MAC address anonymization

Mickael Kovel[1]
[1] *MA1 Cybersecurity, Université Libre de Bruxelles (ULB), Bruxelles, Belgique*
(Dated: 2024-11-11)

CONTENTS

I. INTRODUCTION

bla bla bla

II. WHAT IS A MAC ADDRESS ?

In the context of networking, we need to identify devices on a network. To exchange information between two devices, we point to an identifier. In the context of the internet, we use IP addresses to identify devices. However, IP addresses are dynamic and can change each time the device connects to the network. This is why we use MAC addresses to uniquely identify devices, in the same way that we use postal addresses to identify houses and send mail.

We can define a MAC address as a unique identifier assigned to a network interface controller (NIC) for use as a network address in communications within a network segment.

A. How is a MAC address provided ?

MAC addresses are provided by the manufacturer of the network interface controller (NIC). They are typically assigned at the factory and are hardcoded into the hardware. They cannot be changed by the user and are globally unique.

1. What is the structure of a MAC address ?

A MAC address is a 48-bit number that is typically represented as a 12-digit hexadecimal number. The first 24 bits (6 digits) are the Organizationally Unique Identifier (OUI), which identifies the manufacturer of the NIC. The last 24 bits (6 digits) are the device identifier, which is assigned by the manufacturer. The OUI is assigned by the Institute of Electrical and Electronics Engineers (IEEE) and is unique to each manufacturer. The device identifier is unique to each device and is assigned by the manufacturer.

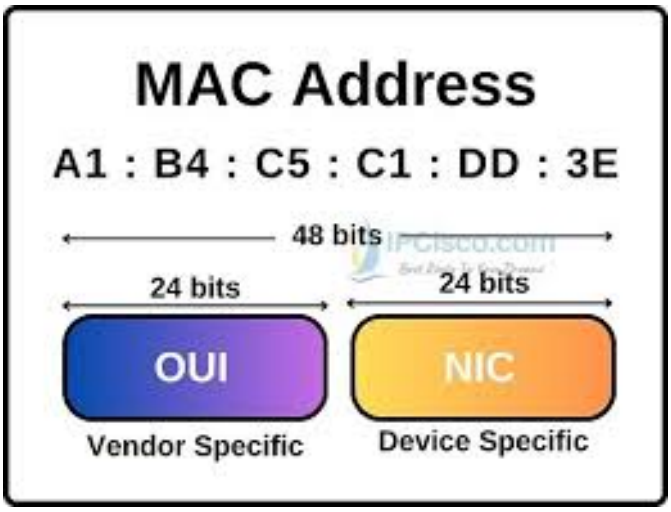

FIGURE 1 – MAC address structure [1]

B. What are the privacy implications of a MAC address ?

C. Where are MAC addresses used ?

1. *Wi-Fi - Ethernet*

In the context of computer networks, MAC addresses are used to identify devices on a local network. They are linked to the IP address of a device and are used to route data packets to the correct destination. The Address Resolution Protocol (ARP) is used to associate an IP address with a MAC address.

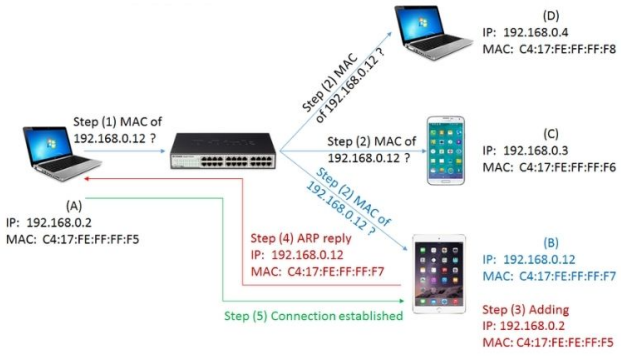

FIGURE 2 – ARP protocol [2]

We can see in the figure 2 that the MAC address is used to identify the device in the network.

2. *Bluetooth*

Bluetooth devices do not use MAC addresses in the same way as Wi-Fi devices. Instead, they use a 48-bit Bluetooth device address (BD_ADDR) that is similar to a MAC address. BD_ADDRs are also globally unique and could be used to track devices in the same way as MAC addresses.

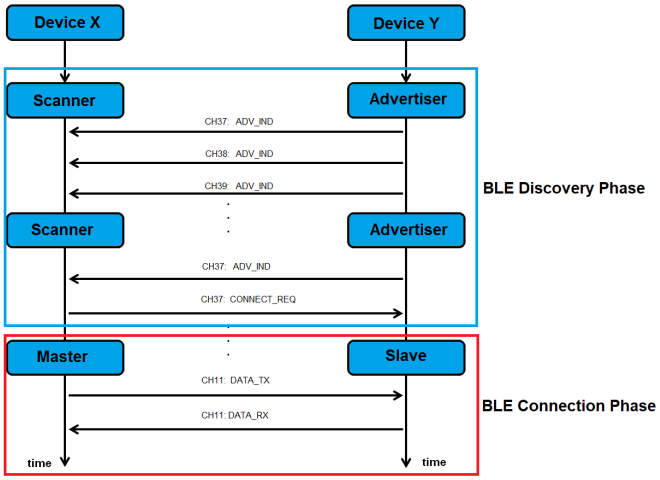

FIGURE 3 – BTLE Link Layer [3]

We can see in the figure 3 that even during the discovery phase, there is exchange in both directions, which means that the BT_ADDR have already been exchanged before even connecting to the device.

III. WHAT DOES THE LAW SAY ?

A. GDPR

"The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and

technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.“ [4]

B. European Convention on Human Rights

34. “Individual applications The Court may receive applications from any person, non- governmental organisation or group of individuals claiming to be the victim of a violation by one of the High Contracting Parties of the rights set forth in the Convention or the Protocols thereto. The High Contracting Parties undertake not to hinder in any way the effective exercise of this right.“
35. “Admissibility criteria
    1. [...]
    2. The Court shall not deal with any application submitted under Article 34 that
       (a) is anonymous ; or [...]“ [5]

C. Convention 108+

18. “The notion of “identifiable” refers not only to the individual’s civil or legal identity as such, but also to what may allow to “individualise” or single out (and thus allow to treat differently) one person from others. This “individualisation” could be done, for instance, by referring to him or her specifically, or to a device or a combination of devices (computer, mobile phone, camera, gaming devices, etc.) on the basis of an iden- tification number, a pseudonym, biometric or genetic data, location data, an IP address, or other identifier. The use of a pseudonym or of any digital identifier/ digital identity does not lead to anonymisation of the data as the data subject can still be identifiable or individualised. Pseudonymous data is thus to be considered as personal data and is covered by the provisions of the Convention. The quality of the pseud- onymisation techniques applied should be duly taken into account when assessing the appropriateness of safeguards implemented to mitigate the risks to data subjects.“
19. “Data is to be considered as anonymous only as long as it is impossible to re-identify the data subject or if such re-identification would require unreasonable time, effort or resources, taking into consideration the available technology at the time of the processing and technological developments. Data that appears to be anonymous because it is not accompanied by any obvious identifying element may, nevertheless in particular cases (not requiring unreasonable time, effort or resources), permit the identification of an individual. This is the case, for example, where it is possible for the controller or any person to identify the individual through the combination of different types of data, such as physical, physiological, genetic, economic, or social data (combination of data on the age, sex, occupation, geolocation, family status, etc.). Where this is the case, the data may not be considered anonymous and is covered by the provisions of the Convention.“
20. “When data is made anonymous, appropriate means should be put in place to avoid re-identification of data subjects, in particular, all technical means should be implemented in order to guarantee that the individual is not, or is no longer, identifiable. They should be regularly re-evaluated in light of the fast pace of technological development.“[6]

D. Belgian Law

Articles 101, 134, and 164 of the Belgian Act of 30 July 2018 on the protection of individuals with regard to the processing of personal data stipulate that the personal data referred to in Articles 99, 132, and 162 must be anonymized before they can be accessed. These articles primarily concern the processing of personal data for historical, scientific, or statistical purposes.

99. “outlines conditions under which personal data from intelligence and security services can be consulted for such purposes. The consultation is authorized only if it does not conflict with the service’s mandate, obligations under the Act of 30 November 1998, ongoing investigations, or international relations. Any request for further processing of such data for other purposes will be refused unless deemed legitimate by the concerned service.“
132. “allows the consultation of personal data held by authorities or appeal boards for historical, scientific, or statistical purposes. However, it is conditional on ensuring that it does not harm any interests protected under the Act of 11 December 1998, particularly regarding the confidentiality and protection of personal data.“
162. “similarly authorizes the consultation of personal data from CUTA (the Coordination Unit for Threat Analysis) for historical, scientific, or statistical purposes. Again, this is contingent on ensuring no harm to CUTA’s mandate, ongoing investigations, or in-

ternational relations. Any requests for further processing of such data for different purposes will be denied unless the processing is considered legitimate and does not interfere with the protected interests.“

[7]

IV. WHO CAN ACCESS A MAC ADDRESS ?

In fact, anyone can access a MAC address. It is transmitted in clear text over the network and can be easily intercepted by anyone with the right tools. The privacy risks associated with MAC addresses are not limited to a specific group of people or organizations. However, the real issue arises when a specific location is associated with a MAC address at a particular time, such as when connecting to a Wi-Fi network. The problem becomes even more serious when this data is recorded, as recording real-time location information can enable tracking of an individual over an extended period.

The critical question then becomes : who would have an interest in recording this data, and why ?

A. Internet Service Provider

Internet Service Providers (ISPs) can access MAC addresses when users connect to their network. By definition, they have access to all the data that passes through their network, including MAC addresses. This data can be used for various purposes, such as network management, troubleshooting, and targeted advertising. However, ISPs are subject to data protection regulations, and the use of MAC addresses for tracking or profiling purposes is generally prohibited.

B. Audience Measurement

Because of the probes and probe requests, MAC addresses can be collected by anyone with a Wi-Fi-enabled device. This includes audience measurement companies, which collect data on the behavior of users in public spaces. So audience measurement companies can access MAC addresses when users walk by their Wi-Fi sensors. This data can be used to analyze foot traffic, measure the effectiveness of advertising campaigns, and provide insights to retailers and other businesses. However, the use of MAC addresses for audience measurement is subject to data protection regulations, and the data must be anonymized to protect the privacy of users.

V. WHAT ARE THE RISKS OF USING A MAC ADDRESS ?

A. Tracking

B. Profiling

C. Re-identification

D. Spoofing

VI. HOW TO ANONYMIZE A MAC ADDRESS ?

A. Hashing

B. Truncation

C. Encryption

VII. NOTABLE INCIDENTS

A. Privacy violation

B. Google Street View

C. Retail and public space tracking scandals

D. Data Leaks and Security Breaches Involving MAC Data

E. Attacks Exploiting MAC Addresses

F. WhatsApp Security Vulnerability

En 2012, une faille de sécurité a été découverte dans WhatsApp, permettant à un attaquant d'usurper l'identité d'un utilisateur si l'adresse MAC de son appareil pouvait être obtenue. À cette époque, WhatsApp n'utilisait pas de chiffrement de bout en bout et s'appuyait sur l'adresse MAC de l'appareil pour authentifier les utilisateurs. En exploitant cette faille, un attaquant pouvait : Se connecter au compte WhatsApp de la victime depuis un autre appareil. Lire et envoyer des messages comme si c'était la victime. Exploitation de la faille : Les attaquants pouvaient obtenir l'adresse MAC via différents moyens, notamment par des attaques sur des réseaux Wi-Fi publics ou via des applications malveillantes capables de lire l'adresse MAC d'un appareil. Ensuite, ils pouvaient utiliser des outils spécifiques pour se connecter à WhatsApp en se faisant passer pour la victime. Conséquences et correction : Ce problème a mis

en lumière les risques d'utiliser des identifiants statiques comme les adresses MAC pour authentifier les utilisateurs. WhatsApp a finalement renforcé ses mesures de sécurité, notamment en adoptant le chiffrement de bout en bout, et a cessé d'utiliser l'adresse MAC pour l'authentification. Ce cas montre l'importance de ne pas se fier aux adresses MAC pour authentifier les utilisateurs, en raison de la facilité avec laquelle elles peuvent être usurpées.

bla bla

VIII. CONCLUSION

A. Why should we anonymize a MAC address ?

End

[1] IPCisco. What is a mac address ? — mac address lookup — mac examples. Accessed : 2024-11-11.

[2] Stanley Arvey. Address resolution protocol (arp) : Everything you should know about, 2022. Accessed : 2024-11-11.

[3] EmbeddedCentric. Introduction to bluetooth low energy / bluetooth 5, n.d. Accessed : 2024-11-11.

[4] European Parliament and Council. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation), 2016. Official Journal of the European Union, L 119, 4.5.2016.

[5] Council of Europe. Article 34,35 of the convention for the protection of human rights and fundamental freedoms, 1950. Official Journal of the European Union, 4 November 1950, as amended.

[6] Council of Europe. Convention 108+ for the protection of individuals with regard to the processing of personal data, 2018. www.coe.int/dataprotection.

[7] Belgian Government. Act on the protection of natural persons with regard to the processing of personal data. Unofficial translation, 2018. Unofficial translation, Belgian Official Journal, 05 September 2018, Consolidated version (05/09/2018).