A Stealthy Location Identification Attack Exploiting Carrier Aggregation in Cellular Networks

Nitya Lakshmanan and Nishant Budhdev, *National University of Singapore;* Min Suk Kang, *KAIST;* Mun Choon Chan and Jun Han, *National University of Singapore*

https://www.usenix.org/conference/usenixsecurity21/presentation/lakshmanan

This paper is included in the Proceedings of the 30th USENIX Security Symposium.

August 11–13, 2021

978-1-939133-24-3

A Stealthy Location Identification Attack Exploiting Carrier Aggregation in Cellular Networks

Nitya Lakshmanan
National University of Singapore

Nishant Budhdev
National University of Singapore

Min Suk Kang*
KAIST

Mun Choon Chan
National University of Singapore

Jun Han
National University of Singapore

Abstract

We present the SLIC that achieves fine-grained location tracking (e.g., finding indoor walking paths) of targeted cellular user devices in a passive manner. The attack exploits a new side channel in modern cellular systems through a universally available feature called carrier aggregation (CA). CA enables higher cellular data rates by allowing multiple base stations on different carrier frequencies to concurrently transmit to a single user. We discover that a passive adversary can learn the side channel — namely, the number of actively transmitting base stations for any user of interest in the same macrocell. We then show that a time series of this side channel can constitute a highly unique fingerprint of a walking path, which can be used to identify the path taken by a target cellular user. We first demonstrate the collection of the new side channel and a small-scale path identification attack in an existing LTE-A network with up to three CA capability (i.e., three base stations can be coordinated for concurrent transmission), showing the feasibility of SLIC in the current cellular networks. We then emulate a near-future 5G network environment with up to nine CA capability in various multi-story buildings in our institution. SLIC shows up to 98.4% of path-identification accuracy among 100 different walking paths in a large office building. Through testing in various building structures, we confirm that the attack is effective in typical office building environments; e.g., corridors, open spaces. We present complete and partial countermeasures and discuss some practical cell deployment suggestions for 5G networks.

1 Introduction

LTE, the global de facto standard of mobile broadband services [1], is designed to prevent direct leakage of private location information of its mobile user devices [2]. However, recent studies demonstrate that adversaries can infer the locations of targeted individuals [3, 4]. Fortunately, existing location privacy attacks in cellular systems have limitations such as inferring only coarse-grained (e.g., macrocell level) location information [3], or requiring the installation of malware on victims' phones [4].

In this paper, we present SLIC[1], a novel location inference attack that overcomes such limitations of the prior attacks and accurately identifies the walking path taken by a target user. To the best of our knowledge, SLIC is the first attack on cellular networks shown to be effective in indoor, multi-story building environments *without* requiring any malware on the target user's phone. SLIC exploits a new *side channel* in the multi-carrier transmission (also known as carrier aggregation or CA; see §2) — i.e., the *number of actively transmitting base stations for concurrent multi-band transmission* — in modern cellular networks.

The proposed SLIC attack relies on two crucial observations. First, cellular networks produce a specific time-series of the above-mentioned side channel when an individual walks a path. This pattern is path specific and is distinguishable from the patterns observed when walking other paths. We call such unique patterns *path fingerprints*. Each path exhibits a highly unique fingerprint because the *invariant* physical environment (e.g., building architecture) surrounding each walking path affects the radio-frequency (RF) signal quality at each location on the path. As a result, if an adversary records in advance a path fingerprint as she walks a specific path, she can identify with high probability whether a target user travels the same path. Second, the side channel of any active users in a macrocell is *publicly* available since it is broadcast unencrypted; see 3GPP TS 36.321 [5]. Hence, any *passively-monitoring* adversary with commodity tools (e.g., open-source LTE tools [6]) in the same macrocell (e.g., 0.4 to 2 kilometers radius) can obtain the fingerprints, rendering our attack highly stealthy and easily accessible. We have reported our findings to GSMA through the coordinated vulnerability disclosure (CVD) program and they are under review as of September 2020.

To design and evaluate accurate location inference, we

*Corresponding author.

[1] **S**tealthy **L**ocation **I**dentification Attack exploiting **C**arrier Aggregation

address three main technical challenges. First, the comparison between path fingerprints is not straightforward when applying standard techniques for comparing two time series data. The standard dynamic time warping (DTW) technique to compute the similarity between two fingerprints removes the absolute values in fingerprints by normalizing them. The problem is that the normalization of the side channel values can also remove the physical location information of a target user. For example, if a side channel value is seven at a certain time, it means the user is at one of the few spots where seven cells are available for concurrent download. Normalizing this absolute value loses such critical location information. Thus, instead, we use an *absolute-value* DTW technique and preserve the location information during classification (§4).

Another technical challenge is to handle the potential noise in the side channel. The side channel we exploit is shown to reliably capture the number of available base stations around a target user when the user is actively downloading (see §3 for several practical attack strategies to trigger downlink activities of a target user). However, this may not be always possible. The side channel may become *noisy* when a target user requires fewer than the maximum available base stations around him; e.g., only a single base station might be required regardless of where on a path a user is located when downlink activity is low. We address this by *explicitly modeling* the limited downlink activity of a target user. To be specific, we use a single integer-value parameter to model the maximum number of activated secondary cells required for a user. We then calibrate each fingerprint record with respect to this parameter and match with the target's fingerprint (§4). We show that with this calibration, our attack still identifies paths with only modest degradation in attack accuracy because the calibrated fingerprints reliably capture the unique dead spots (i.e., locations with only a small number of available base stations) in many walking paths in practice (§8).

The last challenge is that the *full extent* of the SLIC attack cannot be evaluated in the existing cellular networks. Although our small-scale experiment in an existing LTE-A network with only three CA capability (§5) shows promising results (e.g., 50.1% of path identification accuracy among eight different outdoor paths), our attack is expected to be most effective in a near-future cellular network (e.g., 5G) with highly dense small-cell deployments [7] for larger downlink bandwidth. Such highly dense small-cell networks, however, do not yet exist as of 2020. To that end, we develop a novel Wi-Fi-to-5G evaluation framework that translates a real-world experiment with existing densely-deployed Wi-Fi access points into the emulated 5G cellular network experiment results (§6). We argue that this cross-technology conversion ensures a realistic evaluation of dense 5G small cells because they are often designed to be indistinguishable from the existing Wi-Fi systems when the two systems coexist in the same frequency band; see TR 36.889 [8] and a related whitepaper [9].

Our extensive evaluation with an emulated cellular network with nine CA capability in multi-story buildings shows that the SLIC is highly effective in typical office buildings (§7). When we exhaustively search and select 100 different walking paths in a large building, we achieve up to 98.4% of path-identification accuracy. We extend the experiment to various types of buildings (e.g., corridors, open spaces, shared floors) and show that the SLIC is highly effective in the first two office building types (§8). Additionally, we empirically confirm that the fingerprinting mechanism is robust to minor perturbation (e.g., transmit-power control) of RF measurements.

Finally, we provide a number of countermeasures against the SLIC, including some suggested changes to the 3GPP standards and two partial (but readily available) countermeasures (§9). We also briefly discuss cell deployment suggestions for 5G networks so that the risk of location information leakage can be considered early in the cell planning stage.

2 Carrier Aggregation and New Side Channel

In this section, we first present carrier aggregation technology that is required to understand the SLIC attack. We then introduce the new side channel and its real-world examples.

2.1 Carrier Aggregation for Higher Rates

Traditionally, cellular network users are served by a single base station, called a primary cell. In 2010, to keep up with increasing data consumption, 3GPP introduced a new feature called *carrier aggregation* (CA). CA allows users to connect to one or more additional base stations, called *secondary cells*, to achieve higher data rates [10]. With CA, a user is connected to a primary cell for control messaging as well as connected to multiple secondary cells for downlink transmissions.

In the initial releases, 3GPP specifications supported a maximum of four secondary cells [11]. More recently, it has been extended to a maximum of 31 secondary cells (including aggregation of the unlicensed spectrum [12]). We observe flagship phones in the market with increasing CA capabilities (e.g., seven CA supported in Galaxy S20 [13] compared to five CA in Samsung Galaxy S8 [14]). Operators around the world deploy and extensively utilize CA with several secondary cells (e.g., three CA and four CA capabilities in Singapore and South Korea, respectively). The trend of supporting higher CA certainly exists.

Figure 1 illustrates how CA is configured and activated for a user device.[2] We divide the CA operation into three components: configuration, activation, and transmission.

CA configuration. First, a user equipment (UE) measures the signal strength of all nearby secondary cells and sends these

[2]For brevity and easier understanding, we abuse the terminology a little and refer to the configuration and activation of secondary cells as *CA configuration* and *CA activation*, respectively. For example, three CA configuration/activation refers to two secondary cell configuration/activation in addition to one default primary cell.

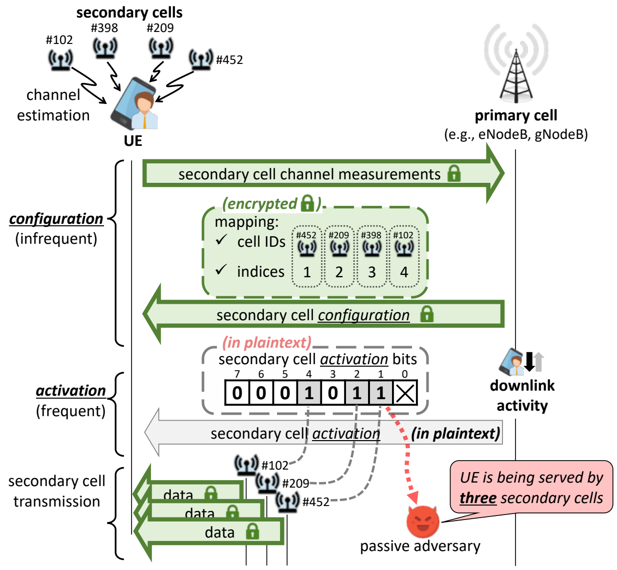


Figure 1: A simplified illustration showing how secondary cells are configured and activated for higher capacity.

measurement reports back to its primary cell. The primary cell applies the secondary-cell configuration algorithm (omitted from Figure 1; see §6.2 for more details) to determine the secondary cells that can be used for downlink transmission. The primary cell assigns each configured secondary cell a unique index so that they can be easily activated later for data transmission (see Section 5.3.10.3(b) in 3GPP TS 36.331 [15]); see the example in Figure 1 where four secondary cells are assigned four indices. Note that the configuration messages are encrypted and thus an unauthorized adversary *cannot* learn the CA configuration (see Appendix A6 in 3GPP TS 36.331 [15]).

CA activation. After the secondary cells are configured for a UE, it can be activated for downlink transmission at any time. In the example shown in Figure 1, as the primary cell receives some downlink traffic for the UE, it activates a subset of the configured secondary cells. The activation is made via a compact (8-bit or 32-bit) activation bitmap MAC control element (MAC CE) where each bit corresponds to the configured secondary cell index (see Section 6.1.3.8 in 3GPP TS 36.321 [5]). Bits set as 1 indicates the activation of the corresponding configured secondary cell and set as 0 indicates deactivation. Note that activation bitmaps are sent to each UE frequently (e.g., 4-8 ms) whenever the secondary cell activation changes. The activation bitmaps are sent *in plaintext* unlike the configuration messages; therefore, any unauthorized adversary, who is in the communication range of a primary cell, can easily learn the number of activated secondary cells for a UE simply by counting 1-bits in a bitmap.

Downlink transmission. Following the activation bitmaps, the UE decodes the scheduling information from the control channel and receives the downlink transmission from the

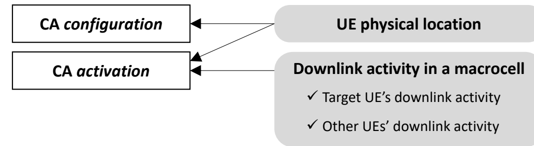


Figure 2: A diagram that illustrates the logical dependency between the CA operations and the UE physical location/downlink activity.

corresponding cells, resulting in a larger aggregate data rate.

2.2 The New Side Channel

The new side channel found in the CA operation is *the number of activated secondary cells* for each UE, as shown in Figure 1. This side channel information can be easily obtained by unauthorized passive adversaries since it is broadcast unencrypted. The adversary can modify open-source tools [6] and utilize commodity software-defined radio devices [16] to decode the control and data channel to obtain the activation bitmap.

Location dependency. Perhaps the most desired property of this side channel for the SLIC attack is its dependency on a target UE's *location*. Figure 2 visualizes this dependency. CA configuration exclusively depends on a target UE's location because it is determined by the distance from nearby secondary cells.

CA activation is also affected by the UE location because the activated cells are strictly a subset of the configured cells; yet, it is also dependent on the downlink activities in the primary cell. First, it is dependent on the target UE's downlink activity because secondary cells are activated only when there is a need for downlink transmission for the target UE. Moreover, CA activation can be additionally affected by the overall load of the cellular system; e.g., the CA activation for the target UE may vary depending on how the secondary cells are already used for other UEs' downlink activities.

It is imperative to ensure that CA activation is *dependent largely on a target UE's location* despite this multi-dependency. In this paper, we show that it is possible in practice. First, we suppress the dependency on a target UE's downlink activity by triggering the target UE's download during the path identification; we present a few practical attack strategies in Section 3. Second, we empirically show that the CA activation in the existing LTE-A networks is strongly dependent on the UE location and less on the load[3] of the cellular system; see our real-world experiments conducted on one of the cellular networks later in this section.

Side channel and path fingerprints. Although this new side channel shows a great potential to leak some UE location information, the side channel itself is insufficient for user location identification attacks. For instance, learning that a

[3]To test in varying system loads, we perform experiments in different times of the day and in different days of the week.

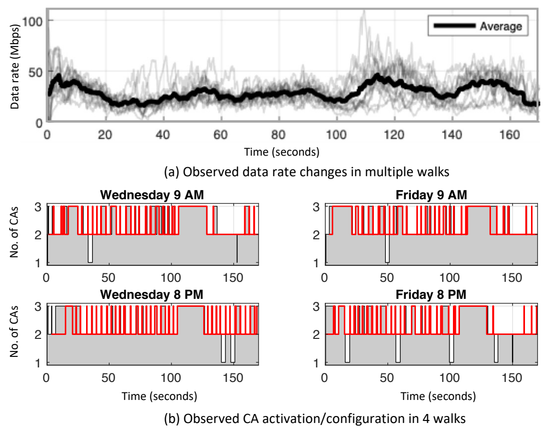


Figure 3: (a) Data rate changes observed in multiple walks on the same 260-meter walking path. (b) CA configuration (red line) and activation (grey area) changes observed in four walks on the same path.

certain UE is being served by three secondary cells does not leak much location information when there exist many other locations that have three (or more) secondary cells.

Instead, we build a *time series* of this side channel information and use it as a *fingerprint* of a walking path when there exist different secondary cell availability along the path. To be specific, as a user walks a certain path, the number of activated secondary cells for the UE may change over time depending on how secondary cells are deployed along the path. We expect unique fingerprints for different walking paths since the deployment of secondary cells and the surrounding physical environment can be highly unique in indoor/outdoor setting.

Real-world evidence. We briefly show the feasibility of this side channel and the fingerprints through our small-scale experiment in an LTE-A network in a metropolitan city, where up to three CA is available (i.e., one primary cell and two secondary cells are available). We walked a 260-meter outdoor pedestrian path fourteen times with a Sony Xperia XZ1 phone. We particularly choose two different times (i.e., 9 AM and 8 PM) in all seven days a week to conduct experiments when the LTE-A network experiences widely different levels of cellular traffic load [17,18]. Figure 3(a) shows how the data rate changes during an entire walk when we keep downloading files. We see that a ***clear pattern emerges***. For example, when a user passes by a certain spot, consistently higher data rates are offered. This means that a user can expect a similar downlink rate change pattern along the path and the pattern seems to be highly independent of when he walks the path.

Figure 3(b) explains why such a clear pattern emerges. The four figures of the CA configuration and activation changes show a highly consistent pattern. This shows that the number of nearby secondary cells at specific spots on the path is highly consistent for different walks at different times. Also, most of the available nearby secondary cells get activated when there exist active downloads during the walks. We observe reliable three CA activation at an average data rate of about 40 Mb/s or higher in our experiment when downloading a large file. We also observe frequent three CA activation when streaming popular YouTube music videos at a moderate 360p video resolution; see more detailed evaluation in Appendix B.

The results clearly show that the side channel is real and, more importantly, has a great potential to be used to form consistent and unique path fingerprints. Similar observations are found in another cellular provider; see Appendix C. More comprehensive evaluations in multiple paths in existing LTE-A networks are found in Section 5, and in a synthesized near-future network in Section 7 and Section 8.

3 Threat Model

Attack goals and capabilities. A SLIC adversary aims to identify the path taken by a target user among all the walking paths that have been fingerprinted in advance. SLIC works for both indoor and outdoor paths as long as they can be fingerprinted. An adversary should be able to fingerprint (e.g., walk with her phone) the paths that are potentially taken by a target user in advance (i.e., during the reconnaissance step (§4.1)).

During the path identification, SLIC requires minimal attack capability of passively monitoring the scheduling channel for a targeted UE. A commodity low-cost software-defined radio device (e.g., USRP [16]) with open-source cellular projects (e.g., srsLTE [6]) is sufficient (§4.2). Note that the SLIC does not require any app installation on the target's phone.

For the path identification, a SLIC adversary can be located anywhere in the radio coverage of the primary cell in which a target UE is located, where the typical radio coverage of a primary cell is 0.4–2 kilometers.

Scope and assumptions. In this paper, we consider that our target user travels on foot to show the feasibility of SLIC. Faster-moving users (e.g., users on scooters or other vehicles) are out-of-scope of this paper as they often move away from short-range secondary cells even before they are activated for CA. We leave attacks on fast-moving users for future work.

We also assume that our target user has some cellular downlink activity because only then the presented side channel is available. Note that it is common to see cellular downlink activity while walking. With the abundance of cellular downlink bandwidth (and the availability of unlimited data plans in many countries [19,20]), it has become a norm to stream music in the form of music 'videos' [21,22]. Moreover, a non-negligible portion of the population watches video content while walking (dubbed as 'Netflix-and-Stroll') [23].

We consider that a SLIC adversary has some basic context information about when her target user is *on the move* and performs the path identification when the target is moving. For example, an adversary may learn the commuting pattern or meeting schedules via the target's public calendar (e.g., a public online calendar with "busy/free" marks).

We consider a single passive adversary device within a primary cell. In principle, with multiple passive adversary devices in adjacent primary cells, an adversary can obtain path fingerprints across adjacent primary cells with minor discontinuity in side-channel measurements.

We use the Temporary Mobile Subscriber Identity (TMSI) of the device to refer to the identity of a user. We assume that adversaries can conduct a separate attack proposed by Shaik et al. [3] (or a similar attack that is recently shown to be feasible in 5G networks [24]) to link the TMSI to the real-world identity of users.

Strategies to activate the side channel. As discussed in Section 1 and demonstrated in Section 2.2, the side channel reliably captures the number of secondary cells at each location when a target user is utilizing the downlink bandwidth. A SLIC adversary may have several options to achieve this:

1) The adversary can encourage a target user to begin downloading certain content for a couple of minutes. For example, a link to any interesting high-volume content (e.g., a BBC breaking news clip) can be sent to a target user to trigger the user's download while walking one of the fingerprinted paths. Additional context information about a target user can be helpful (e.g., sending legitimate work-related file contents). Note that it is far *different* from phishing attacks (where a target user is tricked into clicking a link to malicious content) because the suggested contents are completely benign and malware-free.
2) The adversary can initiate a high-volume interactive session with a target user via emerging technologies, such as augmented reality (AR) or virtual reality (VR) conference calls or multi-player AR games [25, 26]. Even if a target turns off all his video/AR/VR data streams except his voice, the adversary can still activate the target's download by streaming her high-volume video/AR/VR data.

Even if the above options are unavailable, an opportunistic attack is possible.

3) The adversary can opportunistically wait until the target starts some downlink activities. This is possible as the adversary can keep track of the target UE's TMSI over time and monitor the target's real-time downlink rates.

Note that the above strategies do *not* make the SLIC any less stealthy because they are all deemed legitimate behaviors (e.g., sending a legitimate YouTube link, authentic work-related files, or making AR/VR sessions with a target user).

4 Attack Design and Implementation

We begin with the overview of SLIC (§4.1) and explain the detailed attack steps for the main path-identification attack (§4.2), where a SLIC adversary is able to fingerprint all plausible paths. We extend the main attack by considering the case when a SLIC adversary misses fingerprinting some plausible paths (§4.3).

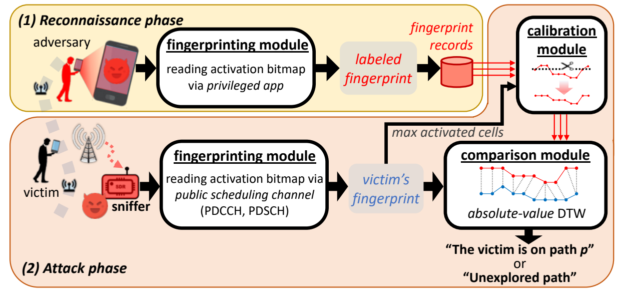


Figure 4: System model of SLIC, which is divided into two phases, namely the Reconnaissance and Attack phase.

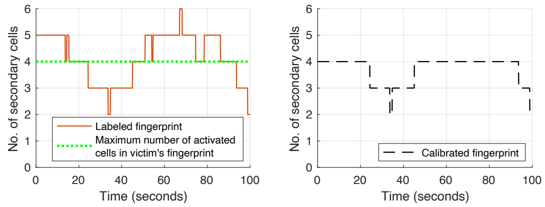
4.1 Overview

SLIC is divided into the following two phases as illustrated in Figure 4. (1) *Reconnaissance phase*: An adversary collects the fingerprints of candidate paths that can be potentially taken by the victim. Specifically, the adversary reads the activation bitmap assigned to a device via a privileged app and obtain the number of activated secondary cells. As the adversary walks a path, the number of activated secondary cells may vary and this change is recorded with the annotated ground-truth location information to obtain a *labeled fingerprint* of the path. The adversary may walk the same path several times to obtain multiple labeled fingerprint records for higher confidence as each walk of the same path will yield similar but slightly different fingerprints. (2) *Attack phase*: The adversary monitors the side channel (i.e., the number of activated secondary cells) of the target victim. Specifically, the adversary utilizes a sniffer tool in the primary cell that can read the activation bitmap broadcast in the public scheduling channel (i.e., by accessing the Physical Downlink Control Channel (PDCCH) and Physical Downlink Shared Channel (PDSCH)). The adversary starts measuring the changes in the side-channel values for a certain duration (e.g., a couple of minutes) to obtain the victim's fingerprint. The adversary calibrates all the labeled fingerprints by limiting the maximum number of activated cells to be the same as the victim's fingerprint to match the victim's data usage or device capabilities. She then compares the victim's fingerprint with all the calibrated labeled fingerprints to identify the path taken by the victim.

4.2 Identifying Fingerprinted Paths

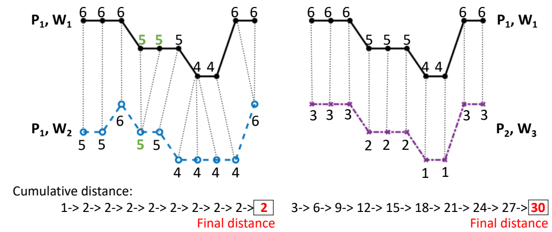
The adversary's goal is to identify the path taken by a victim user among the set of fingerprinted plausible paths. Specifically, the adversary compares the *victim's fingerprint* captured during the attack phase to a set of all *labeled fingerprints* collected during the reconnaissance phase.

Notation. We denote f_p^w in the reconnaissance phase as the labeled fingerprint of a candidate path, p, in a walk, w. Each fingerprint $f_p^w = \{x_1, x_2, \ldots\}$ is a series of non-negative integer side-channel values. We denote the victim's fingerprint in the attack phase as f_u to represent the identity of the path



(a) An example of calibration of labeled fingerprint (b) Resultant calibrated fingerprint

Figure 5: (a) The labeled fingerprint (orange solid line) is calibrated to have at most four activated cells (green dotted line) at any given point of time. (b) The resultant calibrated fingerprint with at most four activated cells.



(a) two walks on the same path (b) two walks on different paths

Figure 6: Illustration of absolute-value DTW. With the absolute-value DTW, we correctly identify the two walks on the same path; i.e., the distance between W_1 and W_2 in (a) is smaller than the distance between W_1 and W_3 in (b).

unknown to the adversary. Hence, the main goal of this attack is to identify the path on which the fingerprint f_u is obtained.

Fingerprinting module *(in (1) Reconnaissance phase).* In the reconnaissance phase, the adversary uses her mobile phone with an app [27] to directly access the messages received from the primary cell. She simultaneously downloads a large file to activate a maximum number of secondary cells and the app reads the activation bitmap along with the configuration messages. By counting the number of activated cells in the activation bitmap while walking a designated path, the adversary obtains the labeled fingerprint of the path.

Fingerprinting module *(in (2) Attack phase).* In the attack phase, the adversary obtains the fingerprints differently by reading the public scheduling channel of the primary cell. As the public scheduling channel is broadcast in clear by the primary cell, the adversary can passively read the activation bitmap information as long as she is in the reception coverage of the victim's primary cell. An adversary can utilize the open-source projects (e.g., srsLTE ([6]) to record the entire downlink signal for the duration (e.g., 2 minutes) of a walk, and read the bitmaps in an offline manner. We use the Matlab LTE Toolbox [28] to read the LTE samples and decode PDCCH and PDSCH. It takes 90 milliseconds on average to decode a PDCCH and a corresponding PDSCH (i.e., one subframe) for a single UE. Overall, it takes about 5.6 minutes for reading a fingerprint of a single victim UE when decoding multiple subframes in parallel with a 32-core machine.

Calibration module. Before this comparison though, we first calibrate all the labeled fingerprints with the maximum number of activated secondary cells observed in the victim's fingerprint. For this, we *cap* all the labeled fingerprints with the maximum CA value of the victim's fingerprint. For example, if the maximum number of activated secondary cells in the victim's fingerprint is four, then we calibrate all the labeled fingerprints to have a maximum of four activated secondary cells at any given point of time; see an example in Figure 5(a) and 5(b). We perform this calibration to adjust the labeled fingerprints as close to the victim's fingerprints in the cases of the victim's low downlink usage or restricted device capabilities.

Comparison module. Upon calibration, we input the calibrated labeled fingerprints and the victim's fingerprint into the comparison module. This module has two main tasks. First, it computes each of the *comparison distances*. Second, it *selects the smallest comparison distance* and outputs the corresponding path as the identified victim's path, $\hat{p}$. That is, we compute $\hat{p} = \arg\min_p \mathsf{dist}(f_p^w, f_u)$.

We implement the distance function $\mathsf{dist}(\cdot)$ for comparing the distance using a modified version of DTW [29]. This enables SLIC to adequately compare a pair of fingerprints of the same path that vary slightly due to several reasons, including non-identical walking speeds and patterns. It finds the best alignment between the two fingerprints by warping or stretching the time axis. It works by computing a cumulative distance between each pair of side-channel values of the two fingerprints to finally obtain DTW distance between the pair. For example, Figure 6(a) depicts the mapping between side-channel values of two walks, W_1 and W_2. It depicts that the 4^{th} side-channel value (i.e., 5) in W_2 is stretched to map with the 4^{th} and the 5^{th} value in W_1. The plot also depicts the minimum cumulative distance computed for each matching pair to finally obtain the minimum final distance of 2 (in red).

Note that, unlike the standard DTW technique where the amplitude of the data is considered as noise, the amplitude of our data (i.e., the number of transmitting secondary cells) is the important side-channel information. Hence, we do *not* normalize the amplitude of our data because, otherwise, we may lose the most critical information in the side channel; namely, the physical location information. We call this *absolute-value* DTW. Figure 6 illustrates the absolute-value DTW mapping between fingerprints of walk W_1, W_2 on the same path P_1 and W_1, W_3 on different paths P_1 and P_2. The figure depicts that the cumulative distance computed for each pair of side-channel values of W_1, W_2 does not increase due to the high similarity in their values. This similarity is correctly reflected in the final distance which is a low value of 2. However, the final distance of W_1 and W_3 is 30 (>2) indicating the walks to be less similar than W_1, W_2. Thus, W_1 and W_2 are correctly identified to be belonging to the same path. If the data were to be normalized, W_1 and W_3 will be identified to be more similar (as their shapes are similar) which may lead to incorrect path identification. As this example illustrates, SLIC utilizes the absolute values of our fingerprints as critical information.

4.3 Handling Unexplored Paths in SLIC

We design an extension of the aforementioned SLIC attack when the adversary *misses* fingerprinting some plausible paths during the reconnaissance phase. We call such paths *unexplored*. The comparison module handles the explored paths in the following ways. It first computes the average DTW distance between all the labeled fingerprints per path to obtain a distance threshold. This threshold determines the average similarity of labeled fingerprints per path. We consider the victim to be walking an unexplored path if the minimum DTW distance between the victim's fingerprint and labeled fingerprints of all paths is greater than the distance threshold (+δ). If the victim's path is already fingerprinted, then the corresponding path is identified as above.

5 Feasibility of SLIC in Existing LTE-A Networks

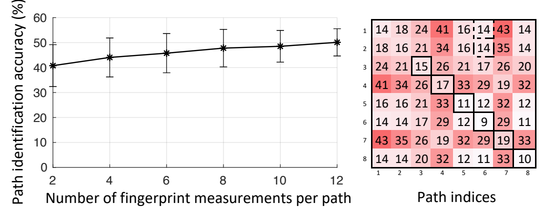
We first aim to show that the SLIC is feasible in an *existing* LTE-A network in a metropolitan city. By showing the attack feasibility in a cellular network with only three CA capability, we want to discuss the promising potential of the SLIC in emerging cellular networks with much higher CA capability.

5.1 Experiment Setup: SLIC in LTE-A

We test the LTE-A network that supports up to three CA capability (i.e., one primary cell and up to two secondary cells) and we use two Sony Xperia XZ1 phones for the experiments. We run MobileInsight [27] on the phones to decode LTE-A messages, just as the adversaries would do for their reconnaissance step. We have walked eight non-overlapping outdoor paths with an average distance of 260 meters and measured their path fingerprints. We have conducted experiments for seven days at three different times of a day (morning, afternoon, and evening) to obtain eighteen walks per path (144 instances of walks) totaling 37 kilometers in 10 hours.

5.2 Path-Identification Results in LTE-A

We partition our dataset into the labeled fingerprints collected during the reconnaissance phase and the victim's fingerprints collected during the attack phase. For the labeled fingerprints, we use one Sony phone to collect data whereas for the victim's fingerprint we use another Sony phone. We consider data collected on the five days (i.e., twelve walks per path, a total of 96 instances) as the labeled fingerprints and in the other two days (i.e., six walks per path, a total of 48 instances) as the victim's fingerprint.



(a) Path identification accuracy of SLIC (b) Correlation matrix

Figure 7: Accuracy of path-identification attack. The identification accuracy increases as the number of labeled fingerprints increases. The errorbar denotes the standard deviation.

We measure the number of configured secondary cells as a proxy for the number of activated secondary cells due to practical constraints in measuring activated secondary cell at a large scale; e.g., our LTE-A providers have a per-month total bandwidth limit for each SIM, and only a finite number of SIMs can be purchased for academic research. In practice, SLIC adversaries would *not* be bound by these constraints because they would be able to purchase large numbers of SIMs for the reconnaissance step. We model the number activated secondary cells using the measured configured cells by adding random noise to the 13% of all measured samples of the configured cells. This is based on our observation that the number of configured and activated secondary cells matches by 87% in the experiment conducted over seven days (§2.2).

We randomly select a varying number of labeled fingerprints per path ranging from 2 to 12 for all the eight paths. We also randomly select thirty fingerprints as the victim's fingerprints. We iterate this process one hundred times for ten random noisy data. Figure 7(a) shows the path-identification accuracy of SLIC with the increasing number of labeled fingerprints per path from 2 to 12. The overall accuracy increases from 40.7% (for two labeled fingerprints per path) to 50.1% (for twelve labeled fingerprints per path).

To further investigate, we analyze the similarities between all fingerprints in all the paths in our dataset. Figure 7(b) shows this in a correlation matrix of the eight paths. The number in a box (i, j) represents the average of all the pairwise DTW distance between all the fingerprints in the i-th path and the j-th path. For the majority of the paths (i.e., path indices 3, 4, 5, 6, 7, and 8), the fingerprints measured for the same path show the smallest DTW distances. For the other cases, the average DTW distances between the ones measured for the same paths are still close to the smallest.

5.3 Promising Potential of SLIC

While overall accuracy is far from ideal, the result in the existing LTE-A system shows a promising potential, particularly considering that the fingerprints are generated with coarse-grained side channels (i.e., only three values in a three CA). First, the result shows that similar fingerprints are expected in two different walks on the same path. Our samples in Figure 7 are collected across seven days in the mornings, afternoons, and evenings, and the fingerprints are quite consistent in general. Second, two distinctive fingerprints are expected in general when measured on different paths.

With these promising results, we envision much higher path-identification accuracy with a finer-grained fingerprinting in a cellular network with higher CA capability. In this paper, therefore, we aim to evaluate SLIC in a near-future network with densely deployed small cells. The next few sections describe how we experiment SLIC in a near-future network. In particular, Section 6 describes our Wi-Fi-to-5G evaluation framework where we use an existing WiFi deployment to emulate a densely deployed cellular network.

6 Wi-Fi-to-5G Evaluation Framework

How can we test the SLIC attack in a cellular network with a slightly more (e.g., 7–9) CA capability — a typical near-future 5G deployment scenario? To demonstrate the SLIC attacks in such a not-fully-deployed dense small cell cellular system, we design and implement a *Wi-Fi-to-5G evaluation framework*. Before we describe the rationale behind the design of our framework, we briefly explain why the two typical evaluation approaches, namely, early-deployment/testbed sites and computer simulations, are inappropriate.

- *Early-deployment sites?* As of January 2020, there are only some cities/countries that have deployed 5G infrastructure [30]. In addition, to the best of our knowledge, none of the real deployment (or testbeds) has dense secondary cell deployment at a large scale yet.
- *Computer simulations?* With some advanced computer simulation tools (such as ns-3 [31], OMNeT++ [32], OPNET [33]), one can model and simulate various *physical environments* (e.g., walls, moving objects) for point-to-point wireless channel experiments. Yet, it would be extremely costly computationally (if not infeasible) to simulate large-scale experiments with tens of small cells and buildings with many walking paths, which is crucial for realistic attack evaluation.

6.1 Rationale Behind Using Wi-Fi for 5G Evaluation

The high-level idea of our 5G evaluation framework is to utilize an existing, densely-deployed *Wi-Fi* system to evaluate the behavior/deployment of 5G secondary cells. We argue that this seemingly unconventional idea is indeed a sound approach to evaluating our attacks in realistic 5G environments. The rationale behind this approach is that one of the most promising dense secondary cell technologies, *License Assisted Access (LAA)* which can aggregate unlicensed spectrum, is specifically designed to be *indistinguishable* from the existing Wi-Fi systems for the "effective and fair coexistence" with existing Wi-Fi systems [8, 9].

Table 1: Comparison of Wi-Fi and LAA

Layer	Features	Wi-Fi	LAA
Physical Layer	Frequency Band	5 GHz unlicensed	5 GHz unlicensed
	Max. Transmit Power	1 Watt	1 Watt
	Modulation	OFDM	OFDM
	Transmit Power Control	Yes	Yes
Medium Access Layer	Channel Access	CSMA/CA	Listen-Before-Talk (LBT)
	Contention Window	Exponential increase	Exponential increase
	Transmitter Detection	Beacon Signal	Discovery Signal
	CCA slot duration	9 μsec	9 μsec
	Transmission Duration	5 to 10 msec	2, 3, 8, or 10 msec

Nearly identical physical and medium-access layers. We first illustrate how LAA is designed to be similar to the Wi-Fi system in the physical and medium-access layers in Table 1. First of all, the physical-layer characteristics of the two systems are nearly identical. The two systems use the same 5 GHz unlicensed frequency bands with the same maximum transmit power of 1 Watt [8, 34] and the same OFDM signal modulation. Second, the medium-access layer of LAA is particularly designed to behave like the Carrier Sense Multiple Access/Collision Avoidance technique in the Wi-Fi standard [35]. To be specific, the LAA secondary cells follow Listen-Before-Talk (LBT) procedure in which the cells have to wait for the medium to be free before it can transmit on it [36]. The two systems have the same slot time duration and their contention windows increases exponentially (see Section 15 in 3GPP TS 36.213 [37]). Finally, they transmit periodic signals in a similar manner and have the maximum transmission duration.

The *indistinguishability* requirement. The LAA secondary cells operate in the same frequency as the existing Wi-Fi systems. Naturally, concerns for fair and effective coexistence have been raised by the Wi-Fi standard community, industry, and regulatory bodies since as early as 2015 [38, 39]. After several years of intense discussion, the 3GPP's solution is to enforce strict constraints for the LAA secondary cells and make them *indistinguishable* from Wi-Fi systems; that is, from the Wi-Fi devices' point of view, these LAA cells should look nearly the same as another Wi-Fi system. Their design principles are well explained in the standard document (see Section 7 in 3GPP TR 36.889 [8]) and even explicitly used as a performance evaluation criterion in Intel's whitepaper on their 4.5G systems [9]. This indistinguishable system design and deployment lead us to believe that the future 5G LAA secondary-cell deployment is likely to be similar to the existing Wi-Fi deployment. Hence, this constitutes the justification of our attack evaluation in Wi-Fi networks.

6.2 Overview of Design and Implementation

Although the idea of evaluating SLIC in Wi-Fi networks may seem straightforward, in fact, it requires a number of non-trivial emulations of 5G physical and medium-access layers as well as the user behavior models to obtain the accurate

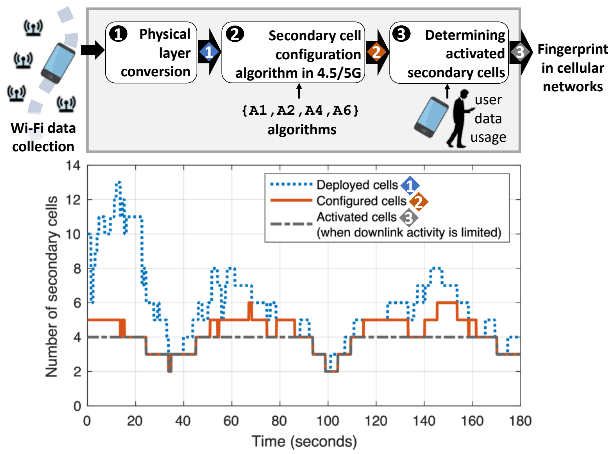


Figure 8: Wi-Fi-to-5G evaluation framework with an example outcome at each stage. The plot represents the change in the number of secondary cells at different stages of the evaluation framework for a single walk on an indoor path.

side-channel information in 5G networks.

Figure 8 depicts our Wi-Fi-to-5G evaluation framework from the collection of Wi-Fi received signal data to the fingerprint. There exist three stages. In ❶, the Wi-Fi signals are converted to 5G signals. Then the secondary cell configuration algorithm, such as `A1`, `A2`, `A4`, `A6` (see Section 5.5.4 in 3GPP TS 36.331 [15]) is applied in ❷ to emulate the configuration of secondary cells for the device. We consider the above algorithms because it mainly utilizes the signal strengths from secondary cells which we emulate using Wi-Fi APs. The dotted blue line in Figure 8 represents the total number of secondary cells deployed at the user location. We can see that only some of them are actually configured by the evaluation of the secondary cell configuration algorithms; see the solid orange line. In ❸, the number of activated secondary cells is determined based on the data usage to finally obtain the fingerprint. When a user's demand for downlink data stream is limited to a certain extent (e.g., watching videos with the low-resolution setting), the cellular system would not utilize more secondary cells than needed; see the dashed grey line in Figure 8 for the case when a target user needs only up to four secondary cells. We present an in-depth evaluation in Section 8.2 regarding the effects of fingerprints with the limited downlink activities. Refer Appendix A for the details of the three implementation steps.

6.3 Limitation in Wi-Fi-to-5G Evaluation Framework

Our Wi-Fi-to-5G evaluation framework demonstrates realistic emulation of SLIC in certain probable 5G deployment scenarios (e.g., densely deployed unlicensed small cells) but not all possible 5G scenarios. In some 5G deployment scenarios (e.g., millimeter waves, sub-6 GHz spectrum), one may expect different SLIC attack performance due to (1) the different radio frequency ranges, and (2) different indoor/outdoor environments. The range of Wi-Fi APs can be shorter than the range of some 5G cells. Hence, the activation/deactivation of the secondary cells may happen more slowly in some 5G deployments compared to the Wi-Fi deployment leading to a coarser fingerprint. One can also expect different fingerprint quality for outdoor deployments as they may consist of cells with a higher range as compared to indoor deployments. All these differences may affect the performance of the SLIC attack, for example, the attacker may need to obtain a longer victim's fingerprint to identify the path.

7 Evaluation of SLIC in Emulated 5G Networks

In this section, we demonstrate the performance of SLIC in our emulated 5G network with up to nine CA capability using the technique introduced in Section 6. We first describe a detailed experiment setup (§7.1) and present the main performance evaluation of SLIC in two common attack cases (§7.2).

7.1 Experiment Setup for Large-Scale Measurements

Apparatus. For the experiment with the Wi-Fi-to-5G evaluation framework (§6), we develop an Android app on a smartphone that collects received signal strengths indicator (RSSI) values of the Wi-Fi access points (operating in both 2.4 and 5 GHz). We also collect other auxiliary information including Wi-Fi SSIDs and their operating frequencies. We further preprocess the collected data by removing data from 2.4 GHz band to emulate the unlicensed 5G secondary cell deployed in the 5 GHz band [8].

Data collection. We invite four participants to walk on a total of 46 different indoor paths, with an average distance of 56.4 meters. We collect the data using two Nexus 6 phones running the aforementioned data collection app. The paths include corridors and staircases of a large office building with five floors in our institution. Walking different directions on the same path (i.e, 1 to 2 and 2 to 1) are considered as two different paths. We ask each participant to walk 10 times per path, yielding a total of 460 instances of walks totaling 25.9 kilometers across the paths in five hours. The walks are distributed between different times of day (i.e., mornings, afternoons, and evenings) across a duration of 21 days to demonstrate SLIC's robustness across different times.

Path synthesis for large-scale experiments. The number of possible distinct paths in a typical multi-story office building may, in fact, be much higher than the number of paths we choose to walk (i.e., forty-six). Office buildings have several indoor intersections, entrances, exits, and staircases and, thus, there could exist several hundred or even a thousand different

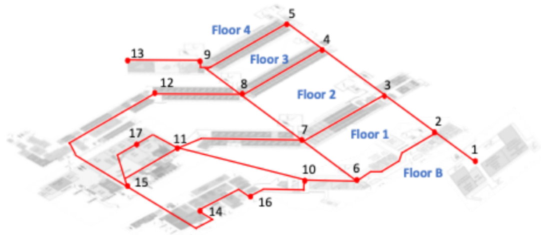


Figure 9: A graph of path segments constructed for a large office building.

Table 2: Number of synthesized paths for path distances

Path distance (meters)	Number of synthesized paths
200-250	291
300-350	624
400-450	1,044
500-550	1,307
600-650	1,073

paths in a building when considering 1-2 minute short walks. Walking all possible 1-2 minute paths (perhaps, multiple times per path) in practice, therefore, would result in an extremely labor-intensive experiment. Instead, we carefully *synthesize* all possible paths in the building using the 46 short paths (or path segments) we measure separately. The synthesized paths represent the large number of plausible paths taken by the victim within primary cell coverage.

For path synthesis, we construct a graph that is made of the above forty-six short path segments as the edges. Figure 9 shows the graph for a large building we test. We superimpose the floor plan of the building for easier interpretation. The graph has 23 edges (46 directed edges) and 17 vertices where every edge connecting any two vertices are collected paths.

We first construct as many distinct paths as possible by finding all random walks without repetition (i.e., no walk on the same segment in a path regardless of the direction). Table 2 shows the number of synthesized paths for different path length ranges. As expected, we can create more distinct paths for longer path lengths. When enumerating for the path length of $[500, 550)$ meters, we can find 1,307 different paths in the building, after which the number gradually decreases for even longer paths due to the limited building size. Then, we obtain fingerprints for the new synthesized paths. We randomly pick the walks from ten collected fingerprints for each path and form the new synthesized fingerprint by concatenating them.

7.2 Attack Evaluation

We present the main evaluation results of the SLIC, focusing on two attack cases: (1) a SLIC adversary fingerprints *all* plausible paths of a victim user, and (2) a SLIC adversary fingerprints *some* plausible paths of a victim user.

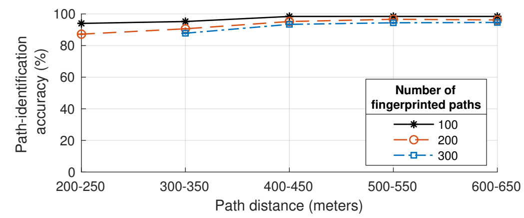


Figure 10: Path-identification accuracy of SLIC. The overall accuracy increases with increasing path distance for a different number of fingerprinted paths.

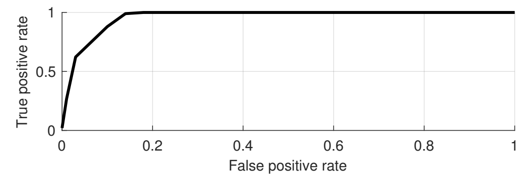


Figure 11: The ROC curve when some of the victim's paths are unexplored. The curve shows that the adversary can achieve high accuracy in identifying fingerprinted paths with low false positives.

7.2.1 Attack Case 1: Fingerprinting *All* Plausible Paths

This attack case models a resource-rich adversary where it is feasible to fingerprint all plausible paths of the victim within primary cell coverage. This is, in fact, similar to our experiment setup where we exhaustively search all possible indoor paths in the large buildings as shown in Figure 9.

Figure 10 depicts the path-identification accuracy of SLIC for varying numbers of fingerprinted paths (100, 200, and 300) with increasing path distance. We evaluate the path-identification accuracy for varying path distance (from 200 to 650 meters). We consider nine labeled fingerprints per path and 100 victim's fingerprints and perform five iterations each time by randomly choosing both labeled and victim's fingerprints. We consider that the victim is fully utilizing all the available secondary cells at each location. Figure 10 indicates that the accuracy increases when the victim walks longer paths. For example, the identification accuracy of 100 fingerprinted path increases from 94% to 98.4% as the distance of the path increases from 200-250 meters to 600-650 meters. Note that we omit the data point at the path distance of 200-250 meters for 300 fingerprinted paths because only 291 paths are available; see Table 2.

7.2.2 Attack Case 2: Fingerprinting *Some* Plausible Paths

In the second attack case, we model a resource-constrained adversary whose best effort reconnaissance may fingerprint the majority of (but not all) plausible paths that could be taken by a victim. Thus, a victim may walk a path that is not fingerprinted by an adversary (or unexplored path).

In this evaluation, we randomly select 300 fingerprinted

Figure 12: Multi-story building types: (a) corridor (e.g., office), (b) open space (e.g., library), and (c) shared floor (e.g., shopping mall).

paths from a total of 624 possible paths with a path distance of 300-350 meters whereas the remaining 324 paths (i.e., 624-300 = 324) are unexplored paths. We then evaluate for 100 victim's fingerprint in which 80 are fingerprinted paths and 20 are unexplored paths. We perform five iterations each time randomly choosing both labeled and victim's fingerprints. Figure 11 shows the ROC curve for 300 fingerprinted paths with the x-axis as the false positive rate (i.e., the ratio of unexplored paths falsely identified as fingerprinted paths) and the y-axis as the true positive rate (i.e., the ratio of fingerprinted paths correctly identified as a fingerprinted path) for varying distance thresholds. The large area under the curve indicates a high true positive rate for a relatively low false positive rate.

Note that care must be taken in this evaluation because some unexplored paths may share some path segments with the fingerprinted paths. For example, an unexplored path 6-7-11-17 shares two path segments with a fingerprinted path 6-7-11-15. If we use such unexplored paths and fingerprinted paths with the shared path segments, our results can be biased due to the artifact of path synthesis. To address this, we remove such unexplored paths from the data set and finally obtain only 87 unexplored paths for the experiment above.

8 How Reliable is the SLIC Attack? — Additional Evaluations

One remaining question is whether the SLIC can reliably identify the path taken by a victim in various environments and operating scenarios. In this section, we perform two additional evaluations to answer: (1) whether the SLIC attack effectiveness depends on different physical environments, particularly various types of building structures (§8.1); and (2) whether the SLIC attack can still identify the path when a victim UE does not fully utilize its downlink bandwidth (§8.2). In addition, we also analyze the effect of minor perturbation of the received radio signal indicator (RSSI) values due to environment changes and transmit power control (TPC) algorithms in the modern wireless systems; refer Appendix D for details.

8.1 SLIC in Various Building Types

Simplified experiment setup. For these extra evaluations, we collect fingerprints from five buildings and use them *directly* (i.e., not going through the path synthesis) for path identification. The measured paths are thus longer than the short path segments we choose for Section 7. We have walked a total of 45 unique paths in the five buildings, collecting 10-14 fingerprints per path with an average distance of 128 meters, totaling 63 km in 12 hours.

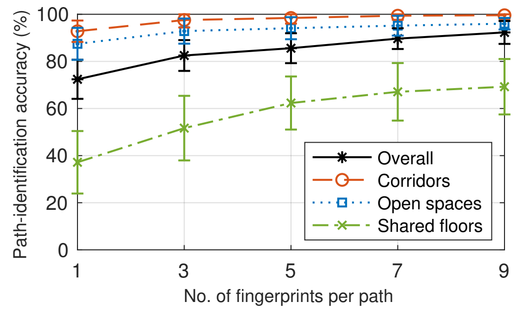


Figure 13: Path-identification accuracy in three different building types. SLIC in the corridor and open spaces results in higher accuracy than shared floors. The errorbar denotes the standard deviation.

We classify the five buildings we have walked into three types of building structures to demonstrate the effect of building structures on the uniqueness of the paths and ultimately on SLIC's performance. The three types are *corridors* (e.g., offices), *open spaces* (e.g., libraries), and *shared floors* (e.g., malls), as illustrated in Figure 12.

To evaluate the performance of SLIC for different building types, we vary the number of labeled fingerprints per path (for all 45 distinct paths) from one to nine (yielding a range of 45 to 405 fingerprints). We also select 30 victim's fingerprints and evaluate for 100 iterations each time picking labeled and victim's fingerprints randomly. Figure 13 depicts the overall identification accuracy for all building types when varying the number of labeled fingerprints per path. The overall accuracy increases to 92.3% with nine labeled fingerprints per path, as depicted by the black solid curve. Figure 13 also depicts the individual accuracy for corresponding labeled and victim's fingerprints collected for specific building types.

We observe that the corridors (orange dashed line) and the open spaces (blue dotted line) yield higher accuracy compared to the shared floors (green dash-dot curve). This is most likely because the first two yield more unique fingerprints than the latter due to their building structure. To confirm our conjecture, we plot the probability density function for the total number of configured cells in our dataset in Figure 14. The plot indicates that the corridors and open spaces have a wider distribution compared to the shared floors, meaning the fingerprints on these buildings have a wider range for variations in their values. Moreover, we compute the average entropy of the fingerprints of paths for each of the building types to check that open spaces and corridors indeed exhibit more variation hence more unique fingerprint. The result-

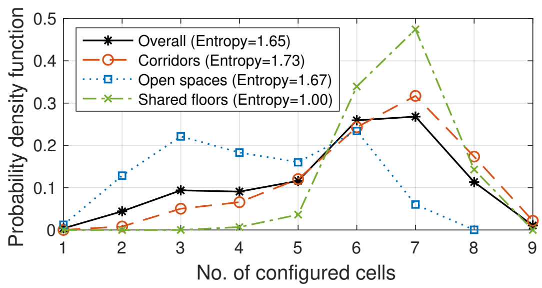


Figure 14: Probability density function for the total number of configured cells. The corridors and open spaces have a wider distribution compared to the shared floor. The entropy indicates that the corridors and open spaces exhibit more variation in their fingerprints.

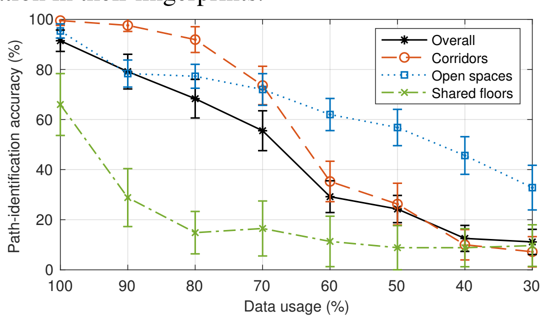


Figure 15: Path-identification accuracy when a victim requires less than the full available bandwidth in the CA configuration.

ing average entropy confirms our conjecture with corridors, open spaces, and shared floors each yielding 1.73, 1.67, and 1.00, respectively. Consequently, the identification accuracy of corridors and open spaces is higher than shared floors.

8.2 SLIC with Limited Downlink Activity

We evaluate the performance of SLIC when the victim has limited downlink data usage, thereby requiring only some of the configured secondary cells to be activated. We also evaluate across different building types. To simulate varying data usage of the victim, we limit the maximum number of configured secondary cells that can be utilized at any given point in time by capping the fingerprint (see Figure 5). We calibrate all the labeled fingerprints based on the victim's fingerprint. We consider nine labeled fingerprints per fingerprinted path and evaluate for 100 iterations each time picking labeled and victim's fingerprints randomly.

Figure 15 depicts the identification accuracy for data usage varying from 100% (i.e., a maximum of nine transmitting cells) to 30% data usage (i.e., a maximum of three transmitting cells). This yields an overall accuracy greater than 50% when the victim's data usage requires at least 70% of the configured secondary cells to be activated. However, it decreases further to only 24.2% for lesser data usage. Hence, we can conclude that SLIC is most effective when the victim's data usage is high triggering most of the configured secondary cells to be activated. We also observe that the open spaces yields higher accuracy even at low data usage compared to corridors or shared floors. The reason can be inferred from Figure 14 which indicates that the open spaces have lower values of side-channel information, which are accurately captured by the calibrated fingerprints leading to better accuracy.

9 Countermeasures

The CA side-channel we discover in this paper is fundamental scheduling metadata of cellular system and thus an end-user cannot remove it without the support from the operator, or, better yet, some changes in the standard. We begin with one countermeasure that *removes* the side channel completely with the modification to the standard. We then discuss how the operators can mitigate the attack without changing the standard. We also discuss some cell deployment suggestions for 5G networks.

9.1 Encrypting Side-Channel Information

This countermeasure makes the secondary cell activation information confidential to any unauthorized parties by encrypting it. In the protocol stack, the resource scheduling is done by the Medium Access Control (MAC) layer which is below the Packet Data Convergence Protocol (PDCP) layer that is responsible for encryption. Implementing the encryption of the bitmap, thus, requires significant changes in the protocol stack. Particularly, the MAC layer must be incorporated with the encryption mechanism to protect the scheduling information. In addition, this change may require a new symmetric key between every UE and the primary cell, which involves changes to the existing key management schemes [2] as it is generally advisable to have separate keys for different purposes.

9.2 Readily-Available Countermeasures

We propose two highly effective countermeasures that require no changes to the standards.

Adding noise to the side channel. This countermeasure exploits the standard operation to make the side channel noisier, rendering the SLIC less effective. According to the specification (see Section 6.1.3.8 in 3GPP TS 36.321 [5]), a device *should ignore* the activation of any secondary cells that are *not* in the list of the configured secondary cells. When some indices in the bitmap are used for the configured cells, the primary cell can activate any of the unused indices to add noise to the side channel. These additional indices are ignored by the device and thus have no effect on the scheduling of the system; yet, these intentionally added indices are wrongly considered as activated cells by an adversary.

Note that the amount of added noise may be limited based on the number of configured secondary cells. If a large number (e.g., close to 31) of secondary cells is already configured, then the primary cell can add only a small number of noise secondary cells.

Changing device identifiers frequently. As we explain in our threat model (§3), SLIC requires linking the real-world identity to the network identity of the devices (i.e., TMSI). A number of attacks, such as Shaik et al. [3] or Rupprecht et al. [40], have demonstrated that such mapping of identities in different layers is possible in 4G networks. As a countermeasure, thus, operators can change the TMSIs frequently (e.g., every connection request), making the mapping difficult in practice. Note, however, that care must be taken when designing and implementing reallocation of TMSIs as new TMSIs can be traceable with new attack strategies; see how one can link the changing TMSIs in the 5G network with the recent attack by Hussain et al. [24].

9.3 SLIC-Aware Cell Planning

Proper *cell planning* is necessary when deploying or upgrading cellular networks. Operators decide numerous high-level system parameters during cell planning, including the basic cell layouts [41], frequency allocations [42, 43], inter-cell interference management [44], etc., to achieve several operational goals [45]: improving radio coverage, maximizing resource utilization, optimizing the system capacity, etc.

A *SLIC-aware* cell planning considers the risk of path identification from the early stage of cell deployment. What we report in this paper is that high variance in the path fingerprints makes the path-identification attacks highly effective. The SLIC-aware cell planning would minimize the variance in the number of nearby cells at different locations, rendering the SLIC highly ineffective. 5G is still at its infancy and, thus, this new SLIC-aware planning goal can be considered for the upcoming network deployment. Also, 5G is expected to be much more heterogeneous (e.g., millimeter waves, sub-6-GHz spectrum, unlicensed bands) than 4G networks and this heterogeneity would make the SLIC-aware cell planning optimization more viable. We leave this as future work.

10 Related Work

Location-privacy attacks exploiting cellular networks. Recent studies demonstrated that adversaries may be able to infer the location of targeted individuals and their traces [3,4]. The closest work to SLIC is proposed by Michalevsky et al. [4] called PowerSpy attack that exploits the fact that mobile devices experience similar power changes when traveling on the same driving path due to the static cellular tower locations. While PowerSpy and SLIC both utilize side-channel information from the cellular network to infer the user's location, they differ significantly in the following two aspects. First, PowerSpy requires *installing malware* on the victim devices, which increases the attack cost. SLIC inherently does not have such requirements as it can simply capture the victim's information by passively monitoring the cellular network. Second, PowerSpy infers driving paths, while SLIC identifies more fine-grained walking paths, even distinguishing across different indoor paths.

Similarly, Shaik et al. [3] demonstrate that in LTE, an adversary can determine whether a target user is in the same cell by probing the targeted user's applications or through silent voice calls. Yet, the attack can only infer whether a target is within the primary cell whereas SLIC can achieve much fine-grained location information.

IMSI catchers [46] can be also used to track cellular users at the cell granularity after the international mobile subscriber identity (IMSI) of a target user is learned via a fake base station. In contrast, SLIC does not require any fake base station and achieves much finer-grained location inference.

Location-privacy attacks exploiting sensors. Researchers proposed side-channel attacks exploiting sensors on a smartphone to infer the victim's location. Han et al. proposes *ACComplice* which utilizes a smartphone's accelerometer to infer the victim's driving routes as well as its starting point [47]. Narain et al. extends the work to also incorporate gyroscope and magnetometer sensors and correctly identify driving paths across 11 cities [48]. Ho et al. relies solely on the barometer of a smartphone to track a car's driving route [49]. However, all of these works requires installing malware on the victim's phone. SLIC inherently removes such constraint, rendering its attack more stealthy.

11 Conclusion

Mobile phone location is undoubtedly highly sensitive information. Our society has reached a strong consensus that phone location data must be handled with extreme care; see the heated discussion regarding contact tracing in the midst of a pandemic. SLIC demonstrates that such highly sensitive phone location information can be leaked to an unauthorized adversary stealthily through benign-looking scheduling metadata in any modern cellular networks. Worse yet, the risk of location information leakage is only expected to grow as cellular networks utilize more frequency spectrum with small cells using heterogeneous physical-layer technologies.

Acknowledgment

We thank our shepherd and the anonymous reviewers for their helpful feedback. We also thank all the participants who helped us collect data. This research was partially supported by Institute for Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No.2019-0-01343, Regional strategic industry convergence security core talent training business).

References

[1] Global Mobile Suppliers Association. Global LTE Subscriptions Forecast to 2020. https://bit.ly/35vSgiI.

[2] 3GPP. TS 33.401 v13.0: 3GPP System Architecture Evolution (SAE), 2016.

[3] Altaf Shaik, Ravishankar Borgaonkar, N. Asokan, Valtteri Niemi, and Jean-Pierre Seifert. Practical Attacks Against Privacy and Availability in 4G/LTE Mobile Communication Systems. In *Network and Distributed System Security Symposium*, 2016.

[4] Yan Michalevsky, Aaron Schulman, Gunaa Arumugam Veerapandian, Dan Boneh, and Gabi Nakibly. PowerSpy: Location Tracking Using Mobile Device Power Analysis. In *USENIX Security*, 2015.

[5] 3GPP. TS 36.321 v13.0: Medium Access Control (MAC) Protocol Specification, 2016.

[6] srsLTE. Open-Source SDR LTE Software Suite from Software Radio Systems (SRS).

[7] Jeffrey G. Andrews, Stefano Buzzi, Wan Choi, Stephen V. Hanly, Angel Lozano, Anthony C. K. Soong, and Jianzhong Charlie Zhang. What Will 5G Be? In *IEEE JSAC Special Issue On 5G Wireless Communication Systems*, 2014.

[8] 3GPP. TR 36.889: Feasibility Study on Licensed-Assisted Access to Unlicensed Spectrum, 2015.

[9] Intel. Alternative LTE Solutions in Unlicensed Spectrum: Overview of LWA, LTE-LAA and Beyond. Technical report, 2016.

[10] 3GPP. TS 36.101 v13.0: Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception, 2016.

[11] 3GPP. Carrier Aggregation Explained, 2013. https://bit.ly/3bWdYNW.

[12] 3GPP. Evolution of LTE in Release 13, 2015. https://www.3gpp.org/news-events/1628-rel13.

[13] GadgetGuy. Samsung Galaxy S20 Ultra 5G – 2020's Best Google Android Phone, 2020. https://bit.ly/2RpWb8l.

[14] Engadget. Samsung's Latest LTE Modem Supports Faster-Than-Fiber Speeds, 2017. https://engt.co/2Fu3rO9.

[15] 3GPP. TS 36.331 v13.0: Radio Resource Control (RRC) Protocol Specification, 2016.

[16] Ettus. USRP B210. https://bit.ly/3huMHn4.

[17] Arun Prabhudesai. Video Viewing Consumes About 60% of Total Mobile Data Traffic in India, 2017. https://bit.ly/3bXHYJa.

[18] Sandvine Intelligent Broadband Networks. Global Internet Phenomena, 2016. https://bit.ly/2Rl1wOi.

[19] Business Insider. 5G is Generating Demand for Korea Telecom's Unlimited Plans, 2019. https://bit.ly/33pgRmL.

[20] Ken's Tech Tips. Vodafone Unlimited Data: Unlimited, Max and Lite Plans With No Download Limits, 2019. https://bit.ly/3bUqa1K.

[21] The Washington Post. How did YouTube Become the Most Popular Music Streaming Site? By Sounding Like the World Itself, 2019. https://wapo.st/2DW3Wjh.

[22] Think with Google. 40 Years After Video Killed the Radio Star, YouTube is Transforming Music. Here's What Brands Need to Know., 2018. https://bit.ly/2RpZBrY.

[23] Mashable SE Asia. People Watch Netflix While Walking, apparently, and the Company Wants Their Data, 2019. https://bit.ly/35x4BmP.

[24] 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 *Network and Distributed Systems Security*, 2019.

[25] ZDNet. Augmented Reality Invades the Conference Room, 2018. https://zd.net/3bYWTmK.

[26] Ericsson. You Need to See Why 5G is the Future of AR Gaming, 2019. https://bit.ly/3iz2brC.

[27] Yuanjie Li, Chunyi Peng, Zengwen Yuan, Jiayao Li, Haotian Deng, and Tao Wang. Mobileinsight: Extracting and Analyzing Cellular Network Information on Smartphones. In *ACM Conference on Mobile Computing and Networking*, 2016.

[28] Matlab. Matlab LTE Toolbox.

[29] Meinard Müller. *Information Retrieval for Music and Motion*. Springer, 2007.

[30] GSA. 5G Market Status: Snapshot January, 2020. https://bit.ly/35wPFF4.

[31] ns 3. 5G-LENA Simulator.

[32] Omnet. OMNet++ - Simulation Models and Tools.

[33] Opnet. Opnet Simulator.

[34] AFAR Communications. FCC Rules for Unlicensed Wireless Equipment Operating in the ISM bands. https://afar.net/tutorials/fcc-rules/.

[35] Digital Guide IONOS. CSMA/CA: Definition and Explanation of the Method, 2019.

[36] Hwan-Joon Kwon, Jeongho Jeon, Abhijeet Bhorkar, Qiaoyang Ye, Hiroki Harada, Yu Jiang, Liu Liu, Satoshi Nagata, Boon Loong Ng, Thomas Novlan, Jinyoung Oh, and Wang Yi. Licensed-Assisted Access to Unlicensed Spectrum in LTE Release 13. In *IEEE Communications Magazine on LTE Evolution*, 2016.

[37] 3GPP. TS 36.213 v13.0: Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures, 2016.

[38] Maravedis. WiFi/3GPP Relations Thaw Over LTE-LAA but not LTE-U, 2015. https://bit.ly/35xAXh7.

[39] Aruba. Wi-Fi First, Hewlett Packard Enterprise's Vision For Mobile Connectivity and LTE in Unlicensed Spectrum. Technical report.

[40] David Rupprecht, Katharina Kohls, Thorsten Holz, and Christina Pöpper. Breaking LTE on Layer Two. In *IEEE Symposium on Security and Privacy*, 2019.

[41] Fan-Hsun Tseng, Han-chieh Chao, Jin Wang, et al. Ultra-Dense Small Cell Planning Using Cognitive Radio Network Toward 5G. *IEEE Wireless Communications*, 2015.

[42] Thomas David Novlan, Radha Krishna Ganti, Arunabha Ghosh, and Jeffrey G Andrews. Analytical Evaluation of Fractional Frequency Reuse for OFDMA Cellular Networks. *IEEE Transactions on wireless communications*, 2011.

[43] Ran Zhang, Miao Wang, Lin X Cai, Zhongming Zheng, Xuemin Shen, and Liang-Liang Xie. LTE-Unlicensed: The Future of Spectrum Aggregation for Cellular Networks. *IEEE Wireless Communications*, 2015.

[44] Min Suk Kang and Bang Chul Jung. Decentralized Intercell Interference Coordination in Uplink Cellular Networks Using Adaptive Sub-band Exclusion. In *IEEE Wireless Communications and Networking Conference*, 2009.

[45] Sami Tabbane. LTE Planning and Dimensioning. In *Workshop on Mobile network planning and security*, 2019.

[46] Stig Mjølsnes and Ruxandra Olimid. Easy 4G/LTE IMSI Catchers for Non-Programmers. In *Computer Network Security*, 2017.

[47] Jun Han, Emmanuel Owusu, Le T. Nguyen, Adrian Perrig, and Joy Zhang. ACComplice: Location Inference using Accelerometers on Smartphones. In *Communication Systems and Networks and Workshops*, 2012.

[48] Sashank Narain, Triet D. Vo-Huu, Kenneth Block, and Guevara Noubir. Inferring User Routes and Locations Using Zero-Permission Mobile Sensors. In *IEEE Symposium on Security and Privacy*, 2016.

[49] Bo-Jhang Ho, Paul Martin, Prashanth Swaminathan, and Mani Srivastava. From Pressure to Path: Barometer-based Vehicle Tracking. In *ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015.

[50] Farhana Afroz, Ramprasad Subramanian, Roshanak Heidary, Kumbesan Sandrasegaran, and Solaiman Ahmed. SINR, RSRP, RSSI and RSRQ Measurement in Long Term Evolution Networks. In *International Journal of Wireless & Mobile Networks*, 2015.

[51] Matthew Gast. 802.11 Wireless Networks: The Definitive Guide, 2005. https://bit.ly/35zmmSz.

[52] 3GPP. TS 36.133 v13.0: Requirements For Support of Radio Resource Management, 2016.

[53] Kworb. Most Viewed Music Videos of All Time, 2020. https://kworb.net/youtube/topvideos.html.

[54] CISCO. Radio Resource Management: Transmit Power Control (TPC) Algorithm. Technical report, 2016.

A Design and Implementation of Wi-Fi-to-5G Evaluation Framework

We first explain the three stages in the Wi-Fi-5G evaluation framework that applies the selection algorithm on the Wi-Fi measurements to obtain the number of configured secondary cells (§A.1). Then, we describe the secondary cell selection algorithm as given in the specification (§A.2) followed by its implementation (§A.3).

A.1 Design of Wi-Fi-to-5G Evaluation Framework

Figure 8 depicts the three stages in Wi-Fi-to-5G evaluation framework.

❶ **Converting Wi-Fi signals to 5G signals.** Even though Wi-Fi and 5G LAA secondary cells share many physical-layer characteristics, there still exist some minor differences that need to be taken into account. For example, we measure the Received Signal Strength Indicator (RSSI) value of the Wi-Fi signals whereas the cellular system measures Reference Signal Receive Power (RSRP) [50], which is a more fine-grained measurement than RSSI. These measurements are made on the beacons or reference signals (i.e., discovery reference signal) transmitted by the Wi-Fi AP and primary cell respectively. However, the periodicity of the transmission of beacons and reference signals is different. On one hand, Wi-Fi transmits beacons approximately every 100 milliseconds [51], whereas the reference signals can be sent every 40/80/120 milliseconds [36]. The limited periodicity of reference signals in Wi-Fi systems leads to coarse-grain measurement. This may degrade the quality of the side-channel information and thus offers the lower-bound performance for SLIC.

❷ **Implementing the secondary cell configuration algorithms.** Unlike 4.5G/5G networks, only a single access point (AP) is assigned to a user at a time in Wi-Fi. To obtain the set of configured secondary cells that can be aggregated for concurrent downlink transmission, we execute the secondary cell configuration algorithms in the specification (Section 5.5.4 in TS 36.331 [15]). The algorithm has the following steps: (1) A primary cell assigns an initial set of secondary cell configuration algorithms and its parameters when a UE first connects to it; (2) the UE measures the secondary cells based on the parameters set in (1) and sends the measurement reports based on the configuration algorithms; (3) Based on the report, the primary cell takes the determined actions, such as selecting a new secondary cell, and informs the UE along with the updated set of algorithms if required; (4) Repeat (2)–(3).

❸ **Determining activated secondary cells.** In this final stage of the evaluation framework, we determine the activated secondary cells based on the user's downlink demand. All the configured secondary cells will be activated if there is a backlog in the UE's downlink. However, if a user's downlink demand is limited the cellular network will activate a fewer number of secondary cells.

A.2 Secondary Cell Selection Algorithm

We reproduce the secondary cell selection algorithm to emulate the behavior observed in traces captured from the real network. The primary cell selects the secondary cells for each user based on the data it receives from the UE. To receive this data, the primary cell first informs the UE about the configuration parameters such as type of measurements, reporting period, configuration parameters, list of frequencies, etc., which are required by the cell selection algorithm [15]. The UE sends measurement reports consisting of radio quality of cells often indicated by RSRP or Reference Signal Received Quality (RSRQ) values. The primary cell uses these reports to select an optimal set of secondary cells, which together with the primary cell form the serving cells.

What to measure. There are two broad categories of radio quality measurements performed by the UE:

1) *Serving and intra-frequency measurements* where the UE measures the serving cells and the neighbouring cells in the same frequency as the serving cells. These neighbouring cells are called intra-frequency neighbouring cells.
2) *Inter-frequency measurements* where the UE measures cells in frequencies configured by the primary cell but not present in the current serving cell frequencies. These cells are known as inter-frequency neighbouring cells. The UE measures the inter-frequency neighbouring cells during a measurement gap configured by the primary cell. During this gap, no transmission or reception is scheduled. This enables the UE to switch to different bands and obtain the signal quality of cells in these bands.

When to send the report. The measurement reports can be sent either periodically or when a condition is triggered. We notice from our captured network traces that the majority of the reports are sent only when a condition is triggered. Hence, we implement a cell selection algorithm where the measurement report is sent by the UE only when one of the events configured by the primary cell is triggered. We discuss the four triggering conditions below.

1) `A1` algorithm (Good-serving-cell): The serving secondary cell radio quality becomes better than the absolute threshold. On receiving this report, the primary cell maintains the secondary cell as a serving cell.
2) `A2` algorithm (Bad-serving-cell): The serving cell radio quality becomes worse than the absolute threshold. This is an indication of bad radio conditions for the secondary serving cell and the primary cell will eventually remove the secondary cell.
3) `A4` algorithm (Good-inter-neighbour-cell): The inter-frequency neighbouring cell radio quality becomes better than the absolute threshold. This will lead to the primary cell adding the inter-frequency neighboring cell to the list of serving cells.
4) `A6` algorithm (Good-intra-neighbour-cell): The report is triggered if an intra-frequency cell's radio quality becomes *offset* better than the current serving secondary cell. This will lead to the removal of the current serving secondary cell and the addition of the better intra-frequency neighbour cell into the serving set.

Configuration. Apart from the thresholds and offsets set by the primary cell, two additional parameters are also configured to avert triggering of unnecessary reports due to sudden fluctuation in the signal strength of cells.

1) *Time-to-Trigger* parameter is the duration of time for which the triggering condition needs to be met to trigger a measurement report.
2) *Hysteresis* is a delta value which makes sure that the measured signal strength is actually better (or worse).

Table 3: Symbol Table

Symbol	Meaning
CID	Cell ID
f	Frequency
$Thres$	Threshold configuration parameter
Hys	Hysteresis configuration parameter
TTT	Time-to-Trigger
$Num_{Serving}$	Number of serving cells
$Config$	Configuration parameters for the triggers
$TimeElapsed$	Time elapsed since the last measurement

A.3 Implementation

We implement a time-driven emulator that emulates the event triggers on the UE and the secondary cell selection in the primary cell. In the emulator, a UE measures the serving and intra-frequency cells every 100 milliseconds and applies Algorithm 1 to checks if the new measurements trigger `A1`, `A2`, or `A6` event. For inter-frequency measurement, our measurement gap is scheduled every 5 seconds to collect inter-frequency measurements and apply Algorithm 2 to check for `A4` event. We set intra-frequency and inter-frequency measurement period to be 800 milliseconds and 1000 milliseconds as per the specification [52] and hysteresis value as two based on the observation in the real network traces. In case an event is triggered, the measurement reports are sent back to the primary cell. The primary cell takes these measurement reports as input for the selection algorithm and outputs the number of configured secondary cells for the UE at that instant.

The Wi-Fi measurement consists of the Wi-Fi MAC, the measured RSSI values, and the frequencies observed by the smartphone every 100 milliseconds. Each MAC address is a unique LAA cell ID, with the RSSI values emulating the RSRP values. Finally, the frequencies of the Wi-Fi APs are assumed to be the operating frequency of the emulated LAA cells. The measurement reports are further divided into serving, intra-frequency, and inter-frequency neighbouring cells. At the beginning of the emulation, we select one Wi-Fi AP from each unique frequency to be the serving secondary cell and the remaining as classified as intra- or inter-frequency neighbours. We then execute our selection algorithm to update the set of serving secondary cells with the configuration parameters set to maximize UE performance.

Algorithm 1 presents the implementation of the `A1`, `A2`, `A6` algorithm. The algorithm takes as input the signal quality of serving and intra-frequency cells, denoted by $Measurement$, the configuration parameters denoted as $Config$, $Num_{Serving}$ which is the total number of serving cells at time $t-1$ and the time elapsed since the previous measurement, denoted as $TimeElapsed$. We apply the algorithm on each of the serving and intra-frequency cell measurements. $Num_{Serving}$ is updated based on the triggered events. Finally, the algorithm outputs the $Num_{Serving}$ at time t and the set of updated serving cells. Algorithm 2 presents the implementation of the `A4` algorithm. Table 3 describes the symbols used in the algorithm.

Algorithm 1: Intra-Frequency Measurement Algorithm

```
Function Intra-Frequency-Measurement-Algorithm(Config, Measurement, Num_Serving, TimeElapsed):
    [CID_Serving, RSSI_Serving, f_Serving] ← getServingCell(Measurement)
    [CID_Intra, RSSI_Intra, f_Intra] ← getIntraFreqCell(Measurement)
    foreach SCell in CID_Serving do
        if RSSI_SCell > A1_Thres + A1_Hys then
            if A1_TTT ≤ 0 then
                TRIGGER A1
            else
                A1_TTT ← A1_TTT - TimeElapsed
            end
        else if RSSI_SCell < A2_Thres − A2_Hys then
            if A2_TTT ≤ 0 then
                TRIGGER A2
                DeleteServingCell(CID_SCell)
                Num_Serving ← Num_Serving − 1
            else
                A2_TTT ← A2_TTT - TimeElapsed
            end
        end
        [RSSI_MaxNeigh, CID_MaxNeigh] ← getBestIntraNeigh(f_Intra, f_SCell)
        if RSSI_MaxNeigh > RSSI_SCell + A6_Off + A6_Hys then
            if A6_TTT ≤ 0 then
                TRIGGER A6
                DeleteServingCell(CID_SCell)
                AddServingCell(CID_MaxNeigh)
            else
                A6_TTT ← A6_TTT - TimeElapsed
            end
        end
    end
    return Num_Serving
```

B CA Activation at a Low Rate

We empirically show that a low data rate application can still trigger up to three CA. We collect fingerprints of a 250-meter outdoor path using Sony Xperia XZ1 phone while streaming YouTube music videos at a 360p resolution. We stream the five most popular videos on YouTube in 2019 and 2020, respectively [53] and show the percentages of one CA, two CA, and three CA activations for all ten videos in Figure 16. We observe an average data rate of 15Mb/s while streaming the ten videos indicating that a moderate data rate triggers three CA as well.

C Effect of Network Load on CA Activation

We empirically show that the CA activation is strongly dependent on the UE location and less on the load of the cellular

Algorithm 2: Inter-Frequency Measurement Algorithm

```
Function
  Inter-Frequency-Measurement-Algorithm(Config,
  Measurement, Num_Serving, TimeElapsed):
    [CID_Inter, RSSI_Inter, f_Inter] ←
      getInterFreqCell(Measurement)
    f_Current ← SelectInterFrequencyToMeasure(f_Inter)
    [RSSI_MaxNeigh, CID_MaxNeigh] ←
      getBestInterNeigh(f_Current)
    if RSSI_MaxNeigh > A4_Thres + A4_Hys then
        if A4_TTT ≤ 0 then
            TRIGGER A4
            AddServingCell(CID_MaxNeigh)
            Num_Serving ← Num_Serving + 1
        else
            A4_TTT ← A4_TTT - TimeElapsed
        end
    end
    return Num_Serving
```

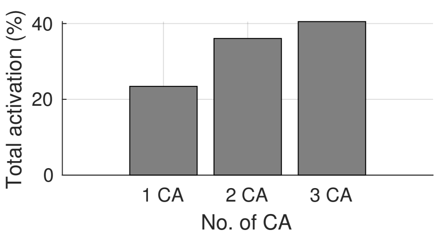


Figure 16: CA activation while streaming ten 360p YouTube music videos.

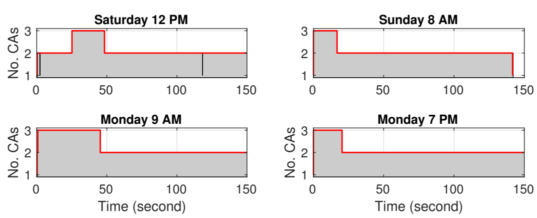
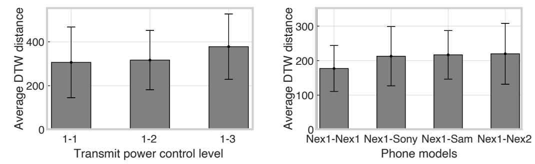


Figure 17: CA configuration and activation observed in four walks on the same path.

system by testing on another LTE-A network in a metropolitan city. We collect four fingerprints of a single 260-meter outdoor pedestrian path using a Sony Xperia XZ1 phone connected to the LTE-A network that supports up to three CA. We particularly choose three different times of the day (i.e., morning, afternoon, and evening) across two days to reflect different network load experienced by the LTE-A network [17, 18]. We run the MobileInsight [27] app while simultaneously downloading files at the maximum rate, to activate a maximum number of secondary cells. Figure 17 shows the CA configuration and activation for four walks on the same path. The plot shows that fingerprints have a consistent pattern across different walks at different times. This indicates that the load of the cellular network has minimal effect on the CA activation.



(a) Average DTW distance for TPC values (b) Average DTW distance for phones

Figure 18: (a) Average DTW distance across three TPC power levels. (b) Average DTW distance between fingerprints collected with Nexus1 and other phones (Nexus2, Sony, and Samsung).

D Perturbation in Received Signal Strengths

RSSI can vary at times due to factors such as transmit power control (TPC), antenna properties, mobile phone models, etc. One question, therefore, is *"How reliable would the fingerprinting mechanism be when RSSI values change across walks?"* From our experiment in the Wi-Fi environment, we show that minor perturbations in the RSSI values do not cause any major difference in the side-channel measurement.

We show the reliability of the fingerprints in the presence of TPC by conducting an experiment in our Wi-Fi environment that implements the Cisco TPC algorithm [54]. Here, we consider the Wi-Fi TPC algorithm to be similar to the algorithm used by the LAA secondary cells(see §6). We retrieve the TPC power level logs of every Wi-Fi AP deployed on a fixed path in our building and collect the path fingerprints every time the TPC power level changes. Specifically, we collect ten fingerprints each when the TPC power level of a Wi-Fi AP on the path is set as 1 (maximum power), 2 (maximum power-3 dB), and 3 (maximum power-6 dB). We then compute the average DTW distance between fingerprints collected when TPC is 1 with fingerprints collected when TPC is 2, and 3 and show it in Figure 18(a). The similar average DTW distance across power levels indicates the fingerprints are similar even in the presence of TPC. Thus, we empirically show that the fingerprints are reliable and have minimum impact.

Similarly, to show the reliability of the fingerprints across different phones, we collect five fingerprints of a fixed path while holding four phones, namely, Sony Xperia Z5, Samsung Galaxy S6, and two Nexus 6 phones (Nexus1 and Nexus2). Figure 18(b) shows the average DTW distance between fingerprints collected using one Nexus 6 (Nexus1) and other phones. The plot shows similar average DTW distances between Nexus1 and the other three phones indicating the fingerprints will be similar across different phone models and hence do not cast a consequential influence on the fingerprint. The aforementioned observation is not surprising because the side channel the SLIC adversaries observe is highly aggregated information that has *lost* most of the rich details of the raw RSSI measurements but contains only the coarse-grain trends. Our attack evaluation in Section 7 shows that this coarse-grained side channel is still highly effective for path identification.