2017

Location privacy in LTE: a case study on exploiting the cellular signaling plane's timing advance

Roth, John D.; Tummala, Murali; McEachen, John C.; Scrofani, James W.

HCSS

J.D. Roth, M. Tummala, J.C. McEachen, J.W. Scrofani, "Location privacy in LTE: a case study on exploiting the cellular signaling plane's timing advance," Proceedings of the 50th Hawaii Conference on System Science, 2017, pp. 6285-6292.
https://hdl.handle.net/10945/55165

Location Privacy in LTE: A Case Study on Exploiting the Cellular Signaling Plane's Timing Advance

John D. Roth*, Murali Tummala†, John C. McEachen†, and James W. Scrofani†
*Department of Electrical and Computer Engineering
United States Naval Academy
Annapolis, MD, USA
Email: jroth@usna.edu
†Department of Electrical and Computer Engineering
Naval Postgraduate School
Monterey, CA, USA

***Abstract*—Location privacy is an oft-overlooked, but exceedingly important niche of the overall privacy macrocosm. An ambition of this work is to raise awareness of concerns relating to location privacy in cellular networks. To this end, we will demonstrate how user location information is leaked through a vulnerability, *viz*. the timing advance (TA) parameter, in the Long Term Evolution (LTE) signaling plane and how the position estimate that results from that parameter can be refined through a previously introduced method called Cellular Synchronization Assisted Refinement (CeSAR) [1]. With CeSAR, positioning accuracies that meet or exceed the FCC's E-911 mandate are possible making CeSAR simultaneously a candidate technology for meeting the FCC's wireless localization requirements and a demonstration of the alarming level of location information sent over the air. We also introduce a geographically diverse data set of TAs collected from actual LTE network implementations utilizing different cell phone chipsets. With this data set we show the appropriateness of modeling the error associated with a TA as normally distributed.**

1. Introduction

With the numbers of Long Term Evolution (LTE) subcribers projected to increase from 1.1 billion in 2015 to 4.3 billion over the next five years [2], this leading cellular technology has never been so nascent a worldwide social force. Additionally, not only are the number of LTE connections increasing, but the frequency in which the individual uses the cellular link is also increasing. Currently, in North America, each subscriber uses about 3.7 gigabytes a month. Over the next five years that number is expected to increase to 22 gigabytes a month [2]. Further, LTE-Advanced (a series of proposed enhancements to LTE such as carrier aggregation, small cells, and MIMO [3]) connections are projected to grow to 500 million by 2018 making the once lofty 3Gb/s wireless links a global reality. LTE is poised to deliver unprecedented improvements in quality of life and worldwide productivity. However, as we move towards these principled ends, we must take care with how we integrate this technology into our social fabric.

Protecting the cellular user's privacy is of special importance. The LTE standard has made great strides in protecting the confidentiality of the user's data plane over older technologies, such as the Universal Mobile Telecommunications System (UMTS), through means such as mutual authentication between the user equipment (UE) and the cellular infrastructure [4]. However, other architectural shifts have left less obvious parts of the radio link, such as the signaling plane, vulnerable to location privacy attacks [1].

Location privacy has received a significant amount of attention in the literature over the last several years [5]–[10]. The seminal definition of location privacy is given by Westin as

> ...the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others [11].

This definition subsumes many modern ideas associated with the use of location information. For instance, anonymized datasets are sometimes recorded in third party servers for the purpose of sociological and market studies, optimal cell tower placement, or traffic monitoring [6]. However, it has been well-known in the scientific community that anonymous location data can be attributed back to specific individuals with remarkable accuracy through computational means such as Markov modeling [6].

More alarming than the knowledge of this possibility is the public's apparent indifference or ignorance to threats to their location privacy. One study reported that 250 users willingly surrendered two weeks of their driving GPS data in return for a 1 in 100 chance of winning a US $200 MP3 player. Moreover, of those 250 individuals, 97 were asked if their data could be shared with third parties and only 20% declined [8]. This general sentiment is indicative of public attitudes toward location privacy and can be found in numerous other studies [12]–[14].

URI: http://hdl.handle.net/10125/41925
ISBN: 978-0-9981331-0-2

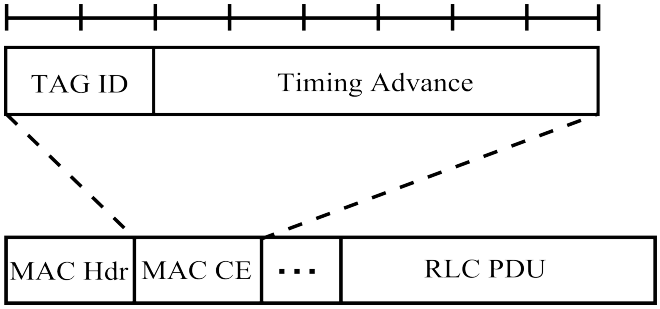


Figure 1: The TA is part of a MAC layer (which encapsulates the radio link control layer protocol data unit) control element [15]. During maintenance the TA is a six bit quantity. LTE release 11+ includes a TA group (TAG) field in order to support multiple TAs from multiple eNBs.

A leading ambition of this work is to raise awareness of concerns relating to location privacy in cellular networks. To this end, we intend to demonstrate how user location information is leaked through a vulnerability, *viz*. the timing advance (TA) parameter, in the LTE signaling plane and how the position estimate that results from that parameter can be refined through a previously introduced method called Cellular Synchronization Assisted Refinement (CeSAR) [1]. Our findings are validated through localization case studies conducted in real-world LTE network deployments.

The remainder of this paper is organized as follows. In Section 2 some preliminaries and theory regarding the use of TA in UE positioning are discussed. In Section 3 we profile the TA behavior in the wild. Section 4 discusses the attack framework, experimental setup, and presents the results of the experiments. Finally, related work and conclusions are presented in Sections 5 and 6 respectively.

2. The LTE Signaling Plane

In this section we illuminate the portion of the signaling plane we will exploit. A full review of the signaling plane is beyond the scope of this paper; however, we refer the interested reader to [4] and [15] for a more in-depth treatment. Our discussion will highlight the TA and how it can be used to provide a UE location, with and without CeSAR. Finally, we discuss the architectural security shift in LTE that makes these attacks possible.

2.1. Frame Timing Management in LTE

LTE manages medium multiple access through orthogonal frequency division multiple access (OFDMA). OFDMA requires that uplink frames arrive at the cell tower, or enhanced-node B (eNB), at the time in which a particular user is scheduled. Any deviation from the scheduled frame arrival time can result in inter-symbol interference which significantly degrades the wireless link. Because user mobility is a inseparable attribute of cellular networks, it is clear that the propagation delay between the UE and the eNB will not be constant. Therefore, in order to incorporate propagation delay into scheduling, and thus prevent inter-symbol interference, a TA parameter is included as a control element (CE) in the medium access control (MAC) layer (cf. Figure 1). This network controlled parameter modifies the UE's uplink burst timing such that it takes into account and adjusts for the propagation delay [15].

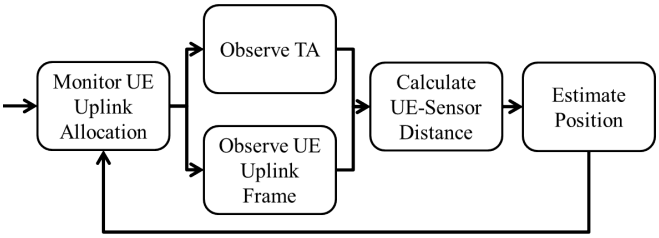


Figure 2: The CeSAR method [1] is a completely passive enhancement to TA-based positioning which can be performed with a simple software defined radio implementation.

During normal maintenance of the wireless link, the TA is a six-bit quantity where each bit represents $16{\times}T_s$ seconds and T_s is the sampling period computed as

$$T_s = (\Delta f \times 2048)^{-1}. \tag{1}$$

Here Δf is the subcarrier spacing (nominally 15kHz) and 2048 is the maximum Fast Fourier Transform size [16], [17].

2.2. CeSAR

CeSAR is a previously introduced and completely passive enhancement to TA-based positioning that has been shown in simulation to improve position estimates by up to 250 meters [1]. The method requires a sensor in the same cell as the UE. The sensor need not be complex, for instance, it could be implemented with a simple processor and a software defined radio. The method also requires *a priori* knowledge of the location of the serving eNB(s) and the sensor. The general steps for the method are outlined here for completeness and are shown graphically in Figure 2. Further implementation details can be found in [1].

1) Monitor UE uplink allocation.
2) Observe TA sent from the eNB(s) to the UE (this step is interchangeable with step 3).
3) Observe UE uplink burst (this step is interchangeable with step 2).
4) Calculate UE-sensor distance.
5) Estimate UE position.

The contribution of CeSAR is essentially an additional equation added to the total system of equations described by the TA annulus(i) which describes the UE-sensor distance.

2.3. Localization Attacks with Timing Advance and CeSAR

Once the TA is known the eNB-UE distance, $\hat{d}$, can be estimated via

$$\hat{d} = T_A \times T_s \left(\frac{16c}{2}\right) \tag{2}$$

where c is the speed of light and $T_A \in [0, \ldots, 63]$ is the binary TA value found in the MAC CE [1]. A straightforward result of (2) is that each TA increment represents 78.125 meters of distance.

The quality of the resulting estimate is a function of the following conditions.

1) The frequency of TA issuance.
2) The number of connected eNBs.
3) The eNB(s)-sensor-UE geometry.
4) The channel quality.

The frequency of TA issuance is lower bounded by the `timeAlignmentTimer` [18]. This is a configurable, implementation specific, parameter which ranges from 500ms to 10s. We have found that the `timeAlignmentTimer` value differs by network provider, but that in practice TAs can be issued much more frequently than specified by the `timeAlignmentTimer` and can be part of the MAC header for nearly every packet sent to the UE from the serving eNB. Further, an active connection is not needed in order to be issued a TA. In practice, TAs are issued periodically even when the UE is in `RRC IDLE` mode. It is straightforward that more frequent TAs provide a better position estimate. Frequent TAs will allow position estimation to recover from outliers more quickly and provide more granularity when tracking.

If the distribution of error, $p(\epsilon|T_A)$, is known then the expected positioning performance can be described via the Cramer-Rao Lower Bound (CRLB) [19], given by

$$\text{var}(\hat{\mathbf{p}}) \geq \sqrt{\text{trace}(\mathbf{I}^{-1})} \tag{3}$$

where $\hat{\mathbf{p}} = [\hat{x}, \hat{y}]^T$ is the position estimate and $\mathbf{I}$ is the Fisher information matrix (FIM). If $p(\epsilon|T_A)$ is approximated as normal ($\approx \mathcal{N}(0, \sigma)$) then it can be shown that

$$\mathbf{I} = \begin{bmatrix} \sum_{i=1}^{N} \frac{(x-x_i)^2}{\sigma_i^2 d_i^2} & \sum_{i=1}^{N} \frac{(x-x_i)(y-y_i)}{\sigma_i^2 d_i^2} \\ \sum_{i=1}^{N} \frac{(x-x_i)(y-y_i)}{\sigma_i^2 d_i^2} & \sum_{i=1}^{N} \frac{(y-y_i)^2}{\sigma_i^2 d_i^2} \end{bmatrix}. \tag{4}$$

A result of (3) and (4) is that as the CRLB varies inversely with the number of eNBs, N. Further, channel conditions can be represented in (4) with σ and geometry is represented in the numerators of the elements of the FIM.

In summary, the best possible environment for positioning would be one in which the UE is connected to multiple eNBs ($N >> 1$), the connected eNBs are not collinear to the UE (more specifically that the UE is within the convex hull of the eNBs), the UE is close to and has a line of sight to the connected eNBs (thus improving channel conditions), and TA issuance is sufficiently frequent. In line with the first observation about relationship between N and $\hat{p}$, since $N_{CeSAR} = N_{TA}+1$ it can be seen from (3) that CeSAR will, in general, improve positioning performance. The expected magnitude of improvement is explored for various scenarios with various N in [1].

2.4. Resolution of Inconsistent Equations

The channel conditions play a significant role in the accuracy of TA. A common channel model for positioning applications is

$$\hat{\mathbf{d}} = \mathbf{d} + \lambda + \eta \tag{5}$$

where $\hat{\mathbf{d}}$ is the set of observed transmitter(Tx)-receiver(Rx) distances, $\mathbf{d}$ are the true distances, λ is a random vector representing the non-line of sight (NLoS) error associated with the signal traveling a non-minimal distance between the Tx-Rx, and η is a normally distributed measurement error. If $N > 2$ the presence of λ and η guarantee that the system of equations will be strictly inconsistent. If $N \leq 2$ then the system may not have a unique solution (and may also still be inconsistent). Achieving an optimal solution with equations of this type is a well-studied field in the literature [19]. This paper only considers scenarios where $N \geq 2$ so we choose the residual error method [20] of selecting a position estimate

$$\hat{\mathbf{p}} = \min_{\mathbf{p}} \left\{ \sum_{i=1}^{N} \left(\hat{d}_i - \| \mathbf{p} - \mathbf{x}_i \| \right)^2 \right\}. \tag{6}$$

Here $\mathbf{x}_i = [x_i, y_i]^T$ is the location of the *ith* eNB (or CeSAR sensor), and $\| \cdot \|$ is the Euclidean norm. The residual weighting method is well-known and accepted technique for finding a parameter when its error distribution is not well characterized [19].

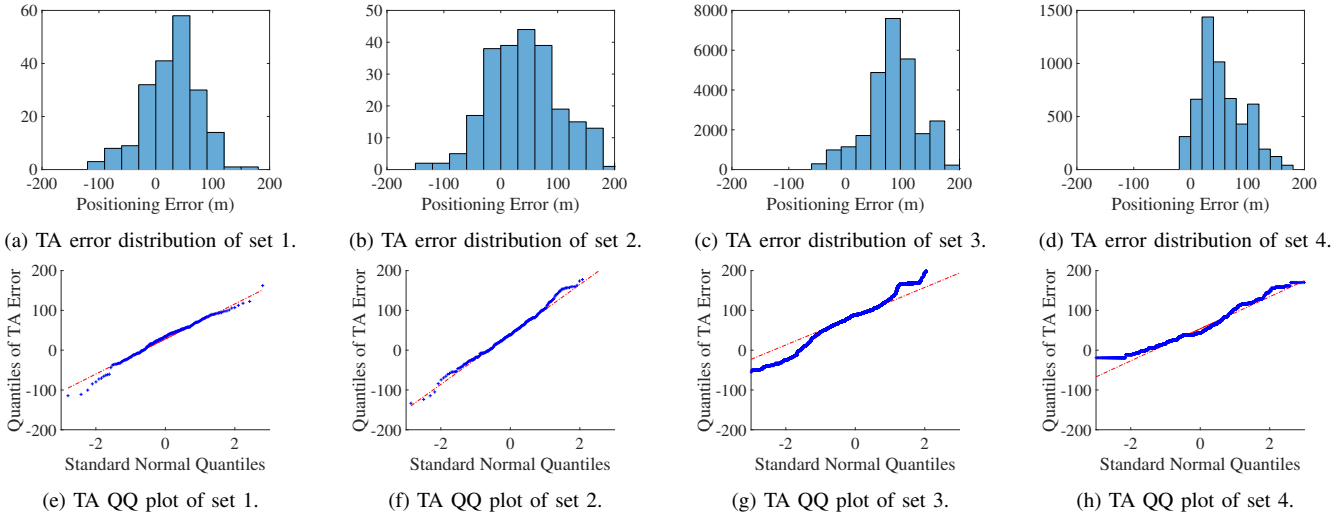
2.5. Timing Advance as a Location Privacy Preserving Mechanism

A LPPM is a method for separating a user, u_i, from that user's location, l, and has two components: obfuscation and anonymization [10]. The act of obfuscating a location will add noise to the actual location, $d' = f_1(\mathbf{p})$, thus an attacker using obfuscated only data, $\langle u_i, d' \rangle$, will have access to user identities, but the associated location data will be imperfect. The act of anonymizing data will replace the user identity with a pseudonym, $u' = f_2(u_i)$, thus an attacker using anonymized only data will have access to exact locations but not identities. A obfuscated and anonymized data set, $\langle u', d' \rangle$, will provide an attacker access to neither piece of information directly.

Formally, the TA can be modeled as a LPPM. The noise added to the data can be modeled with the function

$$d = \| \mathbf{p} - \mathbf{x} \| + \mathcal{U}_{TA} \tag{7}$$

where $\mathcal{U}_{TA}$ denotes a uniform random variable $\in (0, 78.125)$. In other words, The TA obfuscates the actual UE position through a process of spatial quantization. Next, the network anonymizes the UE through assignment of a cell-radio network temporary identifier (C-RNTI) [15]. The C-RNTI is a 16-bit value that uniquely identifies a UE when connected to a specific eNB. The C-RNTI can be thought of as a software address and is assigned dynamically. Therefore, the C-RNTI mapping, $f_{C-RNTI}(\cdot)$, can be thought of as LPPM anonymization.



(a) TA error distribution of set 1. (b) TA error distribution of set 2. (c) TA error distribution of set 3. (d) TA error distribution of set 4.

(e) TA QQ plot of set 1. (f) TA QQ plot of set 2. (g) TA QQ plot of set 3. (h) TA QQ plot of set 4.

Figure 3: The error associated with the TA is presented here along with a QQ plot as a graphical comparison against normally distributed data. All data was collected in suburban environments where a line of sight component is assumed. Tracks 1 and 2 were taken in Monterey, California with the Snapdragon 805 chipset. Tracks 3 and 4 were taken in Annapolis, Maryland with the Snapdragon 801 chipset. All data was collected on the same mobile network carrier.

This LPPM is weak for several reasons. First, the quality of the obfuscation declines rapidly when multiple eNBs are configured. This study will focus primarily on driving this point home. Additionally, cell sectors would also serve to de-obfuscate l'. The quality of anonymity provided by $f_{C-RNTI}(\cdot)$ is also in question [4], [21] although this work will not focus on exploiting this portion of the TA LPPM.

2.6. Confidentiality of the Timing Advance

The TA is a particularly vulnerable parameter since it is sent in clear text on the air interface in LTE. This marks a significant shift in security architecture from earlier technologies like the Global System for Mobile Communications (GSM) where the TA would not have as readily been available. In LTE, the Packet Data Convergence Protocol (PDCP), a layer 2 sublayer, is responsible for ciphering (encryption). Therefore nothing in the lower layers (e.g., the Radio Link Control and MAC layers) is ciphered [22]. Because the MAC CE is the primary bearer of the TA (cf. Figure 1), the TA is sent unencrypted.

2.7. Confidentiality of Uplink Grants

In order for CeSAR to be effective, the sensor must be able to determine what resource elements a UE has been assigned to for uplink (cf. Figure 2). In LTE this information, similar to the TA, is unprotected over the air interface.

LTE has a relatively flat logical channel architecture. The channels are broken into the downlink and uplink subgroups of which the former is of particular interest. In this group there exists a Downlink Control Channel (DCCH) which is a bearer of mainly the Radio Resource Control (RRC) layer information. Also, LTE specifies certain physical channels onto which no logical channel will map. Of interest to this work is the Physical Downlink Control Channel (PDCCH) and the Physical Uplink Control Channel (PUCCH).

Scheduling is the responsibility of the MAC layer and is done dynamically on a frame-by-frame (i.e., 1ms) basis[1] [15]. Therefore, LTE does not assign dedicated control channels. Instead, the information pertaining to uplink scheduling is found in the PDCCH broadcast in the L1/L2 control region of each downlink frame [15].

Consider a UE with information to transmit to the network and without a current valid scheduling grant. The UE will first utilize the uplink L1/L2 control region to indicate to the eNB that it requires uplink resources. As previously discussed, the eNB's scheduling decisions are issued via the PDCCH in the L1/L2 control region. Each scheduling grant is appended with a cyclic redundancy check (CRC) which is calculated with the intended recipient's (or recipients' in the case of multicast) radio network temporary identifier (RNTI). Therefore all grants sent via the PDCCH are checked by each UE with their allocated RNTIs. Grants that do not check out are discarded as either not intended for the UE or invalid [15]. The PDCCH is continuously monitored by each connected UE to update its uplink grant allocation as it is changed dynamically. Therefore, since this information sits below the PDCP it will not be encrypted.

3. Timing Advance Behavior in Modern LTE Networks

In this section we present TA data from three real world LTE cellular network deployments. The first deployment is

[1]It should be noted that the the network can also optionally choose to implement semi-persistent, vice dynamic scheduling.

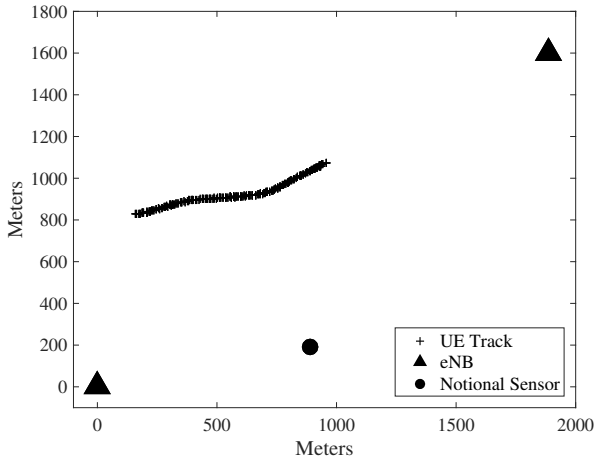


Figure 4: Track 1, used for evaluation of TA-based positioning and CeSAR, is presented in this figure. The track is 830 meters long and includes 323 recorded TA values. Connections were made in this track over the 700 MHz, 1900 MHz, and 2110 MHz bands. Surrounding areas and the exact track location are not presented in order to maintain experimental anonymity.

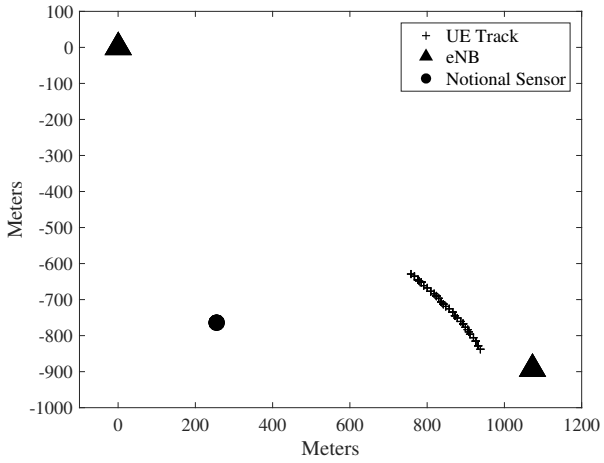


Figure 5: Track 2, used for evaluation of TA-based positioning and CeSAR, is presented in this figure. The track is 277 meters long and includes 73 recorded TA values. Connections were made in this track over the 700 MHz and 2110 MHz bands. Surrounding areas and the exact track location are not presented in order to maintain experimental anonymity.

in Monterey, California and is displayed in Figures 3(a) and 3(b) (set 1 and set 2). These sets represent 197 and 237 TAs respectively and were captured using the Qualcomm Snapdragon 805 chipset. Sets 1 and 2 both utilized the 700 MHz band. Set 3 is shown in Figure 3(c). These data were collected in Annapolis, Maryland and are comprised of 27,200 TAs. Set 4 is shown in Figure 3(d). These data were collected in San Diego, California and are comprised of 5,500 TAs. Sets 3 and 4 were taken with the Qualcomm Snapdragon 801 chipset and utilized the 1900 MHz band.[2]

All the data was collected in suburban settings free from major physical obstructions between the eNB and UE. In tracks 1 and 2 only one building exceeded the typical height of two stories. That structure was six stories tall. Further, track 2 was markedly more rural, however significant variations in terrain height and dense coniferous foliage were noted. Tracks 3 and 4 are similar in layout to track 1.

The most obvious observation is that the error distribution is not uniform as is commonly assumed in TA-based simulations. Rather, the distribution follows some density that is not, strictly speaking, well defined, but is peaked, and unimodal. A comparison of the error distribution with normally distributed data is shown in the QQ plots in Figure 3. Upon inspection, a reasonable fit of the data to the standard normal quantiles is apparent.

To further qualify the goodness of fit to the normal distribution, Pearson's chi-squared test was used. In each of the sets a subset of 50 samples were randomly selected and the p-value was computed with the null hypothesis that the data were drawn from the standard normal distribution. This method was used in order to avoid over sensitivities in the test with large data sets and to provide a standard metric with which to compare the data from sets of a different size. After repeating the above process 10,000 times, sets 1-4 had an average p-value of 0.4097, 0.4227, 0.1587, and 0.1307 respectively. *This suggests that the null hypothesis should not be rejected*. It may therefore be appropriate to approximate the error in suburban TA data as normal. This is especially true when performing a localization attack in which a large number of data are not considered as a whole but rather the working set is limited to a small amount of data at a specific instance in time.

[2]Occasionally, extreme outliers in TA data are noticed. They are not shown in Figure 3 for clarity of presentation.

4. Experimental Validation

In this section we first formalize the attack framework [10]. Next, we describe the experimental setup and the results that follow.

4.1. Attack Framework

Location privacy attacks can be broadly classified into three types of attacks: a meeting disclosure attack, a tracking attack, or a localization attack [10]. In a meeting disclosure attack the adversary is not concerned with the position of the victim in a geodetic sense. Rather, this is a more subtle attack in which the adversary is concerned with learning the nature of the interactions of the victim with other users (e.g., to learn a victim's social network). In a tracking attack the adversary seeks to understand a position history of the victim. Conversely, a localization attack is only concerned with gleaning the current location of the victim. This work

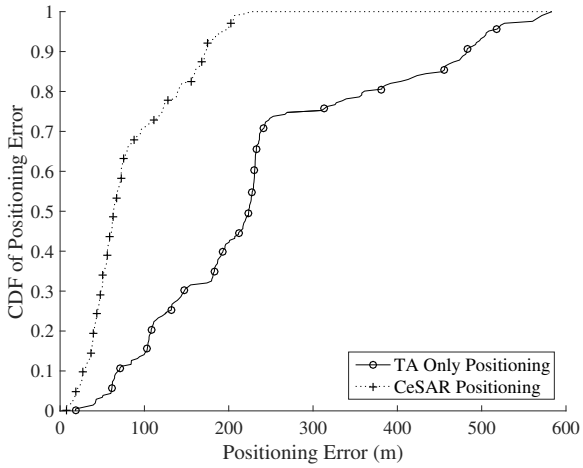


Figure 6: The positioning performance obtained using the track 1 data. TA only positioning is compared to the positioning possible when the TA is augmented with CeSAR.

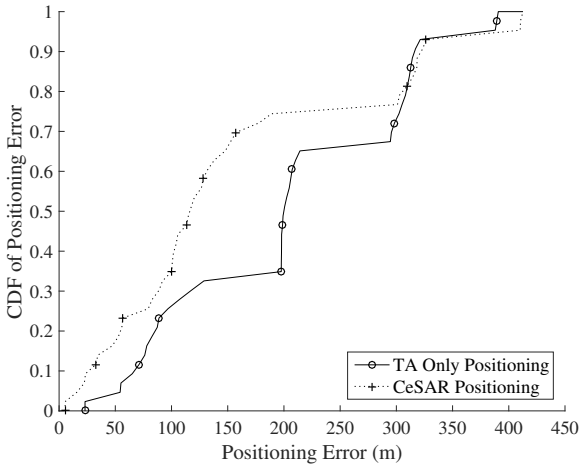


Figure 7: The positioning performance obtained using the track 2 data. TA only positioning is compared to the positioning possible when the TA is augmented with CeSAR.

is primarily concerned with the latter category of attack although the principles of the illustrated vulnerability are applicable to the entire spectrum of attacks.

We assume here that the adversary has real-time access to the information $\langle \mathbf{u}, \mathbf{d}' \rangle$ where $\mathbf{u}$ is the set of attributable user identities and $\mathbf{d}' = f(\mathbf{p})$, where $f(\cdot)$ is modeled as in (7). Therefore perfect knowledge of the identity to C-RNTI mapping is assumed.

4.2. Experimental Setup

In order to quantify the expected positioning performance using the TA parameter and also compare the performance increase possible through CeSAR, we conducted two experiments where a target UE was connected to real network infrastructure in Monterey, California. In each case the UE was driven on a specified track while connected to one neighboring eNB and the TAs issued from that tower were recorded. The UE was then driven a second time through the same track connected to a different neighboring eNB and again the issued TAs were recorded (in the case of track 1 this process was repeated once to increase the data set size). This setup is designed to mimic heterogenous network deployments expected in LTE release 11+ where simultaneous timing management among multiple eNBs will be necessary, thus TAs from multiple eNBs will be available. The setup also mimicks previous suggestions in the literature that TA-based positioning could be improved through forcing an eNB handover [23].

Track 1 and track 2 are shown in Figures 4 and 5 respectively. The surrounding area is not presented in order to preserve experimental anonymity. Track 1 is 277 meters long and includes 73 recorded TA values in the 700 MHz and 2110 MHz bands. Track 2 is 830 meters long and includes 323 recorded TA values with links over the 700 MHz, 1900 MHz, and 2110 MHz bands. Both are conducted in suburban settings free from major physical obstructions between the eNB and UE at approximately 50 kilometers per hour. Position estimates were made *a posteriori* in post-processing.

The CeSAR method was bootstraped into the experiment in post-processing with a notional sensor located as in Figures 4 and 5 to provide reasonable geometry for trilateration. The estimated error in distance resolution from the CeSAR sensor to the UE was modeled as $\mathcal{N}(0, 20)$ where units are given in meters.

As previously discussed in Section 2.4, a position estimate is extracted from the resulting system of equations via the minimization of residual error method [20], given in (6), for both TA only and CeSAR augmented positioning.

4.3. Results

Positioning performance with the TA only method was then evaluated by using the TAs issued from the two different towers to the UE located at a particular location. This process was repeated for each TA pair in the track. Ambiguity resolution in the resulting underdetermined set of equations was assumed in order to provide a more robust benchmark for validating CeSAR.

As expected, TA-based positioning in LTE is more precise than the performance previously theorized in GSM [23], [24]. Using only two TAs an accuracy of 240 meters and 295 meters was found in tracks 1 and 2 respectively in the sense of circular error probable (CEP) 70% [3].

When augmented with CeSAR, TA-based positioning improves to 95 meters and 157 meters in tracks 1 and 2 respectively. This suggests that CeSAR may be able to

[3]CEP 70% is the error upper bounding the realized error 70% of the time. In other words, CEP 70% is X when $\Pr\{\epsilon \leq X\} = 0.7$ and ϵ is a realization of the random variable in question.

provide accuracies on the order of the Federal Communications Commission (FCC) E-911 mandate for network based techniques[4]. In both cases positioning improvement on the order of 150 meters is realized through CeSAR.

5. Related Work

In October of 1994 the United States FCC released a notice of proposed rulemaking [24] requiring cell network providers to locate users who dialed 911 (E-911). Ever since that time there has been significant activity in the field of geolocation in cellular networks. To this end, many techniques were proposed and studied. For example, the IEEE 802.11 standard continues to be a popular enabling technology due to the ubiquity of Wi-Fi. Wi-Fi fingerprinting has been successfully used in indoor positioning, frequently exhibiting room level [26] and even sub-meter accuracy [27].

While the fingerprinting method is popular in indoor and urban environments, the TA has a long and storied application in the literature for outdoor application. In [23], the authors discuss the possibility of using TA as a mechanism for positioning. They note poor accuracy, and suggest forcing base station handover in order to get a second TA to improve positioning. They conclude that the accuracy is not sufficient for TA to be seriously considered by itself as a method for positioning.

Accuracy concerns are echoed in [24] where it is estimated that the accuracy of the GSM TA is theoretically 550 meters and practically 2,200 meters. Nevertheless, it is noted that cell tower location in conjunction with TA is used in many countries around the world as a means for subscriber localization. This is also a "fallback" GSM localization technique in the United States if a subscriber cannot be located with other, more accurate, means.

The authors in [28] suggested taking multiple TA measurements from the same tower and averaging them in order to improve distance estimation. An analysis of the method is presented, but no real-world experimentation was conducted. It was noted that their method will only result in a distance from the cell tower and any further improvement in accuracy will result from other means.

In [29], the authors propose the use of GSM TA for traffic state estimation, not for precise user localization. However, their evaluation oversimplifies the TA behavior in simulation. Again no real-world data is used.

The authors in [30] represent the only study we are aware of that uses actual field recorded TA from a GSM network, although their application was in finding GSM base stations and not user location. Their study was still largely simulation based and they only presented one real-world example.

The largely unsuccessful first forays into using the TA as a parameter for localization are probably to blame for the limited amount of research in GSM TA positioning since initial simulations were not promising. With an accuracy as low as 550 meters to 2.2 kilometers it is not unsurprising that the TA was largely abandoned by the community for a time.

It was not until Jarvis et al. [16] recognized the potential in the TA parameter in LTE networks that researchers reopened their study of the TA as a means to positioning. Although again a simulation only approach, the authors showed viable positioning accuracy in three dimensions when using a TA from three and four eNBs. The authors did not address how using more than one eNB would be possible nor did they assume there was any error associated with the eNB issuing the correct TA to the UE. Similar investigations were conducted using WiMAX technology in [31].

In [32], Wigren uses the LTE TA as a complimentary database feature when performing localization via fingerprinting, a method of comparing *a priori* measurements with real time measurements to improve accuracy. Using a heuristic approach to modeling the behavior of the TA, he noted accuracies on the order of his TA error and suggested his algorithm as an appropriate fallback technology for positioning for E-911 in LTE if Assisted-GPS was not available.

The authors in [33] used LTE TA as a means for proximity discovery in device to device (D2D) communications. They showed through simulation that errors as low as 50 meters were possible for certain eNB geometries. However, their modeling of the TA was also heuristic, and did not account for any error in the eNB issuing an incorrect TA.

The work represented by [34] is the only published work we are aware of that uses actual field measurements to validate TA-based approaches in LTE to positioning. Their approach did not, however, focus on characterizing the TA. Rather, similar to Wigren's approach, they used it as another feature in a fingerprinting approach to localization with the aim of minimizing the cost of training their fingerprint database. They also made no attempt to characterize how the TA value correlated with the true distance of the UE.

In summary the corpus representing the TA parameter in the literature is conspicuously sparse. Even more absent are studies conducted with real-world data thus, making modeling in simulation largely a product of conjecture. We believe this is the first study done with real-world data characterizing the behavior of the TA value. The work here will allow future studies to accurately model TA error and thus explore true maximum likelihood algorithms. We are also the first to formalize this method as a localization attack and apply real-world TA data to to show its efficacy.

6. Conclusions

In this study we have examined the signaling plane of the LTE/LTE-A protocols in the context of a location privacy attack. Specifically, the TA was used as a means to this end. We propose that more should be done to protect the individual user's location privacy as the current protocol implementation almost continually broadcasts a user's location information in the clear.

In developing this argument we first presented the TA as a LPPM in a formal attack framework. Specifically, we

[4] $\Pr(\epsilon \leq 100m) = 0.67$ and $\Pr(\epsilon \leq 300m) = 0.95$ [24], [25]

showed how the TA can be modeled as a obfuscation mechanism and the C-RNTI can be modeled as an anonymizing mechanism. The scope of this work specifically at the TA as an LPPM obfuscation mechanism, and user attribution was assumed.

Next, we presented a geographically diverse data set that also spanned several uplink frequencies and hardware. We showed through graphical and quantitative means of statistical inference that it may be appropriate to model the TA error as Gaussian. This is significant as the literature has traditionally treated this error as uniformly distributed. To the best of our knowledge this is the first time this type of data has been introduced into the literature.

Finally, we showed the type of accuracy possible from TA-based positioning inside an actual network implementation. We further showed how, when augmented with CeSAR, positioning accuracy can approach levels required by the FCC's E-911 mandate. In general, we showed CeSAR added approximately 150 meters of accuracy (CEP 70%).

In summary, we demonstrated that an accurate localization attack can be performed on the LTE protocol due to vulnerabilities in the signaling plane. Further, as the protocol moves toward an implementation with denser infrastructure (e.g., heterogeneous networks) this vulnerability becomes even more exploitable. This complication could be remedied by moving the TA into the portion of the ciphered text similar to the GSM implementation.

References

[1] J. Roth, M. Tummala, and J. Scrofani, "Cellular synchronization assisted refinement (CeSAR): A method for accurate geolocation in LTE-A networks," in *Proc. 49th Hawaii Int. Conf. Syst. Sci.*, 2016, pp. 5842–5850.

[2] Ericsson, "Ericsson mobility report," *White Paper*, 2016.

[3] A. Bleicher, "4G gets real," *IEEE Spectrum*, vol. 51, pp. 38–62, 2014.

[4] J. Cao, M. Ma, H. Li, Y. Zhang, and Z. Luo, "A survey on security aspects for LTE and LTE-A networks," *IEEE Commun. Surveys Tutorials*, vol. 16, no. 1, pp. 283–302, 2014.

[5] T. Murakami and H. Watanabe, "Localization attacks using matrix and tensor factorization," *IEEE Trans. Inform. Forensics Security*, vol. 11, no. 8, pp. 1647–1660, 2016.

[6] S. Gambs, M.-C. Killijian, and M. del Prado-Cortez, "De-anonymization attack on geolocated data," in *Proc. IEEE Intl. Conf. Trust, Security, Privacy Comput. Commun.*, 2013, pp. 789–797.

[7] C.-Y. Chow and M. Mokbel, "Trajectory privacy in location-based services and data publication," in *Proc. ACM SIGKDD Explorations Newsletter*, 2011, pp. 19–29.

[8] J. Krumm, "A survey of computational location privacy," *Pers. Ubiquit. Comput.*, vol. 13, pp. 391–399, 2009.

[9] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive Comput.*, vol. 2, pp. 46–55, 2003.

[10] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *Proc. IEEE Symp. Security Privacy*, 2011, pp. 247–262.

[11] A. Westin, *Privacy and Freedom*, 1st ed. New York: Atheneum, 1967.

[12] L. Barkuus and A. Dey, "Location-based services for mobile telephony: a study of users' privacy concerns," in *Proc. 9th Intl. Conf. Human-Computer Interaction*, 2003, pp. 709–712.

[13] G. Iachello *et al.*, "Control, deception, and communication: evaluating the deployment of a location-enhanced messaging service," in *Proc. ACM Intl. Conf. Pervasive Ubiquitous Comput.*, 2005, pp. 213–231.

[14] E. Kaasinen, "User needs for location-aware mobile services," *Pers. Ubiquit. Comput.*, vol. 7, pp. 70–79, 2003.

[15] E. Dahlman, S. Parkvall, and J. Sköld, *4G LTE/LTE-Advanced for Mobile Broadband*. Academic Press, 2011.

[16] L. Jarvis, J. McEachen, and H. Loomis, "Geolocation of LTE subscriber stations based on the timing advance ranging parameter," in *Proc. Military Commun. Conf.*, 2011, pp. 180–187.

[17] 3GPP TS 36.211, release 10, (v10.7.0), "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation," Feb. 2013.

[18] 3GPP TS 36.331, release 10, (v10.16.0), "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification," Mar. 2015.

[19] I. Güvenç and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surveys & Tutorials*, vol. 11, no. 3, pp. 107–124, 2009.

[20] J. Caffery and G. Stuber, "Overveiew of radiolocation in CDMA cellular systems," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 38–45, 1998.

[21] I. Bilogrevic, M. Jadliwala, and J.-P. Hubaux, "Security and privacy in next generation mobile networks: LTE and femtocells," *Femotcell Workshop*, 2010.

[22] 3GPP TS 33.401, release 9, (v9.7.0), "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3GPP System Architecture Evolution (SAE); Security Architecture (Release 9)," Jun. 2011.

[23] C. Drane, M. Macnaughtan, and C. Scott, "Positioning GSM telephones," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 46–54, 1998.

[24] J. Bull, "Wireless geolocation," *IEEE Veh. Tech. Mag.*, vol. 4, pp. 45–53, 2009.

[25] "Revision of the commission's rules to ensure compatibility with enhanced 911 emergency calling systems, third report and order," *9 FCC Rcd 17388*, 1999.

[26] H. Yoon, R. Shiftehfar, S. Cho, B. Spencer Jr, M. Nelson, and G. Agha, "Victim localization and assessment system for emergency responders," *J. Comput. Civil Eng.*, vol. 30, 2015.

[27] Y. Mo, Z. Zhang, Y. Lu, and G. Agha, "A novel technique for human traffic based radio map updating in wi-fi indoor positioning systems," *KSII Trans. Internet Info. Syst.*, vol. 9, pp. 1881–1903, 2015.

[28] G. Yost and S. Panchapakesan, "Improvement in estimation of time of arrival (TOA) from timing advance (TA)," in *Proc. IEEE Intl. Conf. Universal Personal Commun.*, 1998, pp. 1367–1372.

[29] J. Jin, Z.-J. Qui, and B. Ran, "Intelligent route-based speed estimation using timing advance," in *Proc. IEEE Intell. Transportation Syst. Conf.*, 2006, pp. 194–197.

[30] M. Raitoharju, S. Ali-Löytty, and L. Wirola, "Estimation of base station position using timing advance measurements," in *Proc. Intl. Conf. Graphic Image Process.*, 2011.

[31] R. Whitty, M. Tummala, and J. McEachen, "Precision geolocation of mobile wimax subscribers using timing adjust measurements," in *Proc. 45th Hawaii Int. Conf. Sys. Sci.*, 2012, pp. 5639–5648.

[32] T. Wigren, "Fingerprinting localisation using round trip time and timing advance," *IET Commun.*, vol. 6, pp. 419–427, 2012.

[33] T.-H. Ngo and Y. Kim, "Using timing advance to support proximity discovery in network-assisted D2D communication," in *Proc. Intl. Conf. Ubiquit. Future Net.*, 2015, pp. 926–928.

[34] T. Hiltunen, J. Turkka, R. Mondal, and T. Ristaniemi, "Performance evaluation of LTE radio fingerprint positioning with timing advancing," in *Proc. Intl. Conf. Info. Commun. Sig. Process.*, 2015.