A tutorial on privacy, RCM and its implications in WLAN.

Journal:	*IEEE Communications Surveys and Tutorials*
Manuscript ID	COMST-00137-2023
Type of Manuscript:	Tutorial
Date Submitted by the Author:	20-Mar-2023
Complete List of Authors:	Ficara, Domenico; CISCO, CH-1024 Garroppo, Rosario; University of Pisa, Dipartimento di Ingegneria dell'Informazione Henry, Jerome; CISCO, Cisco’s campus in Research Triangle Park (RTP)
Keywords:	802.11bh, 802.11bi, Randomized And Changing MAC address (RCM), Privacy

A tutorial on privacy, RCM and its implications in WLAN

Domenico Ficara, and Rosario G. Garroppo, and Jerome Henry

***Abstract*—MAC address randomization has been widely enabled in 802.11 for about a decade now. Due to the proliferation of wireless devices and consequently privacy becoming a major concern, in the recent couple of years, more aggressive randomization approaches have been implemented and deployed by mobile device vendors, urging the IEEE to formalize Randomized And Changing MAC address (RCM) operations through 802.11aq Task Group. RCM generates problems for services and procedures of the first two levels of the protocol stack that use the MAC address as an identifier of the host. To analyze and fix these problems, IEEE set up the 802.11bh Task Group, while at the IETF, the MAC Address Device Identification for Network and Application Services (MADINAS) working group has examined the consequences of RCM on the upper layer services. An additional task group 802.11bi has started at the same time with the focus on data privacy protection. This tutorial describes the reasons for RCM, its implementation and its evolution in these years, pointing out its impact on network procedures and services. Then, the tutorial covers the latest progresses in 802.11bh, 802.11bi and MADINAS.**

***Index Terms*—802.11bh, 802.11bi, RCM.**

I. Introduction

The term “privacy” is multivalent. It is used in many contexts, and each of them defines *privacy* in ways that may be domain-specific (e.g. regulatory, social anthropology, etc.) Therefore, any work on privacy is guaranteed to cause endless discussions about its intended scope, until a bounded and clear definition of the term is agreed upon for the target domain.

The world of network communications does not escape this issue. Luckily, as the concerns for privacy started permeating the public discourse during the 2000s and 2010s, several organisations, including ISO [1] and the IEEE [2] introduced a set of definitions that helped refine the notion of privacy in a networking context. Both define privacy as *“the fair and authorized” “processing”* of Personally Identifiable Information (PII). Formally, PII is any data that directly or indirectly identifies an individual or from which the identity or contact information of an individual can be derived. In other words, PII can be a direct identifier (your name, for example sent in the clear over the network) or a value that is not directly an identifier for a person, but becomes that identifier if it can be tied back to a person. For example, the MAC address of your personal tablet, if it is unique, does not directly carry your name, but if an observer can make the association between that MAC address and the owner of the tablet, then seeing that MAC address immediately reveals that you are in the room. Thus a MAC address can also be PII, although it is not the only issue. Other elements transmitted by a device may be PII. IEEE [2] also considers Personally Correlated Information (PCI), that is data gathered about an identified person or small group thereof, by observing activities (e.g., communications) or events associated with those people. For example, if your tablet is the only one in the room sending an ordered sequence of voice packets (whose structure can be observed and understood), it does not matter if your tablet has a new MAC address for every packet: the mere observation of the sequence of voice packets is sufficient to deduce that the source is still your tablet, and tie the activity back to you. PCI is typically found within the mechanics of the transmission, lurking in information or capabilities elements, scrambler seeds, location or timing of beacons, that repeat in predictable manner or display a structure unique enough that allows over time a sufficient narrowing down of the source such that a single device (or group of devices) can be identified.

The exposure of privacy depends very much on ’who’ is behind a device communicating over a network, and [2] distinguishes two types:

- **Personal Device :** a device associated with a user or a group of users (such as a family unit) e.g. your phone, your home Wi-Fi access point (AP).
- **Shared service device:** a device used by a group of people large enough that identification of the device does not easily allow identification of individual users or groups of users, e.g. a work tablet in a warehouse, an enterprise distribution switch, a Wi-Fi AP in a large airport.

Privacy is naturally concerned with the first type of devices, and in particular when personal device communications can be observed. Such observation is possible when the personal device is connected to a wired network accessible to the observer. It becomes trivial if the communication uses radio waves. In that case, any eavesdropper in range of the transmitted signal can attempt to find PII or PCI within the transmission.

This property became a growing concern among personal device vendors, as smartphones and tablets slowly started to dominate the wireless landscape at the end of the 2000s, and as privacy regulations started to appear all over the planet. The vendors efforts focused on suppressing obvious

This work was partially supported by the Italian Ministry of Education and Research (MIUR) in the framework of the FoReLab project (Departments of Excellence).

D. Ficara is with CISCO, CH-1024, Lausanne, Switzerland (e-mail: dficara@cisco.com).

R. G. Garroppo is with the University of Pisa, 56126 Pisa, Italy (e-mail: rosario.garroppo@unipi.it).

J. Henry is with CISCO, 27560 Research Triangle park, USA (e-mail: jerhenry@cisco.com).

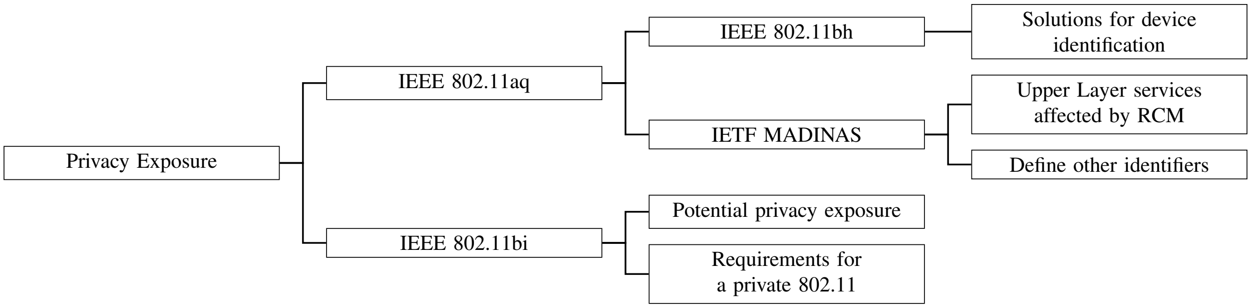

Privacy Exposure
IEEE 802.11aq
IEEE 802.11bi
IEEE 802.11bh
IETF MADINAS
Solutions for device identification
Upper Layer services affected by RCM
Define other identifiers
Potential privacy exposure
Requirements for a private 802.11

Fig. 1. Evolution of IEEE Task Group (TG) and IETF Working Group (WG) related to privacy concerns in WLANs.

PCI, before several of them decided that implementing a Randomized and Changing MAC address (RCM) scheme was also a good way to suppress at least one key PII. However, this change is not innocuous, as it affects the ability of other networking elements to know which device is the intended source or destination of this or that message, with the risk that communication would simply break. Thus different groups, in the IETF and the IEEE, set to provide rules and guidance to achieve the contradicting goals of improving privacy in wireless communications, while at the same time ensuring that device identification would continue in a manner that would allow communication stability.

A. Contribution and Organization

It is worth noting that the IEEE TG 802.11bh and 802.11bi, as well as the IETF WG MADINAS, have recently begun their work. They have produced draft documents that describe the problems and issues related to privacy concerns in WiFi networks (802.11bi [3]) and identify a set of layer 2 services (802.11bh [4]) and upper layer services (MADINAS [5]) that need to be revised to address the issues posed by the activation of RCM. Although the development process is not yet complete and many open issues need to be addressed before finalization, the presentation of the set of WiFi privacy issues discussed in the IEEE TG and IETF WG can encourage the research community to study solutions even if they are not directly involved in the standardization process. The approaches discussed within the 802.11bh TG are summarized and classified to provide the reader with a general understanding of the discussion in this group. We believe that now is an appropriate time to provide an update on the current state of WiFi privacy issues and explore potential solutions and approaches being considered. Our goal is to present this information in a way that is easily understood by the broader wireless networking community.

In this tutorial paper, we leverage our direct participation in the 802.11bi and MADINAS activities to achieve three goals:

- Provide a snapshot of the major privacy issues in WLAN that have been discussed in standardization work and/or presented by the research community.
- Present issues or caveats related to the impact of RCM on services and procedures, which may require further support from the research community, such as further ideas to overcome the issues identified by the standardization work.
- Describe the key research challenges under consideration in the standardization groups.

To the best of the authors' knowledge, this work is the first tutorial to present the privacy issues of WLAN, taking into account the standardization work aimed at solving the negative impact of the most commonly used solution, RCM, on services and procedures at the WiFi layer and upper layer. Furthermore, this tutorial summarizes the standardization work aimed at defining the concerns of WiFi for enhancing privacy and defining the framework under consideration in the TG 802.11bi.

This paper summarizes these efforts and identifies related research challenges, following the structure shown in Figure 1, which summarizes the recent IEEE Task Group (TG) and IETF Working Group (WG) efforts to address privacy concerns in WLANs. The paper is structured as follows: Section II presents how privacy is exposed on wireless (in particular 802.11) communications, Section III presents and classifies the topics related to RCM discussing selected related work, Section IV details the RCM solution and the boundaries posed by the 802.11aq amendment, Section V details the effort of the 802.11bh group to 'fix' the issues introduced by RCM, Section VI underlines the work of 802.11bi to increase the privacy in 802.11 communications, Section VII presents the work of the MADINAS group in IETF, targeting the issue of privacy and RCM for upper layer services, and Section VIII concludes this paper.

II. Privacy exposure in WLANs

IEEE 802.11 was designed from the very beginning with the intent to facilitate communication among wireless devices that might support different options and capabilities, thus making the protocol highly susceptible to privacy attacks by mere fingerprinting through capability exposure.

To understand this problem of privacy exposures in WLANs, a good starting point is the definition of its targets: any element which exposure provides PII or PCI.

Several actors have shown interests in extracting PII from IEEE 802 frames with different objectives, such as:

- Surveillance. Attackers may conduct passive tracking by observing the location and connection time of a target on a network. If PII can be collected from several network connections, it is known as Pervasive Surveillance.
- Probing. Packets may be directed towards a target or its recipient in order to coerce the disclosure of PII.

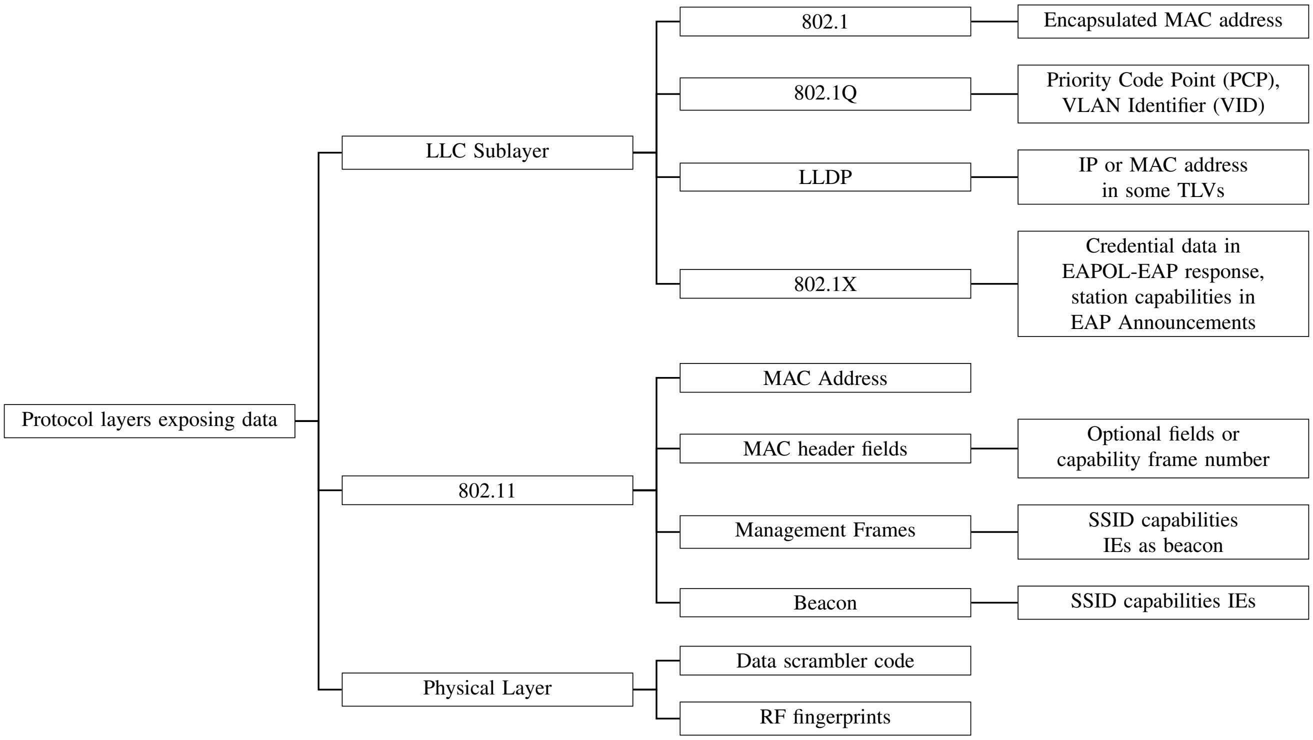

Protocol layers exposing data
LLC Sublayer
802.11
Physical Layer
802.1
802.1Q
LLDP
802.1X
MAC Address
MAC header fields
Management Frames
Beacon
Data scrambler code
RF fingerprints
Encapsulated MAC address
Priority Code Point (PCP), VLAN Identifier (VID)
IP or MAC address in some TLVs
Credential data in EAPOL-EAP response, station capabilities in EAP Announcements
Optional fields or capability frame number
SSID capabilities IEs as beacon
SSID capabilities IEs

Fig. 2. Privacy Exposure in WLANs.

- Modification. Frames exchanged with a target may be modified to induce the exposure of PII.

An obvious target for privacy exposure is the MAC address of a personal device, either as source or destination address (SA or DA hereafter). A key point is that when the target MAC address is an universal stable address (i.e.: that never changes), it is possible to correlate it across multiple networks in terms of time and location. This includes situations where the MAC address is visible as either the SA or DA in a frame, but also when it is present in another network header, such as an encapsulated Ethernet header, IEEE 802.1Q I-TAG, or an IPv6 header.

Correlation on any target MAC address can be used to track its locationwhen it is mobile and to collect frames sent to and from it. Further analysis may involve identifying MAC addresses linked to individuals or determining which individual a MAC address is linked to. Correlation of a target MAC address does not always pose a threat to privacy. An individual may choose to allow the correlation for their own benefit, for example, by explicitly opting in to the correlation after being offered special treatment by the network owner. However, unauthorized correlation may be viewed as an attack.

MAC address is central to this paper as it is the main target of RCM. However, there are other PII and PCI exposures in communications using IEEE 802 Standards, either through 802.11 or, as the frame transits over a wired network, through other protocols. Figure 2 summarizes the main PII/PCI exposure related to the IEEE 802 Standards domain, which covers the two lowest layers of the protocol stack.

1) 802.1: The IEEE 802.1 standard defines a set of functions that can extend the frame header with information that can be used by attackers for acquiring PII. For instance, encapsulated MAC address are used by some IEEE 802.1 protocols, such as IEEE 802.1Q Congestion Notification Message PDU, IEEE 802.1AB Chassis ID, IEEE 802.1AB Port ID, IEEE 802.1AE SecTag. Moreover, a bridge located at a network edge providing access to a group of personally-owned devices, such as a residential gateway, can expose some correlated information. Its MAC address, which needs to be a universal address, used to locate host addresses (e.g., those embedded in a Stream Identifier), is PCI.

2) 802.1Q: In several IEEE 802.1Q protocol elements, the Priority Code Point (PCP) is used to mark frames that should be prioritized because they have special latency requirements (such as voice or video frames). In a similar way, the VLAN Identifier (VID) defined in IEEE 802.1Q is often used to segregate traffic, such as to differentiate that addressed to different organizations or individuals with different roles in the organization. Target classes (such as endpoints transmitting voice or video traffic, Organization, Role, etc.) can be identified based on the PCP and VID either directly using well-known configuration (PCP) or after correlation analysis necessary to map the VID with the organizations or individuals with different roles in the organization.

3) LLDP: Link Layer Discovery Protocol (LLDP) frames (IEEE 802.1AB) deliver information about a station as a set of TLVs (Type-Length-Value). TLVs can contain network address (IP or MAC), system name, system capabilities which can be used by an attacker to identify a target by address or by domain name, to give PCI by identifying a class of target (e.g., Telephone, DOCSIS cable device). Furthermore, Organizationally Specific TLVs may be defined to contain PII or PCI.

4) 802.1X: The Port-Based Network Access Control (IEEE 802.1X) defines a set of message types that can be used, some of these can contain PII, such as Extensible Authentication Protocol (EAP) response message of EAPOL-EAP that may contain login credentials (user, host or both), or EAPOL Announcements that include capabilities for the station. A passive adversary between the target and EAP authenticator can observe any information that an EAP method passes without confidentiality protection or observe Announcement data, and identify or deduce the class of target. Active adversary between the target and EAP authenticator may be able to spoof a legitimate respondent in an EAP method to the point where the target presents its identity (e.g., the subject name in a client certificate).

5) 802.11: Other than in the Address fields, the IEEE 802.11 frame contains a wealth of information that helps correlating a target with a personal device.

For example, any client station (STA) attempting to connect to an access point needs to clarify which options it supports (such as fragmentation, power saving and management behavior, etc.), through optional fields or capability elements. In the 4000+ pages of the 802.11 Standard, there are hundreds of such options, which support depend on the client Wi-FI driver, its generation, the operating system above and the STA vendor choices. The simple observation of unusual elements (compared to the other STAs in the same Wi-Fi hotspot) can be used to identify a specific device.

These options are visible in the STA association requests and other frames of "management" type, whose role is to allow for connection to happen and be maintained. In general, because of their crucial role in making connection possible, 802.11 management frames contain a lot of information. For instance, APs send beacon frames to advertise their wireless network. When the AP is associated to some individuals (e.g.: in a house or small store) beacon data can become PII, uniquely identifying the AP and thus tying any traffic through that AP to the network owner. An adversary can use the beacon Capability field, SSID, RSN IE, EDCA Parameter Set, QOS Capability IE, QoS Traffic Capability to fingerprint the AP. An adversary can actively also emit a Transmit Power Control (TPC) Request and observe the related response from the AP in the subsequent beacon to fingerprint the AP and its location. By examining the Association ID (AID) within the Traffic Indication Map (TIM) included in a beacon frame, it is possible to determine which STAs have data buffered by the AP. This data can also be used by a potential attacker to identify and distinguish the connected clients. Furthermore, well-known attacks (evil-twin) are based on advertising (by means of beacons) an SSID identical to that of another system, causing the victim STA to unknowingly attempt to connect to that rogue AP and allowing the attacker to be able to collect PII from the STA's frames.

Other management frames, such as Probe Request/Response, Re-/Association Request/Response, De-/Authentication, Action, Disassociation and etc. can also be exploited to gather information and PII or PCI. For example, the Probe Response structure is very similar to that of the beacon. The Probe Request can contain the SSID that the client is looking for, which may expose PII. The Disassociation frame contains a Reason code, one or several vendor-specific elements, and optionally a Management MIC Element when Protected Management Frame (PMF) is enabled and the frame is addressed to a group. All these elements can be used to uniquely identify the sender and its position in the infrastructure.

6) Physical layer: Passive radio-frequency analysis can be used at the physical layer to detect the source network interface card (NIC) of an IEEE 802.11 frame. This method relies on exploiting small flaws or peculiarities in the hardware of the transmitter that occur during manufacturing and are present in all NICs, despite their overall similarity. These flaws, which are unique to each transmitter, result in artefacts in the emitted signals and represent a fingerprint of the device [6] [7]. Another feature that might give allow for device fingerprinting is the seed of the scrambler. The purpose of scrambling in IEEE 802.11 is to improve the performance of the physical layer. Scrambling involves xor'ing the binary payload of a frame with a pseudo random sequence to produce an uncorrelated binary sequence with uniformly distributed bits, increasing the information content. However, the seeding of the scrambler can be a potential attack vector, as correlating scrambler state seeds across multiple messages can allow for the re-identification of devices and drivers, bypassing all current privacy-preserving mechanisms [8].

III. Topics related to RCM

This section presents a selection of the related work to RCM, organized taking into account the classification of the topics shown in Figure 3.

The possibility of fingerprinting personal devices has been demonstrated many times. In fact, even without malicious intent, several companies have started tracking personal devices to monitor shopper behavior in stores [9] [10], as well as tracking people in urban areas [11] [12]. Naturally, this practice raises serious privacy concerns as the unique mobility patterns of individuals can be easily identified using just four spatio-temporal points [13]. These issues are precisely the ones that have led to the introduction of MAC randomization, first by device manufacturers and then by the IEEE.

Maltoni et al. [14] present a low-cost and low-power people counter based on the Espressif ESP8266 board, analyzing the overall cost of the solution. The device can collect MAC addresses from Wi-Fi packets, bypassing MAC address randomization, which is demonstrated through practical experiments. The study highlights that as IoT devices become more affordable, even a single person could set up a personal people counting system for malicious purposes in urban areas or indoor environments.

Recently, Abedi et al. [15] demonstrated the security and privacy vulnerabilities of the 802.11 Wi-Fi protocol by showing the ability to reveal the accurate location of Wi-Fi devices in an indoor environment without any cooperation, instantaneously, and surreptitiously. The authors exploited the power-saving mechanism in 802.11 to identify all devices

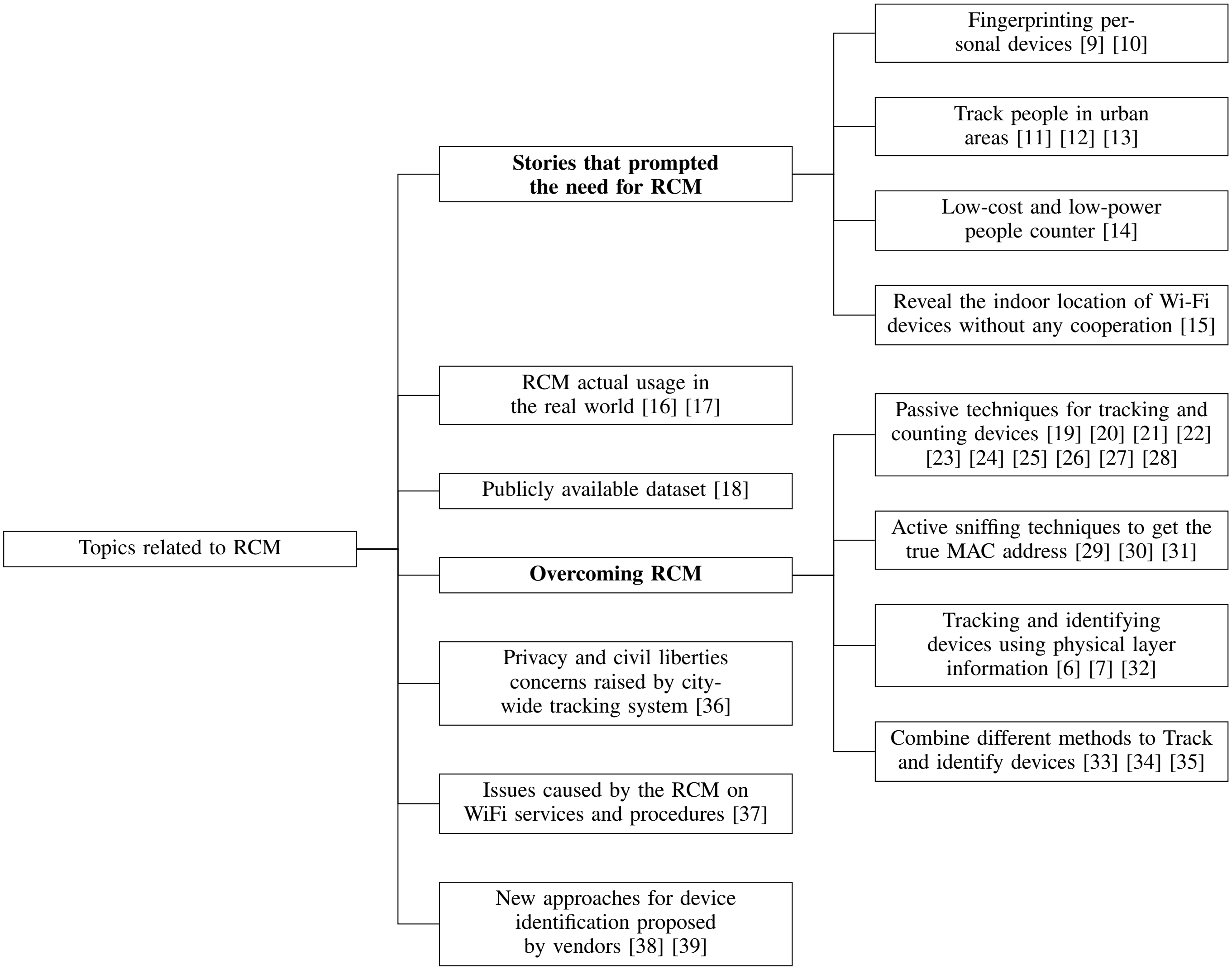

Topics related to RCM
Stories that prompted the need for RCM
Fingerprinting personal devices [9] [10]
Track people in urban areas [11] [12] [13]
Low-cost and low-power people counter [14]
Reveal the indoor location of Wi-Fi devices without any cooperation [15]
RCM actual usage in the real world [16] [17]
Publicly available dataset [18]
Overcoming RCM
Passive techniques for tracking and counting devices [19] [20] [21] [22] [23] [24] [25] [26] [27] [28]
Active sniffing techniques to get the true MAC address [29] [30] [31]
Tracking and identifying devices using physical layer information [6] [7] [32]
Combine different methods to Track and identify devices [33] [34] [35]
Privacy and civil liberties concerns raised by city-wide tracking system [36]
Issues caused by the RCM on WiFi services and procedures [37]
New approaches for device identification proposed by vendors [38] [39]

Fig. 3. Topics related to RCM.

in the network and use targeted packets to perform location estimation. An eavesdropper can use this information for malicious purposes, such as tracking individuals or finding out the occupancy of rooms in a hotel.

Musa et al. [11] proposed a monitoring system and trajectory estimation algorithm, which were reasonably accurate. However, they noticed that many MAC addresses were unlisted and that MAC randomization was not yet widely used. Despite the introduction of MAC randomization by some vendors, subsequent works showed that it was still possible to distinguish and track modern devices. The following works show that MAC randomization is a factor to consider in tracking modern devices, but it does not render the tracking technique obsolete [26] [29].

Since the implementation of RCM, various studies have been conducted to evaluate its actual usage in the real world. The first finding is that it is not widely adopted by users [16], [17]. However, several de-randomization methods have been proposed to cluster frames from the same device observed in a given time period. The methods can be classified into passive sniffing, active sniffing and physical layer analysis.

As concerns passive sniffing, Matte et al. [19] proposed a method to tackle the randomization process by using an algorithm that takes a set of Probe Request frames as input and associates them with different devices. They use an algorithm to create signatures based on inter-frame time, inter-burst time, and frame sending frequency, and compare similarities using the Franklin distance. MAC addresses with a similarity distance below a certain threshold are aggregated. Scheuner et al. [20] present Probr, an open-source system for monitoring room utilization, device localization, person tracking, and device statistics, through monitoring Wi-Fi traffic. Nitti et al. [21] proposed a method to determine if two Probe Request frames came from the same device. A score is calculated based on the difference in arrival time and sequence numbers. These differences are computed for all possible pairs of MAC addresses with the same Information Element (IE) IDs, regardless of the length or content of the IDs. Another system for analyzing urban mobility utilizes a machine learning algorithm for detecting people's flow while preserving privacy [22]. Similarly, in [23], Gebru et ali. propose two methodologies for people counting and mobility detection. However, recent privacy-preserving approaches randomize the sequence number of Probe Request entirely, rendering the previously mentioned methods less reliable. In spite of the increased level of randomization, an efficient crowd monitoring system was developed by the authors of [24], which is based on passive detection of Probe Requests. The algorithm uses a statistical estimator to count devices by measuring the rate of Probe Requests received in 10 ms. Another solution for crowd monitoring is presented in [25], which uses the SSID information of the preferred WiFi access points included in Probe Request, in combination with the information of existing WiFi access points, to determine the daily number of visitors at different locations through post-processing. Another study relying on Probe Request fingerprinting is [26]. The authors proposed a device tracking algorithm using IE IDs fingerprinting. The algorithm first groups Probe Requests into clusters based on their IE IDs, and then attempts to distinguish devices in the same cluster by analyzing the predictable behavior of the sequence number. Two Probe Requests are considered from the same device if their arrival time difference is less than 500 seconds and the sequence number difference is less than 64. In [27], M. Ribeiro et al. introduced a counting and tracking method that uses automatic classification techniques. The study was based on 4 years of probe request data and 7 unsupervised classification algorithms were used to calculate the average accuracy. The study found that over time, the use of factory MAC addresses decreased from over 50% to around $5 - 10\%$. The same method was applied in [28] to monitor passengers using public transportation, resulting in origin-destination matrices and information about which bus stops were the most used. However, the analysis was done without considering random MAC addresses and the solution was not adequate for estimating the number of passengers. Nevertheless, the information provided insight into unusual situations and was useful for communication among stakeholders. As for sniffing, Vanhoef et al. in [26] analyzed 8 million probe requests using active techniques and found around 170000 MAC addresses. They found the best method was through the use of WiFi Protected Setup (WPS) parameters on devices that support it, where the device connection provides a Universally Unique Identifier-Enrollee (UUID-E) parameter. This parameter was linked directly to the device's factory MAC address, enabling the accurate tracing of the real MAC and reducing device counting error.

Martin et al. [29] analyze various techniques that can be used on a large scale to be able to trace random MAC addresses to a single device. In particular, active sniffing methods exploit various vulnerabilities and made attacks such as KARMA attack [30] (creation of fake Access Point from the list of probe BSSIDs) and RTS/CTS attack [31] in order to obtain the true MAC address during the negotiation of the connection with an AP. The attacker must possess knowledge of the device's SSID in order to utilize this method.

Another methods class leverages the unique characteristics of different device drivers to identify the type of device that sent the frames, allowing them to differentiate between devices even if their MAC addresses are randomized. For example, V. Brik et al. [6] were able to accurately classify the device drivers with high accuracy, using a combination of machine learning algorithms and signal processing techniques. The results of this study provide a new approach for tracking and identifying devices in a wireless network, even if the devices use randomized MAC addresses. Recent works, such as [7] [32], provide a survey on approaches for radio frequency fingerprinting and identity recognition via data fusion and feature learning, respectively.

Recently, Uras et al. [33] propose a new MAC address derandomization algorithm that groups Probe Requests generated by the same device. The algorithm combines the features that have been previously considered in isolations, such as the content and the length of optional fields in sent frames and the rate at which the frames are numbered over time. The frames are associated with the same device using density-based clustering algorithms, using these features [34]. The performance evaluation of the proposed clustering schemes is evaluated using a publicly available dataset [18]. In [35], He at ali. propose Cappuccino, a novel privacy-preserving approach that captures the association of probe requests under MAC address randomization. Cappuccino is a system that estimates pairwise frame correlation and associates frames over time. It uses a self-supervised estimator to determine frame correlation, considering multiple modalities such as information elements, sequence number, and received signal strength.

Spiess [36] discusses the impact of tracking systems in Seattle, where WiFi/Bluetooth MAC address collection is used for real-time travel calculation in traffic engineering. However, the deployment of this technology raises privacy and civil liberties concerns, with five broad categories of gaps found by the public: missing information, technical implementation, contractual, data management, and the bigger picture. Alternative technologies, MAC address randomization, challenges with MAC address anonymization, and mobile device model identification are also discussed. At the time of the article's submission (Jan. 15, 2021), the technology was deployed in Seattle but the Seattle City Council had not yet approved its use.

To the best of the authors' knowledge, no scientific work has been published to address the issues caused by the RCM on WiFi services and procedures. Only vendor-written blogs and white papers are available on the topic. For example, for a set of important services and procedures, in [37] the author discusses if they will still function if the MAC address is altered or changes regularly (e.g. daily). He suggests to start making adjustments now to prevent any unexpected issues in the future.

Commercial solutions offer a new approach to device identification. For example, the LEVL platform [38] provides a unique identifier (denoted as LEVL-ID) for each device on the network. LEVL-ID is derived without requiring any input from the user and is completely zero-touch. It does not compromise user privacy as it doesn't rely on user-data or PII for identification. LEVL-ID is a privacy-friendly alternative to the MAC address because is a derived identifier, not stored on the device, and relies on device characteristics only known to the internal network. The DUID (Device Unique ID) [39] is another example of a proprietary solution for providing a unique identity to devices using RCM.

IV. The RCM solution

The concern that a MAC address would provide location and activity information about a person grew with the explosion of the smartphone market at the end of the 2000s. In the early 2010s, several workshops [40] centered around Global Surveillance and Privacy topics noted this exposure and underlined how it was an unintended consequence of the MAC address design in IEEE Standards. Discussions in IEEE 802.1 (architecture group) and 802.11 (WLAN group) concluded that the MAC address was expected to be unique and stable for the duration of a network session, so as to maintain continuity of the network connection and other related services. However, outside of that session, no expectation was to be made on a client station identity, as the station was not requesting any service from the network. This last provision was not clear in the Standards texts, and the IEEE 802.11 Working Group decided to insert this clarification in the 802.11aq amendment [41], that was under development at the time and focused on pre-association discovery exchanges. Meanwhile in 2014, following a mechanism available on Linux and BSD operating systems, Apple introduced the use of randomized MAC addresses in iOS 8 (for recent iPhones and iPads) for unassociated Wi-Fi message discovery (probe requests). Other vendors also implemented the same model as they developed new firmwares for 802.11 chipsets.

The practice received different names (many vendors tending to give a different name, some of them hoping that their choice would become the de facto naming standard, allowing them to claim the invention of the practice), but the most common term in standard organisations soon became Randomized and Changing MAC address (RCM).

A. RCM in IEEE 802.11aq

The text in IEEE 802.11aq was the result of intense conversations between supporters of privacy-first approaches, pushing for the Standard to liberally and explicitly allow the STA to change its MAC address at will, and the supporters of network stability, who felt that the 4000+ pages of the Standard should be carefully combed-through before allowing schemes that might break network stability or backward compatibility. IEEE 802.11aq recognizes (clause 4.5.4.10) that the MAC address, along with other parameters (such as frame sequence numbers, scrambler data or information in the probe request) may allow unique identification of a STA.

To limit this risk, the Standard naturally recommends using non-consecutive sequence numbers and scrambler data, and avoiding querying for specific SSIDs in Probe Requests. In addition (clause 12.2.10), 802.11aq recommends that a STA should periodically change its MAC address to a random value while not associated to an AP. The random MAC should follow the structure of locally administered addresses defined in IEEE 802-2014 [42] and 802c-2017 [43]. Figure 4 shows the structure of the MAC address pointing out the least and the second least significant bits of the initial octet of a MAC address. These bits are indicated as `M` and `X` bit respectively, and give information on the kind of address, i.e. unicast or multicast and globally unique or locally administrated. The set of local unicast addresses is in the following form: `x2-xx-xx-xx-xx-xx`, `x6-xx-xx-xx-xx-xx`, `xA-xx-xx-xx-xx-xx` and `xE-xx-xx-xx-xx-xx`.

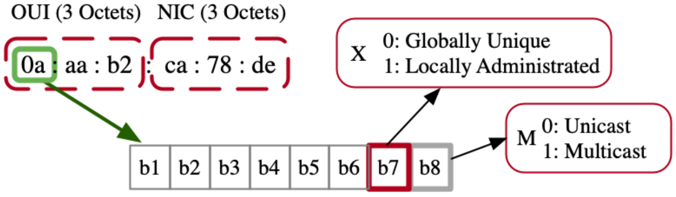


Fig. 4. MAC address format.

In 802.11aq, changing "periodically" the MAC address does not mean “at regularly spaced intervals”, but simply that multiple changes are expected over time (i.e, the STA does not keep a randomized MAC address for long). However, a STA should not change its MAC address as soon as it enters an exchange that causes a state to be generated on an AP. Any exchange where the STA expects a reply from the AP requires a stable MAC address throughout the whole dialog. The stable MAC address can be locally-generated (RCM). An association to an AP is typical of this type of scenario, but pre-association exchanges are targeted too. For example, the STA may query an AP (pre-association) to retrieve a list of services accessible through that AP (IEEE 802.11aq describes many such services). The exchange is carried over the Access Network Query Protocol (ANQP), where the AP can return partial or summarized responses with a flag indicating that additional information can be provided. The STA can then ask for more details on the elements of interest. Throughout this entire exchange, the STA should keep the same MAC address.

The case of roaming across APs is also considered, and 802.11aq clarifies that the STA should keep the same MAC address for the duration of its connection to the entire Extended Service Set (ESS), i.e. including while moving between APs with the same SSID within the same venue and during the same session. The term ESS itself is clearly defined and the 802.11 Standard, but this clarity creates, on purpose, an ambiguity. An ESS is a set of one or more interconnected basic service sets (BSSs, the area covered by a single AP). To the Logical Link Control (LLC) layer, ESS appears as a single BSS at any one of the STAs associated with one of the BSSs. In 802.11, the LLC is the Layer above the MAC Layer (and below the IP Layer). Therefore, the ESS appears as soon as the STA associates to an AP, and exists as long as the STA maintains its connection to that AP, or moves (its connection, and thus its PHY and MAC Layers) between APs in range of one another (thus without ever terminating its connection). The STA leaves the ESS when it disassociates. This definition means that, as soon as the STA disassociates, it is no longer connected to a BSS or an ESS. It can therefore change its MAC address and start a new connection.

Therefore, with 802.11aq, the STA has the choice to use its real (burned in) MAC address for all associations to all

SSIDs. The STA can also use burned in and RCM MAC addresses, and tie a specific MAC address to a given SSID (understanding an ESS as all APs advertising that SSID), using different MAC addresses for different SSIDs (which is what most station vendors have implemented). But the STA can also consider its location, and decide that two non-contiguous APs offering the same SSID are not a single ESS. The STA can therefore generate different MAC addresses for different sites and the same SSID. For example, this possibility means that a smartphone could use one MAC address for the SSID of a coffee shop, and then another MAC address for the same SSID, but in another coffee shop of the same chain. On paper, this possibility supposes that the STA (in the operating system layers) has location awareness (to distinguish two shops of the same chain announcing the same SSID). However, this requirement is practically not true. As the association stops when the STA disassociates, and as the ESS only appears once the STA is associated, the STA can also use different MAC addresses for the same SSID and different sessions. Practically, this means that your smartphone could use a MAC address for your home SSID as you read the news during breakfast, disconnect as you take the smartphone to the office, then generate and use a new MAC address for your home SSID as you come back from work in the evening. More aggressive behaviors are also conceptually possible. The STA can observe that the network connection is idle (no active traversing the Wi-Fi card for "some" arbitrary duration), disassociate (thus leaving the BSS and the ESS), change its MAC address, and immediately start a new association to the same AP. 802.11aq does not provide a view on these scenarios, limiting the scope of its normative text to the scope detailed above.

B. Impact of RCM on Wi-Fi procedures and services

Most operating systems implement a MAC-to-SSID mapping structure, where the STA generates a new MAC address for any association to a new SSID, then keeps that MAC address for any subsequent connections to that same SSID, irrespective of location considerations [40]. However, it is clear to all actors in the field that this mode only provides limited privacy protection. As soon as some PII is identified in a given SSID, using the same MAC address means that the identified user can be tracked each time their device connects again to that same SSID. This RCM scheme only prevents the tracking of the user as the device moves across SSIDs (thus using different MAC addresses). Therefore, some vendors are attempting to limit the lifetime of that MAC address, by changing it after a while. Apple iOS release 16 changes the MAC if the STA has not associated to the SSID for more than 6 weeks [44]. In Android 12 or higher, the Wi-Fi module can use non-persistent randomization type for some networks, re-randomizing the MAC address at the start of every connection or using the existing randomized MAC address. Re-randomization can occur when the DHCP lease duration has expired and more than 4 hours have passed since the last disconnection, or if the current randomized MAC for the network profile is more than 24 hours old. In this case, re-randomization only happens at the start of a connection and the Wi-Fi won't actively disconnect to re-randomize the MAC. If neither of these situations apply, the previously randomized MAC address is used to connect to the network [45]. Microsoft Windows 10 flexibly allows the user to turn on or turn off the RCM for each SSID, as well as to configure a daily change of MAC address [46].

However, these aggressive schemes may have detrimental effects on the network infrastructure and thus, by reactance, on the STA user experience. The assumption that a STA could be identified uniquely via its MAC address had contributed to the development of mechanisms that break when the MAC changes.

Figure 5 provides an overview of these issues, while a detailed discussion follows in the subsequent paragraphs.

- Infrastructure-based location services: the signal from a STA (primarily from Probe Requests) had long been used to deduce the position of a given STA [11]. This side effect of the 802.11 discovery mechanism was used to localize assets in enterprise environments. In public settings (e.g., a mall), location client STAs allowed the venue owner to collect useful statistics as mentioned above (dwell time in different areas, typical path across a store, etc.), and modify the layout accordingly for business (advertisement and promoted goods positions) or safety purposes (escape routes signaling and locations) [9] [10]. For users opting-in to targeted marketing, advertisements or coupons could be sent to the phone (on an app) when the user would reach critical points in a store, even if the STA is not associated to the Wi-Fi network. Without stability in the MAC address during the discovery exchanges, venues can no longer implement these mechanisms, which may not be an issue for unwanted customer tracking, but may be an issue for opt-ins and enterprise asset tracking scenarios.
- Automated re-authentication when a captive portal is used: some venues (e.g., hotels) charge for premium Wi-Fi services and used the MAC address to identify an already authorized device. If the MAC address changes (e.g., every day), then users end up being charged multiple times for services that they bought once for their entire stay.
- Coverage hole detection: in enterprise environments, floor topology changes over the lifetime of a Wi-Fi network. Walls are added or removed, desks are moved, and holes in the covered areas appear. These holes are identified when an STA is detected to scan (send Probe Request messages) at the edge of a Wi-Fi hotspot, then fails to appear as associated at acceptable signal levels on any AP. If the probing MAC address is RCM, then coverage holes are no longer found, and the quality of the coverage suffers. This may affect operations, as Wi-Fi has commonly become business-critical.
- MAC authentication: systems that use MAC addresses to identify devices (for example to assign policies, or band steering and client steering for optimizing client connectivity in a multiple AP environment) can no longer operate if the MAC address is randomized and changes over time. When devices have very simple network stacks

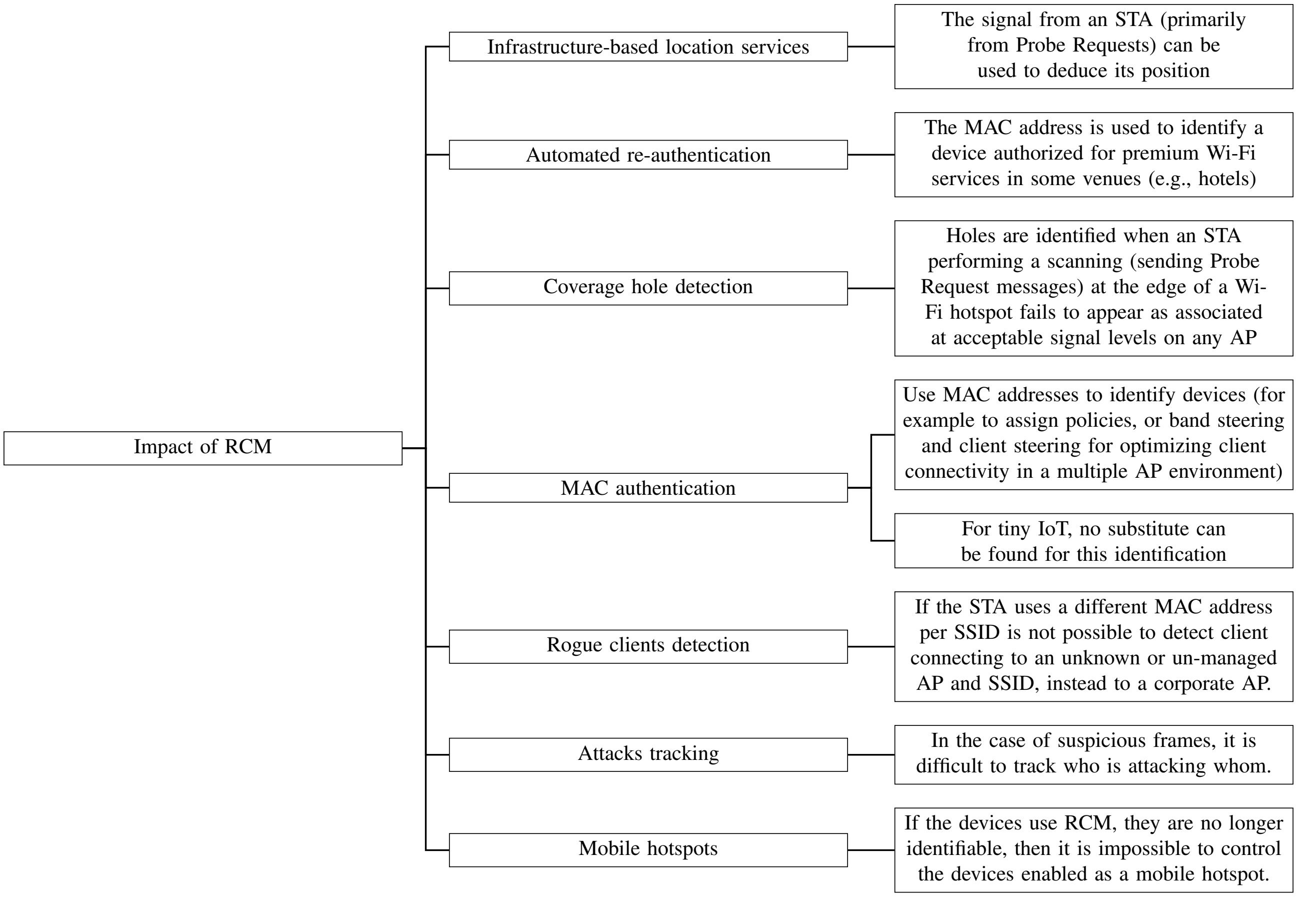

Impact of RCM
Infrastructure-based location services
The signal from an STA (primarily from Probe Requests) can be used to deduce its position
Automated re-authentication
The MAC address is used to identify a device authorized for premium Wi-Fi services in some venues (e.g., hotels)
Coverage hole detection
Holes are identified when an STA performing a scanning (sending Probe Request messages) at the edge of a Wi-Fi hotspot fails to appear as associated at acceptable signal levels on any AP
MAC authentication
Use MAC addresses to identify devices (for example to assign policies, or band steering and client steering for optimizing client connectivity in a multiple AP environment)
For tiny IoT, no substitute can be found for this identification
Rogue clients detection
If the STA uses a different MAC address per SSID is not possible to detect client connecting to an unknown or un-managed AP and SSID, instead to a corporate AP.
Attacks tracking
In the case of suspicious frames, it is difficult to track who is attacking whom.
Mobile hotspots
If the devices use RCM, they are no longer identifiable, then it is impossible to control the devices enabled as a mobile hotspot.

Fig. 5. Impact of RCM on Wi-Fi procedures and services.

(e.g., some IoT objects), no substitute can be found for this identification.

In the enterprise environment, MAC addresses were also used to detect and manage rogue actors.

- Rogue clients: a rogue client is a valid client of the enterprise Wi-Fi network, that suddenly connects to an unknown or un-managed AP and SSID instead of connecting to a corporate AP. Tracking such abnormal behavior is no longer possible if the STA uses a different MAC address per SSID.
- Attacks tracking: attackers are expected to hide their identity (including their MAC address). RCM makes the process automatic for them. When the valid clients MAC address is also randomized, it becomes difficult to track who is attacking whom, when suspicious frames are detected by Intrusion Detection Systems (IDS).
- Mobile hotspots: in some enterprise environments, users may be allowed to use one of their devices as a mobile hotspot for offering connectivity to other devices (especially in areas of known spotty corporate Wi-Fi coverage). The list of devices allowed to perform hotspot functions is usually strict. If the devices use RCM, they are no longer identifiable, imperiling the enterprise security structure.

These issues are directly related to 802.11 operations. As will be seen below, issues also appear at the upper layers. One major ambiguity in the lists above is the role of the user. RCM was intended to increase the user privacy, when the STA is a personal device. In some cases, the STA identification was done without the user consent. But in some other cases, the device is not personal (e.g., a company-provided tablet or barcode scanner), and the notion of consent and privacy becomes difficult.

In order to reflect on these ambiguities, the IEEE 802.11 group created in early 2020 the RCM Topic Interest Group (TIG). The group met for a year, and closed on two key conclusions [47]:

- RCM, as implemented by vendors and allowed by 802.11aq, causes issues. In at least 11 use cases, functions performed in 802.11 networks that were possible with a unique MAC address become difficult or impossible with RCM. Some of these use cases describe scenarios where the user may want a STA to be identified (e.g., auto-detection of a user arriving home or at a working desk from the STA MAC address, and automatic related actions, such as light or temperature adjustment; user opening a support ticket in an enterprise when Wi-Fi issues were experienced, etc.). Some other cases may or may not relate with the user preferences (e.g. a store measuring foot traffic density based on unique MAC address counts, or an airport measuring wait times [and the need to open more counters] based on unique MAC address dwell time). In some other use cases, RCM is a welcome change (e.g., limitation of global surveillance based on individual device MAC address detection over different places and times). In order to remediate the RCM issue in scenarios where the user may want and the network may need some unique STA identification, the 802.11bh group was formed, in the hope of designing a simple replacement for the unique MAC, that could operate over current 802.11 networks.

- IEEE 802.11 in general was initially designed in days where the idea did not exist that 802.11 parameters expressed by a personal device could be used to violate user privacy. As the Standard evolved over the years, the accent was put on operational efficiency and security, but not so much on privacy. As a result, the MAC address, but also many capabilities and parameters found in 802.11 frames, could be used to uniquely identify a device. There is a need to improve the privacy of 802.11 operations. To achieve that goal, the 802.11bi group was formed.

V. IEEE 802.11bh

With the introduction of RCM, general tracking of radio MAC address became unfeasible, and sniffing and eavesdropping got much less relevant. On the other hand, in order for networks to be able to provide correct level and quality of service and apply related feature set to STAs, a certain level of cooperation is needed, at least for the STA's identification. When STAs use randomized MAC addresses, identification (and related services) must come through other methods, which require an additional level of complexity. The focus of IEEE 802.11bh is around such newly introduced problems, proposing approaches for fixing them, in a context of client and network cooperation.

A number of approaches have been proposed to provide the network with an identification method for cooperating clients. The following subsections cover these proposals at high level. The interested reader can refer to [4] for more details.

A. IRM

Identifiable Random MAC (IRM) scheme proposes the use of a secure hash (i.e., HMAC-based Extract-and-Expand Key Derivation Function, HKDF [48]) of an IRM-Address (IRMA) through an IRM-Key (IRMK). The result represents the Output Keying Material (OKM), which is shared as additional IE in the association phase between STA and AP, and is obtained as OKM=HKDF(IRMA,IRMK). As shown in Figure 6, which depicts a few interactions of the n-th association and n-th 4-way handshake (HS) for IRMA, OKM is sent in the Association Request of the STA. In the case of an STA known to the ESS, IRMA prevents third-parties from tracking the STA while still allowing trusted parties to recognize it. To this aim, the IRMK is used to resolve the identity of the STA. Indeed, during the n-th 4-way HS, $IRMK_{n+1}$ is shared either by the AP (message 1, M1, of the 4-way HS) or STA (message 2). The AP stores $IRMK_{n+1}$ in its internal database. At the next $(n+1)$-th) association phase for the same STA, it computes OKM=HKDF(IRMA,$IRMK_{n+1}$) and shares it using the IRM IE. Before the association, the AP may check the stored IRMK(s) in order to determine the IRMK that, together with the IRMA, produces the OKM sent by the STA. In this manner, the AP can therefore identify the STA as being the same. As mentioned above, both AP and STA can take up the role of generating the keying material, which makes this method widely flexible. In the example, we assume that IRMK is exchanged in the 4-way HS messages, but this information can be exchanged using other approaches, such as defining new messages, e.g. IRMK Request/Response frame.

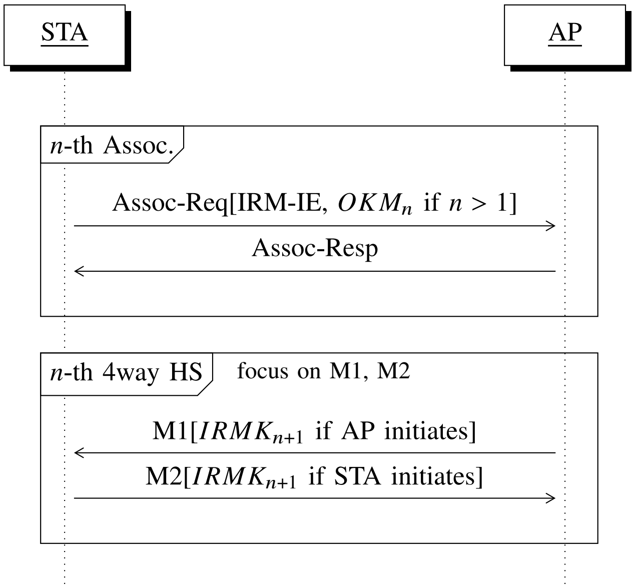

STA
AP
n-th Assoc.
Assoc-Req[IRM-IE, OKM_n if $n > 1$]
Assoc-Resp
n-th 4way HS
focus on M1, M2
M1[$IRMK_{n+1}$ if AP initiates]
M2[$IRMK_{n+1}$ if STA initiates]

Fig. 6. IRMA sequence diagram

B. MAAD

MAC Address Designation (MAAD) has evolved a bit in its proposed version within 802.11bh. It started with the unique peculiarity of being an asynchronous scheme: a STA would initiate a MAAD request asynchronously as an action frame, therefore not requiring changes to the 4-way handshake. However, in its most recent form, the AP is in charge of sharing two new MAC addresses (for the two phases of STA's relationship with AP): MAAD MAC 1 to be used by the STA as TA at next association and MAC 2 to be used for any probe in pre-association phase, if the STA desires to be recognized. Rogue AP spoofing is avoided by an additional ID that is provided by the AP together with the 2 MAAD MACs and then provided again in probe responses. A diagram for interesting parts of the pre-association, association and 4-way handshake phases is presented in Figure 7. In particular, the figure shows that in the case of n-th pre-association, the STA uses the MAC 2 assigned by MAAD, while MAC 1 is used in the n-th Association Request. In message 3 (M3) of the n-th 4-way HS, the AP communicates to STA the ID and the two MAAD MAC addresses to use in the successive iteration with the ESS.

C. Device Identifier

This scheme, currently known as Device Identifier (DID, although this is prone to renaming to titles such as Persistent

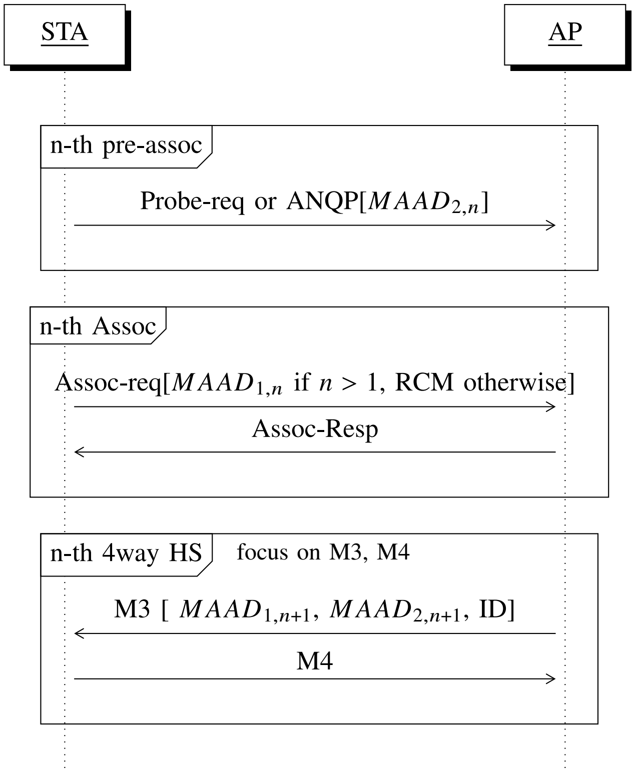

STA
AP
n-th pre-assoc
Probe-req or ANQP[$MAAD_{2,n}$]
n-th Assoc
Assoc-req[$MAAD_{1,n}$ if $n > 1$, RCM otherwise]
Assoc-Resp
n-th 4way HS
focus on M3, M4
M3 [$MAAD_{1,n+1}$, $MAAD_{2,n+1}$, ID]
M4

Fig. 7. MAAD sequence diagram

Opaque Identifier) uses a generic identifier of non-specified form in the 4-way HS. As shown in Figure 8, the similar ping-pong as above is used. In the message M2 of the n-th 4-way HS, the STA send its DID assigned by the AP in the previous iteration, if STA desires to be identified. The message M3 is used by the AP to provide the opaque identifier to be used at next association (i.e., $(n + 1)$-th). As the identifier is opaque, no observer can use this exchange to track the STA. If the STA wants to be recognized at the next session, it sends the identifier (in an opaque IE) upon (re)association. This scheme was the first to be voted into the 802.11bh draft text. It should be noted that the STA is not recognized until it decides to send the opaque identifier.

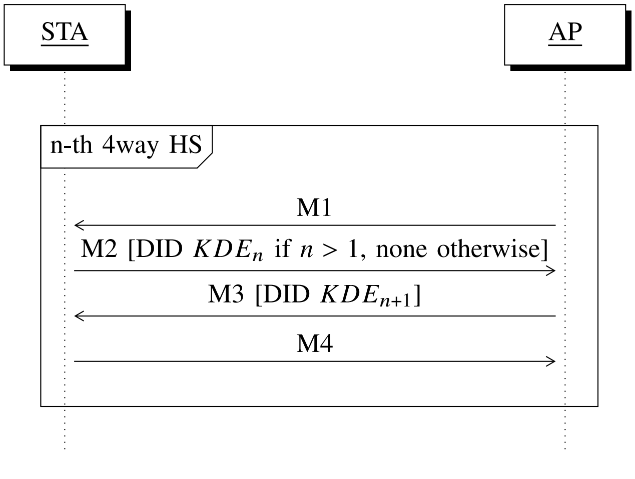


Fig. 8. Device Identifier sequence diagram

D. RRCM and e-RRCM

Rule-based RCM (RRCM) is an interesting proposal which requires additional computation power but extends the horizon of STA identification to more than just next one association. It also allows the STA to be recognized before the next association (i.e., during the scanning phase). This is achieved via the exchange of keying material during the 4-way HS and the use of a seed and increasing numbers as salt for HKDF to compute N MAC addresses to be used in the next N associations. As both sides use the same algorithm, the AP can compute the MAC address that the STA will use in next association (or during next scanning phase), without the need for keys or credentials to be exchanged over the air.

In a variant, enhanced e-RRCM, the STA also computes a Validation Information Element (VIE) that it sends to the AP upon next contacts (scanning or other exchanges). The VIE is a form of hash that shows that the STA is legitimate, limiting the risk of replays attacks (with RRCM, an observer can capture a frame sent by a STA and replay it; the AP has no mechanism to recognize the impersonation).

Classification parameters of 802.11bh schemes

None of the 802.11bh schemes is designed to contradict the 802.11aq rules, as 802.11bh was intended to be a “quick fix”. Therefore, with 802.11bh, changing the MAC address during an association session is not allowed. However, the STA may decide to disconnect, change its MAC address, and then reconnect, as often as desired. In general, all of these schemes allow the STA to opt out of the identification procedure: if the STA desires it can always not "play the game" and behave as a new STA at next association (or pre-assoc) phase. Apart from that, there are some common patterns in the currently proposed schemes which make them classifiable according to a few categories. The classification parameters are summarized in Figure 9 and detailed in the following.

1) Identifier Type: The STA’s identifier has to be a verifiable bitstring. Some approaches proposed in the context of 802.11bh still use random *MAC addresses* as identifiers, based on the assumption that some algorithm shared between STA and AP makes the mapping possible between such MAC address and an identity. IRM, MAAD and IRMA fall into that category. The STA can then associate with RCM, share its next MAC address over a secured link to the AP, then disconnect often and change its MAC. For observers, the MAC has changed and the STA can no longer easily be tracked. But the AP recognizes the STA each time, ensuring continuity of operations. If the STA no longer wants to be tracked, it simply does not share its next MAC address. In other instances (RRCM), the next-association MAC address is computed through some algorithm such as cryptographically secure pseudo-random number generator schemes. In such cases, not only the next-association MAC is known to both parties, but a whole sequence can be computed offline and stored to be used when needed, which is advantageous for the computational workload of AP and STA at association time. This extends the horizon of one algorithm run to mulitple sessions.

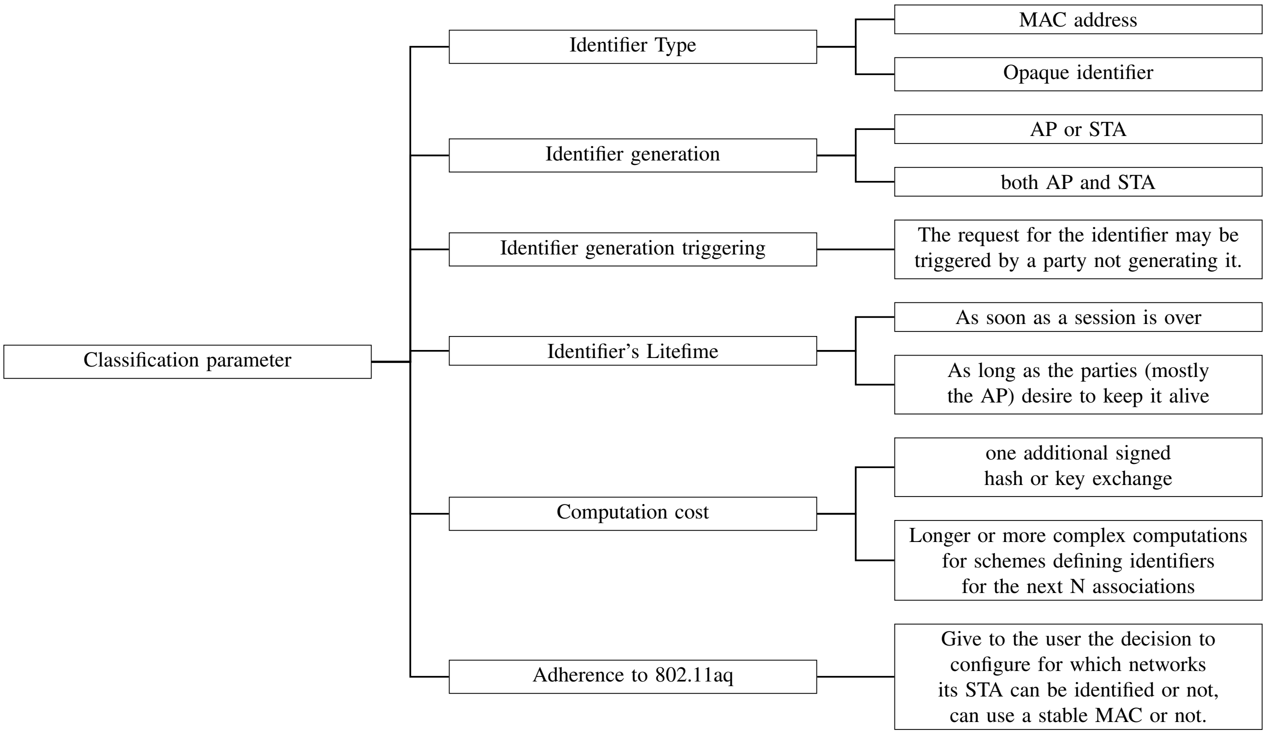

Classification parameter
Identifier Type
MAC address
Opaque identifier
Identifier generation
AP or STA
both AP and STA
Identifier generation triggering
The request for the identifier may be triggered by a party not generating it.
Identifier's Litefime
As soon as a session is over
As long as the parties (mostly the AP) desire to keep it alive
Computation cost
one additional signed hash or key exchange
Longer or more complex computations for schemes defining identifiers for the next N associations
Adherence to 802.11aq
Give to the user the decision to configure for which networks its STA can be identified or not, can use a stable MAC or not.

Fig. 9. Classification of 802.11bh Schemes.

Another class of schemes proposes proposes *opaque identifiers* not in the form of a MAC address, but a value in an information element (IE). The STA is free to use any random MAC address at any association, but it can send an IE to remind the AP of its past identity, allowing the AP to trace the STA across associations, and accordingly maintain continuity of service. The STA needs to send the IE only when it determines that maintaining continuity with a prior association is necessary.

In these approaches, the STA's next-association MAC address is unknown to the AP until the STA decides to share its ID.

This last form (the identity in additional IE) comes in various forms, some of which directly include the identifier for the STA. In that case, the IE is sent over a protected frame (e.g., Action frame sent using Protected Management Frames (PMF), and therefore encrypted), so that only the AP can decipher the frame and read the value. Within the decrypted IE, the identifier is not encrypted. Some other forms recognize that PMF may not be implemented everywhere, and use clever schemes in the IE encoding, so that no one else than the AP can understand its content. The IE can then be a hash of a value that was transmitted or known to the AP before (in previous sessions), and that would not mean anything to non-AP observers.

2) Generation and Allocation of Identifiers: In the two-party relationship we call wireless association, both sides can take the lead and decide what the identifier for a STA is. Approaches have been proposed for both sides to lead and generate the identifier: AP or STA. In a few schemes, both options are possible, according to negotiation or configuration. One interesting variation of this theme is the definition of who triggers the procedure: the request for an identifier can also come from the party that does not generate the identifier.

3) Identifier's lifetime: As mentioned above some identifiers have a very short lifetime: as soon as a session is over, the identifier is gone and any effort spent in tracking it is lost. Other approaches (such as Device ID) propose an identifier for the device that is shared through verifiable secure signatures and stay for as long as the parties (mostly the AP) desire to keep it alive. In other words, no requirement is present for it to be destroyed at each session end.

4) Computation cost: All of the schemes propose slight additional messages or computations at association or pre-association phases. If the horizon of the identifier's lifetime is limited, the same can be said of the computation: one additional signed hash or key exchange. On the other hand, RRCM and e-RRCM define the next N associations and require therefore longer or more complex computations.

5) Adherence to 802.11aq: These schemes provide different levels of adherence to 802.11aq. They suggest that what is needed from 802.11bh is text to clarify in the main Standard the conditions in which a STA can or should change its MAC address. The text would provide control to the user, who would decide on networks for which the user would want the STA to be identified across sessions (e.g., work Wi-Fi) and for which a stable (possibly the burned in) MAC address would be used. For some other networks, the user would configure slow-changing RCM schemes, and fast rotations (with short disconnections, as often as possible without disrupting the user activity).

Table I shows the comparison of the presented 802.11bh schemes.

Parameter	IRM	RRCM	MAAD	Device-ID
ID Type	MAC	MAC	MAC	Opaque
Generation	STA/AP	AP	AP	STA/AP
Lifetime	1	1	1	N
Horizon	1	N	1	1
Computation cost	Low	High	Low	Low

TABLE I
802.11BH SCHEMES COMPARISON ACCORDING TO TYPE, GENERATION, LIFETIME AND HORIZON OF IDENTIFIERS, TOGETHER WITH ITS COMPLEXITY.
LIFETIME, AND HORIZON ARE EXPRESSED IN NUMBER OF SESSIONS.

E. 802.11bh research challenges and future directions

802.11bh was intended to be a short-lived group, inserting a form of stable identifier into the 802.11 Standard to overcome the possible risks of RCM. However, the group work has faced challenges that, in hindsight, were to be expected: implementation challenges, delegation of trust and pre-association discovery.

1) Implementation challenges: The schemes proposed to 802.11bh and based on MAC addresses are conceptually attractive because they are simple. The STA uses a MAC address (of its choice or suggested by the AP), just like it did with RCM and pre-RCM schemes. No need for an additional IE that may hint that the STA is returning to the network (and thus was previously known). The STA decides if it wants to be identified by the AP across association by using the MAC suggested or announced in the previous session, and uses a different MAC otherwise. However, implementing such a scheme is not trivial. The STA needs to keep a map of current vs. future MACs, possibly per SSID, possibly across multiple associations (to avoid reusing the same MAC twice) and on the fly replace its own MAC with a new one if it decides to disconnect and reconnect. None of the low-level drivers in the market today are tailored for this type of tracking and fast Link Layer identifier swap.

The schemes based on an Information Element suppress this difficulty, but may cause compatibility issues that can be hard to solve. The presence of a specific IE can alert attackers to the fact that the STA is returning to the network (causing the attacker to use that information as input into a probabilistic engine to determine if parameters sent by the STA resemble parameters associated to a previously observed MAC address). Thus, the presence of the IE may provide additional hints to an attacker. Additionally, the format of the IE needs to be carefully designed to allow for multiple scenarios (e.g., IoT, enterprise, retail, hospitality) where the identifier string may be different. Thus, an open format may sound preferable. However, such open format also presents the risk of allowing malformed strings (values that would either be not understood by the AP in a particular context, or would be used as special characters to attempt some form of attacks). The IE schemes based on PMF will fail in non-PMF networks. The IE schemes based on smart value encoding or encryption cause both the STA and AP to proceed to additional computation, which may not scale easily (either for the STA wishing to rotate its MAC address at faster pace, or for the AP needing to compute values for possibly hundreds of STAs entering the same venue at the same time).

Complexity: Related to implementation is also the complexity aspect of such schemes. History has proven that complex algorithms tend to be translated in suboptimal real-life implementations in chipsets firmware and software, which eventually bring more harm than benefit to the original intent (think of WPA2 and KRACK attack [49]). While one may argue that these are just bugs, an argument can be made on the real added value for such additional complexity, which eventually advocates for simple and self-evident schemes to have an edge in the long run for success.

2) Delegation of trust: The very fact that the protocol defines an identifier for the STA means that the value generation needs to be carefully designed. For example, suppose that a low-level AP generates always the same identifier for a given STA, or a value that could easily be reversed engineered (e.g., the current identifier value plus a fixed index). The whole purpose of RCM would be broken. Thus letting the AP generate the value would be accepted by a STA manufacturer only if the generation scheme were to be both strictly defined and very difficult to reverse-engineer. Such a design goal is possible, but takes time. Therefore, it may be preferable to let the STA generate its own values. Vendors that are privacy-conscious would implement a robust scheme, without being at the mercy of a poor implementation, and without having to wait years for the 802.11bh group to design a robust scheme accepted by all. However, such a possibility also introduces its own set of challenges. The AP would need to accept and recognize the identity value shared by the STA with or without PMF. There is no guarantee that the AP would recognize and accept any value returned by the STA (e.g., because some special characters are not implemented or forbidden). Even if the AP recognizes the format, there is no guarantee that the value would not have already been sent by another STA on the network, causing identity collisions. By contrast, if the AP generates the identity value, it would discard any generated value that is seen as already in use. But an individual STA cannot have a view of the list of STAs and their identifiers already active on a given AP.

3) Pre-association discovery: The group is also debating the validity of some of the use cases that led to its formation. One of them is pre-association discovery. For example, in some networks, devices of a certain type need to be steered to specific bands or radios (e.g., IoT devices to the 2.4 GHz APs or radios, AR/VR to 6 GHz, etc.) This action needs to happen before association (e.g., APs in the 5 GHz or 6 GHz bands would not respond to the IoT object's Probe Requests). The same requirement appears when the STAs roam between APs, and probe to discover the next AP. Using one-time RCM addresses does not allow for such identification. Conceptually, the issue of pre-association identification seems to be the realm of 802.11bh, as it is one function of 802.11 that RCM broke. However, the relationship between a STA and the AP exists when both exchange frames. 802.11aq clarifies that the STA MAC address should stay static as long as this exchange continues. But nothing in 802.11 mandates that the AP should recognize a STA before the STA attempts to associate. The Standard also does not describe any pre-association steering mechanism. In fact, the Standard states somewhat the opposite [50] in clause 11.1.4.3.4, that the AP must respond with a Probe Response to any Probe Request it receives (with a few logical and procedural exceptions). Ignoring a Probe Request in a given band for steering purposes is therefore not in the Standard, and the question is vividly debated in 802.11bh: should 802.11bh “fix” a common industry practice that happens to violate the Standard? Proponents of the use case observe that some APs respond to all probe requests: they just respond faster in the bands where the device should be steered, and delay their response in the other bands, leading the device to believe that the only available bands are the ones where the device is to be steered. Delaying a response does not violate the Standard. Similarly, opponents to the pre-association identification examine the use case where the MAC address of a device is used to set lighting or heating levels, and possibly other elements (e.g., command the coffee machine to start, auto-open doors etc.), and they underline how unsafe it is to solely rely on a MAC address to perform all these actions. They claim that 802.11bh should not offer solutions to implement an unsafe behavior that happens to have been rendered impossible by RCM. Proponents of the pre-association identification note that the solution would only apply to returning STAs (and therefore use protected frames that would prove each side’s identity to the other), and that 802.11bh could achieve therefore two goals in one solution: solve the "arriving MAC" issue, while removing the security weaknesses of the pre-RCM solution.

All these issues can be solved, and the 802.11bh work continues. However, a short-lived action to implement a fix has become a deeper reflection on what an identifier is and who should generate it.

VI. IEEE 802.11bi

802.11bi was formed with the explicit goal of standardizing user privacy solutions in IEEE 802.11. This is a vast and complex endeavour, because many of the 802.11 pre-association discovery procedures, with their rich exchange of capabilities and support for optional features, lend themselves to fingerprinting, with or without RCM. In the 2020 revision of the IEEE 802.11 Standard [50], the Probe Request, by which a STA discovers the AP (and indicates along the way the protocols it understands, so the AP can craft its Probe Response accordingly), allows for 34 optional elements, the Probe Response allows for 92, the Authentication frame 25, the Association Request frame 46, and the Association Response 57 optional elements. Even after association, when the exchanges are expected to be encrypted, the headers of the frames are not protected, simply because all receivers in range need to know if they are the intended recipient or not. An observer can easily follow the dialog between any network entity and a particular STA, and deduce a lot of information from elements of the header, or the simple observation of the STA address (as source or destination) and the structure of the frame (e.g., its size, the pace of subsequent frames, etc.).

Solving all these elements together may be difficult, if a consequence is to re-design the entire Standard. Therefore 802.11bi is a project targeting several years of work, and which first task is to list the potential privacy exposures and the requirements for a private (or more private) IEEE 802.11 Standard. More than 50 requirements have been identified so far, that can be grouped in 9 categories. It is important to note that, at this early stage of the group work, these are simply requirements, i.e., goals to achieve. The protocol definition will come in the next phase. Many of the requirements target the STA, because it is the side where a personal device is commonly expected. Mechanisms that focus on the STA are labelled CPE (Client Privacy Enhancements). However, the group also envision protection mechanisms for the AP privacy, for two reasons:

- The AP may be a soft AP on a personal device (e.g., a smartphone tethering for a personal laptop), or a home AP. In both cases, it may be a personal device, and its traffic may reveal information about a private user.
- Even in the cases where the AP is a shared service device, it works in tandem with the STA. Communication patterns visible on the AP may reveal information about the STA on the other end of this communication. Thus protecting the STA may also mean protecting the AP.

Naturally, protecting the AP makes no sense if the STA is not also included in the protection mechanism (if the AP is a personal device, there is little doubt that the STA connecting to it is also a personal device). Requirements that protect both sides are labelled BPE (BSS [Basic Service Set, the area covered by an AP] Privacy Enhancements). BPE features extend CPE features. Figure 10 summarizes the different directions explored.

A. Pre-association protection

The pre-association case is fundamental to the idea of privacy-supporting 802.11. The STA and the AP need to agree on parameters allowing them to negotiate the possibility of an association. It may be possible for a STA to only communicate to the AP standardized parameters, sufficient for the AP to know how to answer to the STA (e.g., which protocols and data rates the STA would understand, without expression of any particular additional capability). Such simplification is an obvious improvement that would reduce the STA privacy exposure, as most STAs would likely express the same support.

However, there are two difficulties that still need to be addressed.

- The STA may need to indicate support for particular options. In many cases, this support does not change the STA decision to associate or not, and simply expresses services that the STA expects or desires. In such scenario, the support could be negotiated post-association, through PMF. In other cases, however, support for optional features may condition the STA decision to associate or not. A typical scenario is a STA, running a delay-sensitive application (e.g., a Wi-Fi-enabled AR/VR headset), in range of 2 APs with comparable signal level. The STA may then want to query the APs for their support

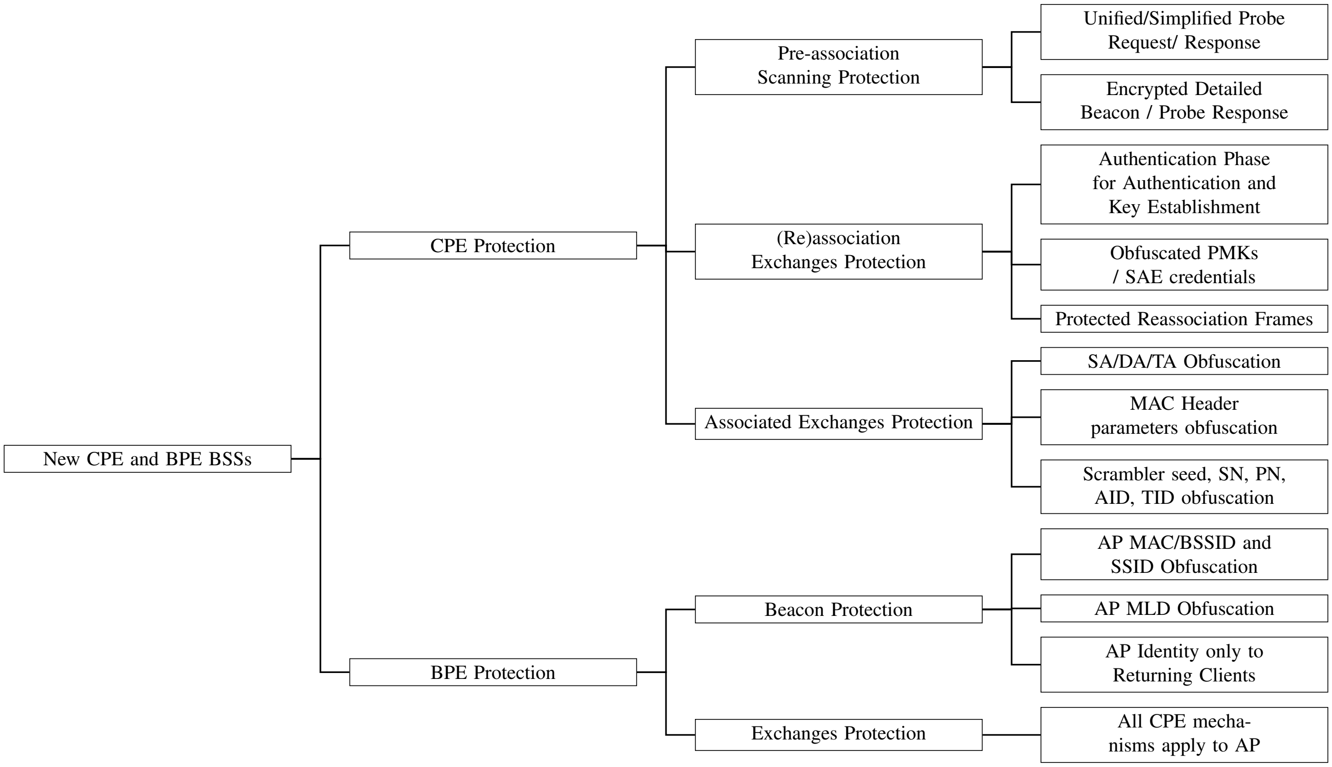

New CPE and BPE BSSs
CPE Protection
BPE Protection
Pre-association Scanning Protection
(Re)association Exchanges Protection
Associated Exchanges Protection
Beacon Protection
Exchanges Protection
Unified/Simplified Probe Request/ Response
Encrypted Detailed Beacon / Probe Response
Authentication Phase for Authentication and Key Establishment
Obfuscated PMKs / SAE credentials
Protected Reassociation Frames
SA/DA/TA Obfuscation
MAC Header parameters obfuscation
Scrambler seed, SN, PN, AID, TID obfuscation
AP MAC/BSSID and SSID Obfuscation
AP MLD Obfuscation
AP Identity only to Returning Clients
All CPE mechanisms apply to AP

Fig. 10. 802.11bi Requirements Groups.

for QoS parameters before deciding which AP is more likely to provide a good user experience. Surfacing the requirements for QoS parameters is necessary for the association decision and may expose information about the STA. One possible direction is to create an association state where a STA can join more than one AP (in the current IEEE 802.11 Standard, a STA can associate to only one AP at a time), use this secured link to negotiate additional parameters, before opting to stay on one of the APs and start its data flow. Such a mechanism would add some overhead, compared to the existing association choreography (as the STA would need to negotiate keys with both APs), but the benefits may be higher than the cost.

- With 802.11be/Wi-Fi 7 Multi Link Operations (MLO), a STA may establish 2 links to an AP. In the current version of the Standard, a STA can announce on one link its multi-link identifier, allowing any observer to map the individual links to the same STA. In other words, observing one link allows an eavesdropper to know what other link the STA is using. A good solution could be to modify the multi-link negotiation structure, and mandate that the STA and the AP should negotiate second link parameters over unicast and encrypted frames in the first link.

B. *(Re)association protection*

The 802.11 association process has three steps (after the STA has established a list of potential APs to join through passive or active scanning).

- The poorly named **authentication phase**, which goal is for the AP to validate that the STA is a valid 802.11 device. Because the term “authentication” has a security connotation, that was raised by many when the first version of 802.11 came out, the early days of 802.11 added to this phase a form of weak mutual authentication (Wired Equivalent Privacy, WEP) that would allow both sides to confirm that they had the same pre-shared key. WEP was inefficient and flawed, and got deprecated in 2007, pushing the authentication phase back to its original goal. However, the value of that phase leaves much to debate. Today, all STAs and all APs implement the entirety of the basic protocol required for these initial exchanges. There is no reason to imagine that a STA would only implement a part of it, and the authentication phase has become a redundant exchange. The 802.11ai-2016 [51] amendment attempted to make this phase more useful, by using it to effectively negotiate authentication keys. But without this additional provision, the authentication phase is similar to these dialogs in video conferences, where one part says "good morning", the other replies "good morning to you", and the first party asks "can you hear me?" (obviously yes, otherwise the other would not have replied). Similarly, the STA could start with the second step below (association), and understand that the lack of response signifies that the STA did not identify a proper 802.11 frame.
- The **association phase**, whose goal is to assign an association identifier (unique client number) to the STA on the AP, and exchange on optional parameters support.

One key issue of course is that the exchange is not encrypted, exposing the STA's (and the AP's) parameters.

- In non-protected networks, the connection is completed after the association phase, and the STA can start sending and receiving data. Most networks implement a form of encryption, however, and the association phase is followed by an exchange aimed at performing **mutual authentication and exchanging encryption parameters**. The outcome is a secured (authenticated and encrypted) connection, mapped to an identifier (e.g., the Pairwise Master Key IDentifier, PMKID, that uniquely identifies a passphrase or a password, without sending its value over the air, in essence "we used password number 64"). This identifier is useful when the STA briefly leaves the network and comes back, with a reassociation request: the STA can directly say "I am back now and will use password number 64". The AP knows directly how to decrypt the STA traffic, and the STA implicitly proved that it was already properly authenticated as a valid client.

This choreography shows its age, with the limited benefits of the authentication phase, and the lack of protection of the association phase. Thus a first requirement is to better protect these exchanges, by making the authentication phase a real authentication phase, through which the 802.1X/EAP exchange, used to perform individual STA and AP mutual authentication and key derivation, would be conducted, as shown in Figure 11. The outcome would be that encryption keys would be available at the end of the authentication phase. The association phase, with its cohort of exchanged parameters, could then be protected from view. As the authentication would also be completed at this point, the STA would also know that it is exchanging with a real AP (and not an attacker impersonating an AP), further limiting the risk that a third party could attempt to collect STA parameters by pretending to be one more AP in the network.

By modifying the authentication procedure, it could also be possible to obfuscate the connection identifier that the STA shares upon coming back. By design, keys are set to be unique for a STA-AP pair (this is true for the operational keys derived from a hotspot-wide shared passphrase, true for other mechanisms like Simultaneous Authentication of Equals (SAE), and naturally true for STA-specific authentication (with 802.1X/EAP). For all these mechanisms, the STA shares the PMKID or the SAE Identifier upon returning, and it is trivial to identify the STA by watching for the PMKID or SAE ID as the STA moves about. The identifier does not expose the STA security, but uniquely identifies the STA. By modifying the authentication phase, it may be possible, upon reassociation, to obfuscate somehow the PMKID or SAE ID, and hide from eavesdroppers information about the STA (and the fact that it is returning).

C. Over the AIR MAC and parameters protection of associated exchanges

As shown in Figure 12, the 802.11 header includes 4 MAC address fields, whose meaning depends on the flags FromDS and ToDS of the Frame Control field. When

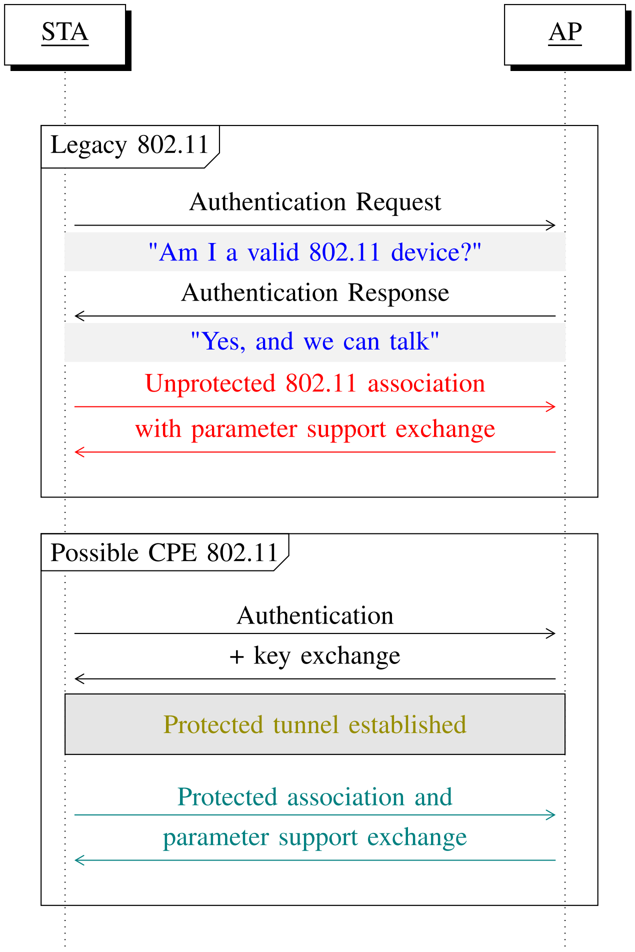
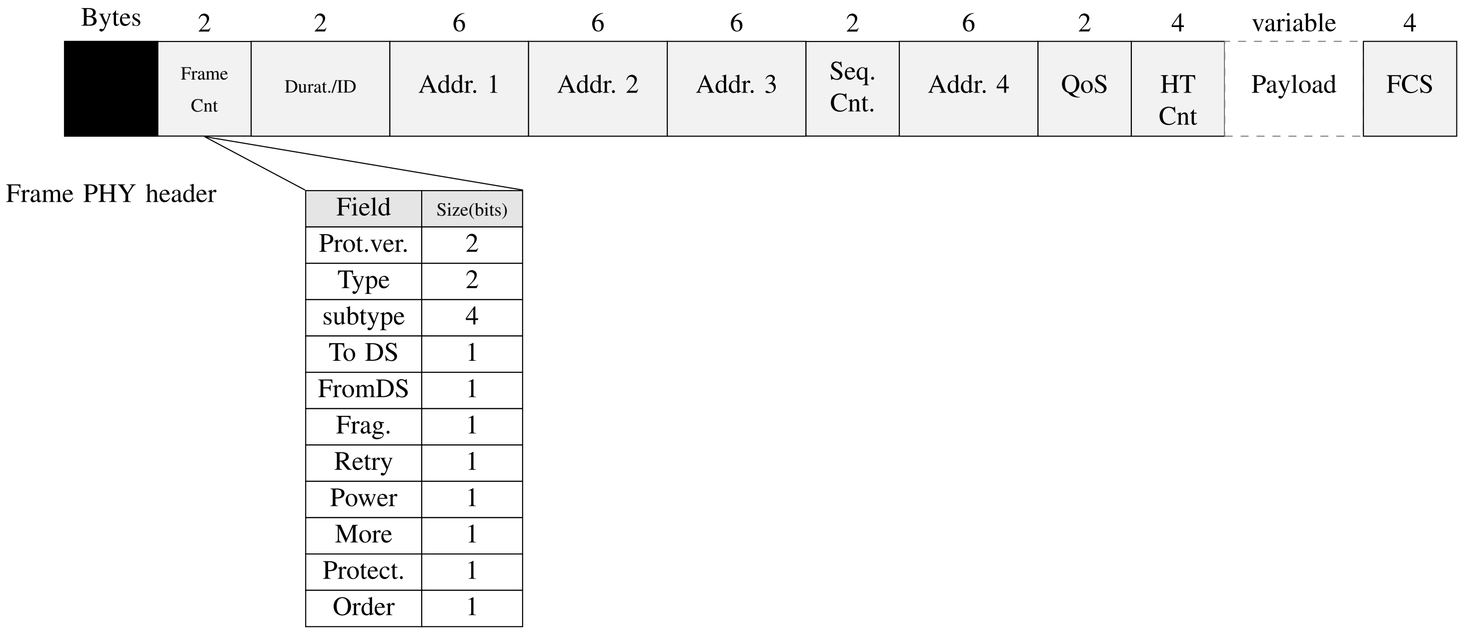

STA
AP
Legacy 802.11
Authentication Request
"Am I a valid 802.11 device?"
Authentication Response
"Yes, and we can talk"
Unprotected 802.11 association
with parameter support exchange
Possible CPE 802.11
Authentication
+ key exchange
Protected tunnel established
Protected association and
parameter support exchange

Fig. 11. 802.11bi: Replacing Legacy Design With Improved Association

the STA is transmitting towards the AP (i.e., FromDS=0 and ToDS=1), the first address is the MAC address is the AP to which the frame is sent (the receiver address, RA), a second address is the STA MAC address (as source address [SA] and transmitting address [TA] entity), and the third address is the end destination (DA, destination address) of the frame The destination can be another STA in the BSS, or a target on the wire behind the AP. Symmetrically, when the AP is transmitting to a STA (i.e., FromDS=1 and ToDS=0), the first address field contains the STA address (as intended receiver and destination, RA and DA), the second the AP address (as the address of transmitting entity, TA), and the third contains the address of the source of the frame (SA, this time the other STA in the BSS, or target on the wire, responding to the STA). In both cases, the fourth address field does not contain useful information. This field is used when the frame is forwarded over the distribution system connecting two APs of the ESS. In this case, FromDS=1 and ToDS=1 and the four address fields contain respectively the RA, TA, DA and SA. The header also contains additional information, such as the frame type (management, data or control), derived from the fields Type and SubType of the Frame Control, and other procedural parameters that are of interest to the intended RA, once the frame has been received and demodulated.



Field	Size(bits)
Prot.ver.	2
Type	2
subtype	4
To DS	1
FromDS	1
Frag.	1
Retry	1
Power	1
More	1
Protect.	1
Order	1

Fig. 12. The standard 802.11 MAC header

In traditional 802.11, the header is not encrypted, allowing all to see the surface of the exchange but not the content. The intent behind excluding the header from the encryption is again the result of 802.11 long history. With WEP, the header also included key elements allowing the receiver to understand how to decrypt the frame payload. Later mechanisms, defined in 802.11i-2004 [52] and merged in the main standard in 2007, aimed at providing a better protection of the payload, but were built with backward compatibility in mind, and thus also left the header unprotected. The keying material for what is commonly known as WPA2 and WPA3 is in fact exchanged using the 4-way handshake messages using the EAPOL protocol. In other words, this information is part of the payload of the Data MAC frame, thus it is inserted after the QoS field of the header.

Today, it makes sense to leave the RA in the clear, which would be *address 1* in most cases, as such visibility is needed for all STAs to start receiving the frame, read the header, stop if the RA is not their own address, but continue reading the frame if the RA is their address or a multicast/broadcast value. However, no observing party has any good reason to need to know who the source of the frame is (SA or TA), or the final destination (DA), if it is not also the RA. Also, the Frame Control part of the frame is only of interest to the correct receiver, and to no one else. Therefore, one direction could be to include these elements in the encryption scheme to obfuscate them from view. Such broad change may cause backward compatibility issues, resulting in the creation of new types of networks where only CPE/BPE (as defined above) devices would operate. For hybrid networks (with legacy devices, unaware of privacy protection enhancements), a simpler (but less effective) method could be to simply obfuscate any address outside of the STA-AP pair. For upstream flows, the RA would be visible as required, but the DA would be an obfuscated identifier, looking like a MAC address, that the STA and AP would agree upon through a protected initial exchange. The TA/DA could also be encrypted, in theory. In practice however, the AP likely has a different key for each associated STA, and needs to know which key to use to decrypt the traffic. There needs to be some information about the sender that the AP understands so as to pick the right key. The easiest way is to let the TA visible. Another way may be to use a form of frame sequence number prediction, that would allow the AP, from a value in the header, to know which STA sent the frame, without revealing the numbering scheme to observers. An obfuscated identifier would also be used for the SA for downstream traffic, thus hiding from view the device with which the STA is dialoguing through the AP.

When a STA needs to roam to another AP, IEEE 802.11r (Fast Transition, FT) allows the STA to pre-negotiate keys and other parameters with the next AP before fully switching its connection. This precaution allows the STA to immediately start transmitting and receiving data upon connecting to the next AP, without wasting time negotiating credentials. Depending on the FT flavor that the STA uses, it either dialogues with its current AP, indicating the MAC address of the target next AP, or with the next AP, indicating the MAC address of the AP it is currently associated to. Here again, these MAC addresses should be obfuscated from view.

Along with the MAC address, the other elements identified in Section II should be protected. Some of them can be obfuscated (integrated within an encrypted header), such as the frame sequence number (indicating the frame number, so duplicates can be identified), the Association ID (AID) by which the STA indicates its registration number (as client on the AP for this particular session), or the Traffic Identifier (TID) that indicates the QoS category to which the data frame belongs. These elements are indicated in many frames. Here again, it is unclear why any observer should be allowed to know another STA AID, or that a particular STA is sending voice vs. video traffic. Other elements are difficult to obfuscate with encryption, as they reside outside of the MAC layer. It is the case of the scrambler function used to generate uncorrelated binary sequences with uniformly distributed bits. This operation allows the same data payload to be mapped to different binary sequences (and hence physical signals). However, some devices implement simple and deterministic algorithms to seed the scrambler, such as running the scrambler seed continuously from frame to frame without a reset or incrementing by one the seed each frame is sent [8]. These simple implementations of scrambler provide information to attackers and allow for the re-identification of transmitters even if MAC addresses or pseudonyms are changed. A good mechanism could be to design a method so that the scrambler seed is randomly changed for each frame.

D. Protected Management Frames

In 2009, the 802.11w amendment [53] (integrated into the 2012 revision of the Standard) instantiated the possibility to encrypt some management frames (Protected Management Frames, PMF). The goal was primarily to avoid AP impersonation, by which an attacker would spoof the AP MAC address, and send a frame to the victim STA, for example unduly terminating the session. Many management frames created by amendments developed in the subsequent years received PMF protection when the risk of impersonation could lead to a detrimental user experience. However, many other management frames stayed unprotected, because the risk or the consequences of AP impersonation was very limited, or because the management frame would come from the STA. However, these frames can surface information about the STA, and possibly help in its fingerprinting. For example, for multi-user transmission, or for beamforming, the AP may send test (sounding) frames to the STA, over a representative subset of the channel, and the STA would return a matrix showing how the signal was received (allowing the AP to know how to best beamform to the STA, or which STAs can be grouped together). An observer can use these frames to roughly understand the STA distance to the AP, if the STA is moving, if its environment is stable, and possibly deduce details about the way the STA receives the AP signals. If the AP groups STAs together, they are assigned a group ID that the STA can also return upon request (here again surfacing information about the STA).

There is no good reason why these management frames should be visible to all observers, and the IEEE has determined a group of 10 management frames that should be encrypted.

E. The special case of the AP and its beacons

The AP can be a BPE because it is a personal device, or at least a device whose privacy needs protection. The case of a home AP or a soft AP running on a smartphone is obvious, but there are also Enterprise cases where the APs are seen as private assets that the enterprise may not want the world (e.g., people passing in the street) to have detailed information about. In that case, many of the elements above (RA/DA and other parts of the header, sequence numbers etc.) can be protected the way they are protected in the client case.

A major difficulty is that IEEE 802.11 expects the AP radio to broadcast multiple times per second a beacon message that advertises the existence of the AP and its parameters. Such message is somewhat necessary, as the STA needs to detect the AP existence before attempting to connect. Each new amendment developed through the 25 years of the 802.11 Standard existence has added information elements to the beacon, to inform passing STAs about new supported capabilities. The 2020 version of the Standard [50] describes close to 80 different capabilities that the AP must, or can, indicate. From a STA privacy standpoint (CPE context), it is indeed better to let the STA ask a general question ("which features do you support?") and let the AP provide the details in a Probe Response. It is even better to let the STA listen passively, and receive the list of supported features in the next AP beacon.

This logic naturally stops when the AP also becomes an object whose privacy is to protect. In that case, it becomes necessary to divide the STAs into two groups: those that are returning to the AP (e.g., your smartphone when you come back home from the office), and those that discover the AP for the first time (your smartphone when you attempt to connect to the Wi-Fi of a friend's home, you have never been to before).

In the first case, a possible protection could be to shorten the beacon to only express trivial information (e.g., basic data rates supported by the AP, that any other AP would support anyway), along with a key, that a returning client would have learned during a previous association. This key could be used to deduce or compute additional elements. Another possibility is of course to encrypt the "non trivial" part of the beacon.

The second case is more difficult to solve, as users expect the see a list of network names (SSIDs), but the SSID itself may be sensitive information if it is unique or identifiable. The intent to protect the AP privacy, in this case, is in the group requirements, but it is likely that different use cases will need to be distinguished, to determine what type of protection can apply to each of them.

The AP MAC address is another element of vivid exchanges. If the beacon stops advertising sensitive information but the AP MAC address remains the same over time, then fingerprinting is still possible. This issue is quite obvious for home and soft APs, but also affects larger networks. Many venues accessible to the general public are operated by private entities. However, several companies use crowdsourcing to establish a map of the venue and its Wi-Fi coverage, with the pretext of facilitating indoor navigation (the veracity of this announced intent is often highly debatable). However, this mapping is performed without the venue owners' consent, exposing the identify and the location of their private networking assets, resulting in all sorts of issues. Therefore, one requirement is to be able to change the AP MAC address at any interval. The mechanism needs to be carefully designed, as it should not affect the experience of already connected STAs.

F. 802.11bi research challenges and future directions

The IEEE 802.11bi faces a long journey ahead. Very early in the requirements determination, the group concluded that it would not be possible to keep the "old" 802.11 *modus operandi* and still implement privacy measures. The Standard as it is defined today is too focused on efficiency and parameters negotiation. There is no way to stop STAs and APs from exchanging them without breaking the communication. The conclusion was to design these two new modes, CPE (to protect the STA privacy) and BPE (to protect both the STA and the AP privacy).

However, such a decision is the easy part. In practice, building new modes where all elements are obfuscated from view is not very difficult. In concept, this phase is just about deciding which elements must be in the clear: for data frames, the RA MAC address. The frame then only needs a cryptographically safe element to hint (only to the receiver) on who is the sender (which of the STAs and the AP) to allow the receiver to decrypt the rest of the header (and of the payload), hiding the exchange from everyone. An observer may still use pattern recognition to understand the dialog (e.g., recognize small frames at regular intervals as voice packets). The group may then take inspiration from work in other groups, like the IEEE 802.1Aedk [54], to add padding and empty payload when necessary to further hide the content of the exchange. None of this structure would be understandable for legacy STAs, but 802.11 is used to this type of challenge. The same issue happened when 802.11n was released (with non-802.11n STAs), then 802.11ac, 802.11ax, etc. In all cases, the PHY preamble indicates the duration of the transmission, so legacy STAs can wait during a transmission which content they would not understand.

The equation becomes more difficult to solve when pre-association is factored. If the AP is a shared service device, then the STA could passively listen to the beacons. The Standard already envisions the idea of shorter beacons, containing only essential data like supported data rates for communication, that the AP could send at faster intervals, e.g., every 20 ms, to limit the STA dwell time on each channel of interest. The STA could also use RCM to send a general discovery question, that would use the same format for all STAs. But if the AP is also a personal device, the challenge is more difficult. Certainly, for habitual networks (e.g., your home), the AP and the STA could agree on a key. Then, when the STA returns, it could hear a short beacon with a cryptographic label that the STA recognizes as being the home AP. Other STAs may hear a short beacon from an unknown (and changing over time) MAC address, whose only clear information would be support for basic data rates. Any other information (like the SSID or the other supported parameters) would be encrypted (and only decipherable by the returning STAs).

This nice scheme leaves unsolved the issue of new STAs. When a new STA is invited to a network where a BPE AP operates, it has no way to discover the AP supported parameters or even the SSID. One possibility is of course to let the network owner "open" the AP communication (removing the encryption), but this is not a very promising direction: putting the omen of the Standard limitation on the user is not a very good technical design, and it also opens to door to multiple issues (users never "closing" back the communication, users not knowing what "opening" or "closing" could mean, etc.). In the end, the goal is a network that "just works", not a network where the user has to complete multiple tasks just to get connected. Certainly, there are tools like the Wi-Fi Alliance EasyConnect [55] program, allowing a STA that already has the network parameters to pass them onto another, new, STA, but adopting such a tool would also mean changing the way billions of users connect to Wi-Fi, with no clear promise of a simpler or automated experience.

The scheme also leaves unsolved the question of hybrid networks. The past amendments that implemented new PHY rates (named in the Wi-Fi Alliance by generational terms, like Wi-Fi 4, 5, 6 etc.) were adopted fast because they were backward compatible with previous generations. A Wi-Fi 5 STA could not communicate at Wi-Fi 6 speeds, but a Wi-Fi 6 AP did support Wi-Fi 5 data rates. When a Wi-Fi 6 STA would speak to the AP, its preamble would inform the nearby Wi-Fi 5 STA of the transmission duration, and it did not matter if the Wi-Fi 5 STA did not understand the content of the transmission. However, at any time, any Wi-Fi 6 device could revert back to Wi-Fi 5 modes to communicate with the Wi-Fi 5 STA (whose capability is expressed in the discovery messages, and understood by all). By contrast, it is not reasonable to envision that a BPE AP would revert to legacy mode when a legacy STA is in the area, because the AP would then expose PII and PCI, defeating the whole purpose of the protection mechanism. The same challenge goes unsolved for the STA: it should not protect all frames when discovering a CPE-compatible AP, then expose PII or PCI when discovering legacy APs. This is probably the most difficult part of the IEEE 802.11bi work: solving the long tail of use cases that are not purely greenfield, allowing the BPE and CPE to exist in these environments, while limiting the amount of privacy exposure that may happen. As there are 3 billions new Wi-Fi devices appearing every year on the market, close to 500 millions hotspots, as Wi-Fi has become business critical in about every vertical, the number of use cases, corner cases and exceptions to consider is staggering. Once the 802.11bi group will have designed the main greenfield mechanisms, solving the issue of hybrid networks will probably consume several years of efforts. Choosing not to solve backward coexistence is an avenue that the IEEE 802.11 working group, and the Wi-Fi Alliance, have occasionally explored in the past, and the cost has always been more complexity, and a much, much longer road to adoption.

VII. The IETF MADINAS

802.11bh and 802.11bi focus on key issues related to the bottom 2 layers of the OSI model (L1, Physical and L2, Data Link), where 802.11 operates. The MAC address itself is a Data Link construct. However, privacy is about the user, which is not a L1/L2 construct. The attacker is also a human (also not an L1/L2 construct). In addition, the MAC address is used by multiple upper Layer services (e.g. DHCP, routing and many more). Therefore, the 802.11 approach is unlikely to be sufficient to fully evaluate and solve the privacy and RCM problems. In order to expand the landscape to the upper layers and the network services, the IETF formed in 2020 the MAC Address Device Identification for Network and Application Services (MADINAS) group [5]. The group's goal is to publish recommendations in three fields:

- list upper layer services that may be affected by RCM (in particular services defined or governed by IETF RFCs, Best Practices or other IETF documents);
- evaluate if other STA identifiers could be used for these services (thus removing the need for RCM mitigation);
- evaluate if some services could be provided seamlessly without the need for identification (in cases where the MAC address was used as an identity because the address was available, but for services that did not really require any identification).

As the group is working through the many thousands of IETF RFCs to determine which services may be affected by RCM, it has also published two initial documents. A first document [40] lists the current RCM schemes implemented by the main STA and operating system vendors. The group intends to maintain this document up to date in Github [56], limiting the risk of obsolescence for published RFCs. The second document is intended to be a larger framework [57] for the group work. In many cases, conversations around RCM and privacy are polluted by general and context-less statements ("I do not want "them" to spy on my device traffic"). In this scenario, who is "I", who is "them", what does "spy" exactly mean and why is traffic visibility a problem is very context-dependent.

To help clarify the exchanges, the second document defines different environments where personal device traffic may appear: home (managed by the user, managed by a third party, like a service provider in assisted living facilities), enterprises (with corporate-provided devices and/or user-owned devices), and public venues. The document then notes that the trust that a user may put in these environments varies. In many cases, a user does not worry that other users in the home space have access to systems (e.g., AP or switch) where all users' traffic may be visible. Depending on the home layout, the user may worry that observers outside of the home may be able to observe the home's wireless communications. As such, the user's trust that the home space does not present major privacy exposure may be reasonably high. By contrast, in a public venue, any person sitting nearby, any entity accessing the AP or the LAN is a potential eavesdropper, and the trust that privacy is not at risk is close to or at zero. In between, trust is high in enterprises, especially for corporate-provided devices (as traffic carried by them is not personal), but can be much less for personal devices brought to the enterprise network (where some traffic can be work-related, but other traffic can be at time personal).

At the same time, the trust is not only about the type of traffic sent and received by the device, but also about who the eavesdropper might be. So the framework document also lists the possible actors that may have visibility into the traffic. This part is interesting because the document notes that there are both human-related entities, but also network functional entities. Human-related entities are of course potential observers, in range of the wireless signal (over-the-air), on the LAN or on the Internet, but also people managing the network. These last entities are not only often expected to see the traffic, but sometimes required to (e.g., to support customers reporting issues or difficulties, or for legal reasons). In this last case, the privacy issue enters difficult territories related to user consent and anti-piracy laws. Fortunately, legislators in many countries are working to produce frameworks for these two considerations (half of the world now has privacy-related laws). Unfortunately, these laws may be very different from one country to the other. The document stays away from hard positions in this field, noting that MADINAS can not solve or operate in this space, and can only observe that some regulatory domains will mandate access to some of the traffic while others will not.

The document also considers the network functional entities. These include devices operating at Data Link layer (e.g., APs, switches), and devices concerned with RCM and operating at upper layers (e.g. routers and upper layer services). These are the elements that MADINAS considers to be in in its scope, without opinion on whether they should or should not track individual device MAC addresses and have visibility or not into the associated traffic. Instead, the document suggests an inventory to list which systems and services use MAC addresses (or the carried traffic), for what purpose, whether RCM breaks their functionality or not, and whether a replacement is already possible or not.

A. MADINAS research challenges and future directions

At the time of writing these lines, there are more than 9000 RFCs. Some of them are partially updated by others, many are about topics that are of no concern for MADINAS. But combing through all 9000+ texts and operating triage is going to take a lot of time. The task is complicated by the very nature of the services encountered. DHCP is a typical illustrative example of this complexity. The operations of the Dynamic Host Configuration Protocol are driven by multiple RFCs, primarily RFC 2131 [58], but also RFCs 951 [59], 1531 [60], 1541 [61], 3396 [62], 4361 [63], 5494 [64] and 6842 [65]). The DHCP protocol was designed to efficiently provide an IP address to clients, and manage the lifetime of such allocation. It is clear that all supporting texts were developed in days where there was the expectation of a strict association between a single MAC address and a single device (pre-RCM). As such, the service is designed to assign a given IP address for a certain time (hours and sometimes days). If the DHCP client leaves the network before the end of the lease, the DHCP server has little means to know about this departure. In principle, the client could send a message (DHCP release) to free the IP address, but many clients ignore this phase (or can't send the message for many operational reasons). The DHCP server then tends to keep the address reserved for that client (seen as a MAC address) until the client is expected to communicate again with the server, to renew the address. Even when returning after the expiration of the lease, a client remembering its previous IP address in that network may request to be allocated the same address again. If no other station is assigned that IP address, the DHCP server obliges, as this often facilitates the client session resumption.

Obviously, a client that disconnects and comes back with new (RCM) MAC addresses at short intervals can cause two types of issues: on one hand, a single client can consume all the available addresses in the DHCP pool, exhausting the server resources. Naturally, the client could be mandated to send a DHCP release message, to free the IP address as the client performs RCM. But in a public venue, the personal device does not trust any element of the infrastructure, wants to hide that it is rotating its MAC address and would not send such an obvious clue to the network. On the other hand, the client may change its MAC address and yet request to be assigned the same IP address again for session stability. But from the server viewpoint, it is a new MAC address and therefore a new client, that cannot take an address already assigned to another STA. Some MADINAS members have suggested to reduce the lease duration to a short amount of time, but this approach creates new issues (for battery operated clients attempting to conserve energy by going to sleep when idle, for servers that may see large bursts a renewal requests at a scale that they are not designed to handle, etc.). Thus the DHCP service may have provisions that can alleviate issues introduced by RCM, but the reality of implementations may make these provisions ineffective.

Additionally, the DHCP service includes options that may contradict the RCM privacy goals. For example, the client can include in its messages a Client Identifier field (that may be used for multiple purposes, such as providing a client type to receive client-specific options, assert the client station group membership [and receive custom access rights accordingly], or simply map machines to users). If a STA sends the identifier “Jerome’s phone 5559131”, rotates a MAC address and comes back requesting a new address with that same identifier, there is a high probability that it is the same client as before, not two devices that just happen to have the same identifier. Here, recommendations are hard to make, because they are environment-dependent, and the DHCP client on the STA is not connected to the Wi-Fi driver, and does not necessarily know in which environment it operates. Thus solving the issue of DHCP alone promises to be a long journey, and DHCP is just one example of many protocols using MAC addresses or affected by RCM in one way or another. The group has identified several dozens of similar protocols so far, each in its area and facing the challenges of having been thought and designed before RCM. Each protocol may have provisions that contradict the goals of RCM. Other provisions may present a partial fix allowing the protocol to maintain its operations in an RCM world, but these provisions may not be used widely in the real world, or their implementation may dramatically affect the server (and supporting network) structure and scale.

The MADINAS group continues its work, but it is likely that the outcome will result in a set of recommendations that will bring dramatic changes to how networking operates today, both in terms of industry practice and in terms of client and server-side implementations. As such, the road to adoption is likely to be long.

VIII. Conclusion

This tutorial examined the issues related to privacy and the use of RCM in IEEE 802.11 communications. As personal devices have become dominant, a protocol that primarily focuses on exchanges efficiency presents multiple opportunities for attackers to fingerprint stations, exposing the user privacy as soon as a connection is made between the device and its owner. To limit this risk, station vendors have started implementing Randomized and Changing MAC addresses, as the MAC is traditionally used as the unique identifier for each network entity. However, changing the MAC address can have negative effects on the communication, and Standard bodies are working to address the issue or RCM and its consequences. IEEE 802.11aq clarified that the MAC address should not be changed during exchanges, like association, where a state is expected, for the STA, on the AP. However, the STA could disconnect and reconnect at fast pace and still fulfill 802.11aq requirements. Even in an 802.11aq-compliant context, RCM breaks functions that were possible when the MAC address was static. The IEEE 802.11bh group is working on a replacement identifier that would allow the AP to recognize the STA (with the user’s approval) without exposing the user’s privacy. Meanwhile, the IEEE 802.11bi is undergoing a deeper work to improve privacy protection in 802.11 exchanges. As the MAC address is also used by upper layer services, the IETF MADINAS group is examining the consequences of RCM on these services, with the goal of providing best practices in the use of the MAC and of station identifiers in general for these services, and recommendations on how these services may need to be improved to account for RCM and the need for user privacy. These different efforts will likely need a few more years before coming to completion. Until then, the pressure on station vendors to improve the user privacy continues, and one has to hope that the Industry will come with solutions before vendors bend to that pressure, and start implementing proprietary schemes that might make coexistence and interoperability difficult, causing a fragmentation of the user experience, and an end to the dream of a Wi-Fi that safe to use and “just works”.

References

[1] “ISO/IEC 29100:2011: Information technology — Security techniques — Privacy framework,” Geneva, Switzerland, Tech. Rep. 29100, 2011.

[2] “IEEE recommended practice for privacy considerations for IEEE 802 technologies,” Piscataway, NJ, Tech. Rep. 802E, 2020.

[3] “IEEE P802.11-Task Group BI: Enhanced Service with Data Privacy Protection.” [Online]. Available: https://mentor.ieee.org/802.11/documents?is_group=00bi

[4] “IEEE P802.11-Task Group BH (RCM).” [Online]. Available: https://www.ieee802.org/11/Reports/tgbh_update.htm

[5] “MAC Address Device Identification for Network and Application Services (MADINAS).” [Online]. Available: https://datatracker.ietf.org/wg/madinas/about/

[6] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, “Wireless device identification with radiometric signatures,” in *Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008, pp. 116–127.

[7] A. Jagannath, J. Jagannath, and P. S. P. V. Kumar, "A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges," *Computer Networks*, p. 109455, 2022.

[8] B. Bloessl, C. Sommer, F. Dressler, and D. Eckhoff, "The scrambler attack: A robust physical layer attack on location privacy in vehicular networks," in *2015 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2015, pp. 395–400.

[9] D. Goodin. (2013) No, this isn't a scene from minority report. this trash can is stalking you. Accessed: 13-02-2023. [Online]. Available: https://arstechnica.com/information-technology/2013/08/no-this-isnt-a-scene-from-minority-report-this-trash-can-is-stalking-you/

[10] How stores use your phone's WiFi to track your shopping habits. Accessed: 13-02-2023. [Online]. Available: https://www.washingtonpost.com/news/the-switch/wp/2013/10/19/how-stores-use-your-phones-wifi-to-track-your-shopping-habits/

[11] A. Musa and J. Eriksson, "Tracking unmodified smartphones using Wi-Fi monitors," in *Proceedings of the 10th ACM conference on embedded network sensor systems*, 2012, pp. 281–294.

[12] B. Huang, G. Mao, Y. Qin, and Y. Wei, "Pedestrian flow estimation through passive WiFi sensing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1529–1542, 2021.

[13] Y.-A. De Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the crowd: The privacy bounds of human mobility," *Scientific reports*, vol. 3, no. 1, pp. 1–5, 2013.

[14] N. Maltoni, A. Magnani, and L. Calderoni, "Privacy threats in low-cost people counting devices," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, ser. ARES '20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/3407023.3409195

[15] A. Abedi and D. Vasisht, "Non-cooperative Wi-Fi localization & its privacy implications," in *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*, ser. MobiCom '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 570–582. [Online]. Available: https://doi.org/10.1145/3495243.3560530

[16] E. Fenske, D. Brown, J. Martin, T. Mayberry, P. Ryan, and E. Rye, "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, 2021.

[17] I. Vasilevski, D. Blazhevski, V. Pachovski, and I. Stojmenovska, "Five years later: How effective is the MAC randomization in practice? The no-at-all attack," in *ICT Innovations 2019. Big Data Processing and Mining*, S. Gievska and G. Madjarov, Eds. Springer International Publishing, 2019, pp. 52–64.

[18] L. Pintor and L. Atzori, "A dataset of labelled device Wi-Fi probe requests for MAC address de-randomization," *Computer Networks*, vol. 205, p. 108783, 2022.

[19] C. Matte, M. Cunche, F. Rousseau, and M. Vanhoef, "Defeating MAC address randomization through timing attacks," in *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, 2016, pp. 15–20.

[20] J. Scheuner, G. Mazlami, D. Schöni, S. Stephan, A. De Carli, T. Bocek, and B. Stiller, "Probr - a generic and passive WiFi tracking system," in *2016 IEEE 41st Conference on Local Computer Networks (LCN)*, 2016, pp. 495–502.

[21] M. Nitti, F. Pinna, L. Pintor, V. Pilloni, and B. Barabino, "iABACUS: A Wi-Fi-based automatic bus passenger counting system," *Energies*, vol. 13, no. 6, p. 1446, 2020.

[22] K. Gebru, "A privacy-preserving scheme for passive monitoring of people's flows through WiFi beacons," in *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2022, pp. 421–424.

[23] K. Gebru, M. Rapelli, R. Rusca, C. Casetti, C. F. Chiasserini, and P. Giaccone, "Edge-based passive crowd monitoring through WiFi Beacons," *Computer Communications*, vol. 192, pp. 163–170, 2022.

[24] J.-F. Determe, S. Azzagnuni, U. Singh, F. Horlin, and P. De Doncker, "Monitoring large crowds with WiFi: A privacy-preserving approach," *IEEE Systems Journal*, vol. 16, no. 2, pp. 2148–2159, 2022.

[25] A. Berenguer, D. F. Ros, A. Gómez-Oliva, J. A. Ivars-Baidal, A. J. Jara, J. Laborda, J.-N. Mazón, and A. Perles, "Crowd Monitoring in Smart Destinations Based on GDPR-Ready Opportunistic RF Scanning and Classification of WiFi Devices to Identify and Classify Visitors' Origins," *Electronics*, vol. 11, no. 6, p. 835, 2022.

[26] M. Vanhoef, C. Matte, M. Cunche, L. S. Cardoso, and F. Piessens, "Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms," in *Proceedings of the 11th ACM on Asia conference on computer and communications security*, 2016, pp. 413–424.

[27] M. Ribeiro, N. Nunes, V. Nisi, and J. Schöning, "Passive Wi-Fi monitoring in the wild: a long-term study across multiple location typologies," *Personal and ubiquitous computing*, vol. 26, no. 3, pp. 505–519, 2022.

[28] M. Ribeiro, B. Galvão, C. Prandi, and N. Nunes, "Passive Wi-fi monitoring in public transport: a case study in the Madeira Island," *arXiv preprint arXiv:2006.16083*, 2020.

[29] J. Martin, T. Mayberry, C. Donahue, L. Foppe, L. Brown, C. Riggins, E. C. Rye, and D. Brown, "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.

[30] "Smart living for smart cities," *Smart Living for Smart Cities*, pp. 3–70, 2020.

[31] G. J. W. Kathrine and C. W. Joseph, "Attacks, vulnerabilities, and their countermeasures in wireless sensor networks," in *Deep Learning Strategies for Security Enhancement in Wireless Sensor Networks*. IGI Global, 2020, pp. 134–154.

[32] Z. Qin, P. Zhao, T. Zhuang, F. Deng, Y. Ding, and D. Chen, "A survey of identity recognition via data fusion and feature learning," *Information Fusion*, vol. 91, pp. 694–712, 2023.

[33] M. Uras, E. Ferrara, R. Cossu, A. Liotta, and L. Atzori, "MAC address de-randomization for WiFi device counting: Combining temporal-and content-based fingerprints," *Computer Networks*, vol. 218, p. 109393, 2022.

[34] L. Pintor and L. Atzori, "Analysis of Wi-Fi Probe Requests Towards Information Element Fingerprinting," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 3857–3862.

[35] T. He, J. Tan, and S.-H. G. Chan, "Self-Supervised Association of Wi-Fi Probe Requests Under MAC Address Randomization," *IEEE Transactions on Mobile Computing*, pp. 1–14, 2022.

[36] C. Spiess, "Is that traffic light tracking you? a case study on a municipal surveillance technology in Seattle," *IEEE Transactions on Technology and Society*, vol. 2, no. 1, pp. 15–19, 2021.

[37] What is Wi-Fi MAC randomization and how does it handle privacy? Accessed: 28-02-2023. [Online]. Available: https://www.extremenetworks.com/extreme-networks-blog/wi-fi-mac-randomization-privacy-and-collateral-damage/

[38] D. Zahavi. The MAC address is going away. NOW what? Accessed: 28-02-2023. [Online]. Available: https://levl.tech/resources/mac-address-is-going-away-now-what

[39] S. Nayak. Randomized and Changing MAC (RCM). Accessed: 28-02-2023. [Online]. Available: https://blogs.cisco.com/networking/randomized-and-changing-mac-rcm

[40] J. Zuniga, C. Bernardos, and A. Andersdotter, "MAC address randomization," Internet Requests for Comments, RFC Editor, RFC draft, October 2022. [Online]. Available: https://datatracker.ietf.org/doc/html/draft-ietf-madinas-mac-address-randomization-04

[41] "IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Preassociation Discovery," Piscataway, NJ, Tech. Rep. 802.11aq, 2018.

[42] "IEEE Std 802-2014 (Revision to IEEE Std 802-2001) - IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture," Tech. Rep., 2014.

[43] "IEEE Std 802c-2017 (Amendment to IEEE Std 802-2014 as amended by IEEE Std 802d-2017) – IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture–Amendment 2: Local Medium Access Control (MAC) Address Usage," Piscataway, NJ, Tech. Rep. 802c-2017, 2017.

[44] "Use private Wi-Fi addresses on iPhone, iPad, iPod touch and Apple Watch." [Online]. Available: https://support.apple.com/en-au/HT211227

[45] "MAC Randomization Behavior." [Online]. Available: https://source.android.com/docs/core/connect/wifi-mac-randomization-behavior

[46] "How to Turn On or Off Random Hardware MAC Addresses for Wi-Fi in Windows 10." [Online]. Available: https://www.tenforums.com/tutorials/39022-turn-off-random-hardware-mac-addresses-wi-fi-windows-10-a.html

[47] A. Andersdotter, "RCM TIG Draft Report Outline," 802.11 RCM group contribution, IEEE, 2019.

[48] H. Krawczyk and P. Eronen, "IETF RFC 5869: HMAC-based extract-and-expand key derivation function (HKDF)," Tech. Rep., 2010.

[49] M. Vanhoef and F. Piessens, "Key reinstallation attacks: Forcing nonce reuse in WPA2," in *Proceedings of the 24th ACM Conference on Computer and Communications Security (CCS)*. ACM, 2017.

[50] “Wireless LAN medium access control (MAC) and physical layer (PHY) specification,” Piscataway, NJ, Tech. Rep. 802.11, 2020.

[51] “IEEE Standard for Information technology–Telecommunications and information exchange between systems - Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Fast Initial Link Setup,” Piscataway, NJ, Tech. Rep. 802.11ai, 2016.

[52] “IEEE Standard for information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 6: Medium Access Control (MAC) Security Enhancements,” Piscataway, NJ, Tech. Rep. 802.11i, 2004.

[53] “IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Protected Management Frames,” Piscataway, NJ, Tech. Rep. 802.11w, 2009.

[54] “IEEE Draft Standard for Local and metropolitan area networks-Media Access Control (MAC) Security Amendment 4: MAC Privacy protection,” Piscataway, NJ, Tech. Rep. 802.1Aedk, 2023.

[55] “Wi-Fi Alliance EasyConnect.” [Online]. Available: https://www.wi-fi.org/discover-wi-fi/wi-fi-easy-connect

[56] “MADINAS MAC Adress Randomization gituhub page.” [Online]. Available: https://github.com/ietf-wg-madinas/draft-ietf-madinas-mac-address-randomization

[57] “Randomized and changing MAC address use cases.” [Online]. Available: https://datatracker.ietf.org/doc/draft-ietf-madinas-use-cases/03/

[58] R. Droms, “Dynamic Host Configuration Protocol,” Internet Requests for Comments, RFC 2131, 1997.

[59] B. Croft and J. Gillmore, “BOOTSTRAP Protocol (BOOTP),” Internet Requests for Comments, RFC 951, 1985.

[60] R. Droms, “Dynamic Host Configuration Protocol,” Internet Requests for Comments, RFC 1531, 1993.

[61] ——, “Dynamic Host Configuration Protocol,” Internet Requests for Comments, RFC 1541, 1993.

[62] S. C. T. Lemon, “Encoding Long Options in the Dynamic Host Configuration Protocol (DHCPv4),” Internet Requests for Comments, RFC 3396, 2002.

[63] B. S. T. Lemon, “Node-specific Client Identifiers for Dynamic Host Configuration Protocol Version Four (DHCPv4),” Internet Requests for Comments, RFC 4361, 2006.

[64] C. P. J. Arkko, “IANA Allocation Guidelines for the Address Resolution Protocol (ARP),” Internet Requests for Comments, RFC 5494, 2009.

[65] P. J. N. Swamy, G. Halwasia, “Client Identifier Option in DHCP Server Replies,” Internet Requests for Comments, RFC 6842, 2013.