Fakultät für Elektrotechnik und Informatik
Institut für Softwaretechnik und Theoretische Informatik
Lehrstuhl für Security in Telecommunications

On the Impact of Modified Cellular Radio Equipment

vorgelegt von
Nico Golde (Dipl.-Inform.)

von der Fakultät IV – Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
DOKTOR DER INGENIEURWISSENSCHAFTEN (DR.-ING.)
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Hans-Ulrich Heiß, Technische Universität Berlin
Gutachter:	Prof. Dr. Jean-Pierre Seifert, Technische Universität Berlin
Gutachter:	Prof. Dr.-Ing. Jochen Schiller, Freie Universität Berlin
Gutachter:	Prof. Dr. Yongdae Kim, Korea Advanced Institute of Science and Technology

Tag der wissenschaftlichen Aussprache: 12. September 2014

Berlin 2014
D83

Ich versichere von Eides statt, dass ich diese Dissertation selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

Datum

Abstract

Telecommunication protocols are considerably complex and both hard- and software have been traditionally kept out of reach of security researchers and attackers with regard to understanding its internal details or modifying equipment. While protocol details are open and publicly available via a large set of telecommunication specifications, the inner workings of mobile devices are proprietary and software stacks running on this hardware are closed. Also, the number of system information leaks from this industry has been traditionally very low and to a large extent mobile phones are still a black box. However, this is more and more changing due to an active Open Source and hacker community with the goal to enable free research on open hard- and software in this space.

This has lead to new research and particularly the accessibility of mobile technology to adversaries that do not operate standard compliant equipment. Such research is particularly interesting for several reasons. First, mobile networks are critical infrastructure and a large extent of the population relies on its security and availability under different circumstances. Second, the telecommunication industry is driven by big international corporations and different aspects of monetary interests are at stake. Third, no proximity is required to communicate with the network, which makes it highly inefficient to trace back certain types of attacks to attackers in practice.

The goal of this research is to build up on newly available hard- and software components for radio communication. Therefore, this work investigates the potential and impact of attacks based on modified hard- and software without exploiting application-specific software vulnerabilities directly. We demonstrate that an attacker in such a position can pose a significant threat both to subscribers and to the carrier itself. While we highlight individual security vulnerabilities in this work, the practical impact of these can be attributed towards the inherent lack of accounting for adversaries operating intentionally misbehaving equipment in standards. We exhibit that it is fundamentally wrong to assume a trusted serving network, unmodified baseband firmware, and rely on security through obscurity principles in the context of mobile telecommunication.

We split this challenge into two different aspects, carrier-grade equipment such as used on the network operator side and traditional handsets used by subscribers in the field. On the network side, we conducted a study of the security of low-power cellular base stations called femtocells to attack both subscribers of such technology and the carrier providing access to the cellular network. This is enhanced by studying the impact of an attacker with arbitrary access to over-the-air traffic on certain privacy properties of modern 3G networks. On the phone side, we focused on modifying the baseband software to evaluate the feasibility of attacks against other subscribers in the network. In both cases, we demonstrate that an attacker with access to carrier equipment, subscriber equipment, and baseband software can cause significant damage against other subscribers and the carrier network. This includes possibilities to impersonate other subscribers, hijack service establishments, several aspects of subscriber privacy violations, and denial of service attacks. Furthermore, we determine how the attacks included in this work can be leveraged to cause significant damage on a large scale.

Zusammenfassung

Telekommunikationsprotokolle sind vergleichsweise komplex und Hard- und Software aus diesem Bereich ist traditionell nicht in Reichweite von Sicherheitsforschern gewesen, die ein Interesse daran haben die Details eines solchen Equipments zu verstehen oder zu modifizieren. Zwar sind Protokoll-Details mit der Veröffentlichung einer großen Anzahl von Telekommunikations-Spezifikationen öffentlich zugänglich, jedoch sind die eigentlichen Details zur Funktionsweise von mobilen Geräten priopriëtär und die verwendete Software unbekannt. Die Anzahl an "Leaks" mit System-Informationen aus dieser Industrie war traditionell niedrig und zu einem großen Teil sind Mobiltelefone ein geschlossenes Stück Hardware. Dies ändert sich jedoch zunehmend durch eine aktive Open Source und Hacker-Community. Diese Community hat ein großes Interesse daran freie Forschung und offene Hard- und Software in diesem Bereich zur Verfügung zu stellen.

Aus dieser Bewegung entstanden verschiedene Forschungsarbeiten. Allerdings impliziert dies auch die Verfügbarkeit dieser Technologien für Angreifer, die nicht standardkonformes Equipment verwenden könnten. Eine Vielzahl verschiedener Gründe macht Forschung in diesem Bereich sehr attraktiv. Mobile Netzwerke gehören heutzutage zu kritischer Infrastruktur. Ein Großteil der Bevölkerung verlässt sich täglich auf die Sicherheit und Zuverlässigkeit dieser Technologien in verschiedenen Lebenslagen. Des Weiteren ist dies eine Industrie, die von großen, internationalen Unternehmen getrieben wird. Diese repräsentieren verschiedene finanzielle Interessen, die hier potenziell auf dem Spiel stehen. Außerdem ist für viele Angriffe in der Praxis keine physikalische Nähe erforderlich. Dies macht es oft schwer Angriffe und deren Angreifer zurückzuverfolgen.

Das Ziel dieser Forschung ist es auf der Verfügbarkeit von modifizierbarer Hard- und Software aufzubauen. Folglich studiert diese Arbeit, wie groß das Potenzial und der Schaden ist, der von Angriffen auf Basis von modifizierten Geräten ausgehen kann, ohne dabei auf spezifische Softwarefehler von Herstellern zurückzugreifen. Es wird gezeigt, dass ein Angreifer in einer solchen Position signifikanten Schaden sowohl gegen Mobilfunkteilnehmer, als auch gegen Mobilfunkanbieter ausüben kann. Obwohl wir einzelne Sicherheitslücken aufzeigen, ist der Schaden in der Praxis auf grundlegende implizite Annahmen der Standards zurückzuführen, die einen Angreifer mit absichtlich bösartigem Equipment nicht berücksichtigen. Wir zeigen auf, dass es fundamental falsch ist anzunehmen, dass das Netzwerk vertrauenswürdig ist, baseband-Firmware unmodifiziert ist und dass das "security through obscurity"-Prinzip im Kontext von mobiler Kommunikation sehr gefährlich sein kann.

Die Arbeit beschäftigt sich hier mit zwei Aspekten, wie es Verwendung durch Mobilfunkanbieter findet und traditionellen Mobiltelefone, wie sie normale Nutzer

verwenden. Auf der Netzwerkseite beschäftigt sich die Arbeit mit speziellen Basisstationen, die als Femtozellen bezeichnet werden, um sowohl Subscriber, als auch den Mobilfunkbetreiber anzugreifen. Dies wird durch eine weitere Studie vertieft, in der ein Angreifer unlimitierten Zugriff auf den Funkverkehr hat. Diese Studie untersucht die Auswirkungen auf die Annahmen, die bezüglich der Privatsphäre der Nutzer in 3G-Netzwerken getroffen wurden. Bezüglich der Telefonseite beschäftigt sich diese Arbeit mit der Modifizierung von Baseband-Software und darauf basierend der Durchführbarkeit von Angriffen gegen andere Nutzer des Netzwerks. In beiden Fällen zeigen wir, dass ein Angreifer mit modifiziertem Anbieter-Equipment, Benutzer-Equipment und der dazugehörigen Software massiven Schaden gegen andere Teilnehmer und den Netzwerkbetreiber selbst verursachen kann. Dies beinhaltet die Möglichkeit sich als andere Teilnehmer auszugeben, den Aufbau von Verbindungen für Mobilfunkdienste in bestimmten Situationen zu übernehmen, verschiedene Möglichkeiten die Privatsphäre zu verletzen und sogenannte Denial of Service (DoS)-Angriffe durchzuführen. Des Weiteren beschäftigt sich diese Studie damit, wie Teile solcher Angriffe verwendet werden können, um signifikanten Schaden in großräumigen Gebieten zu verursachen.

Publications Related to this Thesis

The work presented in this thesis resulted in the following peer-reviewed publications:

- *Let Me Answer That for You: Exploiting Broadcast Information in Cellular Networks*, Nico Golde, Kévin Redon, Jean-Pierre Seifert, In the Proceedings of the 22nd USENIX Security Symposium 2013 (USENIX Security)
(see [64] / Chapter 5)
- *New Privacy Issues in Mobile Telephony: Fix and Verification*, Myrto Arapinis, Loretta Mancini, Eike Ritter, Mark Ryan, Nico Golde, Kévin Redon, Ravishankar Borgaonkar, In the Proceedings of the 2012 ACM conference on Computer and Communications Security (ACM CCS)
(see [38] / Chapter 4)
- *Weaponizing Femtocells: The Effect of Rogue Devices on Mobile Telecommunications*, Nico Golde, Kévin Redon, Ravishankar Borgaonkar, In Proceedings of the 19th Annual Network & Distributed System Security Symposium 2012 (NDSS)
(see [63] / Chapter 3)

Contents

1 Introduction

1.1 Motivation and Problem Statement

Looking back at the last 25 years of technology, mobile telecommunication has transformed the way that we communicate and live our daily lives. As of the first quarter of 2013, 6.8 billion mobile subscriptions are divided between the second, third, and emerging fourth generation (2G, 3G, and 4G) of cellular telecommunication networks [70]. At the same time, the actual number of unique subscribers was estimated at 3.4 billion [70]. Mobile devices are no longer just a platform solely used for phone calls, but provide the full range of communication possibilities that go beyond those that were traditionally found in desktop computers. This includes banking through mobile payment systems, web browsing, instant messaging via short messages or additional applications, location based services, and literally all aspects of communication in general. Mobile devices have become a personal article of great value.

This importance is reflected even by technological areas that are traditionally disconnected from the mobile world. Examples here are providers of online banking systems who discovered mobile phones and specifically the Short Message Service (SMS) as being tied so closely to individual users that they make use of this aspect for mechanisms such as two-factor authentication for these applications [45].

Yet, numerous principles used in mobile telecommunication standards originate from a time in which attacks based on and against these systems were either not in the interest of attackers or not feasible. Such attacks were often hard to perform in practice due to the lack of customizable equipment in this space and the rather closed telecommunication industry. Still, the research community has managed to reverse-engineer some of the secrets of this industry and was for example able to break the A5/1 and A5/2 stream ciphers, which are used in the Global System for Mobile Communications (GSM) to encrypt voice calls among other services [39, 40]. Furthermore, various theoretic research (see Section 2.3) as well as attacks against client side vulnerabilities in SMS handling, Near Field Communication (NFC), Bluetooth (BT), browsers, and others has been conducted by the larger research community [67, 87, 88, 124].

However, the number of practical attacks based on the *Over-the-Air (OTA)* telecommunication protocols and radio equipment itself was limited and certain assumptions (e.g., protocol compliance) have been made for both sides of the cellular radio link. One side of the radio link is the cellular *Access Network (AN)*, which connects a mobile subscriber with the cellular carrier network, while the other side of the radio part is the modem or the so-called baseband of a cellular phone.

Traditionally, both cellular radio stacks have been carefully kept out of reach for any kind of malicious activities. We believe that this in turn resulted in a lack of proactive security considerations for telecommunication networks, when it comes to considering active attackers as part of the threat model. In this thesis, we consider an active attacker to be an adversary who is using modified cellular radio equipment in an unintended way to either harm other subscribers, or cause significant damage to the carrier network, or even mobile telecommunication as a whole.

The 3rd Generation Partnership Project (3GPP), which maintains and creates technical mobile telecommunication specifications, was originally founded to evolve the GSM system towards the third generation of mobile networks. It also intended to significantly improve security in newer generations of mobile telephony standards. This resulted in the specifications for 3G networks based on Universal Mobile Telecommunication System (UMTS) and later also for 4G networks using Long Term Evolution (LTE) technology. The "3GPP 3G Security; Principles and Objectives" specification [8] defines security changes in 3G compared to 2G networks and protocols. The following section from the document serves as an example here to illustrate the extent of considerations regarding active attacks or the lack thereof when designing 3G protocols.

> *"The following weaknesses in the security of GSM (and other second generation systems) will be corrected in 3G security:*
>
> *1)* ***active attacks*** *using a "false BTS" are possible;"*

The only active attack considered in this specification and when designing the 3G security principles has been a false base station attack, also known as *"IMSI Catcher"* attack. This attack allows an adversary to impersonate a carrier network by providing radio access and connectivity to a subscriber, by simple means of copying network configuration parameters, namely using the same Mobile Network Code (MNC) and Mobile Country Code (MCC), as the legitimate operator. As a result, it is possible to, e.g., intercept communication once a phone is using the rogue network. In 2G networks, this attack works due to a design flaw, as the mobile phone has no means to verify the authenticity of the carrier network.

Around the year 2000, when the first 3G networks were commercially deployed [42, 117], the practical implication of this attack were however limited to law enforcement agencies and other entities with access to expensive carrier equipment. Almost 10 years later, the larger research community proved that such active attacks can also be mounted in practice with less costly and publicly available equipment [95]. To date, this remains one of the very few named active attacks, if not the only one, which is explicitly addressed in these telecommunication standards. Nonetheless, this is a reactive measure and not the result of proactive security considerations with regard to telecommunication protocols and active attackers. Throughout this work, we will demonstrate that even the false base station attack has not been fully addressed in modern 3G networks. Due to the introduction of specialized commercial network equipment and the respective protocol changes, it is still possible to conduct such attacks. The industry of course paid attention towards standard attacks, e.g., authentication tokens in 3G are protected against replay attacks. However, these considerations almost exclusively focus on traditional and known best practices around certain sensitive operations. For protocol procedures not obviously impacting security, these considerations are missing and communication partners are assumed to be trusted and protocol conform, i.e. those protocols are not designed primarily with security mind. The various certification bodies for telecommunication equipment (FCC, ETSI, IC, KC, ...) may also entrap in practice to make these assumptions.

One of the first noteworthy attacks to challenge assumptions made by the GSM protocol specifications has been the so-called RACH flood Denial of Service (DoS) attack) published by Spaar in 2009 [108]. This attack consists of a malicious handset, which continuously requests channel resources from a base station. As this is a limited resource, requests from legitimate subscribers seeking access to radio services cannot be fulfilled, effectively resulting in a denial of service condition. At that time, the attack has been implemented by modifying a leaked baseband firmware of a commercial mobile phone. However, such leaks that not only include the source code, but also the necessary requirements to run the leaked code on devices have been historically rather rare.

A very contemporary research direction further challenges the fundamental assumptions of these standards with respect to the practical availability of *open* cellular radio equipment, especially for GSM. A booming market for used telecommunication equipment, cheap software-defined radios (SDRs), leakage of some hardware specifications, and a well-trained Open Source community finally broke up this closed cellular world. The overall community work culminated in three Open Source projects: OpenBSC, OpenBTS, and OsmocomBB [49, 118, 122]. These Open Source projects constitute the long sought and yet *publicly available* counterparts of the *pre-*

viously closed radio stacks. Needless to say that those projects initiated a whole new class of so far unconsidered and practical security investigations within the cellular communication research [72, 75, 89]. Although all of them are still constrained to 2G network handling, in the research of this thesis, we also showcase Open Source software to tamper with certain 3G base stations.

This thesis continues the challenge of the mobile security assumption that *certain active attacks can be safely excluded* from the threat model. In this process, we exhibit several novel attacks against various aspects of mobile telecommunication protocols, both modern and traditional GSM based systems. Most of the attacks demonstrated in this thesis are based on design principles of the underlying technology. We will show how an active attacker in possession of modified cellular radio equipment and its software can pose significant damage and security concerns not only to other subscribers, but also to the carrier. This includes impact on monetary interests, but also covers aspects of security such as privacy, integrity, authenticity, and availability.

1.2 Hypothesis and Methodology

The research presented in this thesis aims to prove the following hypothesis:

"Mobile telecommunication protocols operate under the assumption that subscribers do not use modified radio stacks and are not designed to resist attacks from an active adversary who tinkers with off-the-shelf equipment."

In order to prove this hypothesis, we conduct a study to identify novel active attacks based on modified radio equipment, which highlight the problems arising from such implicit assumptions. During this, we focus on both aforementioned sides of a radio communication link, subscriber equipment, as well as carrier-grade equipment, which implements Access Network functionality.

Moreover, we focused on hardware components that are widely available to the public and at relatively low cost (i.e. not more than 100 €). Our threat model considers an active adversary in possession of hardware components that are driven by modified firmware code specifically tailored to conduct various attacks. We introduce several novel attacks based on telecommunication protocols, which are further used in empirical studies in commercial European carrier networks to evaluate the feasibility and impact in practice. The vast majority of these attacks are based on design shortcomings of the protocols and procedures used for communication. These attacks could be mitigated by standards, which do not assume and rely on a trusted communication partner for the respective covered aspects. We will use this to substantiate that telecommunication protocols used these days provide a large space for improvements regarding defense mechanisms for active attackers.

Experimental Research Environment: For our studies, we use commodity hardware that allows us to modify its operating system in order to gain access to low-level telecommunication protocol procedures.

On the subscriber side, we implemented a custom mobile baseband firmware based on the Open Source project called OsmocomBB [118] to design new attacks and perform them in practice. OsmocomBB implements basic support for the GSM protocol stack logic and is not a mere free reference implementation, but runs on actual hardware, e.g., the Motorola C123 mobile phone used in our studies. This combination allows us to study the possibilities provided to a mobile subscriber who attempts to pose a threat towards the carrier network and/or other subscribers.

On the network side, we implemented several modifications to the Linux-based operating system running on a commercial available femtocell device. These devices basically represent a smaller version of cellular base station equipment that is given into the hands of the customer in order to offload traffic from the mobile carrier network. This implies that these devices are usually deployed at facilities that are not owned by the carrier, i.e., homes, business facilities, etc., which makes them an appealing target for conducting attacks that require carrier-grade equipment.

For this, we obtained a commercially available femtocell device sold by the French operator Société Française du Radiotéléphone (SFR) [106]. The concept of femtocells is applicable to all mobile telecommunication standards. Due to the support of this specific device, our efforts focus on 3G/UMTS based attacks. Because of the complexity of the software running on the femtocell device, we additionally considered the remote attack surface of these devices.

Both of these device types are then used to study the possibilities for using a modified version of these to significantly disrupt and attack mobile communication. We furthermore exhibit on how an attacker is able to leverage these weaknesses to launch large-scale attacks against carrier infrastructure or subscribers in sizeable geographical regions.

1.3 Scientific Contribution and Impact

The work presented in this thesis demonstrates that while industry standards and the associated technology evolved over years, the thought processes attached to securing these systems did not evolve consistently in all technology areas at the same speed.

Some of the attacks that we detail are also applicable to other radio communication, such as Wi-Fi. However, these systems merely see the radio link as a transport layer. Secure protocols, which also changed and evolved over the years, were built on top of this. As an operator of such equipment, there is often even a choice between different techniques to achieve the same goal with varying grades of security. This is different in mobile telecommunication networks, as these need to support devices for a very long time, provide backwards compatibility, and firmware updates in the field are often difficult to realize. Furthermore, certain concerns are completely excluded from the threat model due to the limited access to hardware and software.

By modifying commercially available carrier-grade and subscriber equipment, our results show that the industry needs to understand and address threats caused by active attackers with access to and the ability to modify the inner workings of these devices.

In summary, the key contributions of this thesis are as follows:

- **End User Risk Assessment**: We demonstrate several novel and practical attacks based on rogue femtocell devices and a custom firmware implementation for commercial mobile phones. We show that these attacks can easily compromise all relevant aspects of security for mobile phone subscribers, and to some extent carriers. Namely aspects such as integrity, authenticity, confidentiality, and availability. Such attacks include intercepting of traffic, impersonating other subscribers, tracking subscribers, modifying subscriber communication and data, hijacking transmissions, and conducting large-scale denial of service attacks against large metropolitan area. Furthermore, we highlight new privacy issues that were uncovered during this process in UMTS networks.
- **Design Issues in Mobile Networks**: We highlight several design problems in the procedures and deployment of mobile networks and in specific femtocell, GSM, and UMTS networks. For femtocell networks these attacks are inherent in the basic architecture of current femtocells, are independent of the carrier, and we exhibit how some of these issues conflict with some of the basic 3G security principles and requirements. For UMTS networks, we show how privacy issues can arise from the assumption that certain messages sent to mobile subscribers will always originate from a legit carrier. Moreover, we demonstrate vulnerabilities resulting from the lack of a solid protocol security design and thorough authentication in GSM networks.
- **Implementation and Evaluation**: We developed an arsenal of software to both implement the attacks and enable interaction with critical components of operator infrastructure. All of the presented attacks have been experimentally evaluated in practice, by conducting empirical studies in major German carrier networks (Vodafone, O2, T-Mobile, E-Plus) and the French operator SFR. We eventually assess the boundary conditions and requirements for conducting large-scale attacks in order to cause denial of service conditions within a large geographical area of a major city. Berlin/Germany serves as an example here.

We communicated our research results to respective mobile operators as well as the vendor of the particular femtocell equipment used in our studies. The implementation specific flaws have been addressed by the femtocell vendor in firmware version V2.0.24.1.

1.4 Thesis Structure

Based on the aforementioned methodology, this thesis is structured as follows. In Chapter 2 we will introduce the technological background required to understand the environment and basics of the attacks presented throughout this thesis. This particularly covers aspects of GSM, 3G, and femtocell communication. Furthermore, this chapter comprises a study of related work in which we specifically focus on past work around the impact posed by the presented vulnerabilities.

In Chapter 3 we will concentrate on the operation of malicious, carrier-grade equipment and in specific femtocell devices. We detail how attackers can compromise such devices and highlight the resulting impact to subscribers and carriers associated with this technology. During this process we demonstrate several novel attacks based on this technology.

We continue to dissect privacy properties of UMTS networks (3G) and how those assumptions can be violated in Chapter 4. Furthermore, we re-use a modified femtocell device to prove that these protocol properties can be exploited in practice.

Chapter 5 focuses on attacks originating from modified handset equipment and thus completes the picture of modified radio equipment. We highlight the protocol properties leading to the showcased attacks and evaluate the feasibility of large-scale attacks against metropolitan areas.

Chapter 6 briefly concludes our research and provides directions for future research.

2 Background

The following sections will provide a concise background on radio communication protocols and specifically GSM, 3G, and femtocell technology. This information by no means represents a complete overview of these topics. The 3GPP specifications are available publicly and comprise thousands of documents. The interested reader can find an overview in [3]. We complete this chapter by providing an overview of related work.

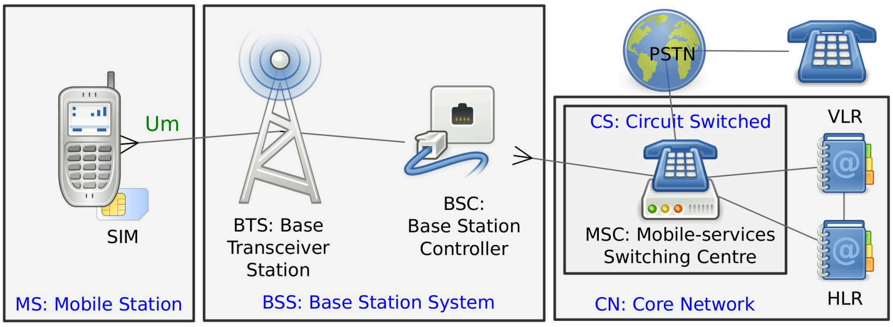
2.1 GSM Network Architecture and Procedures

This section briefly describes the GSM cellular network infrastructure components required to implement a minimal set of features. We continue to explain the important types and functions of logical channels. Furthermore, we depict the protocol details required to understand the basis of some of the attacks presented throughout this thesis.

2.1.1 GSM Infrastructure Components

Despite the complexity of a complete GSM mobile network architecture, only a few entities are relevant in this thesis. We intentionally simplify this to concentrate on the bare minimum of the architecture, which is required to handle subscribers, route services such as calls, and establish channel resources over the air interface. 3GPP TS 03.02 [4] provides a full overview of the components to the interested reader. Following are concrete infrastructure components of relevance to this research.

Figure 2.1: GSM Network Architecture


Um
SIM
MS: Mobile Station
BTS: Base Transceiver Station
BSC: Base Station Controller
BSS: Base Station System
PSTN
CS: Circuit Switched
MSC: Mobile-services Switching Centre
VLR
HLR
CN: Core Network

A simplified version of a typical GSM network architecture reduced to the bare minimum of components participating in mobile communication.

Figure 2.1 illustrates the architecture and connections between relevant components:

- *BTS*: The Base Transceiver Station is a phone's access point to the network. It relays radio traffic to and from the mobile network and provides access to the network over-the-air. A set of BTSs is controlled by a Base Station Controller (BSC) and is part of a Base Station System (BSS).
- *MS*: The Mobile Station is the mobile device interacting with the cellular network provided by a mobile network operator. It comprises hardware and software required for mobile communication (baseband processor, SIM card, and a GSM stack implementation). The MS interacts with the BTS over a radio link, also known as the U_m interface. In this thesis, the mobile phone of a victim is often referred to as MS. Depending on the context, we may use the term MS, user, subscriber, phone, and mobile device interchangeably.
- *MSC*: The Mobile-services Switching Center [9] is a core network entity responsible for routing services, such as calls and short messages, through the network. It utilizes components from BSSs to establish connections to mobile devices, organizes hand-over procedures and connects the cellular network to the Public Switched Telephone Network (PSTN).
- *VLR*: The Visitor Location Register maintains location and management data for mobile subscribers roaming in a specific geographical area handled by an MSC. It acts as a local database cache for various subscriber information obtained from the central Home Location Register (HLR), e.g., the mobile identity. A subscriber can only be present in one VLR at a time. Each of the areas served has an associated unique identifier, the Location Area Code (LAC) [4, 11]. As soon as a phone leaves a certain geographical area called *Location Area (LA)*, it has to perform the Location Update procedure [5] to notify the network of this event.

2.1.2 Logical Channels in GSM

The available GSM frequencies are shared among a number of mobile carriers. Each of the GSM frequency bands is divided into multiple carrier frequencies by means of Frequency Division Multiple Access (FDMA). A BTS serves at least one associated carrier frequencies identified by the Absolute Radio Frequency Channel Number (ARFCN). The ARFCN provides a dedicated pair of uplink and downlink frequencies for receiving and transmitting data over the U_m interface [15]. Because the radio frequency is shared among a number of subscribers, GSM uses Time Division Multiple Access (TDMA) as the channel access method and divides physical channels provided by the ARFCN into 8 time slots. A sequence of 8 consecutive time slots is called a TDMA frame. Multiple TDMA frames form a multiframe. It consists either of 51 or 21 TDMA frames (respectively control frames or traffic frames). Multiframes are further partitioned to provide logical channels.

The two categories of logical channels in GSM are *control channels* and *traffic channels* [6]. Control channels provide means for signaling between the network and the MS. Traffic channels are used as a bearer for voice and data traffic. Because GSM protocol related attacks presented in this thesis are solely based on signaling, we focus on the details of control channels. There are three categories of control channels:

- *BCH*: Broadcast Channels provide a point-to-multipoint, unidirectional channel from the BTS to mobile stations (transmitted on the downlink frequency). Among other functionalities, they act as beacon channels and include logical channels for frequency correction (FCCH), synchronization (SCH), and information about the cell configuration and identity (BCCH) [6, 10].

- *CCCH*: Common Control Channels are used for signaling between the BTS and MS, both on the uplink and downlink. They are used by the MS to request radio resources and to access the mobile network.

- *DCCH*: Dedicated Control Channels carry signaling messages related to handover procedures or connection establishment, e.g., during call setups.

For this thesis and the discussed attacks, we are mainly interested in logical channels that are part of the CCCH and DCCH categories. These categories consist of several logical channels. The logical channels of interest are as follows:

- *PCH*: The Paging Channel is used by the BTS to inform an MS about an incoming service (via paging request messages on the downlink channel). The PCH, which is part of the CCCH, will be monitored by every MS in idle mode unless it is currently using a dedicated channel.

- *RACH*: The Random Access Channel provides a shared uplink channel utilized by the MS to request a dedicated channel from the BTS. Placing a phone call or receiving an incoming service always requires a phone to set up a dedicated signaling channel beforehand.

- *AGCH*: The Access Grant Channel provides a downlink channel used by the BTS to transmit assignment messages that notify mobile stations of assigned channel details. A successful channel request on the RACH will result in an Immediate Assignment message on the AGCH. These assignment messages contain the required configuration parameters that enable the MS to tune to the requested channel.
- *SDCCH*: The Standalone Dedicated Control Channel is used on both uplink and downlink. It is employed for call setup and signaling between BTS and MS. Furthermore, it can be utilized to transmit short messages to the MS.

It is important to note that both the BCH and CCCH channel types are point-to-multipoint channels. This implies that information on the logical downlink channels is broadcasted to all subscribers served by a specific BTS. Throughout this work we will see how this can be abused to model new attacks.

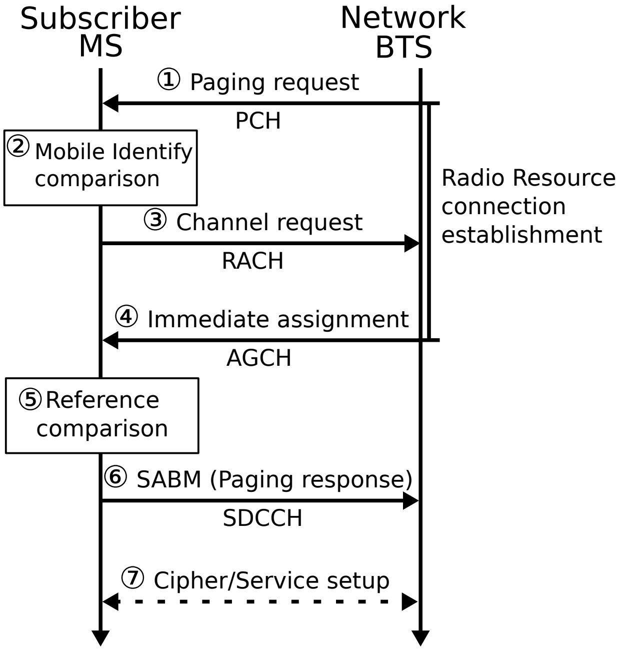
2.1.3 Mobile Terminated Service Procedures and Paging

The GSM specifications differ between traffic originating or terminating at a mobile phone. This is referred to as Mobile Originated (MO) and Mobile Terminated (MT) traffic. The paging related attacks presented during this thesis focus on attacking MT services, such as phone calls or SMS. Thus, in the following we concentrate on the underlying protocol procedures associated with MT services [5].

In order to deliver a service to a phone, the MSC needs to determine the location of the respective subscriber. This has to be done for two reasons. First, mobile phones will be idle most of the time to save battery power and so will not be in constant contact with the network. Thus, the operator does not always know the specific BTS that provides the best reception level to the MS. Therefore, it must broadcast this signal of an incoming service through at least the entire location area. Second, broadcasting this information through the whole operator network would impose a huge performance overhead and possibly overload the paging channel [105].

As the first step, the core network determines the responsible MSC/VLR for the target subscriber with the help of the HLR. Next, the MSC obtains the location information for the destination subscriber from the VLR and sends a *paging message* to all BSCs in the subscriber's location area. This message includes a list of cell identifiers/base stations serving the specific location area [30]. The message also contains the mobile identity of the subscriber, which is usually either an *International Mobile Subscriber Identity* (IMSI) or a *Temporary Mobile Subscriber Identity* (TMSI).

Figure 2.2: Mobile Terminated (MT) Paging



> The MT paging procedure is used by the network as a notification mechanism to inform a device of an incoming service such as a phone call.

We illustrate the remaining protocol logic using a successful MT phone call as depicted in Figure 2.2.

1. The BSC sends a *paging command* message which includes the subscriber identity to all base stations within the location area. All base stations re-encapsulate the mobile identity and transmit it as part of a *paging request* message on the downlink PCH.
2. When receiving a paging request on the PCH, each MS compares the Mobile Identity (MI) included in the request with its own. The result determines whether the message is addressed to itself or a different subscriber.
3. In case of an identity match, the MS needs to acquire access to Radio Resources (RR) in order to receive the MT service. To do so, it sends a *channel request* including a random reference number on the uplink RACH.
4. Upon receipt of the channel request, the network allocates radio resources and a dedicated channel. Next, it acknowledges the request and sends details of the allocated channel to the MS in an *immediate assignment* message on the AGCH downlink. To allow the MS to identify its assignment, the message contains the random reference of the requester.

5. The AGCH is a shared downlink channel. Therefore, an MS receiving an assignment message compares the included reference with the one sent in the request. If the reference matches, the MS tunes to the dedicated signaling channel included in the assignment.

6. After this step succeeded, the mobile station establishes a signaling link, usually over the SDCCH, by sending a GSM Layer 2 *SABM* frame containing a Layer 3 *paging response* message.

7. Following this, the MS and BTS undergo an authentication, ciphering and service setup procedure. Details of this procedure are not relevant for the paging attack presented in this thesis. We skip these details here.

The GSM standard specifies [5] three types of paging requests – type 1, 2, and 3. The type stipulates the number of subscribers that can be addressed with the paging request. Type 1 can page one or two subscribers, type 2 two or three subscribers, and type 3 paging requests are directed towards four subscribers at once. A recent study [75] suggests that in real operator networks the vast majority of paging requests is of type 1. During our early experiments, we verified that 98% of all paging requests that we observed are type 1 requests. Therefore, we ignore type 2 and type 3 paging requests in this thesis.

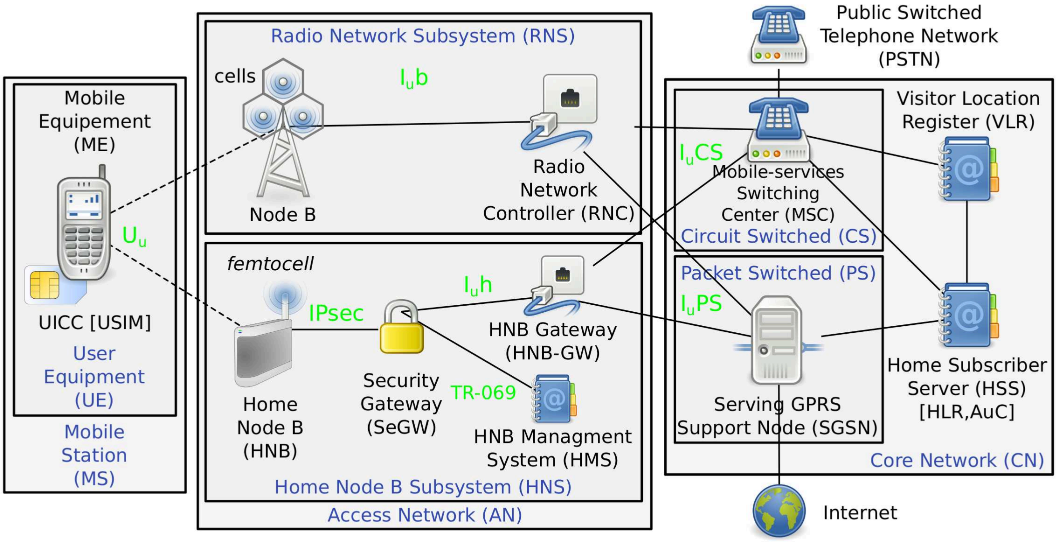
It is important to note that the protocol details around the paging procedure are equivalent in newer telecommunication standards as UMTS and LTE. In fact, the basic request/response nature of this protocol as well as the information included in exchanged messages is the same and has been adopted from GSM [22, 31, 32].

2.2 3G Architecture and Femtocell Network Infrastructure

We continue with a brief overview of the 3G architecture as defined by 3rd Generation Partnership Project (3GPP) [21] and as illustrated in a simplified version in Figure 2.3. 3G networks were introduced in 1999 with the birth of UMTS. Several reasons lead to this development, but the main interest had been to offer increased data rates and to lower costs of mobile data communications. Additionally, 3G offers a significantly more thought-through security architecture with respect to previous mobile telecommunication systems such as GSM

Nowadays, so-called femtocell access points provide 3G over-the-air capabilities for the masses. Therefore, we complete this background section with an introduction to femtocell technology, its network infrastructure, and details on how this technology is integrated within the existing architecture.

Figure 2.3: 3G/Femtocell Network Architecture


Mobile Equipement (ME)
UICC [USIM]
User Equipment (UE)
Mobile Station (MS)
Uu
Radio Network Subsystem (RNS)
cells
Iub
Node B
Radio Network Controller (RNC)
femtocell
IPsec
Iuh
HNB Gateway (HNB-GW)
Security Gateway (SeGW)
TR-069
Home Node B (HNB)
HNB Managment System (HMS)
Home Node B Subsystem (HNS)
Access Network (AN)
Public Switched Telephone Network (PSTN)
IuCS
Mobile-services Switching Center (MSC)
Circuit Switched (CS)
Visitor Location Register (VLR)
Packet Switched (PS)
IuPS
Serving GPRS Support Node (SGSN)
Home Subscriber Server (HSS) [HLR,AuC]
Core Network (CN)
Internet

A simplified representation of a typical 3G network architecture. To incorporate femtocell technology, the Node B has been replaced by the femtocell access point, the Home Node B. The femtocell communicates over a secure IPSec connection through the HNB gateway with the core network. The HNB Management System has been added as an additional femtocell specific component.

2.2.1 3G Architecture

A classical 3G network is divided into three main parts. Firstly, the subscriber part consists of Mobile Stations (MSs), most notably mobile phones, smartphones, and 3G modems. Secondly, the Access Network (AN) is responsible for connecting the wireless devices to the operator back-end network, which is known as the Core Network (CN).

The AN usually consists of multiple Radio Network Subsystems (RNSs). An RNS accommodates base stations called Node Bs (NBs) (equivalent to the BTS in GSM) and a Radio Network Controller (RNC) (equivalent to the BSC in GSM) managing them. The RNC acts as the gateway towards the CN and forwards all traffic originating from the MS that passes through the NBs. Both the original GSM and current 3G architectures trust the base stations: although the phones use link-encryption to communicate with the base station and the base stations could use link-encryption to communicate with the operator's back-end network, there is *no* end-to-end confidentiality or integrity between the user's phone and the carrier's network.

The third part is the CN that further contains two subsystems. As in GSM networks, the Mobile-services Switching Center (MSC) residing in the Circuit Switched (CS) subsystem is mainly responsible for call routing related tasks and maintains the circuit-switched model of conventional telephony, i.e. the Public Switched Telephone Network (PSTN). Conversely, the Packet Switched (PS) network subsystem primarily comprises a Serving GPRS Support Node (SGSN) to route data traffic and resembles a conventional packet-switched network.

The CS and PS subsystems both share a number of common infrastructural components such as the Visitor Location Register (VLR) and the Home Subscriber Server (HSS) which are critical components of access control and billing. As in GSM, the VLR acts as a database for temporary subscriber information belonging to subscribers within a dedicated geographical area served by this register.

The HSS incorporates the Home Location Register (HLR) and the Authentication Center (AuC) required to manage and authenticate subscriber information. Analogous to GSM, HLR acts as a central database storing all data associated with each registered subscriber of the operator. The AuC assists other components within the network by providing authentication and cipher key material needed to establish communications.

2.2.2 Femtocell Introduction

In recent years, there has been significant growth in Fixed-Mobile Convergence (FMC), mostly due to the popularity of Wi-Fi networks, inexpensive mobile data rates, and the increasing usage of smartphones. At the end of 2010, one fifth of more than five billion mobile subscriptions globally had access to mobile broadband [34]. Subsequently, mobile generated data traffic is rapidly increasing and has been predicted [43] to reach a compound annual growth rate (CAGR) of 92% between 2010 and 2015. Therefore, Mobile Network Operators (MNOs) started to explore different solutions to offload the increasing data traffic and bandwidth requirements towards networks such as Wi-Fi and femtocells, instead of expanding their existing expensive Third Generation (3G) networks. One estimate suggests [44] that 31% of the global smartphone traffic in 2010 was offloaded to fixed-line networks through dual-stack handsets (radio communication over Wi-Fi) or femtocells.

A femtocell is a small cellular base station that is typically deployed in home or business environments. It is connected to the operator's network via a broadband connection such as DSL. By deploying these inexpensive devices, carriers are offloading mobile data and voice traffic from their infrastructure to a fixed broadband line provided by the customer.

Furthermore, the customer takes care of the device installation and maintenance by simply attaching it to a local network. This enables operators to both reduce their costs and solve targeted reception problems in indoor environments. Unlike techniques where the phone directly offloads the connection through a Wi-Fi network, femtocells do not require handsets to operate in a dual-stack mode, as the femtocell is acting as a normal base station, offering improved radio coverage, high mobile data rates, and high voice quality to subscribers.

2.2.3 Femtocell Network Infrastructure

In order to integrate femtocells, or in technical parlance Home Node Bs (HNBs), into existing operator networks, a new subsystem has been added, namely the Home Node B Subsystem (HNS). This subsystem shares most major functionality present in the RNS of a conventional 3G network and connects to the same carrier back-end network. The femtocells act as a small cellular base station. Each femtocell communicates to the carrier network via the HNB GateWay (HNB-GW), which has similar functionality to the RNC component of a conventional 3G network. In particular, the HNB handles radio management functions whereas the HNB-GW acts as an interface to provide core network connectivity. As we explain in Section 3.3.1, the deployed HNB actually is a combination between the NB and the RNC.

Unlike the traditional telecommunication equipment, carriers deploy femtocells in environments that are not under their control. Thus, the standard introduced new network services in order to enforce security requirements and to allow operators to remotely control these devices. The Security GateWay (SeGW) component enables the femtocell to communicate with the carrier network in a secure way over an untrusted, shared broadband connection using a separate link-encryption layer to prevent eavesdropping or modification of traffic.

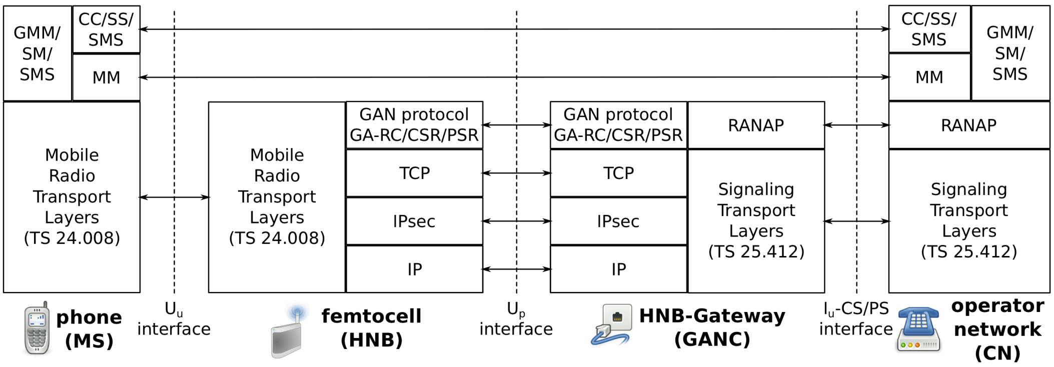
In order to allow the operator to remotely control the femtocell, the standard introduced the Operation, Administration, Maintenance, and Provisioning (OAMP) [20] server as a part of the HNB Management System (HMS) [12]. It acts as the central management entity within the network. While the main components and interfaces of the HNS are defined by the 3GPP, the implementation details are left to vendors and may differ among the various operator networks.

Just like a normal base station, the femtocell is effectively trusted: although there is link-layer encryption between the mobile devices and the femtocell, and between the femtocell and the carrier's network, there is no end-to-end integrity or confidentiality between the phone and the carrier.

2.2.4 Femtocell Core Network Communication: GAN Protocol

The integration of femtocells into the existing telecommunication network over the public Internet is a challenge for operators and vendors. The 3GPP defines the following three approaches to connect these devices to the core network over the so-called I$_u$h interface: I$_u$b over IP, SIP/IMS, and Radio Access Network (RAN) gateway based [73]. While these protocols differ in details, the architecture and

Figure 2.4: GAN Protocol Stack



The Generic Access Network (GAN) network protocol stack is used to communicate to the core network via the GAN controller implemented in the HNB gateway server. GAN protocol messages encapsulate the actual protocol messages.

supported features are all similar. We focus on the third approach based on a RAN gateway as our device uses this protocol. This technique utilizes the 3GPP Generic Access Network (GAN) protocol as the interface between the femtocell and the gateway [17].

The GAN protocol, formerly known as Unlicensed Mobile Access (UMA), was originally designed to allow mobile communication over Wi-Fi access points, enabling the phone to connect to the operator network over an IP network. This protocol was first standardized by MNOs in 2004 and lead to the GAN specification [16, 17] in 2005. The protocol transparently encapsulates all traffic generated by the phone and forwards it to the HNB-GW. This gateway is referred to as GAN Controller (GANC, see Figure 2.4). Similar to the RNC in traditional 3G networks, it is linked to the operator's CN using the I_uCS/I_uPS interface. For compatibility with the HNB architecture, the GAN protocol has been slightly extended (see Section 3.3.1).

The HNB acts as a gateway between the MS and the GAN Controller (GANC) as depicted in Figure 2.4. As in a classical 3G network, mobile phones connect to cells (in this case the HNB) via the U_u interface. Therefore, the presence of GAN is transparent to the subscriber's phone. Our femtocell device supports this protocol to enable mobile telecommunication via the customer's broadband connection. The GAN protocol implementation running on the HNB maps all 3GPP Layer 3 (L3) radio signaling to TCP/IP based GAN messages and passes them to the GANC. This enables the HNB to perform signaling tasks by sending encapsulated L3 messages to the GANC.

Details of these GAN messages are as follows. To map radio signaling from a specific subscriber to GAN messages, the femtocell maintains a TCP connection with the GANC for each individual subscriber. The connection management is based on Generic Access Resource Control (GA-RC) messages. The CS traffic is encapsulated in Generic Access Circuit Switched Resource (GA-CSR) messages, while PS traffic is covered by Generic Access Packet Switched Resource (GA-PSR) messages. We discuss the specific roles of these messages in Section 3.3.2. Additionally, GAN supports Mobile Application Part (MAP)) based signaling to control the telecommunication circuit and to manage the network [24]. This provides the necessary protocol functions for all Mobile Terminated (MT) as well as Mobile Originated (MO) services and thus supports full 3G functionality.

2.3 Related Work

In the last years, various attacks against cellular networks and their protocol stacks have been published. The work presented in this thesis highlights different types of practical attacks that are based on modified radio equipment. All the covered issues share common ground as they are based on architectural shortcomings and design problems in mobile telecommunication standards. Additionally, all of the resulting attacks allow an adversary to disrupt or manipulate services used by other subscribers within the same mobile network or even to compromise other network components.

We separate related work into four parts, femtocell security, modified equipment enabling IMSI catcher functionality, attacks that impersonate subscribers, and denial of service attacks. The first two types focus on attacks carried out from modified carrier-grade equipment, while the impersonation of subscribers and DoS attacks detail attacks originating from subscriber equipment.

2.3.1 Femtocell Security

Not only is it necessary to secure the femtocell subscribers and the device itself, but also the carrier's infrastructure against potential threats. While designing the femtocell security architecture and standards, the GSM Association and 3GPP have addressed such threats. In particular, they discuss security issues and potential attack vectors experienced during the life cycle of femtocell deployments [25, 81]. However, this specification frames various security aspects abstractly and does not address some of the new threats that we present. In the academic context, a few security groups analyzed the femtocell security challenges and requirements [99] and proposed new protection mechanisms [65]. Our research contributes to this, providing extensive experimental results by evaluating a commercially deployed femtocell.

Despite the security requirements, a few researchers have demonstrated weaknesses in such systems to gain root access [36, 41, 57, 71, 115]. Therefore, it shows that such a threat is real and these crucial security issues are not only present in a single device or operator network. However, most of this research has solely targeted the femtocell's security shortcomings and does not address practical attacks and their impact against the carrier infrastructure and end users. Our work differs from these groups by measuring the impact of integrating compromised femtocells into the mobile network infrastructure, which deals with security and privacy issues affecting both the end users and mobile operators.

2.3.2 IMSI Catchers

Some design properties of the mobile telecommunication protocol procedures come with security and privacy implications that have been exploited since years to perform tracking and interception of end users. A prominent example is the identification procedure, consisting in the request of the user identity by the network followed by a cleartext reply containing the user identity (such as the IMSI), is acknowledged in the 3G standard as a breach of the user identity confidentiality [18, p. 19, s. 6.2].

This procedure is exploited by the well-known "IMSI catcher" attack, which is the best known attack to mobile telephony users' privacy. It consists in forcing a mobile phone to reveal its identity (IMSI) [61, 111] by triggering the identification procedure from a fake operator base station (configured with the corresponding mobile network and country code settings). Furthermore, IMSI catchers are used to intercept and monitor communication.

These devices tend to be very expensive on the market. Wehrle and Paget demonstrated [95, 119] that such devices can be built using low cost hardware and Open Source software. However, their research exploits well known vulnerabilities of the GSM authentication protocol required to build such a device. Meyer and Wetzel demonstrated that under some circumstances it is possible to perform a UMTS MitM attack by downgrading a victim phone to use GSM [82].

These attacks take advantage of well-known weaknesses of the GSM authentication and key agreement primitives, such as the lack of mutual authentication and the use of weak encryption. These attacks allow an active attacker to violate the user identity confidentiality, to eavesdrop on outbound communications [83] and to masquerade as a legitimate subscriber obtaining services, which will be billed on the victim's account [35]. Instead of exploiting GSM weaknesses, we bypassed the problem of mutual authentication provided by 3G entirely, by turning an HNB into a low cost 3G IMSI catcher device in this work. Our work describes how an HNB can be abused to build a 3G IMSI catcher in practice.

2.3.3 Impersonation of Subscribers

Nohl and Melette demonstrated that it is possible to impersonate a subscriber for mobile originated services in GSM [72]. By first sniffing a transaction over-the-air, cracking the session key K_c, and knowing a victims TMSI, they were able to place a phone call on behalf of a victim. This works at least in networks that do not perform a fresh authentication procedure for every new service transaction. In this thesis we use a femtocell device in order to impersonate a subscriber that is currently booked into the femtocell. By relaying authentication challenges to a victim subscriber, we were able to send SMS messages or establish other services on behalf of the subscriber.

Additionally, we present attacks that do not only target Mobile Originated services, but also Mobile Terminated services. Thus, in our research we also consider, e.g., the called party instead of the caller as the victim. Contrary to attacking MO services, attacking MT services is time critical as it is required to win a race condition in the paging procedure.

2.3.4 Denial of Service Attacks

We consider relevant types of denial of service attacks in mobile networks that affect MT services for subscribers. We determined three types of denial of service attacks that fulfill this requirement: attacks directly targeting the victim phone, attacks targeting the carrier network, and attacks affecting subscribers, but without direct communication.

The first type comprises DoS attacks that target mobile devices directly, most notably phones. These issues are usually baseband/phone specific and caused by implementation flaws. Several vulnerabilities have been discovered in mobile phones that can lead to code execution and denial of service conditions [87, 120]. Particularly, the Curse-of-Silence flaw enabled an adversary to disable the MT SMS functionality of specific Nokia devices [113]. In [98] Racic et al. demonstrate that it is possible to stealthily exhaust mobile phone batteries by repeatedly sending crafted MMS messages to a victim. Consequently, the phone battery will drain very fast, eventually the phone will switch off, and MO/MT services can no longer be delivered to a victim. Our attack is inherently different from these kinds of attacks, because it is independent from the target device type and does not interact with the victim directly at all.

The second category consists of attacks that target the operator itself, and as a consequence also impact MT services for subscribers. These types are caused by design flaws. Spaar proved in [108] that it is feasible to exhaust channel resources of a base station by continuously requesting new channels on the RACH. Unlike our attack, this attack is limited to a single BTS and does not affect subscribers served by a different cell. Therefore, to attack a metropolitan area, an attacker needs to communicate with and attack every BTS in that area. Enck et al. [53] showed that it is practical to deny voice or SMS services within a specific geographical area by sending a large number of short messages to subscribers in that area. Serror et al. [105] exhibit that similar conditions can be achieved in CDMA2000 networks by causing a significant paging load and delay of paging messages via Internet originating packets to phones. A comparable resource consumption attack for 3G/WiMax has been demonstrated by Lee et al. [78].

As Traynor et al. outline [94], it is also possible to degrade the performance of large networks by utilizing a phone botnet and, e.g., repeatedly adding and deleting call forwarding settings. All of these attacks exhaust network resources mostly due to generated signaling load. As a result, services can no longer be reliably offered to mobile subscribers, effectively causing denial of service conditions. This includes MT and MO services. We exploit a race condition in the MT paging procedure and do not attack the core network itself. Our attack does not intend to generate excessive signaling traffic in the network. As a result, it is not prevented by proposed mitigations for these kinds of issues from previous research.

Our attack based on the paging procedure fits into the last of the three types of attacks that result in DoS for MT services. Most network attacks aim to abuse generated signaling to decrease the overall performance of the operator network. Attacks against mobile devices merely use the network as a bearer to deliver a specific payload to the phone. The third category is stipulated by attacks that target the mobile device itself, but do not send any payload to it. The aforementioned IMSI detach attack discovered by Munaut [89] can effectively cause that a service such as a call, will not result in paging requests by the network anymore. As described in Section 5.3.4, this design flaw even supports our attack. Contrary to this vulnerability, the paging response attack allows us to precisely control when and where a victim can be reached or not. After sending a detach indication, an attacker cannot control anymore for how long this state is kept. Mostly, because a phone can undergo the attach procedure again.

Our approach can be used either to hijack a session or to perform a denial of service attacks. We do attack mobile stations but neither by exhausting network resources, nor by directly communicating to the target device. We can target specific geographical areas, specific subscribers or a group of subscribers without the need to build a hit list of phone numbers residing in that area. Depending on the target, the attack can be either distributed or performed from a single phone. Additionally, the involved costs for this attack are as cheap as acquiring the required number of Motorola C1XX phones, which we use for our attack.

It is noteworthy that the attacks itself are not limited to be implemented on these specific devices. With the availability (e.g., [1, 55, 92, 93]) of cheap so-called software defined radios (SDRs), over-the-air-protocols become more and more accessible. Thus, it is now possible to use a single, flexible platform for different purposes and implement decoding and demodulation entirely in software (additionally to the protocol stacks itself). The hardware itself is not tied to only one use case or a specific frequency range for a particular protocol (instead a very broad range is usually supported), but provides a platform with which these aspects can be fully implemented in software. Further research based on this has already started [56] and can be expected to be continued by the larger research community in the future.

3 Malicious Use of Carrier-grade Equipment

Conducting attacks based on carrier equipment can be a very powerful method to compromise several aspects of subscriber security in mobile networks. Having access to carrier-grade technology does not only allow adversaries to attack end users, but such access is also regularly used by law enforcement agencies throughout the world [90]. Moreover, modern mobile telecommunication standards also specifically list requirements for lawful interception [33].

An attacker who manages to penetrate the core network of a carrier likely is able to carry out similar attacks. Alternatively, an adversary could try to tamper with equipment or even attempt to connect her own equipment to the carrier network to, e.g., inject signaling messages. Carriers usually deploy base stations, antennas, and other equipment at so-called cell sites. These are designated locations, e.g., often on buildings, that are connected to the operator network. An attacker who manages to connect equipment to these sites, may pose a significant threat to a mobile network operator and its users. However, to the best of our knowledge, no public records exist regarding such modifications in practice even though rumours exist that indicate that this does indeed happen. As these sites are usually at publicly inaccessible locations (e.g., on rooftops) or even facilities with basic security measures, they do not appear to be the most attractive target for an attacker though. The likelihood of getting caught and the additional criminal charges for breaking into such a cell site when carrying out attacks against subscribers may be one reason for that.

At this point it is important to specify what exactly qualifies as *carrier-grade* equipment. We define carrier-grade equipment as follows and will use this definition throughout the rest of this thesis: *carrier-grade equipment comprises equipment that allows its operator to provide seamless access to a mobile telecommunication* *network without limiting the expected number of available services or their use in practice.* In theory this could be any type of carrier equipment that participates in communication. Nonetheless, in practice this category is mostly formed by Access Network (AN) equipment, which connects the subscriber over the air interface with the carrier core network.

In the context of GSM, operating a rogue BTS qualifies for this definition. An overview of attacks based on false GSM base stations is provided in Section 2.3.2. The lack of mutual authentication in GSM makes it very easy to carry out such attacks. GSM BTS equipment can be bought on the public market and is often available on online auctioning platforms. It is noteworthy though, that this equipment is usually not cheap, which limits the usefulness of the availability for a populace with malicious intentions.

Starting with modern telecommunication standards such as 3G, operating carrier equipment has become more difficult. One reason for this is the lack of readily available Open Source software stacks to operate such equipment. Another reason is the lack of available equipment. 3G and 4G equipment is much more expensive and sparse than 2G equipment. Likewise, changes in the protocol specifications make it more difficult to carry out attacks based on false base stations, because starting with 3G, mutual authentication between the subscriber and the serving network exists and is mandatory. Therefore, an attacker who does not want to break physically into a cell site facility faces some limitations.

A recent trend by carriers is to offload traffic from their access networks to end users. As a result, this also brings carrier-equipment closer to the hands of adversaries. This trend is reflected with the introduction of femtocell technology. As of now, this is one of the few options currently on the market to operate carrier-grade equipment for newer protocol families such as those in 3G. Therefore, this part of the thesis will focus on femtocell technology and attacks based on such modified equipment.

3.1 Operating Carrier-grade Equipment

As of the first quarter of 2011, 19 operators have adopted femtocell technology in 13 countries around the world and many others are running field trials [59]. This number has increased to 46 operators in 25 countries during the fourth quarter of 2012 [59], including high profile operators such as Vodafone, Movistar, AT&T, SFR, China Unicom, and NTT DoCoMo. Informa estimated that carriers have deployed 9.6 million femtocell access points in 2013 [60] and forecasts it to reach 91 million at the end of 2016 [59]. Thus there are now large fractions of operator infrastructure, which communicates over the Internet and that is deployed at locations where users and adversaries have physical access to the equipment. Due to this situation, these devices may become an appealing target to perform attacks on mobile communication, or use them as a stepping-stone for attacks targeting the operator's network. Therefore, security is one of the top priorities for operators during the deployment process of these devices.

We demonstrate that it is a fundamental flaw of the 3G specifications to treat base stations as trusted devices. This becomes even more important in the context of femtocells. Femtocells involve different aspects of security including integrity of the device, access control mechanisms, and protection of the software update process. Among the top threats identified by the industry [14] are: booting the device with modified firmware; software and configuration changes; eavesdropping on user data; masquerading as other users; traffic tunneling between femtocells; Denial of Service attacks against femtocells and core network parts. While these threats are defined abstractly, their practical impact on mobile communication is rather unclear. We aim to measure the scale of such impact in a real operator network, conducting a practical security analysis of a femtocell device available to the public.

Despite the importance of femtocell security, it is well known [36, 41, 71, 115] that it is possible to get root level access to these devices. However, the negative consequences of rogue devices on mobile communication have not been thoroughly analyzed yet. In this thesis, we show that rogue devices pose a serious threat to mobile communication by evaluating the security impact of femtocell-originated attacks. We begin with an experimental analysis of security threats affecting end users; both end users deliberately using such a device as well as those who are not intentionally booked into the cell (e.g., by means of a 3G IMSI catcher). Furthermore, we evaluate the risk of femtocell-based attacks against the mobile communication infrastructure. This includes operator components and femtocells owned by other subscribers. We investigate how these components can be accessed and what type of network-based attacks are possible against them. Moreover, we argue how femtocell features in combination with common software vulnerabilities can provide a suitable environment to perform signaling attacks or allow turning the femtocell network into a global interception and attacking network.

While these devices run flavors of the Linux operating system, large parts of the functionality are provided by undocumented, proprietary binaries. Therefore, we conducted a vulnerability analysis of the femtocell and network architecture using a mixture of reverse engineering and experimental testing. In our work, we concentrate on a device deployed by the operator Société Française du Radiotéléphone (SFR): the SFR Home 3G femtocell [106]. SFR is the second largest mobile phone operator in France and has been among the early adopters of this technology [107]. However, due to the design of the femtocell architecture, most of the attacks presented in this work are not limited to this specific operator or device.

During our analysis, we have found several security critical attack vectors that can be leveraged to the previously mentioned threats defined by the industry. We will outline the risks of the femtocell technology that are caused due to a combination of operator specific configuration mistakes and problems inherent in the design of the femtocell architecture.

3.2 Compromising Femtocells

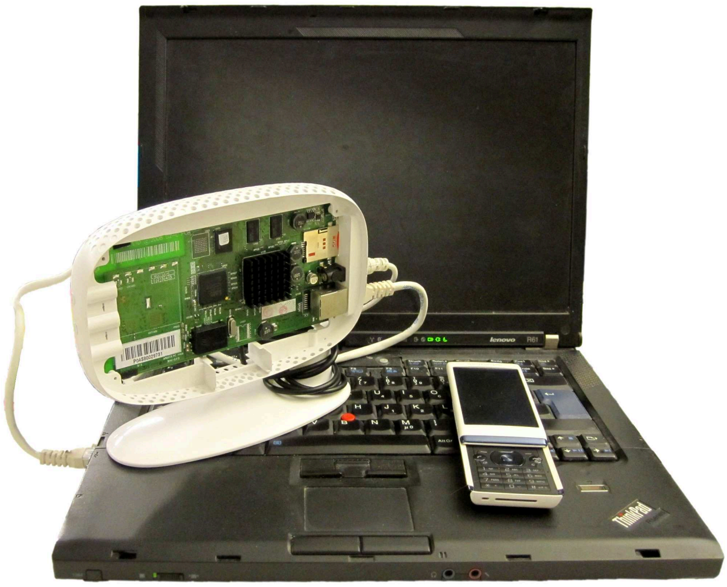
In order to operate the femtocell as a rogue base station and perform the attacks presented in Section 3.3 and 3.4, an attacker must acquire full control over the femtocell. Since femtocells are connected to the carrier's network, they must be secured to protect this critical infrastructure. This effort to secure the femtocell should include functionality such as mutual authentication between the device and the serving network, secure storage, secure network access, and secure communication [25]. However, due to the mass deployment of femtocells, carriers rely on a low cost per unit. Hence, femtocell manufacturers face a trade-off between secure hardware, software security, and low production costs. Consequently, the implementation often includes flaws that can be used to gain control over the device.

Common methods for initial debugging of embedded devices are test-pin probing, packet sniffing, network scanning, and reverse engineering. Attackers use test-pin probing to detect debug ports (i.e., UART or JTAG), or other techniques which researchers have used to gain root access on various commercially deployed devices [36, 41, 58, 71, 115].

Alternatively, as these techniques did not reveal obvious flaws in the device we studied, we examined the recovery procedure in order to compromise our device. If, for any reason, the femtocell is unable to connect to the carrier's network, the recovery mechanism enables the device to repair itself. The recovery procedure fetches and installs the latest working firmware images and configuration settings from the Operation, Administration, and Maintenance (OAM) server. We discovered two critical flaws in the implementation of the recovery mechanism on our device. Firstly, there is no mutual authentication between the OAM server and the femtocell. While the OAM server authenticates the femtocell, there has been no authentication of the OAM server. Thus, we were able to set up our own OAM service and modify its address by spoofing DNS replies.

Secondly, the firmware images provided by the OAM server were signed and encrypted. However, the implementation of this security mechanism included a trivial vulnerability: the OAM server provides the keys used to decrypt and verify the files in the configuration that is fetched by the femtocell. As a result, we were able to use existing images provided by the operator, add additional software and adjust configurations according to our needs, and deploy these modified images via the firmware recovery procedure. By using this method, we gained full control over the HNB [41] and were able to utilize it to perform the attacks presented in the next sections.

Figure 3.1: Experimental Setup



The setup used for our studies includes mobile phone to simulate a victim subscriber and an SFR femtocell running custom software, which is used in combination with a computer that monitors traffic and initiates attacks against the victim device.

Our experimental setup, as depicted in Figure 3.1, essentially consists of a victim phone, a rogue femtocell, and a computer utilized to monitor the network traffic, flash the device, and perform the presented attacks. Although we and others have exploited specific flaws in individual femtocells, due to the absence of substantial and expensive tamper resistance, we must assume that an attacker can always gain root access on any femtocell in their extended physical possession.

3.3 End User Threats

A great advantage for end users of a femtocell is the increased local 3G coverage, and thus higher mobile data bandwidth in their home environment. However, as we demonstrate in the following sections, end users using such a device are subject to several attacks when connected to a rogue femtocell. End users include subscribers who are knowingly using this cell (e.g, by using a femtocell installed in their home environment), as well as those who may not be aware of this, because the attacker has installed his rogue femtocell in an unexpected nearby location.

In both cases, the femtocell architecture has to ensure the confidentiality, integrity, and authenticity of mobile communication as well as the availability of mobile services to the registered subscribers. We show how all these aspects can be violated by a rogue femtocell. It is important to note that although the experimented attacks targeted a specific vendor and a specific protocol (GAN), these attacks rely on the trusted nature of femtocells in the cellular network architecture and thus are adaptable to different vendors, carriers, and devices.

3.3.1 IMSI-Catching and Call Interception (Confidentiality)

The confidentiality of the subscriber data is a very important security aspect in mobile communication networks. It is well known [95, 119] that it is easy to build an *IMSI catcher* device for GSM networks by using radio equipment and Open Source software. This type of attack leverages the fact that the network is not authenticated by the phone in GSM.

The 3G network procedures were not supposed to be vulnerable to IMSI catchers, as the phone authenticates the carrier network. Yet a rogue femtocell can be used to create a 3G IMSI catcher as follows.

Mutual Authentication/Over-the-Air Encryption. In contrast to GSM, the 3GPP defines [27] mutual authentication between the mobile phone and the carrier's network for 3G using a challenge response procedure assisted by the subscriber's Universal Subscriber Identity Module (USIM). If the carrier is not properly authenticating itself, the phone would not attempt to register with the network. Yet since a femtocell is an authorized and authenticated base station with an operator back-end connection, we can use a rogue femtocell as a cheap 3G IMSI catcher, circumventing the problem of mutual authentication without relying on protocol downgrading attacks. Because of roaming agreements, this is not limited to subscribers of the carrier that deployed the femtocell device. Thus, posing a serious threat to the data confidentiality of subscribers being booked into an HNB.

In order to provide mutual authentication, encryption and integrity protection, the femtocell acts as a combination of RNC and NB known from classic 3G networks. To understand call interception, we briefly describe the encryption and authentication procedure. The full details of this procedure are defined in [7, 27].

To guarantee the aforementioned security protections to subscribers, the femtocell receives an Authentication Token (AUTN), an Expected Response (XRES), a random challenge RAND, an Integrity Key (IK), and a Cipher Key (CK) from the carrier's AuC server. The RAND and AUTN values are forwarded to the phone. This AUTN is required by the phone to verify the authenticity of the network. By using a shared secret key K in combination with the random challenge RAND, the subscriber's USIM computes an Authentication Response (RES), IK, and CK. The resulting RES is required to authenticate the phone to the carrier's network and is sent in response to the femtocell, which compares it to XRES (expected response).

The IK and CK keys generated by the CN and transferred to the femtocell are not forwarded to the phone, but kept locally. This key IK is required by both parties to provide integrity protection for authentication and cipher algorithm selection between the femtocell and the phone. In contrast, CK is used as a key to encrypt over-the-air communication between the phone and the femtocell. The femtocell decrypts, encapsulates, and relays the *decrypted* data of mobile subscribers to the operator network. In the case of our device, the signaling traffic is transferred to the HNB-GW via the GAN protocol and the voice call data is encapsulated in an unencrypted RTP stream.

Thus the femtocell has established encrypted connections in two directions based on entirely unrelated cryptographic material. While the connection between the femtocell and the SeGW is encrypted using IPsec, the communication with the phone is encrypted using standard algorithms such as A5/3, with all communication transferred from one encryption scheme into the other on the femtocell. This is a crucial difference to traditional 3G networks in which the encryption and decryption is applied at the RNC.

As Over-the-Air (OTA) encryption support is mandatory and 3G protocols do not provide the necessary functionality for end-to-end encryption, the femtocell has to receive IK and CK from the core network. This in turn means that an attacker in control of the femtocell can sniff and manipulate traffic on the device. In the case of our device, the GAN protocol has been slightly extended for the use with femtocells to provide a *Security Mode Command* message that allows the operator network to transfer key material to the device for OTA encryption support [125].

Access Modes/HNB Configuration. A femtocell usually provides three types of access modes [51]: *open* access, *hybrid* (semi-open), and *closed* access mode. Most femtocells, including our device, default to closed access mode for residential deployments. The device receives a Closed Subscriber Group (CSG) list during the initial provisioning phase of the femtocell. The femtocell applies this list to enforce an access control policy that only allows registered subscribers to connect to it. However, we were able to change this access mode to open access via a hidden vendor web interface that contains basic security flaws. While there is a login page, the configuration pages can be accessed directly and thus bypass the authentication mechanism. The exact location of the specific pages can be determined by analyzing the firmware images. Since the access policy is a software feature, not a hardware restriction, and is enforced on the femtocell, a compromised femtocell can always bypass this control and change modes.

Changing this access mode to hybrid enables the device to allow any subscribers of a specific operator to connect to it, while open access allows any subscriber of *any* operator to access the network through the femtocell. The femtocell firmware usually supports the functionality to enable open access mode for its use in business environments or public areas. Thus carriers supporting open access mode also very likely support roaming between different operators via the femtocells (in particular via the GANC). Moreover, it is possible to change the Mobile Country Code (MCC) and the Mobile Network Code (MNC) of the femtocell to trick a victim subscriber into believing it is connected to the home operator. Thus, providing the basis to enable IMSI catcher like functionality. As roaming is allowed and the subscriber's home operator will provide valid AUTN tokens, mutual authentication is still performed successfully. Additional techniques on how to lure phones into using the IMSI catcher have been presented by Dennis Wehrle [119].

Circumventing IPsec. The femtocells connect to the back-end network via the SeGW, protected using IPsec or similar VPN technology. This means that even though it is possible to use the device as an IMSI catcher, it is not possible to directly eavesdrop on the subscriber traffic. However, there are multiple ways of sniffing this data as an attacker with root access to the femtocell. In our case, a user-space program is in charge of establishing the IPsec connection while a proprietary kernel module encapsulates the network traffic by means of Encapsulating Security Payload (ESP).

To allow the kernel to handle encryption of this tunnel, the user-space program has to pass the cipher material (HMAC, Cipher keys, Security Parameter Index) to the kernel (PF_KEYv2 interface). On our test device, this is performed using the sendto(2) syscall. By hijacking the libc provided wrapper function of this syscall and parsing the message, we were able to grab the key material for exfiltration.

Monitoring Voice Calls. With a successfully exfiltrated session key, we now can construct a sniffer to capture data in the IPSec stream. We built a small helper program, which uses the key material to decrypt the captured traffic. In combination with rtpbreak [50], it reconstructs the unencrypted RTP stream from the packets. The same helper program is also capable of extracting short messages and other user-generated critical data. The voice data is encoded in the RTP stream using the 3GPP AMR [23] speech codec in stream format. We also constructed a small utility based on OpenCORE [110] which can transcode the captured data streams into playable audio waveforms.

This allows an attacker to impersonate any operator by utilizing a rogue femtocell as an inexpensive 3G IMSI catcher and wiretap device. Consequently, adversaries can intercept mobile communication by installing the device in the radio range of a victim. Therefore, completely disregarding the original goal to prevent the active BTS attack in 3G.

This threat is a design problem in the current femtocell architecture since the communication is in-the-clear within the femtocell and a compromised femtocell can always exfiltrate key material. With the current femtocell architecture it is not possible to support end-to-end encryption of critical mobile subscriber data.

Moreover, mutual authentication is always properly performed, as the device is forwarding authentication tokens received from the corresponding CN of the victim's operator.

3.3.2 Modifying Traffic (Integrity)

Data integrity, not just confidentiality, is also critical. Yet we show that a rogue femtocell can also be used to compromise data integrity. We demonstrate such a type of attack by modifying outgoing Short Message Service (SMS) traffic for a phone which is communicating through our rogue femtocell. However, the same approach can be applied to any traffic generated by the phone as well as traffic directed to the phone.

As depicted in Figure 2.4, all traffic generated by the phone is passed to the GANC using the GAN protocol. Nevertheless, the phone is not aware of this protocol being used for the communication. The GAN protocol maps the Connection Management (CM) and Mobility Management (MM) layer messages of the 3G standard to a TCP/IP based network protocol. As soon as the IPsec tunnel is established and the device receives provision data, the femtocell attempts to build a permanent connection to the GANC. This procedure is based on GAN Generic Access Resource Control (*GA-RC*) messages. Additionally, the GAN protocol provides a Generic Access Circuit Switched Resource (*GA-CSR*) layer which is the equivalent to the GSM Radio Resource (*GA-RR*) layer. This layer is in charge of setting up a bearer between the mobile phone and the GANC. To successfully modify the mobile generated traffic, an attacker has to perform a Man-in-the-Middle (MitM) attack on the GAN traffic. Therefore, our attacks on such network-based signaling consist of the following two parts.

GAN Proxy. The first part comprises a GAN protocol proxy that forwards all signaling traffic between the femtocell and the GANC. In addition to this transparent proxying, it enables us to differentiate between GAN messages exchanged via this connection. Consequently, the proxy is able to detect incoming and outgoing GAN messages and in our example, track those that carry SMS data. Since the femtocell is under our control, we can reconfigure the device to communicate with our GAN proxy instead of the real GANC. This provides a simple method to force the HNB to communicate through our MitM proxy.

Attack Client. The second part consists of an attack client program that communicates with the GAN proxy over a slightly extended version of the GAN protocol. This client is able to inject or modify messages exchanged between the femtocell and GANC (thus not requiring OTA key material). To modify outgoing or incoming text messages, our client registers itself with the proxy to indicate that it is waiting for a text message. In the event of an outgoing message by a victim's phone, the proxy forwards its SMS to the registered attack client. Since all authentication is already complete, there is no additional authentication or encryption required. Our attack program is able to decode the forwarded so-called *SMS SUBMIT* message and allows changing either the message content or the destination number. Finally, it re-injects the modified message to the proxy, which subsequently forwards it to the GANC as if it was originating from the victim phone.

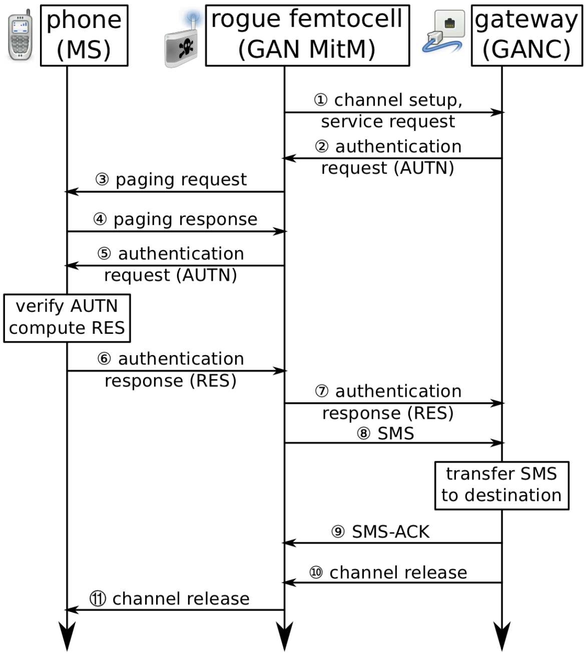
There is no way for the victim phone or the HNB-GW to detect that the message arriving at the operator network is not the same as the one that was originally sent. This demonstrates that, in the current femtocell architecture, it is impossible to ensure the integrity of subscriber data given an HNB is under control of an attacker.

3.3.3 Injecting Traffic/Impersonating Subscribers (Authenticity)

An even higher risk subscribers have to face is the complete impersonation of their subscriber identity. This means that an attacker is able to establish phone calls, send text messages or other data to the network while using a victim's subscriber information, without modifying any phone-generated traffic, allowing the attacker to bill the victim for attacker generated traffic. Reasonable threats include the abuse for social engineering, premium-rate service fraud, or simply the ability to make free phone calls. In this section, we describe how to perform such an attack, abusing subscriber information of a victim booked into a rogue femtocell. For the sake of simplicity, we demonstrate this by injecting an SMS on behalf of a victim using an attacker controlled femtocell.

In general, a phone attempting to use a service needs to issue a service request over a radio channel. To issue such a request on behalf of a victim, the victim's subscriber identity needs to be known to the attacker. Namely, the International Mobile Subscriber Identity (IMSI) or Temporary International Mobile Subscriber Identity (TMSI) needs to be known. For this reason, the developed GAN proxy additionally caches every subscriber information exchanged between phones and operator network. In order to impersonate a subscriber, the attack client registers itself to the GAN proxy and requests a fresh subscriber identity. The proxy returns the identity (depending on the availability either IMSI or TMSI) to the client. Afterwards, the proxy is able to map GAN messages received from the attacking client to the existing TCP connection of the specific subscriber. As service requests are always authenticated in 3G, the attack client and proxy have to additionally circumvent this authentication.

Figure 3.2: GAN Man-in-the-Middle Attacks



> By implementing a MitM attack on the GAN traffic it is possible to impersonate a victim by selectively relaying messages to a victim during the authentication procedure. As a result it is possible for example inject an SMS message into the network on behalf of the victim.

The actual attack, as illustrated in Figure 3.2, is performed as follows:

1. To send a text message, the attacker needs to setup a virtual radio channel over the existing TCP connection between the femtocell and GANC that belongs to the victim's phone. This is performed by sending a *GA-CSR REQUEST* message including an establishment cause indicating the reason for the resource allocation. After receiving, the GANC either accepts or denies this request.
   In case of an accept, the previously gathered subscriber identity is used by the attacking client to transfer a *GA-CSR UPLINK DIRECT TRANSFER* message to the GANC. This message carries the victim's subscriber identity and an L3 message indicating that the client is performing a service request and intends to send an SMS.
2. Since it is not possible to send text messages without being authenticated, the network replies with a *GA-CSR DOWNLINK DIRECT TRANSFER* message encapsulating an authentication request. The attack client can not properly answer this request as the secret key K that is required to compute the expected response (RES) is unknown (stored on the USIM of the victim's phone).

3. The proxy solves this problem by paging the victim subscriber. This paging process is a normal procedure to make a phone aware of an incoming service. The victim phone user does not notice the paging request as it is sent before an actual incoming service is displayed on the device.
4. When the phone replies to the paging request, the proxy forwards the authentication request to the victim phone.
5. Next, the victim phone answers the authentication request. There is no way for the victim to detect that this event has been caused by an outgoing service request from the attacker.
6. The proxy forwards the authentication response to the GANC and stops further communication with the victim device.
7. After this step succeeded, the attacking client continues the process of injecting a fake *SMS SUBMIT* message of our choice to the operator network.
8. The carrier acknowledges the successful SMS transmission and allocated channels are released (see 10-11 in Figure 3.3.3).

As mentioned before, the victim phone as well as the GANC are unable to detect this as long as the victim's phone is currently associated with the rogue femtocell. There is no way for the operator network to identify this message as being spoofed.

The impact of this attack is serious as the resulting billing for the service is based on the victim's subscriber identity. We verified this in a real operator network using prepaid SIM cards. Therefore, this attack clearly violates the principle of message authenticity. Furthermore, this injection attack illustrates that the HNB specification violates one of the basic 3G security objectives (Clause 5, item a) described in [8]. Namely, it fails *"to ensure that information generated by or relating to a user is adequately protected against misuse or misappropriation"*.

We have to stress that this attack vector is not limited to text messages, but can be applied to phone calls or data connections in a similar way. While it is already possible to spoof caller-IDs and short messages (e.g., via external SMS gateway providers offering this as a service), the effect from an attacker's point of view is slightly different. Unlike when using spoofing services, the victim is billed for the usage of the service. Additionally, it is hard if not impossible to track the source of the spoofed message as no external gateway is used and the traffic is originating from within the operator network.

3.3.4 Denial of Service (Availability)

Another threat are Denial of Service (DoS) attacks against subscribers using a rogue femtocell. It was previously discovered that the *IMSI DETACH MM* message is not authenticated in GSM and 3G networks [89]. This type of message is usually sent to the network when a mobile phone powers off. In particular, it represents the signal for the network indicating that the subscriber is no longer using the network and should not be paged for services. Because this message is not authenticated and no confirmation is delivered to the phone, an attacker can fake such *IMSI DETACH* messages to the network. As a result of this, the network assumes that the subscriber is disconnected from the network and thereby mobile-terminated services (incoming calls, text messages, etc.) are not delivered to the phone anymore. Therefore, the phone continues to assume that it is still connected and listens for paging requests.

Consequently, *IMSI DETACH* messages can be abused for DoS attacks against femtocell subscribers. However, since this message is delivered to the VLR, which is usually responsible for a certain geographical area, the attack can not be performed from arbitrary locations. Because *DETACH* messages carry the subscriber identity of the phone to be detached, an attacker additionally has to know the identity (IMSI or TMSI) that currently maps to the victim within the operator network.

While in a typical network this attack is limited due to geographical constraints, it is possible to abuse this behavior in a more serious way in the case of a network of compromised femtocells. Operators usually deploy a dedicated VLR that is used for all femtocell subscribers in the network even though the customer devices are from widely scattered geographical locations. We can only speculate, but we assume that this is done due to the lower number of subscribers in one area compared to traditional networks. In practice this means that it allows us to detach the complete subscriber base of a femtocell network given the knowledge of the mobile identities. The process of gathering these required mobile identities from femtocell subscribers is described in Section 3.4.1. The injection of *IMSI DETACH* messages is based on the capability to send arbitrary L3 messages by utilizing the GAN protocol. In order to attack other subscribers, we implemented a tool that initiates a new connection to the GANC by sending a *GA-CSR REQUEST*. As soon as this "channel" is established, the program continues by sending a *GA-CSR UPLINK DIRECT TRANSFER* message carrying the required L3 message for the detach process. This L3 message consists of a detach indication and the victim's IMSI, because unlike the TMSI, the IMSI is not random.

There may be situations in which the detach would not work, since the subscriber is currently identified within the network using the TMSI rather than the IMSI. However, in practice it is easy to bypass this nuisance by submitting an unknown TMSI to the network. If the network is unable to resolve the TMSI to a subscriber, it requests the client to identify itself using the IMSI. We developed a program to automate this process by looping over a set of mobile identities and sending the required network packets to the GANC. Due to the nature of using a separate channel for each subscriber, this approach can further be optimized by parallelizing this process. Hence, this Denial of Service (DoS) attack has a global impact on the availability of subscribers within the femtocell infrastructure. Our results show that it is possible to perform a large-scale DoS attack from a rogue femtocell against all subscribers currently using the femtocell infrastructure.

3.4 Infrastructure Threats

Rogue femtocells do not only represent a serious threat to subscribers, but also to network operators. Considering that these devices expose a certain part of the operator network to the device owner, it is important that infrastructural components are secured against attacks originating from within the network. Since femtocells are a part of the infrastructure as well, we also focus on their remote attack surface. Femtocells reside in the AN and not in the CN. Thus, not all components of the operator infrastructure are exposed to an attacker. However, several critical infrastructure elements exist within the AN. Additionally, it is possible to access some of the components in the CN by exploiting existing functionality of the femtocell. We conducted an analysis of the network entities exposed to a rogue femtocell in a real operator network and their potential for abuse.

This section is divided into three parts. The first part focuses on the possibility to collect information about other subscribers within the femtocell network. The second part analyses the attack surface of femtocells that may lead to a compromise of a remote device or the disruption of its services. The last part discusses how a combination of our findings can have a critical impact on the operator network as well as on the femtocell infrastructure.

3.4.1 Data Mining Subscriber Information

Given that an attacker can easily gain access to a single femtocell, it is interesting to know what parts of the network are exposed and what kind of information can be gathered. In the following paragraphs, we focus on which kind of information can be collected about mobile phone users via other femtocells within the femtocell's ecosystem.

Scraping. The aforementioned hidden web interface on our device is not only accessible to the device owner or operator, but also from other femtocells within the carrier's network. Thus, anyone connected to the SeGW is able to collect the information provided by the web interface. This interface includes several interesting bits of subscriber information that an attacker can possibly collect. This information includes:

- The International Mobile Equipment Identity (IMEI) and IMSI of the femtocell: This can be used to spoof the identity of HNBs as presented in Section 3.5.
- The IMSI and telephone number (MSISDN) of every user registered in the CSGs: Along with this information, it additionally indicates if a subscriber is connected or performing a call. Collecting this data from all femtocell subscribers is clearly a privacy issue. Furthermore, the IMSI can also be utilized to perform IMSI-detach attacks as discussed in Section 3.3.4.

- The neighbor macrocell list: The femtocell needs to perform a scan for neighbor macrocells to determine which radio frequency can be used by the device. Additionally, the list is used by the OAMP to perform location verification as required by the 3GPP specification [29]. By using this list, attackers are able to geolocate the device with high precision. This reveals the exact location of the device, and due to the limited coverage, also the location of the subscribers using it.

Even if a femtocell does not offer a web interface, the configuration parameters need to be stored on the device. An attacker successfully getting root access to other devices, as shown in 3.4.2, will be able to collect this information as well.

Performance Measurement Server. As specified in the 3GPP requirements for the management and the collection of performance measurement data [19, 20], the femtocell monitors the overall cell activity and submits this data in the form of reports to the PM server. In our case, the device uploads a report every hour using FTP.

The operator we examined used a common FTP account for all femtocells, with no additional file system based security constraints. This enables our rogue femtocell to read all such reports, including those submitted by other femtocells. Due to this flaw, it is possible to collect all measurement reports from all femtocells in the carrier's network. In particular, the reports contain the following information: the IMEI and IMSI of the HNB; the measurement date; the cell ID that is broadcasted over-the-air by the HNB; and the type, duration, and quantity of data transmitted (voice, video, or data traffic).

This enables detailed profiling of femtocells and connected subscribers in the network. While this issue may be an operator specific configuration error, the femtocell specifications require carriers to support this feature [20]. We believe that such information should be well protected and, if saved at all, in an encrypted form. Leakage of the IMEI and IMSI of each HNB exposes serious security threats since this data is used for enforcing access control policies (for example during the provisioning process) within the network. In addition, it certainly endangers the privacy of the HNB subscribers by disclosing their activity.

3.4.2 Gaining Remote Root Access to Femtocells

If an attacker manages to gain root access on other devices within the carrier's network, all the end user attacks described in Section 3.3 become even more serious threats.

Therefore, we conducted a security analysis of the attack surface that our test device exposes to a remote attacker. This includes running services on the device and protocols used to communicate within the network. We identified the NTP and

DNS networking protocols as well as a web server and a TR-069 provisioning service as attack vectors.

Besides NTP and DNS, the femtocell is not making use of any particularly exploitable network protocols. DNS is used to resolve NTP, SeGW, and OAM server names. NTP is often used by femtocells to provide frequency stability on the air interface [14]. Both of these protocols are implemented using standard Open Source software such as ntpdate and glibc functions. It is important to note that these protocols may be subject to spoofing attacks due to their UDP-based nature. Spoofing DNS to provide a firmware recovery server specified by an attacker seemed interesting. Nevertheless, it is not possible for a remote attacker, as this address is not resolved through the IPsec tunnel, but through the LAN interface, thus such an attack would require compromising ISP resolvers. Additionally, our test device is not making use of NTP authentication headers [85] which allows an attacker to spoof NTP server replies. This may impact the availability of the HNB by disrupting frequency adaption. However, none of the network protocols seemed to provide a practical way to gain root access to the device.

Instead, we focused more on the software services provided by the device. The configuration deployment is based on TR-069 [114], which has been adopted by the industry as the de-facto standard for remote provisioning. Both the OAMP services as well as the femtocell run a software stack implementing TR-069, which is based on the Simple Object Access Protocol (SOAP) over HTTP/S. SOAP itself is based on XML. The service providing this functionality is a proprietary daemon running on the femtocell. The provisioning port is also accessible from within the network, enabling the operator to proactively push configuration updates to the femtocell at any time. Additionally, our test device provides access to a web server that is also accessible within the femtocell network. The interface provided by the web server is used by the operator in order to debug customer problems and perform advanced configuration tasks.

Both of these services involve several protocols that are non-trivial to implement (considering the history of bugs in web servers and XML parsers [e.g., 47, 48]). We believe that these services are the most interesting attack vector. They will likely contain software vulnerabilities and poorly reviewed code (compared to the Open Source solutions used by the device). As often on embedded Linux systems, there is no user management on the system and all services run with root privileges. This makes the system services and protocols used by the device an attractive target to attackers.

In order to back up this claim, we conducted an analysis of the web server software by means of reverse engineering the proprietary binaries. As a result, we discovered a buffer overflow in the processing of one of the web server's supported HTTP methods. We were able to successfully exploit this vulnerability and acquire root access to the device. This attack vector enables us to gain control over the femtocells of other customers as the carrier we examined enables any given femtocell to communicate with all other femtocells over the VPN. Thus, it leverages the previously described end user threats in Section 3.3 to other femtocells in the network that are not under our physical control. The vulnerability is registered as CVE-2011-2900.

3.4.3 Leveraging Attacks Against Infrastructure Components

In this section, we describe how it is possible to further leverage the existing flaws to attack the operator infrastructure. We focus on attacks that do not target mobile phone users directly, but the availability and confidentiality of the network.

Signaling Attacks. It is well known that attacks based on signaling pose noteworthy threats to the availability of cellular network systems [54, 105, 116]. Femtocells support radio signaling and communication with back-end networks by design. The femtocell exchanges signaling messages with architectural components such as VLR, HLR, AuC, and SGSN via the HNB-GW to offer mobile services to subscribers. Therefore they also provide potential for abusing this functionality to perform signaling attacks. As described in Section 3.3.3, it is possible to send malicious traffic to the HNB-GW from the femtocell, using our attack client and the GAN proxy. This indicates that if such a device is compromised and configured maliciously by an attacker, it can be used to carry out signaling attacks against classical CN components. While the gateway might apply rate filtering rules, the femtocell is intended to be used by multiple subscribers and thus provides an advantage compared to using a malicious mobile phone for such attacks. Furthermore, the femtocell is communicating via a broadband connection with the back-end and is not subject to additional constraints caused by radio communication (e.g., frequency stability and synchronization). Therefore, it can be used to inject signaling traffic into a network protocol basis at a comparably high rate. A reasonable threat is to use the presented GAN protocol to flood the network with *Location Update Requests* that include different IMSI numbers for each request [116]. As a result, it might be possible to considerably increase the load of the network because it has to generate and store authentication tokens as well as keeping state of these requests. Sending these requests can be performed without any mobile phone and can be automated using the aforementioned attack client to generate the corresponding L3 messages.

Femtocell Botnets Naturally the impact of DoS attacks originating from a single femtocell is limited. Therefore, performing signaling attacks in a distributed manner seems far more practical. A remote root access vulnerability, such as discussed in Section 3.4.2, contributes to this by providing a possibility to build a femtocell botnet.

A number of characteristics that add to this are:

- Communication between femtocells is not filtered. It is important to note that the 3GPP standard explicitly allows communication between two femtocells [29].
- These devices are identical, making it a homogeneous network. Therefore a vulnerability discovered on one of these devices can be applied to all other femtocells within the network.
- Operators are actively deploying femtocells all over the country to extend 3G coverage, thus their number is growing rapidly. We identified around 5000 devices connected to the network of our target operator. The exact number of the deployed femtocells devices is not disclosed publicly by the operator, but as outlined in Section 3.1 these numbers grow steadily. We verified that these devices are indeed femtocells, by testing for the presence of the vendor web interface.
- Due to the fact that it is a small device not intended for direct user interaction, it is hard for users to notice behavioral changes of them.
- Finally, femtocells are supposed to be always reachable and connected to the carrier's network.

Therefore, elevating a remote software vulnerability into a channel to control other femtocells to carry out distributed signaling attacks seems feasible. Moreover, abusing a large number of femtocells allows sending signaling traffic at a low-rate in low-volume and thus evade known detection mechanisms [77].

Global Interception. Since direct communication between femtocells is permitted, it is possible to retrieve information from other devices, as demonstrated in Section 3.4.1. Besides gathering information, it also allows changing the configuration of other femtocells. The easiest way to achieve this is by using the web interface provided by the vendor. Another possibility to taint the femtocell configuration is to utilize a software vulnerability such as the root exploit mentioned in Section 3.4.2 and alter the settings in the device's database.

A crucial point of the femtocell communication is the selected HNB-GW address. By changing this address to an attacker-controlled IP address, it is possible to further extend the local user threats presented in Section 3.3 to a global threat affecting all femtocell-connected subscribers. Consequently, it allows an attacker to redirect signaling traffic of a victim's femtocell to an attacker supplied address. This could be running the previously mentioned GAN proxy. Therefore, attacks such as interception, modification or injection of arbitrary traffic can be leveraged to remote femtocells within the network.

Another important setting is the address of the SeGW. In particular, altering this address allows an adversary to force a remote femtocell to connect to an attacker-controlled SeGW. This can even be a machine outside the femtocell network running an IPsec server implementation. Even though the SeGW is authenticating itself based on a certificate stored on the femtocell, it can be simply replaced utilizing the root access. As explained in the next section, an attacker can connect to the SeGW without a femtocell in order to forward and intercept the traffic. Additionally, it is possible to reconfigure a victim's femtocell to act as an IMSI catcher and to operate in open access mode as explained in Section 3.3.1. As a result, not only registered subscribers, but also mobile phones that are in the radio range of a remote device can be intercepted. Being able to route traffic among femtocells, combined with the ability to reconfigure the devices, enables an attacker to turn the femtocell infrastructure into a global interception network.

We believe that this indicates the inherent need for secure storage in femtocell devices to deploy certificate or other information required for authentication with the SeGW.

3.4.4 Opening Up the Access Network

Traditionally, carrier networks used to be completely separated from other public networks and inaccessible for adversaries. This changes with the integration of femtocells into the mobile telecommunication infrastructure. As explained before, the SeGW has been introduced to provide secure communication between femtocells and the operator's network and restrict access to the AN. It defends against network-based attacks such as eavesdropping, injection, or altering of traffic. Additionally, it ensures authenticity of femtocells connecting to the network. Consequently, femtocell devices can securely communicate with the operator network via a public broadband connection and at the same time a separation between the Internet and the private operator network is maintained.

Due to this inherent requirement of network separation, the integrity of the femtocell is of high relevance. An attacker in full control of the femtocell can overcome this protection mechanism by tunneling traffic through the device. This can be easily achieved by installing, e.g., a SOCKS [79] proxy on the device or by using standard Linux network utilities such as iptables to transform the femtocell into a NAT router. Nevertheless, the necessity of utilizing such a device as a gateway to the operator network poses a serious limitation to attackers. Therefore, direct access to the SeGW would bypass this limitation. We show that it is possible to open up the access to the MNO network without a femtocell and thus increase the possibility to perform network based attacks against operator infrastructure.

The femtocell is communicating to the SeGW by an IPSec tunnel. The authentication procedure is based on EAP-SIM [66], utilizing the Subscriber Identity Module (SIM) placed inside the femtocell. Since the IPsec software on the device is based on a proprietary kernel module and user-space utility, it is difficult to figure out exact configuration details. However, by doing trial and error testing, it was possible to determine them.

Afterwards we applied these details to the configuration of a strongSwan [109] IPsec client running on a computer. Because it is possible to remove the SIM from the femtocell device, we were able to insert it into a smart card reader connected to this computer. As strongSwan directly supports EAP-SIM based authentication, this provides full connectivity to the carrier's network from a normal computer or any device capable of connecting a smart card reader.

Moreover, during our experiments it became clear that the setup does not even require a SIM that is provided within a femtocell (and obviously can be removed). Any valid SIM card from the operator was able to connect to the SeGW. Therefore, our experimental setup overcomes the natural limitations and requirement of utilizing a femtocell device to attack the network, enabling connectivity to the carrier network from any computer using a prepaid SIM card. Thus, paving the way for network attacks.

3.5 Femtocell Security Architecture Vulnerabilities and Analysis

In this section, we discuss the current femtocell security architecture and determine its effect on the existing 3G security principles. We present weaknesses in the authentication process of femtocells that we discovered during our experimental analysis. Additionally, we argue how 3G security concepts contrast with the design of the femtocell security architecture.

Impersonating Femtocells. According to 3GPP requirements [25], the femtocell has to register with the HNB-GW. This is achieved by sending *GA-RC REGISTER REQUESTs* that are based on the subscriber identity (IMSI). However, no additional security measures are applied to this message and the existence of the IPsec tunnel is independent from this procedure. Therefore, it is possible to exploit this functionality to register femtocells using spoofed identities. In our case, this registration process is based on the IMSI stored in the femtocell's SIM, but can be altered either by modifying the system software or by using the presented GAN proxy. In practice this means that by spoofing this message, an attacker can impersonate any HNB that is known to the network. Subsequently, the femtocell security architecture fails to provide adequate authenticity protection during the registration phase of the device. Possible implications of this attack may be the bypass of existing access control policies implemented at the HNB-GW. Furthermore, it may be possible to eavesdrop on communication by subscribers registered at the spoofed HNB.

Key Material/AUTN Handling. The lack of end-to-end confidentiality and integrity protection of the communication between the operator network and the phone is another inherent design problem of the femtocell-enabled security architecture.

To provide OTA encryption, the carrier's network transfers the relevant data to the femtocell including the AUTN, RAND, XRES, CK, and IK values [7, 27, 125]. These enable confidentiality and integrity of signaling and user-generated data between the MS and the HNB. The communication is decrypted on the femtocell device and forwarded unencrypted to the HNB-GW. However, as discussed earlier in Section 3.2, local as well as remote attackers can compromise femtocells in various ways. As a result, total control over the HNB always implies the possibility to violate the confidentiality of subscriber-generated data once the IPsec connection is bypassed. Moreover, authentication tokens are sent to the device during subscriber registration attempts. Because the device supports roaming subscribers, the GAN protocol provides a trivial way to request AUTN tokens from the network for any subscriber by any operator. As demonstrated before, those can be reused (for a certain time) in other attack scenarios [82].

Even though it is required to transfer such authentication information to the femtocell, we believe that this contradicts with the current 3G security architecture, as it affects the principles of integrity and confidentiality. While similar procedures apply to traditional 3G networks, it is important to note that those are not generally accessible by adversaries, physically or by means of network attacks. Given the history of vulnerabilities in various embedded network devices [46], it may be difficult to ensure the physical security of femtocells while at the same time maintaining low production costs. Solely relying on the device to ensure security of subscriber communication seems unpractical.

3.6 Security Considerations

Deploying such unsecured devices at user's premises and considering the sensitivity of the handled user data in light of the practical feasibility of potential attacks against the isolated telecommunication network systems exhibit further need for new additional countermeasures to prevent such risks. In Chapter 1 we argued that protecting against the so-called "*false BTS*" attack has been one of the core principles when designing the 3G architecture and the security thereof. Yet, we have demonstrated that new equipment such as femtocells defeat the purpose of these design goals by essentially giving an adversary the possibility to execute these exact attacks.

In the short-term, the only solution to this problem is applying secure development practices to the femtocell ecosystem, covering update procedures and software security on the device. However, by design it is for example not feasible to remove the session keys of the IPsec tunnel from the device. Moreover, due to a lack of protocol support for real end-to-end encryption between the subscriber and the serving network, it is furthermore not possible to mitigate the discussed attacks. For addressing the presented issues in the long-term, we see no practical solution other than using the femtocell device and in general base station equipment merely as a transport mechanism for an encrypted session with the subscriber instead of always depending on the requirement that the base station establishes the encrypted OTA session.

4

Additional Concerns Regarding Security Principles in UMTS Networks

"Privacy in residential applications is a desirable marketing option. There will be a market need for privacy in a public access service environment."

ETSI EN 300 175-7 Ch. A6,
Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 7: Security features

In the past chapter, we have shown how carrier-grade equipment can be used to mount attacks against subscribers and carrier networks. However, this is just one aspect of the security applications of such modified equipment. As outlined in the introduction of this thesis, the recent developments both in software and hardware for radio enable a new wave of security research in this field and with previously inaccessible telecommunication protocols.

This implies that attacks that were studied before from a purely theoretical perspective can now often also be evaluated in practice. Furthermore, it allows researchers to test explicit and implicit assumptions being made by modern telecommunication standards. In particular, wireless network protocols have to be designed in a way that considers an active attacker with arbitrary access to OTA communication. This is particularly crucial when designing wireless protocols and their privacy preserving characteristics. With the influx of social networks, location-based services, and new technologies, user privacy is at the forefront of current security concerns. It is however important to understand that such new technologies do not necessarily introduce the rapid need for user privacy, but rather highlight the importance and requirement of proper privacy preserving procedures to a larger part of the public.

To some extent, users of mobile networks have to accept the fact that they are giving up a part of their privacy. This can be explained by the need of the carrier network to know a users location simply to deliver a service. Likewise, mail delivery services usually need to know the home location of the recipient. However, a mobile phone user is a moving target and thus has to accept that the carrier is "tracking" the user in order to be able to fulfill the contract. In mobile networks, privacy and the prevention of tracking by third-parties is of particular importance, because of their omni-presence in our daily environment as argued in Chapter 1. Moreover, with more and more deployed base stations in the field to improve the customer experience, tracking by using the physical proximity to a cell site becomes more and more fine-grained at the same time. With the introduction of femtocell technology, this effect is even amplified.

In the following, we utilize a modified a femtocell device in order to demonstrate weaknesses that we discovered, which result in violating the 3G security principles and specifically their privacy properties in modern UMTS networks.

4.1 3G Security Requirements

3G aims to provide authentication, confidentiality of data and voice communication, as well as user privacy [18]. In particular, 3G privacy goals include the following [18]:

- **User identity confidentiality:** the property that the permanent user identity (IMSI) of a user to whom a service is delivered cannot be eavesdropped on the radio access link.
- **User untraceability:** the property that an intruder cannot deduce whether different services are delivered to the same user by eavesdropping on the radio access link.

In order to achieve these two privacy-related properties, 3G (and GSM) relies on the use of temporary identities TMSIs for identifying and paging mobile phones (more precisely mobile stations, MSs) instead of using their long-term identities IMSIs. Indeed, the eavesdropping of the IMSI in plaintext communications would allow the identification of mobile telephony users by third parties. Moreover, the 3G standard requires periodic updates of the temporary identity, to avoid the traceability of a mobile station by third parties.

New temporary identities are periodically assigned by the network through the TMSI reallocation procedure. The newly assigned TMSI is encrypted using a session key, which is established by executing the 3G Authentication and Key Agreement protocol (AKA). The 3G AKA protocol allows MS and network to achieve mutual authentication and establish a pair of shared session keys, namely a ciphering key and an integrity key. These keys are used to ensure the secrecy and integrity of the subsequent communications. While these mechanisms exist, a recent study highlights that those are often purely executed in commercial networks [37].

4.2 Novel Privacy Threats

In this section we describe two breaches of privacy, which expose a subscriber's identity and allow an attacker capable of sending and receiving messages over the air interface to identify the presence of a target mobile phone (MS) in a monitored area, or even track its movements across a set of monitored areas. As we will see, the attacker does not need to know any keys, nor perform any cryptographic operation. This kind of vulnerabilities usually look trivial once uncovered but often remain unnoticed for long time, since they do not involve fancy cryptography but are caused by errors in the protocol logic.

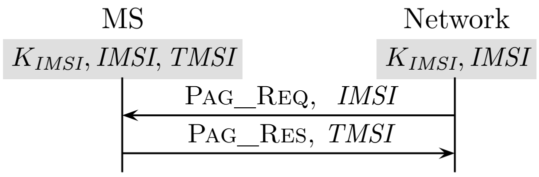
As argued in Chapter 1 and as witnessed by the attacks implementation presented in Section 4.3, a convincing analysis of 3G privacy and security should consider active attackers instead of passive ones. For this reason, we assume that the attacker has unlimited access to the radio link between the mobile station and the base station. He can sniff, inject, replay, and modify messages. This attacker model is the same considered in most of the previous work on GSM/3G security [35, 83, 126].

In the rest of this chapter, we consider a simplified network architecture. This architecture involves simply the mobile stations and the network. The network models both the Base Station (BS) which directly communicates with the MS on the radio link, and the complex structure of databases and servers connected with it and forming the 3G control network. Hence, we abstract away from any communication within the network and model only communication between mobile stations and the network. This abstraction allows us to hide details which are uninteresting for the purposes of our analysis and keep the models used for verification small, but at the same time precisely specify the interactions on the air between MS and network, which are the subject of our analysis.

4.2.1 IMSI Paging Attack

As briefly described in Chapter 2, the paging procedure is used to locate a mobile station in order to deliver a service to it, for example an incoming call. The paging request messages are sent by the network in all the location areas most recently visited by the mobile station in order to locate it and deliver a service to it. The paging request message is sent on a Common Control Channel (CCCH) and contains

Figure 4.1: 3G IMSI Paging Attack



The network can at any time send a paging message to the mobile device using the IMSI as the mobile identity. The phone will respond with its latest TMSI, thus revealing its temporary identity.

the identity of one or more mobile stations. The paging procedure is typically run using the TMSI to identify an MS. However, the IMSI can be used when the TMSI is not known by the network. A mobile station receiving a paging request establishes a dedicated channel to allow the delivery of the service and sends a paging response containing the most recently assigned TMSI (see Figure 4.1).

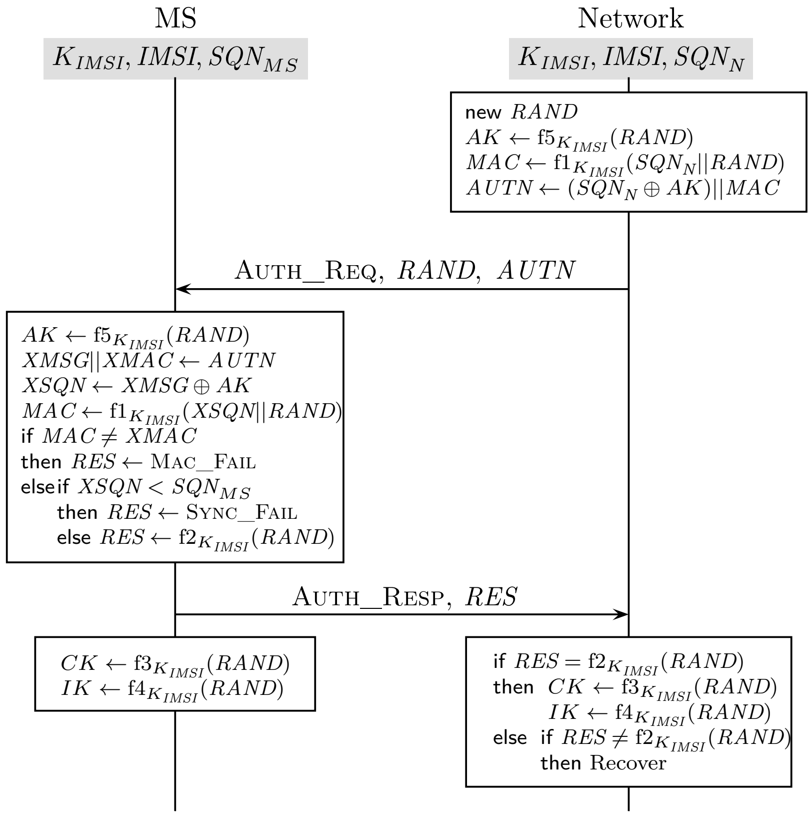
The possibility of triggering a paging request for a specific IMSI allows an attacker to check a specific area for the presence of mobile stations of whom he knows the identity, and to correlate their IMSI and TMSI. As we will detail in Section 4.3.1, in a real setting, the link between the paged IMSI and the related TMSI would need to be confirmed by replaying the attack several times.

4.2.2 AKA Protocol Linkability Attack

The Authentication and Key Agreement (AKA) protocol achieves mutual authentication between an MS and the network, and establishes shared session keys to be used to secure the subsequent communications. The MS with identity *IMSI* and the network share a secret long-term key, K_{IMSI}, assigned to the subscriber by the mobile operator and stored in the USIM. The secret key allows the MS and the network to compute shared ciphering and integrity session keys to be used for encryption and integrity check of communications.

The 3G AKA protocol [18], shown in Figure 4.2, consists in the exchange of two messages: the authentication request and the authentication response. Before sending an authentication request to the mobile station, the network computes the authentication data: a fresh random challenge *RAND*, the authentication token $AUTN$, the expected authentication response (XRES) computed by $\mathrm{f2}_K(RAND)$, the integrity key IK, and the encryption key CK (see Figure 4.2). The functions f1, f2, f3, f4 and f5, used to compute the authentication parameters, are keyed cryptographic functions computed using the shared key K_{IMSI} [26]. The authentication function f1 is used to calculate the message authentication code MAC; f2 is used to produce the authentication response parameter *RES*; the key generation functions,

Figure 4.2: 3G Authentication and Key Agreement (AKA).

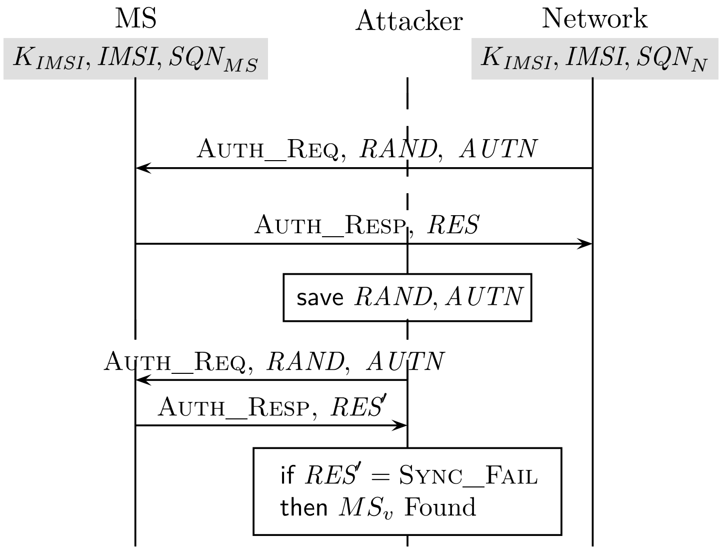

MS
$K_{IMSI}, IMSI, SQN_{MS}$
Network
$K_{IMSI}, IMSI, SQN_N$
new $RAND$
$AK \leftarrow f5_{K_{IMSI}}(RAND)$
$MAC \leftarrow f1_{K_{IMSI}}(SQN_N||RAND)$
$AUTN \leftarrow (SQN_N \oplus AK)||MAC$
AUTH_REQ, $RAND$, $AUTN$
$AK \leftarrow f5_{K_{IMSI}}(RAND)$
$XMSG||XMAC \leftarrow AUTN$
$XSQN \leftarrow XMSG \oplus AK$
$MAC \leftarrow f1_{K_{IMSI}}(XSQN||RAND)$
if $MAC \neq XMAC$
then $RES \leftarrow$ MAC_FAIL
elseif $XSQN < SQN_{MS}$
then $RES \leftarrow$ SYNC_FAIL
else $RES \leftarrow f2_{K_{IMSI}}(RAND)$
AUTH_RESP, RES
$CK \leftarrow f3_{K_{IMSI}}(RAND)$
$IK \leftarrow f4_{K_{IMSI}}(RAND)$
if $RES = f2_{K_{IMSI}}(RAND)$
then $CK \leftarrow f3_{K_{IMSI}}(RAND)$
$IK \leftarrow f4_{K_{IMSI}}(RAND)$
else if $RES \neq f2_{K_{IMSI}}(RAND)$
then Recover

f3, f4 and f5 are used to generate the ciphering key CK, the integrity key IK and the anonymity key AK, respectively.

The network always initiates the protocol by sending the authentication challenge $RAND$ and the authentication token $AUTN$ to the mobile station. $AUTN$ contains a MAC of the concatenation of the random number with a sequence number SQN_N generated by the network using an individual counter for each subscriber. A new sequence number is generated either by increment of the counter or through time-based algorithms as defined in [18]. The sequence number SQN_N allows the mobile station to verify the freshness of the authentication request to defend against replay attacks (see Figure 4.2).

The MS receives the authentication request, retrieves the sequence number SQN_N and then verifies the MAC (condition $MAC = XMAC$ in Figure 4.2). This step ensures that the MAC was generated by the network using the shared key K_{IMSI}, and thus that the authentication request was intended for the device with identity $IMSI$. The mobile station stores the greatest sequence number used for authentication, so far SQN_{MS}. This value is used to check the freshness of the authentication request (condition $XSQN < SQN_{MS}$ in Figure 4.2) to avoid replaying.

Figure 4.2: AKA Protocol Linkability Attack



The attacker obtains RAND and AUTN from over-the-air traffic and then later replays the authentication request to a victim subscriber.

With the help of the USIM, the mobile station computes the ciphering key CK, the integrity key IK and the authentication response RES and sends this response to the network. Next, the network authenticates the mobile station by verifying whether the received response is equal to the expected one ($RES = \mathrm{f2}_K(RAND)$). The authentication procedure can fail on the MS side either because the MAC verification failed, or because the received sequence number $XSQN$, is not in the correct range with respect to the sequence number SQN_{MS} stored in the mobile station. In the former case, the mobile station sends an authentication failure message indicating MAC failure (MAC_FAIL) as the failure cause. In the latter case, the authentication failure message indicates synchronisation failure (SYNC_FAIL) as the failure cause. When a MAC failure occurs the network may initiate the identification procedure. When a synchronisation failure occurs the network performs re-synchronisation.

To detect the presence of a victim mobile station MS_v, in one of his monitored areas, an active attacker just needs to have previously intercepted one legitimate authentication request message containing the pair $(RAND, AUTN)$ sent by the network to MS_v. The captured authentication request can now be replayed by the adversary each time he wants to check the presence of MS_v in a particular area. In fact, thanks to the error messages, the adversary can distinguish any mobile station from the one the authentication request was originally sent to. On reception of the replayed authentication challenge and authentication token $(RAND, AUTN)$, the victim mobile station MS_v successfully verifies the MAC and sends a synchronisation failure message. However, the MAC verification fails when executed by any other mobile station, and as a result a MAC failure message is sent. The implementation of few false BS would then allow an attacker to trace the movements of a victim mobile station, resulting in a breach of the subscriber's untraceability. The proposed attack is shown in Figure 4.2.

4.2.3 Formal Verification

While the paging procedure is obviously a breach of users' privacy, the traceability attack on the AKA protocol is much more subtle. Indeed, the messages exchanged through this procedure contain neither the IMSI nor the TMSI of the MS. So one could think that the AKA protocol provides untraceability by construction. But we just saw that this is not the case. Only careful analysis *w.r.t.* precisely defined privacy requirements could reveal this flaw.

We run the `ProVerif` tool on the IMSI paging procedure and on the AKA protocol[1]. `ProVerif` fails to prove the unlinkability and anonymity of the IMSI paging procedure and exhibits actual attack traces. In the case of the AKA protocol, the anonymity property is proved to hold, while the unlinkability property verification fails. Although, the trace provided by `ProVerif` is a false attack, it does give a hint of the real attack by highlighting the test of the MAC received from the network as the source of the problem. The adoption of formal verification tools during protocol design could have thus revealed design flaws. The interested reader is can find the full details of this formal analysis in [38].

4.3 Mounting the Attacks in Practice with Modified Equipment

In order to test the attacks presented in Section 4.2 in a deployed telecommunication network, we use a commercially available femtocell introduced in Chapter 3. Although, the particular femtocell hardware is tied to the network operator SFR, the proposed attacks are not. Indeed, we tested the attacks using mobile phones registered to different operators, hence just using SFR as serving network. The authentication token $AUTN$ is still provided by the victim's home network.

So by testing our attacks on T-Mobile, O2, SFR, and Vodafone victim MSs, we establish that all these tested networks are vulnerable to the attacks described above. However, we want to stress here that our implementation has the only purpose of showing the feasibility of our attacks and confirm that real cellular networks follow the 3GPP standard specifications and thus are vulnerable to the proposed

[1] The `ProVerif` code is available online [101]

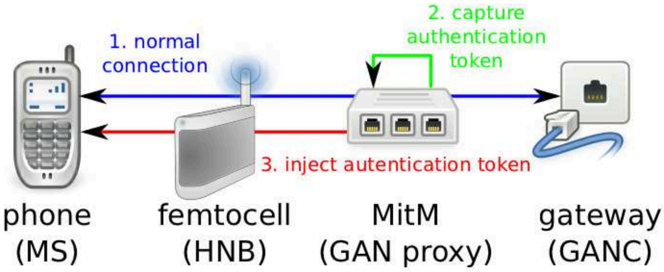
attacks. Due to the need of interoperability and compatibility between different MNOs (Mobile Network Operators), we believe SFR can be considered as a witness of the MNOs compliance with the standards and thus of their vulnerability to our attacks. Although experimentation on different MNOs and using different hardware is needed to confirm this conjecture, further testing is out of the scope of this study. This is not a restriction of our work, whose purpose is to show that our attacks can be mounted in practice, and that commercially available mobile phones do not implement any mechanisms to prevent them. The same attacks could be mounted by appropriately programming a Universal Software Radio Peripheral (USRP) [55], which is a hardware device able to emit and receive radio signals. In this case, one could obtain wider range attack devices in order to monitor larger areas.

4.3.1 Attack Procedure

For the purpose of implementing our attacks (Section 4.2), we use a compromised femtocell like the one described in Chapter 3. More specifically, we reproduce the hacking performed to gain root access of our femtocell and redirect the traffic to a Man in the Middle (MitM) GAN proxy, positioned between the femtocell and the GANC. We use this MitM GAN proxy as entry point for message injection. In particular, using the MitM GAN proxy we can inject messages into the connection between the MNO and the femtocell. The femtocell forwards these messages to the mobile phone, making them appear as if legitimately delivered by the MNO. To perform the attacks, we intercept, modify and inject 3G Layer-3 messages into the communication from the base station to the mobile phone in both directions, GANC-to-femtocell and femtocell-to-GANC. We redirect all the traffic between the femtocell and the GANC to our GAN proxy. The GAN traffic is cleartext travelling over an IPsec tunnel for which we own the key material, thanks to the initial rooting/hacking of the femtocell. Additionally, we developed a set of applications, which allow us to intercept, manipulate or insert selected messages, and distinguish different types of GAN messages. This allows us, for example, to cache subscribers information used to perform the attacks. In particular, we store the random challenge $RAND$, the authentication token $AUTN$, the TMSI and the IMSI of our victim MS. This information is directly extracted from the traffic that is passed through the MitM GAN proxy.

IMSI Paging Procedure Attack. To perform the IMSI paging attack, our software crafts a paging message encoding the necessary paging headers and parameters and a mobile station identity, i.e. one of the previously stored victim IMSIs. The crafted paging request is then sent by the GAN proxy to the femtocell. When the victim mobile phone receives the IMSI paging request, it readily answers with a paging response containing the victim's TMSI. Thus, by injecting a paging request, we can check whether a phone belonging to a designated victim is in the area covered by our device. In case of success, the phone generates the paging response, while

Figure 4.3: Experimental Setup



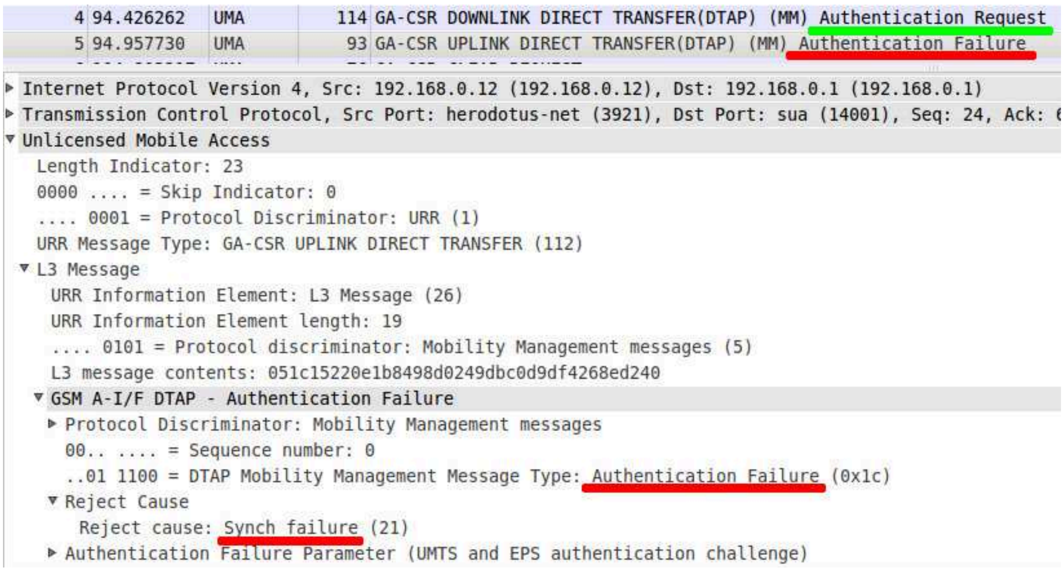
The victim device is connected via the malicious femtocell device with the core network. Via the GAN MitM capability that we implemented, it is then possible to replay previously collected authentication tokens during the AKA protocol attack.

a failed attempt generates no message. In general, it is possible that more than one phone replies to a paging request during the same time slot. However, one can repeat this procedure multiple times and correlate the timing and TMSI usage from the multiple replies as in [52].

AKA Protocol Attack. To perform the AKA attack we replay a given authentication message for a specific target for which the GAN proxy cached the legitimate authentication data, i.e. $RAND$, $AUTN$. This data is sent unencrypted on the radio link and could be captured with any equipment capable of sniffing the radio link. As soon as a dedicated channel is allocated to the MS, e.g., after being paged or when initiating a phone call, our software crafts an authentication request AUTH_REQ using the previously cached $RAND$ and $AUTN$, i.e. replays a previous request. This request is encapsulated into a GAN message and sent to the femtocell. The femtocell takes care of delivering the authentication request message on the dedicated channel assigned to the MS, as illustrated in Figure 4.3. The phone performs a validation of the authentication request and answers with the authentication response. If the response to the replayed authentication is a Synchronisation Failure (see Figure 4.5), then the MS on this dedicated channel is the victim's phone, and the victim is indeed in the femtocell area. Otherwise, the attacker needs to inject the same message to the other mobile stations in his area in order to find out if the victim MS is present or not.

The 3G AKA protocol is performed at each new session in the femtocell setting. This makes the caching of the authentication parameters very easy. Though, we do not have the tools to test if this applies when connecting to a typical Node B, we tested the 3G/GSM interoperability scenario by using the Osmocom-BB software and we observed that in this setting the execution of the AKA protocol can be triggered by calling for example the victim mobile phone a given number of times (by hanging up within a short time window this activity can be made non detectable

Figure 4.5: Linkability Attack



```
    4 94.426262   UMA        114 GA-CSR DOWNLINK DIRECT TRANSFER(DTAP) (MM) Authentication Request
    5 94.957730   UMA         93 GA-CSR UPLINK DIRECT TRANSFER(DTAP) (MM) Authentication Failure

▶ Internet Protocol Version 4, Src: 192.168.0.12 (192.168.0.12), Dst: 192.168.0.1 (192.168.0.1)
▶ Transmission Control Protocol, Src Port: herodotus-net (3921), Dst Port: sua (14001), Seq: 24, Ack:
▼ Unlicensed Mobile Access
    Length Indicator: 23
    0000 .... = Skip Indicator: 0
    .... 0001 = Protocol Discriminator: URR (1)
    URR Message Type: GA-CSR UPLINK DIRECT TRANSFER (112)
  ▼ L3 Message
      URR Information Element: L3 Message (26)
      URR Information Element length: 19
      .... 0101 = Protocol discriminator: Mobility Management messages (5)
      L3 message contents: 051c15220e1b8498d0249dbc0d9df4268ed240
    ▼ GSM A-I/F DTAP - Authentication Failure
      ▶ Protocol Discriminator: Mobility Management messages
        00.. .... = Sequence number: 0
        ..01 1100 = DTAP Mobility Management Message Type: Authentication Failure (0x1c)
      ▼ Reject Cause
          Reject cause: Synch failure (21)
      ▶ Authentication Failure Parameter (UMTS and EPS authentication challenge)
```

As visible in the wireshark dissector trace, the device responds using an authentication failure message, which contains a synchronization failure as the reason. Thus, we can deduce the subscriber identity.

by the victim [52]). For instance, our experiments showed that the execution of the AKA protocol on the UK Vodafone network can be triggered by calling six times the victim mobile phone, and hanging up before it even rings.

To illustrate the use of our attacks, consider an employer interested in tracking one of his employee's accesses to a building. He would first use the femtocell or software-defined radio (SDR) equipment to sniff a valid authentication request. This could happen in a different area than the monitored one. Then the employer would position the device near the entrance of the building. Movements inside the building could be tracked as well by placing additional devices to cover different areas of the building.imilarly, these attacks could be used to collect a large amount of data on users' movements in defined areas for profiling purposes, as an example of how mobile systems have already been exploited in this direction is available in [80]. If devices with wider area coverage than a femtocell are used, the adversary should use triangulation to obtain finer position data.

4.4 Discussion and Potential Mitigation

In this section we briefly discuss potential privacy preserving protocol changes to protect against the presented attacks. We have demonstrated two new violations of 3G privacy principles in UMTS networks. Despite the use of temporary identities to avoid linkability and to ensure anonymity of 3G subscribers, active attackers can rely on the paging procedure to break both anonymity and unlinkability. Moreover, the AKA protocol provides a way to trace 3G subscribers without the need to identify them in any way. While currently the attack depends on the operation of femtocell equipment, it is important to stress that the underlying problem is the lack of privacy preserving procedures in the protocol specification. In order to address the discussed scenarios, two different fixes are required to the protocol specification of which one is slightly more complex in practice.

To protect the IMSI paging procedure, a new shared session key called *unlinkability* key can be introduced. Such a key could be derived by applying a one-way function to the long-term key stored on the subscriber's SIM card and a random challenge that is sent with the paging request. This key should only be used for privacy preserving purposes. An MS receiving a paging request can use this unlinkability key to decrypt the paging message in order to test the contained mobile identity. By doing this, an adversary can observe IMSI paging performed over the air interface, but cannot observe the mobile identity that is contained in the request. Additionally, it needs to be addressed that these messages can be replayed by an attacker to test the presence of a victim in a certain location. For this, a sequence number similar as in the AKA procedure can be used to allow the victim to determine if the paging message was received out of order and may originate from an attacker who is replaying captured messages.

Addressing the AKA linkability attack is slightly more complex. For this we propose to introduce a lightweight public key infrastructure. Each MNO is in possession of a unique pair of private and public key. The public key allows a mobile device to encrypt privacy related information when sending it to the network. The key can be deployed in the mobiles' USIM. In order to address the AKA protocol linkability attack, it is important that the error messages sent to the network in case of an authentication failure do not leak privacy information due to the response that can be observed by an attacker. In case of a failure, the response message will contain all information that is required to either start the authentication procedure and resend a fresh authentication challenge or start the resynchronization procedure.

In the fixed version of the AKA protocol, the initial procedures are the same. The network will send $RAND$ and $AUTN$ to the mobile station and continue to wait for a response. In case the MAC and sequence number checks are successful, the procedure also carries on as expected. In case that either a MAC failure or a synchronization error occurs, the mobile station will send a constant failure message to the network. This message is accompanied by the IMSI of the mobile station as well as the sequence number.

In order to preserve the privacy of this information and specifically the contained IMSI, the mobile station encrypts this message using the public key of the network. In this situation, the network has all the needed information to start the re-synchronization procedure. Furthermore, it needs to be ensured that the network can authenticate this message in order to confirm that it was received from this specific subscriber. Therefore, the sequence number is encrypted with the unlinkability key that relates to the subscriber of the IMSI contained in the response message.

As a result of this, the network is able to deduce the error reason from the sequence number and the IMSI provided by the subscriber and can proceed with the recovery procedure. As Khan et al. [86] also notice, the practical implications of such changes would be immense. It seems almost impossible to address these threats without fully redesigning the entire system and associated protocols.

5 Malicious Use of Modified Subscriber Equipment

Mobile telecommunication protocols almost exclusively focus on the link between the MS and the serving network, the authenticity of the subscriber, and cryptographic means to protect subscriber privacy when it comes to security. The security model is usually very centric around billing, monetary interests, and the intention to preserve financial interests between communication partners. This often means that while the subscriber has to authenticate itself, the phone is essentially treated as trusted by the network. However, there is no guarantee that the underlying hardware and baseband implementation acts in the interest of the network or other subscribers though. This may become a concern in cases where the combination of hardware and software is more powerful and is able to operate outside the boundaries of the specification.

A typical mobile phone contains a GSM modem. This GSM modem is usually comprised of an RF frontend, an analog baseband, and a digital baseband. The RF frontend is in charge of physically receiving and transmitting on GSM frequencies via the GSM air interface (U_m). It also includes the antenna switch, power amplifier, and GSM band filters. The analog baseband is responsible for the modulation and demodulation and interfaces between the digital and the analog domain of the GSM baseband. The digital baseband processes the digital signals and implements the actual GSM protocol stack. The latter completely controls all aspects of the actual communication from Layer1 to Layer3 of GSM [121].

This is a simplified representation of the hardware/software contained in a phone when interfacing with a mobile network. However, the gist of this is that with a couple of constraints (e.g., those posed by filters), a phone is basically just a piece of radio equipment that is not necessarily limited towards the intended use.

5.1 Mobile Phones as a Threat

We believe that modified phones have to be considered as part of the threat model for mobile telecommunication protocols. This includes both hardware and software modifications. Some of the protocol procedures are intended to protect the network against misbehaving or misconfigured devices (e.g., the SIM related procedures of *AUTHENTICATION REJECT* messages). However, this does not consider equipment that is intentionally operating outside of its intended scope.

An example here is the work presented by Munaut et al.[91], which showed that it is possible to take advantage of certain Motorola phones by modifying the digital baseband firmware and remove the uplink filters [96]. As a result, it is possible to utilize the devices as a passive uplink sniffer for GSM or even utilize the devices as a BTS [112]. Using such modifications, it is possible to operate outside the boundaries of the protocol specifications, which assume that mobile devices work in compliance with well-defined specifications. However, the underlying hardware and software is more powerful and modified handset devices have to be considered in the threat model.

Additionally, there is no mechanism that would enforce assumptions made by protocol standards. Secure boot implementations that add cryptographic verification of modem images make it harder for attackers to tinker with protocol implementations. However, this mechanism does not exist on all mobile devices (especially low-end devices) and was not built to address this threat model, but rather allow carriers to fulfill other business requirements, such as implementing region locks. Beyond that, nothing prevents custom built hardware without the same restrictions. Examples are the HackRF [93] and bladeRF [92] projects, which are low-cost hardware platforms for software-defined radio. This means that none of the security critical assumptions made in modern telecommunication protocols can be based on expecting that device hardware and software works within the specified functionality only.

We will continue to show attack scenarios based on a modified digital baseband firmware implementation that is not operating as intended by the GSM specifications and as a result, can be used to attack other subscribers.

5.2 Attack Description

In this section, we will provide the theoretical background of our attack, introduce our experimental setup, and elaborate on the feasibility of such an attack.

5.2.1 Threat Models

Denial of Service Attacks. The first threat comprises an active attacker, interested in significantly *disturbing* mobile terminated services within a specific geographical area, e.g., a district or a part of a city. In certain situations it is desirable to ensure that a person or a device is not reachable via mobile telephony. For example a third-party may want to prevent a specific call from reaching the victim. The effect would be similar to the ability of selectively jamming incoming services for a set of subscribers. This includes individuals and groups of individuals. Such an attack would also have considerable business ramifications. While it would not compromise the general operation of the carrier, it would affect their revenue. The inability to receive a phone call will not only leave angry customers, it further impacts the generated billing as subscribers are charged when a call is connected. If any subscriber is able to place phone calls, but nobody is able to receive services, no profit is created. An exception here are short messages, as SMS operates in store-and-forward fashion and does not create billing on delivery of a message, but on its submission.

Mobile Terminated Impersonation. The second threat considers an attacker who aims to *hijack* a mobile terminated service. As a result, the service would be delivered to the attacker instead of the victim. This turns a passive adversary, who is able to observe air traffic, into an active attacker who can accept the mobile terminated service and impersonate the victim. For example an attacker could be interested in hijacking the delivery of an SMS message. Consequently, it is possible to read its content and at the same time prevent its submission to the victim. In practice this could, for example, allow an attacker to steal a mobile Transaction Authentication Number (mTAN), which is often used as two-factor authentication for online banking, or any other valuable secret from the message. We also consider an attacker who wants to impersonate a victim that is being called. By hijacking the MT call setup, it is almost impossible for the calling person to verify the callee's identity by means other than the voice.

5.2.2 Paging Response Attack Description

Our attack is inspired by two specific properties of GSM networks and its protocols.

Network State. GSM networks involve complex state machines [5] and face high amounts of traffic while operating on tight radio resource constraints. Consequently, it is desirable to keep states as short as possible.

Broadcast Information. The paging procedure is initiated on a broadcast medium, namely the PCH portion of the CCCH, and more importantly is performed before any authentication or cipher setup takes place. This implies that any subscriber, including an adversary phone, is able to observe paging requests for other subscribers, plus the inherent inability of the network to distinguish between a fake paging response and a genuine one.

As a net result, it is possible to exploit these aspects to send paging response messages on behalf of a victim being paged. The network stack can under no circumstances determine which of the replies is the legitimate paging response.

Denial of Service. The GSM documents do not specify the network behavior in such a situation. Therefore, the behavior of such a race condition is implementation dependent and may be exploitable. However, the state machine nature of GSM protocols suggest that if an attacker is able to answer a paging request faster than the intended subscriber, it will no longer be in a state in which it expects a paging response and thus will ignore the message of a victim. Consequently, the victim will receive a channel release message from the network. Next, the service setup will not succeed if the attacker does not provide the correct cryptographic keys required to complete authentication and cipher setup. Accordingly, the service setup cannot proceed and for example, a call will be dropped. The result is a novel and powerful denial of service attack against MT services that 1. does not rely on frequency jamming; 2. does not rely on resource exhaustion; and 3. is very hard to detect.

We verified that it is indeed possible to win the race for the fastest paging response time, as we will demonstrate. We were able to carry out such an attack in all major German operator networks including O2, Vodafone, T-Mobile, and E-Plus.

MT Session Hijacking. Exploiting the paging procedure does not only allow disturbing communication. It is important to note that in certain network configurations, this attack could be abused beyond performing denial of service attacks. Not all countries properly authenticate each service and use encryption. For example, only under 20% of the networks analyzed by the gsmmap project [104] authenticate mobile terminated phone calls 100% of the time. 50% of the tested networks only authenticate 10% of the services [72].

In such a network, an adversary can effectively takeover any MT service that is not authenticated and impersonate a victim. We assume a network without encryption and insufficient authentication as above. If the attacker is able to successfully exploit the race condition on the air interface, it is possible to directly hijack an MT service by following the protocol specifications. The paging response attack proceeds as in the DoS scenario. However, in this case, by winning the race, an attacker can accept, e.g., a victim's phone call or short message.

The victim of such an attack is thus faced with two consequences. For a mobile terminated call, it is not safe to assume that the called party is indeed the desired person. For short messages this implies that a message may not reach the victim, but additionally also that its contents cannot be considered secret.

Even if the network is configured to use encryption, an attacker is merely required to perform an additional step. In an encrypted network without proper authentication, the paging procedure is followed by the cipher setup. During this process to create an encrypted channel, the network sends a *cipher mode command* message to notify the MS of the encryption algorithm to be used. The *cipher mode complete* response from the MS indicates a completion of the cipher setup. In a network that uses encryption, this response has to be encrypted using the session key K_c as input to the A5 encryption algorithm. This session key is derived from a secret key K_i that is stored on the SIM card issued by the operator and a random challenge $RAND$ sent from the network to the MS. Due to the lack of perpetual authentication, an attacker can fully impersonate the victim after cracking the session key K_c and sending the *cipher mode complete* message. The cracked session key then allows to decrypt the subsequent communication that follows the cipher setup.

In practice, essentially both, the A5/2 cipher algorithm as well as the commonly used A5/1 algorithm, have been broken and demonstrated to be cryptographically weak [39, 40, 62, 102]. The session key can be acquired before hijacking the service by sniffing air traffic and using the kraken tool [103]. Also, some networks are configured to still use A5/0 [68], which does not provide any encryption. This further simplifies such an attack in those commercially deployed networks. Furthermore, for the subsequent paging response attack, an attacker does not even require physical proximity to a victim, because, as explained earlier, the carrier network is paging throughout an entire location area. In order to exploit this, an attacker requires a mobile device that enables him to observe traffic on the air interface and send arbitrary messages to the network. Additionally, a practical attack requires the fake response to arrive prior to the victim's message. Therefore, the attack is significantly challenging in terms of timing.

We successfully implemented both, the MT service hijacking and the denial of service attack. For the sake of simplicity, we obtained the session key through the SIM browser in the engineering mode of a Blackberry phone. Nevertheless, as outlined before this step, it can be obtained by a 3rd party by using a tool like kraken [103]. Cracking of K_c is merely a step that has to be performed prior to our attack, but is not part of the problem itself, which is the race condition. Given a known K_c, our code to take over an MT session, can hijack the transmission of a short message delivery in a real network.

It is important to note that the main reason for evaluating the paging race condition in GSM was the availability of freely modifiable hardware and software. However, modern telecommunication standards such as UMTS or LTE are making use of exactly the same paging procedure principles [22, 31, 32]. Insufficient cryptography and authentication further escalate the problem, but the root cause does not only pertain to GSM. We will continue to examine the requirements, boundary conditions, and feasibility of mounting such an attack in practice.

5.2.3 Experimental Setup

Launching such an attack requires hardware and software to interact with GSM base stations. More precisely, the attack relies on a device which allows us to modify its baseband (BB) implementation in order to control its radio communication. Traditionally this has been very difficult due to the closed nature of the GSM industry (phone manufacturers, baseband vendors, infrastructure equipment suppliers). For many years there existed no freely modifiable radio communication hardware with GSM stack implementations. As mentioned briefly before, while the GSM specifications are publicly available (very comprehensive though, over 1000 PDF documents), there are very few manufacturers of GSM equipment who have released any public documentation.

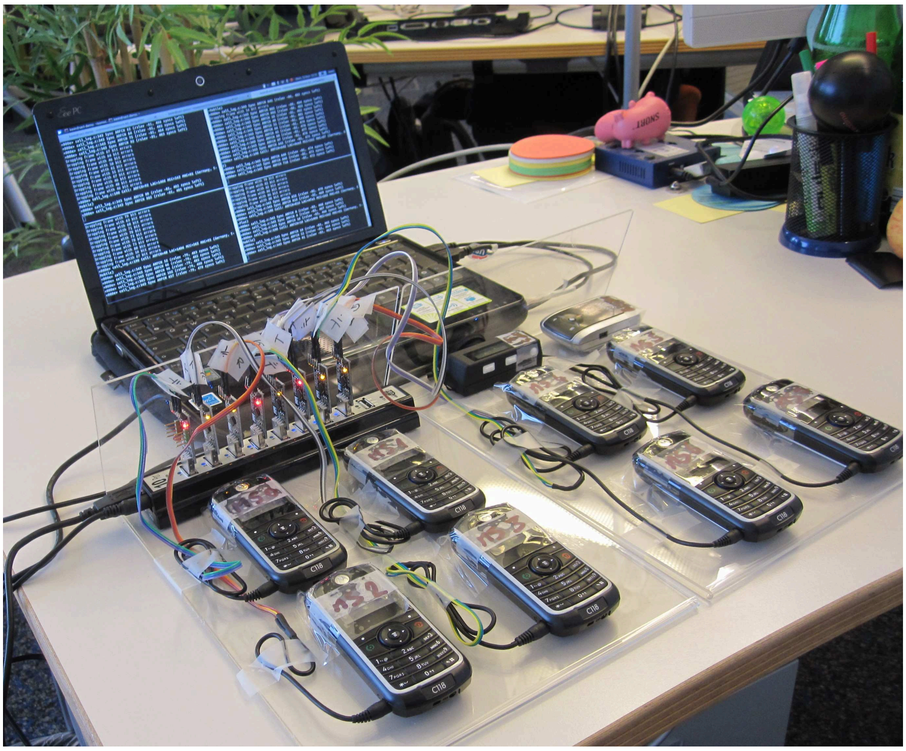
However, this situation has changed in the last years with the availability of inexpensive hardware such as the USRP [55] and various software implementations around the Osmocom [118] project. Additionally, in 2004 the source code of the Vitelcom TSM30 mobile phone was uploaded to a Sourceforge project [97], which allowed a broader audience to study a GSM phone stack for the first time.

Hardware Selection. There are basically three possible choices when it comes to the hardware selection of our desired radio device: *USRP* or other SDR hardware, *Vitelcom TSM30*, and certain *TI Calypso chipset based phones*. All of these devices can be utilized as GSM radio transceivers with software modifications. Yet some of these come with intrinsic disadvantages. First, for SDR hardware such as the USRP there is currently no GSM baseband implementation that allows the device to be used as a handset. While we could have implemented this, it would have been a very demanding task. Second, even though available, the TSM30 source code is a full-featured baseband implementation, which is too complex for our needs. Moreover, the availability of TSM30 devices is sparse and they are not easy to obtain.

Instead we used Motorola C123 and Motorola C118 phones, which are based on the TI Calypso chipset. These phones are inexpensive (around 20 Euros), easy to obtain in quantity, and more importantly can be used in combination with the Free Software baseband implementation OsmocomBB [123]. This enables us to receive over-the-air traffic and send arbitrary GSM frames.

Implementation. OsmocomBB implements a simplified version of the GSM stack. The GSM physical layer (L1) firmware runs on the phone, while the data-link layer (L2) and Layer 3 (L3) run on a computer as an application (layer23). L1 and layer23 communicate with each other via a UART serial connection. Layer 2 implements a modified version of the Link Access Protocol for the D channel (LAPD) used in Integrated Services Digital Network (ISDN) technology, called Link Access Protocol on the Dm channel (LAPDm). Layer 3 comprises three sublayers: Radio Resource management, Mobility Management, and Connection Management.

Figure 5.1: Experimental Setup



We use several Motorola C1XX phones with custom baseband firmware, GPS receivers, and a laptop for serial communication.

As our attack is based on paging, which is part of Layer 3, we required a modified version of the layer23 application. In practice our attack is particularly time critical, because we have to win a race condition on the air interface. It became evident that a layer23 implementation that runs on a computer is far too slow to win the race given the bottlenecks such as queueing between multiple layers, scheduling, and the use of UART serial communication. Consequently, we reimplemented a minimal version of LAPDm and Layer 3 directly in the L1 firmware to run the entire (malicious) software stack solely on the phone. Specifically this includes the paging protocol, which is part of the radio resource sublayer.

Figure 5.1 shows our experimental setup consisting of a laptop and several OsmocomBB phones. The serial cables are required in order to flash the firmware. Using this implementation we can camp on specific ARFCNs, observe paging requests within a location area, and send arbitrary GSM layer2/layer3 messages in a timely manner. Additionally, we used OpenBTS [49] in combination with a USRP as a BTS to test our setup and perform various measurements as later described in Section 5.2.5.

5.2.4 Targeted Attacks

Attacking individual persons requires our OsmocomBB phone to observe air traffic and respond to specific paging requests. In particular paging requests that contain the victims mobile identity. For privacy reasons, most network operators use TMSIs as mobile identities rather than the static IMSI. The TMSI is only valid within a location area and is subject to frequent changes [13]. Therefore, we need to determine the presence and the TMSI of a victim in a given location area.

For this we implemented the method proposed by Kune et al. to reveal the mapping between TMSI and subscriber [75]. We modified OsmocomBB's layer23 mobile application and introduced functionality that issues n (where n is 10-20) phone calls in a row. Next, the application terminates the connection before the target phone is ringing, but late enough so that the network generates a paging request. The victim phone does not ring during this early stage of the protocol flow, because it does not know yet what type of mobile service is incoming (e.g., call or short message). In our tests we empirically determined that, e.g., a time of 3.7 seconds after the *CC-Establishment confirmed* state has the desired effect in the O2 network. The exact timing may differ slightly, depending on the network that is used to initiate the call and the network in which the victim resides.

At the same time, a second phone is monitoring the PCH of any BTS within the target location area for paging requests. All TMSIs contained in the observed paging requests are logged together with a precise timestamp of the event. It makes sense to choose the ARFCN with the best signal reception to minimize errors and possible delays. By first limiting the resulting log to time ranges in which our calls were initiated, we can extract a number of candidate TMSIs. Further filtering the result set for TMSIs occurring in repeating patterns that reflect our call pattern yields to a very small set of candidate TMSIs or even single TMSIs. This process can be repeated to narrow down the set of candidate TMSIs to a manageable number. If the network uses IMSIs for identification, then an attacker could use the same process to determine the subscriber's identity. Alternatively, an attacker could use a Home Location Register query service to obtain the IMSI directly [100].

By default, the monitoring phone does not react to any paging request. After obtaining the victim TMSI, we transfer the TMSI via HDLC over the serial connection to the monitoring phone. This also changes the phone's role from a solely passive listener to an attacker. It starts to compare TMSIs contained in paging request with the supplied victim TMSI. On every match, the attacking phone promptly initiates the previously introduced paging protocol procedure to respond first. As a result, the paging response by the victim will be ignored and the call will be dropped unless we fully accept the service. At this point, it is impossible to reach the victim. To block MT services over a longer period of time, the subscriber identification procedure needs to be reissued due to TMSI reallocations over time [5].

Table 5.1: List of tested phones, their baseband chipset, and the respective baseband vendor.

Device model	Baseband chipset	Baseband vendor
Blackberry Curve 9300	Marvell PXA930	Marvell
iPhone 4s	MDM6610	Qualcomm
Samsung Galaxy S2	XMM 6260	Infineon
Nokia N900	TI Rapuyama	Nokia
Nokia 3310	TI MAD2WDI	Nokia
Motorola C123	TI Calypso	OsmocomBB[1]
SciPhone Dream G2	MT6235	Mediatek
Sony Ericsson W800i	DB2010	Ericsson
Sony Xperia U	NovaThor U8500	ST-Ericsson

[1] Layer1 paging attack code and modified layer23 application.

5.2.5 Feasibility

The success of such an exploit depends essentially on the response time of the attacker and victim devices. To achieve maximum impact, an attacker phone needs to always respond faster than the "average" customer device. The response time of the phone depends on a number of factors that are difficult to measure. This includes signal quality, weather, network saturation, application processor operating system, GSM time slots, and others. Yet, most of these parameters only have very little impact on the overall response time.

As the baseband chipset and its GSM stack implementation handle all radio communication, including the upper layer GSM logic, we suspect it to be a key contributor to a fast response time. We validate this claim by measuring the timing of various phones with different baseband vendors. Referring to a market report [2], Qualcomm and Intel alone accounted for 60% of the baseband revenue in 2011. Yet, relevant baseband chips and stacks that are currently available in mobile phones on the market are Qualcomm, Intel (formerly Infineon), Texas Instruments, ST-Ericsson, Renesas (formerly Nokia), Marvell, and Mediatek. We tested timing behavior for different phones for each of these vendors. Additionally, we also tested the response time for the OsmocomBB layer23 application to back up our claim that this implementation is too slow to perform our attack. Table 5.1 lists the tested phone models, chipset names, and the corresponding baseband vendor.

Timing Measurements. It is not feasible to modify the tested devices itself for measurements, as we only have access to the operating system on the application processor, and not the baseband. Furthermore, the phone could only guess when its response hits the serving network. Thus, in order to estimate the paging response time, we operate our own test GSM BTS based on a USRP and OpenBTS [49].

OpenBTS implements a simplified GSM network stack running on commodity hardware while using the USRP device as a transceiver. We patched OpenBTS to obtain timing information for the different steps during the paging procedure. Specifically, we are interested in the time a phone needs to acquire a radio channel and to send the paging response. This includes two parts of the paging procedure, the time between the initial paging request and the channel request, and the time between the initial paging request and the reception of the paging response. We log both of these timestamps for the relevant baseband vendors in nanoseconds using *clock_gettime(2)*. Additionally, we measure the same for an attack phone running our own lightweight, OsmocomBB-based baseband implementation. To trigger paging activity, we consecutively send 250 short messages, one after each channel teardown, to our test devices.

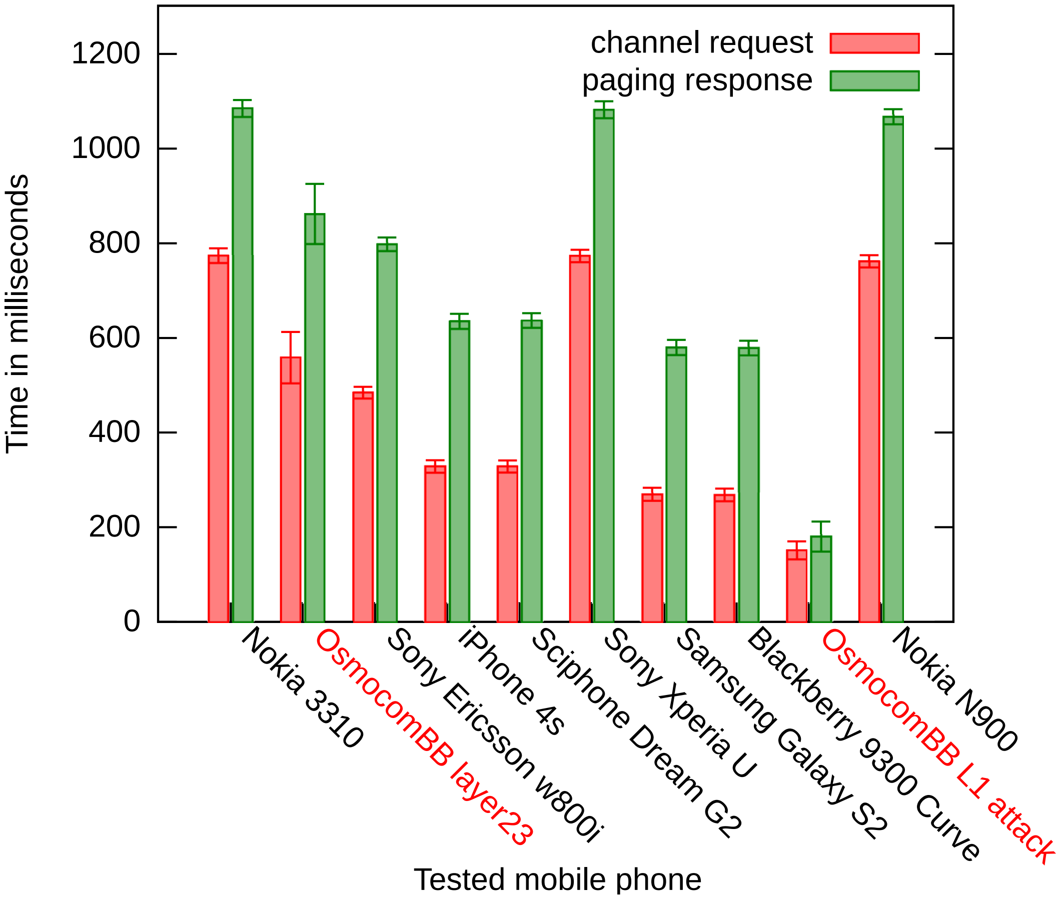
While we could have also used software like OpenBSC [122] in combination with a nanoBTS [69], we decided to utilize OpenBTS to be in full control over the transmission and reception. The nanoBTS is controlled over Ethernet, runs its own operating system, including scheduling algorithms, and cannot be modified. Thus, we used OpenBTS to minimize the deviation that may occur due to the nature of this BTS device.

Timing Observations. Figure 5.2 summarizes the results of our time measurements for each baseband vendor. It shows the elapsed time between the first paging request message sent to the phone, the arrival of the channel request message, and the occurrence of the paging response. Interestingly, the generation of the phone had little influence on the response timing. In our tests, a Nokia 3310, which is 10 years older than the tested Nokia N900, shows almost the same timing behavior. We do not have a definitive answer to explain this observation.

However, a plausible explanation can be found in the age of GSM. GSM was developed in the 1980s and most of the mobile telephony stacks for GSM are of this era. As most baseband vendors nowadays concentrate their efforts on exploring the technical challenges of 3G and 4G telephony standards, we believe that GSM stacks have not been modified for a long time. We do not expect significant modifications of baseband stacks by the respective vendors nowadays. Thus, we assume that timing behavior across different phone platforms using the same baseband will show similar patterns.

The most important observation from Figure 5.2 is that on average, with a confidence interval of 95%, our minimal OsmocomBB-based implementation is the fastest in transmitting the channel request and paging response. For our implementation, there is roughly a 180 milliseconds delay between the paging request and the arrival of the paging response. Thus, on average our attack implementation is able to transmit the final paging response prior to all other major basebands and can be conducted within the duration of a single multiframe (235.4 ms). This includes the OsmocomBB layer23 mobile application, which is significantly slower than our

Figure 5.2: Paging Response Time Measurements



We measured the time difference between initial paging request and subsequent channel request or paging response for different baseband vendors. Confidence interval: 95%.

self-contained layer1 attack software and shows similar timing performance as conventional phones. Therefore, with a very high likelihood, our software is able to win the race. It is also noteworthy that our lightweight stack can transmit the paging response almost immediately after the channel request (and reception of the Immediate Assignment). The test devices show a gap of at least 200ms before the transmission of the paging response. We expect that this is related to internal scheduling algorithms and queuing mechanisms between different layers of the baseband implementation.

5.3 Attacking Location Areas

Besides attacking individual subscribers, we show that it is also possible to leverage this attack to disrupt network service in large geographical regions. As explained in Section 2.1.3, the serving network does not always have the knowledge of the exact location a subscriber resides in. As a consequence, it also does not know which BTS is currently within a good reception of the mobile device. The phone announces a change of the location area by performing the *Location Update* [5] procedure. By monitoring *System Information* [5] messages on the Broadcast Control Channel (BCCH), a phone can keep track of location areas served by the BTSs within reception. The aforementioned lack of knowledge is compensated by the network by distributing paging requests throughout all base stations in the location area. This implies that an adversary is able to observe and respond to paging requests not only transmitted by a single BTS, but within a larger geographical region formed by the location area.

We already showed in Section 5.2.5 that we win the race for the paging response with high probability. Given that an attacker is able to answer all paging requests that can be observed on the PCH, it is possible to perform a denial of service attack against all MT services within the location area. Depending on its size, the impact of this would be massive, e.g., breaking MT calls in areas as large as city districts or even bigger regions. However, in practice there are a few obstacles to consider.

Depending on the paging activity, it is unlikely that service in an entire geographical region can be disrupted by a single attacker phone. In order to send the paging response, the MS has to tune to a dedicated channel. As a result, it would not be able to observe paging requests while being in dedicated mode. After sending the response, the attacker MS has to resynchronize with the BTS to observe CCCH/PCH traffic again. By logging timestamps for the various protocol steps, we measured the time for this procedure on the OsmocomBB side. On average we need 745 milliseconds to resynchronize in order to receive further paging requests after we sent the response. Furthermore, as shown in Figure 5.2, we need on average 180 milliseconds to transmit the paging response. This means that in ideal conditions, with a single phone, we are able to handle up to

$$\frac{60\text{s}}{745\text{ms} + 180\text{ms}} = 64.8 \text{ paging requests per minute.}$$

Depending on the network activity, this may or may not be enough to answer all paging requests. Additionally, we need to examine the different paging activities that can be seen in real operator networks. If the paging activity is very large, then the attacker may need to use multiple phones to perform the attack. Finally, to get an understanding of the impact of such an attack, we need to determine the size of the geographical region covered by a location area.

5.3.1 Location Area Paging Activity

An attack against an entire location area, e.g., in a metropolitan area, requires an adversary to respond to all paging requests in that area. Consequently, the efficiency of a large-scale attack depends on the operator specific paging activities and the allocated resources on the attacker side.

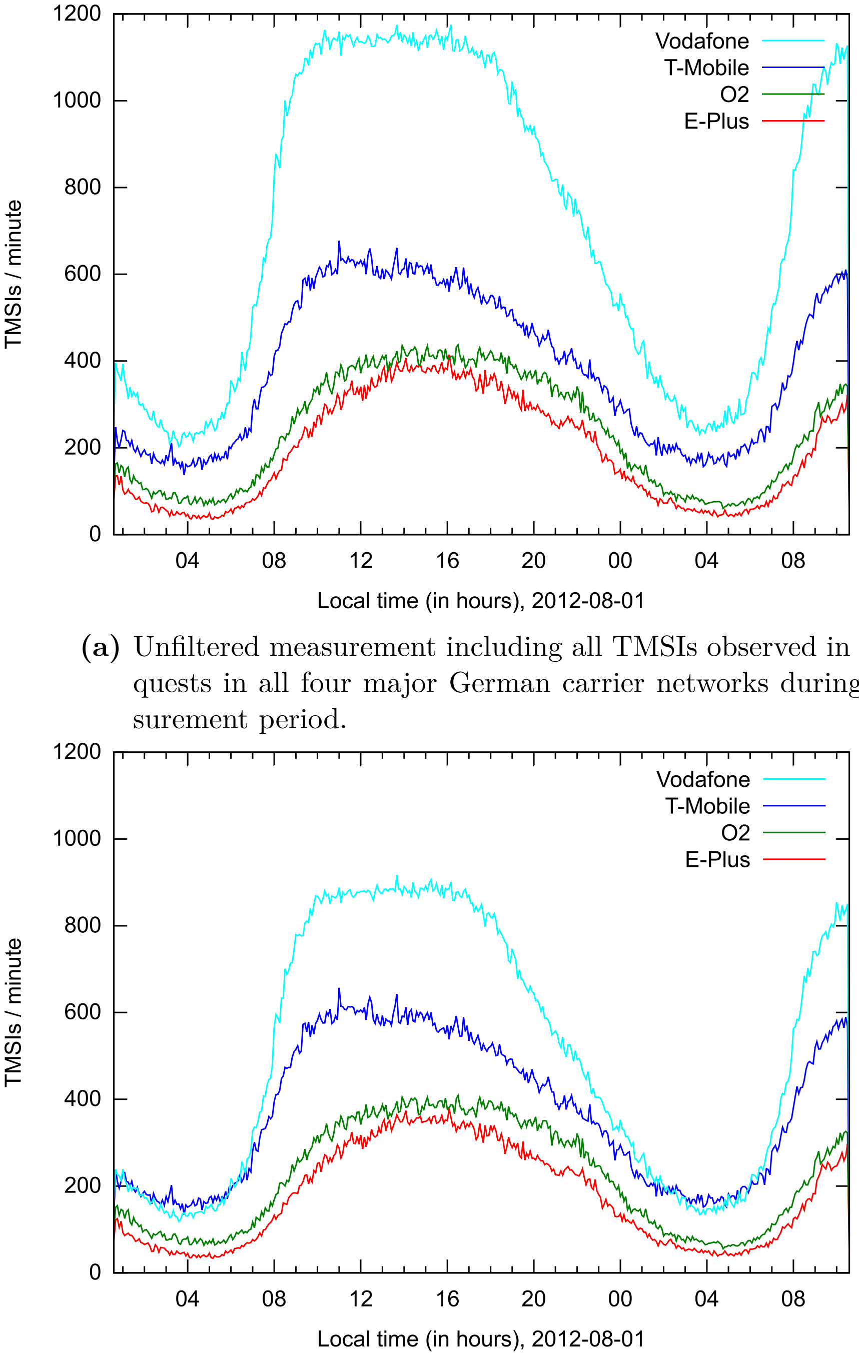
For the purpose of estimating paging activity, we modified our OsmocomBB stack to log all TMSIs in combination with a time stamp of its appearance. Because the paging requests are broadcasted throughout a location area, camping on one operator BTS for that area is sufficient to observe all paging activity for that area on the CCCH/PCH. We recorded the TMSIs in paging requests for all major German operators in a metropolitan area over a time period of 24 hours. The logs were created at exactly the same location, at the same date and time. We observed that in some cases the network is not paging with the TMSI but with the IMSI. For instance, if the subscriber is marked as attached to the network, but cannot be reached using the TMSI, the MSC starts paging using the IMSI. In this case, depending on the operator network configuration, paging may also be performed outside of the location area. However, this type of paging request is the minority and thus ignored in our measurements. Furthermore, assuming that a subscriber is present in the monitored location area, the network very likely already paged using the TMSI in this area. Obviously, it is simple to implement the attack in the case that network pages using IMSIs instead of TMSIs. In fact our code can also handle IMSIs.

Figure 5.3a summarizes the paging observations. The first observation to be made is that paging activity heavily varies throughout the time of the day. The observed pattern is not random, but rather reflects human activity during typical days. It is also interesting to note that the amount of paging requests heavily differs among the various tested operators. While for example E-Plus at peak times has a rate of roughly 415 TMSIs contained in paging requests per minute, Vodafone has almost 1200 in the same time period.

Such differences can be caused for example by the number of active subscribers in the network, or the size of the respective location area. During this measurement, we noticed several reoccurring TMSI patterns. Vodafone is actually always paging each TMSI at least two times. A second paging request is always issued two seconds after the initial paging request. This explains the massive amount of paging requests and we suspect this to be an attempt to improve the overall subscriber availability. Also, our logged data shows that some of these TMSIs are paged at regular intervals. We believe that these requests may partially be directed at M2M devices, e.g., for remote monitoring.

Figure 5.3b shows a filtered version of Figure 5.3a. Specifically, we filtered appearances of TMSIs contained in paging requests that we do not need to respond to. 3GPP TS 04.08 [5] specifies a timer, T3113, that is set on transmission of

Figure 5.3: 24 Hour Paging Measurement



(a) Unfiltered measurement including all TMSIs observed in paging requests in all four major German carrier networks during the measurement period.

(b) Filtered measurement result, removing TMSIs from paging requests caused by the expiry of the T3113 timer.

a paging request. If no paging response was received prior to the expiry of this timer, the network reissues the paging request by paging the mobile subscriber again. However, assuming that we are able to observe and respond to all paging requests, this retransmission would not occur during an attack. Therefore, these can be filtered from the result. By analyzing the logged TMSIs and the respective timestamps, we recorded the reappearance of each TMSI that was originally transmitted as part of a paging request. The vast majority of reappearances in time reach a common maximum which we assume is the timer value. A prevalent value seems to be five seconds. It is also reasonable that this is caused by a triggered timer. A normal call setup takes longer than five seconds [75] and short messages are queued at the SMS service center and likely transmitted over the same channel following one paging request.

As a result, the overall activity of relevance in practice is lower than the general amount of TMSIs contained in observed paging requests. The Vodafone measurements can be reduced by almost 22% during peak times and 33% during low activity times. However, due to the limited memory resources of the attacking phones, we cannot take this into account during an active attack.

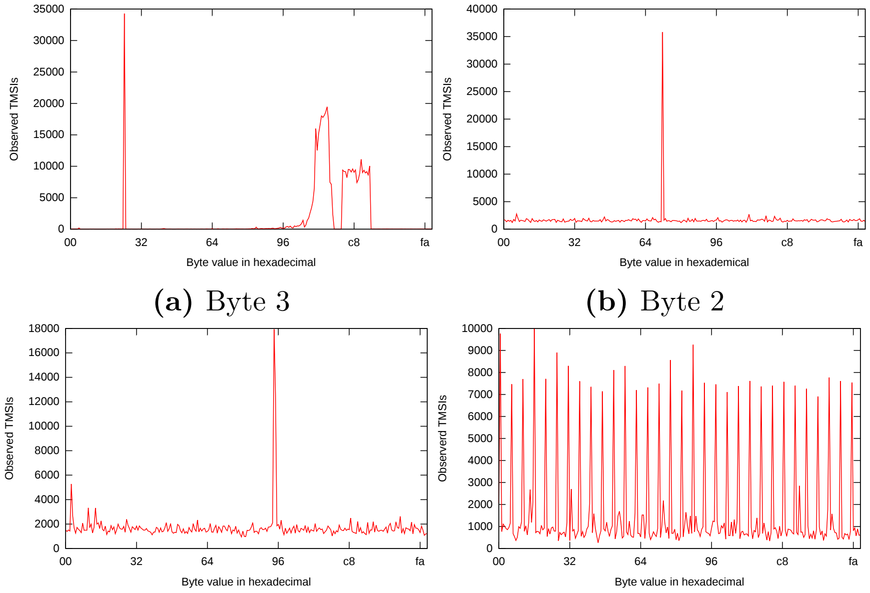
5.3.2 A Randomized Attack Strategy using TMSIs

The measured data in Section 5.3.1 suggests that even in location areas with low paging activity an attacker needs more than a single phone to respond to all paging requests. Thus, paging requests need to be distributed across multiple attacking phones. While serial communication could be used to coordinate these efforts, it also poses a significant slowdown. Consequently, using serial communication would lower the chance to win the race. We therefore decided to not make use of any actual communication between attacking devices, but to use a probabilistic approach instead.

For this, we analyzed the TMSI values to determine the statistical distribution of each individual TMSI byte as contained in respective paging requests. Namely, to prevent the collection of mobile subscriber identities and thus enable tracking, mobile phones are in most cases identified by their TMSI instead of their IMSI. To provide strong anonymity, a network should therefore sufficiently randomize those short term identities to provide unlinkability. A statistically uniform distribution would ease randomly distributing the paging load across multiple phones.

However, an analysis of collected TMSIs highlights that not all bits of the TMSI are sufficiently random or at least uniformly distributed. This may be, because some parts of the TMSI can be related to, e.g., the time of its allocation [11]. We also observed that certain bytes of the TMSI appear more frequently in specific ARFCNs. Thus, we further analyzed the distribution of each individual of the four TMSI bytes, for all tested operators. We use O2 as an example operator here even though nearly identical patterns can be seen for other carriers.

Figure 5.4: Statistical Distribution of TMSI Byte Values based on 437734 Samples from the O2 Network



(a) Byte 3 **(b)** Byte 2

(c) Byte 1 **(d)** Byte 0

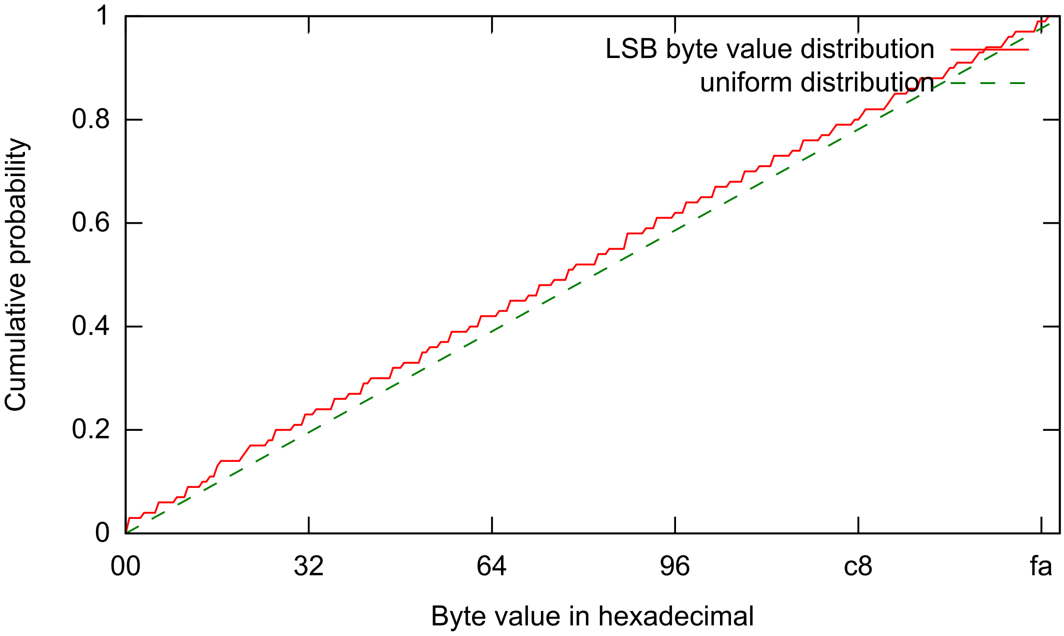
Figure 5.4 shows for each possible byte value how often a specific value was found in TMSIs contained in paging requests that we logged. As visible, all byte values are not uniformly distributed. However, the Figures 5.4a, 5.4b, and 5.4c show a significantly different pattern from Figure 5.4d. Not every possible value of observed values of the least significant byte (LSB) of the TMSI is encountered with equal frequency on the air interface. For example the value `0xff` is not used at all. Some seem to be more likely than others. Nonetheless, Figure 5.4d shows that value ranges are close to a uniform distribution.

This becomes more plausible in Figure 5.5, which compares the cumulative distribution function for the LSB values with the uniform distribution. We make use of this characteristic to delegate specific attack phones to dedicated TMSI LSB byte ranges. Thus, we can distribute the immense amount of paging between several phones by simply using randomization and thus avoid coordination at all.

Outliers for certain value ranges could be compensated by adding more phones to the specific range. To prevent recompilation of our OsmocomBB-based firmware for a distinct value range, we introduced a mechanism to configure the range at runtime. This mechanism is similar to the TMSI setting described in Section 5.2.4 and is based on a HDLC message over serial.

A similar distribution could be achieved by hashing TMSI values and assigning

Figure 5.5: Cumulative TMSI Byte Distribution



We applied the cumulative distribution function to Byte 0 (LSB) of all TMSIs contained in paging requests that were observed in the German carrier network of O2. The result is very close to a uniform distribution.

individual phones to specific hash prefixes. However for simplicity and to reduce the response time as much as possible we decided not to do this.

5.3.3 Mapping Location Areas

When performing a large-scale attack against a geographic region, we have to determine the size covered by the location area. Specifically, this knowledge enables an adversary to precisely plan the affected zone of such an attack. An attacker carefully selects the target location areas for specific regions and operators.

Location areas are not organized to cover equally large areas. As pointed out in Section 5.3.1, this impacts the paging activity that can be observed in a specific location area. Their size differs among operators and specifics of the covered environment. In fact, because of its impact on mobility management, location area planning is an important aspect for mobile network operators. Its size manifests a trade-off between subscriber-induced and network-induced performance degradation. Small location areas can cause a significant signaling overhead in the core network due to frequent location updates. It has already been demonstrated, that this can lead to denial of service like conditions [94]. A large location area causes additional load due to the paging overhead.

The Location Area Code (LAC), which is part of the Location Area Identifier (LAI), is broadcasted by each BTS in regular intervals on the BCCH via a *System Information Type 3* message. To map location areas, we use a slightly modified version of the cell_log application from the OsmocomBB tool-chain. cell_log scans all ARFCNs in the assigned GSM frequency spectrum for a carrier signal. It then attempts to sync to these frequencies and logs decoded system information messages as broadcasted on the BCCH. In combination with off-the-shelf GPS receivers, we determine the geographic location of the observed LAC.

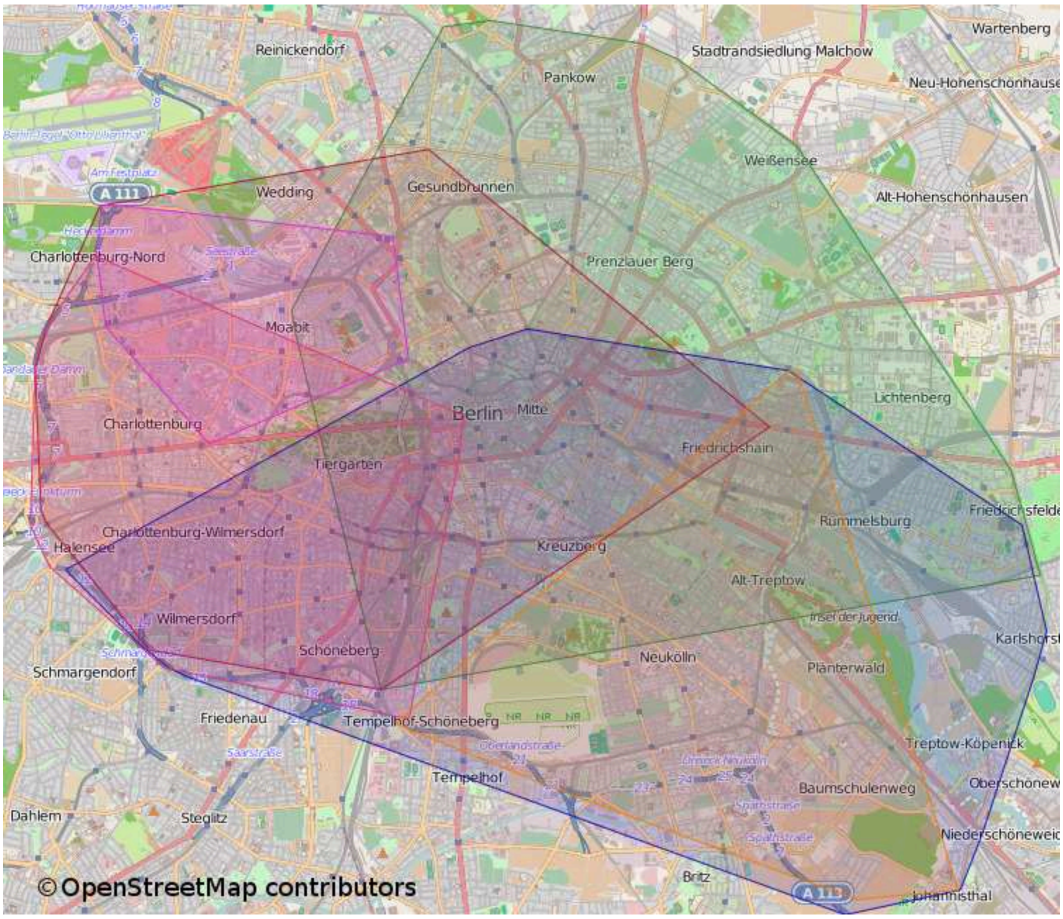
By slowly driving through the city in a car, we collected a number of waypoints and the respective GSM cells in sight. To minimize the loss due to driving speed, the scan process was performed simultaneously on eight OsmocomBB devices. In order to estimate the surface covered by a location area, we calculated the convex hull of points within the same LAC. The size of location areas in a metropolitan area such as Berlin varies from 100km^2 to 500km^2. From our study, the average location area in Berlin covers around 200km^2. Data from OpenCellID and Crowdflow [74, 76] indicate that outside of the city center location areas exist that cover over 1000km^2. Figure 5.6 shows location areas that we mapped for one of the four major operators in Berlin.

Most location areas partially overlap with geographic regions that are part of a different area. These results provide a rough insight on dimensions of location areas in a metropolitan area. It also shows that a large-scale denial of service attack based on the paging procedure has a significant impact to a large number of subscribers.

5.3.4 Amplification of the Paging Response Attack

The attack procedure as introduced in Section 5.2 does prevent MT services from being delivered to a subscriber. However, it does not provide a persistent way to cause denial of service conditions. Access to mobile services is denied as long as an adversary is running the attack. Accordingly, calls reissued by subscribers to reach a person, have to be attacked again, which may further raise the paging load. To prevent this, we make use of another attack that has been publicized before. Munaut discovered that the *IMSI DETACH* message is not authenticated in GSM and 3G networks [89]. As a result, an attacker can easily craft detach messages on behalf of a victim. This message acts as an indication to the network that a subscriber is no longer marked as available for the carrier. As a result, the network marks the mobile station as detached and will no longer page the subscriber until it reassociates with the network. Consequently, this stops the network from delivering MT services. During normal operation, this message is generated by the phone and sent to the network, e.g., when it is being switched off.

Figure 5.6: Location Areas in Berlin



© OpenStreetMap contributors

We mapped the Location Areas of Vodafone Germany in Berlin during a field test by slowly driving through the city center and correlating observed cell ids with GPS coordinates.

The mobile identity contained in the detach indication message is not limited to IMSIs, but can also contain TMSIs. By combining the paging response attack with the IMSI detach attack, it is therefore possible to amplify its effect. After each paging response, our OsmocomBB implementation reuses the collected mobile identity to send a detach message. Accordingly, our attack ensures that an initial call to a subscriber will be terminated and that reissued services such as calls will not cause paging activity again. Thus, by doing so, we effectively reduce the paging load over time.

5.3.5 Large-scale Attack Feasibility

We continue to evaluate the feasibility of a large-scale attack using multiple phones against commercially deployed networks. The transmission of a large number of RACH bursts and SDCCH channel allocations may be limited due to radio resource bottlenecks. We therefore verify, whether or not a single cell provides enough resources, or an attack needs to be conducted in a distributed fashion.

Prerequisites. In the following, we denote the TMSI paging request activity as r_{request} and the number of required phones to handle this respectively as n_{phones}. As discussed in Section 5.3.2, the TMSI LSB value range is used to equally distribute the paging load across multiple attacking phones. Therefore, assigned phones need to wait for a range match. On average this requires $t_{\text{matching}} = n_{\text{phones}}/r_{\text{request}}$ seconds. On a match, the phone sends a channel request on the RACH and a paging response on an SDCCH in t_{response} seconds. It finally synchronizes back to the CCCH in t_{sync} seconds to be prepared for the next run. The required time for an attack is therefore $t_{\text{attack}} = t_{\text{matching}} + t_{\text{response}} + t_{\text{sync}}$ seconds. Thus, for a successful attack, the minimum number of phones an adversary requires is $n_{\text{phones}} \geq r_{\text{request}} \cdot t_{\text{attack}}$.

RACH Resource Constraints. The available resources provided by a single cell depend on its configuration. The GSM specifications defines a number of valid channel configurations [10]. Thus, an adversary is limited by the number of available RACH slots and the number of SDCCHs that a cell provides. In practice cells in metropolitan areas use the *BCCH+CCCH* or *FCCH+SCH+CCCH+BCCH* channel configurations on the first time slot. These are not combined with DCCHs and therefore allow all 51 bursts on the uplink of a 235.4 ms 51-multiframe to be used to transmit channel requests on the RACH. Because the RACH is a shared medium, collisions with requests of other subscribers may occur. According to Traynor et al. [94], the maximum resulting throughput is 37%. As a result, an attacker can transmit up to $r_{\text{RACH}} = 51/0.2354 \cdot 0.37 \approx 80$ channel requests per second in a single cell. Consequently, given that $n_{\text{phones}} \leq n_{\text{RACH}} = r_{\text{RACH}} \cdot t_{\text{attack}}$ is true, a single cell can fulfill the channel request requirements.

SDCCH Resource Constraints. Following the channel allocation, the adversary phone uses an SDCCH to send the paging response. Analogical to the RACH, SDCCHs in medium- or large-sized cells in a metropolitan area are provided on a separate time slot. A typical *SDCCH/8+SACCH/8* channel configuration comprises 8 SDCCHs per 51-multiframe, in theory offering: $r_{\text{SDCCH}} = 8/0.2354 = 34$ SDCCHs per second. Clearly this may make the signaling channel the major bottleneck for this attack. Accordingly, the occupation time of these channels needs to be taken into account. For example according to Traynor et al., a rough estimation of the occupation time of the channel for an Insert Call Forwarding operation is 2.7 seconds [94]. Compared to this, our attack occupies the channel for a very short duration, as shown in Section 5.2.5.

Table 5.2: Example resource requirements for E-Plus.

Variable	Value	Reference
$r_{request}$	415paging/min	Section 5.3.1
$t_{response}$	180ms	Section 5.2.5
t_{sync}	745ms	Section 5.3

Similar to the RACH requirements, the maximum number of attacking phones per cell is therefore bounded by $n_{\text{phones}} \leq n_{\text{SDCCH}} = r_{\text{SDCCH}} \cdot t_{\text{attack}}$.

Example Computation. The following is an example, based on the peak values from our measurements gathered for the E-Plus network and as reflected in Table 5.2. Based on the previous equations, at least $n_{\text{phones}} \approx 10.820$ phones are required to attack a typical location of E-Plus. Given the costs of the Motorola devices, this is a reasonably small amount. Each paging response attack lasts $t_{\text{attack}} \approx 1.564$ seconds. This allows up to $n_{\text{RACH}} \approx 125$ phones without a saturation of the RACH. For the SDCCH, the above formula yields to a maximum of $n_{\text{SDCCH}} \approx 53$ phones. It is also important to note that the number of phones is proportional to the impact. This means that half of the attacking phones would still be able to disrupt service for half of the subscribers of a location area.

A single cell therefore provides enough resources to attack a complete location area of a considerably small operator. In practice these resources are shared with legit MO and MT traffic. The exact traffic patterns and the number of cells per location is unknown. Furthermore, a combination with the IMSI detach attack prevents phones that reside in the location area to generate further MT activity. As we cannot estimate these activities, we do not include this in our calculation. Nevertheless, the results indicate the required resources for a large-scale attack do not extensively exhaust the resources provided by a cell. Additionally, there is no technical limitation of distributing attacking phones across a small number of different cells.

5.4 Countermeasures

In this section we present two countermeasures against the attacks we developed. Specifically, we propose different approaches to resolve both problems. A solution is required to not only fix the denial of service issue, but at the same time the MT service hijacking. Unlike the second prevention strategy, the first solution solves both issues at once, but requires a protocol change.

For the first solution, we propose a change to the paging protocol procedure [5]. To perform authentication, the network is sending a 128 bit random challenge (RAND) to the subscriber. Based on the secret key K_i that is only stored on the SIM card and in the authentication center of the network, the subscriber computes a 32 bit response value using the A3 algorithm. The so-called Signed Response (SRES) value is sent back to the network. In the same fashion, the operator network computes SRES based on K_i as stored in the authentication center. If both SRES values match, the subscriber successfully authenticated itself to the network. However, as mandated in the GSM specification, the authentication is performed after the paging response is processed. As also outlined in the previous chapter, the same principle applies to UMTS [28]. Therefore, the paging response itself is not authenticated.

By adapting the protocol to include the RAND value in the paging request and SRES in the paging response, this can be changed. This implies that all of the paging responses are authenticated, which eliminates session hijacking. At the same time a paging response that includes authentication information can be used by the network to validate the response before changing the state to not expect further paging responses. Thus, also solving the denial of service attack. It is important to note that this requires a fresh RAND for every authentication to prevent replay attacks. This is similar to the protocol change proposed in Section 4.4, which encrypts the paging request using a shared session key called *unlinkability key*. While they use this key to prevent tracking of subscribers via IMSI paging, the same modification also prevents our described attacks. Unfortunately, partly due to the difficulty of updating devices in the field, the industry is reluctant to apply new protocol changes to commercially deployed networks.

Moreover, it is noteworthy that this particular solution only addresses the paging protocol issues, but other parts of the protocol stack will be left potentially vulnerable to similar attacks. For example, the immediate assignment procedure could potentially be abused in a similar way. An attacker could observe immediate assignment messages, immediately tune to the allocated dedicated channel, and subsequently free the radio resource.

The second solution involves no protocol change, but has to dismantle each problem individually. The MT session hijacking issue can be addressed by enforcing authentication for each service request. This would also overcome MO session hijacking. In order to eliminate the denial of service attack, the MSC/VLR state machine needs to be changed. Specifically, the MSC/VLR has to be able to map all incoming paging responses to the correct service as long as no fully authenticated session exists. Accordingly, this circumvents the denial of service attack.

6 Conclusion and Future Work

The trust in the security of cellular networks and specifically the widely used GSM standard has been shattered several times. Yet, the interest in this technology is rising. Deployed 3G femtocells already outnumber traditional 3G base stations globally, and their deployment is increasing rapidly [59]. However, the security of these low-cost devices and the overall architecture seems poorly implemented in practice. Attacks against this type of mobile telecommunication technology are a recurring pattern, yet the industry does little to prevent active attacks by proactively including these types of scenarios in their threat model. Standardization bodies and the industry appear to be mostly operating on a reactive basis in terms of telecommunication security.

The undisturbed operation of telecommunication networks is traditionally based on trust. The inherent trust that each subscriber and participant in communication plays by the rules. Telecommunication network security is traditionally based on secrets, and the fact that it is hard for adversaries to tamper operation equipment. It has become evident in the past (e.g., due to external gateway providers, massive fraud problems, 3G to GSM downgrade attacks) that it is problematic to rely on this trust. Still, as shown in this thesis, the femtocell network architecture heavily relies on a single point of failure, the device itself. Nonetheless, due to several available and modifiable software and hardware projects for telecommunication, including femtocells itself, this trust relationship has to be considered broken.

In this thesis we looked at several aspects of mobile telecommunication security and specifically the involved protocols and their potential for abuse given an adversary in possession of modified radio equipment. This includes the mobile station part as well as access network equipment. We evaluated and demonstrated attacks originating from a rogue femtocell and their impact on end users and mobile operators. They are inherently trusted, able to monitor and modify all communication passing through them, and with an ability to contact other femtocells through the VPN tunnel. Yet when placed in untrustworthy hands, this assumption of trust proves dangerous. It is not only possible to intercept and modify mobile communication, but also to completely impersonate subscribers. Additionally, using the provided access to the operator network, we could leverage these attacks to a global scale, affect the network availability, and take control of a part of the femtocell infrastructure. Furthermore, we demonstrated that widely deployed protocols in 3G networks, which aim to prevent unauthorized parties from tracking individuals, are still not completely providing privacy. Moreover, we evaluated these shortcomings originating from the IMSI paging procedure and the 3G AKA protocol in practice using a modified femtocell device.

We continued to show how to exploit the aforementioned trust, by exploiting the paging procedure on a broadcast medium utilizing a mobile phone running a customized baseband stack. We demonstrated that it is possible to leverage a race condition in the paging protocol to a novel denial of service attack and the possibility to hijack mobile terminated services in GSM. Moreover, we showed that this attack can not only disturb communication for single subscribers, but can also greatly affect telephony in larger geographical regions formed by location areas. A motivated attacker can interrupt communication on a large scale by merely utilizing a set of older, inexpensive consumer devices that are still available on the market. This is considerably more efficient compared to traditional radio jamming due to the broad frequency range of mobile carrier networks and the size of location areas.

As our experimental results demonstrate, this has a considerable impact on mobile telecommunication. By modifying off-the-shelf equipment, we were able to mount denial of service attacks, hijack mobile-terminated transmissions, impersonate other subscribers, attack operator core infrastructure, modify subscriber traffic, and other attacks. In the case of the femtocell technology, we believe that attacks specifically targeting end users are a major problem and almost impossible to mitigate by operators due to the nature of the current architecture. The only solution towards attacks against end users would be to treat the femtocell as an untrusted device and rely on end-to-end encryption between the phone and the operator network. However, due to the nature of the 3G architecture and protocols as well as the large amount of required changes, it is probably not a practical solution.

As this thesis has shown, active attackers running modified hard- and software have to be actively considered by standardization bodies when designing future protocols beyond LTE. As the overall design of these technologies' comes with too many single points of failures, it is even more important to assure complete and secure authentication of all message exchanges as well as exploring options for end-to-end encryption for offloading technologies such as femtocells. In the latter case, the authors would furthermore like to question whether or not the practical advantages of femtocell technology outweigh their potential for critical attacks.

Addressing such problems in practice is difficult due to the large number of stakeholders in this ecosystem and complex relationships between them: handset manufacturers, network equipment manufacturers, carriers, baseband vendors, and standardization bodies. The complexity of such systems is furthermore driven by backwards-compatibility and inter-working requirements. Furthermore, as long as there is no revenue loss for carriers, there is very little incentive to deploy changes. Nonetheless, it is important to recognize that while telecommunication systems have to be considered as part of critical infrastructure for entire countries, security of these systems is not regulated at all. Security settings and protocol changes are left to the industry and carriers. It appears to be an interesting idea that regulation bodies would enforce a set of security requirements for mobile telecommunication networks as well.

Future Work. As the days of locked-down radio equipment and telecommunication stacks are over, especially in the era of software-defined radio, mobile network protocols have to explicitly protect against this particular threat model. However, with the development of more and more complex components for mobile networks, defending those will become more difficult. While we have shown in this thesis that modified equipment can be used to attack subscribers and protocol procedures, open radio equipment also leads the road to a completely different attack surface, the carrier network itself.

It is important to keep in mind that carrier networks are driven by software, which is exposed over the air interface to often anonymous adversaries. Additionally, this software has been artificially kept away from attackers due to the complexity of these protocols and lack of access to hard- and software that can be freely modified. This means that contrary to traditional software, e.g., network server software, it is safe to assume that these services have been kept away from attackers hammering on the software implementation to find exploitable bugs. Software security has many facets of which many are unsolved problems and even the strongest defense mechanisms that we know of have often been bypassed several times in practice.

Of course, the industry is also advancing in several areas. For example the Security Development Lifecycle (SDL) [84] developed by Microsoft has been adopted as the de-facto best practice in many industries related to computers. However, it appears to be a recurring pattern in information security that the strongest software in terms of security is often one that has been the target of attackers for years. With this in mind, the exploration and exploitation of carrier network infrastructure over-the-air appears to be an interesting area for future studies based on modified cellular equipment. We demonstrated the immense impact of maliciously controlled carrier infrastructure on the security of mobile telecommunication.

We continued to prove that an adversary with modified subscriber equipment can attack other subscribers and implicitly cause monetary damage to a carrier network. The natural combination of the two, the ability to modify end user devices to send crafted protocol payloads to exploit software flaws within the operator network and the abilities that an adversary with access to carrier infrastructure has, can be seen as the holy grail in wireless attacks on mobile telecommunication.

Acknowledgements

First, I would like to especially thank my supervisor, professor Dr. Jean-Pierre Seifert for his guidance, encouragement, overall support, plenty of useful discussions, and most importantly, giving me the freedom to work on research that I believe to be impactful and interesting. Without him and the general support of the department for Security in Telecommunication at the Berlin Institute of Technology, none of this would have been possible. I would also like to thank the Telekom Innovation Laboratories for providing financial support during this research.

Special thanks go to my colleagues Kévin Redon and Ravishankar Borgaonkar for the many hours we worked together on the research presented in this thesis and the fun we had around that. It was an honor working with them!

I furthermore would like to extend my sincere gratitude to Dmitry Nedospasov and Patrick Stewin for all the discussions we had, feedback they provided, the times they cheered me up, and the time we spent as the *"PhD team"*. I would probably still be sitting on a half finished thesis without their support.

I would also like to thank Collin Mulliner for his feedback on this thesis and his support and the lessons I learned from him, especially in my early times at SECT. Moreover, I'd like to thank Qualcomm, Alex Gantman, and Renwei Ge for their support on this thesis.

Furthermore, I would like to thank the Free Software movement around mobile telecommunication. Specifically, the Osmocom project, Harald Welte, Holger Freyther, Dieter Spaar, and Tobias Engel for the countless discussions, suggestions, and the immense amount of knowledge shared with me over the last years.

Finally, I'd like to thank my family for all the support over the years and especially my lovely wife Cori for her understanding, encouragement, and love!

Appendix

A Abbreviations and Acronyms

L3	3GPP Layer 3
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
AGCH	Access Grant Channel
AKA	Authentication and Key Agreement
AN	Access Network
AP	Access Point
AS	Autonomous System
ARFCN	Absolute Radio Frequency Channel Number
AuC	Authentication Center
AUTN	Authentication Token
BCH	Broadcast Channels
BS	Base Station
BSC	Base Station Controller
BSS	Base Station System or Subsystem
BTS	Base Transceiver Station

CCCH	Common Control Channel
CGI	Cell Global Identity
CS	Circuit Switched
CSG	Closed Subscriber Group
CI	Cell Identity
CK	Cipher Key
CM	Connection Management
CN	Core Network
DCCH	Dedicated Control Channel
DoS	Denial of Service
EAP-SIM	Extensible Authentication Protocol Method for GSM Subscriber Identity Modules
EAP-AKA	Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement
ESP	Encapsulating Security Payload
FAP	Femtocell Access Point
FCC	Federal Communications Commission
FDMA	Frequency Division Multiple Access
FTP	File Transfer Protocol
FMC	Fixed Mobile Convergence
GAN	Generic Access Network
GANC	GAN Controller
GA-CSR	Generic Access Circuit Switched Resource
GA-PSR	Generic Access Packet Switched Resource
GA-RC	Generic Access Resource Control
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPRS	General packet radio service
GSM	Global System for Mobile Communications
GSM-RR	GSM Radio Resource
HLR	Home Location Register

HMS	HNB Management System
HNB	Home Node B
HNB-GW	HNB GateWay
HNS	Home Node B Subsystem
HSS	Home Subscriber Server
IK	Integrity Key
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IP	Internet Protocol
ISP	Internet Service Provider
JTAG	Joint Test Action Group
LAC	Location Area Code
LAI	Location Area Identification
LAN	Local Area Network
LCS	Location Services
LTE	Long Term Evolution
MCC	Mobile Country Code
M2M	Machine to Machine
ME	Mobile Equipment
MitM	Man-in-the-Middle
MM	Mobility Management
MNC	Mobile Network Code
MNO	Mobile Network Operator
MO	Mobile Originated
MS	Mobile Station
MSC	Mobile-services Switching Center
MSISDN	Mobile Subscriber Integrated Services Digital Network Number
MT	Mobile Terminated
NAT	Network Address Translator
NB	Node B

OAM	Operation, Administration, and Maintenance
OAMP	Operation, Administration, Maintenance, and Provisioning
OS	Operation System
OTA	Over-The-Air
PCH	Paging Channel
PLMN	Public Land Mobile Network
PLMNO	Public Land Mobile Network Operator
PM	Performance Management
PS	Packet Switched
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RACH	Random Access Channel
RAND	Random Challenge
RES	Authentication Response
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RRLP	Radio Resource LCS Protocol
SeGW	Security GateWay
SFR	Société Française du Radiotéléphone
SGSN	Serving GPRS Support Node
SDCCH	Standalone Dedicated Control Channel
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SMS	Short Message Service
SOAP	Simple Object Access Protocol
SSH	Secure SHell
SQN	Sequence Number
TDMA	Time Division Multiple Access
TMSI	Temporary International Mobile Subscriber Identity
UART	Universal Asynchronous Receiver/Transmitter

UE	User Equipment
UMA	Unlicensed Mobile Access
UMTS	Universal Mobile Telecommunications System
USIM	Universal Subscriber Identity Module
UTRAN	Universal Terrestrial Radio Access Network
VLR	Visitor Location Register
VoIP	Voice over IP
VPN	Virtual Private Network
WLAN	Wireless LAN
Wi-Fi	Wireless Fidelity
XRES	Expected Response
ZAP	Zone Access Point

B

List of Figures

C

List of Tables

D

Bibliography

[1] OsmoSDR - rtl-sdr. http://sdr.osmocom.org/trac/wiki/rtl-sdr. [Accessed March 1st, 2014].

[2] Gartner Says Worldwide Smartphone Sales Soared in Fourth Quarter of 2011 With 47 Percent Growth. http://www.gartner.com/it/page.jsp?id=1924314, February 2012. [Accessed March 1st, 2014].

[3] 3GPP. Specification Numbering. http://www.3gpp.org/specifications/79-specification-numbering. [Accessed March 1st, 2014].

[4] 3GPP. Digital cellular telecommunications system (Phase 2+); Network architecture (GSM 03.02 version 7.1.0 Release 1998). Technical report, 3rd Generation Partnership Project, 2000. 3GPP TS 03.02 V7.1.0.

[5] 3GPP. Digital cellular telecommunications system (Phase 2+); Mobile radio interface layer 3 specification (3GPP TS 04.08 version 7.9.1 Release 1998). Technical report, 3rd Generation Partnership Project, 2001. 3GPP TS 04.08 V7.9.1.

[6] 3GPP. Digital cellular telecommunications system (Phase 2+); Multiplexing and multiple access on the radio path (3GPP TS 05.02 version 8.9.0 Release 1999). Technical report, 3rd Generation Partnership Project, 2001. 3GPP TS 05.02 V8.9.0.

[7] 3GPP. Universal Mobile Telecommunications System (UMTS); 3G Security; Integration Guidelines. Technical Specification TS 33.103 v4.2.0, 3rd Generation Partnership Project, September 2001.

[8] 3GPP. Universal Mobile Telecommunications System (UMTS); LTE; 3G Security; Security Principles and Objectives. Technical Specification TS 33.120 v4.0.0, 3rd Generation Partnership Project, March 2001.

[9] 3GPP. Digital cellular telecommunications system (Phase 2+); Base Station System - Mobile Services Switching Centre (BSS-MSC) Interface - Interface Principles (3GPP TS 08.02 version 8.0.1 Release 1999). Technical report, 3rd Generation Partnership Project, 2002. 3GPP TS 08.02 V8.0.1.

[10] 3GPP. Digital cellular telecommunications system (Phase 2+); Mobile Station - Base Station System (MS - BSS) Interface Channel Structures and Access Capabilities (3GPP TS 04.03 version 8.0.2 Release 1999). Technical report, 3rd Generation Partnership Project, 2002. 3GPP TS 04.03 V8.0.2.

[11] 3GPP. Digital cellular telecommunications system (Phase 2+); Numbering, addressing and identification (3GPP TS 03.03 version 7.8.0 Release 1998). Technical report, 3rd Generation Partnership Project, 2003. 3GPP TS 03.03 V7.8.0.

[12] 3GPP. Service requirements for Home Node B (HNB) and Home eNode B (HeNB). Technical Specification TS 22.220 v11.2.0, 3rd Generation Partnership Project, June 2005.

[13] 3GPP. Digital cellular telecommunications system (Phase 2+); Security-related network functions (3GPP TS 03.20 version 8.6.0 Release 1999). Technical report, 3rd Generation Partnership Project, 2008. 3GPP TS 03.20 V8.6.0.

[14] 3GPP. Security of H(e)NB. Technical Report TR 33.820 v8.3.0, 3rd Generation Partnership Project, December 2009.

[15] 3GPP. Digital cellular telecommunications system (Phase 2+); Radio transmission and reception (3GPP TS 45.005 version 9.1.0 Release 9). Technical report, 3rd Generation Partnership Project, 2010. 3GPP TS 45.005 V9.1.0.

[16] 3GPP. Generic Access Network (GAN); Mobile GAN interface layer 3 specification. Technical Specification TS 44.318 v9.2.0, 3rd Generation Partnership Project, 2010.

[17] 3GPP. Generic Access Network (GAN); Stage 2. Technical Specification TS 43.318 v9.0.0, 3rd Generation Partnership Project, February 2010.

[18] 3GPP. Technical specification group services and system aspects; 3G security; security architecture (release 9). Technical report, 3rd Generation Partnership Project, 2010. 3GPP TS 33.102 V9.3.0.

[19] 3GPP. Telecommunication management; Performance Management (PM); Concept and requirements. Technical Specification TS 32.401 v9.1.0, 3rd Generation Partnership Project, October 2010.

[20] 3GPP. Telecommunications management; Home Node B (HNB) Operations, Administration, Maintenance and Provisioning (OAM&P); Concepts and requirements for Type 1 interface HNB to HNB Management System (HMS). Technical Specification TS 32.581 v9.2.0, 3rd Generation Partnership Project, April 2010.

[21] 3GPP. Universal Mobile Telecommunications System (UMTS); LTE; Network architecture. Technical Specification TS 23.002 v9.2.0, 3rd Generation Partnership Project, January 2010.

[22] 3GPP. Universal Mobile Telecommunications System (UMTS);Physical channels and mapping of transport channels onto physical channels (FDD)(3GPP TS 25.211 version 9.2.0 Release 9). Technical report, 3rd Generation Partnership Project, 2010. 3GPP TS 25.211 9.2.0.

[23] 3GPP. Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions. Technical Specification TS 29.090 v10.0.0, 3rd Generation Partnership Project, April 2011.

[24] 3GPP. Mobile Application Part (MAP) specification. Technical Specification TS 29.002 v10.3, 3rd Generation Partnership Project, January 2011.

[25] 3GPP. Security of Home Node B (HNB) / Home evolved Node B (HeNB). Technical Specification TS 33.320 v11.2.0, 3rd Generation Partnership Project, 2011.

[26] 3GPP. Technical specification group services and system aspects; 3G security; cryptographic algorithm requirements (release 10). Technical Report TS 33.105 V10.0.0, 3rd Generation Partnership Project, 2011.

[27] 3GPP. Universal Mobile Telecommunications System (UMTS); LTE; 3G security; Security architecture. Technical Specification TS 33.102 v9.4.0, 3rd Generation Partnership Project, January 2011.

[28] 3GPP. Universal Mobile Telecommunications System (UMTS); LTE;3G security; Security architecture(3GPP TS 33.102 version 9.4.0 Release 9). Technical report, 3rd Generation Partnership Project, 2011. 3GPP TS 33.102 V9.4.0.

[29] 3GPP. UTRAN architecture for 3G Home Node B (HNB); Stage 2. Technical Specification TS 25.467 v10.2.0, 3rd Generation Partnership Project, June 2011.

[30] 3GPP. Digital cellular telecommunications system (Phase 2+); Mobile Switching Centre - Base Station system (MSC-BSS) interface; Layer 3 specification (3GPP TS 48.008 version 9.8.0 Release 9). Technical report, 3rd Generation Partnership Project, 2012. 3GPP TS 48.008 V9.8.0.

[31] 3GPP. LTE;Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode(3GPP TS 36.304 version 9.9.0 Release 9). Technical report, 3rd Generation Partnership Project, 2012. 3GPP TS 36.304 V9.9.0.

[32] 3GPP. Universal Mobile Telecommunications System (UMTS);User Equipment (UE) procedures in idle mode and procedures for cell reselection in connected mode(3GPP TS 25.304 version 9.8.0 Release 9). Technical report, 3rd Generation Partnership Project, 2012. 3GPP TS 25.304 V9.8.0.

[33] 3GPP. Technical Specification Group Services and System Aspects; 3G security; Lawful Interception requirements (3GPP TS 33.106 version 12.1.0 Release 12). Technical report, 3rd Generation Partnership Project, 2013. 3GPP TS 33.106 V12.1.0.

[34] ABI Research. One Billion Mobile Broadband Subscriptions in 2011: a Rosy Picture Ahead for Mobile Network Operators, February 2011.

[35] Zahra Ahmadian, Somayeh Salimi, and Ahmad Salahi. New attacks on UMTS network access. In *Conference on Wireless Telecommunications Symposium*, WTS'09, 2009.

[36] Richard Allen and Doug Kelly. Gaining root on Samsung Femtocells. http://rsaxvc.net/blog/2011/7/17/GainingrootonSamsungFemtoCells.html, July 2011. [Accessed March 1st, 2014].

[37] Myrto Arapinis, Loretta Mancini, Eike Ritter, and Mark Ryan. Privacy through Pseudonymity in Mobile Telephony Systems. In *Proceedings of the 21st Network and Distributed System Security (NDSS) Symposium*, February 2014.

[38] Myrto Arapinis, Loretta Mancini, Eike Ritter, Mark Ryan, Nico Golde, Kevin Redon, and Ravishankar Borgaonkar. New Privacy Issues in Mobile Telephony: Fix and Verification. In *Proceedings of the 19th ACM Conference on Computer and Communications Security*, October 2012.

[39] Elad Barkan, Eli Biham, and Nathan Keller. Instant Ciphertext-Only Cryptanalysis of GSM Encrypted Communication. *J. Cryptol.*, 21(3):392–429, March 2008.

[40] Alex Biryukov, Adi Shamir, and David Wagner. Real Time Cryptanalysis of A5/1 on a PC. In *Proceedings of the 7th International Workshop on Fast Software Encryption*, FSE '00, pages 1–18, London, UK, UK, 2001. Springer-Verlag.

[41] Ravishankar Borgaonkar, Kévin Redon, and Jean-Pierre Seifert. Security Analysis of a Femtocell device. In *Proceedings of the 4th International Conference on Security of Information and Networks*, SINCONF. ACM, November 2011.

[42] CDMA Development Group. CDMA History. http://www.cdg.org/resources/cdma_history.asp. [Accessed March 1st, 2014].

[43] Cisco. Cisco Visual Networking Index: Forecast and Methodology, 2010-2015, February 2011.

[44] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010-2015. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html, January 2011. [Accessed March 1st, 2014].

[45] Commerzbank. Safe Online Banking. https://www.commerzbank.de/portal/en/englisch/products-offers/services/secure-internet-banking/banking.html. [Accessed March 1st, 2014].

[46] Ang Cui, Yingbo Song, Pratap V. Prabhu, and Salvatore J. Stolfo. Brave New World: Pervasive Insecurity of Embedded Network Devices. In *12th Annual International Symposium on Advances in Intrusion Detection*, RAID '09, pages 378–380, Berlin, Heidelberg, September 2009. Springer-Verlag.

[47] CVE Details. Apache Security Vulnerabilies. http://www.cvedetails.com/vulnerability-list/vendor_id-45/Apache.html. [Accessed March 1st, 2014].

[48] CVE Details. Libxml2 Security Vulnerabilies. http://www.cvedetails.com/vulnerability-list/vendor_id-1962/product_id-3311/Xmlsoft-Libxml2.html. [Accessed March 1st, 2014].

[49] D. Burgess et al. OpenBTS. http://openbts.org. [Accessed March 1st, 2014].

[50] Michele Dallachiesa. rtpbreak. http://dallachiesa.com/code/rtpbreak/. [Accessed March 1st, 2014].

[51] Guillaume de la Roche, Alvaro Valcarce, David López-Pérez, and Jie Zhang. Access control mechanisms for femtocells. *Communications Magazine, IEEE*, 48:33–39, January 2010.

[52] Nicholas Hopper Denis Foo Kune, John Koelndorfer and Yongdae Kim. Location leaks over the gsm air interface. In *Annual Network & Distributed System Security Symposium*, NDSS, 2012.

[53] William Enck, Patrick Traynor, Patrick McDaniel, and Thomas La Porta. Exploiting open functionality in SMS-capable cellular networks. In *Proceedings of the 12th ACM conference on Computer and communications security*, CCS '05, pages 393–404, New York, NY, USA, 2005. ACM.

[54] William Enck, Patrick Traynor, Patrick McDaniel, and Thomas La Porta. Exploiting Open Functionality in SMS-Capable Cellular Networks. In *Proceedings of the 12th ACM conference on Computer and communications security*, CCS '05, pages 393–404, New York, NY, USA, 2005. ACM.

[55] Ettus. USRP. http://www.ettus.com/products, 2009. [Accessed March 1st, 2014].

[56] Evrytania LLC. LTE Cell Scanner. http://www.evrytania.com/lte-tools/lte-cell-scanner. [Accessed March 1st, 2014].

[57] fail0verflow. AT&T Microcell FAIL. https://fail0verflow.com/blog/2012/microcell-fail.html, March 2012. [Accessed March 1st, 2014].

[58] Zack Fasel and Matt Jakubowski. Infrastructure Weaknesses in Distributed Wireless Communication Services. http://www.shmoocon.org/, February 2010. [Accessed March 1st, 2014].

[59] Femto Forum. Femtocell Market Status - December 2012. http://www.smallcellforum.org/resources-reports, December 2012. [Accessed March 1st, 2014].

[60] Femto Forum. Femtocell Market Status - February 2013. http://www.smallcellforum.org/newsstory-public-access-small-cell-market-to-hit-us-16-billion-in-2016, February 2013. [Accessed March 1st, 2014].

[61] Dirk Fox. IMSI-Catcher. *Datenschutz und Datensicherheit (DuD)*, 21:539–539, 1997.

[62] Frank A. Stevenson. [A51] The call of Kraken. http://web.archive.org/web/20100812204319/http://lists.lists.reflextor.com/pipermail/a51/2010-July/000683.html, July 2010. [Accessed March 1st, 2014].

[63] Nico Golde, Kévin Redon, and Ravishankar Borgaonkar. Weaponizing Femtocells: The Effect of Rogue Devices on Mobile Telecommunications. In *Proceedings of the 19th Annual Network & Distributed System Security Symposium*, February 2012.

[64] Nico Golde, Kévin Redon, and Jean-Pierre Seifert. Let me answer that for you: Exploiting broadcast information in cellular networks. In *Proceedings of the 22nd USENIX Security Symposium*, Washington, D.C., USA, August 2013.

[65] Chan-Kyu Han, Hyoung-Kee Choi, and In-Hwan Kim. Building Femtocell More Secure with Improved Proxy Signature. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6, Honolulu, HI, November 2009.

[66] Henry Haverinen and Joseph Salowey. Extensible Authentication Protocol Method for Global System for Mobile Communications (GSM) Subscriber Identity Modules (EAP-SIM). RFC 4186, 3rd Generation Partnership Project, January 2006.

[67] HP Security. Mobile Pwn2own 2013. http://h30499.www3.hp.com/t5/HP-Security-Research-Blog/Mobile-Pwn2Own-2013/ba-p/6202185, October 2013. [Accessed March 1st, 2014].

[68] Infosecurity Magazine. Indian company hacks GSM and usurps IMSI. http://www.infosecurity-magazine.com/view/24680/indian-company-hacks-gsm-and-usurps-imsi/, March 2012. [Accessed March 1st, 2014].

[69] ip.access Ltd. nanoBTS 1800. http://www.ipaccess.com/en/nanoGSM-picocell. [Accessed March 1st, 2014].

[70] International Telecommunication Union (ITU). The World in 2013 - ICT Facts and Figures. http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf, February 2013. [Accessed March 1st, 2014].

[71] Nick Jacobsen. Samsung-Femtocell. http://code.google.com/p/samsung-femtocell/, March 2011. [Accessed March 1st, 2014].

[72] Karsten Nohl and Luca Melette. Defending mobile phones. http://events.ccc.de/congress/2011/Fahrplan/events/4736.en.html, December 2011. [Accessed March 1st, 2014].

[73] Kineto Wireless Inc. UMA Today - UMA Jumpstarts FEMTOCELL Market. http://kineto.com/inter_uma_mag_femtocell/document.pdf, 2007. [Accessed March 1st, 2014].

[74] Michael Krell. Crowdflow. http://crowdflow.net. [Accessed March 1st, 2014].

[75] Denis F. Kune, John Koelndorfer, Nicholas Hopper, and Yongdae Kim. Location leaks over the GSM air interface. In *Proceedings of the 19th Annual Network & Distributed System Security Symposium*, February 2012.

[76] Thomas Landspurg. OpenCellID. http://opencellid.org. [Accessed March 1st, 2014].

[77] Patrick P. C. Lee, Tian Bu, and Thomas Woo. On the Detection of Signaling DoS Attacks on 3G Wireless Networks. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 1289–1297, Anchorage, AK, May 2007.

[78] Patrick P. C. Lee, Tian Bu, and Thomas Woo. On the detection of signaling DoS attacks on 3G/WiMax wireless networks. *Comput. Netw.*, 53(15):2601–2616, 2009.

[79] M. Leech, M. Ganis, Y. Lee, R. Kuris, D. Koblas, and L. Jones. SOCKS Protocol Version 5. RFC 1928, March 1996. [Accessed March 1st, 2014].

[80] Path Intelligence Ltd. Path intelligence - retail measurement technology. http://www.pathintelligence.com. [Accessed March 1st, 2014].

[81] Robindhra Mangtani. Security Issues in Femtocell Deployment. Technical Report 1.0, GSM Association, July 2008.

[82] Ulrike Meyer and Susanne Wetzel. A Man-in-the-Middle Attack on UMTS. In *Workshop on Wireless Security*, pages 90–97, 2004.

[83] Ulrike Meyer and Susanne Wetzel. A man-in-the-middle attack on UMTS. In *ACM Workshop on Wireless Security*, WiSe, 2004.

[84] Microsoft. Security Development Lifecycle. http://www.microsoft.com/security/sdl/default.aspx, 2008. [Accessed March 1st, 2014].

[85] David L. Mills. Network Time Protocol (Version 3) - Specification, Implementation and Analysis. RFC 1305, March 1992. [Accessed March 1st, 2014].

[86] Mohammed Shafiul Alam Khan and Chris J Mitchell. Another Look at Privacy Threats in 3G Mobile Telephony . https://pure.royalholloway.ac.uk/portal/files/19681499/alapti.pdf. [Accessed September 1st, 2014].

[87] Collin Mulliner, Nico Golde, and Jean-Pierre Seifert. SMS of Death: From Analyzing to Attacking Mobile Phones on a Large Scale. In *Proceedings of the 20th USENIX Security Symposium*, San Francisco, CA, USA, August 2011.

[88] Collin Mulliner and Charlie Miller. Injecting SMS Messages into Smart Phones for Security Analysis. In *Proceedings of the 3rd USENIX Workshop on Offensive Technologies (WOOT)*, Montreal, Canada, August 2009.

[89] Sylvain Munaut. IMSI DETACH DoS. http://security.osmocom.org/trac/ticket/2, May 2010. [Accessed March 1st, 2014].

[90] NBC NEWS. 'StingRay': Records Show Secret Cellphone Surveillance by Calif. Cops. http://www.nbcnews.com/tech/security/stingray-records-show-secret-cellphone-surveillance-calif-cops-n52181, 2014. [Accessed March 14th, 2014].

[91] Karsten Nohl and Sylvain Munaut. Wideband gsm sniffing. http://events.ccc.de/congress/2010/Fahrplan/attachments/1783_101228.27C3.GSM-Sniffing.Nohl_Munaut.pdf. [Accessed March 1st, 2014].

[92] Nuand. nuand bladeRF Software Defined Radio. http://nuand.com/. [Accessed March 1st, 2014].

[93] Michael Ossmann. HackRF - low cost software radio platform. http://greatscottgadgets.com/hackrf/. [Accessed March 1st, 2014].

[94] P. Traynor, M. Lin, M. Ongtang, V. Rao, T. Jaeger, T. La Porta, P. McDaniel. On Cellular Botnets: Measuring the Impact of Malicious Devices on a Cellular Network Core. In *ACM Conference on Computer and Communications Security (CCS)*, November 2009.

[95] Chris Paget. Practical Cellphone Spying. https://www.defcon.org/html/defcon-18/dc-18-speakers.html#Paget, August 2010. [Accessed March 1st, 2014].

[96] OsmocomBB Project. Motorola Filter Replacement. https://bb.osmocom.org/trac/wiki/Hardware/FilterReplacement. [Accessed March 1st, 2014].

[97] PurpleLabs. Tsm30 firmware. http://web.archive.org/web/20090325133430/http://sourceforge.net/projects/plabs, November 2004. [Accessed March 1st, 2014].

[98] Radmilo Racic, Denys Ma, and Hao Chen. Exploiting MMS Vulnerabilities to Stealthily Exhaust Mobile Phone's Battery. In *Securecomm and Workshops, 2006*, pages 1 –10, 28 2006-sept. 1 2006.

[99] R Rajavelsamy, Jicheol Lee, and Sungho Choi. Towards security architecture for Home (evolved) NodeB: challenges, requirements, and solutions. *Security and Communication Networks*, 4(4):471–481, April 2011.

[100] Routo Messaging / TeleSign Mobile. Bulk SMS Gateway. http://www.routomessaging.com. [Accessed March 1st, 2014].

[101] Mark Ryan. Proverif code for formal verification of umts privacy threats. http://www.markryan.eu/research/UMTS/. [Accessed March 1st, 2014].

[102] Security Research Labs. A5/1 decryption project. http://opensource.srlabs.de/projects/a51-decrypt. [Accessed March 1st, 2014].

[103] Security Research Labs. Decrypting GSM phone calls. https://srlabs.de/decrypting_gsm/. [Accessed March 1st, 2014].

[104] Security Research Labs. GSM security map. http://www.gsmmap.org. [Accessed March 1st, 2014].

[105] Jérémy Serror, Hui Zang, and Jean C. Bolot. Impact of paging channel overloads or attacks on a cellular network. In *Proceedings of the 5th ACM workshop on Wireless security*, WiSe '06, pages 75–84, New York, NY, USA, 2006. ACM.

[106] SFR. SFR Home 3G : pour une couverture 3G optimale à domicile. http://www.sfr.fr/vos-services/equipements/innovations/sfr-home-3g/. [Accessed March 1st, 2014].

[107] SFR press release. SFR lance le service Femtocell, SFR Home 3G, pour offrir la meilleure couverture 3G au domicile de ses clients. http://www.sfr.com/presse/communiques-de-presse/sfr-lance-le-service-femtocell-sfr-home-3g-pour-offrir-la-meilleure, November 2009. [Accessed March 1st, 2014].

[108] Dieter Spaar. RACH flood DoS. http://security.osmocom.org/trac/ticket/1, November 2009. [Accessed March 1st, 2014].

[109] Andreas Steffen, Martin Willi, and Tobias Brunner. strongSwan IPsec solution. http://www.strongswan.org/, June 2011. [Accessed March 1st, 2014].

[110] Martin Storsjo. OpenCORE. http://sourceforge.net/projects/opencore-amr/. [Accessed March 1st, 2014].

[111] Daehyun Strobel. IMSI Catcher, 2007. Seminar Work, Ruhr-Universitat Bochum.

[112] Sylvain Munaut. Further Hacks on the Calypso Platform. http://events.ccc.de/congress/2012/Fahrplan/events/5226.en.html, December 2012. [Accessed March 1st, 2014].

[113] T. Engel. Remote SMS/MMS Denial of Service - Curse Of Silence. http://berlin.ccc.de/~tobias/cursesms.txt, December 2008. [Accessed March 1st, 2014].

[114] The Broadband Forum TR-069. CPE WAN Management Protocol. http://www.broadband-forum.org/technical/download/TR-069_Amendment-3.pdf, November 2010. [Accessed March 1st, 2014].

[115] The Hacker's Choice. Vodafone Access Gateway. http://wiki.thc.org/vodafone, June 2011. [Accessed March 1st, 2014].

[116] Patrick Traynor, Michael Lin, Machigar Ongtang, Vikhyath Rao, Trent Jaeger, Patrick Mcdaniel, and Thomas La Porta. On Cellular Botnets: Measuring the Impact of Malicious Devices on a Cellular Network Core. In *Computer and Communications Security*, pages 223–234, 2009.

[117] UMTS World - News and Information about 3G mobile networks. UMTS / 3G History and Future Milestones. http://www.umtsworld.com/umts/history.htm. [Accessed March 1st, 2014].

[118] Various contributors. Osmocom project. http://osmocom.org. [Accessed March 1st, 2014].

[119] Dennis Wehrle. Open Source IMSI-Catcher für GSM. https://www.ks.uni-freiburg.de/lehre/abschlussarbeiten/master-arbeiten/opensource-imsi-catcher-fuer-gsm, October 2009. [Accessed March 1st, 2014].

[120] Ralf-Philipp Weinmann. Baseband Attacks: Remote Exploitation of Memory Corruptions in Cellular Protocol Stacks. In *Proceedings of the 21st USENIX Workshop on Offensive Technologies*, Bellevue, WA, USA, August 2012.

[121] Harald Welte. Anatomy of contemporary GSM cellphone hardware. http://laforge.gnumonks.org/papers/gsm_phone-anatomy-latest.pdf, April 2010. [Accessed March 1st, 2014].

[122] Harald Welte, Holger Freyther, Dieter Spaar, Stefan Schmidt, Daniel Willmann, Jan Luebbe, Thomas Seiler, and Andreas Eversberg. OpenBSC. http://openbsc.osmocom.org. [Accessed March 1st, 2014].

[123] Harald Welte, Sylvain Munaut, Andreas Eversberg, and other contributors. OsmocomBB. http://bb.osmocom.org. [Accessed March 1st, 2014].

[124] ZDNET. Exploit beamed via NFC to hack Samsung Galaxy S3 (Android 4.0.4). http://www.zdnet.com/exploit-beamed-via-nfc-to-hack-samsung-galaxy-s3-android-4-0-4-7000004510/, September 2012. [Accessed March 1st, 2014].

[125] Jie Zhang and Guillaume de la Roche. *Femtocells: Technologies and Deployment*. John Wiley & Sons, Ltd, March 2010.

[126] Muxiang Zhang and Yuguang Fang. Security analysis and enhancements of 3GPP authentication and key agreement protocol. *IEEE Transactions on Wireless Communications*, 4(2):734–742, 2005.