# The attacks aren't alright: Large-Scale Simulation of Fake Base Station Attacks and Detections

Thijs Heijligenberg
thijs.heijligenberg@ru.nl
Radboud University
Nijmegen, The Netherlands

David Rupprecht
david@radix-security.com
Radix Security
Bochum, Germany

Katharina Kohls
katharina.kohls@rub.de
Ruhr University
Bochum, Germany

## ABSTRACT

Fake base stations are a well-known threat to pre-5G mobile networks and are one of the most common primitives for mobile attacks that are used in the real world. However, despite years of research we only have limited knowledge about their performance spectrum and how well detection mechanisms work in practice. Consequently, mobile network operators and vendors struggle to identify, implement, and deploy a practical solution in the form of detection mechanisms. For the first time, we systematically study fake base station attacks and their main influencing factors. We use a specification-conform simulation model that lets us analyze fake base station attacks on a large scale, and test detection mechanisms on the generated data. The simulation environment allows us to test diverse scenarios with a large measure of control and insight, while providing realism in the aspects that matter. We study detection mechanisms from academic work and ongoing 3GPP discussions. Our experiments reveal the influencing factors of the success of fake base station attacks and detection, and provides nuances for performance that is missing from existing work.

**ACM Reference Format:**
Thijs Heijligenberg, David Rupprecht, and Katharina Kohls. 2024. The attacks aren't alright: Large-Scale Simulation of Fake Base Station Attacks and Detections. In *Workshop on Cyber Security Experimentation and Test (CSET 2024), August 13, 2024, Philadelphia, PA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3675741.3675742

## 1 INTRODUCTION

Fake base stations have far-reaching privacy and security implications for the users of pre-5G mobile networks. They pose a real-world threat in countries like China, India, or the US, where they generate a remarkable daily revenue with SMS spam [23, 24]. Other attacks use fake base stations to localize and track victims by learning the users' identities [18–20]. Just lately, fake base stations were also considered as a way to deploy the Pegasus malware, compromising a victims' phone [12]. Fake base station attacks will be around as long as pre-5G networks are in use, which is likely to be for the foreseeable future. With no solution in sight there remains a need for evaluation of this threat.

Although fake base stations are well-studied and also occur in real-world settings, the literature is focused on feasibility studies in a lab setup. The behavior of fake base stations in the wild, and how users interact with them in the presence of large-scale network infrastructure in the real-world, is not described in academic literature. This is mainly due to practical and ethical limitations in testing. This exposes an important gap in the current understanding of the actual threat posed by fake base stations.

In response to the threat of fake base stations, academia, industry, and the 3GPP have developed and discussed approaches for detection. They can be classified into sensor-based [11, 15], app-based [22, 24], and network-based [6, 9, 14] mechanisms. Sensor-based detection runs on dedicated hardware and requires the deployment of an *additional* infrastructure. App-based mechanisms run on the end devices of users and, therefore, do not require any network-side changes or additional infrastructure. While such an easy setup facilitates crowd-sourced detection, sensor and app-based detections are limited by the amount of information a *single* sensor or phone can gather, which makes detection challenging and error-prone [18]. These approaches have seen limited adoption. On the other side, network-based approaches have a broader view of the infrastructure and potential incidents. While this contributes to a more informed state and possibly higher detection rates, such mechanisms need to be *integrated* into the existing infrastructure. There is no data on the adoption of these methods.

Existing defense concepts have in common that they often assume a specific attack aim and defined technical capabilities at the fake base station. However, we see a wide diversity of these assumptions throughout the literature, reflecting the diversity of attacks. For example, Zhang et al.build an on-phone SMS-spam detection mechanism assuming an attack involves sending SMS-spam [24]. At the same time, similar work in the same context assumes an attack aim focused on identifying a victim rather than sending out spam messages [9]. Consequently, we learn more about the performance of *specialized* countermeasures but have a remaining blind spot when it comes to the effectiveness against the *basic functioning* of fake base station attacks. Due to the diverse nature of real-world attacks based on fake base stations it is paramount to study defenses at this most basic level.

Likewise, we only have limited knowledge about the performance spectrum of fake base station attacks. Experimental setups can demonstrate feasibility of specific attacks in a small setup. However, no existing work focuses on the performance of attacks in a large-scale setup. Improving the understanding of success chance of fake base station attack strategies and their contributing factors is essential to properly evaluate the threat posed by fake base stations and gain full insight into performance of detection mechanism.

To gain insight into the performance of fake base station attacks we propose a simulation framework. This enables a large-scale

This work is licensed under a Creative Commons Attribution International 4.0 License.

*CSET 2024, August 13, 2024, Philadelphia, PA, USA*
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0957-9/24/08
https://doi.org/10.1145/3675741.3675742

evaluation of attack characteristics. We model user behaviors, network components, and attackers that interact during a specification-conform simulation. This allows us to observe the large-scale performance of attacks and identify previously unknown challenges and limitations. Our simulation framework is available at [1].

Following the evaluation of fake base station attack performance we analyze a set of network-based defensive approaches that are currently under discussion in the 3GPP [4] or studied in prior work [9, 15]. We analyze their performance in our large-scale simulation settings and identify limitations that affect the detection capabilities of a defense. Both steps of our experimental evaluation reveal insights into the performance of attacks and defenses, including challenges for attackers and limitations on detection mechanisms in all circumstances.

With our work, we provide a first measure of the impact of fake base station attacks and defenses in a large-scale setting. Our new insights contribute to the ongoing discussion of practical detection mechanisms and help to overcome previous misconceptions regarding the effectivity of fake base stations. To this end, we make the following contributions.

- **Technical Simulation Model.** We provide a specialized simulation model conforming to mobile network specifications that complements existing frameworks. We share this model to contribute to future research.
- **Attack and Detection Performance.** We use the simulation model to analyze the performance of fake base stations as a crucial building block for over-the-air attacks. Based on these findings we analyze the performance of various methods of fake base station detection.
- **Real-World Implications.** We discuss the effectivity of fake base station attacks and defense mechanisms together with considerations concerning real-world implementation. We hope to contribute to the ongoing discussions and enable operators to make a balanced decision about fake base station detection.

## 2 TECHNICAL BACKGROUND

In a 4G mobile network, a User Equipment (UE) establishes a connection to the core network (Evolved Packet Core (EPC)) through the Radio Access Network (RAN). The RAN consists of base stations (Evolved NodeB (eNodeB)) that provide connectivity at a certain frequency and within the covered cell. In the context of fake base stations, we are mainly interested in the messages exchanged on the Non-Access Stratum (NAS) layer between the UE and the EPC, and the Radio Resource Control (RRC) layer between the UE and the eNodeB. In its most basic form, a *fake base station* is an eNodeB that is not part of the legitimate network while still establishing some sort of connection with a UE.

### 2.1 Idle procedures

For a fake base station to succeed the UE must decide to switch from a legitimate eNodeB to the fake base station. This can only happen when the UE is *idle*. In the idle state the UE has established a connection and security context with the network previously, but is not currently actively communicating. There are four characteristics that are particularly relevant for this transition.

**SIB**: SIBs are periodic broadcast messages that spread information for cell selection and reselection. This includes the frequencies on which the UE should search for candidate cells and the rules for selecting cells, such as priorities for each frequency. Both characteristics have an important effect on fake base station attacks, as they influence whether the UE recognizes and selects the attacker's base station.

**Cell Selection**: The cell selection procedure defines how a UE finds the best suitable cell in its current location when it has no network registration. For each frequency that it supports the UE will find cells matching its network provider, and for each it will compute the quality level. The quality level depends on the physical-layer PHY characteristics and the information from the System Information Block (SIB). The UE will connect to the first suitable cell it finds.

**Idle Cell Reselection**: When the UE camps on a cell in the idle state, it will periodically re-evaluate other cells to select a better candidate if possible. Such updates compensate for varying connection characteristics, e. g., caused by users moving around. The procedure relies on a frequency's *priority*, which is an integer set in a SIB message. For the reselection, the UE accomplishes the following checks:

(1) Search for cells on frequencies at a higher priority than the current frequency's priority.
(2) If no higher-priority cells are found, and the current signal is below a certain threshold, the UE will switch to the best intra-frequency cells or equal-priority cell ranked by to signal strength or quality.
(3) If the above does not succeed and the current signal is bad, search on frequencies of lower priority.

**Tracking Area Updates**: A Tracking Area (TA) combines multiple neighboring cells in the network infrastructure, with TA codes sent in a SIB. TAs are used to estimate the whereabouts of idle users within the infrastructure which is used for, e. g., targeted paging (§2.2). To keep the network up-to-date, the UE sends a Tracking Area Update (TAU) that informs the network about its updated TA. TA updates trigger the establishment of an active connection, which can lead to a message exchange that a fake base station can exploit.

### 2.2 Active state

In the active state, the UE is much less vulnerable to fake base station attacks, as in this state the network decides what cell the UE connects to. To make these decisions the network request *Measurement Reports* from the UE (§2.2).

**Measurement Reports** Measurement reports convey information about the network that surrounds a UE [7]. The UE receives instructions on what, how, and when to measure using the physical layer PHY. The UE shares this information with the eNodeB that it currently connects to. Among other items, a measurement report can document the signal strength of all cells in its vicinity transmitting on a target frequency band. Measurement reports are either sent due to an event, such as signal strenghth crossing a boundary, or sent periodically. It is up to the operator to configure this in such a way that there is enough data to make correct handover decisions and ensure adequate connectivity. Real-world data suggests that

configurations vary a lot, with measurement reports being sent roughly every few minutes.

**Switching to connected state** There are two ways in which the UE can begin establishing an active connection with the core network. One option is for the UE itself to decide to do this. This happens regularly with modern handheld devices, where background traffic initiates an active connection. The other option is for the network to initiate the connection, for example when there is a voice call or message for the UE. In this case, the network broadcasts a *paging request* to the tracking area the UE is registered to; the UE then initializes network connection to receive data.

# 3 FAKE BASE STATION ATTACKS AND DEFENSES

A number of mobile network attacks require the use of a fake base station to connect with the user. This connection can then be abused for tracking and identification in the case of an International Mobile Subscriber Identity (IMSI)-catcher [10]. Other uses include, for example, SMS spamming [24]. To evaluate the success rate of the fake base station attack we must first understand its goals and what factors influence these goals. After this we classify the different methods which detect the fake base station.

## 3.1 Fake base station

A fake base stations is an eNodeB that is operated by the attacker with the intention to establish a connection with UEs that are serviced by a legitimate network operator in the area. This allows the fake base station operator to send unprotected messages, which can for example be used to request the permanent identifier IMSI for tracking purposes, but does not allow for setting up a data connection. We assume the fake base station uses a fresh Tracking Area Code (TAC) to force the UE to perform a TAU procedure and thus initiate a connection (see § 2.1). Fake base stations vary in the following ways:

- **Physical properties:** an important feature of a base station is the transmission power of its signal. For a fake base station this influences the area in which UEs will see the fake base stations as a viable cell. Besides this the placement of the fake base station is important; UEs will not evaluate other base stations when they are close to a legitimate base station.
- **Cell properties:** the frequency and cell number the fake base station operates on influences how UEs evaluate it when connected to the legitimate network. Since frequency scanning is expensive the fake base station is restricted to frequencies indicated by the legitimate network. The way these frequencies are configured by the legitimate network influence how the fake base station gets evaluated by UEs.
- **Activity:** fake base station attack success rate is equivalent to the chance that an UE connects, which is not influenced by the fake base station's activities. Defenses, however, can make use of the active communication between the UE and the fake base station to detect anomalous behaviour.

## 3.2 Defenses

Since fake base stations use specification-compliant behaviour it is not possible to prevent them using the current infrastructure. However, using encryption and integrity-protection prevents Man-in-the-middle-based attacks using fake base stations which rely on unprotected communication [21]. The introduction of 5G also introduced a prevention for the IMSI-catcher, which is the most prevalent fake base station-based attack. This is done by preventing the cleartext transmission of IMSIs, which invalidates this attack [17].

Since pre-5G networks will be around for the foreseeable future, the best possible option to mitigate these attacks is to detect them. This can be done in a variety of ways:

- The network operator can deploy **sensors** that scan frequencies for new and anomalous cells.
- **Users** can use their UE to detect anomalous behaviour of the base stations they interact with. This can also be aggregated by a third party.
- The **network** can use the data that it handles for the connection of UEs to detect anomalous patterns.

In this paper we are primarily interested in the network-based detection. Network-based detection does not require additional infrastructure in contrast to sensor-based detection, which makes it more attractive for network operators. User-based detection, as the alternative, is only beneficial over network-based detection if there is data that the UE records that the network does not also receive. In the case of general fake base stations the most important data is the measurement of cells in the UE's neighbourhood, which the network receives through measurement reports (see § 2.2). Users, or third parties aggregating detection data, also do not always have access to ground truth about base station placement, which we can assume the network to have.

# 4 SIMULATION MODEL

Our simulation model implements components of a 4G mobile network and focuses on the most relevant characteristics related to fake base station attacks and their detection. We first introduce these components and define their functionality according to the technical specification of 4G. We then present the parameters of the simulation model.

## 4.1 High-Level Architecture

In our model, we define static characteristics like the underlying map, the distribution of base stations, and the signal model, as well as dynamic features including the movement of users, the communication between components, and an adversary in different attack scenarios.

*4.1.1 Static Characteristics.* The static characteristics represent features of the simulation that are fixed for each scenario.

- **Map.** We use a square map with a mostly-filled grid of spaced-out square buildings (cf. Appendix 8). This layout represents a generic city with building blocks and streets.
- **Base Stations.** We place base stations in strategic locations on the map such that all of the map has coverage. Tooling and techniques exist for performing this in practice but in our simple maps this is straightforward.

- **Signal Model.** We use a raytracing model to approximate the signal model that incorporates interaction (reflections, shadows) with the buildings on our map. The basis for this is the formula for signal loss based on a distance $d$:

$$L(d) = 20 \log_{10} \left( \frac{4\pi d}{w} \right) \quad (1)$$

  Here, $w$ refers to the wavelength of the signal being simulated. Signal strength is determined by the base station's transmission power, the free-space loss according to the travel distance of a ray, and any signal loss incurred by buildings the ray passes or bounces off. Each station also possesses a small number of beams in which its signal is stronger, mimicking a commercial setup with multiple directional antennas.

*4.1.2 Dynamic Characteristics.* The dynamic characteristics of our model represent features that change during a simulation run.

- **Mobility.** Users follow random trajectories along the streets of the map. Users move independently.
- **Communication.** We simulate two active components of a mobile network, namely, the users (UEs) and the infrastructure (EPC and eNodeB). The core network in our simulation performs the tasks of the Mobility Management Entity (MME) and the RRC component of the eNodeB. We do not simulate physical-layer communication, but rather focus on the abstract interaction between components.
- **Attacker.** Similar to the benign base stations, we configure and control the malicious eNodeB through its core network components. The attacker can be toggled to simulate the moment when a fake base station appears/disappears.

## 4.2 Specification Conformity

To achieve realism, our simulation model conforms to the Long Term Evolution (LTE) specification (cf. Appendix 2). Note that these procedures are highly similar in 3G and 5G.

- **Phy.** While we do not simulate actual transmissions, our model focuses on the received signal strength according to our signal model.
- **SIB.** We model a subset of SIB messages necessary for distributing the values needed for the Cell selection and reselection procedures.
- **Cell Selection and Reselection.** We implement *all* primary features of the cell (re)selection in idle mode to imitate how a user (re-)evaluates available cells and picks one of them for the primary connection.
- **RRC Procedures.** For our detection techniques, we use RRC characteristics, e. g., by using measurement reports for a signal-based anomaly detection.
- **NAS Procedures.** The attach and detach procedures are essential for the connection establishment; The TAU mechanism is part of one of the detection techniques in which the network recognizes anomalous update behavior.

## 4.3 Model Parameters

We divide the most relevant parameters of our simulation model into the three groups *Execution*, *Network*, and *Attacker*. The used values for all experiments can be found in the appendix in Table 1.

*4.3.1 Execution.* The *Execution* parameters summarize the general setup of a simulation run. Within this we count the *map size* and *run time*, which are constant. We do vary the *base station* and *user* positions; base stations are placed as described above, and users are distributed randomly. We randomize each run of the simulation using a seed. This allows us to reproduce specific runs while achieving diverse random repetitions overall.

*4.3.2 Network.* The *Network* parameters define the capabilities of the legitimate base stations. Each eNodeB in the network receives its own configuration including the following parameters. For each base station we set the *frequency* and its *transmission power*. Frequencies have a global *priority*.

These parameters, together with base station placement, are set to replicate realistic scenarios. Our simulation offers the opportunity to choose unrealistic parameters to test scenarios that would be unachievable in real life, which could be interesting for future work.

*4.3.3 Attacker.* An attacker deploys a fake base station with the same standard characteristics as a legitimate eNodeB in our model, which we will assume is synonymous with the attacker. We only run experiments with one attacker at a time. An attacker has the same defining parameters as a legitimate eNodeB, cf. 4.3.2. The attacker's position is studied in Section 5.3. We will study the priority of an attacker in Section 5.4. The attacker's signal strength, which we study in relation to the legitimate eNodeBs, is studied in Section 5.5.

## 4.4 Model logic validation

To verify aspects of the simulation model, we conducted practical experiments using commercial equipment. We verified the cell reselection model, as this is vital to our experiments.

To study the reselection procedure, we used an Amarisoft EPC with corresponding eNodeB as the cell the UE is camping on. A second eNodeB based on `srsran` acts as the reselection target. The UE is provided by a Oneplus Nord 2 5G and Huawei P40 Lite 5G. We placed the setup inside a shielding box. We changed the gain of the serving eNodeB at runtime to mimic a change in signal reception. We determined the lowest gain value at which reselection occurs through the bisection method. This gain value is translated to the Reference Signal Received Power (RSRP) based on the value reported by the android system which is calibrated to the base station gain.

The details of the performed test cases can be found in Table 3. We see that commercial the UEs do follow the specifications, although the Huawei UE did not perform intra-frequency reselection as expected in our setup. This would be of limited impact on our results.

This validates that commercial equipment uses the reselection algorithm from the specification. Similar test cases were made for the model code, which validates the logic used to implement the algorithm in the model.

# 5 EXPERIMENTS: ATTACKS

In the first phase of experiments, we analyze what factors influence the performance of fake base stations. We first describe the experimental setup of our simulation model and introduce the metrics that enable us to assess the potential success of an attack in a large-scale setting. Within this model, we analyze key strategical aspects of an attack. We use our experiments to estimate a fake base station's performance range and dependencies in a large-scale environment.

## 5.1 Experimental Setup

In our experiments, we have a fixed set of benign base stations on a map of 2 km × 2 km. Each simulation covers 2000 discrete steps and involves 500 independent users, each representing one random repetition. The benign base stations operate on a fixed setup and in the same locations on the map. We focus on three strategic aspects that affect the success of an attack:

(1) Placement of the fake base station
(2) Frequency priorities
(3) Offered signal strength

As these strategic aspects affect each other, we guide through them step-by-step and always fix two of them while inspecting the effects of varying the third parameter. Our goal is to gain insights into the performance spectrum of fake base stations and to identify an optimal strategy within our network setup.

## 5.2 Success metric

We define the *attack success* as the relative number of users that connect to the fake base station at least once during the runtime. Please note that our metric cannot predict the arbitrary success of a real-world attack, as it is purely focused on an abstract assessment, but is useful for comparing different parameter sets for the same experiment.

## 5.3 Impact of Placement

Within the three strategic aspects of our experiments, the adversary has the highest degree of freedom for the *placement* of a fake base station. It influences how close any potential victims are to the malicious station and how strong the benign stations' signals are in the region it targets.

Given a fixed setup of benign base stations, an adversary can estimate the signal strength of all stations in reach. This allows the attacker to estimate the optimal placement locations for the attack. The quality of a location mainly depends on its distance to the benign base stations and the surroundings (buildings and other obstacles). This becomes visible when relating the attack success to the distance to the closest benign station (cf. Figure 1). In this experiment, the trend of success probability ranges from around 7% to 13%, which we remind the reader does not indicate success chance of a real attack. It does allow us to conclude that moving 100 m further away from the legitimate base station gives an 8.8% success rate increase.

**Conclusion.** The placement of a fake base station is an essential factor for its effectiveness. Our experiments show that, on average, a position further away from the legitimate network is almost twice as effective. The absolute best position, however, is not simply the furthest away. Far-away positions are rare in a well-distributed network.

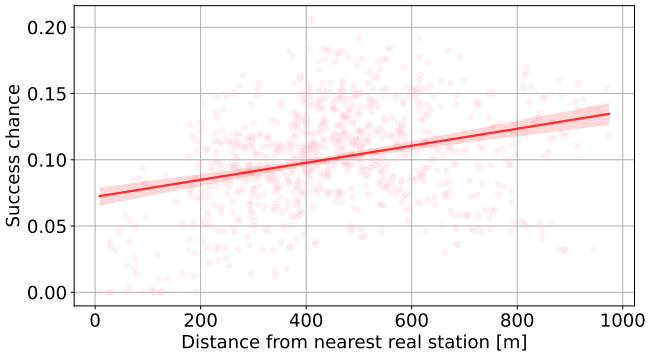


**Figure 1: Effect of fake base station placement on attack success chance. We summarize the attacker's choice of position by the minimum distance to any base station belonging to the legimate network.**

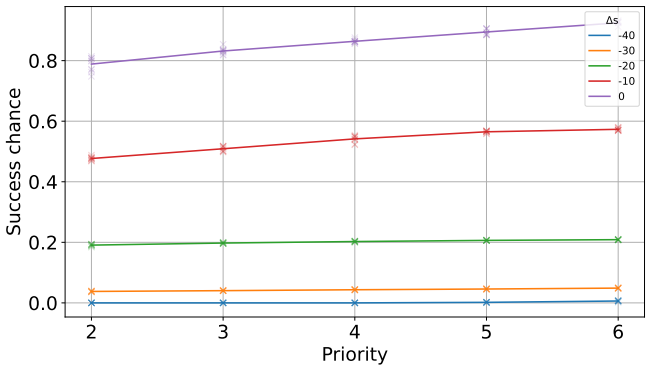


**Figure 2: Success chance of fake base station attacks where the attacker uses an otherwise unused frequency. We show this success chance for different values of the fake base station strength.**

## 5.4 Impact of Priority

We run our experiments in a setting where the legitimate base stations occupy two frequencies of different priorities (3 and 5 respectively). We then study the success rate of an attacker that operate on a different frequency for which we vary the priority:

- Lower than the surrounding (2)
- Equal to the low-priority frequency (3)
- Between the other priorities (4)
- Equal to the high-priority frequency (5)
- A higher priority than the surrounding (6)

Values above or below the given values do not alter the effect of the priority ranking. Figure 2 shows the impact of the priority of the attacker's frequency on the attack's success. As signal strength is a critical influencing factor to the attack's success, we compare the success of different priorities for individual signal strengths. We

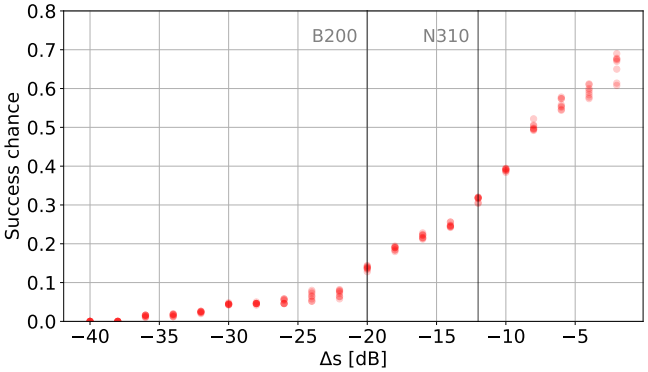


**Figure 3: The impact of the signal strength on fake base station attack success rate.**

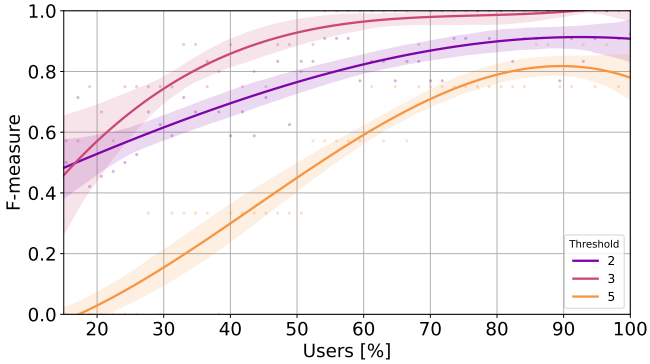


**Figure 4: Polynomial interpolation of F-measure for TAU-based detection for different percentages of users sending dummy TAC. Colors indicate an increasing detection threshold. We show the lowest, highest, and optimal (3).**

observe that impact of priorities is limited compared to the other factors, giving a relative increase of up to 10% between the most extreme scenarios.

**Conclusion.** The choice of frequency for a fake base station is, in terms of the priority that it provides, the relatively least important factor in our experiments since the performance range is the most narrow. While our experiments show an advantage to occupying a higher-priority frequency, this is, in practice, not always possible due to a lack of spectrum space or capable hardware.

## 5.5 Impact of Signal Strength

A significant factor for the success of an attack is the power at which the attacker transmits relative to the surrounding base stations, as shown in Figure 3. Our experiments show that transmitting with a lower power can still yield a successful attack but with a significant performance loss compared to a signal strength similar to the benign station. Mobile network operators usually operate base stations with a strength of around 30 dBm as seen in public base station databases for different countries (cf. [8]). Commercially available antenna equipment fit for use in a eNodeB is less powerful, e. g., the Ettus USRP B200 achieves 10 dBm while the more sophisticated USRP N310 will output up to 18 dBm [16]. According to our evaluation, these values yield effective attacks against users 200 m to 400 m away.

**Conclusion.** Signal strength is the most significant influencing factor in our experiments. Attackers with simple off-the-shelf hardware can already be relatively successful but generally only attain a small fraction of the success of attackers with signal strengths closer to that of the legitimate network's eNodeBs.

Overall, our experiments provides nuance to previous work where a fake base station is just assumed to work. We see that attacks in general are possible, but the wide range of performance in different scenarios indicate that luring a victim into the malicious connection cannot be achieved by default.

# 6 EXPERIMENTS: ATTACK DETECTION

After assessing the performance spectrum of attacks we continue with an evaluation of detection mechanisms, which with the lasting prevalence of pre-5G systems are the only mitigation for fake base stations. In particular, we focus on the two network-based approaches, for which we extend the results of prior work and ongoing discussions by a large-scale performance analysis.

## 6.1 Attack Setup

As a baseline for our defense evaluation, we use a parameter setup that is representative of a capable attacker, as introduced in Section 5 (see Table 1 in the appendix for details). In particular, the attack uses a distance-based placement heuristic, a frequency and thus priority used by surrounding cells, and a signal strength representative of relatively low-cost off-the-shelf equipment (cf. Table 1). The fake base station is active in the second half of the simulation time, which provides us with ground truth information about the detection results. Based on this ground truth, we define the following detection decisions:

- **True Positive TP.** Correct detection of an attack.
- **False Positive FP.** Incorrect detection of a non-existent attack.
- **True Negative TN.** Correct missing detection due to the attack absence.
- **False Negative FN.** Incorrect miss of an ongoing attack.

For the deployment of a detection mechanism, we define **TP** incidents as the primary detection priority. We initially investigate how a mechanism can achieve a reliable detection rate. We then focus on adjusting parameters to improve the trade-off with other aspects of the detection, e. g., reducing the collateral damage through **FP** detections.

## 6.2 TAU Detection

It is often possible to detect fake base station attacks on the NAS layer when monitoring Tracking Area Update (TAU) messages [9]. More precisely, a UE affected by an attack reconnects to the benign network once the connection with the attacker inevitably fails. As part of this process, it sends a TAU containing a *dummy* TA code for the previous connection. The detection is particularly successful for *many* users but is limited to detecting only successful attacks.

*6.2.1 Parameters.* There are two potential limitations to the TAU detection. First, it has been observed in older generations that some messages contain a dummy TA in practice even without an adversary being present [9]. Second, not all UEs send a TAU after an active fake base station attack. We consider three additional characteristics in our simulation model to analyze the performance of the TAU detection.

(1) Do UEs send a TAU after an attack?
(2) Do UEs send a dummy TA in normal operation?
(3) How many TAUs do UEs send?

In our experiments, we focus on the number of UEs that send a TAU (first item). This rate depends on the baseband implementations of devices. According to prior work, we set the number of UEs that send a dummy TA to 13 % [9] and set the number of TAUs sent to a constant in all experiments.

*6.2.2 Detection Mechanism.* The detection mechanism uses an anomaly metric that identifies suspicious TAU messages in the network. To this end, we derive a histogram from the amount of TAU messages with a dummy TA code recorded over time and then identify values that exceed a threshold multiple of the average. An example scenario can be seen in Figure 6.

*6.2.3 Results.* Figure 4 shows the success rates of the detection mechanism. We deem the detection successful if it occurs within 30 s of the attack as this should give time to complete the attack; any other detections are false positives. The threshold values we test range between 2 and 5, with the best score achieved by a value around 3. The threshold that is used in the procedure is the threshold value multiplied by the average. This shows the trade-off of this detection method: lower thresholds will detect more attacks and lead to more false positives. In comparison, higher thresholds have few false positives at the cost of a lower detection rate. With a threshold of 3, results above 0.9 can be achieved when over 50 % of users send a dummy TA code after an attack. The data used by [9] showed only 16 % of users sending a dummy TA, at which point the detection method is much less accurate. We further observe that, with a small fraction of UEs sending a TAU, the detection is overall less reliable.

*6.2.4 Conclusion.* To reach a reliable detection rate, a high percentage of users must send a TAU. At least for 2G and 3G, this was not the case according to [9]. A real-world empirical study is necessary to assess the current state of this phenomenon in baseband implementations.

## 6.3 Measurement Report Detection

Measurement reports contain dynamic information about the network status (§ 2.2). Such information can be used to detect anomalies in the infrastructure [14] or to localize a fake base station [13]. In our experiments, we benefit from the scalability of the simulation model and focus on a specific variant of anomaly detection that incorporates spoofed cell identities [4]. Note that detection when non-spoofed cell identities are used is relatively trivial.

*6.3.1 Parameters.* To simulate this detection method, we use the following setting:

(1) The network operator has access to a coarse signal strength distribution, i. e., it knows the signal strength for points on a grid. We vary the granularity of this grid.
(2) The attacker spoofs the identity of a cell that is close by. In reality, it could also choose a cell much further away, making sure that there is no overlap in coverage. This contributes to the attack performance, but also leads to clearer anomalies and a clearer detection.

In our network setup, UEs are configured to send measurement reports on events that can indicate the need for a handover, and upon significant increase or decrease in serving signal strength, see 2.2. Configurations in which more measurement reports are sent are possible, but this does incur some overhead.

We assume that when the user detects signals for the same cell identity from multiple sources it reports the strongest. This is not made exact in the specifications.

*6.3.2 Detection Mechanism.* In the case of a spoofed cell ID, the adversary uses the identity of a benign cell in the network. During an attack, the measurement report sometimes contains the attacker's signal alongside legitimate cells. This can be detected if the reported signal strengths contradict expected values. An example of this can be seen in Figure 7.

In our simulation, the network operator compares the data from a single measurement report against all points in its granular view of the signal strength map. It tries to find a location where a maximal number of signal strength entries in the measurement report are within some error margin of those in its signal map at a given point. It reports an anomaly if it is impossible to find a sufficiently matching point. Detection can then be performed using either the number of anomalies occurring or further data extracted from the anomalous report.

*6.3.3 Results.* In Figure 5, we present the results of detection based on the number of observed anomalies. It shows for different threshold values what proportion of experiments resulted in a number of anomalies greater than the threshold. The threshold values used for the detection decision depend on the length of the period the measurement reports are analyzed and the number of measurement reports sent. In our simulation the configuration with the best detection (true-positive) rate while having negligible false-positive rate only detects 15% of attacks. This means the detection chance is relatively low, but with enough data this can still be reliable.

*6.3.4 Conclusion.* Different influencing factors challenge the reliable detection of measurement report-based approach. Signals are noisy and cannot be predicted in a perfectly reliable way, which can lead to false positive detections.

Furthermore, the amount of active UEs and the resulting number of measurement reports influence how much information a network can use. Overall, we conclude that the measurement reports can be only used for attack detection in a limited capacity.

# 7 DISCUSSION

Our simulation provides valuable background for works based on fake base station attacks, and shows how challenging these attacks and defenses are in practice.

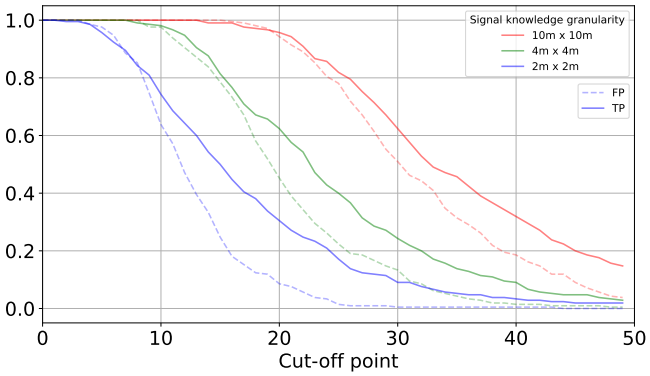


**Figure 5: Success chance of detection using the number of anomalous measurement reports over time. We give results for different levels of signal knowledge.**

## 7.1 Simulation

We choose a simulation approach to be in full control of a large-scale setup with hundreds of active UEs. It allows us to gather a data set that covers a wide performance spectrum. We extend the evaluations of prior work by a large-scale study that measures the performance of attacks and defenses, and shows that we cannot simply expect a successful fake base station attack in all but the most favourable circumstances. We document the most important limitations of fake base station modeling.

**User Mobility.** Users are independent of each other and walk a random trajectory. In a real-world setting, groups and crowds would lead to a different distribution of UEs in an area, which could influence detection. Detection mechanisms do not currently take this into account.

**Signal.** Our signal model approximates real-world signal distributions but ignores moving obstacles and other environmental factors. This leads to less noise in the signal map.

Overall, we find that a simulation approach is the right method to test the assumptions of prior work. Analyzing attacks and countermeasures in a large-scale environment highlights challenges that pose limitations for potential real-world deployment of attacks and detection.

## 7.2 Challenges of Detection

We discuss limitations of fake base station detection mechanisms, which are crucial to consider before implementing these in a real-world environment.

*7.2.1 Overhead.* All detection mechanisms add overhead. Network-based mechanisms in our evaluation (§6.2 and §6.3) rely on existing mechanisms and do not require any additional data exchange. They can still introduce computational overhead for the network. Other network-based detections introduce protocol changes while requiring less computation [4].

*7.2.2 False positives.* Most detection mechanisms introduce a level of false positive detections due to noise and the dynamic transmission characteristics of a network. This leads to a trade-off between reliable detection and collateral damage through incorrect alarms.

# 8 CONCLUSION

In mobile networks, fake base stations are a crucial stepping-stone for attacks that invade the privacy and the security of users. Although prior work introduces different attack and detection mechanisms, limitations in their evaluation leave a blind spot regarding their large-scale capabilities. In this work, we introduced a structured overview of fake base station attacks and defenses, and use a specification-conforming simulation model that enables us to analyze their performance in a large-scale setting. Our results revealed misconceptions regarding the performance of attacks and pointed out challenges in the deployment of detection mechanisms. With these findings, we aim to contribute to the ongoing discussions of the 3GPP, and overcome limitations of prior work.

# REFERENCES

[1] 2024. Fake BTS Framework. https://github.com/thijshberg/fake-bts
[2] 3GPP. 2012. *ETSI TS 136 304 V9.11.0 (2012-07)LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode.* 3rd Generation Partnership Project (3GPP).
[3] 3GPP. 2013. *ETSI TS 124 301 V9.11.0 (2013-04)LTE;Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode.* 3rd Generation Partnership Project (3GPP).
[4] 3GPP. 2014. *ETSI TS 136 331 V9.18.0 (2014-07)LTE; Evolved Universal Terrestrial Radio Access (E-UTRA);Radio Resource Control (RRC);Protocol specification.* 3rd Generation Partnership Project (3GPP).
[5] 3GPP. 2018. *ETSI TS 136 133 V15.3.0 (2018-10) LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management.* 3rd Generation Partnership Project (3GPP).
[6] 3GPP. 2020. *ETSI TS 133 501 V16.3.0 (2020-08)5G; Security architecture and procedures for 5G System.* 3rd Generation Partnership Project (3GPP).
[7] 3GPP. 2022. *3GPP TR 33.809 V0.18.0 (2022-02) Technical Report; 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects Study on 5G Security Enhancement against False Base Stations (FBS).* 3rd Generation Partnership Project (3GPP).
[8] Antennebureau. 2022. Antenneregister. https://antenneregister.nl
[9] Adrian Dabrowski, Georg Petzl, and Edgar R. Weippl. 2016. The Messenger Shoots Back: Network Operator Based IMSI Catcher Detection. In *Research in Attacks, Intrusions, and Defenses: 19th International Symposium, RAID 2016, Paris, France, September 19-21, 2016, Proceedings.* Springer International Publishing, 279–302. https://doi.org/10.1007/978-3-319-45719-2_13
[10] Adrian Dabrowski, Nicola Pianta, Thomas Klepp, Martin Mulazzani, and Edgar Weippl. 2014. IMSI-Catch Me If You Can: IMSI-Catcher-Catchers. In *Proceedings of the 30th Annual Computer Security Applications Conference* (New Orleans, Louisiana, USA) *(ACM Annual Computer Security Applications Conference'14).* Association for Computing Machinery, New York, NY, USA, 246–255. https://doi.org/10.1145/2664243.2664272
[11] GSMK. 2021. *GSMK Overwatch.* https://www.gsmk.de/solutions/network-operator/
[12] Amnesty International. 2021. *NSO Group spyware used against Moroccan journalist days after company pledged to respect human rights.* https://www.amnesty.org/en/latest/news/2020/06/nso-spyware-used-against-moroccan-journalist/
[13] Leyli Karaçay, Zeki Bilgin, Ayşe Bilge Gündüz, Pinar Çomak, Emrah Tomur, Elif Ustundag Soykan, Utku Gülen, and Ferhat Karakoç. 2021. A Network-Based Positioning Method to Locate False Base Stations. *IEEE Access* 9 (2021), 111368–111382. https://doi.org/10.1109/ACCESS.2021.3103673
[14] Prajwol Kumar Nakarmi and Karl Norrman. 2018. Detecting false base stations in mobile networks. https://www.ericsson.com/en/blog/2018/6/detecting-false-base-stations-in-mobile-networks
[15] Peter Ney, Ian Smith, Gabriel Cadamuro, and Tadayoshi Kohno. 2017. SeaGlass: Enabling City-Wide IMSI-Catcher Detection. *Proc. Priv. Enhancing Technol.* 2017, 3 (2017), 39.
[16] NI. 2022. Ettus Research. https://www.ettus.com/
[17] Karl Norrman, Mats Näslund, and Elena Dubrova. 2016. Protecting IMSI and User Privacy in 5G Networks. In *Proceedings of the 9th EAI International Conference on Mobile Multimedia Communications* (Xi'an, China) *(MobiMedia '16).* ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Brussels, BEL, 159–166.
[18] Shinjo Park, Altaf Shaik, Ravishankar Borgaonkar, Andrew Martin, and Jean-Pierre Seifert. 2017. White-Stingray: Evaluating IMSI Catchers Detection Applications. In *USENIX Workshop on Offensive Technologies (WOOT '17).* USENIX Association, Vancouver, BC.

[19] Shinjo Park, Altaf Shaik, Ravishankar Borgaonkar, and Jean-Pierre Seifert. 2019. Anatomy of Commercial IMSI Catchers and Detectors. In *Workshop on Privacy in the Electronic Society* (London, United Kingdom) *(WPES '19)*. ACM, New York, NY, USA.
[20] Stephanie K Pell and Christopher Soghoian. 2014. Your secret stingray's no secret anymore: The vanishing government monopoly over cell phone surveillance and its impact on national security and consumer privacy. *Harv. JL & Tech.* (2014).
[21] David Rupprecht, Katharina Kohls, Thorsten Holz, and Christina Pöpper. 2019. Breaking LTE on Layer Two. In *IEEE Symposium on Security & Privacy (SP)*. IEEE.
[22] srlabs. 2021. *SnoopSnitch*. srlabs. https://opensource.srlabs.de/projects/snoopsnitch
[23] Huahong Tu, Adam Doupé, Ziming Zhao, and Gail-Joon Ahn. 2016. Sok: Everyone Hates Robocalls: A Survey of Techniques Against Telephone Spam. In *IEEE Symposium on Security and Privacy (SP' 16)*. IEEE, San Jose, CA, USA, 320–338.
[24] Yiming Zhang, Baojun Liu, Chaoyi Lu, Zhou Li, Haixin Duan, Shuang Hao, Mingxuan Liu, Ying Liu, Dong Wang, and Qiang Li. 2020. Lies in the Air: Characterizing Fake-Base-Station Spam Ecosystem in China. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security* (Virtual Event, USA) *(ACM Conference on Computer and Communications Security'20)*. Association for Computing Machinery, New York, NY, USA, 521–534. https://doi.org/10.1145/3372297.3417257

## ACRONYMS

**3GPP** 3rd Generation Partnership Project
**DoS** Denial-of-Service
**EMM** EPS Mobility Management
**eNodeB** Evolved NodeB
**EPC** Evolved Packet Core
**GSM** Global System for Mobile Communications
**IMSI** International Mobile Subscriber Identity
**LTE** Long Term Evolution
**MME** Mobility Management Entity
**NAS** Non-Access Stratum
**PLMN** Public Land Mobile Network
**Phy** Physical
**RAN** Radio Access Network
**RSRP** Reference Signal Received Power
**EUTRAN** Evolved Universal Terrestrial Radio Access Network
**RRC** Radio Resource Control
**SIB** System Information Block
**TA** Tracking Area
**TAC** Tracking Area Code
**TAU** Tracking Area Update
**TMSI** Temporary Mobile Subscriber Identity
**UE** User Equipment
**UL** Uplink
**UP** User Plane

## SPECIFICATION FEATURES

In the simulation, we follow the specification and implement those features that are relevant for a network-focused evaluation. We denote fully implemented features with ● and all features that we partially implement with ◐. A partial implementation means that we only cover those elements that are relevant in the context of fake base station attacks and defenses. We do not implement physical transmissions through a wireless channel (denoted as ○). All features of interest are summarized in Table 2.

## NETWORK SETUP

In our network setup, we use fixed locations for base stations and pre-compute a signal map for the entire space of the simulation area. Within the map, we use square buildings to represent obstacles like buildings in an urban area. As the signal model incorporates reflections, the resulting signal map provides a non-uniform distribution of signal strengths. Figure 8 illustrates an example of our map. Lighter areas indicate higher signal strength and squares represent obstacles.

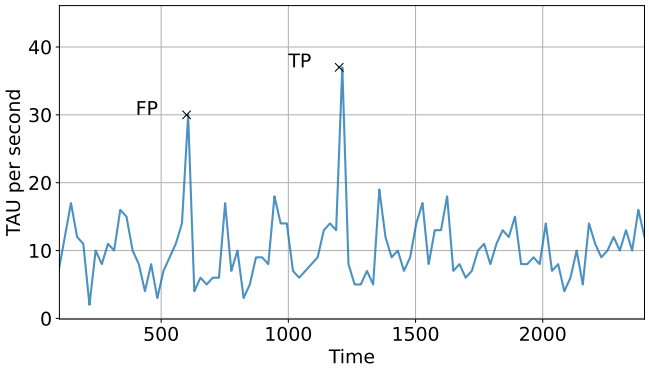


**Figure 6: Example of TAU messages with the dummy TAC value over time. The line indicates the message frequency; the highlighted points show both true positive (TP) and false positive (FP) detections of anomalies.**

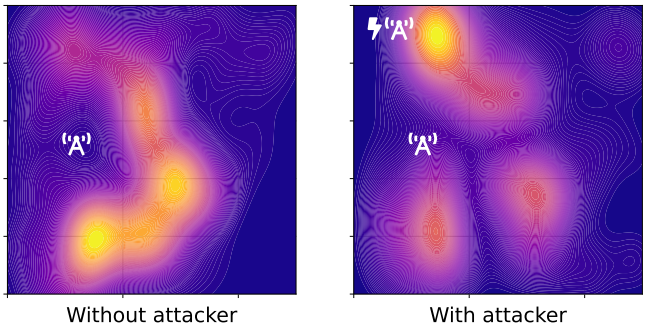


**Figure 7: Signal strengths as given in Measurement Reports for a single base station, with the effect of the attacker spoofing the station's cell identity.**

**Table 1: Overview of Model Parameters.**

| Category | Parameter | §5.3 | §5.4 | §5.5 | §6.2 | §6.3 |
|---|---|---|---|---|---|---|
| Execution §4.3.1 | Map | 2000 | 2000 | 2000 | 2000 | 2000 |
| | Time | 2000 | 2000 | 2000 | 1200+1200 | 1200+1200 |
| | eNodeB | 3 | 3 | 3 | 3 | 3 |
| | Users | 500 | 500 | 500 | 500 | 500 |
| Network §4.3.2 | Frequency | 1 GHz/2 GHz | 1 GHz/2 GHz | 1 GHz/2 GHz | 1 GHz/2 GHz | 1 GHz/2 GHz |
| | Signal Str. | 30 | 30 | 30 | 10 | 10 |
| | Priority | 3/5 | 3/5 | 3/5 | 3/5 | 3/5 |
| Attacker §4.3.3 | Fake BTS | 1 | 1 | 1 | 1 | 1 |
| | Frequency | 2 GHz | 1.5 GHz | 2 GHz | 2 GHz | 2 GHz |
| | Priority | 5 | 2−6 | 5 | 5 | 5 |
| | Rel. Strength | −20dB | 0 dB to 40 dB | 0 dB to 40 dB | 20 dB | −10 dB to 30 dB |

**Table 2: Covered Functionality of 4G Specification.**

| Layer | Feature | Function | Part | Status |
|---|---|---|---|---|
| Phy | Transmission | * | - | ○ |
| | Signal[5] | RSRP | 9.1.2 | ◐ |
| RRC | SIB[7] | SIB2 | 6.3.1 | ◐ |
| | | SIB3 | 6.3.1 | ◐ |
| | | SIB5 | 6.3.1 | ◐ |
| | Cell Selection[2] | PLMN Selection | 5.1 | ◐ |
| | | `idle` mode | 5.2.2 | ◐ |
| | | Cell selection | 5.2.3 | ● |
| | | Cell reselection | 5.2.4 | ● |
| | | Cell camping | 5.2.6 | ● |
| | | Leaving `connected` | 5.2.7 | ◐ |
| | | TA registration | 5.4 | ● |
| | Connection control[7] | System information | 5.2 | ◐ |
| | | Paging | 5.3.2 | ● |
| | | RRC establishment | 5.3.3 | ● |
| | | RRC reconfiguration | 5.3.5 | ● |
| | | RRC release | 5.3.8 | ● |
| | | UE leaving `connected` | 5.3.12 | ● |
| | | Measurements | 5.5 | ● |
| NAS[3] | | Attach procedure | 5.5.1 | ◐ |
| | | Detach procedure | 5.5.2 | ◐ |
| | | TAU procedure | 5.5.3 | ◐ |

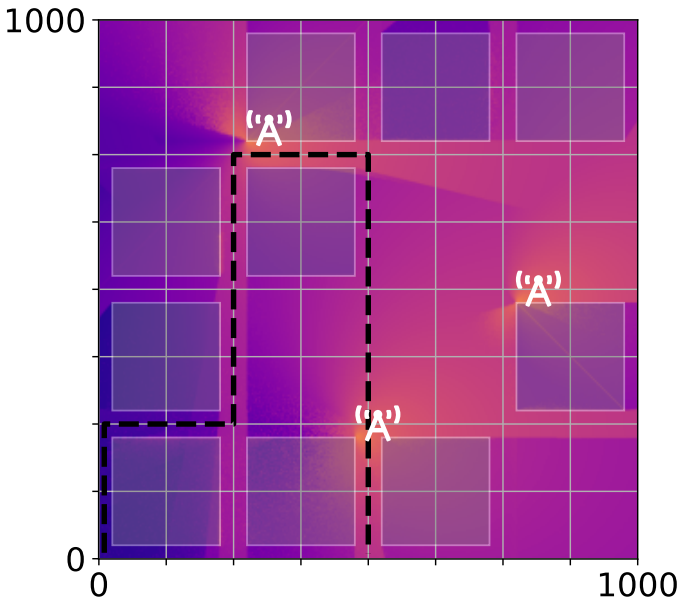


**Figure 8: Example of possible trajectory for a user, with base stations placed at centered positions. Signal strength is shown in the background.**

**Table 3: Cell reselection validation tests and results**

| Description | Oneplus | Huawei | Expectation |
|---|---|---|---|
| Signal loss | ● | ● | lose coverage when srx_lev <0 |
| Intracell | ● | ○ | reselect when target is stronger |
| > q_offset_cell | ● | ○ | reselect when target plus offset is stronger |
| > q_hyst | ● | ○ | reselect when target plus offset is stronger |
| > s_intra_search | ◐ | ◐ | no reselection if srx_lev>s_intra_search |
| Equal priority | ● | ● | reselect when target is stronger |
| > q_freq_offset | ● | ● | reselect when target plus offset is stronger |
| > q_cell_offset | ● | ● | reselect when target plus offset is stronger |
| Lower priority | ● | ● | reselect when srx_lev <0 |
| > thresh_serving_low | ● | ● | reselect when srx_lev <thresh_serving_low |
| Higher priority | ● | ● | reselect always |
| > s_non_intra_search | ◐ | ◐ | no reselection if srx_lev>s_non_intra_search |

●= Test passed, ○= Test failed, ◐= Test failed but effect is not mandatory