Embeded security

MAC Address Anonymization

Mickael Kovel

Université Libre de Bruxelles

10 December 2024

Introduction

- The rise of connected devices in the Internet of Things (IoT) has led to the widespread collection of MAC addresses, which are used to identify devices on a network.
- While these addresses play a crucial role in network communication, they can also be exploited for tracking purposes, raising concerns about user privacy.
- This project explores the implications of MAC address usage from both a privacy and legal perspective, focusing on potential risks and protection mechanisms.
- The goal is to understand how MAC addresses can be anonymized and what legal measures are in place to ensure data protection and respect for privacy.

What is a MAC Address?

- Device identification on a network
  - Using MAC addresses to identify devices
  - Similar to postal addresses for identifying houses
  - Enables communication on a network segment
- Definition
  - A unique identifier assigned to a network interface controller (NIC)
  - Used for communication within a network

How is a MAC Address Provided?

- Provided by the manufacturer of the network interface controller (NIC)
- Typically assigned at the factory and hardcoded into the hardware
- Cannot be changed by the user
- Ensures global uniqueness

What is the Structure of a MAC Address?

- 48 bits long, represented by a 12-digit hexadecimal number
- First part: OUI (Organizationally Unique Identifier)
  - Identifies the NIC manufacturer
  - Composed of 24 bits (6 hexadecimal digits)
- Second part: Device Identifier
  - Assigned by the manufacturer, unique to each device
  - Composed of 24 bits (6 hexadecimal digits)

Bluetooth and MAC Addresses

- Bluetooth devices use a 48-bit Bluetooth device address (BD ADDR)
- Similar to a MAC address, but specifically for Bluetooth devices
- BD ADDRs are also unique and can be used to track devices

Privacy Implications of a MAC Address

- MAC addresses can be used to track devices on a network
- Privacy issues arise if MAC addresses are exposed
- Privacy protection techniques, such as MAC address anonymization

Where Are MAC Addresses Used?

- Wi-Fi and Ethernet networks
  - Used to identify devices on a local network
  - Associated with IP addresses using the ARP protocol
- Used in data exchanges
  - Helps route data packets to the correct destination

What Does the Law Say?

- Role of MAC addresses in the network
  - Using MAC addresses to identify devices on the network

GDPR and MAC Addresses

- The GDPR applies to any information that can directly or indirectly identify a person
- Pseudonymized data remains considered personal data if it can be re-identified
- Anonymized data is excluded from the scope of the GDPR, but anonymization must be irreversible

European Convention on Human Rights

- Right of individual petition before the European Court of Human Rights
- Anonymous requests are inadmissible
- Protection of personal data and the absence of violation of privacy are paramount

Convention 108+ and Identifiability

- The concept of ”pseudonymization” does not lead to data anonymization if the person can still be identified
- Pseudonymized data is considered personal data and protected under the Convention
- Re-identification of data requires suitable technical means and must avoid any risk to privacy

Belgian Law on Data Protection

- Belgian Law of July 30, 2018, on personal data protection
- Personal data must be anonymized before being consulted for historical, scientific, or statistical purposes
- Consultation of data is subject to strict conditions to prevent any breach of confidentiality and data protection

Risks Related to MAC Address

- **Tracking**
- **Audience Measurement**
- **Profiling**
- **Re-identification**
- **Spoofing**

Who Can Access a MAC Address?

- **Anyone can access a MAC address.** It is transmitted in clear text and can be easily intercepted with the right tools.
- The main risk arises when a MAC address is tied to a specific location or time, such as connecting to a Wi-Fi network.
- The privacy risk increases when this data is recorded over time, enabling the tracking of individuals.

Internet Service Providers (ISPs)

- ISPs can access MAC addresses when users connect to their network.
- ISPs have access to all data passing through their network, including MAC addresses.
- This data can be used for network management, troubleshooting, and advertising.
- The use of MAC addresses for tracking or profiling purposes is generally prohibited under data protection regulations.

Audience Measurement Companies

- MAC addresses can be collected by audience measurement companies using Wi-Fi sensors in public spaces.
- This data is used to measure foot traffic and assess advertising effectiveness.
- The use of this data is regulated by data protection laws, and it must be anonymized to protect user privacy.
- Concerns arise when data is used for unethical purposes, such as surveillance or tracking without consent.

Profiling Using MAC Addresses

- MAC addresses allow the creation of detailed user profiles by correlating data such as connection times, locations, and usage patterns.
- This information can reveal personal habits, routines, and preferences.
- Companies use this data for targeted advertising, while malicious actors may exploit it for social engineering attacks.
- This type of profiling is often invisible to users but constitutes a significant invasion of privacy.

Re-identification Risks

- Even randomized MAC addresses can be re-identified when combined with other data sources.
- Usage patterns such as connection duration or frequency can link a temporary address to its original one.
- This compromises the effectiveness of anonymization and exposes users to risks if data is leaked or poorly secured.

MAC Address Spoofing

- Spoofing allows attackers to falsify their MAC address to bypass network access controls or steal identities.
- This is particularly dangerous in systems reliant on MAC addresses for security, such as IoT devices.
- Spoofing illustrates the limitations of using MAC addresses as secure identifiers.

Anonymizing a MAC Address

- **Truncation**
- **Hashing**
- **Encryption**

Truncation

Truncation involves shortening the MAC address by keeping only part of it, such as the first 6 octets (manufacturer's identifier), or a random portion.

- **Advantages:**
  - Simple to implement.
  - Reduces the risk of re-identification as the remaining portion doesn't contain enough identifying information.
- **Disadvantages:**
  - Loss of information if finer identification is needed.
  - May not be sufficient for higher anonymity as the remaining address part may still enable tracing.

Hashing

Hashing applies a mathematical function to the MAC address, producing a fixed-size output (hash). The process is irreversible.

- **Advantages:**
  - Irreversible, meaning the original MAC address cannot be recovered.
  - Protection against rainbow table attacks with the use of salt.
  - Easy to implement when access to the original MAC address is not needed post-anonymization.
- **Disadvantages:**
  - Loss of the original MAC address, which may be problematic if recovery is required for auditing or re-identification.

Salt: A random value added to the MAC address before hashing to enhance security, preventing attackers from using precomputed hashes.

Encryption

Encryption secures the MAC address using symmetric or asymmetric algorithms, allowing recovery with a decryption key if necessary.

- **Recommended Encryption Scheme: DHIES**
  - Diffie-Hellman Integrated Encryption Scheme (DHIES) combines Diffie-Hellman key exchange and symmetric encryption (e.g., AES).
  - Suitable for systems needing high security with controlled access to the original MAC address.

Encryption (contd.)

- **Advantages:**
  - High security with double-layer protection (key exchange + symmetric encryption).
  - Allows recovery of the original MAC address with the decryption key.
  - Ideal for contexts where anonymization must be reversible under strict conditions.
- **Disadvantages:**
  - More complex implementation compared to other methods.
  - Key management can be challenging, especially in large systems.
  - Performance overhead due to encryption/decryption processes.

Notable Incidents

- **Privacy Violation**
- **Google Street View**
- **Phones with Wi-Fi On**
- **Data Leaks and Security Breaches**
- **Attacks Exploiting MAC Addresses**
- **WhatsApp Security Vulnerability**

Privacy Violation: Nordstrom

Nordstrom implemented technology to track customer movements in its stores through their Wi-Fi connections. The goal was to enhance the customer experience and optimize operations, such as adjusting staffing levels and rethinking department layouts. Sensors in stores collected information on the time customers spent in departments. However, after testing the technology, Nordstrom discontinued it in 2013 due to customer feedback, even though the data was intended to be anonymous and aggregated.

- **Impact:** Although the technology aimed to be anonymous, it raised privacy concerns regarding the collection of customer movements.
- **Outcome:** Nordstrom decided to halt the use of the technology following the trial.

Google Street View: Data Collection

Since 2007, Google's Street View cars inadvertently collected data from open Wi-Fi networks while photographing streets. This raised concerns about the security of personal data transmitted over these networks. In an audit, Germany's data protection authority found that Google had gathered fragments of personal web activity. Although the data was never used in products, it highlighted vulnerabilities in unsecured Wi-Fi networks.

- **Impact:** Google admitted to collecting personal data unintentionally.
- **Outcome:** Google halted the data collection and plans to delete the data under third-party supervision.

Phones with Wi-Fi On: Renew London Project

In the Renew London project, data was collected from over 530,000 unique devices to analyze movement patterns, directions, and speeds. The collected data was aggregated and anonymized, but it raised concerns about privacy violations under the Data Protection Act, as MAC addresses could be considered personal data.

- **Impact:** The project demonstrated the potential for targeted advertising based on location and behavior.
- **Outcome:** It is still unclear whether this data collection violated privacy laws, as the devices were not tracked individually.

WhatsApp Security Vulnerability

In 2012, a vulnerability in WhatsApp allowed attackers to impersonate users by obtaining their MAC address. The app relied on the MAC address for authentication, and attackers could exploit this by acquiring the MAC address via public Wi-Fi networks or malicious apps. Once in possession of the MAC address, attackers could log into WhatsApp as the victim and send messages.

- **Impact:** The use of static identifiers like MAC addresses for authentication posed significant security risks.
- **Outcome:** WhatsApp strengthened its security by implementing end-to-end encryption and ceasing to use MAC addresses for user authentication.

Conclusion

- MAC addresses are vital for device identification, but their use poses significant privacy risks.
- Legal frameworks such as the GDPR and European Convention on Human Rights aim to protect individuals from misuse of their data.
- Various anonymization techniques, like truncation, hashing, and encryption, offer ways to mitigate these risks.
- Ensuring the security and privacy of data requires both effective technical measures and adherence to legal standards to safeguard users' rights.

Thank you for your attention

Questions?