile devices and the wireless networks that support them provide a variety of benefits to their users, these networks were not designed with privacy in mind. Wi-Fi radios constantly send *Probe Requests* as a means of scanning for available access points. This has the side effect of announcing their presence and potentially trackable identity information to every nearby eavesdropper. In response to these privacy concerns mobile device manufacturers and software developers deploy a class of techniques called *Media Access Control (MAC) address randomization* for probe requests in which previously persistent network identifiers are randomized to prevent user tracking.

MAC address randomization saw its first major deployment in Apple's iOS version 8.0, released in September of 2014. Other manufacturers followed suit but at a slower and more sporadic pace [23]. In 2017, Martin et al. [21] performed the first study of the effectiveness of Wi-Fi MAC address randomization in the wild and discovered that although Apple had deployed randomization across its lineup of mobile devices, the majority of Android devices did not use any randomization. In addition, several weaknesses in the implementations of randomization were highlighted, including the novel discovery that Request to Send (RTS) and Clear to Send (CTS) frames could leak the MAC address of all known devices at the time.

Recent research has continued to use these identified weaknesses to make broad claims regarding the ineffectiveness of MAC address randomization [12, 28, 31]. However, [21] also made recommendations that could be implemented by manufacturers and standards bodies to realize the potential privacy assurances of MAC randomization. Since then, there have been significant changes in the deployment of randomization that make it uncertain how vulnerable modern devices are to being tracked. More manufacturers include the ability to randomize MAC addresses. Previously discovered issues have been fixed or mitigated. Interestingly, a new type of randomization is being used by some devices called *post-association random-*

*Corresponding Author: Ellis Fenske: USNA, E-mail: fenske@usna.edu
Dane Brown: USNA, E-mail: dabrown@usna.edu
*Corresponding Author: Jeremy Martin: MITRE, E-mail: jbmartin@mitre.org
*Corresponding Author: Travis Mayberry: USNA, E-mail: mayberry@usna.edu
Peter Ryan: MITRE, E-mail: peterryan@mitre.org

Using metadata to fingerprint devices

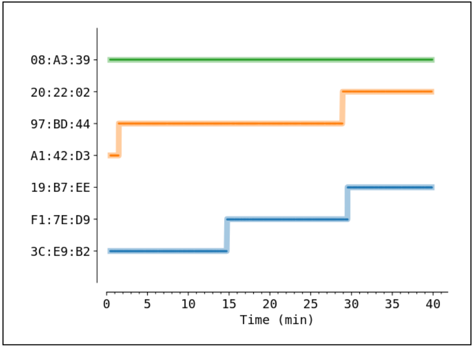
Multiple kinds of metadata are available through passive communication listening:

- MAC address randomization time intervals;
- Emitting power through the Received Signal Strength Indicator (RSSI);
- Device-specific information, sent along with necessary data while probing/advertising.

MAC address randomization through time

Devices randomize their MAC addresses at quasi-constant and non-standardized time intervals.

It is unlikely that a device will randomize its MAC address at the same time as another one, several times.



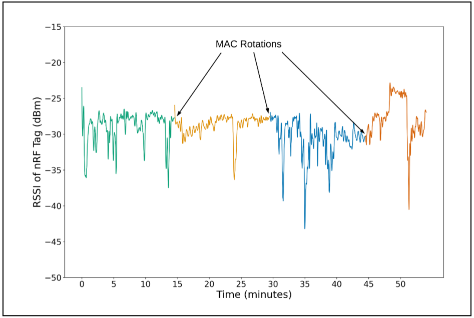
Source: L. Jouans *et al.*, "Associating the randomized bluetooth MAC addresses of a device," in *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, ISSN: 2331-9860, Jan. 2021, pp. 1–6. DOI: 10.1109/CCNC49032.2021.9369628. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9369628 (visited on 11-01-2025)

Use of RSSI for device de-anonymization

When randomizing their MAC adresses, devices are still advertising themselves at the same frequency.

RSSI is unlikely to change during that time, allowing device de-anonymization.

Useful applications against unwanted tracking devices (e.g. AirTags).



Source: T. Despres *et al.*, “DeTagTive: Linking MACs to protect against malicious BLE trackers,” in *Proceedings of the Second Workshop on Situating Network Infrastructure with People, Practices, and Beyond*, ser. SNIP2+ ’23, New York, NY, USA: Association for Computing Machinery, Sep. 10, 2023, pp. 1–7, ISBN: 979-8-4007-0304-1. DOI: 10.1145/3609396.3610544. [Online]. Available: https://dl.acm.org/doi/10.1145/3609396.3610544 (visited on 28-11-2024)

Possible countermeasures

- Randomizing emitting power, thus RSSI;
- Re-using previously generated MAC addresses;
- Introducing silent periods to decrease accuracy of probabilistic methods;
- Minimizing the amount of data sent, like unique identifiers.

Challenges & New Directions

- Metadata-related attacks (e.g. based on RSSI) are not well-researched at the moment: combining those with existing attacks may reduce the effectiveness of existing countermeasures;

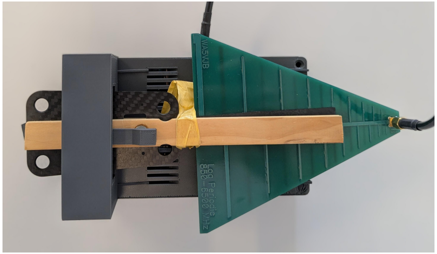
- Reducing metadata, or randomizing it, could impede tracking through upcoming attacks but could also impede devices' usability;

- Upcoming research should consider helpful applications of those de-anonymization techniques, e.g. against unwanted tracking devices, as much as their other use cases.

Upcoming research

We're currently planning to try metadata-related attacks, along with existing ones, in crowded environments.

It will allow us to measure both the effectiveness of those attacks and of countermeasures.



Picture of one of the tracking devices.

Questions?