NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

POSITION ESTIMATE FIDELITY FROM TAG MULTILATERATION ATTACK WITHIN THE 5G ENVIRONMENT

by

Alexander W. Schacht

March 2021

Thesis Advisor: John D. Roth
Co-Advisor: John C. McEachen

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** March 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE POSITION ESTIMATE FIDELITY FROM TAG MULTILATERATION ATTACK WITHIN THE 5G ENVIRONMENT	**5. FUNDING NUMBERS**
6. AUTHOR(S) Alexander W. Schacht	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The advent of 5G promises a new age of speed and connectivity of mobile devices. Location-based services will reach a new state of accuracy as well. 4G/LTE implemented the timing advance group (TAG) to increase throughput by allowing user equipment (UE) to connect to multiple base stations (BSs). Timing advance (TA) commands are utilized in order to maintain time synchronization between each servicing BS by directing when the UE should transmit based on the distance to each associated BS. For 4G/LTE, each TA is a multiple of 78.125 meters. As the subcarrier spacing increases in 5G, this distance resolution drops proportionately. These TA commands are sent frequently as the UE moves throughout the environment and are unencrypted. This opens the concern that if an adversary were to collect and correctly associate the TAGs of a specific target, they may be able to ascertain a position estimate using multilateration. The TAG exploit has been examined for 4G/LTE and has been shown to be significant, but the new subcarrier spacing for 5G theoretically will increase the fidelity with which locations can be determined. The focus of this thesis is to establish a Cramér-Rao Lower Bound (CRLB) for position estimates based on the TA commands for each of the 5G sub-carrier spacing and to implement and test an algorithm for finding a position estimate based on target TAs through simulation.

14. SUBJECT TERMS 5G, timing advance group, TAG, location privacy, Cramér-Rao Lower Bound, CRLB, user equipment, UE, base stations, BS, timing advance, TA	**15. NUMBER OF PAGES** 67
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

POSITION ESTIMATE FIDELITY FROM TAG MULTILATERATION ATTACK WITHIN THE 5G ENVIRONMENT

Alexander W. Schacht
Lieutenant, United States Navy
BS, Oregon State University, 2013

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
March 2021

Approved by: John D. Roth
Advisor

John C. McEachen
Co-Advisor

Douglas J. Fouts
Chair, Department of Electrical and Computer Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The advent of 5G promises a new age of speed and connectivity of mobile devices. Location-based services will reach a new state of accuracy as well. 4G/LTE implemented the timing advance group (TAG) to increase throughput by allowing user equipment (UE) to connect to multiple base stations (BSs). Timing advance (TA) commands are utilized in order to maintain time synchronization between each servicing BS by directing when the UE should transmit based on the distance to each associated BS. For 4G/LTE, each TA is a multiple of 78.125 meters. As the subcarrier spacing increases in 5G, this distance resolution drops proportionately. These TA commands are sent frequently as the UE moves throughout the environment and are unencrypted. This opens the concern that if an adversary were to collect and correctly associate the TAGs of a specific target, they may be able to ascertain a position estimate using multilateration. The TAG exploit has been examined for 4G/LTE and has been shown to be significant, but the new subcarrier spacing for 5G theoretically will increase the fidelity with which locations can be determined. The focus of this thesis is to establish a Cramér-Rao Lower Bound (CRLB) for position estimates based on the TA commands for each of the 5G sub-carrier spacing and to implement and test an algorithm for finding a position estimate based on target TAs through simulation.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

3GPP	3rd Generation Partnership Project
3G	third generation
5G	fifth generation
5GPP	5G Infrastructure Public-Private Partnership
AI	artificial intelligence
AR	augmented reality
BBU	baseband unit
CapEx	Capital Expenditures
CDF	cumulative distribution function
CEP	circular error probable
CeSAR	Cellular Synchronization Assisted Refinement
CPRI	common public radio interface
C-RAN	Cloud Radio Access Network
CRLB	Cramér-Rao Lower Bound
C-RNTI	cell-radio network temporary identifier
DOD	Department of Defense
E-911	Enhanced 911
eCPRI	enhanced CPRI
FCC	Federal Communications Commission

FFT Fast Fourier Transform

GPS global positioning system

GSM Global System for Mobile Communications

IEEE Institute of Electrical and Electronics Engineers

IoS Internet of Skills

IoT Internet of Things

IP Internet Protocol

KPI key performance indicator

LBS location-based services

LTE Long-Term Evolution

MCN Mobile Core Network

MLE maximum likelihood estimate

mmWave millimeter wave

MSE mean squared error

NLLS Non-Linear Least Squares

OFDM Orthogonal Frequency-Division Multiplexing

OFDMA Orthogonal Frequency-Division Multiple Access

OpEx Operational Expenditures

OSI Open Systems Interconnection

PLS physical layer split

PSS/SSS primary and secondary synchronization signals

RAN radio access network

RMSE root mean square error

RRH remote radio head

RSS received signal strength

SCS sub-carrier spacing

SDO standards development organization

TA timing advance

TAG Timing Advance Group

ToA time of arrival

UD ultra-dense

UE user equipment

URLL Ultra-reliable Low Latency

V2X vehicle-to-everything

VNF Virtualized Network Functions

VR Virtual Reality

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

First and foremost, thanks is owed to God, for He is the source of all inspiration and perseverance required in any scientific endeavor.

Second, my undying gratitude goes to my wonderful wife, Victoria, for successfully corralling our boys and allowing me a space to work, even while the world was going to chaos around us. The pandemic was trying for everyone, but especially for parents who had nowhere to go to expend some of their children's inexhaustible exuberance. Thank you, my dear, for all that you do and put up with.

Third, I thank my advisor, Dr. John Roth. Thank you for your constant encouragement and thoughtful guidance throughout this process. I appreciate your flexibility, positivity, and the sincere judgement-free method with which you instill knowledge. Coming into this program with a Biology degree as my foundation was daunting to say the least. I had always had an interest in cyber, communications, and networking, but you truly made that interest flourish into a passion.

Finally, to my two sons, Demitri and Declan. Thank you both for the constant gifts of joy, patience, and growth that you give me each and every day. You are my most cherished gift from God.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

1.1 Motivation

Fifth generation (5G) wireless technology is poised as the catalyst of the fourth industrial revolution, describing a nigh inseparable and indistinguishable blending of the physical, digital, and biological realms [1]. This will be accomplished through a myriad of ancillary technologies, such as vehicle-to-everything (V2X), artificial intelligence (AI), Internet of Things (IoT), Internet of Skills (IoS), augmented reality (AR), and Virtual Reality (VR), resulting in an unprecedented level of connectivity encompassing every part of one's public and private life, with the primary enabler being the fully realized 5G network. The number of connected devices globally is increasing every day. By 2023, projections indicate there will be 29.3 billion devices connected to the global internet, with smartphones accounting for 23%, or roughly 6.7 billion, of those. 1.4 billion connections will be facilitated by the 5G network, an over 100-fold growth from 2019 [2], starkly showcasing the rapid rate of adoption expected. As more devices are connected, with uses ranging from the expected (smart phones, tablets, wearables) to the outlandish (homes, fridges, scales, aquariums), more of an individual's personal data is shared online, with or without their express knowledge, causing the issues of privacy and security to become ever more salient. A niche area encompassing the aforementioned, and the concern this thesis will address, is that of location privacy and security in 5G.

Location privacy and security in Long-Term Evolution (LTE) networks have been established as being vulnerable in [3]. The authors of [3] show that unencrypted timing management signaling can be intercepted, and the information inherent exploited to discern the location of a target user equipment (UE) within the network. The methods to conduct timing management in 5G have changed little since their use in Global System for Mobile Communications (GSM). In sustaining the vulnerability of plaintext transmission, and through the introduction of the novel 5G sub-carrier spacings (SCSs), coined "numerology," this problem may have been exacerbated. We believe the goal of understanding the vulnerability's utility within the 5G realm is exactly in step with the directives laid out in the

National Strategy to Secure 5G [4], the *National Cyber Strategy* [5], and the *Department of Defense (DoD) 5G Strategy* [6], to either buttress our defenses or make use of it against our adversaries.

1.2 Objective

In this thesis, we attempt to show that the timing advance (TA) command in 5G is vulnerable to a multilateration localization attack. Furthermore, we show this results in higher fidelity position estimates in comparison to a similar attack possible in LTE and is thus an increased vulnerability. Once the vulnerability is explained, we will show that the results are statistically efficient by establishing the Cramér-Rao Lower Bound (CRLB) for each new SCS. Lastly, we will complete a comparative analysis on position accuracy with and without the utilization of a localization enhancement algorithm, as presented in [7].

1.3 Chapter Breakdown

We discuss the previous works in multilateration localization and TA exploitation as well as give the apposite 5G technical background to provide the details for understanding and establishing the novelty of our efforts in Chapter 2. In Chapter 3, we will showcase our simulation model, following with our results and analysis in Chapter 4. We finish in Chapter 5 with our closing remarks and recommendations for continued efforts.

CHAPTER 2:
Background

In this chapter, we will discuss the relation of our work to the current overarching body of research surrounding 5G and localization vulnerabilities as well as present the technical overview of 5G necessary to best understand our assumptions, simulation design, and analysis of results.

2.1 Literature Review

The development of techniques for discerning the location of a UE within the network has been an ongoing process for nearly three decades. In 1996, the Federal Communications Commission (FCC) mandated that a standardized accuracy requirement for localization, primarily during Enhanced 911 (E-911) calls, be met [8]. These requirements have been expounded upon, made more robust by further tightened position precision requisites throughout the years. 2021 will mark the six-year benchmark for the fourth FCC report and will call for nationwide providers to achieve 50-meter horizontal accuracy, 3-meter vertical accuracy (z-axis capable devices), or a dispatchable location for 80% of all wireless 911 calls [9]. Localization edicts will continue to enlarge the scope and use of an individual's location data, setting the standard for industry within the United States and abroad.

There is an immense amount of research from industry and academia alike on methods for positioning [10]–[13], leading to a plethora of procedures for various cases. Among these, however, five fundamental techniques using radio signals can be broken out: trilateration, triangulation, proximity, scene-analysis, and hybrid [14]. We focus exclusively on trilateration, or more generally, multilateration for this investigation. Multilateration computes a position estimate based on the intersection or distances between geometric structures (in our case, rings) created by distance measurements (time of arrival (ToA), received signal strength (RSS), etc.) between a target UE and transmitter/receiver. Figure 2.1 clearly displays this premise with respect to our particular circumstances.

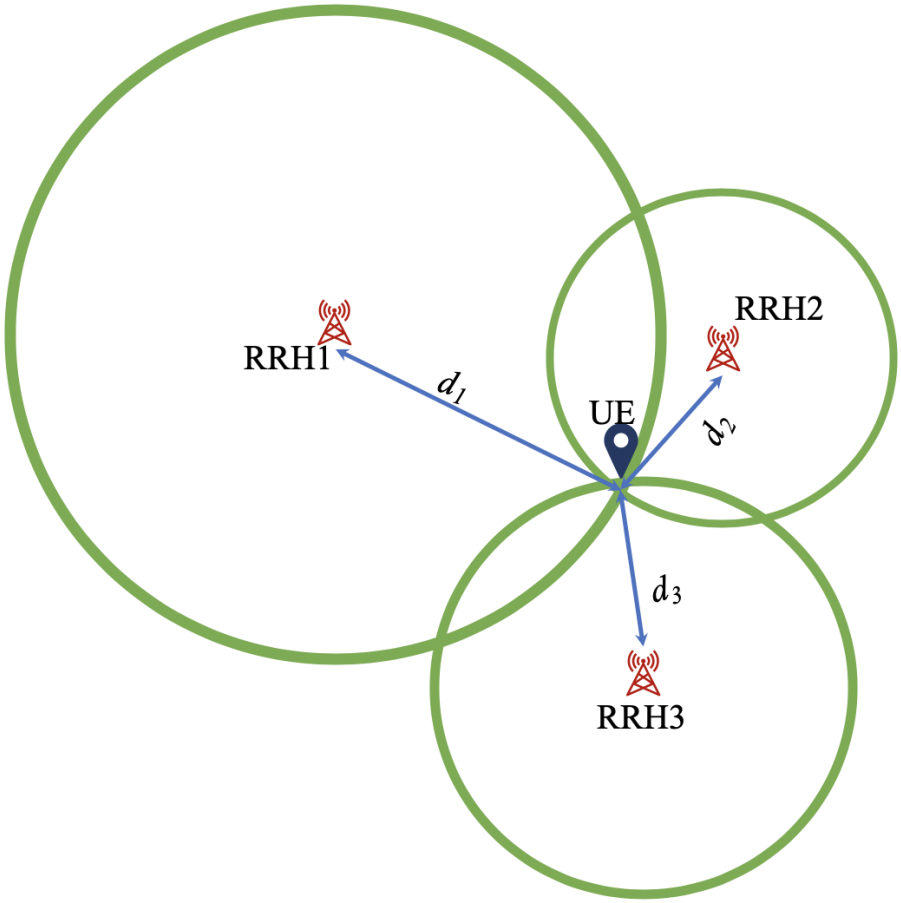

Figure 2.1. An ideal visualization of trilateration in a wireless network, where the intersections of the rings of uncertainty around each remote radio head (RRH) (a 5G network access point) give the approximate position for the target UE. Source: [15].

To match the ever-increasing localization requirements, operators have researched mobile and network-based solutions alike. Through this, multilateration has been validated as a real-world solution to meet regulations in a network-based context [16]. Using the data available from the mobile network timing management schema has been an obvious practice from the beginning of the examination of UE localization and the rise of location-based services (LBS). There has been starkly less focus in the security of the timing management configuration, even following the revelation of the vulnerabilities, as discussed in Chapter 1. This lack of interest may have come from the perceived severity of the vulnerability, with

estimate error minimums reaching approximately 40 meters and maximums exceeding 120 meters, depending on network geometry and the use of localization enhancing techniques [17]. With the protocol remaining unchanged as the world moves to 5G deployments, our research will show that the severity of this privacy concern drastically increases.

Our chief concern, as stated in Chapter 1, is that this increase in position fidelity that can be obtained by a motivated adversary, is directly linked to the decrease in privacy, specifically location privacy, of any other user on the network they so choose to target. The concern for privacy and security of the network is one shared by many. Security is featured as a key performance indicator (KPI) for all standards development organizations (SDOs), and is an objective of all governmental documents concerning the development of 5G. The questions of what defines privacy, how it is changing in the era of 5G, how it is vulnerable, and how to protect it are being constantly mulled about in the literature [18]–[21]. The U.S. Department of Defense (DOD) Defense Innovation Board published [22], lauding 5G for the opportunities it provides and in contrast, plainly warning of an insecure future without the concerted efforts of those pursuing these issues. Expanding further into the militaristic realm, the potential for damage is clearly stated in [6]:

> *With persistent access to an ally's 5G network, an adversary could potentially engage in widespread espionage, threaten the privacy and rights of citizens globally, prepare the operational environment to provide an advantage in armed conflict, conduct information operations, and/or disrupt critical infrastructure.*

2.2 5G Technical Background

The following is a relation of the standard 5G architecture, novel numerologies, as well as timing management methodology.

2.2.1 Architecture

5G aspires to reach unparalleled levels of Ultra-reliable Low Latency (URLL), capacity, cost effectiveness, and environmentally friendly communications [23]. To enable the employment of new spectral gains and take full advantage of the next-generation protocol,

the legacy LTE radio access network (RAN) must be overhauled. To wit, the Cloud Radio Access Network (C-RAN) architecture was conceived, differing from the incumbent architecture in two predominant ways: physical layout and interface design. The physical makeup of a C-RAN is composed of numerous disaggregated, "dumb" RRHs, acting almost singularly as transceivers, connected via a fronthaul network to a grouping of baseband units (BBUs), known as a "BBU pool." Each pool will service a specific geographic region and communicate with the Mobile Core Network (MCN), using the backhaul. This physical separation lies in stark contrast to the LTE network, where the BBU and RRH are deployed together as the base station, or eNodeB, as shown in Figure 2.2. Due to the distribution of relatively cheap hardware, the consolidation of more expensive, energy-dependent hardware, and aggregation of data and control, the C-RAN framework results in greater resource efficiency and sharing, a decrease in Operational Expenditures (OpEx) and Capital Expenditures (CapEx), and acts as an enabler to future use cases such as Virtualized Network Functions (VNF), which are crucial to the advanced network management techniques (i.e., network slicing and on-demand functionality) 5G is capable of [24].

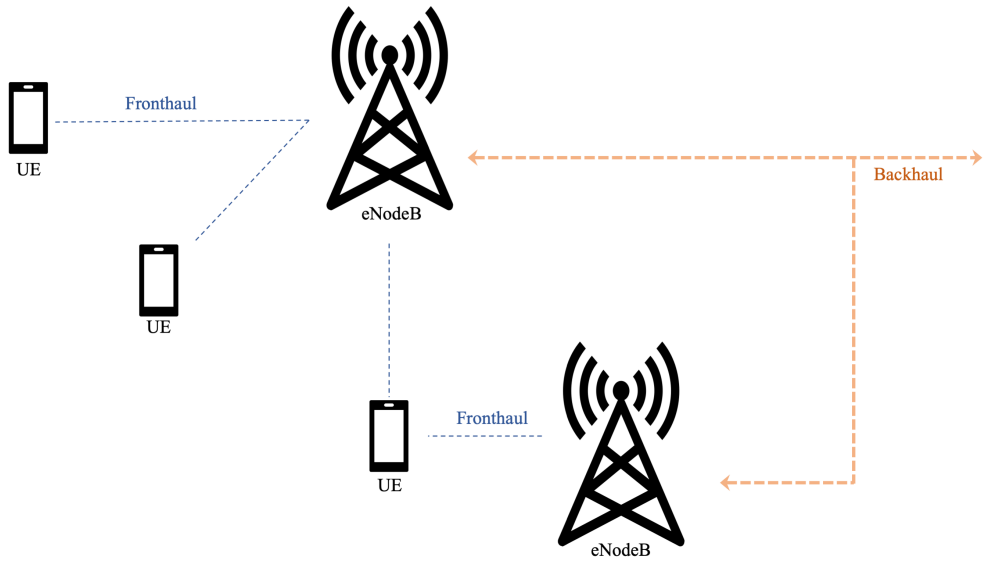

Figure 2.2. Simplified view of the LTE RAN fronthaul/backhaul.

Interface design in 5G is of critical importance to enabling the key metrics and desired disaggregation. Common public radio interface (CPRI), the current design, is responsible

for the digitized and serial *internal* base station interface between the BBU and RRH, as well as transport, connectivity and control, covering the Physical and Data Link layers (Layers 1 and 2) of the Open Systems Interconnection (OSI) model. Essentially, CPRI is the method in which the controlling unit (the BBU) communicates with and directs the radio unit (the RRH). Updates to CPRI are necessary for use in a 5G C-RAN, as the version utilized currently in LTE would require an unfeasible increase in capacity to maintain the proposed data rates [25]. Original interfaces are being proffered by many of the major industry leaders like Institute of Electrical and Electronics Engineers (IEEE), 3rd Generation Partnership Project (3GPP), and 5G Infrastructure Public-Private Partnership (5GPP), to provide a workable solution. CPRI, the industry partnership responsible for the aforementioned LTE interface, has responded with enhanced CPRI (eCPRI) [26]. All of the above efforts strive to find a balance between latency, efficiency, and customizability, founded on a concept put forth by the 3GPP: the physical layer split (PLS) architecture.

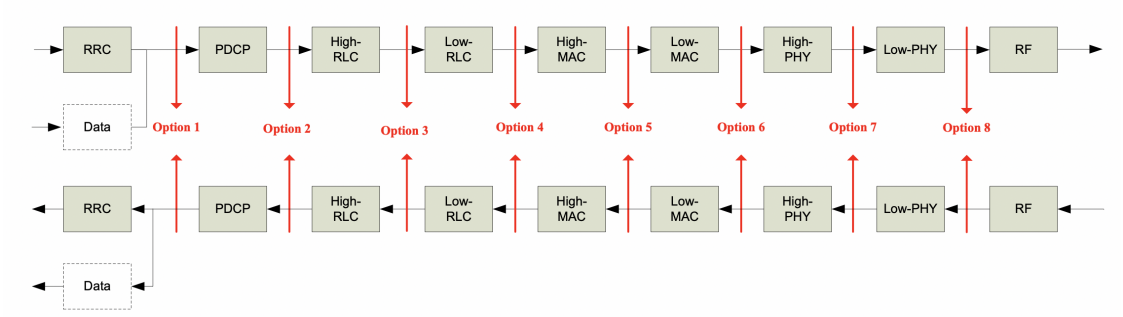

Figure 2.3. Proposed Physical Layer splits, where each box denotes layers of the radio protocol stack, as well as some of their sublayers, that reside specifically within the RAN. Source: [27].

As seen in Figure 2.3, PLS gives enumerated options for different network engineers and operators to decide where they want to distribute various responsibilities; either the BBU or RRH [27]. This allows for each network manager to choose which metrics of performance are best suited for their network. In this thesis, we assume that our target network is operating an eCPRI option 7/2 fronthaul/backhaul. Referencing Figure 2.3 again, a 7/2 option would mean that all functions to the right of option 7 reside with the RRH, those that reside between option 2 and 7 reside with the BBU, and those to the left of option 2 lie with the MCN. Essentially, this option relegates all radio link control and media access control layer

functions to the BBU, while lower functions (modulation, Fast Fourier Transforms (FFTs)) will be carried out by the RRH [28]. This effectively results in the BBU carrying out the majority of the digital baseband radio functions, while the RRH focuses on analog frequency tasks, allowing for cost-effective geographic separation. Figure 2.4 gives a combined view of the architecture and interface of the target network.

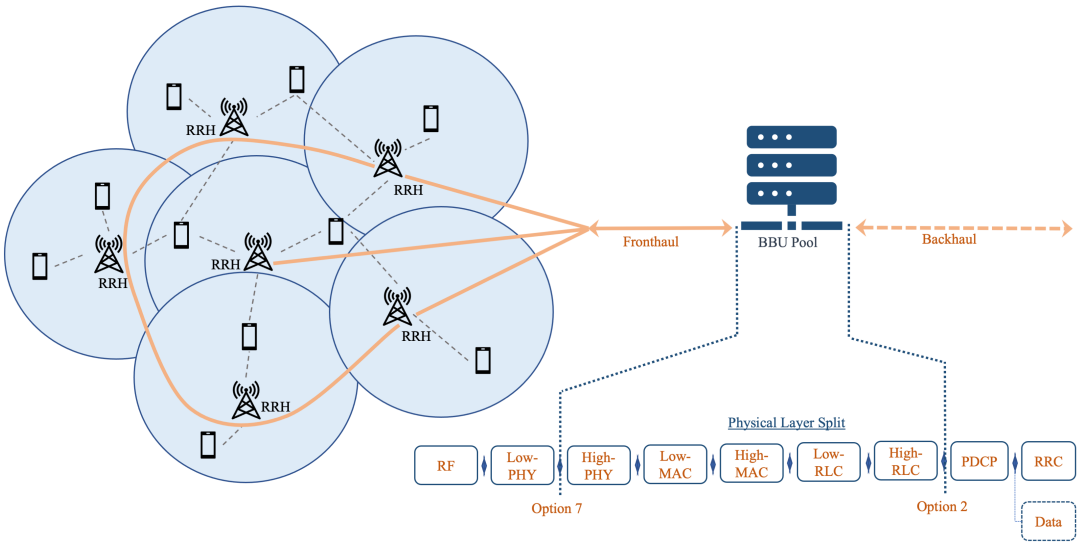

Figure 2.4. Simplified view of the 5G C-RAN fronthaul/backhaul and associated PLS. Source: [15].

2.2.2 Use of the Spectrum

Beginning with third generation (3G), each successive generation of telecommunications has expanded the usable spectrum to increase capacity and alleviate congestion. 5G is no different, offering two different spectrum allocations based on use case; “Sub-6” and “mmWave”. “Sub-6” refers to carrier frequencies below 6 GHz, ranging from 450 MHz to 6 GHz. “Sub-6” will service less dense environments, with larger distances between RRHs, and allow easy transition from, or be backward compatible with, current LTE serviceable areas. The now colloquial “mmWave” stems from the wavelength of the transmitted signal, covering spectrum from 24 GHz to 52 GHz, providing increased bandwidth and reduced latency (see Figure 2.5).

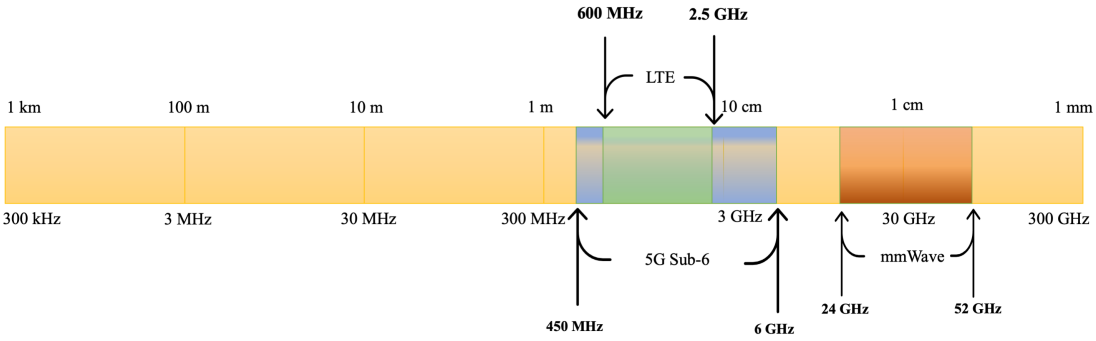

Figure 2.5. 5G frequency allocation and comparison to LTE (show in log-scale).

Just as in LTE before it, 5G will utilize Orthogonal Frequency-Division Multiplexing (OFDM) and Orthogonal Frequency-Division Multiple Access (OFDMA) modulation schemes. OFDM uses multiple orthogonal subcarriers to transmit data symbols in parallel (see Figure 2.6). These subcarriers are separated in LTE by a static SCS of 15 kHz; this is where 5G differs greatly. 5G introduces flexible SCSs and symbol lengths as "numerologies", represented as μ, which ranges from 0 to 4 [29]. SCS is calculated as

$$SCS = 15 \times 2^{\mu} \text{ kHz for } \mu \in [0, 4]. \tag{2.1}$$

allowing for novel SCSs of 15 to 240 kHz[1], visually shown in Figure 2.7. This increased flexibility empowers efficient and novel use of the available spectrum, where generally, higher SCS is used for shorter transmission times alongside mmWave, while lower SCS is optimal for high throughput performance [30].

[1]480 kHz is the maximum SCS as of 3GPP Release 16; however, as it is meant for future study and is not actually used at this time, it is not considered during this investigation.

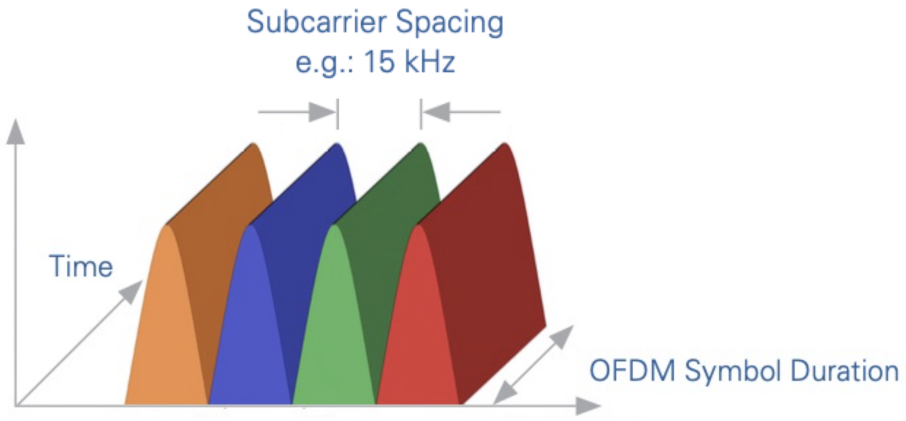

Figure 2.6. Illustration of OFDM principle and overlap and interaction of symbols. Source: [31].

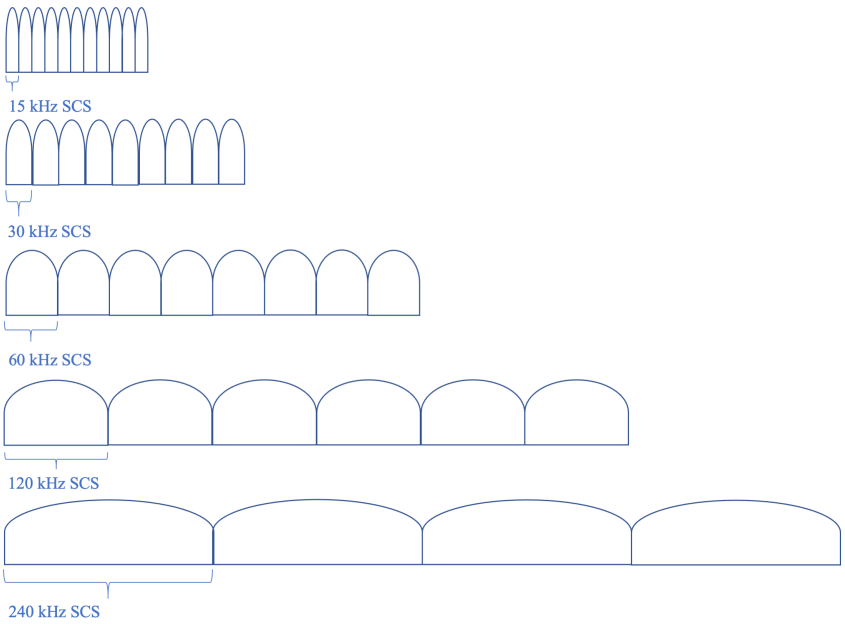

Figure 2.7. Comparison of SCS size across novel numerologies.

2.2.3 Timing Management

Control over time-domain resources within the OFDMA construct is pivotal in ensuring proper network operation, especially when a single UE is communicating with multiple RRHs. 5G, just as LTE before it, enacts this control via the TA command. The TA command is made up of two notable parts, the TA value N_{TA} and the Timing Advance Group (TAG) [32]. The purpose of the command is to ensure that as a UE moves within the serviceable environment, its transmissions arrive at the RRH during its given time slot. As UE distances vary, so too do the propagation times for its transmissions, often significantly. This requires some method of synchronization.

The TAG is a 2-bit field that allows for the unique association of a TA command to one of a maximum of four RRHs, to account for communication with multiple RRHs through carrier aggregation [33] and the likely eventuality that they are not equidistant from the UE. N_{TA} has an associated fixed time value due to the static time unit in LTE, T_s, detailed in Equation (2.2),

$$T_s = \frac{1}{\Delta f_{ref} \times N_{f,ref}} = \frac{1}{15 \times 10^3 \times 2048} \approx 32.6 \text{ nsec} \tag{2.2}$$

where Δf_{ref} is the LTE SCS and $N_{f,ref}$ the maximum number of subcarriers. TA values, represented as integers, N_{TA}, account for 16 time units such that

$$N_{TA} = 16T_s. \tag{2.3}$$

From Equations (2.2) and (2.3), the one-way distance resolution, r, can be calculated as shown

$$r = \frac{cN_{TA}}{2} = 78.125 \text{ meters}. \tag{2.4}$$

In order to employ 5G numerologies, a new base unit of time was developed in [34]

$$T_c = \frac{1}{\Delta f_{max} \times N_f} = \frac{1}{480 \times 10^3 \times 4096} \approx .51 \text{ nsec} \tag{2.5}$$

where Δf_{max} is the maximum SCS and N_f the maximum number of subcarriers. The

relationship between T_c and T_s is found in [29], where κ is introduced and defined as

$$\kappa = \frac{T_s}{T_c} = 64. \tag{2.6}$$

The basic time unit is redefined for 5G as

$$T_s = \frac{1}{\Delta f_{ref} \times N_{f,ref}} = \frac{1}{15 \times 10^3 \times 2^{\mu} \times 2048} \tag{2.7}$$

where Δf_{ref} is the standard LTE SCS this time multiplied by 2^{μ} in order to achieve the various numerologies, and $N_{f,ref}$ is the same as in (2.2). This results in N_{TA} now being defined as

$$N_{TA} = \frac{16\kappa T_c}{2^{\mu}}. \tag{2.8}$$

As to be expected, if $\mu = 0$, the result is the legacy LTE T_s. Based on Equation (2.6), T_c is cancelled out leaving

$$N_{TA} = \frac{16 T_s}{2^{\mu}}. \tag{2.9}$$

Therefore, making use of Equations (2.4) and (2.9), the new distance resolutions are calculated as

$$r = \frac{c N_{TA}}{2} = \frac{78.125}{2^{\mu}} \text{ meters} \tag{2.10}$$

directly conveying the dependence of 5G TA distance resolutions on the associated numerology. Table 2.1 summarizes the new resolutions while Figure 2.8 admits a visual understanding of the TA.

Table 2.1. 5G Numerology Distance Resolutions. Adapted from [35].

μ	Distance Resolution (m)	Subcarrier Spacing (kHz)
0	78.125	15
1	39.06	30
2	19.53	60
3	9.77	120
4	4.88	240

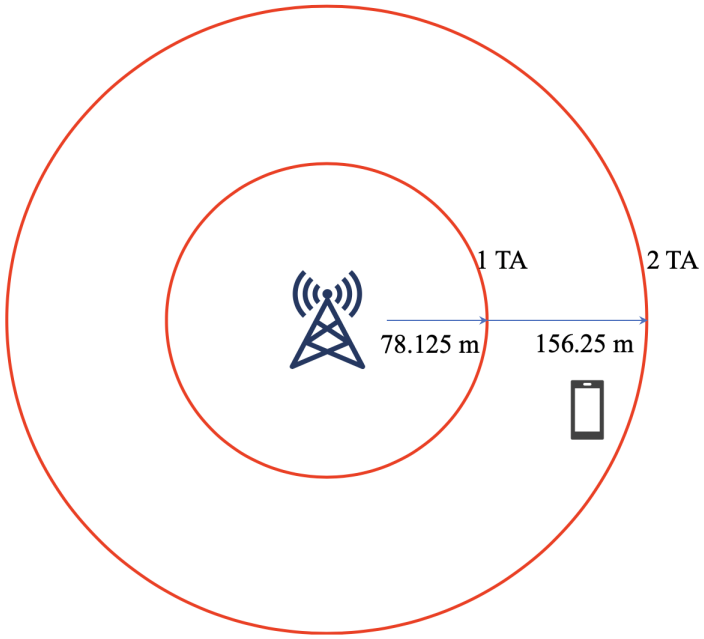

Figure 2.8. Visual representation of a TA command with TA = 2 and the associated distances with $\mu = 0$. Source: [15].

2.3 Statistical Efficiency

A keystone portion of our investigation is to determine whether the method of utilizing the TA to localize a target UE results in a position maximum likelihood estimate (MLE). To this end, we employed the CRLB: a mature method for ascertaining the minimum variance of an unbiased parameter estimate. The bound is defined as

$$\mathrm{Var}_\theta\{T\} \geq I^{-1}(\theta) \tag{2.11}$$

where T is an unbiased estimator of the parameter θ and $I(\theta)$ is the Fisher information with respect to θ. In [17], the CRLB applicability in LTE TA localization has been established, where T is the UE position MLE and θ the parameter to be estimated. It also shows that for TA distance resolution, r, the root mean square error (RMSE) is bounded by the CRLB for

$$r \lesssim 3.4\sigma \tag{2.12}$$

where σ is the standard deviation of distance error estimates. The work presented here is the logical extension of this theory to 5G. Based on (2.12), Table 2.2 summarizes the values of σ for which we expect this to remain true. For instance, suppose a RRH operating with $\mu = 0$ as in LTE can measure UE distance with a standard deviation of 20 meters. This being the case, we can expect the distance error estimate to always be above the CRLB derived from the distance estimation fidelity. However, the same bound may not apply for higher numerologies. This will be further investigated through simulation in Chapters 3 and 4.

Table 2.2. 5G numerologies and their associated σ values based on (2.12)

μ	σ
0	23
1	11.5
2	5.7
3	2.9
4	1.4

2.4 Localization Refinement

The Cellular Synchronization Assisted Refinement (CeSAR) algorithm is a technique for improving the position MLE by way of the TA multilateration method, through the introduction of secondary equipment for receipt of the target UE uplink burst to the RRH. This provides another known distance to further increase position fidelity.

First, the sensor will listen for the primary and secondary synchronization signals (PSS/SSS) from each RRH to establish downlink synchronization with the network. Downlink frame timing is then estimated based on propagation delay between the sensor and the RRH. The sensor now continues to listen to downlink frames until the one with the target cell-radio network temporary identifier (C-RNTI) is identified. The C-RNTI is fundamentally a unique temporary software address issued by the network to each UE, obfuscating the UE identity. Once the target frame has been identified, the associated TA is stripped away and converted to a distance measurement. This measurement is used to estimate the UE's uplink

transmission time. The sensor then observes the uplink burst timing information from one of the serving RRHs. The algorithm will then compare the estimated UE transmission time, with the observed timing of the transmission, and use the determined difference to calculate the distance to the target UE from the sensor. Distance measurements are found for all serving RRHs and finally a position estimate is found for the UE. This algorithm, as defined in [7] is presented in Table 2.3 for brevity.

Table 2.3. Cellular Synchronization Assisted Refinement. Source: [7].

Step	Procedure	Description
1	CeSAR ($\boldsymbol{p}_{RRH}$, $\boldsymbol{p}_{sensor}$, target C-RNTI)	Passive enhancement procedure utilizing sensors, RRH's and target cell-radio network temporary identifier (C-RNTI).
2	**function PSS/SSS SYNC**	Primary and secondary synchronization signal synchronization.
3	sensor $\leftarrow$ RRH downlink	Sensor listens for the PSS/SSS from serving RRH.
4	**end function**	Steps 2-4 are required for synchronization of the sensor to the communicating RRH to allow it to decode cell data.
5	**repeat**	Repeat steps 2-4 for each RRH
6	$x \leftarrow$ observed C-RNTI	Sensor will continue to decode packets.
7	**until**	Sensor continues to search for target C-RNTI.
8	x == target C-RNTI	The target UE downlink frame has been identified.
9	$\hat{d}_i \leftarrow$ TA $\times$ 78.125m	The associated TA is stripped out and converted to a distance.
10	$t \leftarrow$ TA estimate tx	The TA is used to estimate the target UE's uplink transmission time.
11	$t' \leftarrow$ observed	observe true uplink timing information from one of the serving RRHs.
12	$\Delta t \leftarrow t' - t$	Utilize information from step 10 to measure the propagation delay from the UE to the sensor.
13	$\hat{d}' = \Delta t \times c$	Convert to a distance measurement.
14	$\hat{\boldsymbol{d}} \leftarrow [\hat{d}, ..., \hat{d_N}, \hat{d}']^T$	Additional distance measurement is now added to the distance measurement obtained from the serving RRHs.
15	$\hat{\boldsymbol{p}}$ = arg min $p(\hat{\boldsymbol{d}}\|\boldsymbol{d})$	Uses step 14 to find a MLE.
16	**end procedure**	END

2.5 Summary

In this chapter, we have discussed the fundamental requisite topics for a more complete comprehension of the simulation parameters, testing method, and results. Particularly, we gave an overview of 5G architecture, novel use of the spectrum and introduction of numerologies, and timing management method. We then review statistical efficiency, the CRLB, and the CeSAR algorithm. Next, we introduce the four guiding questions of our research, and the simulations derived there from.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3:
Simulation

In an effort to thoroughly elucidate our objective as denoted in 1.2, we ran 4 similar, yet distinctly separate simulations aimed at answering the following questions:

1. How do the number of RRHs collecting and the numerology in use affect the location estimates?
2. What effect does σ have on localization efficiency?
3. How does the positioning algorithm perform in three dimensional space and how does the use of the CeSAR algorithm affect positioning across all numerologies?
4. What would a multilateration attack within a C-RAN architecture look like in a real-world deployment?

In this Chapter, we will discuss the setup of each of these simulations in turn, with their respective results and analysis being discussed in Chapter 4.

3.1 Simulation of RRH and Numerology Effects

The experimental setup for this simulation is established in [35], and reiterated here for clarity and ease of understanding. To answer the first question, we select our two main parameters for testing, specifically, the number of RRHs the UE is in communication with and the numerology in use. We chose to limit the number of RRHs to 2, 3, or 4, which can be supported by the use of the TAG. We then send each of the possible combinations of number of RRHs and numerology in use through our simulation. First, we generate a $100,000\,\text{km}^2$ area to place our RRHs and UE. The UE is always placed at (0,0) (the center point of the area), while the locations of the RRHs were modeled as uniform random variables and placed within the testing area. Second, we calculated the TA value assigned to the UE from each RRH. For example, if the UE is in communication with 2 RRHs, then there will be 2 TA values, each associated to their own RRH. If 3 RRHs are in use then 3 TA values will be present, and so on. Now that there is an associated TA value for each supporting RRH, we have effectively created rings of uncertainty, with a width equal to the distance resolution of the particular numerology in use, around each RRH. These

rings account for all possible locations of the UE. From here it can be discerned that the UE would be inside, or on the boundary of, the intersection of each ring (see 2.1 for a visual representation of this concept). To estimate the position of the UE, $\hat{\boldsymbol{p}}$, we utilized the Non-Linear Least Squares (NLLS) method presented in [10], [36]. This involves the minimization of $\boldsymbol{x} = [x, y]^T$ in the following

$$\hat{\boldsymbol{p}} = \arg\min_{\boldsymbol{x}} \sum_{i=1}^{N} [d_i - \|\boldsymbol{x} - \boldsymbol{x_i}\|]^2 \tag{3.1}$$

where $\hat{\boldsymbol{p}} = [\hat{x}, \hat{y}]^T$, d_i is the distance from each RRH to the center of its TA ring, $\boldsymbol{x_i} = [x_i, y_i]^T$ are the positions of each RRH, i an integer value from 1 to N representative of the total number of supporting RRHs. We ran the NLLS algorithm with the starting position being the actual position of the UE (0,0), and then measured the deviation from that point once it converged. This distance between the NLLS solution and the UE is our final distance error. For simplicity and clarity, this model assumes no error due to noisy conditions. This means that the perceived distance of the UE to the RRH is not shifted due to noise, and the only error accrued is due to the quantization of the distance measurements into the TA values. This focuses the effect of SCS on distance resolution and localization. We simulate the above across each combination of number of RRHs to SCS, and then generate their respective cumulative distribution functions (CDFs) to compare performance (as shown in Section 4.1).

3.2 Simulation of Statistical Efficiency

In Chapter 2, we discuss the CRLB, as well as its use for determining efficiency of localization estimates within the 5G network. Here, we will observe the effect of varying σ across a range of values on localization efficiency, based on the difference between estimate mean squared error (MSE) and the CRLB. For this simulation we begin by establishing our wireless architecture in much the same fashion as the first, however we've reduced the deployment area[2] to 1 km^2, and use the maximum number of RRHs (4), for each simulation. We chose to use the maximum number of RRHs for consistency across values of σ, as well

[2]Due to our assumption of a lossless transmission, the larger distances do not have any effect on the simulation, so we shrank the deployment area for ease of examination.

as to see the results aligned with the best case scenario. The true distances from the RRHs to the UE were computed and introduced to Gaussian noise, to form our distance estimates, $\hat{d}_i$. These were then quantized into their respective TAs. We again find $\hat{\boldsymbol{p}}$ using (3.1). Once found, the squared distance error, error_d^2, is determined by squaring the assessed distance between $\hat{\boldsymbol{p}}$ and the true UE position, $\boldsymbol{p}$. Now that the position error for that trial has been evaluated, we find the CRLB. Applying (2.11), we state that the MSE for $\hat{\boldsymbol{p}}$ is bounded by

$$\text{Var}\{\hat{\boldsymbol{p}}\} \geq \boldsymbol{I}^{-1}(\boldsymbol{x}) \tag{3.2}$$

where $\boldsymbol{I}$, for the purpose of multilateration-based position estimates, is given by [10]

$$\boldsymbol{I} = \begin{bmatrix} \sum_{i=1}^{N} \dfrac{(x-x_i)^2}{\sigma_i^2 d_i^2} & \sum_{i=1}^{N} \dfrac{(x-x_i)(y-y_i)}{\sigma_i^2 d_i^2} \\ \sum_{i=1}^{N} \dfrac{(x-x_i)(y-y_i)}{\sigma_i^2 d_i^2} & \sum_{i=1}^{N} \dfrac{(y-y_i)^2}{\sigma_i^2 d_i^2} \end{bmatrix}, \tag{3.3}$$

and CRLB $= tr(\,\boldsymbol{I}^{-1})$, where $tr()$ is the trace of the matrix, defined as the summation of elements along the main diagonal from upper left to lower right. Next, we calculate the difference between our MSE and the CRLB normalized[3] by the CRLB as

$$\text{ediff}_{\text{CRLB}} = \frac{\text{error}_d^2 - \text{CRLB}}{\text{CRLB}}. \tag{3.4}$$

We then conducted this procedure per each numerology, for values of σ ranging from 1 to 30. This range was chosen in order to show how the estimates trend for each numerology as they approach and exceed the values presented in Table 2.2. Lastly, the average $\text{ediff}_{\text{CRLB}}$ is found for each value of σ.

[3]Due to the CRLB being particular to each iteration based on the relative distances from RRHs to UE, normalizing by the CRLB is required to compare across trials.

3.3 Simulation for Determining $\mathbb{R}^3$ and CeSAR Applicability

To test the potential gains of CeSAR enhancement paired with the novel 5G numerologies, as well as showing the localization effectiveness in $\mathbb{R}^3$, we modified our scheme by uniformly distributing and selecting 4 RRHs and 1 UE within our deployment area, with heights also being uniformly distributed in the range of 0 to 400 feet. We then follow a similar procedure as Section 3.2, solving (3.1), without CeSAR enhancement, while selecting values of σ that meet the CRLB for each numerology (further explanation can be found in Section 4.2). These findings were then enhanced by adding a single CeSAR sensor, via the same method as described in Table 2.3. For the purpose of this simulation, steps 1-13 of the algorithm were assumed to have been completed, and we applied the true distance measurement between sensor and UE, vice the distance measurement calculated in step 13. This additional distance measurement is combined with those calculated previously and the now enhanced MLE is determined as before.

3.4 Simulation of Real-World Use Cases

The final simulation conducted was of a localization attack on the island of Kauai (the location was chosen to facilitate discussion for the 54th Hawaii International Conference on System Sciences, which was held in Kauai). The simulation was broken in to two different wireless architectures: the first being a current in-place setup, the second being a vision of the ultra-dense (UD) 5G deployment to come (as seen in Figures 3.2 and 3.3, respectively). The choice to separate architectures is based on the particular use cases for the numerologies, where in the "in-place" case numerologies 0-2 will be utilized, while 3 and 4 will be used for the UD deployment. This is justified in that the distance of the RRH to the target for the "in-place" case is ≈ 2.96 km which is not conducive to millimeter wave (mmWave) communications. The "in-place" case uses the location of an operational 5G tower, with CeSAR sensors placed based on topography and angle from the projected position of the UE. The UD architecture is somewhat arbitrary, having an emphasis on having near line-of-sight conditions within the observable area as would be necessary for a mmWave UD network.

The UE, RRHs, and CeSAR sensors were plotted in Google Earth to obtain global posi-

tioning system (GPS) coordinates, which were then converted to their values in meters. The CeSAR enhanced algorithm was used to find position estimate for the current architecture ($\mu = 0 : 2$), while the standard NLLS algorithm solved for the UD deployment ($\mu = 3 : 4$). The reason for not running the UD network through the enhanced algorithm is the low probability of properly positioning a sensor to receive mmWave communications. Without already having a solid estimate of the UE position, the extremely small beam-width and controlled directionality inherent to mmWave communications, would make positioning non-trivial. Also, even if correct positioning were achieved, the sensor would have to be practically co-linear to the communicating RRH. This would result in little to no extra information being provided, as the distance estimate created would be nearly the exact same as that created without the CeSAR sensor due to the sensor and the RRH being in-line. An angular difference between the placement of the sensor and the RRH will result in greater information as the angle increases. Figure 3.1 gives an example of how this process works, and shows visually why the case of an offset sensor is desirable.

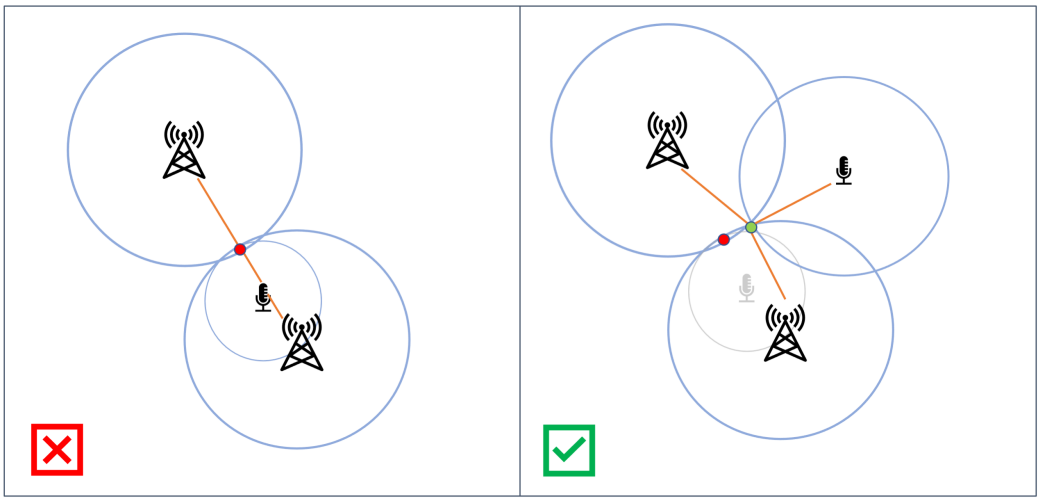

Figure 3.1. The left-hand image shows that when the sensor, represented by the microphone, is co-linear to the receiving RRH, the estimate it provides is no different than that provided by the two RRH on their own. This results in an erroneous position, shown as the red dot. The right-hand image shows that if we move the sensor to a location that is offset from the RRHs, the estimate shifts as well. We can see from the now greyed out TA ring and sensor that the new estimate, highlighted in green, does still fall along that original estimate as it should, though its true position was different then originally estimated.

In order to find the MLE of the target position, this procedure must then be repeated numerous times.

Figure 3.2. Real-world deployment on Kauai. Source: [15].

Figure 3.3. UD deployment on Kauai. Source: [15].

3.5 Summary

We have now introduced our four guiding questions, and the simulations constructed to answer those questions. Specifically, the questions are: how do the number of RRHs and numerology in use affect positioning accuracy, how does σ affect localization efficiency, how does the algorithm perform in $\mathbb{R}^3$ with and without CeSAR enhancement, and finally how does our algorithm perform in actual wireless deployments? Next, we will discuss the results obtained and the inherent implications for the 5G era.

CHAPTER 4:
Results

Now that our simulation environments have been established, in this chapter we will go through the results and answers to our four guiding questions. Monte Carlo trials for simulations were run no less than 100,000 times.

4.1 Simulation of RRH and Numerology Effects

For this simulation, we expected that position accuracy would increase as the number of RRHs increase and/or as the SCS becomes larger, and our results showed just that. The CDFs shown in Figures 4.1 – 4.3, illustrate how sharp a performance increase comes with the novel numerologies introduced in 5G.

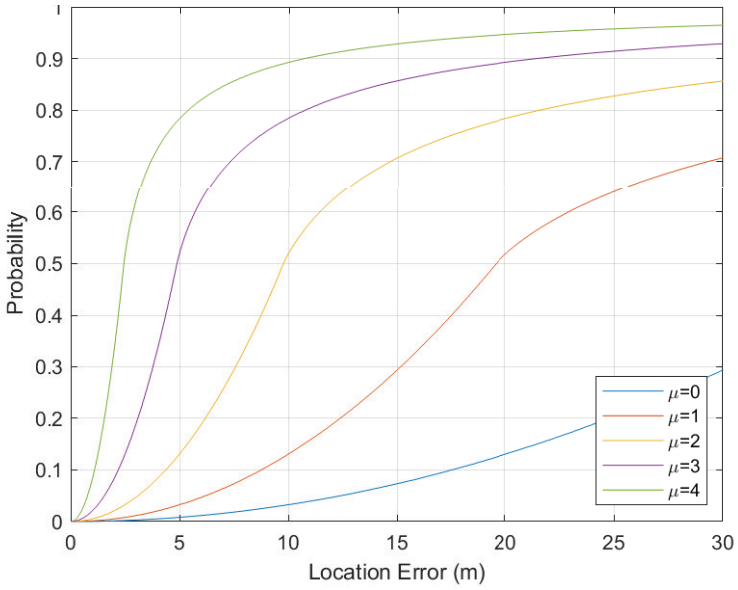

Figure 4.1. CDFs of location error across all numerologies when receiving TA data from *two* RRHs communicating with the target UE. Source: [35].

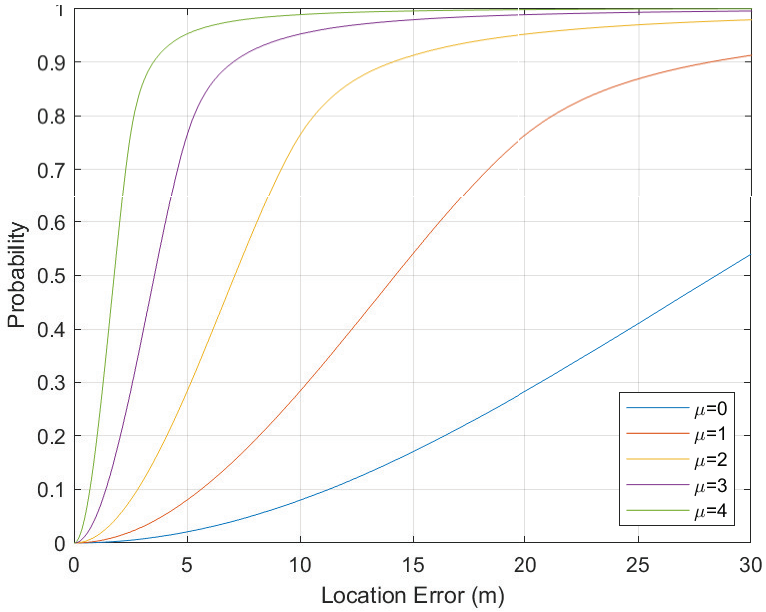

Figure 4.2. CDFs of location error across all numerologies when receiving TA data from *three* RRHs communicating with the target UE. Source [35].

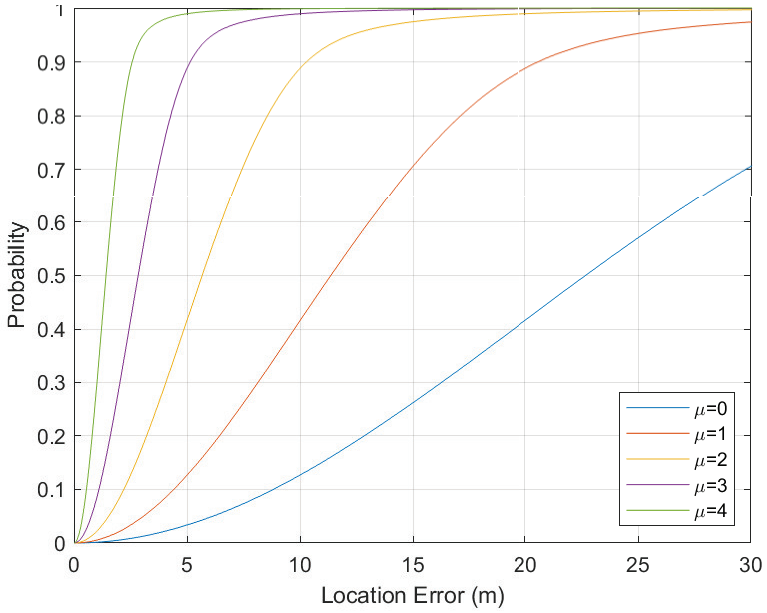

Figure 4.3. CDFs of location error across all numerologies when receiving TA data from *four* RRHs communicating with the target UE. Source: [35].

Delving further, in Tables 4.1 – 4.3 we tabulated the 90% and 95% circular error probable (CEP) values for each of the RRH combinations. It can be seen that the trend from one numerology to the next, for every number of RRH, is approximately halving the location error. The least accurate scenario occurs with 2 RRHs and $\mu = 0$, whereas the most accurate scenario is 4 RRHs and $\mu = 4$, where a target UE could be localized to within 3.1 meters with 95% confidence. Comparing performance between the 2 RRH scenario with that of both 3 and 4, we see that at a 90% confidence there is an average performance increase of 67% and 75.95%, respectively. At 95% confidence this further escalates to 76.4% and 85.3%, demonstrating that in either case there is dramatic improvement in localization with the addition of a RRH. Comparing solely between 3 and 4 RRH cases, we see a surge in performance at both the 90% and 95% confidence levels of 27.1% and 37.5%, respectively.

Table 4.1. Location Error Probabilities using 2 RRHs across all numerologies. Adapted from: [35].

μ	**Subcarrier Spacing (kHz)**	**90% Confidence Location Error (m)**	**95% Confidence Location Error (m)**
0	15	168.72	317.46
1	30	85.41	165.71
2	60	42.98	84.42
3	120	21.50	42.48
4	240	10.75	21.36

Table 4.2. Location Error Probabilities using 3 RRHs across all numerologies. Adapted from: [35].

μ	**Subcarrier Spacing (kHz)**	**90% Confidence Location Error (m)**	**95% Confidence Location Error (m)**
0	15	56.29	78.57
1	30	28.21	39.33
2	60	14.14	19.71
3	120	7.06	9.85
4	240	3.53	4.92

Table 4.3. Location Error Probabilities using 4 RRHs across all numerologies. Adapted from: [35].

μ	**Subcarrier Spacing (kHz)**	**90% Confidence Location Error (m)**	**95% Confidence Location Error (m)**
0	15	41.18	49.12
1	30	20.57	24.61
2	60	10.29	12.27
3	120	5.14	6.14
4	240	2.57	3.08

The CEP for six different cases are presented in Figures 4.4 – 4.6. The CEP relates confidence location error and location distance error, and is another simple method of depicting location accuracy [37]. The different cases are a combination of 2, 3, or 4 RRHs and the lowest and highest numerologies, thereby comparing the highest and lowest position fidelity for each number of RRHs in use. The UE's position is normalized to the center of the graph and is displayed as a red asterisk. Blue markers indicate estimated positions over 1,000 trials (this number was chosen arbitrarily for demonstration). The red rings delineate the area within which we can say with 95% confidence that the target resides. Graphs were scaled to properly show distinction of the circle and surrounding points. Differences in accuracy become clear when observing this difference in scale per graph.

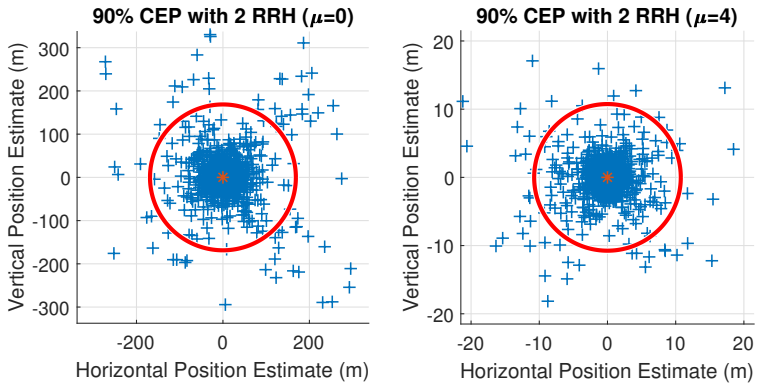

Figure 4.4. 90% CEP using 2 RRH with numerologies $\mu = 0$ and $\mu = 4$. Axes have been scaled appropriately based on the SCS in use for ease of viewing. Source: [35].

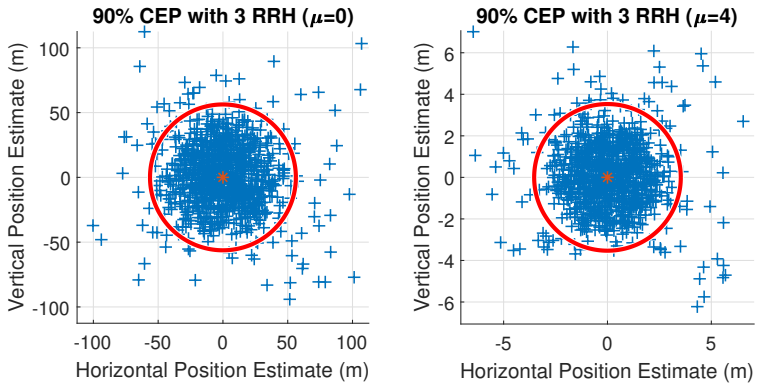

Figure 4.5. 90% CEP using 3 RRH with numerologies $\mu = 0$ and $\mu = 4$. Axes have been scaled appropriately based on the SCS in use for ease of viewing. Source: [35].

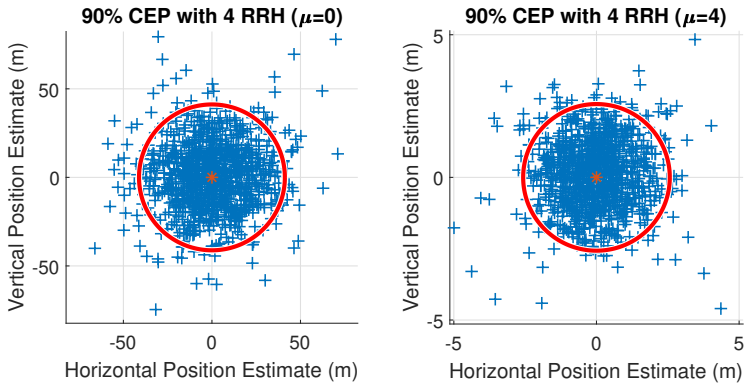

Figure 4.6. 90% CEP using 4 RRH with numerologies $\mu = 0$ and $\mu = 4$. Axes have been scaled appropriately based on the SCS in use for ease of viewing. Source: [35].

4.2 Simulation of Statistical Efficiency

In order to determine the statistical efficiency of the localization algorithm, we chose to compare the MSE of our distance estimates to the CRLB at varying values of σ, as discussed in Section 3.2. Figure 4.7 presents the results of this investigation, where the y-axis is the difference between the MSE and the CRLB for every σ value along the x-axis. The figure shows that for each numerology as σ increases, $\text{ediff}_{\text{CRLB}}$ (determined by (3.4)), approaches 0. Vertical dotted lines indicate the expected value of σ from Table 2.2 to which the MSE for that numerology would meet the CRLB, based on the theoretical bounding of $r \lesssim 3.4\sigma$ (as described in Section 2.3) The results add weight to the theory as each numerology performed closely, and met the CRLB, producing baseline values of σ that return the most statistically efficient location estimate per numerology chosen. What is perhaps even more interesting, is that these results suggest that the CRLB will be the lower bound for the localization MSE for all values of σ, and that as σ becomes large, the MSE meets the CRLB. This differs from the inequality discussed for LTE. We believe the reason for this divergence is based on

simulation differences between this work and that of [17]. The latter performed an analysis based on many trials where the position of the UE does not change and is either in the center of a TA ring, or on the edges. When the UE is exactly in the center of a TA ring, the distance estimate matches that of the quantized TA value. For example, assuming $\mu = 0$, if the UE is located exactly in the center of the TA = 2 ring, then the actual UE distance from the RRH of 156.25 m will exactly equal that of the TA = 2 bin, resulting in no shift and no loss in accuracy of the follow-on estimate. If however, the UE was located on the edge of the ring, then the actual distance of $\approx$ 118 m would end up being shifted 38 m when quantized by the TA = 2 TA command. A UE located on the edge of a TA ring can be considered the worst case scenario for the algorithm to estimate, while being located in the center can be considered the best; therefore, the comparison of both is effectively examining the two extremes. The authors' results show that in these two cases, as the ratio of τ to σ increases, where τ is the estimated value of a random variable[4], the RMSE for each case split away from one another concluding with one increasing and the other decreasing past the CRLB at $\approx 3.4\sigma$ (see Figure 4.8). This bifurcation supports (2.12). However, in our examination the location of the UE within each TA annulus was random based on the uniform distribution of RRHs for each trial. We believe that this randomness resulted in the CRLB being a valid lower bound for all values of σ. In fact, giving Figure 4.8 a more thorough examination, it would appear that if one took the average of the normalized RMSE between each case, the result would be values that were above or nearly equal to the CRLB for any ratio of τ to σ. These results are formally defined in the following conjecture.

Conjecture. *If the location of a target UE within each of the N TA annuli is random, then*

$$\lim_{N\to\infty} \tfrac{1}{N} \sum_{i=1}^{N} \left(\frac{e_i^2 - \mathrm{CRLB}_i}{\mathrm{CRLB}_i} \right) \geq 0 \; \forall \; \frac{r}{\sigma}$$

This presents a novel understanding of the relationship between the CRLB and TA localization estimation that demands further evaluation than what is simply provided here.

[4]This is the equivalent of r, or the distance resolution for each numerology, in our case

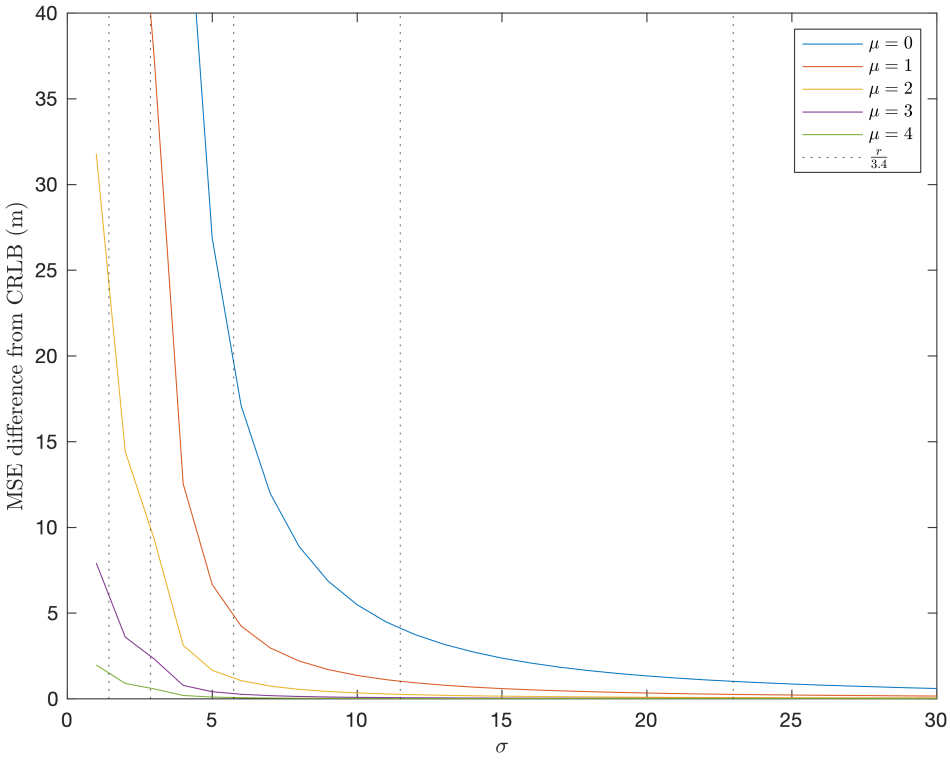

Figure 4.7. Difference between MSE and CRLB for each SCS across σ values. Source: [15].

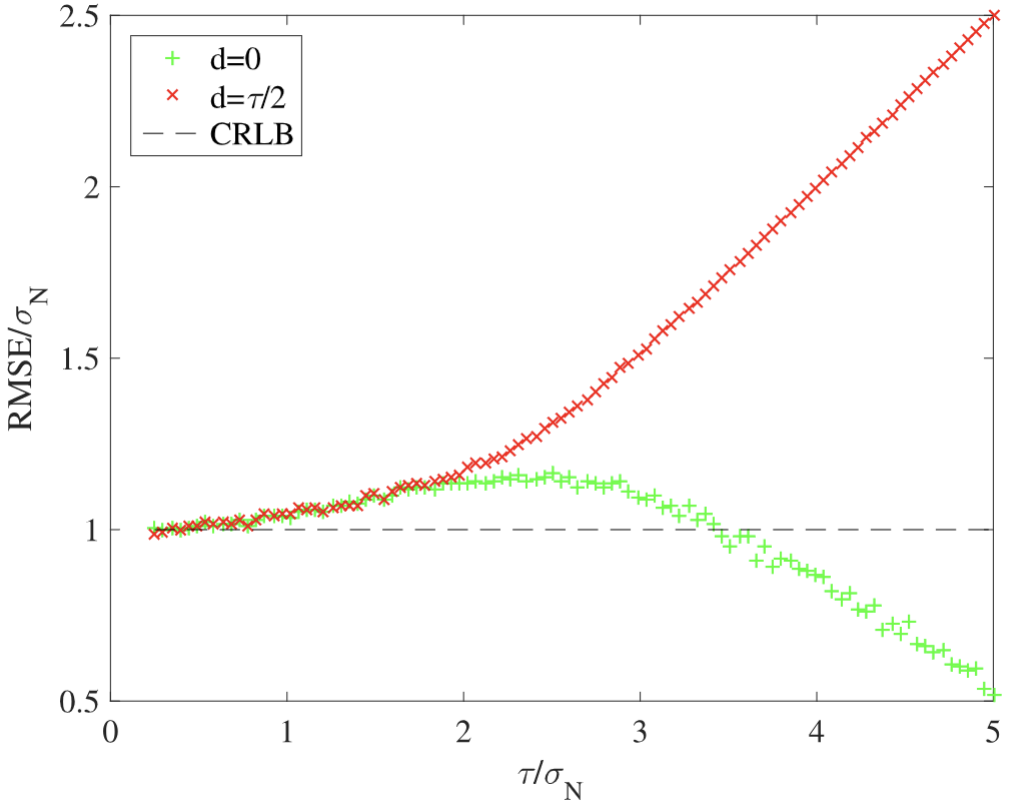

Figure 4.8. A numerical study where the mean value of a latent random variable is estimated with the quantized realization of the latent random variable. The study is repeated across two different cases, where there is no shift due to quantization (d=0) and where there is the worst case shift (d=$\tau/2$). Source: [17].

4.3 Simulation for Determining $\mathbb{R}^3$ and CeSAR Applicability

Table 4.4 summarizes the results for this simulation. The effectiveness of the algorithm to facilitate localization within $\mathbb{R}^3$ is shown under the "Standard" column. To properly frame these results we must compare them to the results for Simulation A, in Section 4.1, which examines the algorithms effectiveness in $\mathbb{R}^2$. When comparing Tables 4.4 and 4.3, it can be seen that, though there were 4 RRHs in use for both simulations, the algorithm effectiveness

was greatly reduced in three dimensional space, due to the addition of the height parameter to be estimated. In fact, if we next compare Tables 4.4 and 4.1, we observe that the results are similar to that of when only two RRHs are used. This equates to nearly an order of magnitude decrease in localization accuracy for all SCSs. It is noteworthy, however, that when going from 90% CEP to 95%, the position error increase ($\approx$ 27%) was substantially subdued in comparison to the increase that occurs when only 2 RRHs are used ($\approx$ 88%). The CDFs for $\mu = 0$ and $\mu = 4$, as seen in Figure 4.9, portray a marked increase in accuracy when using the CeSAR algorithm. Taking a closer look at the 90% and 95% CEPs for each numerology presented, at the worst point accuracy increases by $\approx$ 20% ($\mu = 0$) and $\approx$ 40% ($\mu = 4$). Both[5] of these surges in fidelity are significant and support the use of CeSAR to augment TA-based multilateration attacks in 5G.

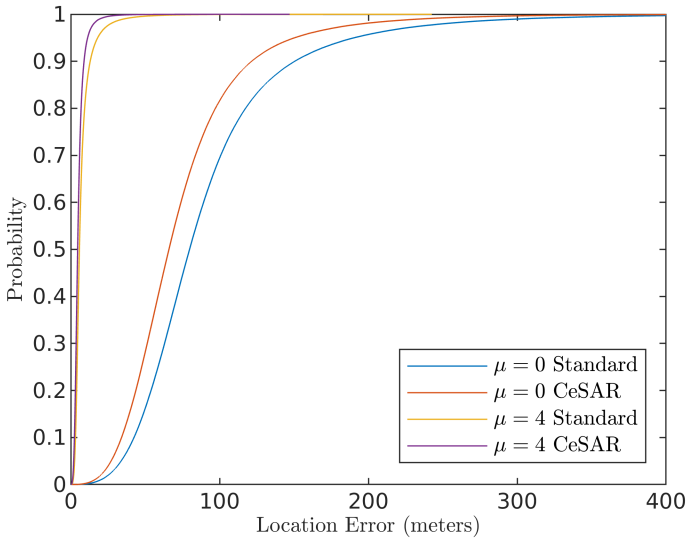

Figure 4.9. CDFs for both standard and CeSAR implementations for $\mu = 0$ and $\mu = 4$. Source: [15].

[5]For this simulation we looked at all numerologies, but as mentioned in Section 3.4 we believe that the use of CeSAR at numerologies greater than $\mu = 0 : 2$ would be non-trivial in practice at its best.

Table 4.4. CEPs across each numerology in $\mathbb{R}^3$, depicting a substantial decrease in localization accuracy for the standard algorithm, while highlighting the boon to accuracy provided by the CeSAR enhancement. Source: [15].

μ	**Standard**		**CeSAR enhanced**	
	CEP = 90% (m)	CEP = 95% (m)	CEP = 90% (m)	CEP = 95% (m)
0	150.42	190.38	121.47	150.39
1	86.68	112.60	65.35	81.84
2	47.41	62.93	34.17	43.08
3	25.12	34.06	17.45	22.11
4	13.00	18.02	8.87	11.28

4.4 Simulation of Real-world Use Cases

For the Kauai-based deployment, the position estimate errors for $\mu = 0 : 2$ were 30.12 m, 12.36 m, and 5.40 m on average. In the case of the UD deployment they were 8.64 m and 4.02 m. The final location estimates are rendered in Figure 4.10, with cyan markers denoting current system estimates, and pink UD system estimate. The spacing of the cyan markers as nigh linear and equidistant which begins at the $\mu = 0$ marker was an interesting result. We believe this is due entirely to the CeSAR sensors' distance estimates independence from noise, as well as how the sensors were placed in the environment. The UD deployment errors were consistent with expectations from the three previous results.

Figure 4.10. Position estimates for target UE. Source: [15].

4.5 Summary

Referring back to our four guiding questions in Chapter 3, we can now provide the following insight:

1. The number of RRHs and numerology in use affect the location estimates. As either increase, so too does the accuracy of the estimates. The most accurate localizations occur when 4 RRHs are in communication with the UE using the largest numerology ($\mu = 4$).
2. As σ increases, localization statistical efficiency increases as well, where the MSE of position estimates come to meet the CRLB across all numerologies. We also show numerically that in the case presented in the Conjecture, this is true across all values of σ.

3. The positioning algorithm performs significantly worse in $\mathbb{R}^3$ than in $\mathbb{R}^2$ (see Figure 4.4), especially at the lower numerologies. At higher numerologies, this method can still be considered effective with position errors being within 20 meters. The CeSAR algorithm will improve positioning across all numerologies. However, the use of CeSAR at numerologies greater than $\mu = 0 : 2$ is not practical, due to the physical limitations of signal collection at higher frequencies, in a setting using the procedure described in Table 2.3.
4. Our simulated environment showcased the effectiveness of our methods using actual 5G deployment locations on the island of Kauai mixed with simulated CeSAR receivers, as well as a simulated UD environment. The results from this were consistent with expectations laid out by the previous simulations' findings, giving credence to the algorithms efficacy when used with real wireless network deployment geometries.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5:
Conclusion

Here we have extended work presented in [17] to 5G. We now conclude with recommendations for future work.

5.1 Final Thoughts

This thesis has shown that in using the aggregated unencrypted TA commands of a target UE, an adversary can calculate a MLE of the target position, in short order. The speed and accuracy of this attack leads us to believe it would be quite possible to utilize this procedure for not only a localization attack, but tracking the UE as it moves through the environment. Therefore, we believe that as 5G technologies continue to come online, an end user's location privacy will be at even greater risk.

5.2 Recommendations for Follow-on Research

The major limitation of this study is that it had to be done entirely in simulation. The next logical step would be to conduct this attack in an operational 5G test bed, using both static UEs as well as those dynamically moving at varying velocities, testing how well the algorithm can localize a target from real signals. It would also be worthwhile to continue evaluating the algorithm efficacy as a tracking attack. Similarly, further investigation of the effects of σ on the efficiency of the algorithm is needed. We speculated that when observing a randomized UE location, the CRLB applies for all values of σ; however, a mathematical proof is needed to cement this notion.

Another worthwhile endeavor would be to investigate whether or not location accuracy will saturate based on the number of RRHs and/or CeSAR sensors in use in $\mathbb{R}^3$. This is of particular interest as in $\mathbb{R}^2$, the benefit of adding extra sensors past the maximum of 4 utilized by the network gives little increase to localization accuracy, and thus saturates very quickly. However, as shown in Simulation C (Section 4.3), the addition of a sensor resulted in a moderately large increase in fidelity in $\mathbb{R}^3$. We postulate that the addition of the vertical axis allows for many more possible sensor locations that do not result in co-linear

arrangements with the other sensors, providing solid information and enhancing the end estimate for each sensor added, up to a point. Defining what that point is would be valuable in determining proper deployment of the algorithm for each real-world use case.

A final area would be to investigate mitigation techniques. The difficulty here lies in that obvious efforts to mitigate this vulnerability reduce some of the major boons of 5G. If the TA is encrypted, this will increase latency, especially at times when the UE is moving through the environment and is receiving TAs often. If the PLS split is altered so that Layer 2 functionality is included at the RRH, this would allow for eCPRI built-in security measures associated with Internet Protocol (IP) (IPsec) and Ethernet (MACsec) to be implemented when sending data across the fronthaul. However, this added complexity will increase CapEx and OpEx, resulting in an increase in cost for network managers and end users alike, dispelling support for this option. Other, clever modalities for securely transferring fronthaul data, or specifically the TA information would be a major accomplishment.

List of References

[1] The Fourth Industrial Revolution: what it means, how to respond. (2016). World Economic Forum. [Online]. Available: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/. Accessed Oct. 27, 2020.

[2] "Cisco annual internet report (2018-2023)," White Paper, Cisco, March 2020.

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

[4] White House, "National Strategy to Secure 5G," 2020. Available: https://www.whitehouse.gov/wp-content/uploads/2020/03/National-Strategy-5G-Final.pdf

[5] White House, "National cyber strategy of the united states of america," 2018. Available: https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf

[6] Secretary of Defense, "Department of Defense (DOD) 5G Strategy," 2020. Available: https://www.cto.mil/wp-content/uploads/2020/05/DoD_5G_Strategy_May_2020.pdf

[7] J. D. Roth, M. Tummala, and J. W. Scrofani, "Cellular synchronization assisted refinement (CeSAR): A method for accurate geolocation in LTE-A networks," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 2016, pp. 5842–5850.

[8] Federal Communications Commission, "Report and order and further notice of proposed rulemaking on revision of the FCC rules to ensure compatibility with enhanced 911 emergency calling systems," pp. 96–264, 1996.

[9] Indoor Location Accuracy Timeline and Live Call Data Reporting Template. (2020). Federal Communications Commission. [Online]. Available: https://www.fcc.gov/public-safety-and-homeland-security/policy-and-licensing-division/911-services/general/location-accuracy-indoor-benchmarks. Accessed Nov. 2, 2020.

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

[11] J. J. Caffery, "A new approach to the geometry of TOA location," in *Vehicular Technology Conference Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conference (Cat. No.00CH37152)*, 2000, vol. 4, pp. 1943–1949 vol.4.

[12] V. Lazarev, G. Fokin, and I. Stepanets, "Positioning for location-aware beamforming in 5G ultra-dense networks," in *2019 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*, 2019, pp. 136–139.

[13] Z. He, Y. Ma, and R. Tafazolli, "Indoor TDOA mobile positioning with clock drift and its Cramer-Rao bound," in *European Wireless 2013; 19th European Wireless Conference*, 2013, pp. 1–5.

[14] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018.

[15] A. Schacht, K. Foster, and J. Roth, "Location Privacy in the Era of 5G," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021.

[16] F. C. Commission, "Fourth report and order on wireless E911 location accuracy requirements," pp. 15–9, Jan. 2015.

[17] J. D. Roth, M. Tummala, J. C. McEachen and J. W. Scrofani, "On location privacy in LTE networks," in *IEEE Transactions on Information Forensics and Security*, 2017, pp. 1358–1368.

[18] B. Liu, W. Zhou, T. Zhu, L. Gao, and Y. Xiang, "Location privacy and its applications: A systematic study," *IEEE Access*, vol. 6, pp. 17 606–17 624, 2018.

[19] X. Yang, L. Gao, J. Zheng, and W. Wei, "Location privacy preservation mechanism for location-based service with incomplete location data," *IEEE Access*, vol. 8, pp. 95 843–95 854, 2020.

[20] M. Liyanage, J. Salo, A. Braeken, T. Kumar, S. Seneviratne, and M. Ylianttila, "5G privacy: Scenarios and solutions," in *2018 IEEE 5G World Forum (5GWF)*, 2018, pp. 197–203.

[21] M. Haus, M. Waqas, A. Y. Ding, Y. Li, S. Tarkoma, and J. Ott, "Security and privacy in device-to-device (D2D) communication: A review," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1054–1079, 2017.

[22] Defense Innovation Board, "The 5G Ecosystem: Risks & Opportunities for DOD," DIB, Tech. Rep., 2019.

[23] China Mobile Research Institute, "C-RAN: the road towards green RAN," Tech. Rep., 2013.

[24] S. Perrin, "Evolving to an open C-RAN architecture for 5G," in *Heavy Reading*, 2017.

[25] A. M. Corporation. (2018, Feb.). Cloud RAN and eCPRI fronthaul in 5G networks. [Online]. Available: https://medium.com/5g-nr/cloud-ran-and-ecpri-fronthaul-in-5g-networks-a1f63d13df67

[26] eCPRI specification V1.2. (2018). Common Public Radio Interface; eCPRI Interface Specification. [Online]. Available: http://www.cpri.info/downloads/eCPRI_v_1_2_2018_06_25.pdf

[27] 3GPP TR 38.801 (V14.0.0), "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on new radio access technology; Radio access architecture and interfaces (Release 14)," Mar. 2017.

[28] C. Ranaweera, E. Wong, A. Nirmalathas, C. Jayasundara and C. Lim, "5G C-RAN architecture: A comparison of multiple optical fronthaul networks," in *2017 International Conference on Optical Network Design and Modeling (ONDM)*, 2017, pp. 1–6.

[29] 3GPP TS 38.213, release 16, (v16.1.0), "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical layer procedures for control," Apr. 2020.

[30] N. Patriciello, S. Lagen, L. Giupponi, and B. Bojovic, "5G new radio numerologies and their impact on the end-to-end latency," in *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2018, pp. 1–6.

[31] National Instruments. (2018). 5G New Radio: Introduction to the Physical Layer. [Online]. Available: http://www.ni.com/en-us/innovations/wireless/5g/new-radio.html#whitepaper

[32] 3GPP TS 36.321, release 16, (v16.0.0), "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification," Apr. 2020.

[33] E. Dahlman, S. Parkvall, and J. Skold, *5G NR the next generation wireless access technology*. Elsevier Ltd., 2018.

[34] 3GPP TS 38.211, release 16, (v16.1.0), "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical channels and modulation," Apr. 2020.

[35] A. Schacht, J. Long, and J. Roth, "Timing management in 5G and its implications for location privacy," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.

[36] J. J. Caffery and G. L. Stuber, "Overview of radiolocation in CDMA cellular systems," in *IEEE Commun. Mag.*, 1998, pp. 38–45.

[37] K.W. Cheung, H.C. So, W.-K. Ma, and Y.T. Chan, "Least squares algorithms for time-of-arrival-based mobile location," in *IEEE Transactions on Signal Processing*, 2004, pp. 1121–1128.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California