



Forensic Identification of GSM Mobile Phones

Jakob Hasse

dence GmbH

c/o Technische Universität Dresden

jakob.hasse@dence.de

Thomas Gloe

dence GmbH

c/o Technische Universität Dresden

thomas.gloe@dence.de

Martin Beck

Institute of Systems Architecture

Technische Universität Dresden

martin.beck1@tu-dresden.de

ABSTRACT

With the rapid growth of GSM telecommunication, special requirements arise in digital forensics to identify mobile phones operating in a GSM network. This paper introduces a novel method to identify GSM devices based on physical characteristics of the radio frequency hardware. An implementation of a specialised receiver software allows passive monitoring of GSM traffic along with physical layer burst extraction even for handover and frequency hopping techniques. We introduce time-based patterns of modulation errors as a unique device-dependent feature and carefully remove random effects of the wireless communication channel. Using our characteristics, we could distinguish 13 mobile phones at an overall success rate of 97.62% under real-world conditions. This work proves practical feasibility of physical layer identification scenarios capable of tracking or authenticating GSM-based devices.

Categories and Subject Descriptors

K.4.2 [Social Issues]: Abuse and Crime Involving Computers; C.2.0 [General]: Security and Protection

Keywords

mobile phone identification; digital forensics; GSM; radio fingerprinting

1. INTRODUCTION

The currently most used mobile telecommunication system GSM lacks reliable mechanisms to identify end user mobile devices. Identification using the device identification number IMEI is considered insecure. Available hardware flashers allow to change and manipulate a mobile phone's software including the IMEI number. In consequence, law enforcement agencies focus on monitoring the SIM identifier IMSI, which can be changed easily by switching SIM cards.

Another method to identify GSM devices could evaluate characteristics of the transmitted wireless signal. GSM

networks use radio transmissions as communication channels, which naturally represent a shared medium. The radio transmissions can therefore be captured passively by third party receivers located within the communication range of the sending device and do not rely on the cooperation of the sender. Inaccuracies in the manufacturing process and allowed tolerances of the radio hardware are likely to introduce identifying traces in the signal. Recent studies document possibilities to identify IEEE 802.11 devices using characteristics of the transmitted signals. Brik et al. [2] measures errors in the modulation domain to generate a fingerprint of a IEEE 802.11 based device. B. Rasmussen et al. [1] use 5 transient features of the signal amplitude to identify short range devices. Focusing on the same class of devices, D. Zanetti et al. [9] propose features of the link frequency, which is allowed to deviate between 4% and 22% according to the specification of the UHF Class 1 communication protocol. These works demonstrate the feasibility of RF device identification on the physical layer but are less important for selecting physical features for GSM. First investigations targeting GSM device identification are reported by Reising et al. [8]. Evaluating the instantaneous frequency and phase responses of the raw RF signal in transient and midamble regions, notable identification performance was achieved for a small set of 3 devices using a fixed location during training and test.

Threshold based burst extraction as performed by Reising et al. is only applicable under laboratory conditions, but not in realistic scenarios. Our work aims to evaluate identification performance in a practical environment. Because of the complex GSM radio access scheme, burst extraction is a complicated task. To extract the communication stream of an individual phone, we interpret the captured radio signal according to the GSM standard. Instead of using coarse statistical moments of the raw captured signal, we analyse the signal in the well defined domain of the GMSK modulation, similarly to Brik et al. Based on the interpreted radio signal, we are able to derive signal features independent from the transmitted data, allowing to evaluate entire bursts instead of using only midamble and transient regions. Because GSM is a high precision communication system allowing only marginal inaccuracies of the RF hardware, modulation accuracy features used for IEEE 802.11 are not directly applicable. We propose to evaluate the modulation errors in respect to the time of one burst transmission, targeting both time dependent and constant inaccuracies of the radio hardware. Using signal processing, we carefully diminish random effects introduced by different locations or

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IH&MMSec'13, June 17–19, 2013, Montpellier, France.
Copyright 2013 ACM 978-1-4503-2081-8/13/06 ...\$15.00.

variable radio parameters. Reising et al. employed a fixed location to capture training and test data, while we prove our features are location independent. In summary our key contributions are:

- the first practical identification system of GSM devices based on radio frequency fingerprints,
- a software receiver for full extraction of bursts allocated to individual mobile devices,
- proposal of time based device characteristics based on modulation accuracy, and
- a comprehensive evaluation of our proposed methods and comparison to state-of-the art mobile device identification measures.

Our identification system allows law enforcement agencies to identify mobile phones under surveillance without the cooperation of the network. Further, these methods might be integrated in GSM base stations to improve authentication procedures and to find phones reported to be stolen. Equally, a mobile phone can identify a base station to authenticate the communication partner on the network part. When integrated in this way, implementation difficulties relating to encryption and burst extraction will become easier or obsolete.

The remainder of the paper is structured as follows. In Section 2, we introduce the most important technical details of a GSM network detailing the physical layer air interface. In the following Section 3, we describe the procedure of recording physical radio signals and common problems relating to GSM features. The signal processing applied on each burst is described in Section 4 along with an introduction of the proposed signal features. Section 5 describes the test setup for our experiments performed in Section 6. We conclude in Section 7.

2. GSM FUNDAMENTALS

A GSM network provides land based mobile communication. In contrary to satellite based systems, the communication counterpart is a base station on the ground at a fixed position. The network is run by an operator and can be connected to other networks like public switched networks or other mobile networks using gateways. Figure 1 shows an overview of the entities in a GSM network. The mobile phone communicates with the base station over the air using the Um protocol. The base station provides the RF link in a fixed geographical location, known as the cell. One or more base stations are managed by a base station controller, taking care of radio channels and some features like handovers. The mobile switching center acts as a switch between all components of a fixed network, connecting to multiple base station controllers, other switching centers, databases or different phone networks.

For a seamless operation, databases like the visitor or home location register store data of subscribers currently using the network. This includes authentication data and identifiers like the IMSI or IMEI. While the IMSI identifies a SIM module with the corresponding contract between a subscriber and the network operator, the IMEI identifies the mobile phone or a similar device which is able to gain access to a GSM network. The handling of the IMEI depends on the network. Many networks query the IMEI on

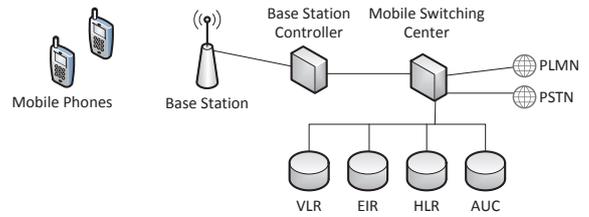


Figure 1: GSM Network Architecture

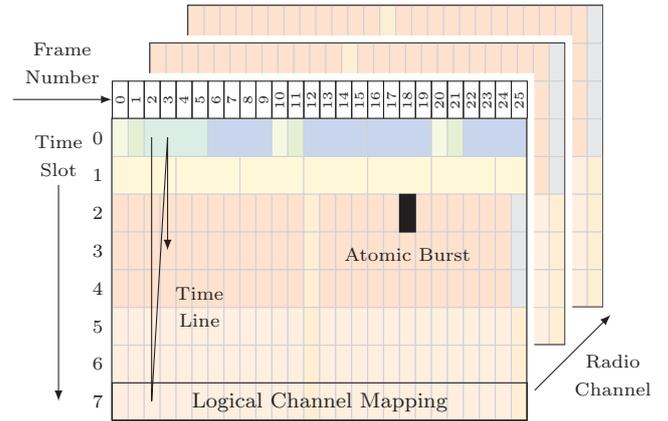


Figure 2: GSM Multiplexed Access Scheme

every connection of a mobile phone to perform a lookup in the equipment identity register, which holds the IMEIs of stolen, malfunctioning phones or IMEIs which are known to be used by criminals. The storage and transmission of the IMEI is handled by the software of the mobile phone and can be altered with a simple software patch.

When making a call, a mobile phone connects to the most powerful base station in range to establish a communication channel to the network. After authentication at the authentication center, the current location will be updated and stored in the home or visitor location register. A temporary replacement for the IMSI may be assigned to prohibit the plain transmission of the IMSI on every connection attempt. Finally, the requested call service is performed by establishing a transmission channel through multiple switching centres to the target communication partner.

To identify a mobile phone, we focus on the air interface (Um) between the mobile phone and the base station. The protocol is divided in three layers: physical, link and networking/signalling layer. For our identification method, the understanding of the physical layer including the wireless transmission channel is most important. Because of the complex access scheme for radio resources, it is a challenge to capture RF signals of an individual GSM mobile phone. On the most basic level, RF resources of a base station are multiplexed using frequency and time (see Figure 2). The frequency bands allowed to be used by the government are split into radio channels with a bandwidth of 200 kHz each.

The base station decides how many radio channels to use and whether to switch between radio channels over time (frequency hopping) or not. When employing frequency hopping, a communication stream can be spread over all radio channels the base station supports. The state of the time multiplex is defined by the frame number and time slot. Each frame consists of 8 time slots, each time slot lasts for approximately $577 \mu\text{s}$. The frame number is incremented continuously until reset after 3.5 h. The combination of radio channel, frame number and time slot defines a dedicated physical radio resource during which a mobile phone sends one radio burst. Based on the physical multiplex, logical channels are defined by fixed mappings to create communication subchannels for specific uses.

As a consequence of the access system, the mobile phone almost never transmits two bursts consecutively. Each burst is transmitted independently, so the procedure of sending a ‘normal burst’ (see Figure 3) is always the same for the RF hardware. At the start of a burst, the sending power ramps up to the desired signal strength. The tail bits at the beginning and end of a burst are set to zero. The data payload is split into two data segments of 57 Bit each, with a training sequence in between. The training sequence is one of the available bit sequences defined in the GSM specification [3]. The base station decides which training sequence to use for the radio transmission. The two additional stealing flags have a special meaning in some logical channels and are ignored in others. After the last tail bits, the RF hardware powers down. The burst transmission of the next time slot starts at the end of the guard period, which lasts for 8.25 Bits at GSM symbol rate. The described normal burst is modulated with the Gaussian Minimum Shift Keying (GMSK) algorithm. There are specialised burst versions for specific applications which will not be analysed in detail in this work. The observed normal burst transmission performed by the mobile phone is the data source for the identification procedure described in Section 4.

When a mobile phone communicates over the air, it often changes logical channels. To extract the burst transmissions performed by an individual phone, a receiver must be able to interpret the higher protocol layers in order to follow the communication stream. Figure 4 shows the protocol flow when making a call, different logical channels are represented by different arrow colors. A mobile requests radio resources on the random access channel (RACH). The network assigns dedicated resources using the common control channel (CCCH) and switches the subsequent communication to a standalone dedicated control channel (SDCCH). After requesting call related services, encryption is activated. The base station may require to send the IMEI before the call is initiated. For the upcoming speech transmission, a traffic channel (TCH) is established. Now all signalling packets are sent over the fast associated control channel (FACCH). The mobile ends the call with the disconnect message and the communication channel will be released.

3. RECORDING GSM SIGNALS

The first step towards identification of mobile phones over the air is to capture the radio bursts emitted by one or more targeted mobile phones operating in a GSM network. There are two general possibilities to access these signals. The first one is to perform the identification algorithm at one communication endpoint, i.e. the base station analyses the signals

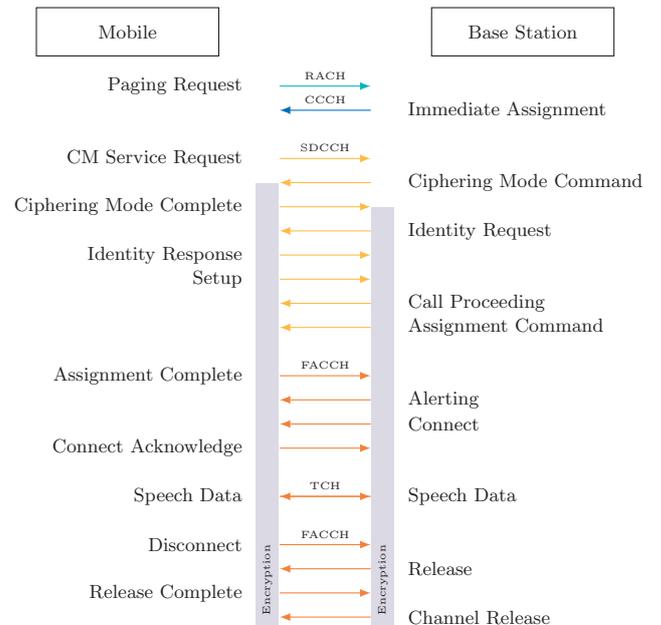


Figure 4: Protocol Flow of Making a Call

emitted by the mobile phone or vice versa. The identification algorithm can be implemented as an extension to the existing GSM transceiver. The receiver’s burst extraction algorithm can be reused minimising additional implementation of the GSM protocol. However, to perform experiments on public mobile networks, access to public base stations is required, but was not available in this study. The second possibility is to capture the radio signals as a third party, using a receiver which does not take part in the communication.

Reising et al. [8] capture the radio signals of mobile phones located next to the receiver using a threshold to extract transmitted bursts from these phones. Although quite simple and straight forward, this burst selection process can not be used in a practical environment with a bigger distance between the mobile phones and the receiver. Other phones might interfere and the extracted bursts of an individual phone may become mixed up with other communication streams. Instead, we used a two way receiving software defined radio (SDR) to capture GSM signals from the mobile phone and from the base station at the same time. The SDR needs to be placed in receiving range of both communication entities. That way the complete communication stream can be observed in both directions without interfering or disrupting the ongoing GSM communication.

We developed a specialised receiver software able to act as a mobile phone for the frames received from the base station and as a base station for the remaining frames. The signals of both directions are interpreted according to the GSM specification to extract the correct communication stream and the corresponding bursts of each mobile phone under test. As a side effect, our software extracts the whole protocol flow of the communication and we are able to identify the mobile phone on protocol basis to assure the correctness of the identification algorithm. Because of the special scenario, we could only reuse a small part of existing open

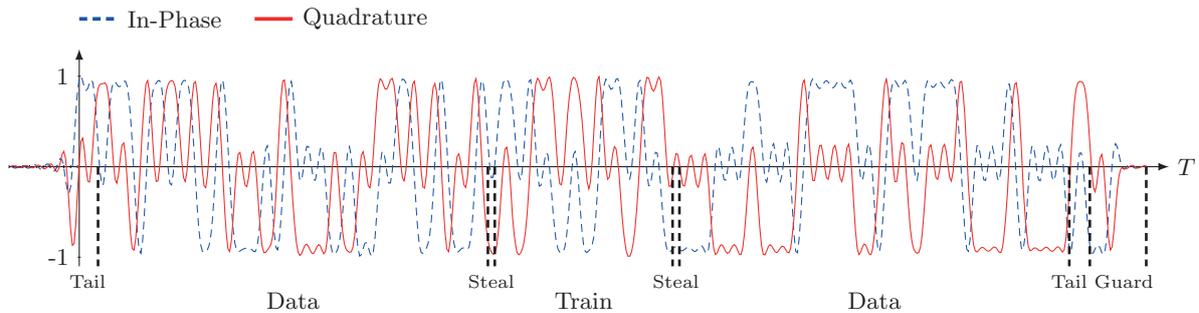


Figure 3: Structure of a Captured Normal Burst

source projects like Airprobe¹ and OpenBTS². The following remarks describe some interesting aspects we encountered during implementation.

3.1 Synchronisation

The receiver must be able to synchronise to a base station. Every base station sends a special burst containing only zeros on a regular basis on the main radio channel. After modulation, the signal contains a sine wave at a frequency of approximately 67.7 kHz. Deviations from this frequency are used to calculate clock offsets in our receiver and in mobile phones connecting to the base station. The clock of the base station acts as a common reference.

Perfect synchronisation is not possible, which introduces random frequency offsets for the mobile and our receiver on each connection. To reduce the random influence, our receiver works with a constant frequency offset which is determined once for the acquisition device. The offset stays constant even when monitoring different base stations. This is possible, because the base stations are required to have a very precise clock with frequency deviations of 0.05 ppm (parts per million) or less. Still remaining frequency offsets are compensated in the signal preprocessing stage. Time synchronisation is established using a special synchronisation burst sent by the base station, containing the current state of the time multiplex system.

3.2 Encryption

As described in Section 2, some part of the communication stream is usually protected by encryption. For every attempt to make a call, an observer can obtain at least 32 bursts from the same mobile phone before the communication switches to an unknown logical channel. For practical reasons, we want to capture a lot of packets within a short period of time. The transmitted speech data of a voice call produces up to 217 bursts per second. When encryption is activated, the logical channel and the underlying mapping to physical resources of the traffic channel is unknown. To be able to observe the bursts, the mobile phone must be able to reveal the temporary encryption key or the encryption needs to be broken.

In our experiments, all public base stations activated encryption with the A5/1 cipher algorithm. We used a known-plaintext attack proposed by Karsten Nohl [7] to calculate the temporary key with an exhaustive search for all mobile

phones under test, which were not able to reveal the key due to software limitations. To use this attack on encrypted GSM bursts, the keystream has to be determined. A plaintext predictor for special GSM messages was implemented as described by Sylvain Maut³. This cipher attack was successfully applied on the captured traffic of all mobile phones under test, carefully leaving out other transmissions not relating to our experiments. Although this attack is capable of calculating one key in 4 seconds on specialised hardware, our software receiver is not able to find the key in real time of the communication.

Recall that the proposed identification algorithm does not require decryption of the communication stream in principle. In case 32 bursts are not enough for identification, the correct logical channel and physical mapping of the speech data transmission could potentially be guessed for base stations which are rather idle. The decryption method was chosen for practical reasons.

3.3 Frequency Hopping

Base stations may decide to employ frequency hopping, which results in changing the radio channel for each time-frame during communication. While experimenting, some of the analysed base stations never used frequency hopping, whereas some hopped every time and others seemed to decide randomly whether to use hopping or not. Depending on the configuration of the cell, the communication stream can be spread over a wide bandwidth. To follow the hopping sequence, it would be possible to retune the receiving device in real time, like a mobile phone does, which however highly depends on the capabilities of the acquiring device. We chose to capture the whole bandwidth the communication might be spread over to be independent of real time tuning operations. However recording two bands simultaneously with a bandwidth of several MHz each requires a careful setup of the recording device to cope with the high data throughput and to avoid dropping frames.

3.4 Receiver Signal Processing

The raw captured signal needs to be preprocessed in order to demodulate the GSM signal. After recording at a wide bandwidth, the recorded signal is split into individual GSM radio channels using a polyphase filter bank channeliser. This filter is implemented in software running on a computer, which decreases performance dramatically. In later versions this task could be performed in real time by

¹<https://svn.berlin.ccc.de/projects/airprobe/>

²<http://wush.net/trac/rangepublic>

³<http://web.archive.org/web/20100808001500/http://lists.lists.reflexor.com/pipermail/a51/2010-July/000804.html>

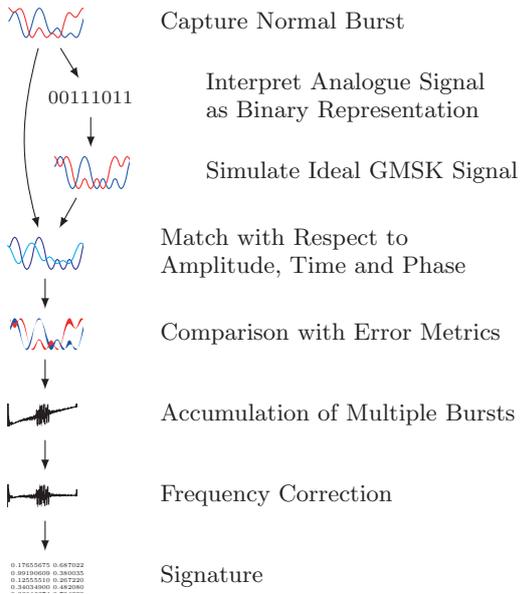


Figure 5: Overall Procedure of Feature Extraction

specialised hardware like an FPGA system. For synchronisation with the target base station, frequency offsets are corrected using a finite impulse response filter. Due to technical limitations of the acquisition device, we had to include a fractional resampler, aligning the sample rate to a multiple of the GSM symbol rate. The performed signal processing is the same for every run of the receiver and for every individual channel to minimise the introduction of new characteristics to the signal.

4. PHYSICAL CHARACTERISTICS

Every mobile phone has a RF frontend which contains hardware components working on the analogue signal. When a mobile phone sends a burst, the digital signal passes through a digital-analogue converter, a band pass filter, mixers and an amplifier. Inaccuracies in the manufacturing process result in minor physical differences of these components. Even components coming from the same stack of a manufacturing procedure do have different properties caused by random effects, slight differences in material or sub-components.

Because ideal operation of these components is not possible, they are manufactured and sold in classes of error tolerance. The tolerances in the processing chain add up to a reasonable amount of unintentionally introduced errors in the resulting RF signal. When designing a RF communication system, the allowed error tolerance of the RF signal is well specified to ensure proper operation. These errors can be measured and used to generate a unique fingerprint of a mobile phone’s RF components. This fingerprint is not easy to forge, because it would involve replacing hardware components fixed on the circuit board of the mobile phone. The ‘normal bursts’ collected from a communication stream by the GSM receiver software are the basis for the calculation and detection of the fingerprint.

Rather than analysing the raw RF signal as done by Reising et al. [8], our analysis interprets the signal according to the Gaussian Minimum Shift Keying (GMSK) modulation.

For every extracted burst, the receiver demodulates the signal to produce a binary representation which is employed to create a mathematical ideal simulation of the modulated burst. The differences between every observed and ideal sample are used to estimate error metrics which can be employed for identification. Brik et al. [2] shows this procedure for the QPSK modulation of IEEE 802.11 and confirms applicability for robust physical device identification. The following describes the overall procedure (c.f. Fig. 5) to extract suitable characteristics from GSM normal bursts based on modulation errors.

4.1 Simulation

The simulation of GMSK is quite complex. Testing several software based implementations of GMSK modulators, we observed common systematic differences when comparing an ideal signal to the realisation of a mobile phone. The OpenBTS implementation showed the best similarity, but still required slight changes to the Gaussian pulse in order to match the general shape of the collected signals for our over-sampling rate. The modulation was optimised exemplarily using the tested mobile phones as a reference and remained constant throughout all experiments. Even with this optimisation, a small portion of systematic differences remains, which results in content dependent fluctuations of the error metrics. These variations are expected to result from different practical hardware implementations of a GMSK modulator, compared to the mathematical model.

4.2 Matching

For device identification, it is essential to extract characteristics which remain stable over different dimensions and especially over time. We try to identify and remove all aspects which introduce random behaviour or which can be attributed to side effects, which the identification should be invariant of. The first varying aspect is the sending power. The base station can order the mobile phone to use different levels of sending power, depending on the current reception strength. Additionally, the reception strength varies depending on environmental conditions.

For an invariant comparison, the captured in-phase and quadrature signal is normalised. A filter detects the maximum and minimum peaks by comparing 3 neighbours for each side of a sample. The detected peaks are ordered by absolute value and a threshold decides which samples to include in averaging. The average of the selected peaks is normalised to 1, thus the maximum and minimum peaks match 1 and -1, respectively. This is based on the assumption, that the changes in amplitude due to movement of the target phone is negligible for the duration of a single burst, i.e. $577 \mu s$. According to the GSM specification [5], a mobile phone is allowed to send a burst with an arbitrary phase offset. In our experiments, we confirmed for every mobile phone under test, that the overall phase offset of a burst behaves randomly. We remove this effect by aligning the captured signal to a simulation with a zero phase offset for all experiments. Because of the interdependency to time, we approximate the correct time t_o and phase offset ϕ_o simultaneously with an increasing precision using overall correlation as the optimisation metric. Employing prior knowledge about time alignment from the receiver software, we first match the signal on sample level and try different phase offsets of $\phi_o = \pi$ and $\phi_o = \pi/2$. After matching the

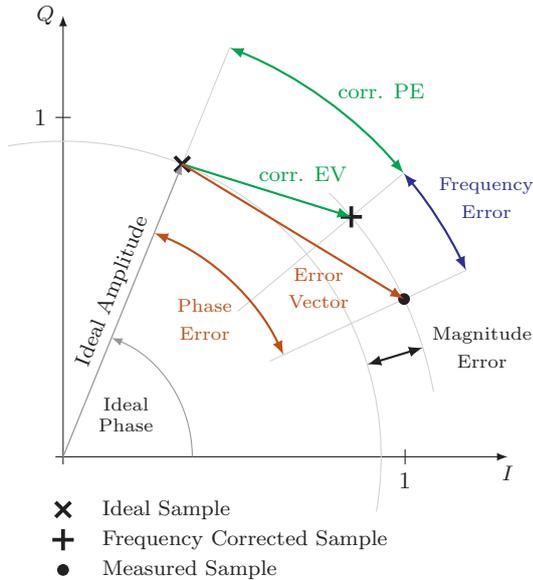


Figure 6: Error Metrics with Frequency Correction in I/Q Diagram

signals on coarse precision, a more precise ϕ_o is calculated by the average difference between the ideal phase ϕ_i and observed phase ϕ , N being the count of samples in the burst:

$$\phi_o = \frac{1}{N} \sum_{n=0}^N (\phi(n) - \phi_i(n)). \quad (1)$$

The methods described here to achieve phase and time alignment can be exchanged with more sophisticated algorithms available in latest RF technology to improve computational performance. As the last preprocessing step, we align the observed and simulated signal in time on a fractional sample precision. With a fractional timing offset in the range $(-1, 1)$, a linear optimiser selects the best fractional t_o , maximising correlation. Note that the timing offset can not be used as a device identification feature, because it is dependent on the location of the mobile phone. The identification procedure should be invariant of the current location, as long as the mobile phone resides in the reception range of our receiver.

4.3 Error Metrics and Frequency Correction

When the observed signal is aligned with the simulation, different error metrics can be employed to quantise the differences between the signals. Depicting an observed sample in a constellation diagram, the length of the position vector represents the current amplitude and the angle to the I axis represents the current phase at a certain point in time (c.f. Fig. 6). When compared to an ideal sample, the difference in amplitude is called Magnitude Error (ME) and the phase difference is called Phase Error (PE). The vector between the observed and ideal sample is the Error Vector (EV) and its length the Error Vector Magnitude (EVM). These are common error metrics describing the precision of a modulated signal. In GSM, phase errors may occur up to 5° RMS⁴

⁴Root Mean Square, $\text{RMS}(\mathbf{v}) = \sqrt{\frac{1}{N} \sum_i^N v_i^2}$. (N denotes the number of elements in \mathbf{v} .)

and up to 20° peak. The EVM may deviate 9-10% RMS and up to 30% peak [6, 4]. The ME is limited implicitly by the EVM and is not specified separately as a modulation error metric.

Following the approach of Brik et al., we initially tried to quantify the accuracy of a mobile phone's RF hardware using the average of the aforementioned metrics over a single collected burst to measure the accuracy of a mobile phone's RF hardware as an identification characteristic. This characteristic was not stable enough for identification, because the random part of PE and ME did not allow to detect the deterministic errors of the observed signals. Taking the high precision of GSM into account, we propose to use characteristic error patterns over the time of a normal burst as a device-dependent feature.

As the process of sending a normal burst is always the same, the RF hardware introduces deterministic deviations at specific times of a burst, e.g. fluctuations of the power amplifier. With consideration of the error metrics in respect to the time of a burst, it is possible to evaluate both time-dependent and modulation-dependent characteristics. Systematic differences introduced by the modulation algorithm of different mobile phones consolidate in the training sequence and tail bits during accumulation, because the modulated bits do not change, assuming a constant training sequence. Time dependent fluctuations of the radio hardware are captured in all regions of the burst.

To improve identification performance, we accumulate the error metrics for each sample position over all available bursts of one signature. This is possible, because all bursts were aligned in time to match the individual simulation, which makes them independent of the individual time offset. When t is the sample position of a burst and M the total count of bursts contributing to one signature, the accumulated ME trajectory \mathbf{a}_{ME} is determined with the average at each sample position:

$$\mathbf{a}_{\text{ME}}(t) = \frac{1}{M} \sum_{i=0}^M \text{ME}_i(t). \quad (2)$$

The \mathbf{a}_{PE} and \mathbf{a}_{EVM} are accumulated likewise. After accumulation, the \mathbf{a}_{PE} trajectory reveals a dependency to a linear model. The slope of this model is a remaining frequency error, which is attributed to imperfect synchronisation mechanisms in mobile phones and our receiver.

In initial experiments, we tested the frequency error as a possible characteristic for identification, but found the characteristic was unstable due to randomness of the synchronisation procedure and noise effects, compared to the required precision of the mobile phones RF hardware after synchronisation. We determine the linear frequency model with a least squares approximation and deduct the frequency related part from PE in respect to time, resulting in the frequency corrected phase error, \mathbf{a}_{PEfc} .

As can be seen in Figure 6, the frequency error also influences the EV. Thus, we need to correct the observed samples to compensate the errors introduced by the detected frequency offset resulting in the $\mathbf{a}_{\text{EVMfc}}$. Note that the frequency error has only negligible effect on the ME⁵. The \mathbf{a}_{PEfc} , $\mathbf{a}_{\text{EVMfc}}$ and \mathbf{a}_{ME} are the error patterns we can use as input for a classification algorithm. Each of these time de-

⁵The frequency error only influences the ME in combination with strong I/Q imbalance, which we did not observe.

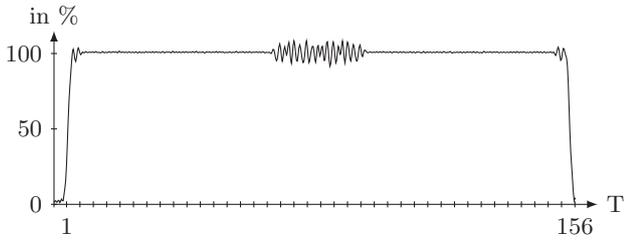


Figure 8: \mathbf{a}_{PWR} Trajectory of a Motorola C118 Mobile Phone

pendent patterns contain the typical errors at specific sample positions of a normal burst. Because we used an oversampling rate of 4 at the GSM symbol rate for one burst, we end up with $148 \text{ Bit} \cdot 4 = 592$ error metrics for each trajectory.

To justify the selection of these patterns, we show two different examples relevant for device identification. In the first comparison, we calculate the \mathbf{a}_{PEfc} trajectory of the same mobile phone at two different locations, accumulating over 1000 bursts each for demonstration purpose. As shown in Figure 7, the \mathbf{a}_{PEfc} is very similar for the same device, regardless of the current location. In the area of the training sequence, the two trajectories are almost identical. The second comparison shows the same trajectory for two mobile phones of the same brand and model. The differences of the two trajectories represent different characteristics of the RF hardware of the individual phone. For the two phones, the differences are visible especially in the data sequences. Apart from two big and two minor spikes, the training sequences are similar. Note that systematic errors depending on the modulated bits add up in the training sequence and the tail bits, but average out in the data sequences, because the modulated bits change randomly for each burst⁶. The following experiments will evaluate the performance of these characteristics in an identification scenario.

4.4 Power Trajectory

Unlike other digital modulation schemes, GMSK has a constant envelope, i.e. the signal power does not drop to zero and remains level in the transition between different states of the constellation diagram. The GSM specification requires the mobile to keep the power level constant during the transmission of a normal burst [4]. We can use this property to obtain the amplitude related errors more easily. With the aforementioned normalisation defining the average peaks of the in-phase and quadrature signals to be 100%, the observed signal only needs to be aligned in time in order to obtain a power trajectory PWR. In this work we used the whole burst in combination with the simulation for time synchronisation. However, time synchronisation can be achieved using the training sequence only, as done in almost all GSM receivers. This makes the PWR trajectory independent of the simulation and renders the computational intensive matching procedure obsolete for this feature. This can be crucial for identification scenarios requiring real time operation. When accumulating over M bursts, the accu-

⁶For the collected bursts, the encryption algorithm was active which produced a pseudo-random bit stream for the data sequences on physical layer.

mulated power trajectory \mathbf{a}_{PWR} can be calculated for each sample position t of a normal burst using the normalised in-phase signal I and quadrature signal Q as follows:

$$\mathbf{a}_{\text{PWR}}(t) = \frac{1}{M} \sum_{i=0}^M \sqrt{I_i(t)^2 + Q_i(t)^2}. \quad (3)$$

The resulting trajectory looks like depicted in Figure 8. Before and after the burst, additional time is granted to power the amplifier up and down and to adjust to the target power level just before the burst transmission starts. This trajectory is similar to the \mathbf{a}_{ME} , but represents the general sending power of an RF signal instead of magnitude errors in the modulation domain. By not interpreting the signal according to the modulation, the processing time is greatly reduced for amplitude based errors. Similarly to the \mathbf{a}_{ME} , the intentional power fluctuations of GSM add up in the training sequence whereas they average out in the data sequences. The following experiments will evaluate the performance of the presented features in an identification scenario.

5. TEST SETUP

In the following, we describe the practical details of our experiments. All parameters are summarised in Table 2. For signal acquisition, we used two USRP N210 devices operating in synchronised MIMO mode. Each of them is equipped with a daughterboard covering an analogue bandwidth of 40 MHz. The device recording the uplink frequency band was equipped with a +3 dB 900/1800 MHz GSM antenna, the device capturing the downlink band used a general purpose antenna covering a wide frequency range. The acquisition site was inside of an office building, without any arrangements like shielding. The site was exposed to other common radio signals present in office buildings such as wireless networking at the time of acquisition.

For the selected acquisition radios, commodity hardware can record radio signals with a sample rate up to 10 MHz for two simultaneous channels. Recording with higher bandwidths is only possible with optimised hardware. When selecting a base station to monitor, we had to take the limitations of the recording capabilities into account. We selected a base station operated by T-Mobile, because of the strong signal and a recordable combination of radio channels used for frequency hopping. For the selected base station, the main radio channel and the channels used for the hopping sequence were separated by 10-15 MHz. To cover the main and the hopping channels in one recording with an available recording bandwidth of 10 MHz, we had to split each band in two different recording blocks, one capturing the main radio channel (30) and the other capturing the hopping radio channels (82-102). So recording was performed on four channels in total with a sample rate of 5 MHz each, based on two analogue radio sources. With the same technique of splitting an analogue source into two different recording blocks, it is also possible to record communication of a GSM base station with a single USRP and daughterboard only, assuming frequency hopping is deactivated and the base station sends strong enough signals in the E-GSM band.

For our experiments, we used a total of 13 mobile phones of 4 different manufacturers and 9 models (c.f. Tab. 1). Note that the Sony Ericsson J100i and all Motorola phones are designed and manufactured by Compal Inc. and sold under a branded name. They share the same system de-

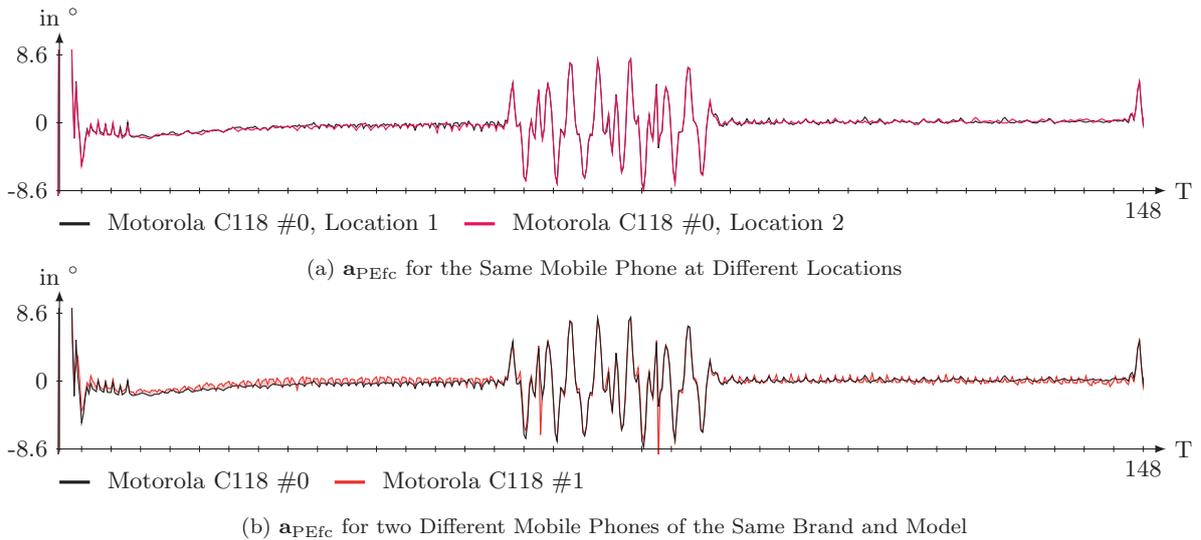


Figure 7: Two Examples of \mathbf{a}_{PEfc} Trajectories

Table 1: Mobile Phones Used in Tests

Manufacturer	Model	Chipset	#
Motorola	C115	Calypso G2 (D751749ZHH)	1
Motorola	C118	Calypso G2 (D751749ZHH)	4
Motorola	C123	Calypso G2 (D751749ZHH)	1
Motorola	C139	Calypso G2 (D751749ZHH)	1
Motorola	C140	Calypso G2 (D751749ZHH)	1
Sony Ericsson	J100i	Calypso G2 (D751749ZHH)	1
Nokia	6100	UPP8M/MJOELNER S2006	1
Nokia	E51	BB5 SL2 RAPIDO	1
HTC	TyTNII	Qualcomm MSM7200	1
Palm	Pre	Qualcomm MSM6801A	1

sign and use very similar or the same RF hardware chipsets. The four Motorola C118 phones can only be distinguished by the IMEI. This selection of mobile phones was chosen to provide the most difficult identification problem due to the high amount of identical hardware chipsets. For the training stage, the phones were placed next to the receiver. While the acquisition was running, each mobile phone was voice called twice, the calls were answered and the phones transmitted for 45 seconds each call. For the test stage, the mobile phones were placed at a different location, 4 metres away from the receiver. Because the new location used in the test set is completely unknown in the training stage, the performance of the identification algorithm is evaluated location independent. Because of different ranges, the captured signals have different signal to noise ratios with 13.3 dB and 5.4 dB for train and test set, respectively. Note that our identification system ignores bursts containing massive amounts of bit errors. In order to produce a valid signature, the simulation needs to be correct. Bursts with an EVM error of more than 60% RMS were skipped. Approximately 25% of all frames were dropped because of this requirement due to a very basic demodulation algorithm. This can be improved by using a more sophisticated demodulator able to determine the binary representation more precisely. An

alternative would be to implement bit correction for the simulation.

For classification, each signature was accumulated using 30 passed bursts, with training and test counts of 150 and 100. The influence of these parameters is analysed in Section 6. The signatures were put into a linear Support Vector Machine (SVM) for classification. To calculate the SVM model, the training signatures were split in two equally sized parts. The first group was used to estimate the parameter of the linear kernel using a grid search with cross validation. The model was then determined using the second group of signatures. Thus, the identification model was not optimised to the new location of the test set and optimisation of the kernel was done without manual correction.

6. EXPERIMENTS

We start our practical investigations with individual evaluations of each proposed feature. The True Acceptance Rate (TAR) is used as a performance indicator. The TAR is the probability of detecting a given device correctly. The average TAR over every device under test symbolises the overall success rate of an experiment. The results in Table 3 indicate that the \mathbf{a}_{PEfc} is the most successful characteristic, identifying a device correctly with an overall probability of 96.67%. Amplitude-based features do not work as good as phase-based error metrics. The \mathbf{a}_{ME} performs better than the \mathbf{a}_{PWR} feature, because the \mathbf{a}_{ME} can remove content dependent effects in the data sequences more efficiently by comparing to the ideal simulation instead of simple averaging of the signal power. If the computational resources are available to produce and match an ideal simulation, the \mathbf{a}_{ME} should be preferred to the \mathbf{a}_{PWR} . The mixed characteristic \mathbf{a}_{EVMfc} which is based on both amplitude and phase characteristics, performs almost as good as the \mathbf{a}_{PEfc} .

To further analyse synergy effects between these features, we test every possible combination and evaluate the performance using the average TAR (c.f. Tab. 3). The obvious best performing combination is \mathbf{a}_{PEfc} with \mathbf{a}_{EVMfc} virtually matching the performance of the individual \mathbf{a}_{PEfc} . Except

Table 2: Parameters of Test Setup

Acquisition	Device	USRP N210, SBX USRP N210, WBX
	Sample Rate	5 MHz-4
	GSM Oversampling	4
Cell	Provider	T-Mobile
	Uses Hopping	Yes
	Main Channel	30
	Hopping Channels	82-102
Phones		13
Locations	Training	At Receiver (13.3 dB)
	Test	4 m away (5.4 dB)
Classification		Linear SVM
	Bursts per Signature	30
	Training Signatures	150
	Test Signatures	100

Table 3: Overall Performance of Individual and Combined Features

a_{PWR}	a_{ME}	a_{PEfc}	a_{EVMfc}	
Individual				
60.75 %	68 %	96.67 %	93.33 %	TAR
Combinations				
		•	•	96.50 %
•		•	•	92.25 %
	•	•	•	89.50 %
•	•	•	•	89.00 %
•		•	•	87.83 %
•	•		•	87.83 %
		•		86.33 %
	•	•		86.17 %
	•		•	85.75 %
•	•	•		85.33 %
•	•			69.83 %

for the combination of a_{PWR} and a_{ME} , the identification rates of the best individual feature was better compared to a combination with other features. Nevertheless, we used the combination of the best individual performing feature a_{PEfc} and a_{EVMfc} for the following experiments to include as many relevant device dependent characteristics as possible.

Within our second experiment we analyse the interdependency between the two classification parameters accumulation count and total amount of bursts per phone in training. The accumulation count varies from 10 to 50, while the total amount of bursts for one phone varies between 500 and 4500. The amount of training signatures is calculated by dividing the total bursts count per phone by accumulation count, e.g. for an accumulation count of 50 and a total training burst count of 4500, the number of training signatures equals 50. The count of test signatures remains constant at 100 signatures per phone for all training parameter combinations.

The results are illustrated in Figure 9 and document an overall increase of performance when using more bursts in training for each phone. When 3500 bursts or more are available for the training procedure, the chosen accumulation count has only marginal influence on the overall success rate. For lesser amounts of training bursts, it is generally

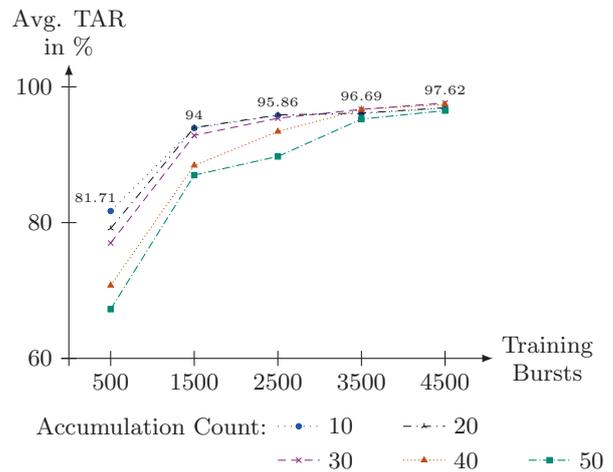


Figure 9: TAR as a Function of Accumulation Count and Total Amount of Bursts per Phone in Training

better to use a lower amount for the accumulation count in order to produce a higher number of training signatures. For 1500 and 2500 available training bursts per phone, an accumulation count of 20 is optimal while an accumulation count of 30 produces best results for 3500 and 4500 training bursts. Collecting a training set of 4500 bursts would require monitoring a voice call for 21 seconds. Note that with a low accumulation count of 30 or less, it is possible to identify a known mobile phone without the need to break the encryption of the GSM traffic. For a mobile phone to establish a communication channel, at least 30 bursts have to be transferred on a publicly known physical channel. This is enough to produce one or more test signatures, possibly using sliding windows. This way, any phone with a previously recorded model can be identified very easily.

In the last experiments, we determined the optimal configuration for a classification system. Using the best performing parameters of 30 for accumulation and 4500 for training burst count, we show the detailed results of the identification procedure in Table 4 omitting all zeroes for visual clarity. For 13 mobile phones and a total of 1300 test signatures, only 31 signatures were not classified correctly. The Motorola C123 was detected with the worst TAR of only 87%, mistakenly detecting the Nokia 6100 in 7% of cases. Also, the Motorola C123 and Motorola C118 #3 seem to share similarities in the selected hardware characteristics, because they are confused with each other in 4% of cases. Six devices were identified perfectly, i.e. every available test signature was matched to the correct class. With the overall average TAR of 97.62%, the classification performance is slightly better compared to the work of Reising et al. [8] at our noise level and clearly outperforms in the amount of classes showing the practical relevance of our approach.

7. CONCLUSION

This work is a first step in mobile forensics to identify mobile devices in a GSM network without relying on traditional identifiers like IMEI or IMSI. By targeting the air interface of GSM on physical layer, it is possible to identify mobile phones without the interaction with or recognition by the sender. We proposed to detect physical properties

Table 4: Confusion Matrix for Twelve Mobile Phones (in Percent)

	Original Phone	Identified As														
		C115	C118 #0	C118 #1	C118 #2	C118 #3	C123	C139	C140	J100i	6100	E51	TyTNII	Pre		
	Motorola C115	100														
	Motorola C118 #0	1	99													
	Motorola C118 #1			100												
	Motorola C118 #2			1	99											
	Motorola C118 #3					96	4									
	Motorola C123	2				4	87									
	Motorola C139				2			98								
	Motorola C140								100							
	Sony Ericsson J100i									100						
	Nokia 6100										100					
	Nokia E51											100				
	HTC TyTN II	1	2										94			
	Palm Pre														96	

of a mobile phones RF hardware to create a unique fingerprint for every individual device. The environment of the performed experiments was chosen to be a most realistic one, passively monitoring the communication on a public GSM network. The radio signals were captured according to the GSM specification at varying locations. Based on previous work by Brik et al. [2], features of modulation accuracy originally targeting IEEE 802.11 were adopted for GSM to be used as a device identifier. These features have been improved, resulting in time based features describing a characteristic pattern of modulation deviations over the time of a normal burst. Using the improved features, a total of thirteen mobile phones have been correctly identified with an overall success rate of 97.62%. This included four identical and nine almost identical phones, which proves the selected features to be unique for an individual device.

Compared to previous work by Reising et al. [8], we were able to improve the overall detection performance using a practically relevant implementation of the burst selection process. At the same time we increased the amount of mobile phones under test considerably. Training and test signatures were obtained at different locations assuring location independence. Our proposed model uses the well defined domain of GMSK modulation, backed by the modulation requirements of the GSM specification. We showed that identification based on RF hardware inaccuracies is possible even for a high precision communication system like GSM.

Future work should analyse the influence of environmental conditions such as temperature or movement of a target device, as well as the robustness against noise and a bigger distance between the mobile phone under test and the listening receiver. The features can potentially be improved by using additional statistical parameters other than the average during accumulation. The burst extraction process may be improved to work in real time employing digital signal processing hardware. Further analysis should target the robustness against potential attacks trying to forge a device signature.

8. ACKNOWLEDGEMENTS

Accommodation and travel was partly funded by the Federal Ministry of Economics and Technology of Germany, the European Union and the European Social Fund.

9. REFERENCES

- [1] K. Bonne Rasmussen and S. Capkun. Implications of radio fingerprinting on the security of sensor networks. In *Third International Conference on Security and Privacy in Communications Networks and the Workshops (SecureComm' 2007)*, pages 331–340, 2007.
- [2] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom '08)*, pages 116–127, 2008.
- [3] ETSI TS 100 908 V8.11.0 (2003-06). *Digital cellular telecommunications system (Phase 2+) — Multiplexing and Multiple Access on the Radio Path — 3GPP TS 05.02 version 8.11.0 Release 1999*, 2003.
- [4] ETSI TS 100 910 V8.20.0 (2005-11). *Digital cellular telecommunications system (Phase 2+) — Radio Transmission and Reception — 3GPP TS 05.05 version 8.20.0 Release 1999*, 2005.
- [5] ETSI TS 100 959 V8.4.0 (2001-11). *Digital cellular telecommunications system (Phase 2+) — Modulation — 3GPP TS 05.04 version 8.4.0 Release 1999*, 2001.
- [6] P. Kimuli. Introduction to GSM and GSM mobile RF transceiver derivation. *RF DESIGN*, 26(6):12–21, 2003.
- [7] K. Nohl. Attacking phone privacy. https://srlabs.de/blog/wp-content/uploads/2010/07/Attacking.Phone.Privacy_Karsten.Nohl1.pdf, 2010.
- [8] D. R. Reising, M. A. Temple, and M. J. Mendenhall. Improved wireless security for GMSK based devices using RF fingerprinting. *International Journal of Electronic Security and Digital Forensics*, 3(1):41–59, 2010.
- [9] D. Zanetti, P. Sachs, and S. Capkun. On the practicality of UHF RFID fingerprinting: How real is the RFID tracking problem? In *Privacy Enhancing Technologies*, volume 6794 of *LNCS*, pages 97–116, 2011.