Mobile subscriber WiFi privacy

Piers O'Hanlon, Ravishankar Borgaonkar
Department of Computer Science
University of Oxford, United Kingdom

Lucca Hirschi
LSV, ENS Paris-Saclay
Université Paris-Saclay, France

***Abstract*—This paper investigates and analyses the insufficient protections afforded to mobile identities when using today's operator backed WiFi services. Specifically we detail a range of attacks, on a set of widely deployed authentication protocols, that enable a malicious user to obtain and track a user's International Mobile Subscriber Identity (IMSI) over WiFi. These attacks are possible due to a lack of sufficient privacy protection measures, which are exacerbated by preconfigured device profiles. We provide a formal analysis of the protocols involved, examine their associated configuration profiles, and document our experiences with reporting the issues to the relevant stakeholders. We detail a range of potential countermeasures to tackle these issues to ensure that privacy is better protected in the future.**

I. Introduction

Most of the world's smartphones now have a WiFi interface which provides another way for the user to access the Internet and associated services such as mobile operator run WiFi networks and WiFi-Calling. Indeed, given that many of the mobile subsystems now depend upon services located on the Internet, today's smartphones are highly dependent on a connection. There are now also many high volume Internet based services, such as streaming video, which require substantial bandwidth and stretch the mobile networks to their limits so many operators have sought to offload data transport from the mobile networks to WiFi wherever possible. To make this process easy for the user the connection to, and use of, WiFi services has been automated as much as possible. However to grant authenticated access to such operator run WiFi networks, and WiFi-Calling services, requires the use of authentication protocols to negotiate access. These authentication protocols are also used to grant access for Femtocell devices but since WiFi has now become much more prevalent, WiFi based solutions are seeing more rapid growth. According to Cisco's analysis [1], by 2020 it is projected that there will be over three times as many WiFi connected mobile devices, taking the total to around 1.7B.

In this paper we highlight the privacy issues around the current use of these authentication protocols. To authenticate a connection one typically requires the use of some form of user identifier coupled with a secret token. In the case of mobile devices the main user identifier is the International Mobile Subscriber Identity (IMSI) which is a globally unique identifier used to authenticate a mobile subscriber on the mobile network.

We have noticed widespread use of specific authentication protocols (EAP-SIM and EAP-AKA) that have failed to employ identity privacy features and transport the IMSI unencrypted over WiFi. Crucially, we have discovered that these protocols are preconfigured on many smartphones so that devices automatically attempt authentication. During the course of the protocol exchange the IMSI may be observed or, with an active attack, can be forcibly revealed. The use of these protocols has recently seen rapid growth as they enable automatic authenticated WiFi connection in public spaces (such as the London Underground). Furthermore we have also developed another technique that performs an active attack that enables the extraction of the IMSI from phones that have enabled the WiFi-Calling feature which is currently being offered by a growing number of operators.

The leakage of the IMSI means that an attacker can potentially link it with the user's other hardware addresses, such as the WiFi MAC address, so tracking may then be performed independently of IMSI extraction attacks. Whilst a number of newer mobile operating systems do now provide for randomisation of the MAC address this is often limited to certain phases of network attachment or may be circumvented through the use of other protocols [2], [3].

We have implemented both the active and passive attacks which may be deployed on a laptop, or embedded device, with a suitable WiFi interface.

We carry out a comprehensive analysis of the privacy implications of these protocols, resulting in the following contributions:

- **WiFi-Based IMSI Catcher:** We have discovered flaws in the deployment of widely used authentication protocols that allow for the creation a new WiFi-based IMSI catcher.
- **Low-cost IMSI Catcher PoC:** We demonstrate the practical feasibility of our attacks using a low-cost proof of concept platform. In particular, we implement attacks labelled as *A1, A2*, and *A3* to show various methods of tracking subscribers (Section VI).
- **Formal analysis of EAP-SIM/AKA:** We leverage the automatic cryptographic protocol verifier ProVerif [4], to provide a security analysis of the privacy issues of these protocols, which enabled us to find an attack (*A4*) on EAP-SIM, using known GSM triplets, in the presence of pseudonyms and encrypted IMSIs.
- **Mobile OS Industry impact:** After following the responsible disclosure process, our research findings were acknowledged by the device manufactures (Apple, Google, Microsoft, and Blackberry) and the Global System for

Mobile Communications Association (GSMA)[1]. As a direct result of our reporting these issues to Apple they decided to develop and deploy the identity privacy *conservative peer* mode for EAP-SIM/AKA into iOS10.

These privacy threats have arisen from a number of competing issues around the trade-offs and compromises made in the design and deployment of these protocols in the mobile networks. The complexity of the issues means that no one party can actually fix the problems we have exposed. As a result we had prolonged discussions with both the mobile OS manufactures and mobile operators in an attempt to address the issues before we went public. In the paper we detail our experiences, examine the issues and attempt to address them.

The remainder of this paper is organized as follows. In Section II, we outline the background on the relevant technologies and protocols. We then describe our experimental setup in Section III, followed by an explanation of our adversary model in Section IV, which lays the groundwork for an explanation of the privacy issues in Section V, followed by description of the attacks we have developed in Section VI. We analyse the security issues in Section VII. In Section IX, we discuss related work. Finally, in Section X, we draw conclusions.

II. Background

There are a range of protocols involved when a device attempts to attach and utilise services over WiFi, which we detail in this section.

The type of WiFi Access Point security ranges from open-access, to the use of manually entered statically assigned passwords, to fully automated systems that rely upon some form of security device or token. There has been a large growth in the deployment of 'auto connect' (IEEE 802.1X based) WiFi Access Points where the security keys are automatically negotiated using an authentication protocol based upon the credentials in the smartphone's SIM card.

During authentication one requires the use of some form of user identifier coupled with a secret token. In the case of mobile devices the main user identifier is the International Mobile Subscriber Identity (IMSI). It is stored in the SIM, or USIM, on a Universal Integrated Circuit Card (UICC) in the mobile device and also within the operator in their subscriber database, or Home Subscriber Server (HSS). The secret token in this case is the subscriber's secret authentication key, K_i, which is securely stored within the USIM. The K_i may only be accessed indirectly by asking the USIM to perform certain cryptographic operations on the K_i for the purposes of authentication. The K_i is also stored within the operator's HSS.

A. EAP-SIM/AKA Protocol and Identities

The authentication is performed by one of two protocols; EAP-SIM [5] and EAP-AKA [6] which are methods based upon the Extensible Authentication Protocol (EAP) [7]. EAP is an authentication framework which is standardised by the Internet Engineering Task Force (IETF) and provides for a range of EAP-based authentication methods. The EAP-AKA and EAP-SIM standards do not themselves specify any form of encryption for the transport of identifiers (*e.g.,* IMSI). Whilst these protocols can be protected through the use of temporary identifiers and tunnelling we have found that current implementations do not provide sufficient protection, resulting in leakage of the IMSI. Furthermore the deployment of preconfigured automatically installed WiFi profiles on many mobile phones results in large numbers of smartphones being potentially vulnerable to privacy compromise.

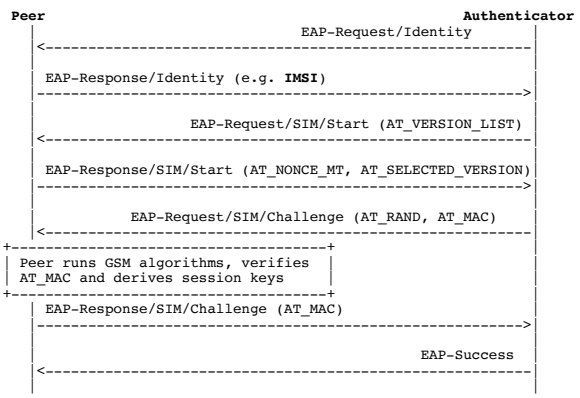

Fig. 1. EAP-SIM Full Authentication

The basic operation of EAP-SIM and EAP-AKA involves the exchange of an identity followed by an authentication exchange, as outlined in Figure 1 showing the full authentication exchange for EAP-SIM. In the case of EAP-SIM the authentication exchange is based upon the use of GSM authentication triplets, which are generated by the operator given knowledge of the secret key K_i, and consist of the Signed Response (SRES), Random number (RAND), and Ciphering Key (Kc). The protocol only exchanges the RAND as the other quantities may be derived from the K_i by the SIM in the mobile device as part of the authentication. In the case of EAP-AKA the authentication involves the use of a quintuplet authentication vector, which is generated by the operator given knowledge of the secret key K_i and an associated sequence number, consisting of a random number (RAND), an authenticator (AUTN) used for authenticating the network to the identity module, an expected result (XRES), an Integrity check Key (IK), and a Ciphering Key (CK). The EAP-AKA protocol exchange utilises stronger keys and uses both the RAND and AUTN which encodes the sequence number providing for stronger protection than EAP-SIM.

With EAP-SIM/AKA there are three basic identity types

[1]"The GSMA represents the interests of mobile operators worldwide, uniting nearly 800 operators with almost 300 companies in the broader mobile ecosystem, including handset and device makers, software companies, equipment providers and internet companies, as well as organisations in adjacent industry sectors." - http://www.gsma.com/

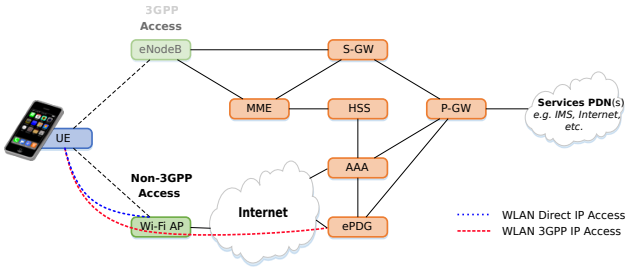

Fig. 2. Mobile network architecture

employed by the protocols:

Permanent identity (IMSI) The IMSI is used for the full authentication phase when neither of the following temporary identities are available (dependent upon the peer's mode of operation).

Pseudonym identity An identity which is a pseudonym for the IMSI, which is provided by the Authentication, Authorization and Accounting (AAA) server and usually has a limited lifetime.

Fast re-authentication identity A transient single-use identity - a new one is communicated on each reconnection. It provides for faster re-authentication as it relies on cached key material in the UE and at the AAA server, allaying the need to talk to the HSS.

The behaviour of the protocols are affected by the peer's policy configuration, which can operate in one of two modes:

Liberal peer The current default where the peer responds to any requests for the permanent identity.

Conservative peer A more privacy sensitive mode of operation where the peer only responds to requests for the permanent identity when no pseudonym identity is available. It is being proposed for deployment.

The 3GPP standard '3G security; wireless local area network (WLAN) inter-working security' [8] defines two authentication mechanisms, which may be seen in Figure 2[2], by which devices can provide for what is termed 'non-3GPP access' to services over WiFi. We have discovered serious issues with protection of the IMSI in both mechanisms.

B. WLAN direct IP access

The first mechanism is termed 'WLAN direct IP access' which provides users with automated secured access to WiFi networks through the use of IEEE 802.1X. This service has become widely deployed by many mobile operators as shown in Section V-B. Specifically, 802.1X is an IEEE Standard for port-based Network Access Control (PNAC). It provides an authentication mechanism to devices wishing to attach to a LAN or WLAN. The standard defines the encapsulation of the Extensible Authentication Protocol (EAP) over IEEE 802 and is thus known as 'EAP over LAN' (EAPOL). Whilst EAPOL was originally designed for IEEE 802.3 Ethernet, it was later updated for use with other IEEE 802 LAN technologies such as IEEE 802.11 wireless (WLAN) and fibre. Most mobile devices currently utilise EAP-SIM, whilst EAP-AKA is beginning to see deployment, for authentication to such networks. As mentioned above both utilise the IMSI as the permanent identity for the user, which is not encrypted as part of the protocol.

[2]The diagram is based upon a figure by Joe Deu-Ngoc from https://en.wikipedia.org/wiki/System_Architecture_Evolution

C. WLAN 3GPP IP access

The second mechanism is termed 'WLAN 3GPP IP access' which is utilised for the 'WiFi-Calling', or Voice over WiFi (VoWiFi), service which has been deployed by a number of operators, and is growing in popularity. In this case the phone will attempt an Internet Protocol security (IPsec) connection to their mobile operator's Evolved Packet Data Gateway (ePDG) over any connected WiFi network.

IPsec is a protocol suite, standardised by the IETF, for secure Internet Protocol (IP) communications that operates by authenticating and encrypting every IP packet of a communication session. The core protocols that define IPsec are the Authentication header (AH) [9], the Encapsulating Security Payload (ESP) [10] which provide for authentication and confidentiality respectively, and the Internet Key Exchange (IKEv2) [11] protocol which negotiates the key and session set up. IKEv2 initiates the connection in two phases:

1) **IKE_SA_INIT:** Negotiates security parameters for the IKE Security Association (SA), sends nonces, and sends Diffie-Hellman values.
2) **IKE_AUTH:** Transmits identities, proves knowledge of the secrets corresponding to the two identities, and sets up an SA for the first AH or ESP child SA.

It is during the IKE_AUTH phase that EAP-AKA is utilised to verify the identities. As mentioned in the security considerations section of the IKEv2 standard [11], whilst the SA set up is protected from endpoint impersonation through the use of a certificate, the exchange of identities is not protected, despite being encrypted.

III. Experimental Environment

The hardware requirements for these attacks is minimal, which underscores their importance. For the attacks we utilised a laptop running Linux with a WiFi interface capable of monitor mode, but we have also tested it on the Raspberry Pi, a low cost single-board computer. The most recent Raspberry Pi comes with built in WiFi and costs around £33. The attacking skills and expertise required to build and operate such a WiFi based IMSI catcher are lower than compared with traditional IMSI catchers mostly due to the fact that WiFi devices operate in unlicensed spectrum and are widely available, as compared to the more limited availability of suitable devices that operate in the various licensed mobile bands (although suitable Software Defined Radio (SDR) devices are becoming increasingly available and affordable).

For passive attacks to capture link-layer packets on WiFi the interface needs to be put into monitor mode and set to listen on the frequency of the specific Access Point under observation. Also the WiFi interface and driver for capturing the traffic needs to support the same modes of operation as the targeted Access Point so that all traffic may be monitored. We utilised *libpcap* based capture software including *Wireshark* and *tcpdump*.

The same hardware configuration is required for the active attacks with addition of the appropriate software. We utilised a modified version of *hostapd* to enable the laptop to function as an Access Point which performed the active attack. For the attack that impersonates the ePDG we utilised the *StrongSwan* IPsec server.

A. Ethical Considerations

Given that the hardware required to perform these attacks is readily available one must consider the ethical implications of such systems. The range of conventional mobile spectrum IMSI catchers may be measured in Kilometres. In GSM, theoretical range between base station and mobiles is about 35km [12], whilst the range of WiFi is typically less than 100 Metres, although this can be extended to Kilometres with specialist equipment. Furthermore, with the use of disassociation attacks [13], a target device may be forced to attach to a malicious access point.

This means that WiFi-based IMSI catching can reach into people's homes so it has a substantial invasive capability and should not be misused. Also since the problems are due to a combination of factors this means that the issues cannot be fixed quickly by one party.

It is for these reasons that we have withheld the details of the attacks from publication for over six months. We worked closely with the OS manufacturers to inform them of the issues, reporting them to Apple, Google, Microsoft, and Blackberry who all acknowledged the issue in their platforms. During this time they informed us of plans and we discussed potential countermeasures. Furthermore, we also reached out to the operators and gave a talk on the issues to the GSMA's Fraud and Security Architecture Group (FSAG).

IV. Adversary Model

In this section, we outline the adversary model for our attacks. We consider following two assumptions for this model:

1) The primary goals of the adversary against WiFi subscribers using cellular services is to learn presence of a subscriber and track their location. This is also facilitated by linkage of hardware identities (*e.g.,* linking the IMSI with WiFi MAC address).
2) The adversary is in the same geolocation area as the victim, or is capable of installing and/or compromising suitable systems in the same area.

The adversary model is divided into following two types:

Passive A passive adversary is able to deploy a system to silently sniff over-the-air broadcast channels of a targeted WiFi access point. In particular, they have access to a hardware device (for example a PC capable of running WiFi interface in monitor mode) and associated software needed to sniff and read messages.

Active An active adversary is able to set up a fake WiFi access point (or take control of one) to lure subscribers to attach and intercept the communication traffic. The PC with WiFi interface card or a separate WiFi access point device could be used to intercept the traffic. The device could also perform a dissociation attack on the existing WiFi infrastructure to effectively force clients to join to the malicious network.

As mentioned in the previous section the equipment necessary for an attacker is readily available as the attacks may be performed on a standard laptop with an appropriate WiFi interface.

V. Privacy Leaks

In this section, we describe the privacy leaks and attacks when a device utilises mobile operator run WiFi hotspots and WiFi-Calling technologies. In particular, we explain protocol and configuration issues responsible for leaking sensitive information.

As we have outlined in the Section II the currently deployed protocols do not provide sufficient protection for the IMSI. Whilst this is something that has been mentioned in the protocol documents themselves, the issue is exacerbated by current configuration and deployment choices.

A. Direct IP Access privacy

In the case of the Direct IP access, the basic problem is that the EAP-SIM/AKA interaction is not encrypted, and during the course of the protocol exchange the IMSI is revealed when the device first connects to the network so it may be passively observed. Also due to the fact that the current approach does not yet utilise pseudonyms, it is possible to perform an active attack to reveal the IMSI from a device attempting to connect.

B. Smartphone Configuration Profile issues

We have discovered that these problems are greatly amplified by the fact that many smartphones are now preconfigured to automatically connect to a list of operator supplied WiFi network SSIDs. Thus many smartphones will attempt authentication when in the range of their preconfigured list of SSIDs. Given that there is no way to verify the authenticity of an advertised SSID, the phone has no way to know whether it is talking to a malicious or genuine mobile operator Access Point.

These preconfigured profiles may either be installed automatically or manually. The automatic profiles are provided by the mobile operators for use on iOS, Android and Windows phones.

Specifically we examined iOS9 which contains preconfigured profiles for many operators around the world. We analysed these profiles and found 60 profiles spanning 41 different countries, which contained 66 unique SSIDs. The profiles also contain other configuration information such as

specification of the supported authentication mechanism and EAP type(s). These profiles are configured for use when a corresponding SIM card is inserted into the device. The profiles may also be updated via iTunes or directly downloaded from Apple by the phone. The profiles are signed which protects the devices from installation of maliciously crafted profiles. We have also analysed some of the latest profiles available online from Apple and we note that some operators are now beginning to offer EAP-AKA in preference to EAP-SIM.

On other devices, users may manually configure the connection to these operator run WiFi networks. The procedure for doing this is generally outlined on the operators websites, which usually consists of the user manually choosing the authentication technique (*e.g.,* EAP-SIM). Once the user has manually configured their phone the configuration will be stored allowing the device to automatically connect from then on, and thus potentially be tracked.

C. WiFi-Calling privacy

The issue with this method is that whilst the connection to the mobile operators Edge Packet Data Gateway (EPDG) is encrypted during the setup phase of the IP security (IPSec) protocol, unfortunately, cryptographic certificates are not used to protect the actual IMSI exchange. This means that the exchange is susceptible to a man-in-the-middle attack and thus the IMSI may be revealed. Thus a malicious attacker can set up an IPsec server to impersonate ePDG which would be capable of participating the IMSI identity response of EAP-AKA interaction of the IKE_AUTH phase.

D. Device identity linkage

The leakage of the IMSI means that an attacker can potentially link the user's other hardware addresses, such as their WiFi MAC address. Whilst there has been other work [14] attempting to correlate other mobile identifiers, such as the International Mobile Equipment Identifier (IMEI), with WiFi but they were not aware of the attacks we have discovered. This correlation may be performed in either direction so that a target of interest may have a known WiFi MAC address which could then be linked to their IMSI identity, or another target may have a known IMSI which could then be linked to their MAC address. Thus, once the IMSI has been associated with the WiFi MAC address, tracking can then be performed without further IMSI extraction attacks. Whilst a number of newer mobile operating systems do now provide for randomisation of the MAC address, this is often limited to certain phases of network attachment (*e.g.,* during the WiFi probing phase) or may be circumvented through the use of other protocols [2]. Once an attacker obtains a target's IMSI it can then potentially be looked up, on Internet based services commonly referred to as 'HLR (Home Location Register) lookup' services, to reveal the user's actual phone number.

VI. Attacks

This section details the attacks we developed to highlight the issues we have raised. We have implemented a Proof of Concept (PoC) of all the attacks below except attack *A4*.

- A1 **Direct IP access: IMSI observe** This attack allows a passive attacker to capture the IMSI over the air. It operates using monitor mode WiFi capture to capture the EAPOL packets. The WiFi interface needs to be set to the appropriate channel and in range of the mobile device of interest. Specifically the *EAP-Response/Identity* and *EAP-Response/SIM/Start* responses may be observed for the presence of the IMSI in the Identity field or AT_IDENTITY attribute, respectively.
- A2 **Direct IP access: IMSI extract** The attack involves setting up of an Access Point which advertises an SSID matching the configuration of the mobile device of interest. The malicious AP issues requests for the mobile device's full identity by *EAP-Request/SIM/Start* containing an AT_PERMANENT_ID_REQ to which the mobile device of interest should reply with its IMSI, provided it is liberal peer mode. To oblige the mobile device to connect to the malicious AP, it can also send out dissociation packets for any other competing APs.
- A3 **3GPP IP access: IMSI extract** This attack requires that the attacker sets up an IPsec server to impersonate the ePDG. The attacker also needs to either have control over the DNS responses returned to the mobile device or control over the IP routing between the device and the ePDG. Control over the DNS may be obtained either by taking control of the AP that the device is connected to, or by spoofing the DNS replies to the device (since DNS is not generally authenticated). Otherwise the IP routing may be manipulated using ARP attacks on the local AP network, or by taking control of the AP. Once these two preconditions have been met, when the mobile device of interest attempts connection to the ePDG this will diverted to the imposter ePDG which will participate in the IKEv2 authentication phases up until the IMSI has been exchanged in the IKE_AUTH phase, after which the connection will fail silently (without notifying the user). If the mobile device is configured in conservative peer mode then this attack may fail to obtain the IMSI.
- A4 **GSM Triplet attack: IMSI match** This attack is outlined in detail in Section VII, and requires that the attacker has obtained n GSM authentication triplets (where n is usually 3) corresponding to a specific $\texttt{IMSI}_0$ of interest. This attack can function when pseudonyms or encrypted IMSIs are in use. The attacker needs to set up a malicious AP in a similar fashion to the second attack. When a mobile device attempts to authenticate to the attacker AP it replies to the *EAP-Response/SIM/Start* message with a challenge calculated based on its knowledge of the GSM triplets and the provided AT_NONCE_MT. The answer to this response allows the attacker to verify whether the mobile device is the targeted one (*i.e.,* the one having

$\texttt{IMSI}_0$) or not.

A5 **IMSI and WiFi correlation** This attack simply forms a correlation between the obtained IMSI and the WiFi MAC address.

A. Impact

The attacks have differing impact and applicability. The first two attacks require that the attacker be local to the mobile device of interest or to have compromised equipment locally. The third attack may be performed also entirely remotely if the DNS responses can be manipulated so the impostor ePDG can also be positioned anywhere on the Internet. The TTL of the DNS responses can also be set to a long time out so the device may continue to use the cached address, thus extending the duration of the tracking. It should also be noted that whilst the attacks allow for tracking of a device through its globally unique IMSI they do not allow for the cloning of a device, nor decryption of traffic content.

The attacks could be compared to other tracking attacks as mentioned in the related work Section IX which rely on knowledge of the WiFi MAC address of the mobile device of interest. As mentioned the MAC address is now randomised on a number of newer mobile OSs so devices are harder to track using just WiFi Probe packets. So WiFi tracking needs to resort to devices that are connected to APs. There is little added information from a MAC address beyond the organizationally unique identifier (OUI) part of the address which can be used to identify the manufacturer of the WiFi interface. However the IMSI can potentially be used to lookup the user's mobile number and can also potentially be tracked if the SIM card is transferred to a different device. Although the user can also replace their SIM card which will then contain a new IMSI.

The range of the WiFi IMSI catcher could also be extended using commercial range extenders [15] or custom equipment. Although such range extension methods will increase the attack setup cost and may require a special set of expertise to deploy the fake WiFi access points.

VII. Security Analysis

In this section, we first investigate tradeoffs between security and criteria like usability and scalability; and discuss how they impact on privacy aspects of the subscriber. Further, we analyse deployment issues, cost, and complexity of various protection features such as TLS based approaches. Later, we apply formal verification techniques against EAP-SIM and EAP-AKA protocols and discuss our findings.

A. Trade-offs and Deployment Issues

As discussed in section V-B, privacy attacks are amplified due to that fact that many smartphones automatically connect to a list of preconfigured operator supplied WiFi networks. We consider the following points to explain the trade-offs between security and usability. From a usability perspective, it is good for smartphones to discover Internet supported WiFi networks automatically and connect to them silently. However, the feature of being silent and automatic makes them susceptible to an attacker who can setup a fake WiFi Access Point.

Although, luring smartphones to attach to fake WiFi networks would not introduce privacy leaks if there were adequate protection measures. For example, operators and smartphone OS manufacturers could deploy a tunnelled authentication. The smartphone can initially establish a secure tunnel to the operator controlled server. Such a secure tunnel could be established using TLS based approaches such as EAP-TTLSv0 or EAP-TLS. The EAP-SIM or EAP-AKA would be transported within the tunnel for authenticating subscribers to the mobile operator.

B. Formal Verification

We conducted automated formal analysis of EAP-AKA and EAP-SIM using a symbolic model based upon applied π-calculus [16]. We modelled[3] EAP-AKA and EAP-SIM in the ProVerif input language (*i.e.,* Blanchet, Abadi & Fournet's dialect [17] based on applied π-calculus). We used those models in order to formally verify *unlinkability* of the USIM for two users. In a nutshell, this unlinkability property is expressed as an indistinguishability property (from the attacker's point of view) between: (i) a situation where the same USIM is executing two sessions and (ii) a situation where two different USIMs are executing one session each. In practice, this can be done using the notion of *diff-equivalence [17]* that ProVerif can automatically verify. As a result, we were able to automatically verify unlinkability for those two protocols and some variations and different threat models.

a) **Analysis of EAP-SIM:** The EAP-SIM protocol (excluding the fast re-authentication mechanism) is essentially a stateless protocol using standard cryptographic primitives that can be defined using a sub-term convergent theory [18] (*i.e.,* convergent and such that the right-hand side of each rewriting rule is actually a syntactic sub-term of the left-hand side). Therefore, the full authentication part of EAP-SIM can be modelled in the ProVerif's dialect quite easily. We did so to analyse the unlinkability property as explained above. Note that error messages in EAP-SIM do not leak new information except the binary information failure/success [5]. Our model does not explicitly describe them but note that, given the way diff-equivalence is defined, ProVerif already considers an attacker knowing conditional's outcomes. Whilst there was an early pen-and-pencil informal analysis of the protocol [19], to the best of our knowledge, this is the first formal analysis of EAP-SIM in a symbolic model.

We ran ProVerif and, as expected, the tool finds an attack (in 0.35s) based on the IMSI leakage. We also tested the same model after having hidden the exchange of the IMSI from the attacker (we model this exchange on a private channel corresponding to either encrypting the IMSI, or the use of a pseudonym). ProVerif did not find any attack for this version. However, as soon as we give to the attacker only n

[3]Our ProVerif models of EAP-AKA and EAP-SIM are freely available at https://sites.google.com/view/models-eap.

arbitrary GSM triplets (n is a security parameter of EAP-SIM and is usually set to 3) corresponding to one IMSI, we denote $\texttt{IMSI}_0$, then PROVERIF finds a different attack (in 4.58s). Indeed, in such a case, an active attacker is then able to trace the mobile device (USIM) having $\texttt{IMSI}_0$ forever as follows:

1) *EAP-Request/Identity, EAP-Response/Identity:* The attacker sets up a malicious AP which requests, using EAP-SIM, the identity of the mobile device. As discussed above, we consider that the attacker is not able to deduce the IMSI from the answer (*i.e.,* EAP-Response/Identity), such as when pseudonyms or encrypted IMSIs are utilised.
2) *EAP-Response/SIM/Start:* The attacker receives the challenge sent by the mobile device.
3) *EAP-Request/SIM/Challenge:* The attacker computes an answer to this challenge using the GSM triplet known for $\texttt{IMSI}_0$ and sends back this response. This is possible because the expected answer to this challenge can be built with GSM triplets that *are not bound to* the challenge.
4) *EAP-Response/SIM/Challenge:* By analysing the response, the attacker learns whether the mobile device accepted the challenge or not. If the challenge is accepted then the attacker learns that he is communcating with a mobile device having $\texttt{IMSI}_0$. Otherwise, he learns he is communicating with a device having $\texttt{IMSI} \neq \texttt{IMSI}_0$.

Note that this attack can be carried out independently of the IMSI protection: be it encryption of the IMSI, use of a temporary IMSI or use of pseudonyms as is the case in future proposals for protecting the IMSI exchange. Also this attack exploits a lack of mutual authentication for the specific threat model we described (*i.e.,* when the attacker knows n GSM triplets). As already pointed out in [19], lack of session independence stems from this flaw. We have shown above that a traceability attack is also enabled by the same flaw.

We stress the advantage of having such symbolic models that allow for quick analysis of such variations of protocols and threat models in order to better understand to what extent such protocols are secure.

b) **Analysis of EAP-AKA:** As already mentioned we also modelled and analysed EAP-AKA. However, contrary to EAP-SIM, EAP-AKA combines different features causing serious problems for existing methods and tools. Let us briefly discuss what are the main challenges that arise from the formal verification of the EAP-AKA protocol:

- the modelling of the exclusive-or operator which cannot be handled by existing tools[4],
- the presence of a state (*i.e.,* the Sequence Number whose value must be stored from one session to another), and,
- basic arithmetic (*i.e.,* the Sequence Number is basically an integer and integer additions and comparisons are carried out by the USIM).

Each of these features of the AKA protocol constitutes a major challenge to existing techniques. Note that previous work involving the modelling of the EAP-AKA or AKA protocol as a symbolic model (*e.g.,* [21]) already acknowledge those difficulties and thus had to greatly approximate the protocol as well. Finally, note that unlinkability is defined as an indistinguishability property rather than a reachability property. Indistinguishability, often formalized through observational equivalence, is notoriously difficult to verify [22] compared to much more common reachability properties. This makes even more complex formal verification of unlinkability for EAP-AKA.

Nevertheless, we modelled EAP-AKA (excluding the fast re-authentication and re-synchronization mechanism) by using the same kinds of over-approximation (*i.e.,* by replacing the exclusive-or by a different construct having simpler algebraic relations and by replacing Sequence Numbers by fresh values for each session). For such a model, PROVERIF found the expected attack based on IMSI leakage in 0.25s.

VIII. COUNTERMEASURES

There are a number of countermeasures that may be deployed to mitigate these attacks. We break these down into two main deployment categories; firstly those that may be deployed by the operators, which would include the actual mobile devices or services developed by vendors and mobile OS manufacturers. Secondly, those that may be performed by users on their mobile devices. Clearly there needs to be cooperation, and preferably standardisation, in the design of the developed systems to operate correctly as a whole.

A. Operator and Vendor Mitigations

Whilst both EAP-SIM and EAP-AKA specify support for protection of identity privacy in the form conservative peer mode and the use of pseudonyms, most implementations have not provided support for it until very recently. Indeed as a result of our report on these issues to Apple, they developed support for conservative peer mode in iOS10. This support should be enabled in the core and on the devices where available as it improves the resistance of both protocols to IMSI attacks. There still remain questions regarding the exact operation of conservative peer mode and whether roaming and timeouts might lead to other problems.

Whilst there are ongoing efforts to provide for improved authentication privacy such as JFK [23] and secret handshake protocols [24] these protocols have not seen widespread standardisation or deployment. Thus, to improve upon the privacy of EAP based authentication, one needs to transport the authentication over an encrypted tunnel. There are a number of potential EAP-based tunnelling protocols, mostly based around Transport Layer Security (TLS) [25], such as Protected Extensible Authentication Protocol (PEAP), which is referenced in the EAP-SIM and EAP-AKA specifications, or others such as EAP-TLS [26], and EAP-TTLS [27]. The issue with EAP-TLS is that it requires the use of client based certificates which makes things more difficult for mobile use as each client would need a certificate deployed to it. However, PEAP and EAP-TTLS only require server based certificates

[4]In PROVERIF, using [20], it is possible to deal with $\oplus$ but only for reachability properties.

which makes things simpler whilst retaining confidentiality of the EAP identities. The choice between PEAP and EAP-TTLS is less clear as they both provide for a good level of security, with EAP-TTLS providing for a wider range of authentication options. The use of PEAP is somewhat complicated by the fact that two versions exist, PEAPv0 [28] from Microsoft and PEAPv1 from Cisco. The final issue to consider is level of support in AAA platforms and in the mobile OSs, which is pretty broad for both protocols so it will come down to operator choice.

The benefit of EAP based tunnelling is that it can be utilised to improve the privacy of both the WiFi network attachment and for WiFi calling attacks, since tunnelled EAP may be utilised with both EAPOL, and IKEv2 respectively. For IKEv2 there is also the possibility of utilising 'Multiple Authentication Exchanges in the Internet Key Exchange (IKEv2) Protocol' [29] which would allow performing a certificate-based authentication followed by an EAP authentication of the user.

The deployment of server side certificates leads to more complexity in the operators network, and potentially in the client OS. Clearly the operators would need to deploy the certificates throughout their AAA infrastructure and potentially to edge devices depending on where they decide to terminate the authentication exchanges. These certificates need to be maintained with mechanisms in place to revoke certificates in the case of compromise. The corresponding support in the clients needs to be considered carefully as to how the certificates are verified - in particular as to the roots of trust utilised. A simple deployment might just utilise the OS's existing CA keystore to verify the AAA server's certificate, but given the number of CAs trusted in today's OSs this may not provide for a significant improvement in security. So some OSs may need to consider the use of a separate specialised keystore for verification of such entities.

The issue with introducing yet more layers into security mechanisms is that the additional processing and potentially increased number of round trip times can impact the latency of the connections. Thus it may also be worth exploring other options such as encryption of the IMSI, which has been proposed by the 5G-ENSURE project in the 'Privacy Enhanced Identity Protection' enabler [30], or taking other iniatives [31] further.

In terms of addition measures the mobile OS manufacturers could consider is providing more user control over the use of automatic optional security mechanisms. Furthermore it would be beneficial if the users could have some way of editing their stored WiFi network associations credentials in a similar fashion to that used to modify their credentials stored during web browsing.

B. User Mitigation

The range of mitigation options that are possible by users is somewhat limited and depends upon the mobile device's OS. In terms of control over the WiFi network behaviour on iOS devices it is possible to selectively disable the 'Auto-Join' toggle for networks which will stop the device from automatically attaching, although this can only be done in the presence of the WiFi network(s) in question. On some versions of Android is possible to remove existing Auto-WiFi profiles whilst on others one is limited to configuring the networks in range as with iOS.

On both iOS and Android it is possible to manually disable the WiFi Calling functionality. We are investigating the possibility of developing an Android or iOS app to detect and notify suspicious behaviour of WiFi access points (for example requesting IMSI only and disconnecting). A similar kind of logic is being used in certain Android apps [32], [33] to assist in detecting fake 2G or 3G IMSI catchers. Finally users can also just switch off WiFi in untrusted environments.

IX. Related Work

We divide related work into three categories: WiFi privacy leaks, IMSI catcher attacks, and proposed countermeasures to remedy such attacks.

Various studies have highlighted the risk of loss of privacy whilst using public WiFi hotspots [34], [35], [36], [37], [38], [39]. In particular, these studies identified issues by analysing WiFi probes, MAC addresses, encrypted and unencrypted traffic. In comparison, our study reveals privacy issues in the deployment of cellular authentication protocol over WiFi connections.

The IMSI catcher attack risks the privacy of mobile subscribers and several studies outline the issues responsible for such attacks in GSM networks [40], [41]. Recently researchers demonstrated the feasibility of building an IMSI catcher for LTE networks and tracking subscribers with their social identities [42]. Our work is similar in terms of tracking subscribers, however the cost of the attacker's set up and skill set required is considerably less. Furthermore, our IMSI catcher attacks are not constrained by the legal requirement of using licensed frequency spectrum as compared to traditional IMSI catchers.

Several countermeasures have been proposed to protect privacy of subscribers on WiFi. To mitigate privacy attacks over WiFi networks, Alfredo et al proposed solutions at the link layer [43]. However their mechanism does not cover IMSI protection. Studies have also been performed on wide scale multi-platform WiFi MAC randomisation [44], but these haven't considered the role mobile authentication protocols. Another approach of using anonymity was proposed by Raghunath et al [45], however due to the current architecture of cellular communication systems and lawful interception requirement it is not feasible. Techniques such as symmetric encryption to protect subscriber identity have also been proposed to improve the EAP-AKA protocol [46]. However, such solutions require modification of the existing cellular architecture.

X. Conclusions

In this work we analysed a key set of authentication protocols which are implemented in many of the world's smartphones that currently fail to provide sufficient protection

for subscriber identifiers. Our results help to shed light on the security of deployed authentication protocols and associated systems, and more generally, provide insight into the state of the art in security mechanisms currently deployed by industry. These insights raise questions about what tradeoffs are made and how the industry should ensure that widely-used protocols employ best identity protection practices.

We have shown how insufficient protection of device identities can be exacerbated by device preconfiguration, which can amplify issues that taken alone may not pose such a wide scale risk.

This work also illustrates that to achieve suitable levels of privacy protection the appropriate action needs to be taken at multiple levels. These include improvements in the development of standards so that privacy issues are clearly addressed and the use of privacy enhancing features are mandated. Furthermore such standards need to be appropriately implemented by vendors, operators and mobile OS manufacturers into future 5G networks.

ACKNOWLEDGEMENT

This research has been performed within the 5G-ENSURE project (www.5GEnsure.eu) and received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 671562.

REFERENCES

[1] Cisco, "Cisco VNI mobile forecast (2015 2020)," Cisco Report, 2016. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] M. Vanhoef, C. Matte, M. Cunche, L. S. Cardoso, and F. Piessens, "Why MAC Address Randomization is Not Enough: An Analysis of Wi-Fi Network Discovery Mechanisms," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, ser. ASIA CCS '16. New York, NY, USA: ACM, 2016, pp. 413–424. [Online]. Available: http://doi.acm.org/10.1145/2897845.2897883

[3] J. Martin, T. Mayberry, C. Donahue, L. Foppe, L. Brown, C. Riggins, E. C. Rye, and D. Brown, "A Study of MAC Address Randomization in Mobile Devices and When it Fails," *ArXiv e-prints*, Mar. 2017. [Online]. Available: http://adsabs.harvard.edu/abs/2017arXiv170302874M

[4] B. Blanchet, "An Efficient Cryptographic Protocol Verifier Based on Prolog Rules," in *Proceedings of CSFW'01*. IEEE Comp. Soc. Press, 2001, pp. 82–96.

[5] H. Haverinen and J. Salowey, "Extensible Authentication Protocol Method for Global System for Mobile Communications (GSM) Subscriber Identity Modules (EAP-SIM)," RFC 4186 (Informational), Internet Engineering Task Force, Jan. 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4186.txt

[6] J. Arkko and H. Haverinen, "Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA)," RFC 4187 (Informational), Internet Engineering Task Force, Jan. 2006, updated by RFC 5448. [Online]. Available: http://www.ietf.org/rfc/rfc4187.txt

[7] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson, and H. Levkowetz, "Extensible Authentication Protocol (EAP)," RFC 3748 (Proposed Standard), Internet Engineering Task Force, Jun. 2004, updated by RFCs 5247, 7057. [Online]. Available: http://www.ietf.org/rfc/rfc3748.txt

[8] 3GPP, "3G security; wireless local area network (wlan) interworking security," TS33.234, 2002, Latest release: 13.1.0 (2016-12-20). [Online]. Available: http://www.3gpp.org/DynaReport/33234.htm

[9] S. Kent, "IP Authentication Header," RFC 4302 (Proposed Standard), Internet Engineering Task Force, Dec. 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4302.txt

[10] ——, "IP Encapsulating Security Payload (ESP)," RFC 4303 (Proposed Standard), Internet Engineering Task Force, Dec. 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4303.txt

[11] C. Kaufman, P. Hoffman, Y. Nir, P. Eronen, and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)," RFC 7296 (INTERNET STANDARD), Internet Engineering Task Force, Oct. 2014. [Online]. Available: http://www.ietf.org/rfc/rfc7296.txt

[12] J. Eberspächer, H. Vögel, C. Bettstetter, and C. Hartmann, *GSM - Architecture, Protocols and Services (3. ed.)*. Wiley, 2009. [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470030704.html

[13] J. Bellardo and S. Savage, "802.11 denial-of-service attacks: Real vulnerabilities and practical solutions," in *Proceedings of the 12th Conference on USENIX Security Symposium - Volume 12*, ser. SSYM'03. Berkeley, CA, USA: USENIX Association, 2003, pp. 2–2. [Online]. Available: http://dl.acm.org/citation.cfm?id=1251353.1251355

[14] J. Martin, D. Rhame, R. Beverly, and J. McEachen, "Correlating GSM and 802.11 Hardware Identifiers," in *MILCOM 2013 - 2013 IEEE Military Communications Conference*, Nov 2013, pp. 1398–1403.

[15] EDIMAX, "N150 Wireless Outdoor Range Extender/Access Point with Built-in 12dBi Antenna," Report. [Online]. Available: http://www.edimax.co.uk/edimax/merchandise/merchandise_detail/data/edimax/global/home_access_points_n150_outdoor/ew-7303apn_v2/

[16] M. Abadi and C. Fournet, "Mobile values, new names, and secure communication," in *Proc. 28th Symposium on Principles of Programming Languages (POPL'01)*. ACM Press, 2001.

[17] B. Blanchet, M. Abadi, and C. Fournet, "Automated Verification of Selected Equivalences for Security Protocols," in *20th IEEE Symposium on Logic in Computer Science (LICS 2005)*. Chicago, IL: IEEE Computer Society, Jun. 2005, pp. 331–340.

[18] S. Delaune and L. Hirschi, "A survey of symbolic methods for establishing equivalence-based properties in cryptographic protocols," *Journal of Logical and Algebraic Methods in Programming*, pp. –, 2016.

[19] S. Patel, "Analysis of EAP-SIM Session Key Agreement," IETF 87 Proceedings, 2003. [Online]. Available: https://www.ietf.org/proceedings/57/slides/eap-11.pdf

[20] R. Küsters and T. Truderung, "Reducing protocol analysis with xor to the xor-free case in the horn theory based approach," *Journal of Automated Reasoning*, vol. 46, no. 3-4, pp. 325–352, 2011.

[21] M. Arapinis, L. Mancini, E. Ritter, M. Ryan, N. Golde, K. Redon, and R. Borgaonkar, "New privacy issues in mobile telephony: fix and verification," in *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 2012, pp. 205–216.

[22] S. Delaune and L. Hirschi, "A survey of symbolic methods for establishing equivalence-based properties in cryptographic protocols," *Journal of Logic and Algebraic Methods in Programming*, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S235222081630133X

[23] W. Aiello, S. M. Bellovin, M. Blaze, J. Ioannidis, O. Reingold, R. Canetti, and A. D. Keromytis, "Efficient, DoS-resistant, Secure Key Exchange for Internet Protocols," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*, ser. CCS '02. New York, NY, USA: ACM, 2002, pp. 48–58. [Online]. Available: http://doi.acm.org/10.1145/586110.586118

[24] G. Tsudik and S. Xu, *A Flexible Framework for Secret Handshakes*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 295–315. [Online]. Available: http://dx.doi.org/10.1007/11957454_17

[25] T. Dierks and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2," RFC 5246 (Proposed Standard), Internet Engineering Task Force, Aug. 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5246.txt

[26] D. Simon, B. Aboba, and R. Hurst, "The EAP-TLS Authentication Protocol," RFC 5216 (Proposed Standard), Internet Engineering Task Force, Mar. 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5216.txt

[27] P. Funk and S. Blake-Wilson, "Extensible Authentication Protocol Tunneled Transport Layer Security Authenticated Protocol Version 0 (EAP-TTLSv0)," RFC 5281 (Informational), Internet Engineering Task Force, Aug. 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5281.txt

[28] Microsoft, "MS-PEAP protected extensible authentication protocol (PEAP)," Microsoft, 2016. [Online]. Available: https://msdn.microsoft.com/en-us/library/cc238354.aspx

[29] P. Eronen and J. Korhonen, "Multiple Authentication Exchanges in the Internet Key Exchange (IKEv2) Protocol," RFC 4739 (Experimental),

Internet Engineering Task Force, Nov. 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4739.txt

[30] 5G-ENSURE Project, "5G-ENSURE D3.4 5G-PPP Security enablers documentation (v1.0)," 5G-ENSURE Project, 2016. [Online]. Available: http://5gensure.eu/sites/default/files/Deliverables/5G-ENSURE_D3.4_5G-PPP_Security_Enablers_Documentation.pdf

[31] F. van den Broek, R. Verdult, and J. de Ruiter, "Defeating IMSI Catchers," in *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '15. New York, NY, USA: ACM, 2015, pp. 340–351. [Online]. Available: http://doi.acm.org/10.1145/2810103.2813615

[32] SRLABS, "Snoopsnitch." [Online]. Available: https://opensource.srlabs.de/projects/snoopsnitch

[33] Udar Swapnil and Ravishankar Borgaonkar, "Darshak." [Online]. Available: https://github.com/darshakframework/darshak

[34] M. Cunche, M. A. Kaafar, and R. Boreli, "I know who you will meet this evening! Linking wireless devices using Wi-Fi probe requests," in *WoWMoM - 13th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks - 2012*, San Francisco, United States, Jun. 2012. [Online]. Available: https://hal.inria.fr/hal-00747825

[35] M. Cunche, "I know your MAC Address: Targeted tracking of individual using Wi-Fi," in *International Symposium on Research in Grey-Hat Hacking - GreHack*, Grenoble, France, Nov. 2013. [Online]. Available: https://hal.inria.fr/hal-00858324

[36] P. Rouveyrol, P. Raveneau, and M. Cunche, "Large Scale Wi-Fi tracking using a Botnet of Wireless Routers," in *SAT 2015 - Workshop on Surveillance & Technology*, Philadelphia, United States, Jun. 2015. [Online]. Available: https://hal.inria.fr/hal-01151446

[37] P. Falcone, F. Colone, A. Macera, and P. Lombardo, "Localization and tracking of moving targets with wifi-based passive radar," in *2012 IEEE Radar Conference*, May 2012, pp. 0705–0709.

[38] N. Cheng, X. O. Wang, W. Cheng, P. Mohapatra, and A. Seneviratne, "Characterizing privacy leakage of public wifi networks for users on travel," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 2769–2777.

[39] Atkinson, JS, "Your WiFi Is Leaking: Inferring Private User Information Despite Encryption," Doctoral thesis, UCL (University College London), 2015.

[40] Christine Padget, "Practical Cellphone Spying," Defcon 18, Las Vegas, USA, 2010.

[41] D. Strobel, "IMSI catcher," Chair for Communication Security, Ruhr Universitat Bochum, 2007.

[42] A. Shaik, J. Seifert, R. Borgaonkar, N. Asokan, and V. Niemi, "Practical Attacks Against Privacy and Availability in 4G/LTE Mobile Communication Systems," in *23nd Annual Network and Distributed System Security Symposium, NDSS 2016, San Diego, California, USA, February 21-24, 2016*, 2016.

[43] A. Matos, R. L. Aguiar, J. Giro, and F. Armnecht, "Toward dependable networking: secure location and privacy at the link layer," *IEEE Wireless Communications*, vol. 15, no. 5, pp. 30–36, October 2008.

[44] C. J. Bernardos, J. C. Ziga, and P. O'Hanlon, "Wi-Fi internet connectivity and privacy: Hiding your tracks on the wireless Internet," in *2015 IEEE Conference on Standards for Communications and Networking (CSCN)*, Oct 2015, pp. 193–198.

[45] M. T. Raghunath and C. Narayanaswami, "A Practical Approach to Location Privacy in Public WiFi Networks," IBM White Paper, 2003. [Online]. Available: http://domino.watson.ibm.com/library/cyberdig.nsf/papers/EAAB4C2AC7591C8B85256D9600666E3C/$File/rc22781.pdf

[46] H. Mun, K. Han, and K. Kim, "3G-WLAN interworking: security analysis and new authentication and key agreement based on EAP-AKA," in *2009 Wireless Telecommunications Symposium*, April 2009, pp. 1–8.