Wi-Fi Privacy: Vulnerabilities and Defense Mechanisms
Anirudh Ekambaranathan
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
a.g.ekambaranathan@student.utwente.nl

ABSTRACT
Wi-Fi has become a very popular technology and many wireless devices use it to connect to the Internet. This paper looks into the vulnerabilities which come with Wi-Fi and also looks into defense mechanisms to counter these vulnerabilities. Having created a list of all attacks and defense mechanisms this paper shows that there is no single solution to protect a wireless device when connected to Wi-Fi.

Keywords
Wi-Fi, security, privacy.

1. INTRODUCTION
More and more people nowadays make use of cellular technology and the smartphone is the preferred device for people to utilize this technology [4]. Many smartphone users have an Internet connection on their phone and there are many small programs, called applications or Apps, which rely on Internet technology to work.

There are several ways to connect a smartphone to the Internet. Popular options are Mobile Broadband and Wireless Fidelity (Wi-Fi). The focus of this research is on Wi-Fi (or also called the 802.11 protocol) and looks into some vulnerabilities it creates through the MAC protocol.

When users wish to connect to Wi-Fi, they have the option to enable Wi-Fi on their wireless devices. This will make the device start looking for Wi-Fi signals by emitting packets known as MAC frames. By doing this, the device is looking for Access Points to connect to the Internet. After the device has connected, it will continue searching for signals. The MAC frames, emitted by Wi-Fi enabled devices, are not entirely encrypted. In particular, the header of the MAC frame is unencrypted and thus readable by humans. The MAC frame header contains, amongst other things, the MAC address, which is unique to each device.

There are devices called Wi-Fi sniffers, which keep record of the incoming MAC frames. These sniffers can be at any location to collect data sent by smartphones and other devices [31].

It has been shown before that Wi-Fi has its security problems [8], mainly because of the unencrypted MAC header. The data collected by the Wi-Fi sniffer can be used for various purposes, such as booby-trapping, analysing shopping behaviour and crowd analysis.

1.1.1 Applications
Booby-trapping: When a certain MAC address is associated to an individual, it is possible to know when that individual enters specific areas. In that case, when the Wi-Fi sniffer receives a frame with the MAC address of the device a deliberate action can be triggered. It is also possible to estimate an approximate distance from the device to the sniffer since the signal strength is given.

This can be done for both malicious purposes, such as when a certain individual enters a shop the prices of the commercial products are raised, or it can also be used for educational and research purposes, such as improving security.

Shopping behavior: When the MAC address of people is known, it is possible to track them around a certain location. The eavesdroppers can, for example, be shop owners who would like to analyze the behavior of their shoppers. They might then use the gathered information to make the shop more user friendly or better arrange their products to increase their sales.

Crowd analysis: MAC addresses are also used as a tool for traffic and crowd analyses [3] [4] [38]. This data can then be used to analyze the behavior of drivers. These analyses may lead, for example, to improving traffic related problems. However the gathered information may also be used maliciously, for example, to analyze how often certain users travel by car.

1.1.2 Problem Statement
From the data collected by a Wi-Fi sniffer it is thus possible to reveal information about the behaviour of people [38] or track them over spatial distances [10]. This is often seen as a serious breach of privacy. Given that these sniffers are installed at more places such as stores, gyms and libraries, unwanted parties could gain an insightful view of the behaviour of victims [1]. This paper will explore the problems of Wi-Fi and look at some solutions to these problems and research whether it is possible to completely protect the privacy of users of wireless devices.

The rest of this paper first provides some background information on the Wi-Fi protocol (Section 1.1). Section 2 will look into the vulnerabilities of Wi-Fi and attacks possible because of these vulnerabilities. Section 3 discusses various defence mechanisms to protect wireless devices and its users as well as various advantages and disadvantages to these mechanisms. Section 4 describes the lessons learnt and concluding remarks.

1.2 Background Information
Wi-Fi technology is part of the 802.11 specifications. A Wi-Fi network can exist in two different modes: infrastructure mode and Ad-Hoc mode. Ad-Hoc mode is a method of wireless communication in which separate devices communicate with each other. In this mode, every participating device forwards data for other participating devices. More common in
Wi-Fi is infrastructure mode, which is the focus of this research.

A Wi-Fi infrastructure network consists of Access Points (AP) and Stations. Stations are devices, such as laptops, tablets and smartphones, which need to connect to the Internet through a wireless medium. An AP connects a station to the Internet and is responsible for sending and receiving data to and from stations connected to that particular AP. An AP will forward data to a router and therefore very often a router and AP are integrated into a single unit. In this paper the term Station and (wireless) device are used interchangeably.

All Stations of an infrastructure network and its Access Points (AP) collectively are called Basic Service Set (BSS). A BSS can be identified by a Service Set Identifier (SSID). The SSID is the MAC address of the AP belonging to the BSS. Section 2.2.3 discusses the SSID more extensively.

### 1.2.1 Service Discovery

When a Station needs to connect to Wi-Fi, it will seek to establish a virtual connection to an AP. There are two ways to do this: passive scanning and active scanning. During passive scanning the APs periodically transmit beacon frames containing the AP’s MAC address and SSID.

During active scanning a wireless station periodically broadcasts a probe request frame, which will be received by neighbouring APs. The receiving APs respond with a response probe frame. A probe request frame contains an SSID field so that the wireless station can associate with an AP. Then the station sends its Configured Network List (CNL), which is a list of all the networks the station previously has been associated with. The CNL of a device, therefore, contains information about the users past and this information is unencrypted. Because of this, it is now possible to leave the field containing the CNL empty. The AP must in that case still respond with a probe response frame. This broadcasted request probe is known as the broadcast probe request.

### 1.2.2 MAC Address

The MAC address, also known as the LAN address or physical address, is a link-layer address. For wireless LANs, the MAC address is 6 bytes long, which yields 2^48 possibilities, and is often represented in hexadecimal notation, with each byte expressed as a pair of hexadecimal numbers. The MAC address space is managed by the IEEE, which sells chunks of this consisting of 2^32 addresses to manufacturers. The first 24 bits of the address can thus be used to identify the manufacturer and it is up to the manufacturer to create unique combinations of the remaining 24 bits.

The MAC address does not belong to a hierarchical structure, contrary to the IP address. Wherever a wireless station is, its MAC address will at all times be the same. The same cannot be said for IP addresses. IP addresses must be changed when the host of the address moves to a different subnet.

For this research, the MAC frame header is specifically interesting. The reason is that the MAC frame header is always emitted in plaintext and is therefore not encrypted. During active scanning the station periodically sends an active probe request and anyone with a Wi-Fi sniffer can view these requests.

### 1.2.3 The IEEE 802.11 Frame

The 802.11 contains four address fields out of which three are relevant to the infrastructure mode. Address field 1 contains the MAC address of the destination of the frame. Address field 2 contains the MAC address of the sending station or the AP. Address field 3 contains the MAC address of the router interface which indicates the subnet to which the pertinent BSS is connected. The fourth address field is used when APs forward frames to each other and is thus only used in the Ad-Hoc mode.

### 1.2.4 Wi-Fi Sniffers

For security purposes penetration testing tools have been developed to examine the resilience of software technologies. Many such tools are available on the market and aircrack-ng [2] is an example of this. Aircrack-ng is a program which cracks the keys of 802.11 WEP and WPA-PSK.

There are also other types of software which monitor the Internet traffic and are called packet analyzer or wireless sniffers. The sniffers capture packets of data as they flow across the network.

### 2. WI-FI VULNERABILITIES

Figure 1 shows how Wi-Fi vulnerabilities are categorized in this paper.

The first category which is looked into is attacks on devices. This category deals with techniques which attackers can employ to harm devices of other users, or deny other users of Wi-Fi service. The second category is tracking of users and devices. This category looks at what methods exist in associating identifiers to users in order to track them. It is possible to use addresses, such as the MAC address, or digital fingerprints, discussed in Section 2.2.3.

#### 2.1 Attacks

The following attacks are performed at the network level of devices and not necessarily on a single target. The difference
between these techniques and tracking (discussed in Section 2.2) is that when these attacks are performed, the user of the target device will immediately feel an adverse consequence. The user will, for example, notice that the device has been disconnected from Wi-Fi. When tracking occurs, the user of the target device will often not know.

2.1.1 Spoofing Attacks
Spoofing [23] is a technique whereby an attacker sends packets into the network with a source address which is not his own. By doing this, the attacker can pretend to be someone else. Spoofing is a form of attacking itself, but more often than not it is used as an aid for other attacks such as Denial of Service. Studies have been performed on detecting spoofing [39] with the help of sequence numbers [19] and received signal strength [23] [35]. It is possible to spoof MAC addresses in Wi-Fi networks through software and tools which are available to the public [41].

2.1.2 Denial of Service Attacks
A denial of service attack occurs when an attacker manages to degrade the network quality for devices or deny network service for devices. Two types of Denial of Service (DOS) attacks based on spoofing [35] are discussed:

Deauthentication/Disassociation [22], [23]: In order for a station to connect to an AP, it must first go through a two-step handshaking phase. After this, it is possible for the AP or the station to send a Deauthentication frame which disassociates the AP and the station from each other. When a DOS takes place through spoofing, the attacker pretends to be the AP and connects to a station. As soon as the two are connected, the attacker, who pretends to be the AP, sends the Deauthentication frame and disassociates the device from the (fake) AP. After the attacker disassociates, it reconnects again to the same device. This will make the device constantly connect/reconnect and never make it possible to successfully access the network. The attacker can also impersonate another station, thereby sending a Deauthentication frame with the MAC address of the attacked station. The AP then disassociates with the attacked station whereas it never asked for it. The latter effectively does the same, however the attacker does not require to pretend to be an AP.

Power-saving Denial of Service [19]: A wireless device in a Wi-Fi network can enter a sleeping state, during which the AP saves all the frames meant for that device. The attacker again spoofs the MAC address and pretends to be a device of which incoming frames are stored at the AP. The attacker requests the stored frames and the AP sends the frames to the attacked device. There is nothing wrong with this as such, however, because the victim’s device is constantly receiving its stored frames, it will never be able to enter a sleeping state and this results in a quicker drainage of battery.

2.2 Tracking
Basic tracking [30] can be defined as an eavesdropper simply sniffing packets and keeping an eye on devices. This is a baseline for all other forms of attack. This method of tracking is not sophisticated and does not employ any advanced techniques. Through this method, the eavesdropper may be able to see which devices are emitting frames, their frequency and other information from the MAC frame header. There are various pieces of information, called identifiers, with which an attacker may try to track a device or user and there are several methods an attacker or eavesdropper may employ to associate these identifiers to the victim. These methods are discussed in the following sections.

2.2.1 Stalker Method
The stalker method [4] involves at least two individuals, the attacker and the victim. The victim is in possession of a wireless device or station which is transmitting probe request frames. In this scenario it does not matter whether the victim is connected to an AP or not. The attacker is in possession of a Wi-Fi sniffer and is sniffing the frames sent by the victim. The attacker uses the MAC address to distinguish the device of the user. Upon receiving the MAC address, the attacker can associate the MAC address to the individual.

This technique is easiest to implement when the attacker and victim are together and no other device is in the neighborhood. In that case the sniffer of the attacker will not pick up any addresses from other devices. When there is a group of people and one single device needs to be identified, it is possible to focus on the victim’s device and wait until all the other devices are replaced. This way, after the entire time, only one MAC address would have remained the same.

It is important to keep the victim in the area of monitoring, that is, the area as far as the Wi-Fi sniffer reaches. This can be done by either following the victim with the Wi-Fi sniffer, or by placing multiple sniffers over a large area and thus expanding the monitoring area.

This technique is easy to implement and does not require expensive materials or complex setups. Sniffing software are freely available and can be installed on mobile devices. The difficulty in this technique lies in the fact that the stalker needs to be able to remain stealthy and unnoticed, which may require some skill.

2.2.2 Beacon Replay Attack
The previous technique assumed that the device was using active probing, meaning that the device occasionally sends out a probe request message. This might not always be the case. The Beacon Replay Attack [10] is applicable to devices using passive scanning.

A wireless device will try to connect to a Personally Identifying Wireless Network (PIWN) [10] when it has connected to it before. Such networks may be home networks or networks to which the device regularly connects. The attacker can try to impersonate such a network by spoofing the beacon frame of the PIWN. By doing this, the attacker hopes that the victim’s device recognizes the frame as one coming from one of the APs in its Configured Network List (CNL). In response to this beacon frame, the victim’s device will try to connect and thus reveal information contained in the MAC frame header. A successful execution of this attack yields the same results as the stalker method and thus associates an identifier, such as the MAC address, to the victim. Compared to the stalker method, this setup is slightly more complex. The attacker is required to install software to send out frames. Then the attacker needs to properly configure the frames and send out valid information. Moreover, it may not always be possible to successfully guess the address or SSID of a PIWN.

2.2.3 Fingerprinting
A problem with tracking through the MAC address is that users can easily change their MAC address through publicly available software. Another problem is that some devices may be shared by multiple people, making it hard to bind a device to a single person. This does not render the abovementioned techniques useless, but makes them more difficult to deploy.

Regardless, it is so that other information is leaked during the time a device is trying to or is connected to Wi-Fi. This other information, e.g. time of arrival, CNL, SSID, are sensitive as
well and can be used to identify or profile a user. That information can be collected, for a single person, to combine and create a (unique) digital fingerprint [32]. That person can thenceforth be identified by that fingerprint. Fingerprinting is the subject of the following sections. Two distinct types have been classified. Fingerprints can be created from information specific to devices (device fingerprinting) or from information transmitted by users (behavioral fingerprinting). Both methods will be examined shortly.

2.2.3.1 Device Fingerprints

Device fingerprinting can be done through:

Radio Fingerprinting [20]: this technique identifies the signal characteristics of radio transmitting devices. These signals are different for each transceiver. The setup to capture transients is complex, expensive and requires to be operated by experts. This prevents it from being deployed on a wide scale and is therefore not generally used by attackers. On the other hand, the system can reach a very low false positive rate and high detection.

Passive Data Link Layer Fingerprinting: In 2006 it was discovered that the time between two beacon frames for network detection is up to the manufacturer of the NIC [13]. It is then possible to identify the different firmware versions and thus to create a fingerprint. The details of this technique are a little complex as well and requires the involvement of experts. In this method a large number of active probe requests are needed to successfully create a fingerprint. However, when a device tries to connect to an AP, practically, only several beacon frames need to be transmitted. This method is thus not applicable in case a device does not use active probing and when a device changes or upgrades its firmware this method can also not be used.

The above mentioned fingerprinting methods are not always applicable or reliable. Behavioral fingerprinting, is more realistic as a method for tracking. Behavioral fingerprinting can be as complex as the tracker wants it to be.

2.2.3.2 Behavioral Fingerprints

A behavioral fingerprint is made by combining implicit identifiers. These are simply pieces of information found in the packets or frames transmitted by users, or other patterns made available by the user. This method of fingerprinting combines several different identifiers such as activation patterns in emails [27], network destinations, CNL, SSID, broadcast packet size, request header fields in HTTP sessions [14], MAC addresses etc., to create a unique fingerprint. This technique does not rely on a single source, making it more reliable. It is, however, far more complex than simple MAC address tracking. The analysis associated with behavioral fingerprinting is sometimes mathematically complex and requires to be done by experts. Behavioral fingerprinting is a very powerful tool to uniquely identify a target.

Since behavioral fingerprinting combines data from many different sources it is possible to create fingerprints in many different ways. It is for example possible to relate the broadcast packet size of a device to the frequency of frames sent. It could be possible that such combinations yield a unique fingerprint for every person [32].

AP Fingerprint: A study done by Kumar and Helmy [25] shows that it is enough to use one identifier to create a fingerprint, namely the APs wireless devices connect to. Large organizations or institutions such as universities or companies may set up many different APs. According to Kumar and Helmy [25] a user only connects to 5% of the APs in 90% of the cases. These APs then form a unique combination for each user.

This approach, however, will not always work. The attacker will not always have access to the APs and will therefore not be able to withdraw the list of connected devices. This method works only if the attacker is an organization or person who provides the Wi-Fi service. This might happen in shopping malls, where they want to observe shopping behavior. This fingerprint is only available in the range of the pertinent APs. There is no way to transfer this fingerprint to another location.

The Configured Network List (CNL) is a powerful tool to create fingerprints with. Although it must be mentioned that a CNL alone may not be enough. There is no guarantee that the CNL is unique for each individual. Nevertheless, the CNL can be used, not only to create a fingerprint, but to infer sensitive information about a person’s past as well.

2.2.4 SSID

In a single probe request frame, the SSID field can reveal a lot of information about a single target. This field contains a list of the device’s preferred APs, is not encrypted and often human-readable. The SSIDs are often names of networks, companies, institutions or organizations and it is therefore not always hard to link the SSID to the actual provider [24].

There is a reason for the SSIDs being broadcasted. Active probing allows the device to search for APs which only serve the requested SSID. It is possible to disable this and leave the SSID field empty. The downside of this is that all APs must then respond to the probe requests, which takes up more bandwidth and, in case of multiple responses, makes the process slower. The same applies to passive scanning, in which case the device needs to wait for APs to broadcast their presence. There are some hidden networks which do not broadcast their SSID at all and, in case of passive scanning, the device would never know of the network’s existence [28].

The attacker has two points of attack if he chooses to act upon the SSID vulnerability. The attacker wants to infer the SSID of the network. In case it is not a hidden network the AP will broadcast its own SSID and thus the attacker can simply sniff this information. In case of a hidden network, the attacker can probe the network and thus receive its SSID. The second alternative is for the attacker to wait for another wireless device, which has connected to the AP before, to broadcast its list of SSIDs.

The list of SSIDs captured by an eavesdropper from a single device could possible reveal the identity of the user. This tracking technique is different from the previous mentioned Stalker Method (Section 2.2.1) and Beacon Replay Attack (Section 2.2.2). In this case the attacker knows who the victim is, but wants to associate him to his list of SSIDs. This is easy if the victim is alone. In case the victim’s wireless device is always turned on in crowded areas the attacker will have to analyze the lists. If the attacker has some information about the target, such as hometown or working environment, he can scan the list looking specifically for these related SSIDs.

When the attacker is in possession of the CNL (Configured Network List), containing the SSIDs, of the victim it is possible to infer some traits about his behavior. Since the list reveals information about places the individual has visited in the past it is possible to create a user profile. Using the same reason as is done by Kumar and Helmy [25] regarding the unique combinations of APs people visit, we can infer the same about the CNL. It is probable that the CNL of a person is nearly unique, and thus this CNL can be used to identify the person wherever he or she is.
The problem with the CNL is that it does not reveal the frequency at which places have been visited. For example, it is possible that a certain café has been visited only once by a victim, or it is possible that this is the café the victim visits every Friday night.

After an attacker has set up user profiles, it is possible to compare them and seek for relations [9] [11]. For example it is possible that two profiles which are similar belong to two people who might know each other or to relate multiple devices to a single user.

The vulnerability presented by the CNL is very obvious: it contains sensitive information. The new IEEE 802.11 standard allows for devices to leave the SSID field empty and forces APs to still respond. This solution is very simple, however not all devices have implemented this. Therefore, the CNL still is a serious threat.

3. COUNTERMEASURES

The previous sections explained and showed some of the many problems with the current Wi-Fi implementation. To overcome these problems several countermeasures have been developed.

This section looks into some of these measures and analyses their feasibility. The measures are analyzed from three different perspectives. Firstly from the point of view of the owner: i.e. looks what the person can do, e.g. disabling Wi-Fi. Secondly from the point of view of the device, i.e., what software can be installed on the device to prevent tracking? Lastly from the point of view of the IEEE 802.11 protocol, i.e. how can it be changed to improve the privacy of the users?

3.1 From the Perspective of the User

There are several measures people can take in order to improve their privacy. It is, for example, not advised to connect to unknown APs or surf unreliable websites. These APs or website may save private and personal information. Below are described some other measures which may improve privacy.

3.1.1 Disabling Wi-Fi

Wireless devices are emitting frames when the Wi-Fi function is turned on. This means that when the user has no intention of connecting to a network, the device will still be actively probing. As such this does no harm, however as soon as the device comes in range of a sniffer it becomes susceptible to attacks and tracking.

One solution to this is to disable Wi-Fi on the device as soon as it is not needed. This way the user will prevent sniffers from sniffing data in unknown locations. Though this sounds like a viable solution, it will not always work. There is no guarantee that reliable APs are not under surveillance of sniffers. Trustworthy shops in malls may make use of sniffers to analyze the behavior of their customers.

Another problem with this approach is that modern smartphones and other devices are made for the convenience of the user. Needing the user to constantly modify the Wi-Fi settings considerably cuts down the user experience. This method may work in certain situations, though it does not guarantee complete privacy.

3.1.2 Switching to Broadband

Another option is to avoid Wi-Fi altogether. Cellular internet access, or wireless broadband, is becoming more and more popular nowadays. Within cellular Internet access there are several different architectures, with 3G and 4G being most popular. Every new architecture improves on the older one and employs different technologies.

Wireless broadband takes away the dangers and problems presented in previous sections. There is no open MAC header being emitted, nor are there any active probe requests. There are several disadvantages to wireless broadband though. Wi-Fi is becoming nearly ubiquitous and almost at all places Wi-Fi is available and often for free. Wireless broadband is not free of charge and providers offer a subscription for approximately $40-$50. This results to a yearly cost of $500-$600 as opposed to free Wi-Fi. Many services such a Voice-over-IP make cellular voice and 3G services unnecessary. Moreover, broadband comes with its own set of security/privacy issues, such as Denial of Service and eavesdropping [26].

3.2 From the Perspective of the Wireless Device

The measures described above are impractical and neither Wi-Fi nor smartphones were developed for the user to constantly monitor its security. The owner of a device, therefore, can also take several other measures which automatically take care of some safety aspects. It is possible to run software which may bypass the security problems described in the previous section.

3.2.1 Random MAC Address/Pseudonyms

It is possible to randomize the MAC address from time to time. Despite the MAC address being unique and associated to the hardware of a device, there is various software available which freely change this address. The Apple iPhone is known to have an implementation of such a technique [37]. This technique may serve as a baseline for future analysis.

The implementation works as follows [29]. While the Apple iPhone is not connected to a network it does, under circumstances and certain conditions, emit a randomized MAC address. This only occurs when the device is in a so called Sleep Mode, which occurs when all services are inactive. For some devices, which make use of email services and receive frequent notifications, Sleep Mode is rarely entered. When the device does enter Sleep Mode and Wi-Fi is enabled, random MAC addresses are emitted. As soon as the device receives a notification or leaves Sleep Mode, the original MAC address is used again. Also after the device connects to Wi-Fi, the original address is used.

This implementation by Apple may seem to be the solution to the dangers described in the previous sections. However a study done by Motorola Solutions showed that this method by Apple impacts Wi-Fi tracking very little [29]. The major reason being that many users have applications running which produce notifications and thus prevent the device from entering Sleep Mode. Another reason is that when a device is not connected to Wi-Fi, it only emits active probe requests approximately every five minutes and if the user is moving, it is very difficult to effectively track the device and thus rendering the randomized MAC address technique redundant.

Other techniques for disposing of identifiers have been suggested [18] and software [41] is available to freely run on wireless devices which randomly changes the MAC address. Every time the device starts up, a randomized MAC address is emitted and used for connecting to Wi-Fi. MAC addresses cannot be changed at any given time, this will end all active sessions and this will be bothersome for the user. Grunwald and Gruteser discuss a design and implementation of disposable wireless identifiers [18]. This is a more sophisticated approach to randomly changing the MAC address.

3.2.1.1 Challenges to Random MAC Addressing

The purpose of the design suggested by Grunwald and Gruteser [18] is to reduce the time during which an attacker can track a
single device. However, there are several challenges to randomly changing the MAC address. Firstly the newly chosen address should be unlinkable to the old one and must be valid. To solve this challenge, the following is proposed. Part of the 128-bit address can be generated through an MD5 hash and the remaining bits can then be filled in such that it is a valid IEEE 802.11 address.

Secondly when wireless devices choose random addresses, it is possible that a collision occurs. This means that a device chooses a MAC address which is already in use. Grunwald and Gruteser [18] approach the probability a collision occurs with the birthday paradox and show that for small networks the probabilities for collisions are small [18]. However, when the network of devices expands and spans over hundreds of nodes connected by several APs, the probabilities of address collision become rather high. To solve this problem the following approach is suggested. The device sends request, with a random address in the MAC address field, for another random address. If in this request the original address were to be used, the information that is tried to be hidden will actually be revealed. When a reply is received for this request, it indicates that the new MAC address is already in use. This process is then repeated until an available address is found.

Thirdly some providers and companies have their Wi-Fi set up so that only certain devices can connect. The AP then only allows certain MAC addresses to connect to the network. When a device decides to randomly change its MAC address it loses its rights to connect. To solve this an alternate authentication protocol is suggested based on a symmetric key protocol [18]. This design yields positive results and shows that tracking time by an adversary can be significantly reduced. Without changing addresses, 50% of the clients can be tracked for more than approximately 5 hours, whereas with random address changing this drops down to below 10 minutes, which is a significant change [18].

3.2.2 Correlation Attack
Changing the MAC address seems to offer protection against simple tracking. However, Huang et al. [21] presented a correlation attack which becomes available through randomly changing pseudonyms. Assuming the wireless device is running a software to consistently and constantly update its MAC address it is possible for an attacker or eavesdropper to correlate two addresses or pseudonyms sent by a device travelling in space [21]. Spatial correlation can accurately be used when the device is moving at a constant speed and when other devices are not in the vicinity. Temporal correlation is used when the period between two pseudonyms is small.

3.2.3 Silent period
In order to defend against the correlation attack, Huang et al. have presented an advanced implementation of random pseudonyms implementing a silent period [21] [22]. The idea behind this technique is to wait a random amount of time after disposing of the old pseudonym. This period is known as the silent period. This technique is more effective as more devices deploy it and are active in the same neighbourhood. When two or more devices enter and exit the silent period, and thus possess a new pseudonym, the eavesdropper or attacker will not be able to determine to which device the new pseudonym belongs. Using a silent period is considerably more effective than using a constant time between pseudonyms. Less than half the devices can be tracked for more than 30 minutes, this opposed to the devices being tracked for 70 minutes when not using the silent period [21].

The silent period is not very difficult to implement and does not charge extra overhead in comparison to the regular randomization. It is thus an effective method against tracking (stalker attack, sniffing etc.). Nevertheless, this does not provide protection against fingerprinting. The CNL of a device is often still transmitted and the silent period does not affect that. Collision of addresses remains to be a problem in randomly changing MAC addresses.

3.2.4 Distributed Assignment of MAC Addresses and Mix Zones
The alternative to the device assigning itself a MAC address is a scheme wherein within the network MAC addresses are dynamically assigned to the devices. Such schemes [34] have been proposed for wireless sensor networks, but could also be applied to Wireless LAN. Such a system is however rather complex and requires a lot of computational power since a central system must keep track of all the devices and their activities. Today’s network of wireless devices is very large and dispersed. It is infeasible for different APs to keep track of all MAC addresses that are currently in use and this makes it difficult to deploy it on a large scale.

A notion similar to the above is the mix zone [6]. A mix zone is a spatial area where a group of users exchange pseudonyms. Within a mix zone the devices do not transmit frames and thus APs and sniffers do not receive any data. During the period a device is in the mix zone it exchanges its pseudonym with other devices. APs are not sent the original identity, the MAC address for instance, of the device, but rather this new pseudonym. The device from then onwards uses the pseudonym to communicate to the AP. This distribution of pseudonyms must happen through a trustworthy intermediary. The idea is that eavesdroppers do not know in which way the pseudonyms have changed and this makes tracking not possible or at least less likely. Every time a user enters a mix zone the pseudonyms are mixed up with other users in the mix zone.

This technique yields the same advantages as the Random MAC addressing scheme with Silent Period. However this also poses new challenges. There must be an underlying system, which users can register to and it only becomes effective as the number of users increase. The advantage is that MAC address collision does not occur, however it still does not provide complete protection against fingerprinting. Improvements to the mix zone have been proposed [33] by viewing mix zones through a game theoretical model.

3.2.5 Obfuscation
Basic pseudonym techniques and randomizing MAC addresses are not highly effective against behavioral fingerprinting. Xu et al. present another method based upon injecting bogus traffic into the network and outgoing stream of frames from the device [40]. The assumption is made that pseudonym techniques (such as randomization of the MAC address) are being applied and the eavesdroppers are trying to track devices through the use of implicit identifiers, such as MAC address, IP address, SSID etc. Eavesdroppers cannot tell the difference between the real and fake packets and this makes their fingerprints less accurate. Xu et al. [40] looked into various methods of injecting the bogus data into the network stream. Randomly doing so proved not to be effective, however a method whereby data is injected in a systematic manner reduces identification rates by up to 20% [40].

Considering this technique is used in cooperation with the abovementioned random MAC addressing scheme with Silent Period, it can considerably reduce identification rates. However, the effectiveness and practicality of this system is
still doubtful. Shopping malls which are in control of the APs can fingerprint devices based on the AP fingerprinting technique. In case there is a single attacker it is possible to discard other CNLs and focus the targeted one. This technique is effective when the attacker wants to sniff data in general. In this case the data will be obfuscated and unreliable. The sniffer will register devices which are not there.

3.2.6 Encryption
As an alternative one might consider encrypting all management frames used to connect to an AP. As mentioned by Greenstein et al., such a system will bring with it a new set of challenges [16]. However it is worth considering whether it may solve the problems presented above.

A simple approach would be to use a public/private key pair. The wireless device could use the public key of the AP to send a probe and in it embed its own key. The AP could then respond using the device’s public key. This system however requires the AP to perform computations even if the probe is meant for another AP.

Next to the heavy computations required by the AP and the device, this method is not enough to protect against tracking through the use of implicit identifiers [16]. The frame header would still reveal sequence numbers, protocol version etc. It would not be possible to encrypt the entire frame, including the header, since the frame header contains information, such as the duration field, which is meant to be broadcast. One might argue that all the users in a network could share a key to decrypt the header of the frame. This would only be possible in private networks where not every device could join. However this would not be scalable; if hundreds of users share the same key this could set up potential security risks since it is not possible to keep track all the users and devices. When keys accidentally are leaked the entire system would be compromised. Furthermore if only the fields meant for the AP were to be encrypted the AP would be required to do more heavy computations. This would be an excellent set up for a Denial-of-Service attack, and thus inviting attackers.

Despite encryption being a very intuitive approach and solving some of the abovementioned problems, it seems impractical and introduces new dangers and threats.

3.2.7 Other Techniques
There are other techniques as well, such as hash based anonymization and reducing transmission power. Cunch et al. discuss some practical issues with hashing the MAC address [12]. Jiang et al. [22] extensively discuss reducing the transmission power so that only the corresponding AP of a device is in range to receive frames. This reduces the chance that the attackers and eavesdroppers are in range to sniff frames. The challenge is that this decrease in transmission power should happen without exchanging any messages, otherwise the eavesdropper can infer what has happened to the device and perhaps decide to deploy more sniffers. This decrease in transmission power may prevent frames being sent to unnecessary APs, however it cannot be known where sniffers are located. This method does not provide protection against tracking, fingerprinting or other attacks as such.

3.3 From the Network Perspective
So far the techniques discussed have not been effective against the SSID which is broadcasted by Wi-Fi enabled devices. The following technique provides protection against the SSID field being revealed [24].

3.3.1 Location Aided Probing
When a device wants to connect to an AP it sends out a list of preferred SSIDs; this is a list of APs the device has connected to before and the device is set to automatically connect to these APs. However, when a device wants to connect, there is no need to send out all the preferred APs. Location Aided Probing [24] is based on the idea that only those APs in the vicinity of the device are probed. This system is called LAPWiN (Location Aided Probing for Protecting User Privacy in Wi-Fi Networks) [24].

The device relates a physical location to an AP and the next time the device is in the same area it only probes the APs it related to that area before. APs do not send out their location, therefore it is not possible to obtain the location through the Wi-Fi protocol. However wireless devices often have some sort of GPS functionality or other positioning methods such as accelerometers.

LAPWiN basically filters which APs broadcast and which are not at a given location. When a device enters an area near which it connected to an AP before, it starts sending out probe requests to connect. When it successfully connects this specific AP the current location is added to LAPWiN. This makes connecting to that particular AP more accurate and faster.

3.3.2 SlyFi
So far all the techniques and systems discussed build upon Wi-Fi and add features to make it more secure. Greenstein et al. present a protocol similar to 802.11, called SlyFi [17], and makes use of encryption functions to protect privacy. This protocol manages to overcome some of the problems mentioned in Section 3.2.6. The main goal, besides encrypting the frames, is also to provide unlinkability and protection against profiling. When two frames are submitted subsequently an eavesdropper should not be able to link the two frames to one device.

Public key cryptography yielded heavy computations for the wireless device and AP; and symmetric key cryptography becomes inefficient for APs with many keys. For that reason SlyFi uses a mechanism called Tryst [17] to encapsulate management frames, which builds upon symmetric key but alters the protocol slightly to counter the adverse properties just mentioned. Tryst makes use of the fact that wireless devices only infrequently probe APs and that there are only a few different types of management frames. The details of this implementation is omitted and the interested reader is referred to [17]. This method of encryption also computes a hash over the symmetric key encryption. This hash appears to be random and allows APs to quickly discard frames if the frames are not meant for them through quick hash-table lookups.

The randomness of the hash and the narrowness of the interface allows unlinkability of the management frames. However this is not the case for data frames. A different encapsulation mechanism was developed for that: Shroud [17]. Shroud depends on the fact that data frames are sent after a connection has been established and thus that the receiver will receive the frame. This allows Shroud “to compute a sequence of unlinkable addresses on a per packet basis [17].” Shroud computes a different address for every packet sent. For every next packet the sender and receiver both compute the address + 1. This assumes that the AP and device have some underlying agreements on how this is done. The paper by Greenstein et al. [17] shows the details of this implementation.
4. CONCLUDING REMARKS AND LESSONS LEARNED

This paper initially provided background information on Wi-Fi and explained how the current implementation has several vulnerabilities. An explanation was given on what a Spoofing attack [23] is and how this can lead to a Denial of Service attack [19] [22] [23] [35]. In this attack, the attacker pretends to be someone else by spoofing the MAC address and sends Deauthentication frames, which disconnects the victim’s device from the Wi-Fi network.

This paper then discussed several tracking methods. It was mentioned that with the Stalker Method [4] and Beacon Replay Attack [10] it is possible to associate identifiers in the MAC frame header to individuals. Two different fingerprinting techniques were discussed: device fingerprinting [13] [20] and behavioral fingerprinting [14] [25] [27]. Device fingerprinting uses signals or data unique to devices, such as radio signals or the time between two beacon frames, to create digital fingerprints. Behavioral fingerprinting analyses data created by the behavior of users, such as the list of access points a user has connected to, and uses this to create fingerprints. It is easily possible to create behavioral fingerprints based on the Configured Network List of users [24] [28].

Various defense mechanisms to protect Wi-Fi enabled devices and their users have also been described. The user has the option to turn off Wi-Fi or switch to a broadband subscription. These options are not always advantageous, as asking the user to turn off Wi-Fi every now and then compromises the user experience and switching to broadband can become expensive. Another option is therefore to randomize the MAC address [18] [27] [30] in the frame header. However doing this does not prevent eavesdroppers and attackers from using other identifiers, such as the CNL or Sequence number field. Obfuscation techniques, such as injecting random packets in the network containing bogus data [40], may help when eavesdroppers are tracking crowds, however, when an attacker or eavesdropper is tracking a targeted individual obfuscation techniques are not enough to protect the user’s privacy.

It is then possible to apply techniques which change the infrastructure of Wi-Fi slightly. LAPWiN [24] is a method whereby the CNL is broadcasted only when not in the vicinity of a known AP. This prevents attackers from applying SSID fingerprinting techniques. SlyFi [17] is a protocol similar to Wi-Fi, however it encrypts frames sent over the network, whereas Wi-Fi does not encrypt the MAC frame header.

Table 1 lists the various Wi-Fi vulnerabilities and defense mechanisms discussed in this paper. A cross denotes that the defense mechanism in that column protects the device or user from the corresponding attack/tracking mechanism in that row. It can be seen that SlyFi is the only mechanism which protects against all the listed attacks. Apart from SlyFi there is no single solution to protecting a device and user from all mentioned attackers/eavesdroppers.

A combination of LAPWiN and Randomization of the MAC address with Silent Period offers strong protection against most attacks. However, for the general public it may not be easy to extend their devices with these mechanisms. It is therefore important that manufacturers and Wi-Fi service providers implement these mechanisms in their devices.

4.1.1 Concluding Remarks

Wi-Fi is becoming available at more locations. However, the current Wi-Fi implementation has many vulnerabilities. This work has looked into Wi-Fi vulnerabilities and defense mechanisms. It is difficult for users to easily protect their privacy. Manufacturers of wireless devices can help in improving the situation. The Apple iPhone already has an implementation of MAC address randomization. Even though this implementation is not very strong, it is a step in the right direction. Single improvements suggested before do not seem to be enough and challenges still need to be overcome before Wi-Fi will be completely safe and secure.

5. REFERENCES


| Table 1: Wi-Fi Attacks/tracking in relation to defense mechanisms |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Defence mechanisms** | **Disabling Wi-Fi** | **Random MAC/ no silent period** | **Obfuscation** | **Distributed MAC** | **Mix zone** | **Random MAC/ silent** | **LAPWiN** | **SlyFi** |
| Basic tracking | x | x | x | x | x | x | x | x |
| ID Detection/beacon replay | x | x | x | x | x | x | x |
| ID Detection/stalker method | x | x | x | x | x | x | x |
| Tracking with Correlation | x | x | x | x | x | x | x |
| Device fingerprinting | x | x | x | x | x | x | x |
| Behavioral fingerprinting | x | x | x | x | x | x | x |
| SSID profiling | x | x | x | x | x | x | x |


