Freaky Leaky SMS: Extracting User Locations by Analyzing SMS Timings

Evangelos Bitsikas
bitsikas.e@northeastern.edu
Northeastern University

Theodor Schnitzler
theodor.schnitzler@tu-dortmund.de
Research Center Trustworthy Data Science and Security

Christina Pöpper
christina.poepper@nyu.edu
New York University Abu Dhabi

Aanjhan Ranganathan
aanjhan@northeastern.edu
Northeastern University

Abstract

Short Message Service (SMS) remains one of the most popular communication channels since its introduction in 2G cellular networks. In this paper, we demonstrate that merely receiving silent SMS messages regularly opens a stealthy side-channel that allows other regular network users to infer the whereabouts of the SMS recipient. The core idea is that receiving an SMS inevitably generates Delivery Reports whose reception bestows a timing attack vector at the sender. We conducted experiments across various countries, operators, and devices to show that an attacker can deduce the location of an SMS recipient by analyzing timing measurements from typical receiver locations. Our results show that, after training an ML model, the SMS sender can accurately determine multiple locations of the recipient. For example, our model achieves up to 96% accuracy for locations across different countries, and 86% for two locations within Belgium. Due to the way cellular networks are designed, it is difficult to prevent Delivery Reports from being returned to the originator making it challenging to thwart this covert attack without making fundamental changes to the network architecture.

1 Introduction

Despite the emergence of smartphones and various messaging applications, the Short Message Service (SMS) remains an essential communication channel for sending and receiving text messages. SMS is widely used in marketing campaigns, appointment reminders, short customer surveys, and even as part of two-factor authentication [37], identity verification, and security/emergency alerts [41, 42]. Since its introduction in the GSM standard in the early 1990s, SMS remains a key service across cellular generations, including 5G standards [4].

SMS's prevalence, global reach, and message delivery reliability have made it a significant attack vector in recent years. For example, smishing attacks [25] use an SMS with malicious links to direct victims to a phishing website and deceive them into divulging sensitive information. The Flubot virus [19] in 2021/22 spread via SMS containing links to trojan apps that accessed sensitive data like banking credentials, contacts, and disabled security options. SMSes have also been used for spamming [14]. Simjacker [9] and its variant WIBAttack [48] are other malware examples that use binary-embedded SMS messages.

In this work, we take an orthogonal approach and show how an attacker can subtly exploit SMS and determine a victim's location. Prior location identification and localization techniques in the cellular domain relied on retrieving the temporary and permanent identifiers of a mobile device using false base stations [26, 49], using them to track the user's whereabouts within a certain area. Authorities worldwide have used silent[1] SMSes [31] to uncover their owners' locations, however, these approaches rely on cooperation from the network operators and/or necessitate sniffers.

Unlike the above attacks, our attack does not require access to the network operator's infrastructure or false base stations around the victim's area of interest. Instead, it works by leveraging SMS Delivery Reports, which are transmitted back to the sender when the network delivers the SMS to the recipient. The sender can request these reports, and there is no way for the recipient to prevent them. By measuring the round-trip time, i. e., the time elapsed between sending an SMS and receiving the corresponding Delivery Report, our attack can distinguish various locations of the target recipient and determine their location area after a training phase. The attacker can behave like a regular user and does not require access to advanced equipment, but a typical smartphone device.

Consider the following scenario of a nation's diplomat (victim) giving a press conference from a specific location, e. g., official residence. Given the public knowledge

[1] Silent SMSes are not displayed by the victim's mobile.

of the victim's current location and phone number, the adversary starts sending silent or regular SMSes to the victim and collects their round-trip time measurements, generating *timing signatures* for that specific location. Then, at a later time, when the attacker wants to infer whether the victim is back in their residence, the adversary simply sends a silent SMS and determines whether the timing signatures match. Since Delivery Reports are solidly rooted in protocol specifications across all mobile network generations with no possibility to disable them, the attacker can reach the victim at any time by solely possessing their mobile phone number. It is hard to disable the attack without a significant overhaul of the cellular network specifications.

To the best of our knowledge, this work is the first to identify an SMS Delivery Report-based timing side-channel that leaks location information. Our work makes the following contributions:

1. We enumerate the various cellular network components that contribute to the timing delays and identify six timing-related features to create a robust location signature. Based on the location signature, we design an approach to execute our SMS location inference attack.
2. We perform a large-scale study collecting Delivery Report timing measurements across three continents, nine countries, and ten operators to create our training dataset. We send SMS messages between devices and measure Delivery Reports return times within and across different setups in the US, multiple countries in Europe, and the Middle East.
3. We use the collected measurements to evaluate the performance of our location inference attack. Our experiments show that we can achieve up to 75% and 96% accuracy for location identification in nearby and far countries, respectively. Our model achieves over 70% for many cases within a country or certain region, such as within Germany, Netherlands and Belgium. We analyze factors affecting the accuracy of our location classification and perform network and temporal stability analyses for additional evaluation.
4. We discuss potential countermeasures against SMS timing attacks, including enforcing random or uniform delays in the core network.

In summary, we highlight the effectiveness of inferring location based on SMS Delivery Reports and the challenges associated with mitigating such an attack.

Responsible Disclosure: The privacy issues caused by the SMS timing attack have been recognized by GSMA on the *GSMA Mobile Security Research Acknowledgements* page under *CVD-2023-0072*. GSMA has been considering several countermeasures, including artificial delays and robust SMS filtering.

Longer Version: This version focuses on the primary aspects of the location inference attack. Please refer to the arXiv version for additional results and technical details at https://arxiv.org/abs/2306.07695.

Code Release: The code along with the dataset are publicly available on Github at https://github.com/vagelis-sudo/SMS-Location-Identification-Attack.

2 Background on SMS Networks

First, we explain the various types of network architectures that are used for SMS exchanges. Then, we illustrate how the SMS procedure in such networks works and which timing delays are involved, respectively.

2.1 Cellular network architectures

Figure 1 shows an extended version of 4G/LTE (a)–(b) and 5G standalone (c) architectures for the SMS procedure including the 2G/3G structures. In this work, we focus on LTE and 5G networks.

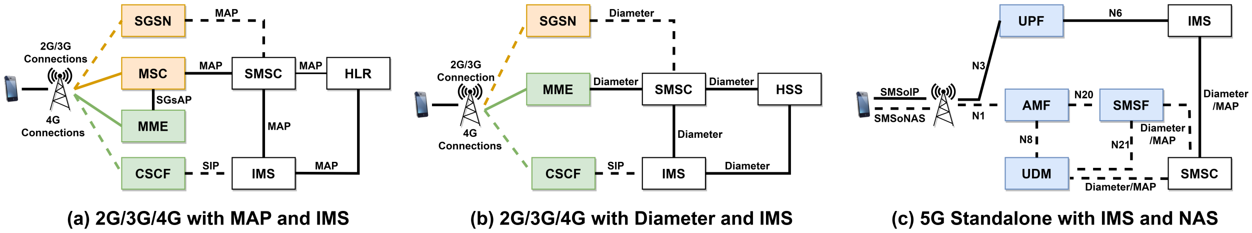
5G has two SMS delivery routing paths and protocols: SMSoIP and SMSoNAS. SMSoIP or IP-based communication (data-plane) leverages the SIP protocol and the IP Multimedia Subsystem (IMS) architecture [5–7] to communicate with the Short Message Service Center (SMSC). SMSoNAS uses the Non-Access Stratum (NAS) protocol for SMS transmission and delivery, providing NAS encryption and integrity-protection [2, 3] through control-plane traffic after establishing the security context [4, 8].

Furthermore, LTE services support chiefly IP-based communication through the IMS (Figure 1), then alternatively the SGsAP interface, which eliminates the need for 2G/3G fallback, and finally, the NAS signaling communication combined with the Diameter protocol [20]. Typically, the IMS incorporates the IP Short Message Gateway (IP-SM-GW), an IMS Application Server that handles SIP-based messaging services for IMS subscribers.

The selection between SMSoNAS and SMSoIP depends on the SMS originator and the network support, even though IP-based communications are more prevalent, as the User Equipment (UE) subscribes to the IMS after completing the Authentication and Key Agreement (AKA) procedure with the Core Network.

2.2 SMS procedure

SMS services are accessible to all network generations (2G-5G) [20] as a process of exchanging short text messages between two network subscribers. The SMS exchange between originator and recipient requires forwarding to the Core Network, where the SMSC man-



(a) 2G/3G/4G with MAP and IMS

(b) 2G/3G/4G with Diameter and IMS

(c) 5G Standalone with IMS and NAS

Figure 1: The SMS architectures based on the protocols and generations.

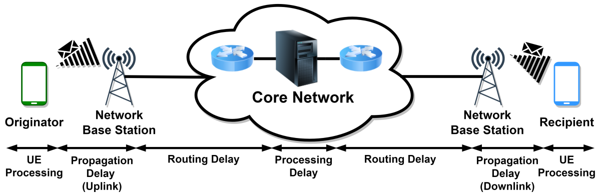


Figure 2: Various delays for one SMS transmission between two users. Similar delays apply for the Delivery Report which is sent back to the originator.

ages the SMS process and delivery (Figures 3 and 4). After receiving the message, the recipient sends a Delivery Report, which is forwarded through the SMSC to the originator acknowledging the delivery. Delivery Reports provide detailed information on the status of every message sent including "Delivered", "Accepted", "Failed", "Undeliverable", "Expired", and "Rejected". Failed deliveries could be due to incorrect phone number, disabled international roaming, unreachable recipient, mobile plan restrictions, etc.. Note that delivery notification is enabled by the originator in their phone's settings on modern smartphones. Delivery Reports are used for data cleansing/updating, improving response rates, audit trail, and systems monitoring.

There are three primary SMS statuses: i) *Sent*, which indicates that the mobile device has sent the SMS to the SMSC and the SMSC has confirmed its reception, ii) the *Delivered*, meaning that the recipient has received the SMS and has responded with the Delivery Report, and iii) *Failed* when errors occur.

2.3 Network Delay Factors

SMS text transmissions and Delivery Reports incur timing delays in the communication channel. Figure 2 illustrates the delays for a single SMS transmission between the originator and the recipient.

(1) UE Processing: This is the time taken by the phone to process the SMS for transmission or reception. The corresponding base station has already completed its transmission at that time. The processing includes the modem and OS procedures, and the user-related services used (e. g., calls, SMSes, mobile data) that occupy uplink and downlink resources.

(2) Propagation Delay: This depends on the RAN network's design, configuration, and deployment including the front-haul of the mobile network, physical properties and quality of the signal, transmission capabilities of the base station, and management of the uplink and downlink communications.

(3) Routing Delay: SMS messages pass through multiple network entities depending on the architecture and the generation (e. g., LTE, 5G, etc.). Routing delays occur in the mobile back-haul, i. e. the transport network that connects the core network and the RAN, as well as within the core network. Apart from the SMSC and the gateways, the SMS may require additional processing, e. g., by the AMF (5G SA), MME (LTE-5G NSA), and IMS, before reaching the destination, thereby also contributing to the routing delays.

(4) Processing Delay: This delay generally includes the SMSC, the IMS, and MSC/MME/AMF processing. The SMSC manages the SMS reception and delivery process and may also deploy congestion, filtering, and prioritization techniques.

3 SMS-based Location Inference Attack

The high-level idea of our attack is as follows: The time elapsed between sending an SMS and receiving the corresponding SMS Delivery Report differs depending on the receiver's current location, implying one can distinguish different receiver locations by observing the elapsed time.

3.1 Attacker Goal and Assumptions

The attacker's goal is to locate the victim receiver's whereabouts, specifically, whether the victim's mobile is in a specific geographic area of interest.[2]

We assume that the attacker knows the victim's mobile number and can send an SMS to that number. The SMS can be regular private messages, undirected mass

[2] We do not tackle the tracking of exact movement patterns of the victim in this paper.

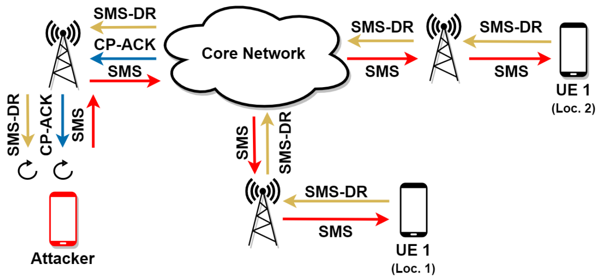


Figure 3: Network flow for SMS transmissions in different locations.

messages (e.g., marketing, advertisements) that the victim will likely ignore, or a silent SMS that victim's device acknowledges without any content or alerts, remaining entirely unnoticed by the victim. We assume the attacker can target any subscriber (victim) with a valid mobile number attached to a cellular provider and maintain a typical connection to send text messages to the victim and receive delivery notifications. The adversary can access any network operator using the corresponding (e)SIM as a normal user.

Additionally, we assume the attacker can collect measurements from locations of interest directly from the victim when located at specific locations/areas of interest (without revealing the attack) or deploy similar devices and connections as the victim at these locations for data collection. The attacker is *not* limited in terms of the number of smartphone devices, (e)SIMs, mobile numbers, or subscription plans. The attack does *not* require physical access to the victim's USIM cards, mobile devices, or any network entities (e. g., base stations, core network, etc.). Finally, the attacker neither obtains nor modifies sensitive information, e. g., cryptographic keys.

3.2 Timing Features

As shown in Figure 3, SMS transmissions to different device locations generate acknowledgments from the core network (CP-ACK) associated with *Sent* notification and Delivery Reports (SMS-DR) from the receivers resulting in the *Delivered* status. Hence, the attacker can leverage three timestamps to execute the attack:

- t_{tx}: `SMS transmit time` as the time when attacker sends the SMS,
- t_{sent}: `SMS sent time` as the time when the attacker receives the "Sent" notification, and
- t_{del}: `SMS delivery time` as the time when the attacker receives the "Delivered" notification.

After the SMS transmission is complete, the real sent duration T_{sent}, the real delivery duration T_{del}, the total delivery duration T_{tot}, and the delivery ratio P can be calculated as follows:

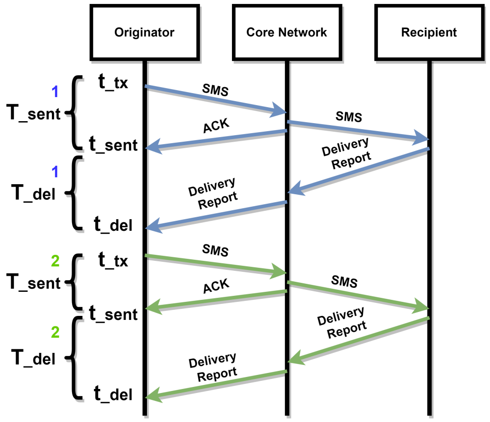


Figure 4: Timing features for each SMS transmission.

$$T_{sent} = t_{sent} - t_{tx} \tag{1}$$

$$T_{del} = t_{del} - t_{sent} \tag{2}$$

$$T_{tot} = T_{del} + T_{sent} \tag{3}$$

$$P = \frac{T_{del}}{T_{tot}} = \frac{t_{del} - t_{sent}}{t_{del} - t_{tx}} \tag{4}$$

These features apply to each individual SMS transmission only. Figure 4 shows the timing features for two SMS transmissions.

To produce robust location signatures and generate a pattern, we consider two consecutive SMS transmissions ($i-1$ and i), and estimate the difference in real sent duration $T_{\Delta sent}$ and real delivery duration $T_{\Delta del}$, respectively:

$$T_{\Delta sent} = (T^{i}_{sent} - T^{i-1}_{sent})/T^{i-1}_{sent} \tag{5}$$

$$T_{\Delta del} = (T^{i}_{del} - T^{i-1}_{del})/T^{i-1}_{del} \tag{6}$$

The *location signature* is a combination of these six features: $(T_{sent}, T_{del}, T_{tot}, P, T_{\Delta sent}, T_{\Delta del})$.

3.3 Attack Concept

The attack is conducted in two phases: (i) A *Preparation* and (ii) an *Attack* phase.

In the *Preparation* phase, the adversary repeatedly sends multiple (silent) SMS, with Delivery Reports enabled, to the victim while observing their respective locations. The attacker collects measurements to identify the timing characteristics of the victim's locations. Despite being aware of the victim's locations at this stage, the victim will not notice that they are being surveilled when the adversary uses silent SMSes. Using these measurements and analyzing the different timing features outlined in Section 3.2, fingerprints for each of the victim's locations are generated.

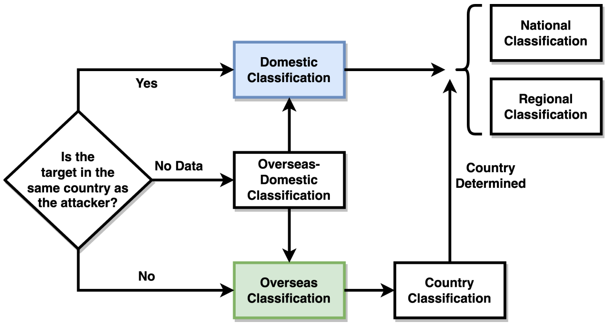


Figure 5: The classification methodology

In the *Attack* phase, the adversary collects new measurements without knowing the victim's location and attempts to determine their current location based on the timings. To do this, the adversary must solve a *classification* problem, i. e., assign the newly observed measurements to one of the previously seen locations by comparing timings with the respective location fingerprints. Depending on the victim's movement patterns and the locations observed in the preparation phase, the classification occurs in multiple iterations. Therefore, the classification problem is partitioned into a step-wise location prediction problem involving several location identification tasks with decreasing granularity levels from classifying international locations to regional (e. g., at city-level).

Classification Methodology. We describe the classification approach that the attacker follows to retrieve a victim's location in multiple iterations (Figure 5). We use the example of a victim moving internationally.

Initially, the attacker may not have sufficient intelligence regarding the victim's current country of residence. Thus, the first step is to determine whether the victim is *Overseas* or *Domestic*. If the victim is overseas, then the attacker proceeds with determining the specific country (country-based classification). Once the country is known, the attacker may choose to perform either a national or regional classification depending on the attacker's objectives and the victim's routine. In the regional classification, the attacker attempts to discover the victim's location within a limited area, while the national classification has a macroscopic view of the country, incorporating cities and towns.

Having knowledge about the victim's general geographical whereabouts such as North America, can help narrow down potential candidate locations making classification more manageable. If there is only one country and one city, the methodology can be simplified to just regional location identification. Therefore, the attacker does not need to adhere to the entire methodology as it primarily depends on the victim's routine.

4 Experimental Validation

In this section, we present our experimental validation of the SMS-based location inference attack. We describe our measurement setup and the different stages of the two attack phases outlined in Section 3. Figure 6 provides an overview of our experimental procedure.

4.1 Setup

We send SMSes between smartphones at different geographical locations to collect measurements for our experiments. Our setup includes *active devices* (phones) controlled via the Android Debug Bridge (ADB) to send SMSes to other devices. These phones are configured to analyze cellular traffic and baseband logs to extract timing and network information such as protocols, connections with the core network, AT SIM commands, etc. Active devices have SMS Delivery Reports enabled to visualize notifications while sending messages. *Passive devices* are used to receive messages. We list all the devices used in Table 8.

Our devices are located across several countries, including the United States (US), UAE (AE), and seven countries in Europe (BE, DE, DK, GR, LU, NL, UK). The experiments cover ten operators and several generation technologies such as LTE, LTE+, 5G NSA/SA. Additionally, we record the approximate channel condition such as strength and quality, for each receiving location. Table 1 presents the relevant characteristics of all locations that appear in our measurements.

We conduct three rounds of measurements serving different purposes:

(i) We conduct long-distance international measurements with devices in multiple countries, with the sender located in AE-1.

(ii) We send messages from a single active device to passive devices at various domestic locations, including multiple cities and locations within them, for AE, GR, DE, NL, BE, and LU. The experiments are conducted from different sender locations, with the sender in AE-1 for AE experiments, GR-1 for the GR experiments, and DE-4 for the rest. The primary objective is to demonstrate a practical and realistic scenario involving a person's natural everyday behavior on a smaller scale including regular commuting to adjacent countries.

(iii) We collected measurements across different operators and roaming devices at several locations. Specifically, we focused on distinguishing between network operators and smartphone devices which assists our location identification.

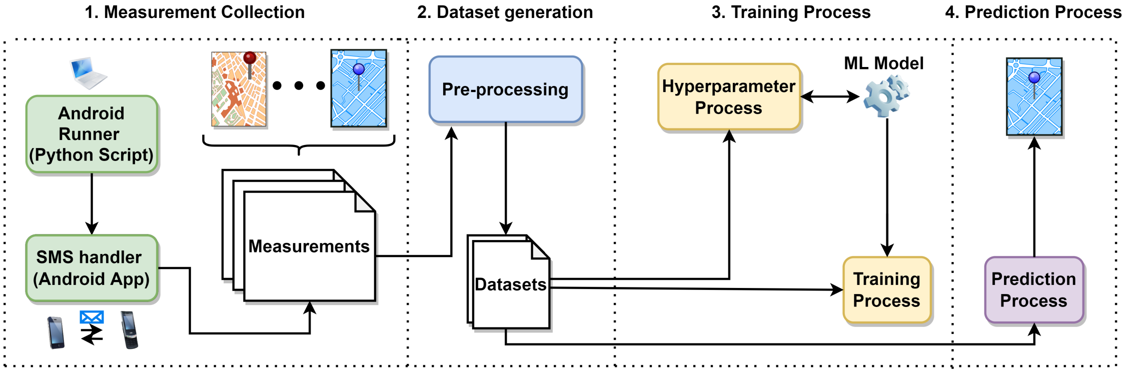


Figure 6: Components and stages of the SMS location identification attack.

4.2 Measurement Collection

Our data collection is sketched in step 1 of Figure 6. We developed an Android application called *SMS handler* that runs on active devices and sends one silent SMS at a time to a target device. Once the SMS is sent, the application waits for the Delivery Report (both *Sent* and *Delivered* notifications) and records all the required timestamps and computes the features (1)-(4) (cf. Section 3.2).

We use a python script, *Android Runner*, to automate SMS transmission to a designated receiver and capture the Delivery Report timings for each SMS. The script interacts with the smartphone through basic ADB commands and key events (to press buttons, fill text input fields, etc.) without requiring device rooting. The script runs on a Dell Latitude E5450 and a regular desktop computer (cf. step 1 in Figure 6) using a *cronjob* for repeated execution.

We schedule *SMS burst*, i. e., consecutive 20 SMS transmissions, on an hourly basis. To distribute the SMSes for each location, we span them over 2 to 3 days to avoid potential SMS spam filtering and prevent network congestion, which may affect the timings. This procedure also helps us collect representative traffic dataset, including various times of the day, potential network configuration changes, and different levels of network loads. Throughout our measurement campaign, we have sent and accumulated around 155,512 SMSes. Refer to Table 9 for SMS numbers per device, per country, and per operator.

We constantly monitor whether the active device sent the silent SMS successfully during our experiments. We use the Android logging tool *Logcat* to investigate the routing methods and connection establishments and track the SMS procedures.

4.3 Dataset Generation

We now describe how we aggregate our collected measurements to generate the evaluation dataset (step 2 in Figure 6). We calculate the timing features from the collected data, generate location signatures, each composed of all six timing features obtained during or derived from a single measurement iteration (Section 3.2).

Our evaluation dataset contains signatures for each candidate location, covering various granularity levels, from domestics and overseas to national and regional classifications (cf. Figure 5). In our data, we also identify the SMS routing modes, i. e., SMSoIP for LTE/LTE+, SMSoIP for 5G, and SGsAP/Diameter for LTE/LTE+.

4.4 Location Classification

We opted for `Multilayer Perceptron (MLP)` using Python's SKLearn libraries as classifier to perform location classification because of its flexibility in parametrization and high performance on large datasets. The model comprises a stochastic gradient descent solver, softmax and sigmoid activations for multiclass and binary classifications respectively, and three layers with 10, 40, and 10 nodes respectively for the input, hidden, and output layers. Additionally, we set the maximum iterations to 5000, the learning rate to be constant, batch size to be 32, and the alpha to 0.0001. We performed automatic and manual parameter tuning to improve the model's accuracy (Appendix A provides more details). We focus on *accuracy* throughout our classifications, measuring the number of correct predictions out of the total predictions made.

The training and prediction procedures correspond to steps 3 and 4 in Figure 6. The datasets are randomly split, while the class with the highest probability is assigned by the MLP classifier as the prediction result. Training and prediction processes utilize the cross-validation methodology with 10 k-folds to prevent over-fitting and pro-

Table 1: Receiver locations and their characteristics. GR-1, AE-1 and DE-4 acted as senders (using LTE/LTE+/5G) and receivers, but the table focuses on the receivers only. Channel conditions show the approximate connection quality from our devices in those locations. Receivers that have ranges in *Distance* column represent an area instead of a specific fixed position.

Rec.	Dist. [km]	Connection Type	Routing	Cond.	Operator
International Receiver Locations					
Int-GR	3266	LTE,LTE+	SMSoIP	▮▮▮▯	C
Int-DE	5460	LTE,LTE+	SMSoIP	▮▮▮▮	E
Int-DK	5880	LTE+,5G NSA/SA	SMSoIP	▮▮▮▮	I
Int-UK	5700	5G NSA/SA	SMSoIP	▮▮▮▮	H
Int-US	10.710	LTE,LTE+	SMSoIP	▮▮▮▯	J
Receiver Locations in the UAE					
AE-1	1-7 meters	5G NSA/SA	SMSoIP	▮▮▮▮	A, C
AE-2	10	5G NSA/SA	SMSoIP	▮▮▮▯	A, B
AE-3	14	LTE,LTE+	SMSoIP	▮▮▯▯	A, B
AE-4	135	LTE+,5G SA	SMSoIP	▮▮▮▮	A
Receiver Locations in Greece					
GR-1	1-5 meters	LTE,LTE+	SMSoIP	▮▮▮▯	C, D
GR-2	8	LTE,LTE+	SMSoIP	▮▮▮▯	C
GR-3	12	LTE,LTE+	SMSoIP	▮▮▮▮	C
GR-4	180	LTE	SMSoIP	▮▮▮▯	C
GR-5	200	LTE	SMSoIP	▮▮▯▯	C
GR-6	290	LTE	SMSoIP	▮▮▮▯	C
Receiver Locations in Germany					
DE-1	11	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
DE-2	45	LTE,LTE+	SGsAP/Diameter	▮▮▮▯	E,F,G
DE-3	2	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
DE-4	0	LTE,LTE+	SGsAP/Diameter	▮▮▮▯	E,F,G
DE-5	31	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
DE-6	0 – 5	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
DE-7	0 – 35	LTE,LTE+	SGsAP/Diameter	▮▮▮▯ – ▮▮▮▮	E,F,G
DE-8	110 – 130	LTE,LTE+	SGsAP/Diameter	▮▯▯▯ – ▮▮▮▮	E,F,G
DE-9	0 – 110	LTE,LTE+	SGsAP/Diameter	▮▯▯▯ – ▮▮▮▮	E,F,G
DE-10	59	LTE,LTE+	SGsAP/Diameter	▮▮▯▯	E,F,G
Receiver Locations in the Netherlands					
NL-1	130	LTE,LTE+	SGsAP/Diameter	▮▮▮▯	E,G
NL-2	125	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	G
NL-3	90	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	G
NL-4	129	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
NL-5	128 – 130	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
Receiver Locations in Belgium					
BE-1	195	LTE,LTE+	SGsAP/Diameter	▮▮▮▯	E,F,G
BE-2	153	LTE,LTE+	SGsAP/Diameter	▮▮▮▯	E,F,G
BE-3	116 – 210	LTE,LTE+	SGsAP/Diameter	▮▯▯▯ – ▮▮▮▮	E,F,G
Receiver Locations in Luxembourg					
LU-1	220	LTE,LTE+	SGsAP/Diameter	▮▮▮▮	E,F,G
LU-3	165 – 225	LTE,LTE+	SGsAP/Diameter	▮▯▯▯ – ▮▮▮▮	E,F,G

Locations (Cities/Regions): *Int-GR*: Athens, *Int-DE*: Bochum, *Int-DK*: Copenhagen, *Int-UK*: London, *Int-US*: Boston, *AE-1,3*: Abu Dhabi, *AE-2*: Saadiyat, *AE-4*: Dubai, *GR-1,2,3*: Athens, *GR-4*: Chania, *GR-5*: Messenia, *GR-6*: Thessaloniki, *DE-1*: Dortmund, *DE-2*: Raesfeld, *DE-5*: Unna, *DE-3,4,6*: Bochum, *DE-7*: Ruhr Area, *DE-8*: Aachen Area, *DE-9*: NRW State, *DE-10*: Borken, *NL-1,4,5*: Veldhoven, *NL-2*: Eindhoven, *NL-3*: Roermond *BE-1*: Bastogne, *BE-2*: Sankt-Vith, *BE-3*: Wallonia Region *LU-1*: Luxembourg City, *LU-3*: Western Regions
Operators: *A*: Etisalat (UAE), *B*: du (UAE), *C*: Cosmote (GR), *D*: Vodafone (GR) *E*: Telekom (DE), *F*: Vodafone (DE), *G*: Telefonica (DE), *H*: Vodafone (UK), *I*: Telenor (DK), *J*: Mint (US)

mote model generalisation. We also compared the performance of a Random Forest Classifier, Decision Tree Classifier, and Recurrent Neural Network with Keras libraries, but the optimized MLP outperformed them all. Therefore, we present our results for the MLP model only.

5 Location Classification Results

We follow the classification methodology outlined in Section 3 proceeding step-by-step from coarse- to fine-grained location classifications and present our results.

5.1 International Classification

For the international classification, we focus on large geographical areas of the victim, primarily attempting to identify locations in different countries. Our results are shown in Table 2.

Overseas-vs.-Domestic Classification aims to determine whether the victim is within the home country or abroad. This binary classification experiment groups the AE locations (home country) together and Int-X locations together. The results indicate that the target can be identified with an accuracy of 96%. The two box plots in Figure 8a show a clear timing difference between the two classes based on the Delivery Report (T_{del}), facilitating accurate identification.

Country-based Classifications aim to determine the victim's location in a specific country. First, we conduct experiments on countries that are far apart to demonstrate the existence of timing differences. We perform multi-class classification for all Int-X locations in different countries and achieve 96% accuracy. The box plots in Figure 8b depict the timing difference in the dataset between GR, DE, DK, UK, and US locations. Next, we select only EU countries for a multi-class classification to identify locations within a smaller geographical area. We used Int-GR, Int-DE, Int-DK locations (sender AE-1, based on another continent) achieving 95% accuracy.

In Figure 7, we present the confusion matrices for the overseas-vs.-domestic and country-based classifications (from Table 2). The figure confirms the high-accuracy results from the table and identifies the predictions that lead to less accurate results, involving classification with sender DE-4 and nearby receiver countries. For operators G and E, LU and NL receiver locations result in higher misclassifications than for DE and BE. The model also shows a loss of accuracy for operator F, where timing characteristics for DE, LU, and NL cause errors due to similarities, but the most likely returned result is still the correct one for each case.

Finally, we performed a country-based classification targeting adjacent and nearby countries to identify even closer geographical locations. The victim traveled to DE-4, NL-4, BE-1, and LU-1 using operators G, E, and F. Our classifiers achieved 75%, 74%, and 62% accuracy

Table 2: Classification results for international experiments.

Classification	Size/Class	Operators	Receiver Locations	Sender Location	Accuracy
Overseas-vs.-Domestic	1200	A, C, E, H, I, J	AE-X, Int-X	AE-1	96%
All Country-based	280	C, E, H, I, J	Int-X	AE-1	96%
EU Country-based	280	C, E, I	Int-GR, Int-DE, Int-DK	AE-1	95%
EU Country-based	257	G	DE-4, NL-4, BE-1, LU-1	DE-4	75%
EU Country-based	319	E	DE-4, NL-4, BE-1, LU-1	DE-4	74%
EU Country-based	313	F	DE-4, NL-4, BE-1, LU-1	DE-4	62%

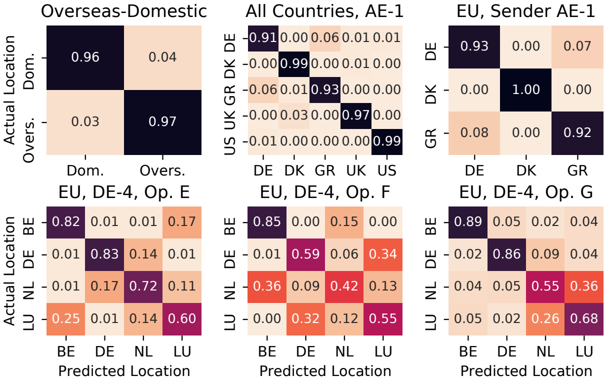


Figure 7: Confusion matrices for international classifications, displaying the results from Table 2 in more detail. Although the model misclassifies more often for operator F, it achieves high accuracy in many cases.

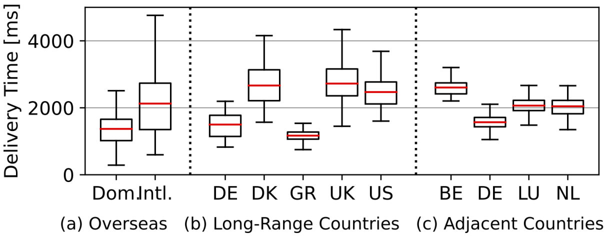


Figure 8: Delivery timings for receivers in various countries.

for these specific locations using operators G, E, and F, respectively. These three EU country-based classifications with four classes have an average accuracy of 70% with the best performing being 75% for operator G and E. Figure 8c shows the timing difference between those countries with NL-4 and LU-1 having similar delivery timings. However, raw delivery timing is only one of the six features we take into consideration in this case.

5.2 National & Regional Classification

In this section, we explore the location characteristics at a regional scale within the same country. Our classifications include receiver locations from Table 1, which can be either fixed locations or areas through which a receiver is moving. We evaluate our model against fixed locations, areas, and their combination within each country. We repeated the classification for every combination of locations in our dataset, with sample sizes varying from 100 to 500. Table 3 summarizes our results, broken down by receiver country and the number of locations, and includes the repetition with the largest sample size, which depends on the available data for each location.

Fixed Locations. Our classification achieves an average performance of 68 % in Germany based on 57 classifications of pairs of two locations. However, performance varies depending on the pairs of locations, so the average must be interpreted carefully. The best performing classification (DE-3 and DE-5) achieves 92 % classification accuracy. Detailed results for all pairs of locations in Germany are presented by the matrix in Table 7. The average performance for the Netherlands across 15 classifications of location pairs is 63 %, with 98 % classification accuracy for NL-2 and NL-3. For Belgium, the overall performance is 86 %, but this only includes four classifications of the same two locations (BE-1 and BE-2) 40 km apart from each other, using different phones.

Our classification scores decrease for larger sets of locations in all countries, but it should be noted that the chance of randomly guessing the correct location is also lower (e. g., 33 % for 3 locations instead of 50 % for 2 locations). Nevertheless, the average classification scores of 76 % and 79 % in the UAE and in Greece, respectively, still indicate a high performance.

Areas with Multiple Locations. Areas can be challenging to distinguish as they are not associated with the attributes of one location only and may overlap. We report area classification results for DE locations in Table 7. In binary classifications, the model achieves an average accuracy of 57 % for 21 classifications, with DE-6 and DE-8 being the best-performing pair reaching 72 %. For three and four classes in DE, the model achieves 41 % and 34 %, respectively. Similar to the fixed locations, performances should be read and understood separately, as each combination has different features.

Mixed Locations. In this scenario, we explore the combinations of fixed locations and areas, which shows that the attacker is not limited to distinct types only. We used

Table 3: Summary of regional/national classifications within the same country.

Type	All Classifications		Best Performing	
	Num*	Avg. Acc.	Loc. Set	Accuracy
Regional classifications with 2 locations (Random: 50 %)				
DE Fixed	57	68 %	DE-{3,5}	92 %
NL Fixed	15	63 %	NL-{2,3}	98 %
BE Fixed	4	86 %	BE-{1,2}	95 %
DE Area	21	57 %	DE-{6,8}	72 %
DE Mixed	80	67 %	DE-{8,10}	88 %
NL Mixed	4	71 %	NL-{3,5}	88 %
BE Mixed	8	77 %	BE-{2,3}	84 %
LU Mixed	4	67 %	LU-{1,3}	72 %
Regional classification with 3 locations (Random: 33 %)				
AE Fixed	1	76 %	AE-{1,2,3}	76 %
GR Fixed	2	79 %	GR-{1,2,3}	82 %
DE Fixed	46	54 %	DE-{2,5,10}	83 %
NL Fixed	7	48 %	NL-{1,2,3}	68 %
DE Area	13	41 %	DE-{6,7,8}	50 %
DE Mixed	252	50 %	DE-{5,6,10}	81 %
NL Mixed	6	59 %	NL-{1,3,5}	78 %
BE Mixed	4	67 %	BE-{1,2,3}	73 %
Regional classification with 4 locations (Random: 25 %)				
AE Fixed	1	58 %	AE-{1,2,3,4}	58 %
DE Fixed	19	47 %	DE-{1,2,5,10}	64 %
NL Fixed	1	53 %	NL-{1,2,3,4}	53 %
DE Area	3	34 %	DE-{6,7,8,9}	38 %
DE Mixed	402	41 %	DE-{2,5,9,10}	67 %
NL Mixed	4	48 %	NL-{1,2,3,5}	58 %
Regional classification with 5 locations (Random: 20 %)				
DE Fixed	3	37 %	DE-{2,3,4,5,10}	50 %
DE Mixed	398	34 %	DE-{2,3,5,8,10}	55 %
NL Mixed	1	42 %	NL-{1,2,3,4,5}	42 %

**Num* denotes the numbers of different classifications, e. g., for different sets of receiver locations and phones.

measurements from DE, NL, BE, and LU for the classification tasks in Table 3. In binary classifications, the model achieves 67 %, 71 %, 77 % and 67 % on average for DE, NL, BE and LU locations, respectively, while reaching up to 88 % in certain classifications. The model scores lower for classifications that include three, four, and five locations. For example, DE has an average accuracy of 50 %, 41 %, and 34 % for three, four, and five classes, respectively. Nonetheless, the large number of classifications with even diverse features should be taken into account cautiously, i. e., 252, 402, and 398 for three, four, and five classes, respectively.

The performances of classifications are highly variant depending on the sets of locations. Figure 9 illustrates the distribution of the performance of all classifications. We also present detailed results for individual classifications for all pairs of locations for each country in Tables 4 – 6 in the Appendix.

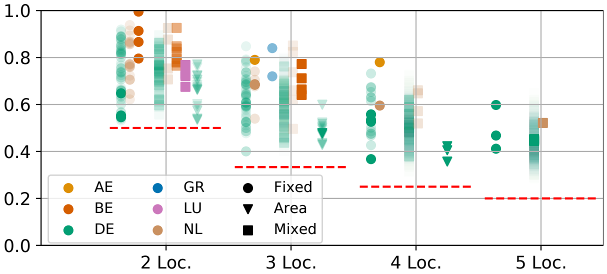


Figure 9: Classification accuracy of regional/national classifications within the same country. Dashed red lines indicate the probability of randomly guessing the correct location.

5.3 Misclassification Errors

In location identification, a misclassification error for an SMS measurement means that the timing pattern is matched to the wrong location, i.e., wrong pattern distribution. False results can arise due to various machine-learning (ML) factors, such as overfitting and model complexity, as well as in the form of outliers due to special network conditions. In any case, more sophisticated and motivated adversaries with more resources and ML expertise may enhance the model to improve the attack.

Country-based classifications are primarily impacted by factors such as adjacency between countries and network homogeneity (including similar operators), making it more challenging to distinguish locations. The impact of these factors can be seen for operators E and F in Table 2 and in Figure 8 for LU and NL. In fixed locations, timing similarities between locations (with the same UE and operator) can make the classification less precise due to congruent variance in network delays. It can be even more challenging when locations are very close and have similar signal conditions, such as NL-1 and NL-2 for operator G (Table 6) with 62 % accuracy. However, this is not always the case, as in the classification of DE-3 and DE-4 for operator E (Table 7) which achieves 87 % accuracy. In addition, areas and mixed classifications can be similarly difficult to distinguish, as they combine measurements from multiple distinct locations and may overlap. Nonetheless, Tables 6 and 7 include high accuracy scores even in such cases.

6 Additional Evaluations

In this section, we present evaluations that provide additional insights, such as the impact of geographic separation of different receiver locations, and show how an open-world classification affects the performance of the attack. We also present results from the temporal stability and network timing analyses in which we collected additional data for several DE and NL locations.

6.1 Temporal Stability

We perform a temporal stability analysis to determine if the attack can still work even after some time has elapsed since the model was trained. For this purpose, we modified the original attack evaluation by training the model on a baseline dataset and testing it on measurements collected X days after the training phase. Therefore, we collected new and protracted data for the same locations with similar operators and devices to accommodate experimentation for up to one month after the training.

Figures 10a and 10b depict eight examples of how the accuracy fluctuates for DE-4/NL-4 and DE-4/NL-2 classifications in a span of 35 days. We used operators G, E, and F with Huawei P8 Lite (p8l), Google Pixel 6a (px6a), Samsung Galaxy A53 (a53), and OnePlus 7 Pro (op7) devices. Each trend represents specific device(s) and operator. According to the graphs, each combination has a distinct trend, as their measurements' characteristics differ. Consequently, increases and decreases in accuracy between various days are also expected for classifications in which the model scores both with high and low accuracies. Furthermore, in Figures 10a and 10b, operator E with the p8l device is more susceptible to degradation than the rest of the combinations, but it takes more than 23 days for the degradation to slowly appear. Operator G with p8l in the DE-4/NL-4 classification shows a small degradation but retains high accuracy after 35 days. As a result, the collection of new data and retraining may not be necessary for all classifications. For classifications that continue to have high scores, the attacker may continue using their data.

6.2 Network Analysis

We evaluate the impact of congestion, potential network changes, and other time-varying characteristics by running the location classification separately for different days and times of the week. The classification process is the same as the regular attack but with specific test data slices for different times of the day and days of the week. We grouped measurements into four sets for different times of the day (0-5, 6-11, 12-17, and 18-23) and seven sets for days of the week. We use data collected at two locations (DE-4 and NL-4) with sufficient measurements in our dataset for separate analyses across time slices, multiple phones, and operators.

Figure 11 shows the classification accuracy for two victim phones with one operator (G) throughout the entire week. The scores remained consistently high, with scores of 88 % and 89 % for OnePlus and Samsung, respectively. Figure 12 also shows the model's performance for different time windows using four phones with three different operators. While performance differs across operators, with classification only working for G achieving around 80 % and above, the scores generally remain stable throughout the day. The experiments illustrate results for specific locations, devices, and operators, and hence do not allow to draw general conclusions regarding the localization accuracy of specific devices. For the purpose of completeness, we acknowledge and report less accurate results as well.

6.3 Distances Between Locations

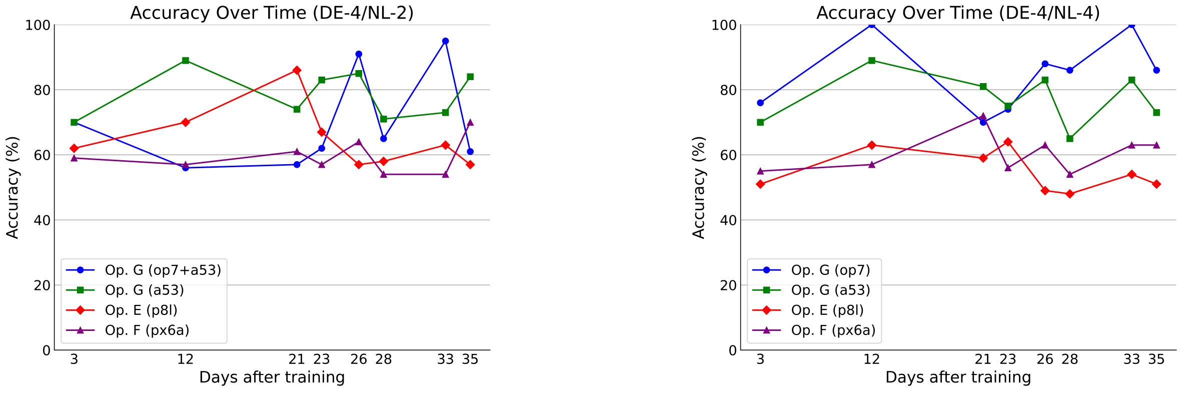
In this part, we analyze the relationship between classification accuracy and distances between the locations using pairs of two *fixed* locations in Germany, The Netherlands, and Belgium. Figure 13 shows the average classification accuracy for all pairs of locations in these three countries. It reflects the impact on accuracy of (a) the geographical distance between the two receiver locations, and (b) the distances between the sender and each of the receiver locations. For the latter, we consider the average of the two distances.

We found no correlation between distances and accuracies, contradicting the assumption, that receiver locations further apart from each other or from the sender would result in more accurate classification. Therefore, distance may not be the main factor affecting classification accuracy.

6.4 Open-world Scenarios

Open-world cases refer to unknown/unseen locations, for which the attacker has not accumulated measurements for model training. We discuss three methods to tackle these cases that can be used separately or in combination.

First, the attacker can utilize outlier/anomaly detection mechanisms and unsupervised one-class classifications to reduce the "nearest neighbor" effect and identify if the data belong to an unknown location. Although this separate study requires thorough experimentation and we consider its comprehensive evaluation as future work, we carried out an experiment using an `Isolation Forest` model. The model was configured with 100 estimators (without parameter tuning) and was trained on the domestic (AE) dataset attempting to identify overseas measurements during the prediction phase. With each class having 1200 samples, as indicated in Table 2, it achieved an 88% accuracy for anomaly detection indicating that the predicted data belong to an unseen location.



(a) DE4-NL2

(b) DE4-NL4

Figure 10: Accuracy trends of the DE4-NL2 (a) and DE4-NL4 (b) classification for various operators and devices until 35 days from the model training.

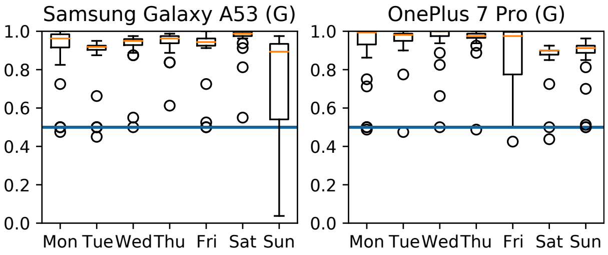


Figure 11: Network analysis of DE-4/NL-4 classification for different days and devices (Operator G).

Second, the attack can be enhanced by modifying the MLP classification model to output the probability of the user being in a specific location instead of the predicted class. We have modified our initial model to run further experiments. Figure 14 illustrates the probabilities (per row) for fixed and area classifications in AE, DE, and GR with three distinct SMS transmissions/samples (i. e., 0, 1, 2), respectively. For AE and GR, the results show that the probabilities do not fall below 80%. For the specific DE area classifications in Figure 14b, the probabilities are more evenly allocated since the model cannot decisively decide the true class, especially in the first SMS transmission. In this case, the attacker may perform further assessments for the top two (DE-9 and DE-7) classes, or conclude that the victim might be located in one of them.

Third, the adversary can reduce the chances of unknown classes by expanding the measurement campaign to more potential locations that are not routinely tied to the victim (e. g., famous landmarks). There are research works (focusing on WiFi) that collect data from various places within cities and areas [46, 47], while targeting either Access Points (APs) or smartphone devices. Additionally, the attacker can focus on utilizing areas instead of fixed positions to expand the coverage. Although this approach may not reveal the exact position (which can be translated to GPS coordinates) of the victim if the area incorporates too many positions, it allows the attacker to still track the victim without relying on the routinely fixed locations. However, the extensive data an attacker needs to collect beforehand may limit the practicality of this approach. In general, the attacker might prefer to resort to a binary decision, i. e., to determine whether or not the victim is at one of their previously seen locations, as described in the first two methods. In Section B in the Appendix, we provide more information on how the attacker can manage large dataset collections in the context of handling unseen locations.

7 Discussion

Our study provides insights into how different locations of SMS receivers can be distinguished based on measuring the time it takes to deliver an SMS. In this section, we discuss potential countermeasures to mitigate the attack at different levels as well as the limitations of our study.

7.1 Countermeasures

UE-based countermeasures. On UE devices, defenses can be implemented at the application layer or become a part of the system firmware which could be suitable for low-level cellular traffic control. To our knowledge, there is no significant progress so far apart from Qualcomm's demonstration of rogue base station detection [39]. On the other hand, application-based defenses elaborate on false base station detection [15–17, 29, 30, 33–35, 40], and on malicious SMS detection (e. g., binary, silent,

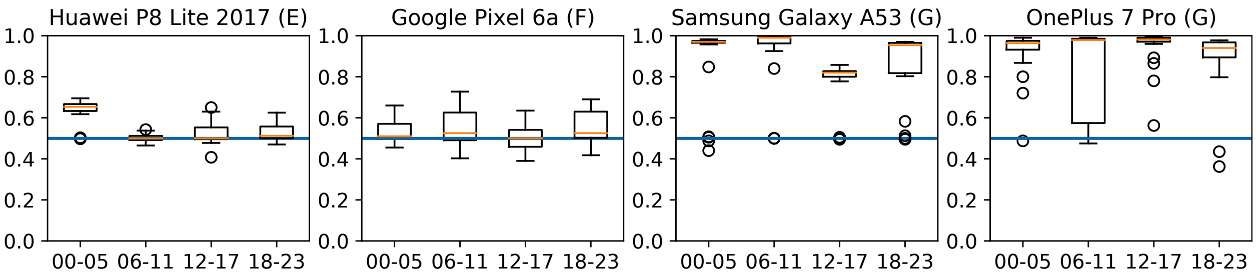


Figure 12: Network analysis of DE-4/NL-4 classification for different time windows and devices (Operator G). The figures show an example of accuracy scores for a certain combination of locations, devices, and operator.

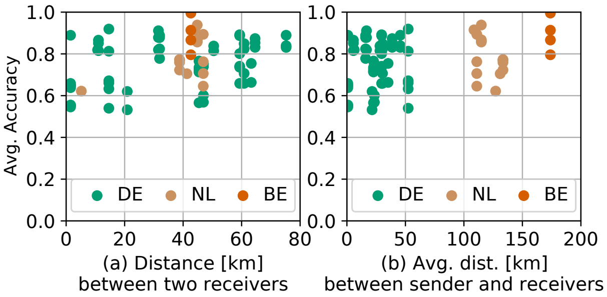
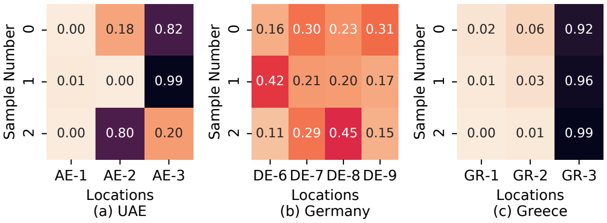


Figure 13: Accuracy of classifications with two locations, depending on (a) distances between both receiver locations and (b) between sender and receivers



Sample Number	AE-1	AE-2	AE-3
0	0.00	0.18	0.82
1	0.01	0.00	0.99
2	0.00	0.80	0.20

Sample Number	DE-6	DE-7	DE-8	DE-9
0	0.16	0.30	0.23	0.31
1	0.42	0.21	0.20	0.17
2	0.11	0.29	0.45	0.15

Sample Number	GR-1	GR-2	GR-3
0	0.02	0.06	0.92
1	0.01	0.03	0.96
2	0.00	0.01	0.99

Figure 14: Probability matrices for fixed locations and areas.

etc.) [13,51,53]. RILDefender [54] expands the SMS attack detection by monitoring the Radio Interface Layer.

Nonetheless, we do not consider that these detection mechanisms are applicable in our case since we do not operate a false base station and do not solely rely on silent SMS. Measurement collection and prediction can happen through regular SMS as well. Therefore, there is currently no actual countermeasure against our timing attacks. Moreover, these approaches have several other drawbacks. They lack preventive countermeasures, which means that the attack has already succeeded by the time the user is potentially alerted. Furthermore, they may rely on the user to manually block potential attacks, while legitimate SMS use cases could be rejected too. Practicality is further decreased as these applications cannot be supported by devices other than Android OS and specific basebands while rooting of the device is required for the application to capture and analyze the traffic. Consequently, the only countermeasures could be to either manipulate the Delivery Reports with a random delay or not send them at all.

Network-based countermeasures. Currently, no countermeasures exist to thwart location identification against a network subscriber. In fact, the network possesses neither the detection nor the prevention mechanisms to hamper or make timing attacks unattainable. However, as a first response, the operators could disable silent SMSs across their network. Although timing attacks are still feasible, the attacker will be forced to use only regular SMSes to collect measurements and interact with the victim, which is less stealthy.

In addition, operators will need to maintain a resilient spamming/flooding filter in the core network, either in the IMS or SMSC, to capture incessant transmissions destined for a specific target. The suspicious communications can either be dropped or intentionally delayed to obstruct the attack. Nevertheless, this approach may significantly impact performance for normal users. As an alternative and more holistic countermeasure, the operators could alter all SMS timings uniformly or randomly to disrupt any side-channel analysis. This could occur during the routing and processing in IMS and SMSC. Once again, this can lead to significant performance degradation which can spread to entire networks.

Finally, a draconian but effective solution would be to eliminate Delivery Reports altogether. Nonetheless, it would necessitate considerable architectural modifications in the core network and smartphone devices (e. g., baseband modems) and re-evaluation of the specifications. Additionally, it is a challenging attempt because it would require worldwide adoption and impede the user experience, network performance testing, and commercial usage (e. g., marketing).

7.2 Limitations

Due to our practical approach involving the same device(s) being physically placed in different geographical locations, we are limited in the amount of data that can be collected within a reasonable time period. However, we consider that the data collected in 34 different locations spread across 10 countries provide sufficient insights to demonstrate the severity of the potential threat.

Our evaluations have demonstrated that the attack works with varying performance, depending on the sets of possible locations of a victim. While classifications can reach high accuracy of 95 % and more, the attack does not work well under *all* circumstances, which is an expected outcome in empirical measurements in the real world. However, as our extensive evaluations at different granularity levels show, performance may not be a matter of distances between locations (cf. Section 6.3). This implies that there are multiple factors contributing to the success of the attack and, thus, this may demand further in-depth evaluations and insights in future work.

While the main part of our evaluations focuses on the performance of the attack in scenarios with distinct sets of known locations of the victim (closed world), cases involving new locations the attacker does not know in advance (open world) might be even more intriguing. Whereas we already peek into this direction in Section 6.4, we argue that the attack working in *closed world* scenarios already poses a significant privacy threat.

8 Related Work

Most related to our research is the work by Schnitzler et al. [47] that explored the feasibility of distinguishing the location of message recipients' in messenger applications. In contrast, our SMS-based timing side-channel has more severe consequences as it relies on a fundamental and universal technology built into every mobile device in the world and cannot be mitigated. Our attack is not limited to specific applications and can remain undetected through silent SMSes.

Several works [21,22,28] attempt to localize the cellular network user actively or passively by capturing identifiers. Shaik et al. [49] (extending on [27]) leverage paging messages and insecure measurement and Radio Link Failure (RLF) reports to reveal GPS locations in certain situations. Other localization approaches capitalize on the MAC layer and timing advance values for localization [38,43]. Ltrack [26] demonstrated improved localization to as much as 20 m with Timing Advance and more passive adversaries, and capitalizing on overshadowing techniques [18, 55]. Lakshmanan et al. [28] demonstrate that sniffers collecting the activation bitmap broadcast in the public scheduling channel and the target's identifiers can identify a path taken by a target among a list of candidate paths, with a scale below 1 Km. The adversary must accumulate sufficient measurements by taking the paths multiple times.

Our work differs in that we do not target the communication channels, use false base stations or sniffers, and are not geographically constrained. Deploying false base stations [24] may prove not only complex in many scenarios but also less stealthy, especially for active attackers that leverage malicious attachments and MitM [11, 12, 44, 45]. Passive attackers may require proximity to the target, especially with 5G SA, where beamforming imposes additional positioning constraints. Therefore, many suggested countermeasures against false base stations ranging from 3GPP studies [1] to PKI mechanisms [23, 50] and detection techniques [15–17, 30, 33–35, 40] are ineffective.

Various SMS attacks have been demonstrated in the past, such as Simjacking [9] that exploits the vulnerabilities of the S@T Browser technology to extract sensitive user information and execute commands. [52] explores spamming, spoofing, DoS, and silent SMS that could impact the LTE network. Furthermore, Mulliner et al. [32] introduced a vulnerability analysis framework that is used to monitor unexpected smartphone behavior leading to large-scale DoS attacks. Apart from SMS, audio call features have been explored [10, 36] for fingerprinting and anomaly detection to detect call redirection/hijacking. Sonar [36] uses the audio latency with the round-trip time due to distance, while PinDr0p [10] leverages the applied codes, packet loss profiles and bit error rates.

9 Conclusion

In this work, we introduced a novel timing side-channel attack for exploiting the SMS procedure, allowing us to distinguish between receivers in locations within a specific region, country, or abroad. We have demonstrated that the SMS procedure leaks timing delays related to the receiver and operator, and we constructed an attack that uses silent SMS to remain stealthy. In addition, we argue that the attack can reach every user who possesses a smartphone device and is subscribed to a network operator. This increases the impact and practicality of the attack as the adversary needs only the victim's phone number to collect measurements from their usual locations of the target. Finally, we clarify that it is hard to enforce countermeasures against timing attacks due to the required architectural modifications, SMS worldwide use, and performance overhead.

Acknowledgements

This work was supported by NSF grant 2144914, by UA Ruhr under the Research Alliance Ruhr program, by the Abu Dhabi Award for Research Excellence 2019 (#AARE19-236), and by the Center for Cyber Security at New York University Abu Dhabi (NYUAD). The authors would like to thank Michel Lang, Philipp Markert, and Lena Schnitzler for their help with data collection.

References

[1] 3GPP. *Technical Specification Group Services and System Aspects Study on 5G Security Enhancement against False Base Stations (FBS) (Release 17)*, 12 2020. Version 0.12.1.

[2] 3GPP. 5G; Non-Access-Stratum (NAS) protocol for 5G System (5GS); Stage 3). Technical Specification (TS) 24.501, 3rd Generation Partnership Project (3GPP), 07 2022. Version 17.7.1.

[3] 3GPP. 5G; Security architecture and procedures for 5G System). Technical Specification (TS) 33.501, 3rd Generation Partnership Project (3GPP), 07 2022. Version 17.6.0.

[4] 3GPP. 5G; System architecture for the 5G System (5GS). Technical Specification (TS) 23.501, 3rd Generation Partnership Project (3GPP), 07 2022. Version 17.5.0.

[5] 3GPP. Digital cellular telecommunications system (Phase 2+)(GSM); Universal Mobile Telecommunications System (UMTS); LTE; 3G security; Access security for IP-based services. Technical Specification (TS) 33.203, 3rd Generation Partnership Project (3GPP), 05 2022. Version 17.1.0.

[6] 3GPP. Digital cellular telecommunications system (Phase 2+)(GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Support of SMS over IP networks; Stage 3. Technical Specification (TS) 24.341, 3rd Generation Partnership Project (3GPP), 05 2022. Version 17.1.0.

[7] 3GPP. Digital cellular telecommunications system (Phase 2+)(GSM); Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Stage 2 . Technical Specification (TS) 23.228, 3rd Generation Partnership Project (3GPP), 05 2022. Version 17.3.0.

[8] 3GPP. Universal Mobile Telecommunications System(UMTS); LTE; 5G; Non-Access-Stratum (NAS) protocol for Evolved Packet System(EPS); Stage 3 . Technical Specification (TS) 24.301, 3rd Generation Partnership Project (3GPP), 07 2022. Version 17.7.0.

[9] Adaptive Mobile Security Limited. Simjacking. `https://f.hubspotusercontent10.net/hubfs/8487362/Reports/AdaptiveMobile_Security_Simjacker_Technical_Paper_v1.01.pdf`.

[10] Vijay A. Balasubramaniyan, Aamir Poonawalla, Mustaque Ahamad, Michael T. Hunter, and Patrick Traynor. Pindr0p: Using single-ended audio features to determine call provenance. In *Proceedings of the 17th ACM Conference on Computer and Communications Security*, CCS '10, page 109–120, New York, NY, USA, 2010. Association for Computing Machinery.

[11] Evangelos Bitsikas and Christina Pöpper. Don't hand it over: Vulnerabilities in the handover procedure of cellular telecommunications. In *Annual Computer Security Applications Conference*, ACSAC '21, page 900–915, New York, NY, USA, 2021. Association for Computing Machinery.

[12] Evangelos Bitsikas and Christina Pöpper. You have been warned: Abusing 5g's warning and emergency systems. In *Proceedings of the 38th Annual Computer Security Applications Conference*, ACSAC '22, page 561–575, New York, NY, USA, 2022. Association for Computing Machinery.

[13] CellularPrivacy. Android IMSI-catcher detector. `https://github.com/CellularPrivacy/Android-IMSI-Catcher-Detector#support`.

[14] Cloudmark. SMS spam overview — preserving the value of SMS texting. `https://www.cloudmark.com/en/resources/white-papers/sms-spam-overview-preserving-value-sms-texting`.

[15] A. Dabrowski, G. Petzl, and E. Weippl. The messenger shoots back: Network operator based imsi catcher detection. In *RAID*, 2016.

[16] Adrian Dabrowski, Nicola Pianta, Thomas Klepp, Martin Mulazzani, and Edgar Weippl. Imsi-catch me if you can: Imsi-catcher-catchers. In *Proceedings of the 30th Annual Computer Security Applications Conference*, ACSAC '14, page 246–255, New York, NY, USA, 2014. Association for Computing Machinery.

[17] Mitziu Echeverria, Zeeshan Ahmed, Bincheng Wang, M. Fareed Arif, Syed Rafiul Hussain, and Omar Chowdhury. Phoenix: Device-centric cellular network protocol monitoring using runtime verification. In *The Network and Distributed System Security Symposium (NDSS)*. Springer, 2021.

[18] Simon Erni, Patrick Leu, Martin Kotuliak, Marc Röschlin, and Srdjan Capkun. Adaptover: Adaptive overshadowing of LTE signals. *ArXiv*, abs/2106.05039, 2021.

[19] Europol. Takedown of sms-based flubot spyware infecting android phones. `https://www.europol.europa.eu/media-press/newsroom/news/takedown-of-sms-based-flubot-spyware-infecting-android-phones`.

[20] GSM Association. Official Document NG.111 - SMS Evolution. Technical Specification (TS) 111-v2.0, GSM Association, 11 2020. Version 2.0.

[21] Byeongdo Hong, Sangwook Bae, and Yongdae Kim. GUTI reallocation demystified: Cellular location tracking with changing temporary identifier. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society, 2018.

[22] Syed Rafiul Hussain, Mitziu Echeverria, Omar Chowdhury, Ninghui Li, and Elisa Bertino. Privacy attacks to the 4G and 5G cellular paging protocols using side channel information. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, California, USA, February 24-27, 2019*. The Internet Society, 2019.

[23] Syed Rafiul Hussain, Mitziu Echeverria, Ankush Singla, Omar Chowdhury, and Elisa Bertino. Insecure connection bootstrapping in cellular networks: The root of all evil. In *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '19, page 1–11, New York, NY, USA, 2019. Association for Computing Machinery.

[24] Roger Piqueras Jover. LTE security, protocol exploits and location tracking experimentation with low-cost software radio. *ArXiv*, abs/1607.05171, 2016.

[25] Kaspersky. What is smishing and how to defend against it. `https://www.kaspersky.com/resource-center/threats/what-is-smishing-and-how-to-defend-against-it`.

[26] Martin Kotuliak, Simon Erni, Patrick Leu, Marc Röschlin, and Srdjan Capkun. LTrack: Stealthy tracking of mobile phones in LTE. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 1291–1306, Boston, MA, August 2022. USENIX Association.

[27] Denis Foo Kune, John Kölndorfer, Nicholas Hopper, and Yongdae Kim. Location leaks over the GSM air interface. In *19th Annual Network and Distributed System Security Symposium, NDSS 2012, San Diego, California, USA, February 5-8, 2012*. The Internet Society, 2012.

[28] Nitya Lakshmanan, Nishant Budhdev, Min Suk Kang, Mun Choon Chan, and Jun Han. A stealthy location identification attack exploiting carrier aggregation in cellular networks. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3899–3916. USENIX Association, August 2021.

[29] Yuanjie Li, Chunyi Peng, Zengwen Yuan, Jiayao Li, Haotian Deng, and Tao Wang. Mobileinsight: Extracting and analyzing cellular network information on smartphones. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, MobiCom '16, pages 202–215, New York, NY, USA, 2016. ACM.

[30] Z. Li, W. Wang, Christo Wilson, Jian Jhen Chen, C. Qian, T. Jung, L. Zhang, K. Liu, Xiangyang Li, and Y. Liu. Fbs-radar: Uncovering fake base stations at scale in the wild. In *NDSS*, 2017.

[31] Matthias Monroy. Significantly more „silent sms“ with german police authorities. `https://digit.site36.net/2019/02/25/significantly-more-silent-sms-with-german-police-authorities/`. Accessed: 2019-02-25.

[32] Collin Mulliner, Nico Golde, and Jean-Pierre Seifert. Sms of death: From analyzing to attacking mobile phones on a large scale. In *USENIX Security Symposium*, 2011.

[33] Prajwol Kumar Nakarmi, Mehmet Akif Ersoy, Elif Ustundag Soykan, and Karl Norrman. Murat: Multi-rat false base station detector. *CoRR*, abs/2102.08780, 2021.

[34] Peter Ney, Ian Smith, Gabriel Cadamuro, and Tadayoshi Kohno. Seaglass: Enabling city-wide imsi-catcher detection. *Proceedings on Privacy Enhancing Technologies*, 2017(3):39 – 56, 01 Jul. 2017.

[35] Shinjo Park, Altaf Shaik, Ravishankar Borgaonkar, Andrew Martin, and Jean-Pierre Seifert. White-stingray: Evaluating IMSI catchers detection applications. In *11th USENIX Workshop on Offensive Technologies (WOOT 17)*, Vancouver, BC, August 2017. USENIX Association.

[36] Christian Peeters, Hadi Abdullah, Nolen Scaife, Jasmine Bowers, Patrick Traynor, Bradley Reaves, and Kevin Butler. Sonar: Detecting ss7 redirection attacks with audio-based distance bounding. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 567–582. IEEE Computer Society, 05 2018.

[37] Christian Peeters, Christopher Patton, Imani N. S. Munyaka, Daniel Olszewski, Thomas Shrimpton, and Patrick Traynor. SMS OTP security (SOS): hardening SMS-based two factor authentication. In *ASIA CCS'22: ACM Asia Conference on Computer and Communications Security, Nagasaki, Japan, 30 May 2022 - 3 June 2022*, pages 2–16. ACM, 2022.

[38] Benjamin A Pimentel. *Passive Geolocation in a 4G WIMAX Single Base Station Scenario*. Phd thesis, Naval Postgraduate School, Monterey California, 2013.

[39] Qualcomm. Qualcomm reveals demo to prevent smartphones from being hacked by connecting to fake base stations. `https://iphonewired.com/news/259588/`.

[40] Cooper Quintin. Detecting fake 4G LTE base stations in real time. In *USENIX*. USENIX Association, Presentation, February 2021.

[41] Bradley Reaves, Nolen Scaife, Dave Tian, Logan Blue, Patrick Traynor, and Kevin R. B. Butler. Sending out an SMS: characterizing the security of the SMS ecosystem with public gateways. In *IEEE Symposium on Security and Privacy, SP 2016, San Jose, CA, USA, May 22-26, 2016*, pages 339–356. IEEE Computer Society, 2016.

[42] Bradley Reaves, Luis Vargas, Nolen Scaife, Dave Tian, Logan Blue, Patrick Traynor, and Kevin R. B. Butler. Characterizing the security of the SMS ecosystem with public gateways. *ACM Trans. Priv. Secur.*, 22(1):2:1–2:31, 2019.

[43] John D. Roth, Murali Tummala, John C. Mceachen, and James W. Scrofani. On location privacy in LTE networks. *IEEE Transactions on Information Forensics and Security*, 12:1358–1368, 2017.

[44] David Rupprecht, Katharina Kohls, Thorsten Holz, and Christina Pöpper. Breaking LTE on layer two. In *IEEE Symposium on Security & Privacy (SP)*. IEEE, May 2019.

[45] David Rupprecht, Katharina Kohls, Thorsten Holz, and Christina Pöpper. IMP4GT: IMPersonation Attacks in 4G NeTworks. In *ISOC Network and Distributed System Security Symposium (NDSS)*. ISOC, February 2020.

[46] Domien Schepers, Aanjhan Ranganathan, and Mathy Vanhoef. Let numbers tell the tale: Measuring security trends in Wi-Fi networks and best practices. In *Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '21, page 100–105, New York, NY, USA, 2021. Association for Computing Machinery.

[47] Theodor Schnitzler, Katharina Kohls, Evangelos Bitsikas, and Christina Pöpper. Hope of Delivery: Extracting User Locations From Mobile Instant Messengers. In *Network and Distributed System Security Symposium*, NDSS '23, San Diego, CA, USA, February 2023. The Internet Society.

[48] Security Affairs. After simjacker, wibattack hacking technique disclosed. billions of users at risk. `https://securityaffairs.co/wordpress/91800/hacking/wibattack-sim-attack.html`.

[49] Altaf Shaik, Jean-Pierre Seifert, Ravishankar Borgaonkar, N. Asokan, and Valtteri Niemi. Practical attacks against privacy and availability in 4G/LTE mobile communication systems. *ArXiv*, abs/1510.07563, 2016.

[50] Ankush Singla, Rouzbeh Behnia, Syed Rafiul Hussain, Attila Yavuz, and Elisa Bertino. Look before you leap: Secure connection bootstrapping for 5G networks to defend against fake basestations. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, ASIA CCS '21, page 501–515, New York, NY, USA, 2021. Association for Computing Machinery.

[51] SRLabs. Snoopsnitch. `https://opensource.srlabs.de/projects/snoopsnitch`.

[52] Guan-Hua Tu, Chi-Yu Li, Chunyi Peng, Yuanjie Li, and Songwu Lu. New security threats caused by IMS-based SMS service in 4G LTE networks. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, page 1118–1130, New York, NY, USA, 2016. Association for Computing Machinery.

[53] Swapnil Udar and Ravishankar Borgaonkar. Darshak. `https://github.com/darshakframework/darshak`.

[54] Haohuang Wen, Phillip Porras, Vinod Yegneswaran, and Zhiqiang Lin. Thwarting smartphone SMS attacks at the radio interface layer. In *30th Annual Network and Distributed System Security Symposium, NDSS 2023, San Diego, California, USA, February 27- March 3, 2023*, 03 2023.

[55] Hojoon Yang, Sangwook Bae, Mincheol Son, Hongil Kim, Song Min Kim, and Yongdae Kim. Hiding in plain signal: Physical signal overshadowing attack on LTE. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 55–72, Santa Clara, CA, August 2019. USENIX Association.

A Neural Network Parameter Tuning

We used manual and automatic parameter tuning. For manual tuning, we mainly experimented with the neural network layers. For automatic parameter tuning we explored the following various setups:

Parameters: [{
hidden_layer_sizes: (10,40,10), (8,8,8), (10,10,10), (8,10,8), (10,50,10), (10,60,10)
activation: (tanh, relu, logistic, identity)
solver: (sgd, adam)
alpha: (0.0001, 0.001, 0.005)
learning_rate: (constant, adaptive)
max_iter: (100, 200, 500, 1000, 2000, 5000)
momentum: (0.2, 0.5, 0.7, 0.9)
},]

B Further Discussion

Can the scalability and feasibility be improved? The process can be automated (including SMS exchange, measurement collection, data processing, training, and prediction) with minimum requirements in terms of equipment. The attacker does not have to be constantly involved in the process, apart from occasional monitoring. Additionally, in our work, we perform this attack with a single device-sender at a time from one location, while multiple devices at the same time could be deployed to expedite the process. It is also possible to increase the SMS exchange rate. Furthermore, the dataset size is not a hurdle for the location identification attack, since silent SMS can be stealthily utilized and the attack works with small and large datasets (Table 9, i.e., 350-15,000 SMS). Hence, the attacker may choose to collect a smaller amount of data for training and prediction.

What if the SMS Delivery Report fails? During the experimentation we very rarely noticed that a Delivery Report failed. As mentioned in Section 4.2, we defined the SMS burst accordingly in order to avoid network congestion, flooding and errors. Furthermore, the slow rates increase the stealthiness of the attack. In case a Delivery Report fails, we ceased the transmission for 10 minutes and then continued without any issue until we collected the complete dataset.

What if the victim is using the device at the moment of the attack? Cellular services typically include calls, SMSs, Internet access (e. g., web browsing, video streaming, social applications, etc.). These activities, especially SIP calls, may culminate in fluctuations in the timings, if the channel is heavily occupied. The attacker can distinguish and filter out such timings since they tend to be statistical outliers.

Table 4: Classification accuracy for pairs of locations in Belgium and Luxembourg.

Receiver Locations	Accuracy
Sender Location: DE-4, Operator E	
BE-1, BE-2	83 %
BE-1, BE-3	80 %
BE-2, BE-3	74 %
LU-1, LU-3	64 %
Sender Location: DE-4, Operator F	
BE-1, BE-2	95 %
BE-1, BE-3	72 %
BE-2, BE-3	80 %
LU-1, LU-3	66 %
Sender Location: DE-4, Operator G	
BE-1, BE-2	86 %
BE-1, BE-3	84 %
BE-2, BE-3	84 %
LU-1, LU-3	72 %

Table 5: Multi-class classification tasks between fixed positions for AE and GR.

Samples	Receiver Locations	Accuracy
Sender Location: AE-1, Operator: A		
300	AE-1, AE-2, AE-3, AE-4	58%
300	AE-1, AE-2, AE-3	76%
Sender Location: GR-1, Operator: C		
300	GR-4, GR-5, GR-6	76%
300	GR-1, GR-2, GR-3	82%

Table 6: Classification accuracy for pairs of fixed locations and areas in the Netherlands.

	NL-2	NL-3	NL-4	NL-5
Sender Location: DE-4, Operator: E				
NL-1		60 %	52 %	
NL-2	–			
NL-3	–	–	52 %	
NL-4	–	–	–	
Sender Location: DE-4, Operator: F				
NL-1		50 %	48 %	
NL-2	–			
NL-3	–	–	54 %	
NL-4	–	–	–	
Sender Location: DE-4, Operator: G				
NL-1	62 %	92 %	49 %	68 %
NL-2	–	98 %	61 %	58 %
NL-3	–	–	88 %	88 %
NL-4	–	–	–	70 %

Table 7: Classification accuracy for pairs of fixed locations and areas in Germany.

	DE-2	DE-3	DE-4	DE-5	DE-10	DE-6	DE-7	DE-8	DE-9
Sender Location: DE-4, Operator: E									
DE-1	74 %		63 %		79 %	63 %			
DE-2	–	77 %	62 %	74 %	65 %	68 %	73 %	60 %	62 %
DE-3	–	–	87 %	76 %	72 %	86 %	44 %	54 %	62 %
DE-4	–	–	–	75 %	67 %	55 %	53 %	64 %	57 %
DE-5	–	–	–	–	72 %	74 %	73 %	77 %	63 %
DE-10	–	–	–	–	–	64 %	61 %	56 %	66 %
DE-6	–	–	–	–	–	–	57 %	62 %	56 %
DE-7	–	–	–	–	–	–	–	63 %	58 %
DE-8	–	–	–	–	–	–	–	–	46 %
Sender Location: DE-4, Operator: F									
DE-1		82 %	58 %	56 %			74 %	74 %	60 %
DE-2	–								
DE-3	–	–	67 %	76 %			60 %	52 %	66 %
DE-4	–	–	–	64 %			67 %	62 %	52 %
DE-5	–	–	–	–			70 %	82 %	68 %
DE-10	–	–	–	–	–				
DE-6	–	–	–	–	–	–			
DE-7	–	–	–	–	–	–	–	62 %	54 %
DE-8	–	–	–	–	–	–	–	–	51 %
Sender Location: DE-4, Operator: G									
DE-1	81 %		78 %	50 %	86 %	74 %	70 %	70 %	68 %
DE-2	–	62 %	56 %	82 %	91 %	64 %	75 %	88 %	78 %
DE-3	–	–	61 %	92 %	82 %	52 %	61 %	77 %	72 %
DE-4	–	–	–	82 %	84 %	68 %	50 %	84 %	74 %
DE-5	–	–	–	–	86 %	81 %	72 %	81 %	70 %
DE-10	–	–	–	–	–	76 %	83 %	88 %	84 %
DE-6	–	–	–	–	–	–	58 %	72 %	58 %
DE-7	–	–	–	–	–	–	–	58 %	62 %
DE-8	–	–	–	–	–	–	–	–	52 %

Table 8: Device Specifications. Except from Google Pixel 4 XL which used eSIM, all devices were equipped with physical SIM cards. The attack worked on all tested smartphones.

Device	Modem	OS	Model	Release
Apple iPhone 13	Qualcomm Snapdragon X60	iOS 15	A2633	2021
One Plus Nord 2 5G	MediaTek Dimensity 1200 5G	Android 11	DN2101	2021
Alcatel 1S	Spreadtrum UNISOC SC9863	Android 11	6025D	2021
Apple iPhone 12 mini	Qualcomm X55 modem	iOS 15	A2399	2020
Nokia 8.3 5G	Snapdragon 765G 5G	Android 10	TA-1243	2020
Apple iPhone 12	Qualcomm Snapdragon X55	iOS 15	A2403	2020
Samsung Galaxy A21S	Samsung Exynos 850	Android 10	SM-A217F	2020
Huawei P40 Pro 5G*	HiSilicon Kirin 990 5G	Android 10	ELS-NX9	2020
Nokia 5.3	Qualcomm Snapdragon 665	Android 11	TA-1234	2020
Google Pixel 4 XL	Qualcomm Snapdragon X24	Android 12	G020J	2019
OnePlus 7 Pro	Qualcomm Snapdragon 855	Android 11	GM1910	2019
Google Pixel 3a	Qualcomm Snapdragon 670	Android 11	G020F	2019
Samsung Note 10 5G	Samsung Exynos 9825	Android 10	SM-N976Q	2018
Huawei P8 Lite 2017	HiSilicon Kirin 655	Android 10	PRA-LA1	2017
Apple iPhone 7	Intel XMM7360	iOS 15	A1778	2016
Apple iPhone 5	Qualcomm MDM9615M	iOS 10	A1428	2013

Table 9: Number of SMS received in our experiments. The * denotes a sender device only.

Device	Countries and Operators											
Int'l, Nat. & Reg.		AE		GR		UK	US	DE	DK			
(Sections 5.1 & 5.2)	A	B	C	C	D	H	J	E	I			
Apple iPhone 13	0	0	0	350	0	0	0	0	0			
One Plus Nord 2 5G	350	350	0	0	0	0	0	0	0			
Alcatel 1S	0	0	0	350	0	0	0	0	0			
Apple iPhone 12 mini	350	350	0	0	0	0	0	0	0			
Nokia 8.3 5G	350	350	0	0	0	0	0	0	0			
Apple iPhone 12	0	0	0	350	350	350	0	0	0			
Samsung Galaxy A21S	0	0	0	350	0	0	0	0	0			
Huawei P40 Pro 5G*	0	0	0	0	0	0	0	0	0			
Google Pixel 4 XL	0	0	0	0	0	0	350	0	0			
Samsung Note 10 5G	350	350	0	0	0	0	0	0	0			
Apple iPhone 7	0	0	0	700	0	0	0	0	350			
Apple iPhone 5	0	0	350	0	0	0	0	0	0			
OnePlus 7 Pro	0	0	0	0	0	0	0	350	0			
Nat. & Reg.		BE			DE			LU			NL	
(Section 5.2)	E	F	G	E	F	G	E	F	G	E	F	G
Nokia 5.3	0	0	798	1350	0	3159	0	0	455	0	0	1419
OnePlus 7 Pro	0	0	839	2021	0	3109	0	0	422	0	0	1411
Google Pixel 3a	0	818	0	1963	2516	1092	0	499	0	0	1399	0
Huawei P8 Lite 2017	804	0	0	3342	1111	1153	513	0	0	1416	0	0
Temporal & Network					DE						NL	
(Sections 6.1 & 6.2)				E	F	G				E	F	G
Huawei P8 Lite 2017				14132	0	0				15607	0	0
OnePlus 7 Pro				0	0	16625				0	0	12799
Google Pixel 6a				0	14752	0				0	16115	0
Samsung Galaxy A53				0	0	11095				0	0	15778