SCTE·ISBE CABLE-TEC EXPO 2019 NEW ORLEANS, LA SEPT. 30-OCT. 3

2019 Fall Technical Forum SCTE·ISBE • NCTA • CABLELABS

MAC Randomization in Mobile Devices

A Technical Paper prepared for SCTE•ISBE by

Carol Ansley
Senior Counsel
CommScope. Inc.
3871 Lakefield Dr. Suwanee GA 30024
+1 404-229-1672
carol.ansley@commscope.com

Table of Contents

List of Figures

List of Tables

Introduction

Privacy is an important topic around the world as consumers discover that much of their personal information has been leaking away from them, particularly with the introduction of technology such as smartphones. For example, mobile devices, such as smartphones, with Wi-Fi® provide a way for companies large and small to track those who visit their premises or even just stroll by on the sidewalk. Privacy advocates have been calling for improved privacy for mobile devices and mobile device companies recently have begun to change their devices' behavior. This paper will discuss the original problem, which is still widespread, and cover some of the improvements that have been implemented or at least promised in upcoming device releases. The paper will close with a discussion of the areas that still need attention.

Wi-Fi and MAC Address Randomization

Privacy and security have been getting headlines and provoking interest in the general public as well as in the business community. Most people carry smart phones and other Wi-Fi enabled devices, but they seldom consider just having those devices on their person as a potential privacy risk. To understand more about this issue, we will first review the features within Wi-Fi that can expose trackable information.

1. A Quick Wi-Fi Primer

IEEE® 802.11 networks, a.k.a. Wi-Fi networks, consist of client devices, often called *stations*, and *access points*, or APs. 802.11 stations are typically mobile devices like tablets, laptops or mobile phones, but they can also be relatively stationary devices such as set top boxes or printers. Access points provide wireless connections between a group of stations and a wider network, such as the Internet. Access points can be standalone products but are also often integrated into more complex devices, such as broadband access gateways. Access points also often provide routing features beyond simple wireless connectivity. Stations connect to access points for access to a wider, usually wired, network through a process called *association*. Once a station is *associated* with an access point, that station can send traffic through the AP to the Internet, for example, and receive content from the Internet.

One access point device may support multiple wireless networks. Service Set Identifiers (SSIDs) are used to identify different wireless networks. An access point device may support multiple SSIDs on a single channel and it may support SSIDs on several channels spread over multiple bands. An SSID is not guaranteed to be unique, so stations (and their users) need to ensure that they associate with the intended network.

Figure 1 – Access Point and Stations

2. How does Wi-Fi make my location and activities less private?

When 802.11 was first designed, a client device or non-AP station was expected to be a device that could move and hence would need the ability to discover new access points as it moved out of range of its current access point. As part of the general protocols of Wi-Fi, a station is given tools to discover what access points serving SSIDs might be near so that it can make a good choice for its next association. A station can send out a query to the surrounding 802.11 wireless world to find out what access points are nearby and access points hearing that request can respond.

This feature is called the Probe Request/Response exchange within 802.11. A station sends a Probe Request identifying itself and asking access points that receive the message to respond with their network information. A Probe Request can be targeted to a specific SSID if a station is looking for a specific network, but more commonly a station sends out a broadcast probe. Any access point hearing a broadcast Probe Request can respond with a Probe Response. The Probe Response identifies the SSID and access point and provides information to the station about the SSID's configuration.

A station may have to repeat a Probe Request on different frequencies to cover the band or bands on which it is operating. For example, a new station might start by sending a probe on channel 1. If a nearby access point is using channel 6, it will not hear that request. If the station does not receive a response, it could move on to another channel, such as channel 6, and send out another probe. The example access point on channel 6 would hear the second probe and respond. The station, once it received the probe response, would determine whether it wanted to continue to probe on other channels, or if it wanted to begin an association exchange with that AP.

2.1. Does This Affect Me?

If you were to open up the Wi-Fi menu in your smartphone or laptop in a new place, you would see the device populate that Wi-Fi menu with available wireless networks. Sometimes it may take a moment or two before the list is complete and the network you expect to see is shown in the menu. You are watching the output of many probe request/probe response exchanges between your device and the surrounding APs to determine the available SSIDs.

All of these exchanges happen over the air for anyone to receive if they are monitoring that channel. Below is an example exchange captured in a wireless lab.

117	4.173079	6a:f9:09:95:1a:76	Broadcast	802.11	151	Probe Request, SN=3421, FN=0, Flags=.....C, SSID=Wildcard (Broadcast)
118	4.173961	ArrisGro_b7:fd:16	6a:f9:09:95:1a:76	802.11	427	Probe Response, SN=143, FN=0, Flags=.....C, BI=200, SSID=SR2-Allion

Figure 2 – Probe Request/Probe Response

In message 117, a station sent out a broadcast probe request and was answered with a probe response in message 118 by an AP with the SSID SR2-Allion.

This exchange was captured with a standard scanning utility running in a laptop. In the field, such captures are simple to set up. In areas with a large number of clients and access points, the airwaves can be very busy as multiple clients probe multiple APs. For example, on a busy market street one of the Colorado students recorded 137 clients probing 372 times generating over 4000 probe responses in about 35 seconds.

The information in these frames can allow a mobile device's location to be tracked. Particularly, included in the probe request message is the source address of the originating station so that an access point can address the probe response to that station.

Frame Control	Duration	Broadcast or target SSID	Originating STA MAC	Sequence Number	Frame Body	Frame Check Sequence

Figure 3 – Management Frame Format

The originating MAC address is repeated in the Probe Response message sent by the access point so that the station that sent the Probe Request can identify the responses it triggered. To go on, we need to understand a bit more about MAC addresses and how they work.

3. Device Identifiers: EUIs, MAC addresses, OUIs and the Local Address Space

In early networks, engineers quickly figured out that it was important to be able to uniquely identify devices on the network. An Extended Unique Identifier (EUI) is a term for a sequence or code that uniquely identifies a device or one of its network interfaces. A MAC address is an EUI-48 code. It consists of 48 binary digits, usually written as a string of 12 hexadecimal digits. The first 6 hexadecimal digits of a MAC address are defined to assist in network management. These digits can uniquely identify the manufacturer of the device. The IEEE Registration Authority manages the distribution of 6 digit OUIs (Organizational Unique Identifiers) to companies. An OUI begins a device's MAC address and is followed by another 6 digits so that the MAC addresses used by any one company will be unique and will not overlap with the MAC addresses used by other companies. The IEEE publishes a database on the Internet allowing any device with a valid OUI MAC address to be traced back to its manufacturer.

The MAC address space is even larger than the part given over to manufacturer OUIs. When MAC addresses were developed, the designers realized that in some circumstances a piece of equipment might need a local or user assigned address for a network interface. The second bit of the second hex digit can be set to 1 to indicate that a MAC address is a locally assigned address and not part of the manufacturer OUI scheme.

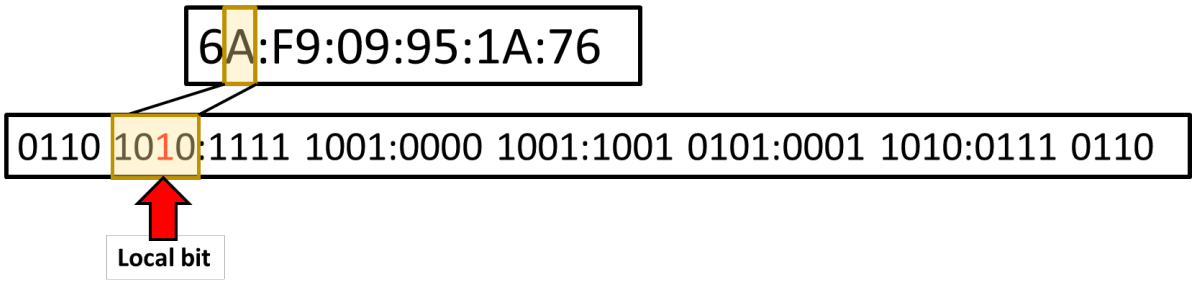

Figure 4 – Local Bit Within MAC Address

If the local bit is set to one then the MAC address has been locally assigned either by the device itself or by a local network authority and is not guaranteed to be unique. Recent standards activity has attempted to add more structure to the local bit by defining its significance in terms of the other nearby bits.

The Local Address Study group within 802.1 proposed a further refinement to the use of the Local bit by adding address classifications based on the next two bits in a MAC address.

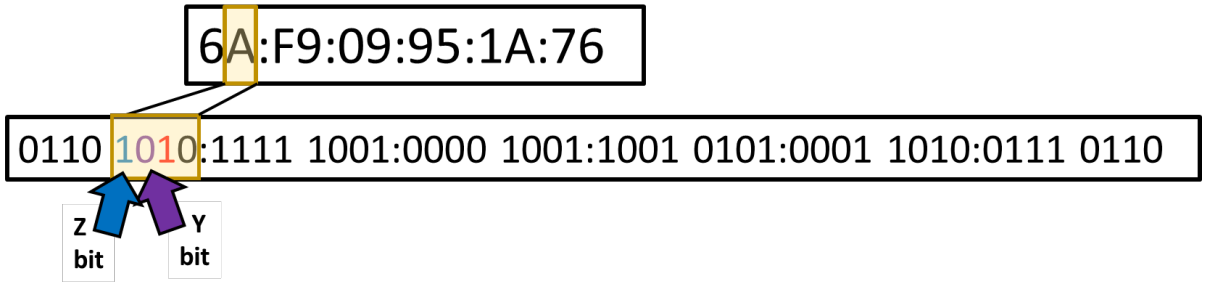

Figure 5 – Y and Z Bits Within MAC Address

These bits are denoted the Y and Z bits and allow the second hex digit to indicate the purpose of the local MAC address assignment. An optional specification was developed called the Structured Local Address Plan or SLAP.

Table 1 - SLAP defined Local MAC Address Quadrants

Second Hex Digit of MAC address	Name
2 (0010)	Administratively Assigned Identified (AAI)
6 (0110)	Reserved
A (1010)	Extended Local Identifier (ELI)
E (1110)	Standard Assigned Identifier (SAI)

An address using AAI (2) indicates that the address is assigned by a local administrator according to any local arbitrary scheme. In a domain with AAI addresses, the local administrator is responsible to ensure uniqueness of assigned addresses. An address in the reserved space, indicated by a 6, can be used for a similar purpose but the specification notes that later versions of SLAP may define other uses for that address space.

An address using ELI (A) indicates that the rest of the leading 6 digit block is a Company ID (CID). CIDs are assigned by the IEEE-RA just like OUIs, and can be used by devices in a similar way.

Finally, an address using SAI (E) indicates that the address has been assigned by a protocol specified within an IEEE 802 standard. Such a standard may make use of the address space to define other more sophisticated uses as appropriate for that standard. While 802.11 has not yet made use of this feature, it is a possibility for later revisions of the standard.

Now that we've reviewed the technical background, lets return to the problem and activity to reduce its effects.

4. Efforts to Increase Location Privacy

Privacy advocates noticed that as people adopted smart phone technology en masse, those users at the same time were at risk of losing some of their privacy. A person who walked through a mall, for example, could be tracked as they walked along, if even some of the mall's access points were networked together. Even if that person did not log into the mall's Wi-Fi network, their progress through the space could be tracked, where they stopped, what shops they entered, etc. simply by tracking the probe requests from their phone as the walked along. If that person had ever associated their device with the mall's Wi-Fi network in the past, their identity could be attached to the information from their current visit by associating the MAC address used in the prior association and the MAC address appearing in the probe request messages.

This situation was noted by many members of the wireless industry and several companies began to make changes to the behavior of their Wi-Fi enabled devices. The IEEE-SA 802.11 standards group also considered what changes could be made to improve the privacy of mobile device users.

The next sections review the current (as of 2019) behavior of mobile devices seen in the field. This data comes from testing in a wireless lab on test units and from data graciously shared by Prof. Jim Lansford and his University of Colorado wireless studies class. The Univ. of Colorado data was gathered in and around their campus. I also had limited access to records of associated clients from a Ruckus deployment.

4.1. Apple® Devices

Apple's devices are very common in the Wi-Fi world. A survey of several recent Wi-Fi deployments found Apple devices making up at least 45% of the devices seen. Because of the large number of Apple devices present in most Wi-Fi deployments, it is instructive to examine the behavior of Apple devices

Apple has taken a proactive approach to location privacy by defaulting their mobile devices to use randomized MAC addresses with the local bit set in non-associated probe requests. In lab testing, unassociated Apple devices were seen to periodically change their randomized MAC addresses as well to keep a randomized address from becoming a new default MAC address. When a user selects an SSID for association, an Apple device was seen to shift back to communicating using a MAC address with its Apple OUI, including in new probe requests.

Looking at data graciously provided by Prof. Jim Lansford and his University of Colorado Wireless students, many probe requests with local MAC addresses can be identified as originating from Apple devices. Additional fields within a probe request can indicate the device manufacturer directly as well as other information such as the Wi-Fi chip set manufacturer. The local MAC addresses originating from Apple devices were seen to set the local bit to one (2, 6, A and E in the second hex digit).

4.2. Windows 10® Devices

Windows 10 provides a set of features related to location privacy and MAC randomization, though they are dependent upon the hardware of the laptop for the exact features supported. Randomized MAC addresses are used for probing while a laptop is unassociated by default, though a user can disable that feature. Also, Windows 10 allows a user to decide whether or not to use randomized MAC addresses when associating with a new SSID. Windows 10 allows the user to also decide whether to keep the same randomized MAC address every time the laptop associates with that network, or if it should change to a new randomized MAC address every 24 hours.

In the real world data discussed earlier, some devices using MAC addresses with the local bit set identified themselves as Windows devices.

4.3. Android® Devices

Android has approached location privacy more cautiously than some others. MAC randomization was added as a developer option at first, and was moved to become a default for non-associated devices only in Release 9 released in late 2018. In Release 10 due out later in 2019, according to a media report, Android will also offer the ability for a device to present a randomized MAC address when associating with an SSID. The device will use the same randomized MAC address for later associations with that SSID.

Because Android allows much more product differentiation than iOS, not all Android devices have implemented access to the same level of privacy features.

4.4. 802.11's activities

The IEEE-SA 802.11 Working Group has been concerned about this topic for some time. The 802.11aq amendment added a new feature to incorporate several changes needed to improve location privacy and provide standards-based recommendations.

In the 802.11aq amendment released in 2018, a new MIB variable called dot11MACPrivacyActivated was created in the amendment to support increased location privacy for Wi-Fi devices. The PrivacyActivated mode requires several changes to reduce the possibility that a station could be tracked easily when it is activated. MAC randomization for at least pre-association messages is required to be in accordance with IEEE SA 802-2014 and 802c-2017, which is to say at least using the local bit address space. The amendment also points out that if only a MAC address is changed, but the associated message sequence count is not randomly reset within that device, a determined observer could still track a device. Because of this possibility, PrivacyActivated mode requires the sequence number to be reset randomly as well. Similarly, the seeds used within the PHY DATA scrambler can also be used to track a device and are also required to be reset randomly when the dot11MACPrivacyActivated variable is true. Finally, the amendment pointed out that if a station probes for a specific SSID instead of using the broadcast address, that activity could also be used to track the station. It is important to note that these changes to the MAC address and sequence number and scrambler seed can only be effective if the device has not started a stateful exchange. For instance, if a station has started an association exchange that has not completed, an intervening MAC address change by the station will cause the association to fail.

The 802.11 Working Group has continued to have discussions in this area even after the adoption of amendment 11aq. A Topic Interest Group is currently examining the issues around randomized and changing MAC addresses (RCM TIG). The reference document, 11-19-0588-01-0rcm, has a very complete and up to date summary of the various activities around the standards-setting world in this area. Within the 802.11 standards processes, the TIG may issue a report and dissolve, or it may lead to the formation of a Study group to determine if a standards amendment would be appropriate.

5. Market adoption of these features

Looking at the data collected from the field, non-associated devices sending probe requests are overwhelmingly using randomized MAC addresses. The majority of those randomized MAC addresses are also using the local bit to signal that they are local MAC addresses.

When one looks at associated devices, the percentage of associations made with randomized MAC addresses is still very low. In a sample of 2 populations totaling 10,000 associated devices, the average percentage of randomized MAC addresses was 1-2%.

One inference that could be drawn from this is that users tend to make use of the default configuration of their mobile devices. When mobile devices switched to randomized MAC addresses for probing by default, that change was picked up quickly and transparently by most users. When the feature is not active by default, and instead requires some knowledge and effort on the part of the end user, it is much less likely to be widely used.

6. Is it enough?

Many privacy advocates appreciate that the current state of affairs is better than it was, but they are still concerned that it does not prevent determined attackers from tracking devices, and by extension their

users. Researchers have pointed out that it is not too difficult for a determined person to still track a device generally, and in some cases associate it back to a specific user.

For example, if one considers a device that changes its MAC address to different random assignments, but does so with a regular pattern, it is not hard to understand that such a device can be tracked. If a first MAC address appears for a few seconds with the device travelling down a hallway, then a second MAC address appears moving down the same hallway with the first MAC address no longer being detected, it would be reasonable to associate the two MAC addresses together. Also, most devices spend a lot of time already associated and there are challenges, as noted before, to making randomization changes while in a stateful relationship. See reference papers from the Univ. of Lyons, iMinds-Distrinet and the US Naval Academy for more in depth research into issues that they researched with respect to MAC randomization.

6.1. Why is MAC Randomization a Concern?

MAC Randomization may be a concern for some parties because a device's MAC address has been used as a definitive way to identify a device for many purposes and that assumption is not necessarily valid in the future.

A possible future approach to increasing privacy would be to constantly change the MAC address and other trackable parameters of a mobile device both while it is un associated and after association. For a single short session at a coffee shop, a device using a different MAC address on each association is probably not a problem. The drawback to going fully anonymous is that many more advanced features depend upon recognizing a device and associating a known history or known authorization with that device. The Wireless Broadband Alliance surveyed the possible features impacted by such a change and found several concerning issues that they raised with the 802.11 Working Group.(footnote to WBA liaison)

For example, a parent may want to place parental control restrictions on the sites a tablet can access. If the tablet constantly presents different MAC addresses to the home network, one can only provide consistent parental control functionality by forcing the user to log in each time the MAC address changes. Similarly, many other network features depend on the larger network being able to associate a specific device with certain permissions or attributes. In a simpler case, an access point or router may have a limit to the number of devices that can associate. Such a limit might use device MAC addresses under the soon to be obsolete assumption that a device's MAC address is a unique identifier. Having a single device register multiple times with different MAC addresses might exhaust that device limit.

Another class of features that often use device MAC addresses is wireless network management. A network manger may store a history and profile for a device to facilitate client steering or band selection decisions. A device might be found to have poorer performance when associated a 5 GHz than at 2.4 GHz. If the network can store that information, then the device will not be steered to a 5GHz channel at a later association. If the network cannot identify that client, then it must rediscover that performance problem with each repeat association. Similarly, a help desk or customer support desk will have to rely on the user providing information, versus being able to acquire information directly from the device over the network.

Another area where constantly changing MAC address may cause user dissatisfaction is user authentication. A user may accept the need to authenticate themselves to the network once, but not repeatedly every time their device's MAC address randomly changes. Even with the current proposal to use the same randomized MAC address with a single SSID, that may still cause issues if the extended network the user wants to access uses multiple SSIDs. Under current proposals and behavior, the same device would show up as multiple devices.

Finally, for some stationary devices, such as a set top box or printer, MAC address randomization may be unnecessary. Some devices are not expected or intended to move from one network to another. Other devices such as a rental scooter or luggage cart, may also not wish to support MAC randomization as there is a reason to readily track their location.

Conclusion

MAC randomization features have been spreading through the landscape of wireless devices over the last couple of years with the goal of providing enhanced location privacy for the users of mobile devices. Not everyone is in agreement that these features have done enough to ensure location privacy for the users of those devices and it is a given that these features will continue to evolve.

As randomized MAC addresses spread into more devices and expand their use into other states beyond pre-association, existing wireless networking features may be affected in ways that will cause user dissatisfaction. MAC randomization features have not yet caused serious problems with existing services, but questions are being raised about the possible effects of the next stages of randomization on existing network level features. Standards bodies are also working in this area, but the marketplace has been moving quickly.

The goal of this paper is to ensure that the development of MAC randomized wireless devices is highlighted and to point out areas that this technology may affect in the future.

Abbreviations

AAI	Administratively Assigned Identifier
AP	Access Point
ELI	Extended Local Identifier
EUI	Extended Unique Identifier
IEEE	Institute of Electrical and Electronics Engineers, Inc.
MAC	Medium Access Control
MIB	Management Information Base
OUI	Organizational Unique Identifier
SAI	Standard Assigned Identifier
SG	Study Group
SLAP	Structured Local Address Plan
STA	Station (802.11 nomenclature)
TG	Task Group
TIG	Topic Interest Group
WG	Working Group

Bibliography & References

Part 11: Wireless LAN Medium access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 5: Preassociation Discovery, IEEE Computer Society, 2018.

IEEE RA Guidelines for Use of EUI, OUI, and CID, IEEE SA, 2017.

11-18-1671-00-0arc, IEEE 802.11 TG ARC, 2018. (WBA liaison and notes)

11-19-0588-01-0rcm, IEEE 802.11 TIG RCM, 2019. (review of random and changing literature)

"Why MAC Address Randomization is not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms," ACM, 2016.

"A Study of MAC Randomization in Mobile Devices and When it Fails," U.S. Naval Academy, 2017.

