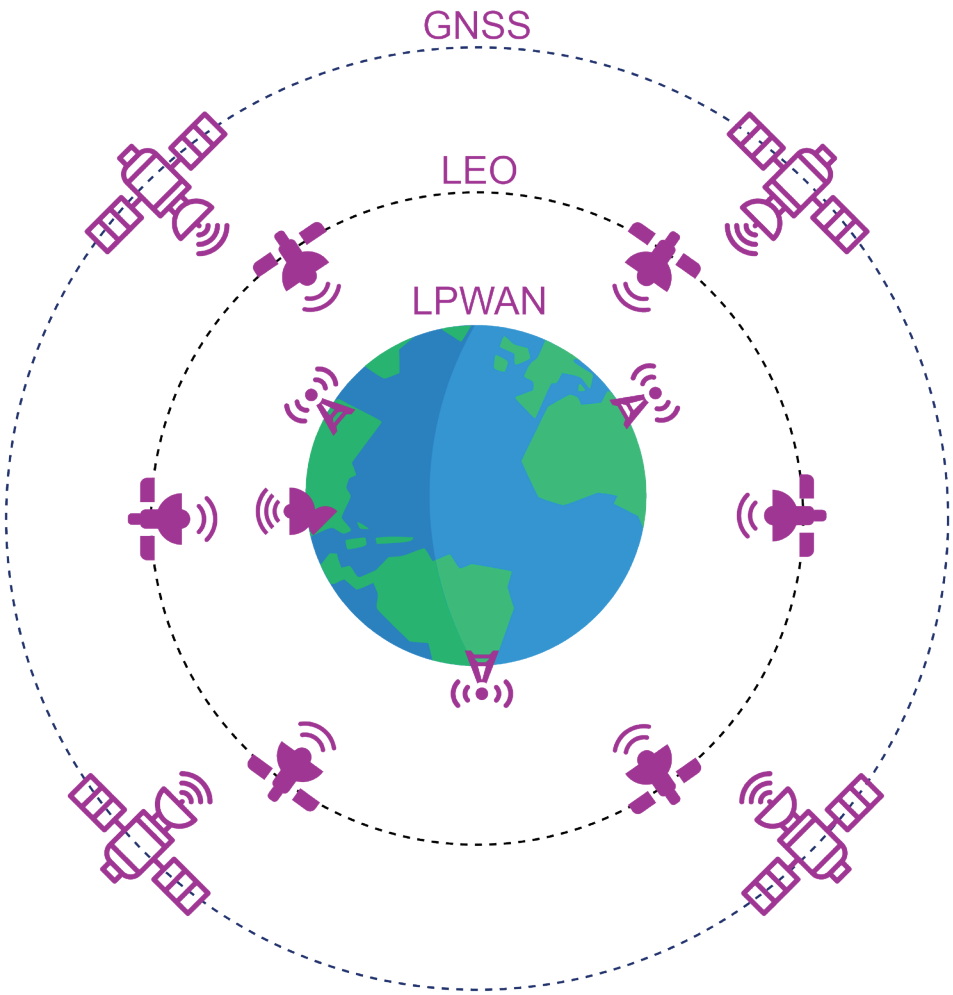
Energy-efficient Positioning for the Internet of Things

Thomas Janssen



Supervisors **Prof. Dr. Maarten Weyn | Dr. Rafael Berkvens**

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Applied Engineering
Faculty of Applied Engineering | Antwerp, 2022

University of Antwerp

Faculty of Applied Engineering

Energy-efficient Positioning for the Internet of Things

Thesis submitted in fulfilment of the requirements for the degree of
Doctor in Applied Engineering
at University of Antwerp

Thomas Janssen

Supervisors
Prof. dr. Maarten Weyn
Dr. Rafael Berkvens

Antwerp, 2023

Jury

Chairman
Prof. dr. Jeroen Famaey, University of Antwerp, Belgium

Supervisors
Prof. dr. Maarten Weyn, University of Antwerp, Belgium
Dr. Rafael Berkvens, University of Antwerp, Belgium

Members
Prof. dr. Johann Marquez-Barja, University of Antwerp, Belgium
Dr. Grigorios G. Anagnostopoulos, University of Geneva & University of Applied Sciences and Arts, Western Switzerland
Dr. Tyler G. R. Reid, Stanford University & Xona Space Systems, United States of America

Contact
Thomas Janssen
University of Antwerp
Faculty of Applied Engineering - Electronics-ICT
IDLab research group
Sint-Pietersvliet 7, 2020 Antwerp, Belgium
M: thomas.janssen@uantwerpen.be

ISBN 978-90-57287-72-5
Wettelijk depot D/2022/12.293/35

Abstract

LOCATION data is required for a plethora of Internet of Things (IoT) applications running on billions of mobile devices worldwide. A few example applications include asset tracking, search-and-rescue operations and the scientific monitoring of air or water quality.

Global Navigation Satellite Systems (GNSSs), such as Global Positioning System (GPS) or Galileo, have been established as the standard for worldwide localization. However, the rapidly increasing need to locate IoT devices in recent years has exposed several shortcomings of traditional GNSS approaches. These limitations include the weak signal propagation in indoor and dense environments, a high energy consumption and the inability to communicate a location to a remote end user. Therefore, several industries have shown an increasing demand for alternative, innovative, and energy-efficient positioning solutions that are more suited in an IoT context.

In contrast to GNSS, Low Power Wide Area Networks (LPWANs) were designed for energy-efficient communication of small sensor readings in a metropolitan area. In this type of terrestrial networks, thousands of IoT devices can send a message to nearby gateways, which in turn deliver the message to a central server. Interestingly, the uplink communication signals can be exploited to locate the mobile transmitter. Such a localization approach benefits from the low-power and low-cost LPWAN communication, as well as from the coverage in indoor environments.

Another very promising alternative to GNSS is the use of satellites in Low Earth Orbit (LEO) for Positioning, Navigation and Timing (PNT). Driven by the recent 'New Space' movement, the commercialization of the space market has opened the door to a myriad of opportunities. The thousands of LEO satellite launches of Iridium, SpaceX, Amazon, OneWeb and many others enable applications such as high-quality satellite telephony, worldwide Internet access and smart agriculture through Earth Observation. Providing PNT services through LEO satellites in an energy-efficient way will only improve the value of these applications in the emerging market of satellite IoT.

The objective of this thesis is to investigate innovative, large-scale, and energy-efficient positioning technologies and techniques in the context of IoT. I examine how wireless networks, either terrestrial or space-based, can be leveraged for locating IoT devices, and how I can improve their positioning performance.

The performance analysis and optimization of localization using LPWAN technologies constitute a significant part of the work in this thesis. To this end, three major LPWAN technologies are investigated: Sigfox, LoRaWAN and Narrowband IoT (NB-IoT). Localization experiments are carried out using real-world measurement data collected in Antwerp, Belgium. Within these experiments, I analyze the performance of Received Signal Strength

(RSS)-based positioning algorithms. More specifically, I evaluate different path loss models in range-based algorithms and apply Machine Learning to optimize the performance of RSS-based fingerprinting methods. Furthermore, I discuss how the positioning performance can be further improved through changes in network infrastructure and User Equipment (UE).

In the final part of this work, I conduct a survey for the European Space Agency (ESA) with the goal to explore innovative space-based PNT solutions, again with a focus on low energy consumption. I analyze the state-of-the-art performance of novel GNSS approaches, such as Assisted GNSS (A-GNSS) and snapshot processing techniques (S-GNSS). When compared to conventional pseudoranging, these techniques significantly reduce the overall energy consumption of the UE. Moreover, my survey covers the potential of Doppler positioning techniques leveraging LEO satellite Signals of Opportunity (SOOP), as well as the promising dedicated LEO-PNT systems under development.

IoT-enabled devices have different constraints and application requirements. Therefore, the important trade-off between positioning accuracy and energy consumption is discussed throughout this work. There exists no one-size-fits-all technology that performs excellent in any use case in terms of these two parameters. Thus, interoperability between technologies is key to enable global energy-efficient communication and positioning applications.

Samenvatting

LOCATIEGEGEVENS zijn noodzakelijk voor een groot aantal Internet of Things (IoT) toepassingen die op miljarden mobiele apparaten draaien. Enkele voorbeelden van toepassingen zijn het traceren van goederen, opsporings- en reddingsoperaties en het wetenschappelijk monitoren van lucht- of waterkwaliteit.

Global Navigation Satellite Systems (GNSSs), zoals Global Positioning System (GPS) of Galileo, zijn de norm geworden voor wereldwijde lokalisatie. De snel toenemende noodzaak om IoT-apparaten te lokaliseren in de afgelopen jaren heeft echter verschillende tekortkomingen van de traditionele GNSS-benaderingen blootgelegd. Deze beperkingen omvatten de zwakke signaalpropagatie in binnen- en stadsomgevingen, een hoog energieverbruik en het ontbreken van de mogelijkheid om een locatie te communiceren naar een eindgebruiker op afstand. Daarom is er in verschillende industrieën een toenemende vraag naar alternatieve, innovatieve en energie-efficiënte positioneringsoplossingen die meer geschikt zijn in een IoT-context.

In tegenstelling tot GNSS zijn Low Power Wide Area Networks (LPWANs) ontworpen voor energiezuinige communicatie van kleine sensorwaarden in een grootstedelijk gebied. In dit soort netwerken kunnen duizenden IoT-apparaten een bericht verzenden naar grondstations in de buurt, die het bericht op hun beurt afleveren aan een centrale server. Interessant is dat deze communicatiesignalen kunnen worden gebruikt om de mobiele zender te lokaliseren. Een dergelijke lokalisatiebenadering geniet van de LPWAN-communicatie met laag vermogen en lage kosten, alsook van de dekking in binnenomgevingen.

Een ander veelbelovend alternatief voor GNSS is het gebruik van satellieten in een lagere baan rond de Aarde (Low Earth Orbit - LEO) voor plaatsbepaling, navigatie en tijdsbepaling (Positioning, Navigation and Timing - PNT). Onder impuls van de recente 'New Space' beweging heeft de commercialisering van de ruimtevaartmarkt de deur geopend naar een groot aantal mogelijkheden. De duizenden LEO-satellietlanceringen van Iridium, SpaceX, Amazon, OneWeb en vele anderen maken toepassingen mogelijk zoals satelliettelefonie van hoge kwaliteit, wereldwijde internettoegang en slimme landbouw door aardobservatie. Het op energie-efficiënte wijze aanbieden van PNT-diensten via LEO-satellieten zal de waarde van deze toepassingen in de opkomende markt van satelliet-IoT alleen maar verhogen.

Het doel van deze thesis is om innovatieve, grootschalige en energie-efficiënte positioneringstechnologieën en -technieken in een IoT context te onderzoeken. Ik onderzoek hoe draadloze netwerken, zowel op Aarde als in de ruimte, kunnen worden gebruikt voor het lokaliseren van IoT-apparaten, en hoe ik hun positioneringsprestaties kan verbeteren.

De prestatie-analyse en optimalisatie van lokalisatie met behulp van LPWAN-technologieën vormen een belangrijk deel van het werk in deze thesis. Daartoe worden drie belang-

rijke LPWAN-technologieën onderzocht: Sigfox, LoRaWAN en Narrowband IoT (NB-IoT). Lokalisatie-experimenten worden uitgevoerd op basis van meetgegevens verzameld in Antwerpen. Binnen deze experimenten analyseer ik de prestaties van positioneringsalgoritmen gebaseerd op de ontvangen signaalsterkte (Received Signal Strength - RSS). Meer bepaald evalueer ik verschillende signaalsterkte modellen in afstand-gebaseerde algoritmes en pas ik Machine Learning toe om de prestaties van RSS-gebaseerde fingerprinting methodes te optimaliseren. Verder bespreek ik hoe de positioneringsprestaties verder kunnen worden verbeterd door veranderingen aan netwerkinfrastructuur en gebruikersapparatuur.

In het laatste deel van dit werk voer ik een onderzoek uit voor het Europees Ruimteagentschap (ESA) met het doel innovatieve ruimtegebaseerde PNT-oplossingen te onderzoeken, opnieuw met de nadruk op een laag energieverbruik. Ik analyseer de prestaties van de nieuwste GNSS benaderingen, zoals Assisted GNSS (A-GNSS) en snapshot verwerkingstechnieken (S-GNSS). In vergelijking met traditionele GNSS technieken, zorgen deze technieken voor een significante vermindering van het totale energieverbruik van het apparaat. Bovendien behandelt mijn onderzoek het potentieel van Doppler-positioneringstechnieken die gebruik maken van opportunistische signalen van LEO-satellieten, alsook de veelbelovende LEO satellietsystemen die specifiek voor lokalisatie-doeleinden ontwikkeld worden.

IoT-apparaten hebben verschillende beperkingen en toepassingsvereisten. Daarom wordt in dit werk de belangrijke afweging tussen positioneringsnauwkeurigheid en energieverbruik besproken. Er bestaat geen 'one-size-fits-all' technologie die het best presteert in eender welke toepassing op vlak van deze twee parameters. Aldus is interoperabiliteit tussen technologieën de sleutel om wereldwijde energie-efficiënte communicatie en lokalisatie toepassingen te verwezenlijken.

Acknowledgements

IN the past four years, I met a lot of incredible people, who helped me shape the person I am today. Therefore, I want to express my gratitude to the following people.

First and foremost, I would like to express my deepest appreciation to my supervisors, Prof. Maarten Weyn and Dr. Rafael Berkvens. They gave me the opportunity to start a PhD and guided me along the way. Through their unique combination of industrial know-how and academic insights, I can hardly imagine a better way of being supervised. Maarten and Raf, thanks a lot for the invaluable feedback and endless support, you are amazing.

I want to thank all other members of my doctoral committee (Prof. Jeroen Famaey, Prof. Johann Marquez-Barja, Dr. Grigorios Anagnostopoulos and Dr. Tyler Reid), for your helpful feedback on this dissertation.

This endeavor would not have been possible without the funding received by the FWO (Fonds voor Wetenschappelijk Onderzoek, grant number 1S03821N). To all taxpayers: Thank you, I hope your money was well spent.

I will always remember the day Wouter Rymenants walked into our office, and asked me if I was interested in a project for the European Space Agency (ESA). This day literally took my research on the ground to a higher level. After a few years, I presented our hard work at ESA. Never, not in my wildest dreams, did I ever imagine I could end up in this situation. Thank you so much, Wouter!

I had the pleasure of working with extremely talented people at IDLab, imec and the University of Antwerp. Many thanks to the people in the LoPow team in particular. Special thanks go to Michiel, Mats and Dragan, for the countless ping-pong matches and after work drinks, successfully putting things in perspective. I would be remiss in not mentioning Guido Van Landeghem, for the educational and fun lab we taught together.

I could not have undertaken this journey without the support of friends and family. Thank you to Liam & Macey, for the countless boardgame nights, keeping me from working in the weekend. Thank you to all makers at iMagineLab, for inspiring me through your creative minds. Last but not least, words cannot express my gratitude to Laura, as well as to my parents. Thank you for always being there for me and for believing in me. I promise to keep my feet on the ground and not go to Mars. At least not without you.

"The journey is the reward" ~ Steve Jobs

Contents

List of Publications

Journal Articles

1. **T. Janssen**, M. Weyn, and R.Berkvens, "Localization in Low Power Wide Area Networks Using Wi-Fi Fingerprints", *Applied Sciences*, vol. 7, no. 9, p. 936, Sep. 2017, https://doi.org/10.3390/app7090936.

2. F. Lemic, V. Handziski, M. Aernouts, **T. Janssen**, R. Berkvens, A. Wolisz, and J. Famaey, "Regression-Based Estimation of Individual Errors in Fingerprinting Localization", *IEEE Access*, vol. 7, pp. 33 652–33 664, 2019, https://doi.org/10.1109/ACCESS.2019.2903880.

3. **T. Janssen**, N. BniLam, M. Aernouts, R. Berkvens, and M. Weyn, "LoRa 2.4 GHz Communication Link and Range", *Sensors*, vol.20, no. 16, p. 4366, Aug. 2020, https://doi.org/10.3390/s20164366.

4. **T. Janssen**, R. Berkvens, and M. Weyn, "Benchmarking RSS-based localization algorithms with LoRaWAN", *Internet of Things*, vol. 11, p. 100235, Sep. 2020, https://doi.org/10.1016/j.iot.2020.100235.

5. **T. Janssen**, R. Berkvens, and M. Weyn, "RSS-Based Localization and Mobility Evaluation Using a Single NB-IoT Cell", *Sensors*, vol. 20, no. 21, p. 6172, Oct. 2020, https://doi.org/10.3390/s20216172.

6. M. Aernouts, **T. Janssen**, R. Berkvens, and M. Weyn, "LoRa Localization: With GNSS or Without?", *IEEE Internet of Things Magazine*, pp. 1–7, 2022, https://doi.org/10.1109/IOTM.001.2200019.

7. **T. Janssen**, A. Koppert, R. Berkvens, and M. Weyn, "A Survey on IoT Positioning leveraging LPWAN, GNSS and LEO-PNT *(Submitted on July 30, 2021)*", *IEEE Internet of Things Journal*, pp 1–25, 2023.

8. C. Baek, R. Halili, **T. Janssen**, N. BniLam, and M. Weyn, "Angle of Arrival Estimation for GPS Signals with Low-Cost Software Defined Radios", in *Transactions on Aerospace and Electronic Systems*. IEEE, 2022 (Submitted on August 17, 2023).

Conference Proceedings

1. **T. Janssen**, M. Aernouts, R. Berkvens, and M. Weyn, "Outdoor Fingerprinting Localization Using Sigfox", in *2018 International Conference on Indoor Positioning and*

Indoor Navigation (IPIN). IEEE, sep 2018, pp. 1–6, https://doi.org/10.1109/IPIN.2018.8533826.

2. **T. Janssen**, R. Berkvens, and M. Weyn, "Comparing Machine Learning Algorithms for RSS-Based Localization in LPWAN", in *Proceedings of the 14th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2019)*. Editors: L. Barolli, P. Hellinckx, J. Natwichai, Springer International Publishing, 2019, pp. 726–735, https://doi.org/10.1007/978-3-030-33509-0_68.

3. **T. Janssen**, M. Weyn, and R. Berkvens, "A Primer on Real-world RSS-based Outdoor NB-IoT Localization", in *2020 International Conference on Localization and GNSS (ICL-GNSS)*, IEEE, Jun. 2020, pp. 1–6, https://doi.org/10.1109/ICL-GNSS49876.2020.9115578.

4. W. Van Uytsel, **T. Janssen**, R. Halili, and M. Weyn, "Exploring Positioning through Pseudoranges using Low Earth Orbit Satellites", in *Proceedings of the 17th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2022)*. Editor: Editors: L. Barolli, Springer International Publishing, 2022, https://doi.org/10.1007/978-3-031-19945-5_28.

List of Figures

List of Tables

Acronyms

*k*NN	*k*-Nearest-Neighbors
3GPP	3rd Generation Partnership Project
5G	Fifth Generation
A-GNSS	Assisted GNSS
A-GPS	Assisted GPS
ADR	Adaptive Data Rate
AI	Artificial Intelligence
AIS	Automatic Identification System
ANN	Artificial Neural Network
AoA	Angle of Arrival
AoD	Angle of Departure
API	Application Programming Interface
ARTIC	Argos Receiver Transmitter with Integrated Control
BLE	Bluetooth Low Energy
BOM	Bill Of Materials
BPSK	Binary Phase Shift Keying
BSSID	Basic Service Set Identifier
CDF	Cumulative Distribution Function
CGI	Cell Global Identity
CL	Cooperative Localization
COST	European Cooperation in Science and Technology
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CRLB	Cramer-Rao Lower Bound
CS	Compressed Sensing
CSI	Channel State Information
CSS	Chirp Spread Spectrum
DAD	Duplicate Address Detection
DOP	Dilution Of Precision
E-CID	Enhanced Cell ID
ECC	Electronic Communication Committee
eDRX	Extended Discontinuous Reception
EKF	Extended Kalman Filter

eNB	Evolved Node B
ESA	European Space Agency
FSCM	Frequency Shift Chirp Modulation
FSPL	Free Space Path Loss
GNSS	Global Navigation Satellite System
GPC	Global Processing Center
GPS	Global Positioning System
GSA	European GNSS Agency
IDP	Indoor Dominant Path
IEEE	Institute of Electrical and Electronics Engineers
IGS	International GNSS Service
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
ION	Institute of Navigation
IoT	Internet of Things
ISM	Industrial, Scientific, and Medical
ITU	International Telecommunication Union
JCNS	Joint Communication And Sensing
JWST	James Webb Space Telescope
LEO	Low Earth Orbit
LoS	Line of Sight
LPWAN	Low Power Wide Area Network
LR-FHSS	Long Range Frequency Hopping Spread Spectrum
LS	Least Squares
LTE	Long Term Evolution
LTE-M	Long Term Evolution for Machine Type Communication
MAC	Media Access Control
MCL	Maximum Coupling Loss
MEO	Medium Earth Orbit
ML	Machine Learning
MSE	Mean Square Error
NB-IoT	Narrowband-IoT
NCL	Non-Cooperative Localization
NFC	Near Field Communication
NLoS	Non-Line of Sight
NN	Neural Network
NORAD	North American Aerospace Defense Command
NR	New Radio
NS3	Network Simulator 3
NTN	Non-Terrestrial Network

OLS	Ordinary Least Squares
OTDoA	Observed Time Difference of Arrival
PCA	Principal Components Analysis
PNT	Positioning, Navigation and Timing
PPP	Precise Point Positioning
PSM	Power Saving Mode
PTT	Platform Terminal Transmitter
PVT	Position, Velocity and Time
QoS	Quality of Service
RAM	Random Access Memory
RF	Radio Frequency
RFID	Radio Frequency Identification
RINEX	Receiver Independent Exchange Format
RMS	Root Mean Square
RMSE	Root Mean Square Error
RSRP	Reference Signal Received Power
RSS	Received Signal Strength
RSTD	Reference Signal Time Difference
RTK	Real-Time Kinematics
RTT	Round-Trip Time
RX	Receive
S-GNSS	Snapshot GNSS
SDR	Software Defined Radio
SF	Spreading Factor
SGD	Stochastic Gradient Descent
SIM	Subscriber Identity Module
SNR	Signal-to-Noise Ratio
SoC	System on Chip
SoOP	Signals of Opportunity
SOTA	state-of-the-art
SPS	Standard Positioning Service
SRD	Short Range Devices
STAN	Simultaneous Tracking And Navigation
SVM	Support Vector Machine
SVR	Support Vector Regression
TDoA	Time Difference of Arrival
TLE	Two Line Element
ToA	Time of Arrival
ToF	Time of Flight
TRL	Technology Readiness Level
TTFF	Time To First Fix
TTN	The Things Network
TX	Transmit

UE	User Equipment
UNB	Ultra-Narrow Band
VHF	Very High Frequency
WAP	Wireless Access Point

Chapter 1

Introduction

THE introduction of the Internet substantially changed the way we live today. We are now able to search for online information, videocall with our family and organize digital conference meetings with people from across the globe in times of a pandemic. A new revolution in the history of the Internet is the Internet of Things (IoT). As the name indicates, the IoT is a network that connects physical things or smart devices to each other via the Internet. By 2025, it is estimated that 152,200 new devices will connect to the IoT every minute, reaching a total of more than 27 billion connected devices and generating $4-11 trillion in economic value [1, 2].

The problem with this massive number of IoT devices worldwide is that they need to be monitored properly. In many IoT applications, location-awareness is of crucial importance. Example applications include the tracking of *things* like industrial assets, moving sensors, animals in the wild or your beloved dog. In each of these applications, energy consumption may be equally or even more important than positioning accuracy, given the limited battery capacity of the device.

Global Navigation Satellite Systems (GNSSs), such as Global Positioning System (GPS) or Galileo, have been used for decades to locate a device on Earth. Using the latest satellite augmentation systems and services, the positioning accuracy of GNSS has increased from a few meters to a few centimeters. However, several problems arise when leveraging GNSS to locate IoT devices. First, GNSS were not designed with low energy consumption in mind. The battery lifetime of a typical IoT device with a GNSS receiver does not exceed a few months [3, 4]. While efforts exist to lower the overall energy consumption, IoT devices often require a lifetime of multiple years without recharging batteries, e.g., when tracking bird migrations over their full lifespans. Obviously, the actual energy consumption highly depends on the environment and the application requirements [5]. A second disadvantage of GNSS is the weak signal reception on Earth. Moreover, the signal from Medium Earth Orbit (MEO) does not penetrate through walls. Therefore, localization in indoor or dense urban environment is limited or impossible. Third, GNSSs suffer from unintentional and intentional interference. The latter is caused by jamming and spoofing attacks, which poses security threats in autonomous driving, for instance [6]. Finally, a GNSS receiver is only able to locate itself. However, in many IoT applications, the user needs to know the position of the mobile device, which consequently requires the device to send the estimated GNSS coordinate over a wireless network to the user application, consuming even more energy.

In order to establish communication between IoT devices and their users, Low Power Wide Area Network (LPWAN) standards such as Sigfox [7], LoRaWAN [8], Narrowband-IoT (NB-IoT) and Long Term Evolution for Machine Type Communication (LTE-M) [9] have been developed. Generally, LPWAN standards enable IoT devices to exchange small messages over long distances while maintaining a very low energy consumption, which allows these devices to work for multiple years on small batteries. This meets the need of an endless number of sensing applications in the domains of healthcare [10], environmental monitoring [11, 12, 13, 14], wildlife monitoring [15], smart agriculture [16], logistics [17] and smart cities [18, 19], amongst many others [20, 21, 22].

As an alternative to GNSS, LPWAN technologies provide a modern solution for both communication and localization. Figure 1.1 lists only a few example IoT applications in which location context is important. LPWAN signals can be used to determine the location of an IoT device [5, 23, 21]. More specifically, we can exploit or derive the characteristics of sub-GHz LPWAN signals such as the Received Signal Strength (RSS), Time of Arrival (ToA) or the Angle of Arrival (AoA) to locate a mobile transmitter. Fingerprint-based and range-based algorithms are commonly implemented in indoor positioning systems. These algorithms can be applied and optimized for use in outdoor environments as well. Due to the star topology of most LPWANs, end devices cannot directly communicate with other nearby end devices. Therefore, Non-Cooperative Localization (NCL) schemes are mostly considered in LPWAN-based localization. In contrast, Cooperative Localization (CL) schemes can lead to higher accuracies; however, they do not outweigh the increased device complexity and energy consumption. For the same reason, it is common to perform the localization in the cloud, instead of on the device itself.

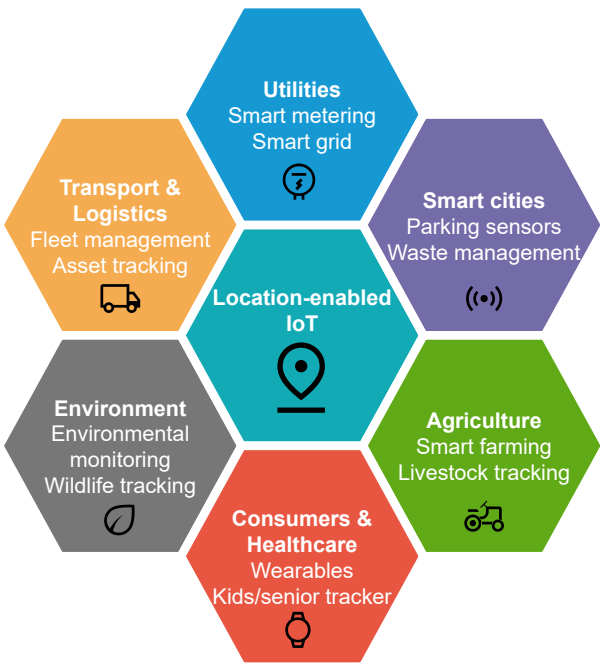


Figure 1.1: Location context is important in many IoT applications.

Terrestrial networks only cover about 10% of the world's surface [24]. Due to the economical and geographical reasons, they simply cannot cover the entire planet (e.g., oceans, jungles, deserts and mountainous regions) [25, 26]. Therefore, we have to look up to the sky. The *New Space* movement focuses on removing barriers to increase the access to space. Where traditional space missions were managed and funded by large governmental agencies, the New Space era has opened the door for independent private space missions. The reusable rockets and ridesharing missions of Elon Musk's SpaceX are an excellent example of this approach. Besides large companies like SpaceX, Amazon and OneWeb, a myriad of start-ups with innovative ideas also find their way to space [27]. Every single week of 2022, multiple satellites and rockets were launched [28]. The commercialization of the space market was the key for enabling a plethora of new opportunities in space. Nowadays, we are able to navigate, make a phone call, browse the Internet, or observe live satellite imagery from almost anywhere on Earth thanks to satellite constellations.

In the last few years, many satellites are launched into the emerging Low Earth Orbit (LEO). Satellites in LEO are around 20 times closer to Earth compared to GNSS satellites. Therefore, a LEO satellite signal reaches the Earth faster, enabling low-latency communication. Moreover, the signal oftentimes arrives more powerful, improving penetration in harsh environments. While most of the currently available systems primarily focus on communication, there are many reasons to explore a LEO Positioning, Navigation and Timing (PNT) system that overcomes the aforementioned limitations of GNSS in MEO. Such a system can provide a backup or even an alternative for GNSS. The estimated cost of a GNSS outage to the UK alone is over £1B per day [29]. Moreover, it was estimated that a single GPS-III satellite costs as much as 300 small LEO satellites with matching visibility and geometry [30]. Finally, LEO satellites provide more robustness to interference, more resistance to spoofing attacks and, similar to LPWAN, the ability to communicate the location of a remote device.

In this thesis, I explore these novel LPWAN and LEO satellite technologies, as well as novel GNSS techniques, and I evaluate their feasibility for localization.

1.1 Motivation

Traditional GNSSs were not designed with low energy consumption in mind. Our planet has a limited number of resources, and at the time of writing the amount of available renewable energy is insufficient. Therefore, new localization technologies must consider energy-efficiency in their design, paving the road towards a sustainable IoT. Moreover, a few years ago, alternatives to GNSS for large-scale localization barely existed. These were two motivating factors to research alternative, large-scale and energy-efficient positioning technologies.

With technological advancements such as autonomous driving and smart agriculture, new positioning challenges arise. Key performance requirements such as positioning accuracy, reliability and energy-efficiency will continue to improve. Every 30 years, the positioning error decreases by an order of magnitude due to new technological infrastructure [30]. LEO satellites are expected to play a major role in this trend [31]. Furthermore, LEO satellites and LPWAN gateways feature energy-efficient communication, which is lacking in GNSS.

Personally, I have always had a passion for technology, especially for technology used to navigate around the world. About a decade ago, I learned about Geocaching. Whenever I have the time, I still explore new places I otherwise wouldn't go to by discovering hidden treasures around the world. What I also find interesting, is how localization technology impacts our everyday lives, aside from being used in an industrial context. All of this inspired me to investigate new ways to locate anything, anywhere, anytime.

1.2 Contributions

The main contributions of this thesis can be summarized as follows:

1. **Improve RSS-based localization with Sigfox and LoRaWAN.** (Chapters 3–6)
   In 2018, Aernouts et al. set a baseline for localization in LPWAN [32]. I improved the existing RSS-based fingerprinting and ranging algorithms leveraging innovative Machine Learning (ML) algorithms. Using the original Sigfox and LoRaWAN data sets, I was able to double the localization accuracy. Moreover, I experimented with sending W-Fi fingerprints over LPWAN to further increase the accuracy. Lastly, I investigated the feasibility of using LoRa modulation in the 2.4 GHz band for localization purposes.

2. **Provide first insights in RSS-based localization with NB-IoT.** (Chapter 7)
   Shortly after the deployment of a public NB-IoT network in Belgium, I collected experimental measurements in a realistic environment in the city of Antwerp. By doing so, I was able to perform the first RSS-based localization experiments with NB-IoT. I shared my data set with interested researchers, who used it to test new positioning algorithms or to apply coverage analyses. Moreover, I raised the issue that changes in User Equipment (UE) hardware and network infrastructure are needed to improve cellular IoT positioning and mobility. I am delighted to announce some UE manufacturers already implemented these proposed changes.

3. **Explore the possibility for PNT services leveraging LEO satellites.** (Chapter 8)
   After having submitted a project proposal for the European Space Agency (ESA), I received funding to bring IoT localization to space. By replacing the LPWAN gateways with LEO satellites and discussing the potential PNT performance, I provided ESA with insights into what strategic steps need to be taken for a fully functional LEO-PNT system. Moreover, I investigated a Signals of Opportunity (SoOP) approach.

4. **Compare energy-efficient positioning technologies and techniques.** (Chapter 8)
   During the aforementioned project for ESA, I analyzed the state-of-the art of large-scale and energy-efficient positioning systems. Through a survey, I compare the performance of LPWAN localization, novel GNSS approaches such as Assisted GNSS (A-GNSS) and Snapshot GNSS (S-GNSS), and positioning techniques leveraging LEO satellites.

For each of these contributions, more details can be found in the following publications:

Contribution 1:

1. **T. Janssen**, M. Weyn, and R.Berkvens, "Localization in Low Power Wide Area Networks Using Wi-Fi Fingerprints", *Applied Sciences*, vol. 7, no. 9, p. 936, Sep. 2017, https://doi.org/10.3390/app7090936.
2. **T. Janssen**, M. Aernouts, R. Berkvens, and M. Weyn, "Outdoor Fingerprinting Localization Using Sigfox", in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, sep 2018, pp. 1–6, https://doi.org/10.1109/IPIN.2018.8533826.
3. **T. Janssen**, R. Berkvens, and M. Weyn, "Comparing Machine Learning Algorithms for RSS-Based Localization in LPWAN", in *Proceedings of the 14th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2019)*. Editors: L. Barolli, P. Hellinckx, J. Natwichai, Springer International Publishing, 2019, pp. 726–735, https://doi.org/10.1007/978-3-030-33509-0_68.
4. **T. Janssen**, N. BniLam, M. Aernouts, R. Berkvens, and M. Weyn, "LoRa 2.4 GHz Communication Link and Range", *Sensors*, vol.20, no. 16, p. 4366, Aug. 2020, https://doi.org/10.3390/s20164366.
5. **T. Janssen**, R. Berkvens, and M. Weyn, "Benchmarking RSS-based localization algorithms with LoRaWAN", *Internet of Things*, vol. 11, p. 100235, Sep. 2020, https://doi.org/10.1016/j.iot.2020.100235.

Contribution 2:

1. **T. Janssen**, M. Weyn, and R. Berkvens, "A Primer on Real-world RSS-based Outdoor NB-IoT Localization", in *2020 International Conference on Localization and GNSS (ICL-GNSS)*, IEEE, Jun. 2020, pp. 1–6, https://doi.org/10.1109/ICL-GNSS49876.2020.9115578.
2. **T. Janssen**, R. Berkvens, and M. Weyn, "RSS-Based Localization and Mobility Evaluation Using a Single NB-IoT Cell", *Sensors*, vol. 20, no. 21, p. 6172, Oct. 2020, https://doi.org/10.3390/s20216172.

Contributions 3 and 4:

1. **T. Janssen**, A. Koppert, R. Berkvens, and M. Weyn, "A Survey on IoT Positioning leveraging LPWAN, GNSS and LEO-PNT *(Submitted on July 30, 2021)*", *IEEE Internet of Things Journal*, pp 1–25, 2023.

1.3 Outline

Figure 1.2 shows the organization of the chapters in this thesis. Each chapter is based on the publications listed in Section 1.2.

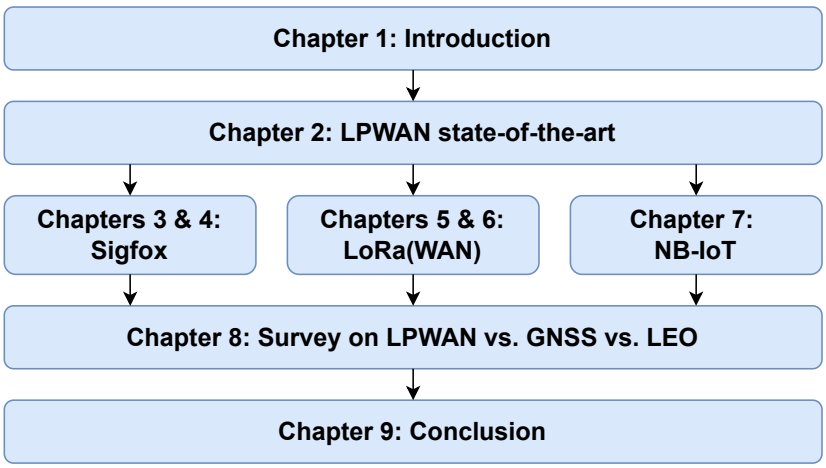


Figure 1.2: Thesis outline

This thesis continues with Chapter 2, which describes the state-of-the-art (SOTA) of LPWAN. I explain LPWAN technologies, concepts and methodologies to calculate the position of a mobile device. The purpose of this chapter is to provide the building blocks of LPWAN upon which I build in the next chapters. Please note that the SOTA of satellite-based technologies and techniques is covered in Chapter 8.

Chapters 3 and 4 describe RSS-based localization experiments with the Sigfox technology. While the experiments in Chapter 3 use Wi-Fi fingerprints transmitted over Sigfox, Chapter 4 focuses on localization using Sigfox only.

Chapters 5 and 6 focus on localization using LoRaWAN and LoRa, respectively. In Chapter 5, I apply ML to improve the performance of RSS-based fingerprinting in a public LoRaWAN network. Chapter 6 explores the feasibility of localization using LoRa modulation in the 2.4 GHz band.

In Chapter 7, I present the first RSS-based positioning results using the NB-IoT technology. I describe the measurement campaign in Antwerp, as well as the proximity, ranging and fingerprinting algorithms. After the evaluation of these algorithms, the chapter ends with a discussion on how to improve the mobility of NB-IoT.

Chapter 8 provides an in-depth survey conducted in the framework of the projects carried out for ESA. The purpose of the survey is to compare LPWAN localization to positioning using GNSS and LEO satellites. In a state-of-the-art analysis, I compare terrestrial and innovative space-based technologies and techniques in terms of positioning performance.

Finally, Chapter 9 summarizes the conclusions of this thesis and indicates directions for future research.

Chapter 2

State-of-the-art of LPWAN

WHEN you want to keep track of something valuable, the required technologies and techniques often depend on the context and the application requirements. For example, a postal service can locate parcels inside a warehouse using Wi-Fi technology. From the moment a parcel leaves the warehouse, however, the delivery van needs to be tracked via different technologies, as the range of Wi-Fi is limited. A track-and-trace device in the van often uses a GNSS receiver to determine the position of the vehicle, while a cellular network technology communicates this position to a remote end user. In many IoT applications, energy efficiency, low cost and low complexity are given higher priorities than meter-level localization accuracy [20]. Moreover, privacy sensitive applications encourage not knowing the exact location of e.g., the postal worker. Thus, we can use LPWANs to estimate the rough location of a mobile device in a large-scale outdoor environment with minimal energy consumption [33, 34].

The purpose of this chapter is twofold. First, I explain the basic concepts of LPWAN, which are required in order to follow the reasoning in the next chapters. By doing so, I hope the reader becomes acquainted with the building blocks of my work. Second, I provide insights on what was already investigated before I started my research. Therefore, I discuss several related works that are relevant in the fields of IoT and LPWAN localization.

This state-of-the-art chapter is structured in the following way. The basic concepts and terminology of LPWAN are described in Section 2.1. Subsequently, Section 2.2 provides an overview of LPWAN technologies from a historical perspective. I explain what LPWAN technologies are, how to classify them and how they have evolved over time. Finally, Section 2.3 describes localization techniques that are and can be used for LPWAN localization. I discuss how LPWAN signal characteristics can be exploited to determine the position of a mobile IoT device and elaborate on the available algorithms in the literature.

2.1 LPWAN Communication & Architecture

Communication in a typical LPWAN scenario takes place in the following way. A general LPWAN architecture is visualized in Figure 2.1. An IoT device, also referred to as UE, sporadically transmits a small uplink message to one or more nearby terrestrial gateways. The gateways, sometimes referred to as base stations, subsequently forward the message to a server of the network provider they belong to. From there, the payload of the message can be obtained via the Internet to end user devices and applications. This uplink communication using a star topology is mostly used in LPWAN for transmission of sensor data. However, downlink communication in the reverse direction is also possible. This is convenient to remotely activate or configure devices in the field.

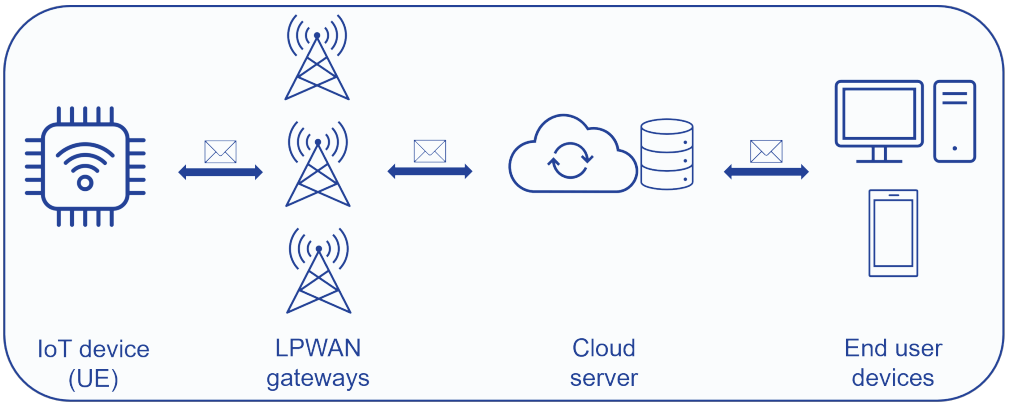


Figure 2.1: General LPWAN architecture

Thus, devices in an IoT scenario have a connection to the Internet through a network of gateways. In theory, this connection can be a continuous data connection. This is the case for devices that are large enough to carry additional batteries or for static devices with a power supply. In practice, many IoT devices are mobile and equipped with small batteries. Therefore, LPWANs propose to reduce the energy requirement of the mobile data connection, albeit at reduced data rates [33].

2.2 A Brief History of LPWAN Technologies

The landscape of LPWAN technologies can be categorized based on the frequency bands at which they operate. Technology operators can acquire a license to exclusively operate at dedicated frequencies. LPWAN standards leveraging licensed frequencies include the cellular NB-IoT and LTE-M. In contrast, LPWAN technologies operating in unlicensed frequency bands need to share the spectrum with other technologies. Most prominent examples of unlicensed LPWAN include Sigfox and LoRaWAN, operating in the 868 MHz band in the EU and the 915 MHz band in the US. Obviously, many other LPWAN technologies exist, such as Weightless, DASH7, Telensa, NB-Fi and the Institute of Electrical and Electronics Engineers (IEEE) 802.15.4 group of standards. The low energy consumption of LPWAN

technologies is extensively investigated in the state-of-the-art [35] and therefore falls outside the scope of this work. For an in-depth technology specification of all aforementioned LPWAN technologies, the interested reader is referred to a set of recent surveys [5, 20, 36, 34].

Ever since their introduction, LPWAN technologies have evolved thoroughly. Cellular LPWAN technologies have known a rapidly increasing economic growth. With 47% of all LPWAN-enabled devices, NB-IoT is currently leading the market, followed by LoRaWAN with 36%. Together with Sigfox and LTE-M, they account for 96% of the global LPWAN market [37]. Therefore, I chose NB-IoT, LoRaWAN and Sigfox for the localization experiments in this work. LTE-M was omitted simply because there was no such network fully available in Belgium at that time. Thus, as the focus in this work is on Sigfox, LoRaWAN and NB-IoT, the key characteristics of these technologies are summarized in Table 2.1 and their evolution is described in the following subsections. Note the fact that the developed algorithms and methodologies are generic and can easily be applied to other LPWAN technologies.

Table 2.1: Key characteristics of LPWAN technologies discussed in this work.

Characteristic	**Sigfox**	**LoRaWAN**	**NB-IoT**
Licensed frequencies	No	No	Yes
Frequency band	868 MHz (EU) 915 MHz (US)	868 MHz (EU) 915 MHz (US)	LTE bands (450-2200 MHz)
Bandwidth	100 Hz	125 - 500 kHz	180 kHz
Max. data rate	100 bps (UL) 600 bps (DL)	0.3 - 37.5 kbps	150 kbps (UL) 127 kbps (DL)
Communication range	3-10 km (urban) 30-50 km (rural)	5 km (urban) 20 km (rural)	1 km (urban) 10 km (rural)
Modulation	DBPSK (UL) GFSK (DL)	LoRa CSS	QPSK
Max. payload size	12 bytes (UL) 8 bytes (DL)	250 bytes	1600 bytes
Max. update rate	Limited (duty cycle restrictions)	Limited (duty cycle restrictions)	Unlimited (until intolerated)

2.2.1 Sigfox

Founded in 2010, the French company Sigfox was one of the pioneers in the LPWAN market. The Sigfox technology employs an Ultra-Narrow Band (UNB) channel of 100 Hz with a Differential Binary Phase Shift Keying (BPSK) modulation operating on the 868 MHz telecommunication frequency in the EU. This frequency does not require a license for operating Short Range Devices (SRD), defined as devices with a limited transmission power. To comply with the EU regulations, a Sigfox device can only transmit up to 140 messages of

12 B each on a single day at 100 bit/s. This stems from a 1% hourly duty cycle regulation, limiting Sigfox to 36 s of signal transmission or six messages per hour. The transmission of a single Sigfox message takes 6 s, during which the same information is transmitted three times for error prevention. The communication range varies from 30–50 km in rural areas to 3–10 km inside cities [33, 38].

In 2017, a nationwide Sigfox network was deployed in Belgium [39]. In the first part of my research, I simulated the transmission of Wi-Fi fingerprints over this network for localization purposes. Around the same time, a colleague in my research group employed this network to collect a large database of Sigfox messages in Antwerp, Belgium [32]. Using these two databases, I improve the localization with Sigfox as discussed in Chapters 3 and 4, respectively.

In January 2022, the French company behind the Sigfox technology filed for bankruptcy protection due to to significant financial challenges [40]. After a few months however, the company with all its assets and contracts with customers and operators was acquired by UnaBiz, a Singapore-based massive IoT service provider [41]. From July 2022 onwards, UnaBiz continued the operations, keeping Sigfox as the name of the technology [42]. Please note that these events do not affect the validity of the results of my localization experiments using Sigfox.

2.2.2 LoRaWAN

With the aim of providing a low-power, long-range wireless communication protocol, the French start-up Cycleo developed the LoRa (Long Range) technology in 2009. Three years later, the California-based company Semtech acquired the startup and further improved the technology. Semtech now owns a patent and therefore sells licenses to LoRa chipset manufacturers. At the same time, the LoRa Alliance was founded in 2015 and maintains the open-source LoRaWAN standardization [43].

Sub-GHz LoRa-based communication has been studied thoroughly in the past decade. The authors in [44] were the first to provide an in-depth analysis of the functional components of both the physical (LoRa) and data link (LoRaWAN) layer. A more recent characterization of the popular LPWAN protocol operating in unlicensed bands is compiled in [45], along with related surveys. Derived from Chirp Spread Spectrum (CSS), the proprietary LoRa modulation achieves high resilience to interference and fading effects. When compared to Sigfox, LoRaWAN in general provides more flexibility through the use of a variable bandwidth, coding rate and Spreading Factor (SF). By configuring a higher SF, the airtime and thus the communication range increases at the expense of a reduced data rate, payload size and battery lifetime.

LoRaWAN is widely used to communicate small messages over large areas. A LoRaWAN network can fully cover a city-scale environment with only a few gateways. Due to the deployment of public and private LoRaWAN networks, both academia and industry performed research about the sub-GHz technology. The mobility of LoRaWAN has been addressed in [46], while enhancements of the Adaptive Data Rate (ADR) mechanism are suggested in [47]. The forwarding of an uplink message by multiple gateways ensures that a handover

can take place without any loss of data. In another study, a robust frame detection algorithm was proposed in order to detect LoRa-modulated frames with minimal complexity [48]. Furthermore, a frame relay strategy was found to be a feasible way to improve the link quality of poorly connected nodes and successfully extend the range of LoRaWAN [49].

More than a decade ago, LoRa was invented to provide a LPWAN protocol operating at sub-GHz frequencies. Because of local spectrum regulations, LoRa hardware modules need to be adapted to operate in different frequency bands. However, a few years ago, Semtech started developing a range of LoRa transceivers operating within the 2.4 GHz band [50]. The motivation to operate in this band was mainly because it is globally available. On the one hand, this tackles the problem of having to develop multiple chipsets operating in different frequency bands, thus paving the way for the development of a universal chipset which can operate anywhere in the world. This is especially valuable for track-and-trace applications, where goods cross different zones worldwide. Moreover, LoRa devices operating at 2.4 GHz can transmit at higher data rates because of the higher available bandwidth [51]. Consequently, this technology can offer a balance for applications that require a higher data rate than sub-GHz LPWANs and a longer communication range than classic 2.4 GHz technologies such as Wi-Fi and Bluetooth. Additionally, a higher bandwidth also allows for more accurate time-based localization [52]. On the other hand, a few concerns are raised with the expansion of LoRa to the widely used unlicensed 2.4 GHz Industrial, Scientific, and Medical (ISM) band. The interference properties of Wi-Fi devices operating in this band were characterized in [53]. In another study, Polak et al. address coexistence issues with Wi-Fi and LoRa 2.4 GHz [54]. The theoretical assumptions about the high robustness of LoRa against interference are confirmed. However, it was found that the robustness highly depends on both the configuration and the properties of interfering technologies. Moreover, Polak et al. used different bandwidths in their evaluation and compared them to the bandwidths that are currently available for LoRa at 2.4 GHz.

The maximum communication range in the 2.4 GHz band strongly depends on the technology and its respective transmission power. For the Bluetooth Low Energy (BLE) 5 standard, the maximum communication ranges in different scenarios are summarized in [55]. Typical ranges are 50 m in an indoor scenario, 165 m in an outdoor Non-Line of Sight (NLoS) scenario and 780 m in a Line of Sight (LoS) scenario. Semtech aims to outnumber these ranges with their LoRa 2.4 GHz chipset, featuring a Time of Flight (ToF) ranging possibility [56]. A proof-of-concept implementation for coherent multi-channel ranging with this LoRa radio chip is provided in [57]. However, the potential benefits of LoRa in the 2.4 GHz band have not yet been investigated thoroughly, which constitutes the main objective of Chapter 6.

2.2.3 NB-IoT

In 2016, the 3rd Generation Partnership Project (3GPP) introduced NB-IoT in Release 13. In contrast to other LPWAN technologies such as Sigfox and LoRaWAN, NB-IoT operates in the licensed spectrum. Therefore, specific frequency bands are reserved, which reduces interference effects. Moreover, there are less duty cycle limitations and regulations on the amount of data to transmit. The technology is popular amongst cellular providers, as they can easily deploy an NB-IoT network on top of existing Long Term Evolution (LTE) infrastructure in a rapid and scalable fashion.

Every new 3GPP release has brought improvements to the NB-IoT technology. A historic overview of releases is provided by Rastogi et al. [36]. Release 14, for instance, introduced increased data rates and transmission block sizes. Other additions include the ability to multicast, a wake-up signal to further reduce the energy consumption, and even positioning features. The specifications of both physical and data link layers of NB-IoT are detailed in [36, 58] and summarized in Table 2.1.

There are a few important architectural differences between NB-IoT and LoRaWAN [59]. After a negotiation phase between the network and the end device, an NB-IoT end device connects to a single cell of a nearby base station, i.e., the serving cell. In contrast, a LoRaWAN-enabled end device transmits an uplink message to all nearby gateways. Another noteworthy difference is that usually the density of NB-IoT base stations is higher compared to the density of LoRaWAN gateways, as NB-IoT leverages ubiquitous LTE infrastructure.

During the real-world NB-IoT localization experiments in Chapter 7, I experienced mobility limitations in the NB-IoT protocol. Because of the aforementioned architectural decision, the UE only reported the currently serving cell, instead of all neighboring cells. According to Radnosrati et al., the number of reported cells is reduced to a minimum in order to minimize the signaling cost and support a large number of devices [60, 61]. Obviously, this behavior is undesirable in localization use cases, since more participating Evolved Node Bs (eNBs) result in a higher location accuracy [62, 63]. Many simulation studies use the LTE module in Network Simulator 3 (NS3) to simulate an NB-IoT network [64]. However, by doing so, they do not face the aforementioned issue. Therefore, I chose to address this issue and evaluate the performance of NB-IoT localization using real-world measurement data. Fortunately, from 3GPP Release 16 onward, a UE may optionally report about strongest neighbors [36].

Another mobility issue in the NB-IoT protocol is the lack of handover support. Moon et al. explain the concepts of Extended Discontinuous Reception (eDRX) and cell reselection of NB-IoT [65]. They acknowledge that cell reselection can only take place while the UE is in idle mode. Moreover, the cell selection process described in [61] suggests that only after failing to connect to the currently serving cell, the cell reselection process is initiated. Additionally, this issue introduces latency which can reach up to 10 s [46]. Furthermore, it is proven that when reselecting a cell more frequently, despite the increased energy consumption, the mobility performance increases significantly [65].

In July 2019, 3GPP Release 16 announced that NB-IoT will be adopted under the umbrella of Fifth Generation (5G) mobile network technologies and will therefore evolve as part of 5G specifications [9].

Finally, a lot of academic research has been devoted to extending the coverage of NB-IoT using satellites in LEO [66, 67, 68, 69, 70]. In conjunction, 3GPP Release 17 adds standardization for Non-Terrestrial Networks (NTNs) to address massive IoT use cases in remote areas via satellites in LEO, MEO and GEO [71]. The survey in Chapter 8 elaborates on the interoperability between terrestrial and satellite networks for positioning purposes.

2.3 LPWAN Localization Techniques

As an alternative to GNSS-based solutions, LPWANs are being used to locate a transmitting device in large-scale environments with minimal energy consumption [33, 34]. In general, localization based on wireless communication can be split into three categories [21, 23], as shown in Figure 2.2. Each category employs a different characteristic of the signal arriving at the nearby gateways. The first category constitutes algorithms that calculate the Time of Arrival (ToA) or Time Difference of Arrival (TDoA) of the signal at the gateways to determine the position of the UE. A second category covers positioning algorithms based on the Angle of Arrival (AoA) of the received signal through an antenna array. In the third category, the Received Signal Strength (RSS) is exploited. This category further distinguishes between RSS-based proximity, ranging and fingerprinting approaches. Oftentimes, these techniques are combined in a single positioning system. A multimodal localization framework enables to switch to an optimal localization method, as well as location-based seamless handover mechanisms [72, 73]. In the following subsections, I analyze the state-of-the-art of these approaches through related works.

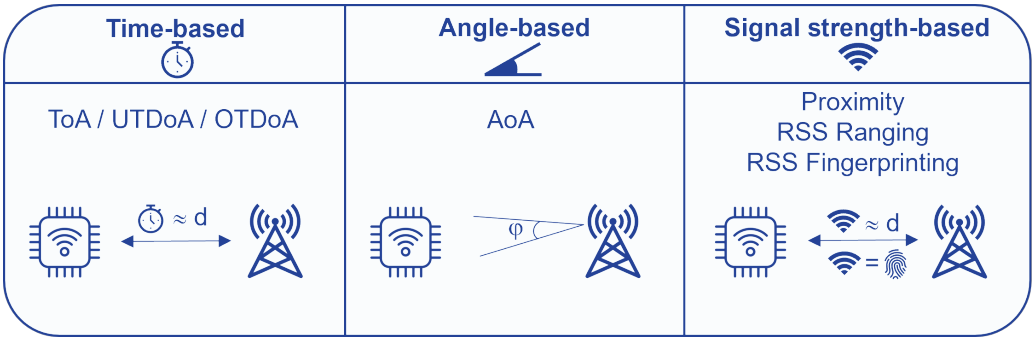


Figure 2.2: Techniques to locate an LPWAN-enabled device, categorized by signal characteristic.

2.3.1 Time-based Algorithms

The principle behind timing-based localization is quite straightforward. In a ToA algorithm, the distance between a mobile device and the gateway is calculated by multiplying the signal propagation time with the speed of light. When three or more gateways receive the signal, the position of the mobile transmitter can be determined through a multilateration algorithm [74]. However, ToA-based solutions require precise synchronization between gateways and the mobile node, increasing device complexity and power consumption. Therefore, the TDoA is often calculated instead [75]. Nonetheless, TDoA-based solutions still require precise synchronization between the gateways themselves, at the expense of the operator.

Timing-based localization solutions have been studied extensively. A comparison of state-of-the-art TDoA algorithms is provided in [76]. Such algorithms are widely applied in LPWAN technologies [77]. The authors in [78] demonstrated the use of TDoA-based positioning with LoRaWAN, leading to a median accuracy of 200 m. This accuracy was significantly improved to 75 m by tracking the movement speed and taking road infrastructure into account.

Similarly, TDoA-based positioning experiments in a public LoRa network lead to a median accuracy in the order of 150 m [79]. Furthermore, a new frequency hopping pattern for the preamble of physical random access channel (NPRACH) in NB-IoT has been designed to further increase the ToA estimation accuracy [80, 81]. In contrast to LoRa and NB-IoT, timing-based localization using Sigfox is considered not feasible due to the extreme low time resolution of the UNB technology [82]).

With 3GPP Release 14 of the NB-IoT standard, two advanced positioning techniques are introduced: Enhanced Cell ID (E-CID) and Observed Time Difference of Arrival (OTDoA) [83, 84]. In E-CID, the position of the UE is estimated by performing measurements on radio signals [85]. These measurements estimate either the Reference Signal Received Power (RSRP), TDoA or the AoA. This enhancement of traditional cell identification enables more accurate localization. The accuracy however still depends on the cell size and can vary significantly.

In contrast to the aforementioned uplink TDoA approaches, OTDoA is a downlink-based localization approach. OTDoA combines Reference Signal Time Difference (RSTD) measurements between neighboring eNB cells and a reference cell to estimate the location of a mobile UE [86]. However, the implementation of OTDoA requires precise synchronization between the eNBs and the presence of a location server, leading to an increased network complexity and eNB cost for operators. Therefore, the OTDoA feature has not yet been implemented in many countries, such as Belgium. Consequently, most performance evaluations of OTDoA-based positioning are carried out in the form of simulation studies [62, 86, 87, 88, 89, 90]. Recently, the first implementation of an OTDoA-capable NB-IoT chip was presented [91]. Measurements carried out in a laboratory setup result in a Root Mean Square (RMS) error of around 50 m. Furthermore, Tong et al. developed a general analytical model to study the performance of OTDoA [62], where the module needs to keep track of the time of arrivals. As an alternative to OTDoA, a 1-bit passive radar was used to reduce the complexity and energy consumption of the localization [92].

Besides the limited number of OTDoA-enabled NB-IoT networks, an OTDoA-capable UE requires the presence of highly accurate clocks, which is often not the case in low-cost IoT devices. This again highlights the need for more energy-efficient, low-complexity and low-cost localization methods in NB-IoT.

2.3.2 Angle-based Algorithms

Instead of estimating the distance between a mobile device and a gateway, one can also estimate the AoA or the Angle of Departure (AoD) between them. In AoA-based localization, the first step is to measure the phase of the arriving signal at multiple points in space. Therefore, the receiver (i.e., a nearby gateway) requires an antenna array system. The second step is to calculate the AoA based on the phase measurements. When the AoA at two or more receivers is determined, one can estimate the position of the mobile transmitter through a triangulation algorithm [75]. To keep the complexity of an LPWAN end device to a minimum, AoA approaches are more attractive compared to AoD approaches.

A state-of-the-art AoA estimation algorithm for use in sub-GHz bands is presented in detail

in [93, 94]. By synchronizing antennas, the AoA of a narrow-band signal was estimated with an absolute mean error well below 2 degrees. These experimental results along with the development of a low cost AoA unit [95] prove the feasibility of AoA estimation techniques for localization of IoT devices in outdoor environments.

Oftentimes, AoA estimation is combined with other techniques. For example, Aernouts et al. devised a probabilistic localization model that combines TDoA and AoA estimations in a sub-GHz LoRa network [52, 96]. Simulations show that adding AoA reduces the median localization error with 57%. These results are validated through the use of real LoRaWAN data, achieving a mean location estimation error of 339 m and 153 m in a LoS and NLoS scenario, respectively. Finally, AoA techniques are also combined with RSS-based techniques [97], which are discussed next.

2.3.3 Received Signal Strength-based Algorithms

In contrast to TDoA-based and AoA-based localization approaches, gateways or antennas do not need to be synchronized in an RSS-based approach. In such an approach, nearby gateways measure the signal strength of the received signal (expressed in dBm) to determine the position of the mobile transmitter. As localization using TDoA [5] and AoA [94] is extensively investigated in literature, this work focuses on RSS-based localization techniques.

Because of the narrowband nature of LPWAN technologies, localization based on LPWAN communication signals is very challenging. Gezici et al. show that the Cramer-Rao Lower Bound (CRLB) for estimating the range between mobile device and gateway through the RSS is proportional to the distance between them [98]. Given the wide area networks, one can generally assume that the CRLB will be rather high. Moreover, RSS-based approaches need to deal with signal fluctuations, interference and multipath effects [99].

RSS-based algorithms can be further classified into three main classes. Proximity-based, range-based and fingerprint-based localization are the most common techniques found in recent literature [100, 101]. While this section provides the state-of-the-art of RSS-based algorithms and how they perform in LPWANs, the actual workings of the algorithms are illustrated and discussed in more depth in Chapters 4, 5 and 7. Hence, the results of related works serve as a baseline for my RSS-based localization experiments.

2.3.3.1 Proximity

Proximity localization assumes that a mobile device is at or near the location of the gateway with the strongest link. The accuracy of the proximity method depends on the range of the communication method. Therefore, proximity localization is often used in Radio Frequency Identification (RFID) scenarios, which have a limited range [102]. If the range increases, the mobile device can be in communication with multiple gateways. This enables more sophisticated location estimation algorithms, such as taking the centroid of the gateway locations and clustering techniques [103].

While an RSS proximity algorithm with map matching in a public LoRaWAN network lead to

a median estimation error of 1000 m, a TDoA-based approach in the same environment lead to a median error of 150 m [79]. Despite its low accuracy, a simple proximity algorithm can provide a rough location estimate in areas with limited coverage, i.e., with a single nearby gateway.

2.3.3.2 RSS Ranging

In RSS ranging localization, the signal strength between a mobile device and a gateway is translated into a distance between them using a radio propagation model, also referred to as path loss model [104]. Similar to timing-based techniques, the range estimates to multiple gateways are subsequently used as the input for a multilateration algorithm. From a geometry perspective, this boils down to calculating the intersection of circles in 2D or spheres in 3D spaces [75]. However, due to inaccuracies in the range estimates, these do not always overlap and more sophisticated algorithms are required. The output coordinate of the multilateration algorithm is the final location estimate of the mobile device.

The way a signal propagates through the air strongly depends on the environment. Therefore, numerous path loss models have been developed for use in free space, indoor, urban and rural areas. For instance, the International Telecommunication Union (ITU) has developed an indoor path loss model for 2.4 GHz technologies [105], a model for free space attenuation and multiple outdoor propagation models [106]. Various models for mobile communication are introduced in [99] and compared by Singh in [107]. Examples include the empirical models of Okumura, Hata and COST-231, which are most widely used in urban areas, but are adapted for use in rural areas as well. The COST-231 Walfisch-Ikegami model is further discussed in [108]. Propagation models have been studied extensively for many technologies, such as BLE [55], IEEE 802.11ah [109] and 4G LTE [110]. More recently, these models have been modified for use in LPWANs [111]. More specifically, the model parameters are tuned in order to match urban [112], suburban and rural scenarios [113]. One goal of these models is to support operators with network planning, which involves coverage analysis [114] and estimation of link quality [115]. For example, given the distance between gateway and receiver, the link budget is analyzed. Opposed to this, the inverse operation can also be used for localization purposes. As I evaluate the feasibility of these models for use in LPWAN localization, more details can be found in Chapters 5, 6 and 7.

Recent works investigated RSS-based ranging localization in LPWAN. A preliminary performance evaluation of a Min-Max ranging algorithm [116] in a public Sigfox network was carried out in [103]. The empirical Hata model resulted in a mean location estimation error of 721 m. Robles et al. improved the Min-Max algorithm by adding weights [117]. Another popular ranging algorithm is ordinary Least Squares estimation. Telles et al. demonstrated that a simple weighted centroid localization algorithm can outperform TDoA techniques [118].

In my research, I analyze these ranging algorithms to improve the overall localization performance of LoRaWAN. However, LoRaWAN ranging [119] can be quite challenging, since the frequency hopping in LoRaWAN significantly impacts the localization accuracy [120]. Consequently, Wolf et. al designed a multi-channel ranging system for LPWAN [100]. Finally, to the best of my knowledge, I am the first to apply RSS ranging in a public NB-IoT network.

2.3.3.3 RSS Fingerprinting

Pattern recognition, most commonly known in the localization community as fingerprinting, consists of comparing a snapshot of the current environment with a database of such snapshots linked to known locations [75]. In contrast to the aforementioned timing-, angle- and other RSS-based approaches, the (sometimes sensitive) gateway locations do not need to be known in a fingerprint-based approach.

More specifically, an RSS-based fingerprint localization approach consists of an offline and online phase. During the offline phase, a measurement campaign is conducted in a predetermined area. Each fingerprint holds RSS measurements of receiving access points or gateways. This fingerprint is stored along with the ground-truth location in a training database. During the online phase, a mobile device is localized by matching its fingerprint measurements against the earlier collected measurements in the training database. This matching procedure can be implemented through probabilistic methods, Machine Learning algorithms or by applying a pattern matching technique such as k-Nearest-Neighbors (kNN) analysis.

Fingerprinting is mostly used in constricted indoor environments with wireless technologies such as Wi-Fi and BLE [121, 122, 123]. Obviously, indoor approaches tend to obtain better accuracies than outdoors because of denser network deployment. One thing to consider in Wi-Fi fingerprinting is how to compare the fingerprint created by the mobile device with the fingerprints in the database. Torres-Sospedra et al. provide an overview of 51 distance metrics that could be used for this comparison [124]. Another consideration in Wi-Fi fingerprinting is the challenge of constructing the fingerprint database. While the traditional approach is to collect the fingerprints during a labor intensive site survey, newer studies have looked into crowdsourcing the database construction [125]. Following the latter approach, Sapiezynski et al. achieved a 15 m median location error for 73% of their Wi-Fi scans in an outdoor environment, similar to the accuracy of Google's crowdsourced geolocation service [126]. Another new trend is to apply Deep Learning algorithms for fingerprint-based indoor localization [127].

To date, there already exist a number of studies on fingerprint-based localization in LPWAN. For example, Song et al. use Channel State Information (CSI) of NB-IoT for indoor pattern matching localization [128]. The majority of studies, however, explore fingerprinting in an outdoor scenario. Sallouha et al. evaluated a fingerprint classification method based on Support Vector Machines (SVMs) [82]. In their implementation, RSS measurements are used to estimate the distance between a transmitting Sigfox device and the responding gateways to classify in which area the transmitter is located. Within such a class, the location estimation is improved by distance estimation between end-devices and ground-truth GPS devices. They achieve a classification accuracy of 78% in fingerprints with a radius of 150 m, separated by 2.5 times this distance from other fingerprints. However, the experiments were conducted in a small area which contains only two classes. To overcome this pitfall, Aernouts et al. collected Sigfox and LoRaWAN datasets with a spatial spread that covers the entire city center (53 km^2) of Antwerp, Belgium [32]. The open-source datasets are subsequently used by other researchers to evaluate and compare the performance of their algorithms [129, 130, 131, 132, 133, 134]. For instance, Sallouha et al. again tested their fingerprint classification algorithm with seven classses spread over the city of Antwerp. They were able to classify the location of a mobile device with an accuracy of 80% [135]. Lastly,

besides the RSS values in the dataset, Etiabi et al. also include the SF in the fingerprint, achieving increased localization accuracy [136].

In order to improve the accuracy of outdoor RSS fingerprinting, the distance metric and RSS representation in a *k*NN algorithm should be optimized, as suggested in [124]. I carried out this optimization process for Sigfox and LoRaWAN in Chapters 4 and 5, respectively. The performance was further improved by Anagnostopoulos et al. in [130, 131]. In Chapter 7, I present the results of the first RSS-based experiments with NB-IoT.

Besides estimating the location of an IoT device in LPWAN, fingerprinting approaches can also provide an estimate of the localization error [137]. By providing such a reliability estimate, end users are aware of the possible size of the area in which their device is located. To quantify this level of uncertainty, a dynamic accuracy estimation method is presented in [138]. In this method, a first model is trained to estimate multiple device locations. Next, a second model is trained but with corresponding location estimation errors (i.e., the difference between estimated and ground-truth location) as target values. Hence, one can predict the accuracy of a new fingerprint in a data-driven way. In [139], the same authors raise the problem that many researchers use their own data and metrics to evaluate their algorithm, which does not allow fair comparison between algorithms. Therefore, the authors use all available indoor and outdoor datasets to apply the most frequently used algorithms and compare their positioning performance using various metrics (mean, median, 95th percentile, standard deviation, etc.). Thus, the study provides a comprehensive overview of what algorithm yields the best results, independent of the data set.

Chapter 3

Localization with Sigfox and Wi-Fi Fingerprints

SUPPLY chain management requires regular location updates of assets, which can be enabled by LPWANs, such as Sigfox. While it is useful to localize a device simply by its communication signals, such an approach may not fulfill the accuracy requirement of the application, given the ultra narrowband nature of Sigfox. On the contrary, installing a GNSS receiver on the device greatly increases its energy consumption.

In this chapter, I present my solution to this localization problem, based on sensing nearby Wi-Fi access points and transmitting over Sigfox. It is important to note that there is no need to connect to the Wi-Fi access points. Due to the 12 byte payload limit of Sigfox, I can transmit only two Basic Service Set Identifiers (BSSIDs), which are six bytes each. This leaves no payload space to transmit any RSS information. While Wi-Fi fingerprinting is well studied, I investigate for the first time the location error when reducing the available information to only two BSSIDs. Usually, anywhere between ten to sixty BSSIDs are discovered, including an RSS value to indicate if it is nearby or far away. Although more information can be transmitted over Sigfox by sending multiple messages, it still has a limit of six messages per hour. Thus, using two messages instead of one increases the update period from ten minutes to twenty minutes. This is why I am interested in using just a single Sigfox message. The aim of the research in this chapter is to establish the expected location error for this approach. Therefore, I use online crowd sourced Wi-Fi access point databases, which are inherently less accurate than the carefully constructed databases in common research. While crowd sourced services have obvious drawbacks, especially being noisy, they create a myriad of opportunities [126].

This chapter continues as follows. Section 3.1 outlines the localization approach, discusses the online Wi-Fi databases, and presents the measurement locations. Section 3.2 shows the results in terms of location accuracy and the frequency of missing an access point. Section 3.3 analyzes and discusses those results. Finally, Section 3.4 draws the conclusion.

This chapter is based on:

Thomas Janssen, Rafael Berkvens, & Maarten Weyn. *Localization in Low Power Wide Area Networks Using Wi-Fi Fingerprints*. Applied Sciences, vol. 7, no. 9, p. 936, Sep. 2017, `https://doi.org/10.3390/app7090936`.

3.1 Methods and Materials

This section first describes the localization principle and the approach for the experiments. Subsequently, I discuss the characteristics of the available online Wi-Fi access point databases. The section ends with an overview of the measurement locations in Antwerp.

3.1.1 Localization Approach

The general approach of Wi-Fi fingerprinting localization can be divided into two steps: the training and operational step. Firstly, the training step involves matching information about the discovered networks or Wireless Access Points (WAPs) (i.e., their BSSID and RSS) to a measurement location in a database. Secondly, the operational step involves requesting from the database the location of newly discovered WAPs during operation. The database will reply with the location of each WAP. I will not construct the database myself, as there already exist many databases that are available online.

To discover nearby access points, a laptop or a smartphone with a Wi-Fi adapter is required. A Python script saves the BSSID and RSS of each discovered access point. In Windows, this can easily be done by running the following command: *netsh wlan show networks mode=bssid*. The Python script then saves the output to a file, together with the current location. The current location is determined by a GPS module. To be more accurate, all locations are chosen to be recognized on a map, so that a visual check of the GPS location can be performed. This procedure is executed on 36 locations in and around Antwerp, Belgium.

When the Wi-Fi data is collected and transmitted over Sigfox to a backend server, this server will perform an HTTP-request to a database for every BSSID. Afterwards, the database will respond with a message, which includes the coordinate of the access point. If the BSSID is not found in the database, an error message will be sent. Subsequently, all possible combinations of two BSSIDs are generated and the mean coordinate of every pair is calculated. If only one of the two BSSIDs was found in the database, the coordinate of that BSSID will be considered as the mean coordinate. Finally, the distance error can be calculated. This is the distance between the coordinate of the position where the measurement took place and the mean coordinate of the pair of BSSIDs. This distance is calculated using the Haversine function, which takes into account the curvature of the Earth [140].

Additionally, I calculate the probability of having no match of the BSSID in the database:

$$P(\neg nB) = \frac{\binom{\neg B}{n}}{\binom{W}{n}}, \tag{3.1}$$

where $P(\cdot)$ is the probability operator; W is the number of access point seen at a location; B are the BSSIDs in W that are found in the database, $\neg B$ are the BSSIDs that are not found in the database; n is the number of BSSIDs that can be used for localization—in this case, this is maximum 2 because of the limitations of Sigfox communication; $\neg nB$ are the

combinations of n BSSIDs that are not found in the database. This probability is interesting, because if no BSSIDs of the Wi-Fi access points that are seen at a location can be found in the database, the location of the device cannot be estimated.

In an attempt to improve the accuracy of the results, I develop an additional strategy: only sending the combination of BSSIDs that have the strongest signal strength. It is not possible to send the RSS values together with the BSSIDs in one Sigfox message, but I can use it in the sending device to determine the right combination of BSSIDs. Although this strategy is likely to consume more energy from the device (to pick the right combination of BSSIDs), it may decrease the mean location errors.

Another improvement is attempted by combining databases. In my experiments, I use the WiGLE and LocationAPI databases (see Section 3.1.2). The most straightforward combination is to take the average of the locations returned by each database. If a database does not return a location for one or both access points, these locations are simply disregarded in the calculation of the average. Since LocationAPI returns a single location for two BSSIDs, I account for that location twice, so that it receives an equal weight compared to the two results from the WiGLE database. In fact, it is impossible to know if LocationAPI knows about only one of the two access points; in this case the weight of that location is erroneously doubled, which cannot be prevented.

This straightforward combination neglects the fact that the location error from the LocationAPI database is generally much better than that of the WiGLE database. Thus, another approach for combination is to choose the location provided by the LocationAPI database, unless the BSSIDs could not be found in that database; in which case the WiGLE database is consulted for a location estimate. While I will show the location error results of the first approach, this second combination approach will be further analyzed.

To summarize, I apply the following localization strategy and improvements when localizing a device with Wi-Fi access points using Sigfox as communication channel:

1. Choose any combination of two BSSIDs for localization;
2. Choose the two BSSIDs with the highest RSS for localization;
3. Combine the two databases through averaging; and
4. Combine the two databases by preferring the LocationAPI database, using WiGLE only when the BSSIDs cannot be found in LocationAPI.

A single Sigfox message allows an update period of ten minutes. Using two messages would increase the period to twenty minutes, which is why I keep the focus on a single message of 12 B. Localization with this constraint has not been researched before.

3.1.2 Wi-Fi Databases

There is a variety of fingerprinting databases to work with. Some companies send their so called wardrivers all over the world to gather information about Wi-Fi access points.

However, there is a trend to create more crowd-sourced databases [126]. People can create a free account and upload data of the access points in their neighborhood. The most popular and largest databases are summarized in Table 3.1.

Table 3.1: Summary of largest Wi-Fi fingerprinting databases.

Database Name	**Number of Wi-Fi Access Points**	**Reference**
Combain Positioning Service	1,296,000,000	[141]
LocationAPI by UnwiredLabs	1,500,000,000	[142]
Mozilla Location Service	30,000,000,000	[143]
Navizon	*unknown*	[144]
WiGLE	4,750,076,107	[145]

Databases with information about Wi-Fi access points are growing at a significant rate. This can be stated by observing that, for example, the Navizon database is growing at a rate of 500,000 data points every day [144]. It is interesting to note that this database also keeps track of the number of cell phone towers. One of the many reasons of this significant rate is the IoT and the world of ubiquitous computing we are living in today. Navizon even pays people [146] to contribute to the database. The one million registered users can just download the app, drive around in their car and earn a pay-out of 15$ for every 10,000 access points.

One database used for this research is called WiGLE [145]. WiGLE is short for Wireless Geographic Logging Engine. The website is an open source platform; everyone can view registered BSSIDs around the world on a map. If a free account is created, requests can be made and data can be uploaded to the database. At this moment, WiGLE wardrivers have discovered over 4.5 billion Wi-Fi access points around the world. This database is chosen because it is easy to work with and has a decent number of access points.

Another database used for this research is LocationAPI [142]. LocationAPI is a website hosted by UnwiredLabs, an Indian corporation focusing on positioning systems. At this moment, LocationAPI wardrivers have discovered over 1.5 billion Wi-Fi access points worldwide. Its database contains a large number of wireless access points in Belgium.

To obtain an objective look at the error or accuracy of Wi-Fi localization through a Sigfox network, the characteristics of both databases are compared. WiGLE has a larger worldwide database. This does not mean that the WiGLE results are more accurate, because there are more access points discovered in the US than in Belgium. WiGLE claims wardrivers have discovered 1.7 million WAPs in Belgium. On the other hand, while LocationAPI contains a smaller number of access points, this database includes 11 million WAPs in Belgium.

A request to the WiGLE database results in a string with a lot of information. It returns the latitude and longitude, but also data like security details and uptime indicators, see Figure 3.1. LocationAPI returns less information, but it is important to note that the data it does return is more useful: each result includes the mean coordinate of multiple BSSIDs, plus the estimated accuracy in meters. A useful extra feature is the 'balance' parameter. This informs the user on how many successful requests are left, see Figure 3.2. However, it is not possible to request the location of a Wi-Fi access point in LocationAPI.

```
"resultCount":1,
"last":1,
"success":true,
"results[{"trilat":51.17709351,
          "trilong":4.41607475,
          "ssid":"UAntwerpen",
          "qos":1,
          "transid":"20110605-00081",
          "firsttime":"2011-05-31T17:11:13.000Z",
          "lasttime":"2013-10-28T11:32:52.000Z",
          "lastupdt":"2013-10-28T09:33:29.000Z",
          "netid":"00:0b:86:26:86:16",
          "name":null,
          "type":"infra",
          "comment":null,
          "wep":"2",
          "channel":6,
          "bcninterval":0,
          "freenet":"?",
          "dhcp":"?",
          "paynet":"?",
          "userfound":false}],
"first":1
```


Figure 3.1: WiGLE returns detailed information about an access point.


```
'status': 'ok',
'balance': 89,
'lat': 51.17786625,
'lon': 4.42454131,
'accuracy': 10
```


Figure 3.2: LocationAPI returns only the average location of multiple BSSIDs, but also an estimated accuracy.

The documentation of both services helps to get started with making requests [145, 142]. However, the documentation of WiGLE is harder to find and understand, since it is still under development. The LocationAPI documentation is very easy to find, easier to read and it contains examples. Table 3.2 presents a summary of all the properties and characteristics of the two databases.

Table 3.2: Comparison of the properties and characteristics of the WiGLE and LocationAPI databases.

Property	WiGLE	LocationAPI
Number of access points worldwide	4,750,076,107	1,500,000,000
Number of access points in Belgium	1,700,000	11,000,000
Accuracy given in request	No	Yes
Balance given in request	No	Yes
Number of BSSIDs in a request	1 BSSID	At least 2 BSSIDs
Number of coordinates in a response	1 coordinate per WAP	1 mean coordinate
Documentation	More complex	More intuitive

3.1.3 Measurement Locations

A measurement is executed at 36 different locations in and around Antwerp. The locations are chosen so that there is a variety of more and less densely populated areas. At every location, a script is run to save all BSSIDs and RSS values of the access points in the vicinity, together with the exact coordinate of the measurement location. Figure 3.3 shows a map of the different measurement locations.

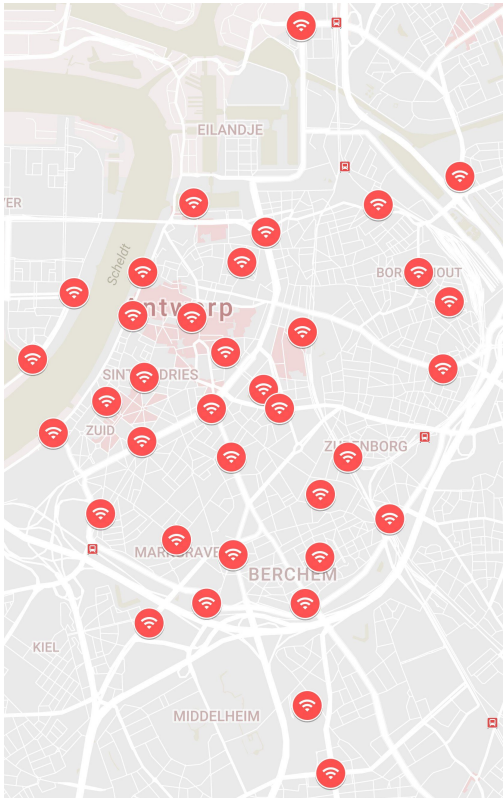


Figure 3.3: Measurement locations in Antwerp, ©2017 Google.

3.2 Results

This section summarizes the location estimate results by database. First the WiGLE database, then the LocationAPI database, and finally the combination of both databases. For each database, I discuss the overall location error, the probability of finding the BSSIDs, and the location error when transmitting only the two BSSIDs with the highest RSS value.

3.2.1 WiGLE Results

The location estimates when localizing using the WiGLE database are prone to large errors, see Figure 3.4. The overall mean location error is 6.24 km, while the overall median location error is 77 m. The 75th percentile location error is larger than 1 km only at locations 12 and 14; at location 14, it is 5.12 km. Ten locations have an error greater than 2 km, which are clipped in Figure 3.4. One of those is location 8. The mean location error is 101 km and the median location error is 45 m. To represent the measurement, a cumulative distribution of the errors at this location is shown in Figure 3.5. There are some combinations of BSSIDs with a significant location error: 802 km and 1633 km. However, more than 90% of the BSSID combinations have an error smaller than 119 m, see Figure 3.5b.

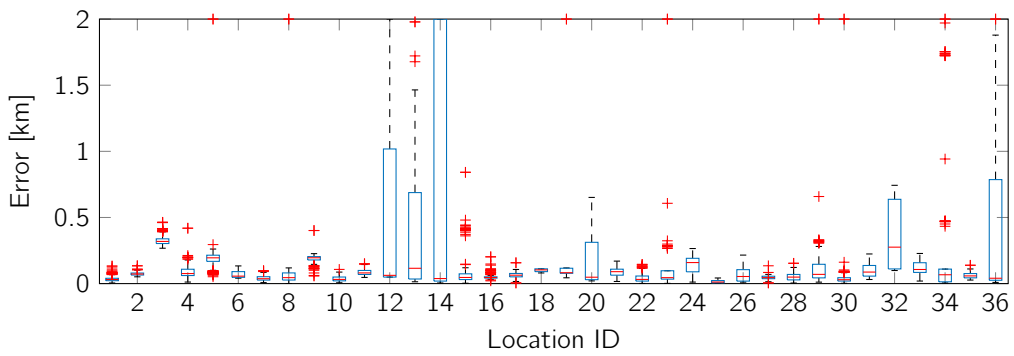
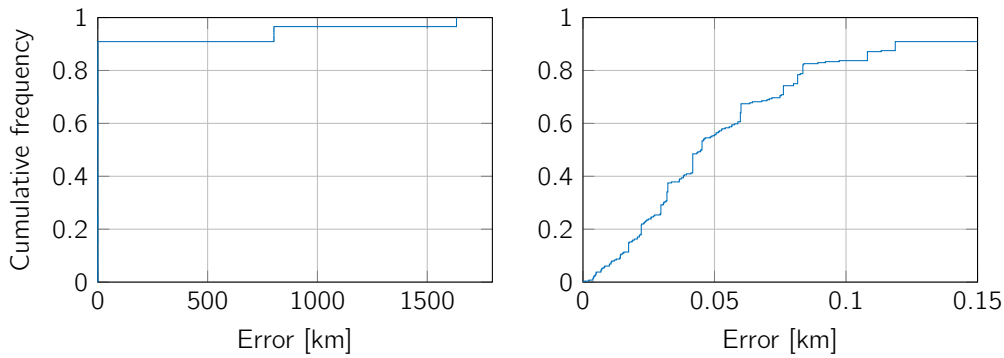


Figure 3.4: Overview of the location error on all 36 locations using the WiGLE database. Errors larger than 2 km are clipped.



(a) *(b)*

*Figure 3.5: Cumulative distribution of the location error at the eighth measurement location, using the WiGLE database. A detail of Figure (**a**) is shown in Figure (**b**). The mean location error is 101 km, the median location error is 45 m.*

The average probability of finding the BSSIDs of a single WAP in the WiGLE database at one of the 36 locations is 32.6%. This is calculated using Equation (3.1) with $n = 1$.

The average probability of finding neither of two BSSIDs in the WiGLE database is 13.2%, which is calculated with $n = 2$ or:

$$P(\neg 2B) = \frac{\binom{\neg B}{2}}{\binom{W}{2}}. \tag{3.2}$$

A graph of the probability of finding neither BSSIDs in the WiGLE database is shown in Figure 3.6.

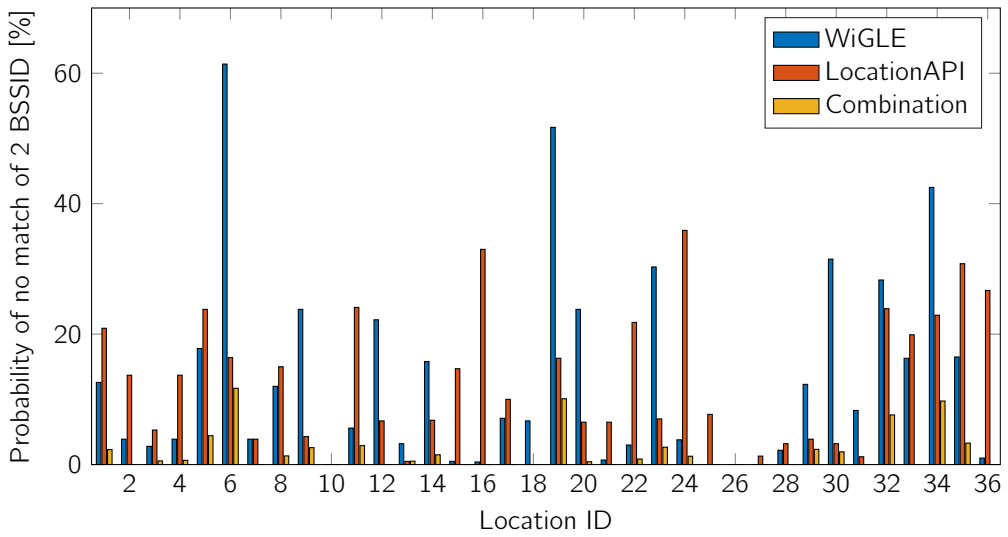


Figure 3.6: Probability of finding no match for either of two BSSIDs in the WiGLE and LocationAPI database. The mean chance of finding no match over all locations is 13.2% for WiGLE, 12.5% for LocationAPI, and 1.9% for the combination.

When sending the combination of BSSIDs with the highest RSS values to the WiGLE database, the mean error of all measurement locations decreases from 6.24 km to only 103 m. The median error decreases from 77 m to 66 m. The probability that neither of those BSSIDs could be found, however, is 11.1%; no location can be estimated at those locations. Figure 3.7 shows the distribution of the error.

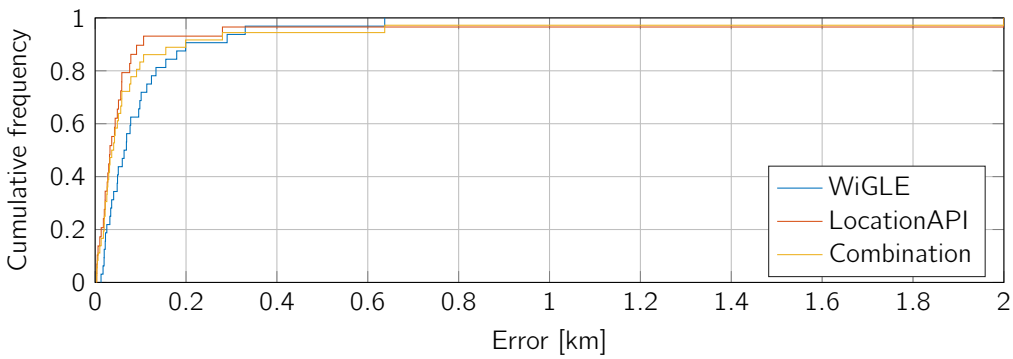


Figure 3.7: Cumulative distribution of the location error when two BSSIDs with the highest RSS are transmitted. One outlier of LocationAPI and the combination is at 2.01 km.

3.2.2 LocationAPI Results

The location estimates when localizing using the LocationAPI database are much less prone to large errors than with the WiGLE database, as shown in Figure 3.8. The overall mean lo-

cation error is 21.06 km, while the overall median location error is 37 m. The 75th percentile location error is never larger than 0.5 km. Seven locations have an error greater than 2 km, which is clipped in Figure 3.8. One of those is location 1. The mean location error is 546 km and the median location error is 23 m. Figure 3.9 shows the cumulative distribution of the errors at this location. There is one combination of BSSIDs with a significant location error: 8543 km. However, more than 93.6% of the BSSID combinations have an error smaller than 43 m, see Figure 3.9b.

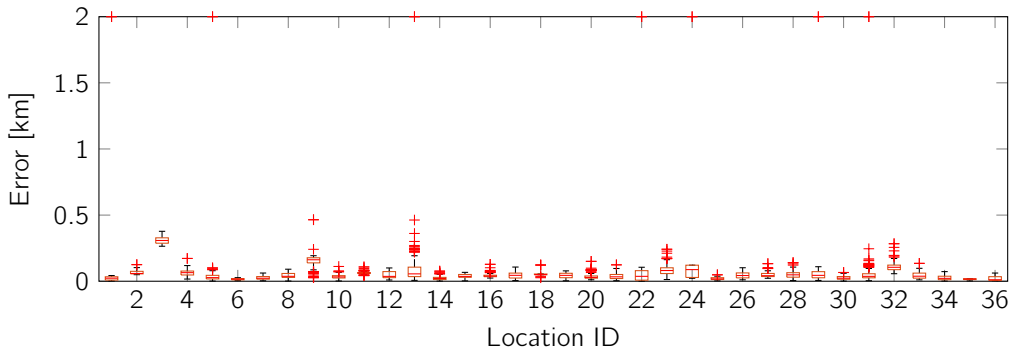
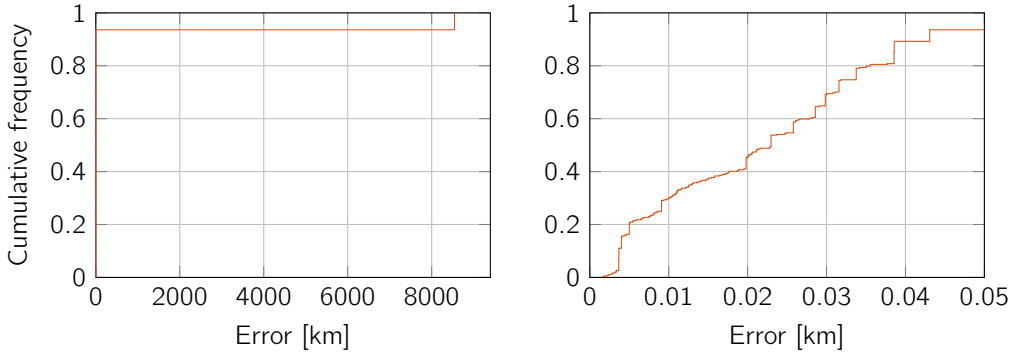


Figure 3.8: Overview of the location error on all 36 locations using the LocationAPI database. Errors larger than 2 km are clipped.



(a) *(b)*

*Figure 3.9: Cumulative distribution of the location error at the first measurement location, using the LocationAPI database. A detail of Figure (**a**) is shown in Figure (**b**). The mean location error is 546 km, the median location error is 23 m.*

The average probability of finding neither BSSIDs in the LocationAPI database of a combination of WAPs seen at one of the 36 location is 12.5%. This is calculated as in Equation (3.2). Since it is not possible to request a single BSSID to the LocationAPI database, the chance of match of a single BSSID cannot be calculated. A graph of the probability of finding neither BSSIDs in the LocationAPI database is shown in Figure 3.6.

When only sending the combination of BSSIDs with the highest RSS values to the LocationAPI database, the mean error of all measurement locations decreases from 16 km to only 114 m. The median error decreases slightly, from 37 m to 33 m. The probability that neither

of those BSSIDs can be found, however, is again 19.4%; no location can be estimated at those locations. Figure 3.7 shows the distribution of the error.

3.2.3 Combination of WiGLE and LocationAPI

A straightforward solution to improve the results is to combine the WiGLE and LocationAPI databases. As a first approach, I propose to combine the location estimates from both databases by taking their average; this result can be seen in Figure 3.10. The overall mean location error is 31.79 km, while the overall median location error is 53 m. The 75th percentile location error is only once larger than 1 km. Fifteen locations have an error greater than 2 km, which is clipped in Figure 3.10.

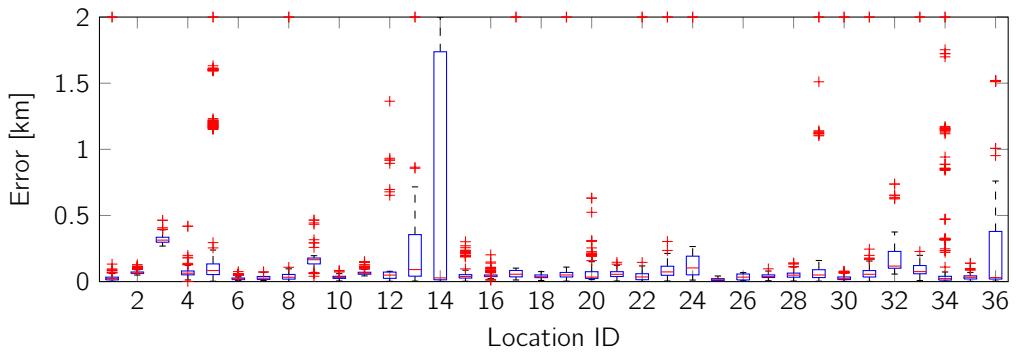


Figure 3.10: Overview of the location error on all 36 locations using the combination of the two databases by averaging. Errors larger than 2 km are clipped.

The location error overview of the second approach, which favors the LocationAPI database since it generally has a lower location error, is shown in Figure 3.11. The overall mean location error is 24.17 km, while the overall median location error is 40 m. The 75th percentile location error is never larger than 0.5 km. Twelve locations have an error greater than 2 km, which is clipped in Figure 3.11.

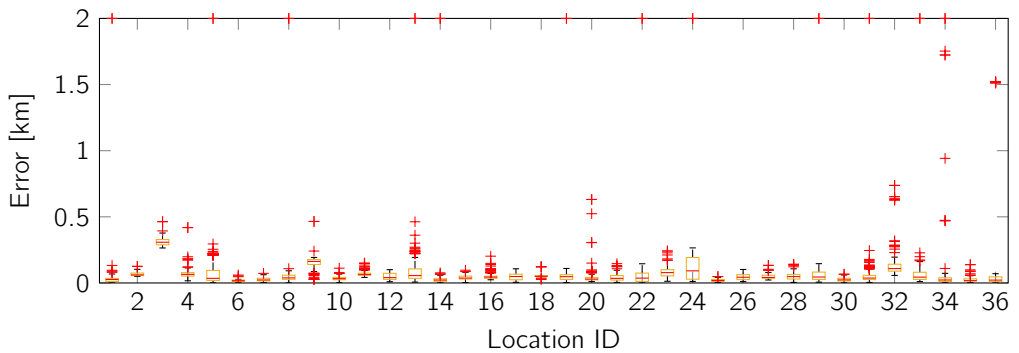


Figure 3.11: Overview of the location error on all 36 locations using the combination of the two databases by favoring LocationAPI over WiGLE. Errors larger than 2 km are clipped.

The average probability of finding neither BSSIDs in the LocationAPI database, nor in the WiGLE database, of a combination of WAP seen at one of the 36 locations is 1.9%. This is calculated as in Equation (3.2). A graph of the probability of finding neither BSSIDs in the combination of databases is shown in Figure 3.6.

When only sending the combination of BSSIDs with the highest RSS values to the combination of databases, the mean error of all measurement locations decreases from 24.17 km to only 125 m. The median error decreases slightly from 40 m to 39 m. Using this method, a location could be estimated at all 36 test locations. Figure 3.7 shows the distribution of the error.

3.3 Discussion

The WiGLE and LocationAPI Wi-Fi fingerprint databases both contain a large number of entries worldwide as well as in Belgium. However, the localization results significantly rely upon these databases.

3.3.1 Database Comparison

When considering the median location errors, the LocationAPI results are usually better than the ones of WiGLE. At only 4 out of 36 measurement locations, the requests to WiGLE yield a smaller median location error, see Figure 3.12. The mean distance errors are in the range of 15 m to 546 km. This variety of mean errors is, as mentioned in the results, caused by outliers in the location estimates.

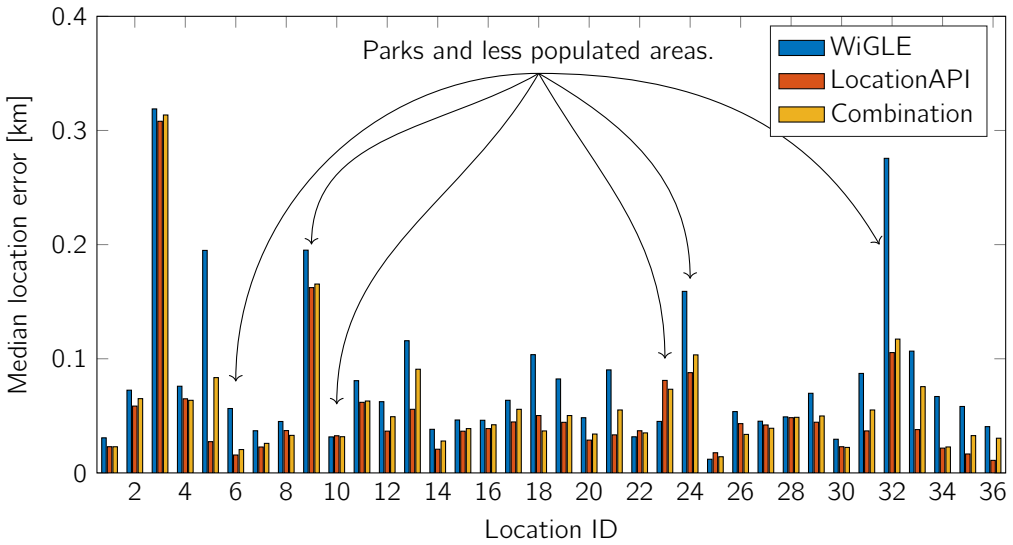


Figure 3.12: Median location error at each test location for the WiGLE and LocationAPI databases, and for their combination. The location error is not consequently larger in less populated areas.

The probability that a BSSID combination is sent that contains two BSSIDs that result in no match is for both databases around 13%. However, there are large differences in this probability at specific locations, see Figure 3.6. This is how the combination of the two databases reduces the probability to 1.9%.

3.3.2 Reasons for Flaws in the Results

When looking at the size of the location errors, I observe that a large error often occurs because a single BSSID leads to a wrong latitude and longitude pair. This causes a large error in every combination which contains that BSSID. When not considering that single 'failed BSSID', the errors will remain quite small and the distance error would decrease. Therefore, I always calculate the median location error, which is more resistant to such outliers.

There will definitely be a fault during the transformation of some BSSIDs to coordinates. Therefore, when localizing a device using Wi-Fi and the WiGLE or LocationAPI database, one should keep in mind that there could be a 'corrupt' BSSID in the databases. When requesting a BSSID, the servers might respond with a wrong coordinate of the wireless access point. The distance between the GPS coordinate and the mean coordinate can vary from a few meters to over several thousands of kilometers. Of course, this influenced the mean error of every measurement where such a flaw occurred.

For example, at location 1, when using the LocationAPI database, the errors of on ore more BSSIDs show a large deviation. Only 5% of all combinations result in an error that is equal to 8543 km. The coordinates that are returned are located in Brazil. The remaining 95% of the errors are always below 44 m. Therefore, the median error is only 23 m, while the mean error is 546 km. This outlier problem is much less apparent when selecting the two BSSIDs with the highest RSS. The error at location 1 is then 3.7 m. Location 29, however, still has an error of 2.01 km. Overall, as shown in Figure 3.7, the mean location error in the WiGLE database is 103 m, the median location error is 66 m, and no location estimate can be found at 4 test locations. In the LocationAPI database, the mean location error is 114 m, the median location error is 33 m, and no location estimate can be found at 7 test locations. When combining the databases, the mean location error is 125 m, the median location error is 39 m, and an estimate can be found at all 36 test locations.

If I select the four BSSIDs with the highest RSS, which requires me to use two Sigfox messages and increases the ten minutes update period to twenty minutes, the location error improves only slightly. In the WiGLE database, the mean location error becomes 56.11 km, due to outliers. The median location error becomes 63 m, which is an improvement of only 3 m. However, no location estimate can be found at only 2 test locations. In the LocationAPI database, the mean location error is 71 m, and the median location error is 31 m. This is a larger improvement than in the case of WiGLE. A location estimate can be found at all 36 test locations. Because the best combination favored the LocationAPI estimate, the results for the combination are equal to those for LocationAPI. Thus, while faults in the results can be alleviated partly by sending two Sigfox messages rather than one, the improvement in location error is not so large as to make it an attractive option.

Additionally, the reliability of the databases can definitely be questioned. For example, anyone can anonymously upload data to the databases. Hopefully, there is a validation step before the data is actually made available to other users. Otherwise users can enter wrong data into the database, for any number of reasons. The managers of the WiGLE and LocationAPI database admit that there could be geographical errors in their databases. However, an administrator of the LocationAPI website stated the following [147]:

> "A lot of effort has gone into enabling our algorithms to auto-adapt based on the type of geography requested, correct for unclean data-sets (in case of crowd-sourced data) by intense cleansing of both contributor and customer data."

Since it is not possible for companies like WiGLE of UnwiredLabs to verify each coordinate of every uploaded WAP, a localization application should have some built-in control mechanisms. Imagine, for example, a user uploading 100 coordinates of access points in a specific, relatively small region. If 99 of the coordinates are only a few meters away from each other, and there is one coordinate a few kilometers further, this could indicate that this single coordinate is the result of the bad resolving in the database. The information of this access point should be left out of the data, to avoid sending a combination with the coordinate of that 'corrupt' BSSID. Another problem which must be tackled is that, even though Media Access Control (MAC) addresses are unique worldwide, there will always be spoofers manipulating their MAC addresses for any number of reasons. Therefore, there should be a Duplicate Address Detection (DAD) mechanism in the localization application. DAD enables eliminating the entry with the 'false location' from the database if the 'real' location is already registered multiple times in the database.

Furthermore, it needs to be mentioned that there could be small distance errors, due to the limitations of the hardware. The GPS and the Wi-Fi adapter of the laptop have a certain refresh speed, so the user location or list of wireless access points is not instantly updated when moving.

3.3.3 Population Density

Usually, a lot of Wi-Fi networks can be discovered in the heart of a city like Antwerp. In the countryside or in open fields or parks, it is harder to detect a Wi-Fi network, because there are just less WAPs available in such areas. Therefore, it is interesting to investigate whether or not there is a relationship between the kind of area and the distance error. Of the 36 locations, there are six locations which can be categorized as less densely populated areas: locations 6, 9, 10, 23, 14, and 32. These locations are indicated in Figure 3.12.

Three out of the five biggest median errors were measured at a less densely populated location. However, the largest error was measured at a densely populated location and there are also errors in less densely populated areas that are smaller than the mean error. So, I cannot confirm if there is a relationship between the distance error and the kind of location. Despite the limitation of using only two BSSIDs, the number of access points in the vicinity does play a role. The chance of having no match in the database increases in less populated areas from 13 to 20%.

3.4 Conclusion

I wanted to know how accurately one can localize a mobile device that communicates over Sigfox by discovering nearby Wi-Fi access points. I used two online databases, WiGLE and LocationAPI, to look up the locations of Wi-Fi access points discovered by the mobile device. Because of the limitations in Sigfox communication, only two BSSIDs can be searched for in a single transmission, with a maximum of 140 transmissions per day.

Using the WiGLE database, I achieved a median location error of 45 m, with a 86.8% chance of finding the combination of BSSIDs transmitted by the device. Using the LocationAPI database, I achieved a median location error of 23 m, with a 87.5% chance of finding the combination transmitted by the device. Both databases have outliers well above 100 km. However, only in the WiGLE database the 75th percentile location error exceeds 1 km, and this only at two of the test locations.

In order to increase the chance of finding the combination of BSSIDs transmitted by the device, I combined the results of both databases. Since the LocationAPI database is more accurate than the WiGLE database, I chose to combine them by only using the WiGLE database when the combination of BSSIDs was not found in the LocationAPI database. With this method, I achieved a median location error of 40 m, with a 98.1% chance of finding the combination transmitted by the device.

Since only 140 transmissions are allowed per day, with only 6 per hour, and only two BSSIDs can be sent in a single transmission, I additionally explored a strategy for selecting the combination of BSSIDs. The strategy consists of selecting the two BSSIDs with the highest RSS value. Using the combination of both databases, the median location error is 39 m, and this combination could be found on all 36 test locations. Moreover, I found that using two messages to transmit four BSSIDs does not decrease the location error significantly.

Future work consists in increasing the localization reliability by building a device that can both sense Wi-Fi access points and transmit two BSSIDs over Sigfox. Such a device would not only validate my results in a closed experiment, but also allow to measure how much energy is exactly required for this type of localization. This would verify if it is a viable way to track and trace assets in an IoT scenario.

Chapter 4

Sigfox Fingerprinting Localization

VARIOUS IoT applications benefit from context awareness. For example, a mobile environmental monitoring device has to send an alert if a certain air quality threshold is exceeded, but the threshold in a dense urban area will be different from the threshold in rural areas. Concisely, the application has to adapt its behavior based on context information such as the current location of the IoT device. Hence, localization methods have to be applied to obtain this context information.

As discussed in Chapter 2, localization based on GNSS is considered not suitable for long-term battery-powered IoT devices that sporadically transmit location updates. Hence, I investigate alternative solutions.

In the previous chapter, I explored the concept of localization by transmitting Wi-Fi fingerprints over a Sigfox network. However, in such an approach, one needs to sacrifice limited payload data for localization purposes. In order to eliminate this need, this chapter investigates localization using the signal strength of the Sigfox communication link itself. By doing so, no additional message or energy is required to obtain a location estimate.

Thus, the current chapter presents my research on outdoor Sigfox fingerprinting localization. More specifically, I use an openly available dataset of Sigfox fingerprints [32] to analyze and optimize the performance of a kNN fingerprinting algorithm. Similar to previous research on indoor Wi-Fi fingerprinting [124], I analyze the effect of 31 distance functions and 4 RSS representations on the location estimation error.

The remainder of this chapter is structured in the following way. Section 4.1 describes the publicly available dataset and how I use it to analyze outdoor Sigfox fingerprinting algorithms. In Section 4.2, I present and discuss the results of this analysis. Finally, Section 4.3 concludes the chapter, summarizing the main observations.

This chapter is based on:

Thomas Janssen, Michiel Aernouts, Rafael Berkvens, & Maarten Weyn, *Outdoor Fingerprinting Localization Using Sigfox*, in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, Sep. 2018, pp. 1–6, https://doi.org/10.1109/IPIN.2018.8533826.

4.1 Methodology

Several steps needed to be taken in order to setup the localization experiments successfully. In this section, I describe the procedure of every experiment and explain why I performed each particular step. Figure 4.1 shows an overview of these steps.

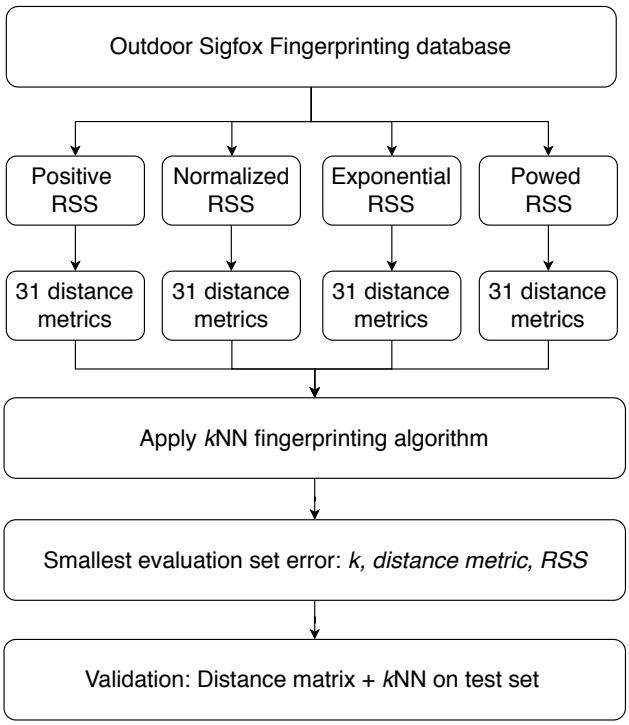


Figure 4.1: Overview of localization experiments

4.1.1 Sigfox dataset

In order to implement fingerprinting in a large outdoor area, an extensive measurement campaign has to be conducted to create a large training database that covers the entire area. Previous work by Aernouts et al. provided the research community with large LPWAN fingerprinting datasets in outdoor areas [32]. One of these datasets holds 14,378 Sigfox messages which were collected in the city center of Antwerp, using the Sigfox network which is deployed by Engie M2M. These messages were transmitted by devices that were mounted on 20 cars of the Belgian postal service. The cars commute through the entire city center on a daily basis, which benefits the spatial spread of the dataset. Every twelve minutes, the current GPS coordinates of such a device are sent via a Sigfox message. Together with the RSS measurements of every Sigfox gateway that received the message, the GPS coordinates of the message were stored in the Sigfox dataset. Consequently, the dataset consists of 14,378 rows where each row represents a unique training sample. The first 84 columns represent all Sigfox gateways which are present in the dataset. If a gateway received

a message, the RSS value for that gateway is filled in in its respective column. Otherwise, an out-of-range value of −200 dBm is inserted. The last columns in the dataset show the receiving time and GPS coordinates of a message.

For the experiments, I divided the Sigfox dataset in a training set, validation (or evaluation) set, and test set. The size of these subsets are respectively 70%, 15%, and 15% of the complete dataset.

4.1.2 Distance functions and RSS representations

In a fingerprinting algorithm, the distance between two points in space needs to be calculated. So far, investigations have been confined primarily to the Euclidean distance and RSS data expressed in dBm. However, various options can be considered to improve the results of the distance calculations. Cha et al. described several distance functions and similarities [148]. Torres-Sospedra et al. implemented all these distance functions for an indoor fingerprinting positioning system using Wi-Fi [124]. Based on this research, I investigate the accuracy of Sigfox in outdoor environments. In total, 31 different distance functions are implemented. Some distance functions are excluded, since some of them are equivalent or irrelevant for the outdoor experiments. Besides, in distance-based methods (such as kNN), some distance functions are equivalent and some similarities occur since two distances can only differ by a single constant. In general, distance functions are categorized into families, based on their similarities. The 31 implemented distances are listed below by family.

- *Minkowski family*: Euclidean, Manhattan, Minkowski-3, Minkowski-4, Minkowski-5, Chebyshev
- *L1 family*: Gower, Sørensen, Soergel, Kulczynski, Canberra, Lorentzian
- *Intersection family*: Intersection, Wave Hedges, Czekanowski, Motyka
- *Inner product family*: Jaccard, Dice
- *Fidelity family*: Hellinger, Matusita, Squared Chord
- *Squared L2 family*: Squared Euclidean, Pearson χ^2, Neyman χ^2, Squared χ^2, Probabilistic χ^2, Divergence, Clark, Additive Symmetric χ^2
- *Combinations*: Kumar-Johnson, Average(L_1, L_∞)

Additionally, Torres-Sospedra et al. list four alternatives to represent the signal strength values: *positive, normalized, exponential* and *powed* values. For the mathematical background of these data representations, the interested reader is referred to the work of Torres-Sospedra et al. [124]. Note that I use the same values for the parameters required in the exponential and powed representations, so $\alpha = 24$ and $\beta = e$. As mentioned before, an out-of-range value of −200 dBm is put into the fingerprinting database for gateways that did not receive a Sigfox message.

4.1.3 Distance matrices

In order to perform a fingerprinting algorithm, distance matrices need to be generated in the next step. Each row in a distance matrix represents a validation sample, while each column represents a training sample. Thus, for every validation sample, the distance (in the signal space, not the geographical distance) to every training sample is stored in the corresponding cell in the distance matrix. This is visualized in Figure 4.2. For every distance function and for every RSS representation, a distance matrix is generated.

Figure 4.2: Representation of a distance matrix

4.1.4 Fingerprinting algorithm

Given the dataset and the distance matrices, I can start processing all this data and apply a fingerprinting algorithm. Many algorithms are available, e.g., *k*NN, SVMs [149] and Neural Networks (NNs). For my experiments, I apply the *k*NN classification algorithm. Note that I evaluate the performance of other algorithms in Chapter 5. In the *k*NN algorithm, the coordinates of the k nearest training samples are used to estimate the position of the device of interest. In practice, a sample of signal strengths is compared to the training samples in the fingerprint database. Using the *k*NN algorithm, I can then obtain a location estimate for the transmitting device. In pursuance of achieving the highest possible accuracy, I search for the optimal k for every distance function and RSS representation, with k ranging from 1 to 16. The optimal k is defined as k where the smallest distance is measured.

4.1.5 Error calculation

After applying the fingerprinting algorithm, I calculate the error between the estimated position and the actual position of the device. This is done by calculating the Vincenty distance, which takes into account a newer and more accurate model of the curvature of the Earth than the Haversine distance. For every distance function and RSS representation, the mean distance error is calculated for every k in the validation set and for the optimal k in the test set, which is discussed in the next section.

4.1.6 Validation

Finally, in order to validate the parameters of the fingerprinting algorithm and the results, a test dataset is used. From this set, a distance matrix is generated for the distance function and RSS representation with the smallest distance error and optimal k. In that way, I can compare the distance error of the test dataset to the distance error of the validation dataset, for the optimal k and distance function. If these errors are similar, the results are assumed to be valid.

4.2 Results

In this section, I present the results of the experiments and compare the different distance functions and RSS representations to each other. In order to validate the results, I use a test set for the case with the smallest error, optimal k and RSS representation.

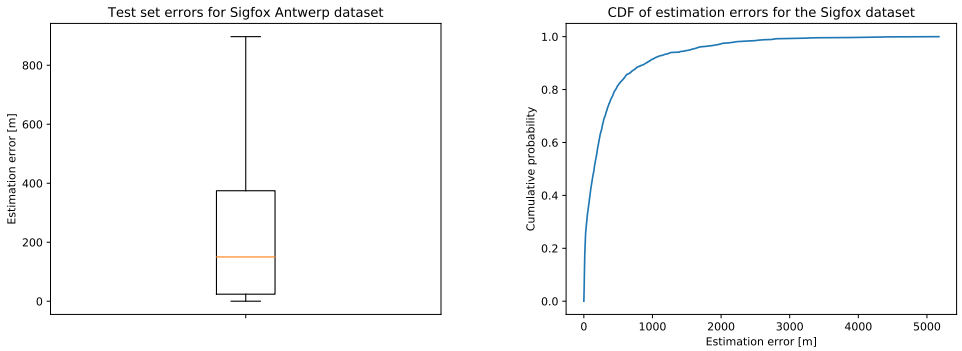
In Table 4.1, the optimal k and mean location estimation error are listed for every distance function and every RSS representation. In almost all cases, the lineal representations (positive and normalized) yield similar results. Furthermore, one can observe a decrease of the mean error for the exponential and powed representations. Using one of these two representations always results in a smaller distance error as when using a lineal representation, except for the Kumar-Johnson distance. When comparing the exponential and powed representations, the exponential RSS representation always yields the smallest distance error. The Sørensen, Soergel, Kulczynski, Czekanowski and Motyka distance functions all yield the exact same errors. Using the exponential RSS representation and $k = 6$, the smallest distance error can be achieved. In these cases, the mean estimation error is 322 m.

Since Sørensen distance requires the least computational power, I choose this distance function for the test set calculations. Thus, the test set distance matrix is created using exponential RSS representation, the Sørensen distance and $k = 6$. When performing the same fingerprinting algorithm as with the validation set, the mean test set error equals 340 m, which is close to the mean error of 322 m obtained with the validation set. Therefore, I can assume the results to be valid. Figures 4.3a and 4.3b show a box plot and Cumulative Distribution Function (CDF) of the test set errors, respectively. The median error is 150 m and the 95th percentile yields an error of 1548 m.

In Figure 4.4, the correlation between the number of gateways that received a test message and the estimation errors for that message is shown. In general, one can observe that if more gateways receive a message, the estimation error decreases. Consequently, Sigfox network operators should take the number of gateways into account to improve the location estimation accuracy.

Table 4.1: Fingerprinting results for the Sigfox urban dataset, showing the optimal value of k and the mean location estimation error in meter for every RSS representation and distance function.

Measure	Positive RSS		Normalized RSS		Exponential RSS		Powed RSS	
	k	Error	k	Error	k	Error	k	Error
Euclidean	10	667	10	667	6	329	9	361
Manhattan	7	537	7	537	6	333	6	356
Minkowski 3	9	729	9	729	6	341	7	373
Minkowski 4	13	752	13	752	6	357	6	394
Minkowski 5	14	762	14	762	6	367	7	409
Chebyshev	7	906	7	906	5	413	4	494
Gower	7	537	7	537	6	333	6	356
Sørensen	9	539	9	539	**6**	**322**	7	346
Soergel	9	539	9	539	**6**	**322**	7	346
Kulczynski	9	539	9	539	**6**	**322**	7	346
Canberra	9	687	9	687	8	488	7	543
Lorentzian	6	460	6	524	6	333	6	357
Intersection	7	537	7	537	6	333	6	356
Wavehedges	7	613	7	613	7	453	7	481
Czekanowski	9	539	9	539	**6**	**322**	7	346
Motyka	9	539	9	539	**6**	**322**	7	346
Jaccard	11	677	11	677	7	325	6	352
Dice	11	677	11	677	7	325	6	352
Hellinger	9	775	9	775	8	417	9	567
Matusita	9	775	9	775	8	417	9	567
Squared Chord	9	775	9	775	8	417	9	567
Squared Euclidean	6	789	10	667	6	329	9	361
Pearson χ^2	3	1122	3	1113	3	575	2	847
Neyman χ^2	4	1056	6	1048	3	739	4	830
Squared χ^2	12	747	12	747	6	391	8	477
Probabilistic Symmetric χ^2	10	758	12	747	6	391	8	477
Divergence	10	707	10	707	9	588	10	682
Clark	9	687	9	687	8	487	7	543
Additive Symmetric	6	955	6	948	7	555	6	939
Kumar-Johnson	6	950	6	940	7	636	6	947
Average(L_1, L_∞)	8	853	6	730	6	390	6	443



(a) *(b)*

*Figure 4.3: (**a**) Box plot and (**b**) Cumulative Distribution Function of estimation errors of the Sigfox urban test dataset.*

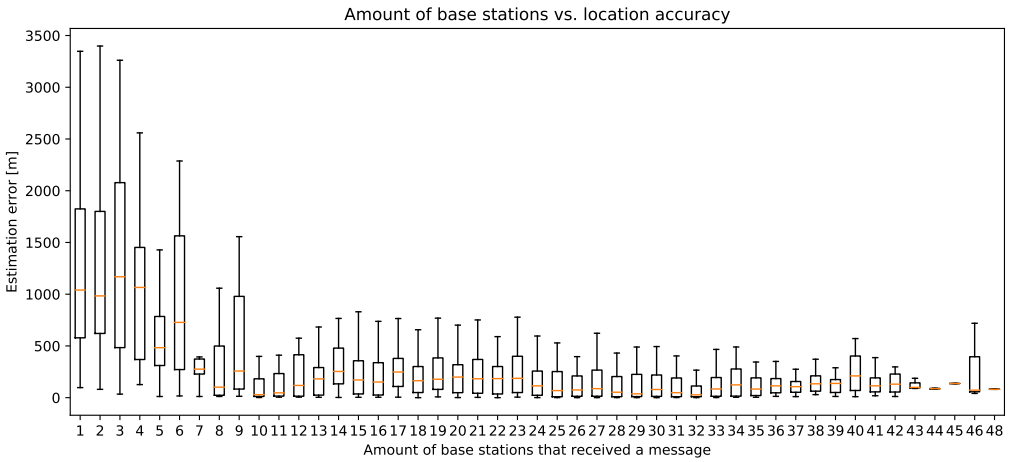


Figure 4.4: Box plots of estimation errors of the Sigfox urban test dataset.

4.3 Conclusion

Due to the proliferation of IoT devices, new energy-efficient technologies are needed to locate them. In this chapter, I carried out experiments to investigate the accuracy of RSS-based fingerprinting localization using Sigfox. A kNN algorithm was implemented to match a new fingerprint to a training dataset. In that way, I was able to estimate the position of the mobile device.

Furthermore, I improved the location estimation by analyzing 31 distance functions and four ways to represent RSS data. In the optimal case, $k = 6$ and the Sørensen distance is used in combination with the exponential RSS representation. Compared to the state-of-the-art in the same network and region [103], my results double the accuracy with a reduction in mean location estimation error from 689 m to 340 m.

The achieved positioning accuracy might still seem very poor compared to GNSS. However, this level of accuracy is sufficient to provide location-awareness in many IoT applications. When tracking an asset worldwide, for instance, a location estimate with an accuracy of 340 m does give you a good idea of where it currently resides. Moreover, remember that no additional energy on the mobile device is required to obtain a location estimate.

In future work, several options can be explored to further improve the location estimation accuracy of the fingerprinting algorithm. First, my results show higher location estimation errors in areas with few receiving gateways. Therefore, increasing the number of gateways in sparsely covered areas will generally increase the localization accuracy in these areas. Second, fine-tuning of the α and β parameters in the exponential and powed RSS representations, respectively, may lead to higher location accuracy. Third, the kNN algorithm can be improved by implementing a weighted variant of this algorithm. Finally, more studies are required to optimize the size of the datasets. In the next chapter, I will repeat this analysis using other ML algorithms on a LoRaWAN dataset that was also collected in the city center of Antwerp [32].

Chapter 5

RSS-based LoRaWAN Localization

WHEN choosing a localization algorithm for any IoT application, a trade-off between location accuracy, deployment cost and computational performance should be considered. Therefore, this chapter examines this trade-off for RSS-based localization in a LoRaWAN network. I use a publicly available outdoor LoRaWAN data set to evaluate fingerprint-based and range-based location estimation algorithms. The data is collected in the same environment as in the experiments of the previous chapter, enabling positioning performance comparison between Sigfox and LoRaWAN. The experimental results presented in this chapter serve as a benchmark for RSS-based LoRaWAN localization methodologies in a real-world large-scale environment.

The major contributions in this chapter can be summarized as follows:

- In an RSS-based fingerprinting approach, I evaluate and compare the performance of ten Machine Learning (ML) algorithms in terms of location accuracy and computational complexity.
- I assess the accuracy of RSS-based ranging approaches by implementing and comparing several outdoor path loss models. In addition, I evaluate a modified version of the E-Min-Max location estimation algorithm using different weight functions.
- I implement different gateway selection strategies in order to improve the range-based localization accuracy.
- I discuss the benefits and limitations of both fingerprint-based and range-based outdoor localization approaches.

The rest of this chapter is organized as follows. Section 5.1 briefly describes the public LoRaWAN data set used to evaluate all localization algorithms. Subsequently, the large-scale localization experiments are outlined. While Section 5.2 presents ML algorithms to apply in a fingerprinting approach, Section 5.3 describes novel combinations of path loss models and location estimation algorithms in a range-based approach. Furthermore, different gateway selection strategies are discussed. Section 5.4 presents the results of the experiments in terms of localization accuracy and performance. Finally, the benefits and limitations of each approach are discussed in Section 5.5 and the main conclusions are drawn in Section 5.6.

This chapter is based on:

Thomas Janssen, Rafael Berkvens, & Maarten Weyn. *Comparing ML Algorithms for RSS-Based Localization in LPWAN*, in *Proceedings of the 14th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2019)*. Editors: L. Barolli, P. Hellinckx, J. Natwichai, Springer International Publishing, 2019, pp. 726-735, https://doi.org/10.1007/978-3-030-33509-0_68.

Thomas Janssen, Rafael Berkvens, & Maarten Weyn. *Benchmarking RSS-based localization Algorithms with LoRaWAN*. Internet of Things, vol. 11, p. 100235, Sep. 2020, https://doi.org/10.1016/j.iot.2020.100235.

5.1 Publicly available LoRaWAN data set

The algorithms presented in this chapter are evaluated on a publicly available data set. This enables other researchers to compare their algorithms on the same data set. Aernouts et al. collected a large number of LoRaWAN messages in the city of Antwerp, Belgium [32]. In total, 130,430 messages are collected by 72 LoRaWAN gateways throughout the city. Each message consists of the RSS in dBm to all gateways, together with a GPS coordinate. The latter is used as a ground truth reference. If a message is not received by a gateway, the RSS value for that gateway is set to −200 dBm. Figure 5.1 displays the messages inside a predefined bounding box on a map of Antwerp, Belgium. The gateway locations are not shown, since these are confidential.

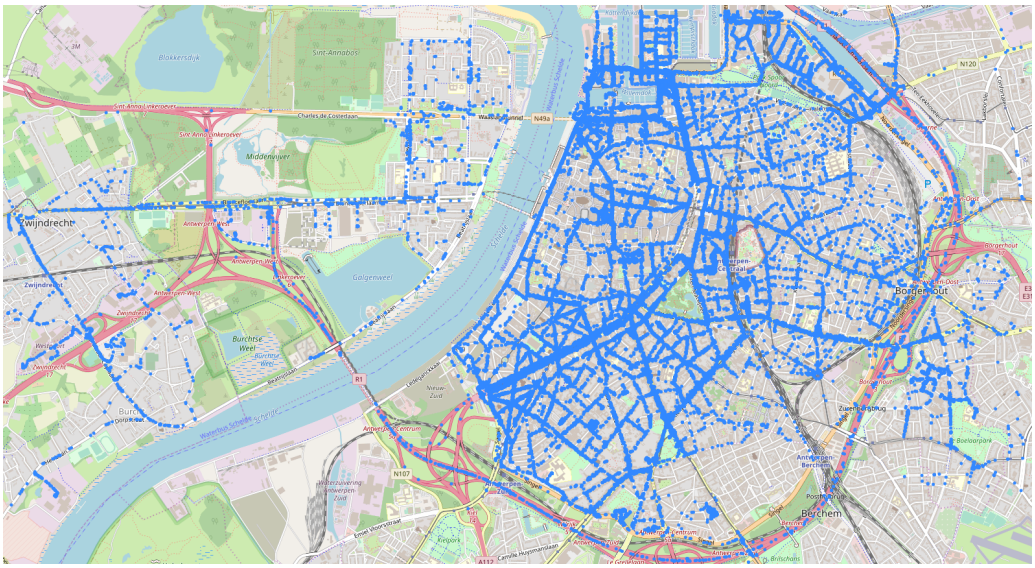


Figure 5.1: The public data set consists of 130,430 LoRaWAN messages collected in the city of Antwerp, Belgium [32]. The blue dots represent the GPS coordinates of every measurement. © 2020 OpenStreetMap contributors

5.2 Machine Learning-based fingerprinting algorithms

The location of a mobile transmitter can be estimated by matching the RSS values from receiving gateways to earlier collected data at known locations. In this section, I describe the steps taken to evaluate a fingerprint-based localization approach, leveraging the power of ML algorithms. First, some preprocessing steps need to be taken in order to represent the data in the most optimal way. Second, since I am predicting continuous-valued output, regression-based algorithms are most suitable for this supervised ML problem. In total, ten regression algorithms are evaluated, which can be classified into four different categories. For every category of algorithms, I clarify how the algorithms work and what parameters are optimized. Finally, the benefits and limitations of each algorithm are discussed.

In my experiments, each algorithm is benchmarked in terms of three performance metrics, defined as follows:

- Location estimation error, i.e., the distance in meter between the GPS coordinate and the estimated coordinate;
- The R^2 score, i.e., a ML performance metric indicating how close the actual target values are to the fitted regression line;
- The computation time, i.e., the elapsed time between the fitting of a model and the estimation of an output coordinate for all validation messages, using a Virtual Machine with 32 GB Random Access Memory (RAM) memory and 10 Central Processing Unit (CPU) cores, of which 6 are used in parallel in every algorithm.

5.2.1 Preprocessing steps

To put the LoRaWAN data set discussed in Section 5.1 into a more ML perspective, each message is related to a single sample and each receiving gateway relates to a feature. Therefore, the data set consists of 130,430 samples and 72 features in total. In ML, data preprocessing is required to prepare raw, sometimes incomplete data for further processing. Hence, before feeding the ML algorithms with the data set, some preprocessing steps are taken.

A common first step in supervised ML approaches is to divide the data set into three subsets. The largest subset (70%) of the data set is used as training data, 15% of all messages is used to validate the localization algorithms and 15% is used as unbiased test data to evaluate the results obtained through the validation set.

Subsequently, the RSS data is transformed into another format. In the raw data, the RSS values are represented in decibels relative to a milliwatt (dBm). Similar to the experiments in Chapter 4, I evaluate the localization accuracy when using the normalized, exponential and powed RSS representation [124]. While some scaling is already introduced when modifying the RSS representations, the data is scaled before passing it to the input of the ML algorithms.

Finally, a Principal Components Analysis (PCA) is performed on the data set. With PCA, I can extract the most relevant features out of the data set. This is highly desired, since not all gateways receive each message, thus reducing the number of features and noise. Moreover, extracting the principal components of the data set is often used to decrease the computation time of each ML algorithm. I performed PCA on the data set with 95% of the variance retained, resulting in a reduction from 72 to 40 components.

5.2.2 Linear regression algorithms

Linear regression algorithms attempt to fit a linear function from the provided training data and estimate the numeric output values, given new input values. By fitting the function, a

linear ML model is created. This model can be represented in the form:

$$y_{pred} = w[0] * x[0] + w[1] * x[1] + ... + w[n] * x[n] + b, \tag{5.1}$$

where x is the feature vector of length n and w and b are the parameters that need to be learned by training the model.

Several variations of linear regression algorithms exist. In this research, I evaluate the following linear algorithms: Ordinary Least Squares (OLS), Ridge, Lasso, Elastic Net, Stochastic Gradient Descent (SGD) and Polynomial regression.

As the name suggests, Ordinary Least Squares is the most basic linear regression algorithm. In this algorithm, the Mean Square Error (MSE) between the predicted and real output values is minimized. If the features of the data are correlated, the number of random errors in the target values increases. This phenomenon is called multicollinearity. Therefore, the independence of the features is very important in the OLS algorithm.

Ridge regression is similar to OLS, with the difference that in Ridge regression the magnitude of the coefficients w is reduced by a factor α, as can be seen in Eq. (5.2). This constraint is known as ℓ_2 regularization. The optimal value of α is found by evaluating Ridge regression with cross-validation. During the cross-validation, the value of α is optimized through approximate value iteration.

$$\min_{w} \|Xw - y\|_2^2 + \alpha \|w\|_2^2 \tag{5.2}$$

In the Lasso regression algorithm, some coefficients in Eq. (5.1) are set to zero. Consequently, some features are ignored by the model. This is called ℓ_1 regularization. Similar to Ridge, the optimal value of α is found through cross-validation. With only a few non-zero weights, the advantage of the Lasso regression is the reduced amount of time needed to train the model.

The Elastic-Net linear regression model performs both ℓ_1 and ℓ_2-norm regularization of the coefficients. In fact, it is a combination of the Ridge and Lasso algorithms, in the sense that there are fewer non-zero weights and the regularization properties of Ridge are maintained.

Stochastic Gradient Descent is a ML algorithm that can be used for classification and regression problems. It is often used for training Artificial Neural Networks (ANNs). However, it can also be used for training linear regression models. Gradient Descent is a method to find the values of a function that minimizes a cost function. To find the optimal values, the initial parameters are constantly updated. Thus, Gradient Descent is an iterative method, resulting in slower training times. In the stochastic variant of this algorithm, one iterates over a few randomly selected samples, reducing the computational complexity in large data sets. Furthermore, different cost functions and corresponding parameters have been evaluated on our data set. By iterative approximation, it was found that the Huber cost function lead to the most accurate results (i.e., smallest location estimation error) with $\epsilon =$ equal to 1×10^{-2} and a stopping criterion equal to 1×10^{-3}. The Huber cost function uses a squared cost function and past a distance of ϵ it uses a linear cost function. Therefore, it becomes less sensitive to outliers in the data set.

The last linear regression algorithm I discuss is the polynomial regression algorithm, in which a linear function is fitted within a higher-dimensional space. Thus, the degree of a polynomial function is lowered. In our case, the localization accuracy was maximized when reducing the degree to 1. This approach allows to fit a much wider range of data and benefits from the relatively fast performance of linear regression algorithms.

5.2.3 Support Vector Regression

Support Vector Regression (SVR) can be used in classification and regression problems, both in lower and higher dimensional cases. Simple regression algorithms attempt to minimize the error rate, while in SVR, the goal is to maximize the margin between the hyperplane and the support vectors (i.e., the data points closest to that hyperplane). In other words, one needs to find a function that has at most ϵ deviation from the actually obtained targets in the training data [150]. Hence, a loss function with a margin of tolerance ϵ is defined because of the real numbered target values. Furthermore, SVR is characterized by a kernel function that maps lower dimensional data to higher dimensional data. I implemented SVR with a third-degree polynomial kernel function and free parameters $\epsilon = 0.01$ and $C = 1000$, determined by iterative approximation on the training subset. The parameter C controls the penalty imposed on observations that lie outside the ϵ margin and helps to prevent overfitting. The main advantage of SVR is that the computational complexity does not depend on the dimensionality of the input space. However, when the number of samples in the data set exceeds a few tens of thousands, the algorithm becomes computationally demanding.

5.2.4 k Nearest Neighbors

The k-Nearest-Neighbors (kNN) algorithm is an intuitive yet effective ML approach which is often used in indoor and outdoor localization applications [128]. During the offline training phase of the fingerprinting algorithm, the feature vectors and target values are stored. In the online validation phase, I estimate the target values (i.e., the GPS coordinates) of a test feature vector. This is done by calculating the distance or similarity between each training vector and the test vector [124]. As a final step, the centroid of the target values of the k smallest distances is used as the estimate for the target values, where k is user-defined. In our experiments, I iterate over k ranging from 1 to 20, thus optimizing the number of nearest neighbors. In the weighted variant of the kNN algorithm, neighbors that are closer to the test sample have a greater influence than neighbors which are farther away. An important advantage of the kNN algorithm is the fact that there is no explicit training phase, i.e., the training messages only need to be stored, instead of used to fit a model as in previously mentioned algorithms. Other benefits of kNN are the simplicity of the algorithm and the variety of distances to choose from. On the contrary, the user-defined value of k, the computational complexity and the high outlier sensitivity are the limitations of the algorithm.

5.2.5 Random Forest

Random Forest is an ensemble technique, i.e., multiple ML algorithms are combined to solve classification or regression problems. The general idea is to construct multiple decision trees during the training phase and output the mean of all individual predictions as an estimated target value. The technique randomly samples the training observations when building the trees. Random Forest has been proven to outperform other ML algorithms in terms of indoor fingerprinting localization accuracy [151]. Despite the computational complexity of each individual decision tree, the overall training and matching speeds are very fast, even for high-dimensional input data. Finally, since multiple algorithms are combined, overfitting is reduced significantly and the stability of the technique increases. In my implementation, I chose 100 estimators that are combined using the bootstrap aggregation (bagging) technique. Increasing the number of trees above 100, vastly increased computational cost and only slightly improved prediction performance.

5.3 Range-based localization algorithms

Besides fingerprint-based methods, range-based methods can estimate the location of a mobile device as well. As discussed in Chapter 2, range-based estimation algorithms use a path loss model to convert an RSS value from a specific gateway into an estimated distance to that gateway. Afterwards, geometrical principles are used to estimate the location of a device. One should choose a path loss model depending on the used technology and the environment under consideration. Given the dynamic and heterogeneous environment of the city of Antwerp, I implement and evaluate the performance of several path loss models, which are specified in Section 5.3.1. The actual localization algorithm and its weighted variants are discussed in Section 5.3.2.

5.3.1 Path loss models

The environment under consideration consists of the urban city center of Antwerp, in combination with more rural and open spaces, such as the river Scheldt, docks, quays and parks. This makes it a challenging task to find an optimal path loss model for this environment, because most outdoor path loss models are either optimized for use in an urban, suburban or rural area. Furthermore, I do not want to create a custom path loss model for Antwerp, since I want to provide a general deterministic algorithm which can be deployed in similar environments.

In the next paragraphs, I describe the path loss models which are used to evaluate the accuracy of range-based LoRaWAN localization. Table 5.1 lists the parameters used in these path loss models.

Table 5.1: Parameters used in the path loss models for the range-based localization algorithms.

Description	Symbol	Value	Unit
Frequency	f	868	MHz
Gateway height	h_b	30	m
Mobile node height	h_m	1.5	m
Building (rooftop) height	h_r	25	m
Street width	w	10	m
Incident angle	ϕ	90	°
Distance between buildings	B	10	m
Distance between Tx and Rx	d	N/A	m

5.3.1.1 Okumura-Hata models

The Okumura-Hata model is an empirical path loss model which can be used in an outdoor context with transmission distances up to 20 km [107]. The path loss equations for urban, suburban and rural environments are:

$$L_{urban} = A + B\log_{10}(d), \tag{5.3}$$

$$L_{suburban} = A + B\log_{10}(d) - C, \text{ and} \tag{5.4}$$

$$L_{rural} = A + B\log_{10}(d) - D, \tag{5.5}$$

with:

$$A = 69.55 + 26.161\log_{10} f - 13.82\log_{10} h_b - 3.2[\log_{10}(11.75h_m)]^2 + 4.97,$$
$$B = 44.9 - 6.55\log_{10} h_b,$$
$$C = 5.4 + 2[\log_{10}(f/28)]^2, \text{ and}$$
$$D = 40.94 + 4.78(\log_{10} f)^2 - 18.33\log_{10} f.$$

5.3.1.2 COST-231 Hata models

The COST-231 Hata model extends the original Hata model and the path loss formula is [107]:

$$L = A + B\log_{10} d + C \tag{5.6}$$

with

$$A = 46.3 + 33.9\log_{10} f - 13.28\log_{10} h_b - 3.2[\log_{10}(11.75h_m)]^2 + 4.97,$$
$$B = 44.9 - 6.55\log_{10} h_b, \text{ and}$$
$$C = 0 \text{ (for suburban environments) or}$$
$$C = 3 \text{ (for urban environments).}$$

5.3.1.3 IEEE 802.11ah models

The IEEE TGah working group proposed several outdoor path loss models for use with IEEE 802.11ah, i.e., in sub-GHz frequency bands [152]. I evaluate two of these models: a near-LoS macro deployment (further referred to as ah-macro) and the ah-pico model. The equations for these models are:

$$L_{macro} = 8 + 36.7 \log_{10} d \tag{5.7}$$

and

$$L_{pico} = 23.3 + 36.7 \log_{10} d + 21 \log_{10} (f/900), \tag{5.8}$$

respectively.

5.3.1.4 COST-231 Walfisch-Ikegami model

Finally, the COST-231 Walfisch-Ikegami model is a widely used path loss model which can be highly configured with parameters such as building heights and street widths [153]. Since I am evaluating the path loss models in a large-scale heterogeneous environment, the parameters set in Table 5.1 are averaged values. The COST-231 Walfisch-Ikegami path loss:

$$L = L_{fs} + L_{rts} + L_{msd}, \tag{5.9}$$

is calculated by adding the free space loss:

$$L_{fs} = 32.4 + 20 \log_{10} d + 20 \log_{10} f, \tag{5.10}$$

rooftop to street and scatter loss:

$$L_{rts} = -16.9 + \log_{10} f + 20 \log_{10} (h_r - h_m) - 10 \log_{10} w + L_{ori}, \tag{5.11}$$

with the orientation factor given by: $L_{ori} = 4 - 0.114(\phi - 55)$ and multi-screen diffraction loss:

$$L_{msd} = L_{bsh} + k_a + k_d \log_{10} d + k_f \log_{10} f - 9 \log_{10} B, \tag{5.12}$$

with factors $k_a = 54$, $k_d = 18$ and $k_f = -4 + 1.5(f/925 - 1)$. Since the gateway height is higher than the average rooftop height, the shadowing gain is given by:
$L_{bsh} = -18 \log_{10} (1 + h_b - h_r)$.

5.3.2 Location estimation algorithm

Through the path loss models presented in the previous section, the distances between the device and the receiving gateways are calculated. A location estimation algorithm leverages

these distances to estimate the location of the mobile transmitter. I modify an extended version of the well established Min-Max algorithm, referred to as E-Min-Max [117]. This algorithm can be explained visually by composing a rectangle based on the estimated distances to every receiving gateway, as shown in Figure 5.2. Given a certain weight function, a weight is calculated for each vertex of the rectangle. The weights are based on both the estimated distances from the gateways and the Karney distance (i.e., an improved version of the Haversine and Vincenty distance) between the gateways and that vertex [154]. Finally, the transmitter location is determined by taking the weighted average of the locations of the vertices.

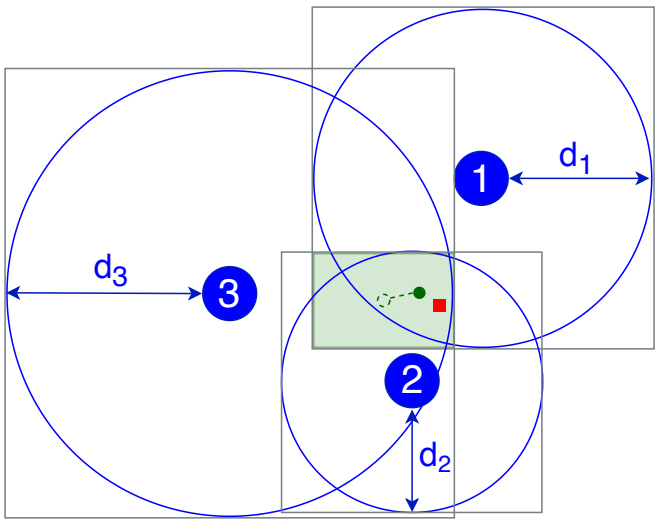


Figure 5.2: A graphical representation of the E-Min-Max algorithm with estimated distances to 3 gateways, represented by the blue circles. The estimated location (indicated with a green dot) is shifted from the center of the Min-Max rectangle depending on the weights to every vertex of the rectangle. The red square indicates the actual location of the mobile transmitter.

Originally, four weight functions were introduced in [117]. Additionally, two new weight functions were introduced in [155]. In order to use these functions, I modify them to work with the WGS-84 coordinate reference system of the Earth. The weight functions for a vertex j are defined in Eqs. (5.13) to (5.19), where W_0 refers to the non-weighted Min-Max algorithm:

$$W_0(j) = 0.25 \tag{5.13}$$

$$W_1(j) = \frac{1}{\sum_{i=1}^{n} \left| K_{i,j} - d_i \right|} \tag{5.14}$$

$$W_2(j) = \frac{1}{\sum_{i=1}^{n} \left(K_{i,j} - d_i \right)^2} \tag{5.15}$$

$$W_3(j) = \frac{1}{\sum_{i=1}^{n} \left| M_{i,j} - d_i \right|} \tag{5.16}$$

$$W_4(j) = \frac{1}{\sum_{i=1}^{n} \left| K_{i,j}^2 - d_i^2 \right|} \tag{5.17}$$

$$W_5(j) = \frac{1}{\sum_{i=1}^{n} \left(M_{i,j} - d_i \right)^2} \tag{5.18}$$

$$W_6(j) = \frac{1}{\sum_{i=1}^{n} \left| M_{i,j}^2 - d_i^2 \right|} \tag{5.19}$$

with d_i the estimated distance from the device to gateway i and $K_{i,j}$ and $M_{i,j}$ the Karney distance and the Manhattan distance on the surface of the Earth as a WGS-84 ellipsoid from gateway i to vertex j, respectively.

After selecting a weight function from the above, the final coordinates of the mobile device are estimated as:

$$[lat_0, lon_0] = \left[\frac{\sum_{j=1}^{4} W_a(j) \cdot lat_j}{\sum_{j=1}^{4} W_a(j)}, \frac{\sum_{j=1}^{4} W_a(j) \cdot lon_j}{\sum_{j=1}^{4} W_a(j)} \right] \tag{5.20}$$

5.3.3 Gateway selection strategies

The number of receiving gateways plays a significant role in localization research [79]. Therefore, I investigate this aspect in more detail in my experiments. One can expect LoRa gateways located tens of kilometers away from the mobile transmitter to be able to pick up the LoRa signal in ideal circumstances and open areas. From a localization point of view, the number of receiving gateways located far away from the transmitter should be limited, as they might decrease the localization accuracy. To that end, I run my experiments with three different gateway selection strategies, listed below.

1. I evaluate the localization performance using all receiving gateways of a given message. Participating gateways can be located throughout the public nationwide LoRaWAN network.

2. I only consider messages that were picked up by three or more gateways. If only one or two gateways receive a message, the range-based E-Min-Max algorithm cannot be evaluated. In that case, the algorithm boils down to a simple proximity algorithm, where the location of the gateway with the strongest link is used as a location estimate of the mobile device. However, I want to evaluate the accuracy of range-based

methods in this experiment. Thus, all messages with less than three measurements are discarded. Obviously, I do keep track of the number of messages with three or more gateways.

3. A last strategy extends the previous one, adding a geographical limitation. The goal in this strategy is to only use messages with at least three receiving gateways located inside a specific bounding box of Antwerp. This bounding box is constructed to remove receiving gateways that are far away and result in a high number of outliers. It is defined by the following minimum and maximum latitude/longitude pairs: (51.171931, 4.297380) and (51.245467, 4.518483). Note that the bounding box is larger than the bounding box for collected messages, as defined in Section 5.1. This is because messages near the edge of the inner bounding box can still be received by nearby gateways outside the inner bounding box, but inside the outer bounding box.

5.4 Results

5.4.1 Fingerprint-based localization performance

The ten ML algorithms presented in Section 5.2 are now evaluated in terms of location estimation error, R^2 score and the time needed to compute the results. Table 5.2 summarizes these metrics for every algorithm and for the lineal, exponential and powed representations of the fingerprinting data set.

Table 5.2: Mean location estimation error (in meters), R^2 score, and computation time (in seconds) for every ML algorithm using the lineal, exponential and powed RSS representation of the LoRaWAN data set.

Algorithm	**Lineal RSS**			**Exponential RSS**			**Powed RSS**		
	Error	R^2	Time	Error	R^2	Time	Error	R^2	Time
Linear OLS	801	0.70	1.19	785	0.73	1.15	786	0.72	1.08
Linear Ridge	800	0.70	**0.82**	785	0.73	0.84	785	0.72	0.83
Linear Lasso	801	0.70	0.88	785	0.73	0.90	786	0.72	0.98
Linear Elastic Net	801	0.70	0.94	785	0.73	0.92	786	0.72	0.98
Linear SGD	799	0.70	11	784	0.73	11	784	0.72	12
Linear Polynomial	801	0.70	1.20	785	0.73	1.20	786	0.72	1.29
SVR	1148	0.57	1168	1206	0.49	1616	1155	0.55	1162
*k*NN	354	0.90	131	345	0.90	517	349	0.90	131
*k*NN weighted	348	0.90	146	344	0.90	484	343	0.90	147
Random Forest	351	**0.91**	56	609	0.84	56	**340**	**0.91**	53

For the set of linear regression algorithms, the exponential RSS representation yields the smallest location estimation errors. By using this representation, the accuracy of all linear algorithms varies around 785 m. The R^2 score, which gives an indication of how close the actual target values are to the fitted regression line, increases accordingly from 0.70 using lineal RSS to 0.73 using exponential RSS.

Out of all evaluated algorithms, the SVR performs the worst. Due to the use of kernels, it took an extraordinary amount of time to compute the results. Furthermore, the R^2 score of 0.49 for exponential RSS indicates the model could not find a good fit, leading to mean location estimation errors above 1 km.

In contrast, the (weighted) kNN and Random Forest algorithms yield the best results regarding localization accuracy and R^2 score. The weighted variant of the kNN algorithm is slightly more accurate than the non-weighted version, leading to a mean location estimation error of 343 m using powed RSS and $k = 15$. Similarly, the Random Forest ensemble technique results in an accuracy of 340 m, using powed RSS as well. Figure 5.3 shows a box plot of localization errors for every algorithm using the powed RSS representation. While it takes 147 s to compute the results in the kNN algorithm, Random Forest takes advantage of its bagging technique, reducing the computation time to 53 s.

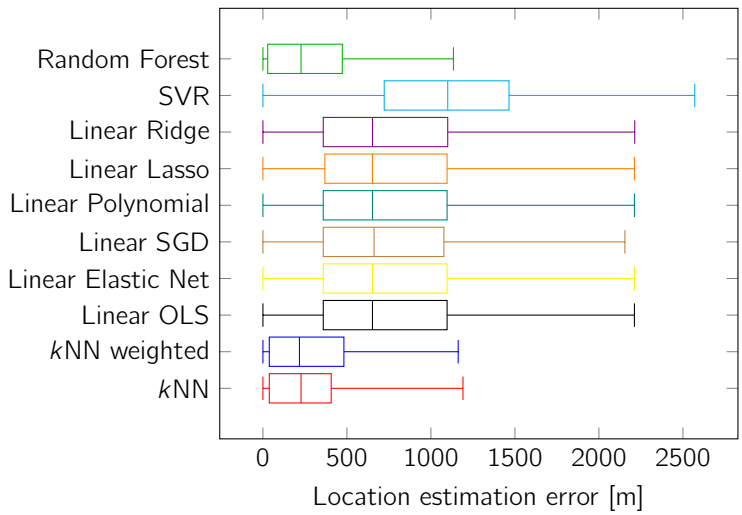


Figure 5.3: Location estimation errors for every regression algorithm using the powed RSS representation of the LoRaWAN data set

5.4.2 Range-based localization performance

As defined in Section 5.3.3, I evaluated three gateway selection strategies. The first strategy is to use all receiving LoRaWAN gateways. Since 58% of all messages were received by only one or two gateways, this strategy leads to mean localization errors above 1 km for all path loss models. By eliminating these messages from the experiments in the second gateway selection strategy, a mean location estimation error of 869 m was achieved for the rural Okumura-Hata model. In 95% of all messages, the estimation error is below 2217 m. However, some extreme outliers were observed. For example, a message transmitted in the city center of Antwerp was received by a gateway in Brussels, located 43 km farther away. A third strategy, which applied a geographical filter on participating gateways, successfully reduced the number of outliers and thus increased the overall localization performance. It should be highlighted that in the following, all range-based localization results are calculated

using the third gateway selection strategy. Thus, I use 55,259 out of 130,430 LoRaWAN messages with three or more receiving gateways located within a predefined bounding box to evaluate the performance of range-based localization.

Table 5.3 lists the mean location estimation errors for all path loss models and weight functions of the E-Min-Max algorithm. It shows that in 4 out of 8 evaluated path loss models, the third weight function, W_3 (5.16), achieves the lowest location estimation error. With an average localization error of 729 m, the IEEE 802.11ah macro model is the most accurate path loss model using this weight function. However, the non-weighted Min-Max algorithm leads to a smaller estimation error of 700 m and 722 m when using the suburban and rural Okumura-Hata models, respectively. The reason for this is discussed in Section 5.5. Figures 5.4 and 5.5 show the location estimation errors for every path loss model in a box plot, for the non-weighted (W_0) and weighted (W_3) Min-Max algorithm, respectively. In order to visualize the outliers in the results, the location estimation errors are shown on a map for the Min-Max (W_0) algorithm with the suburban Okumura-Hata model and the E-Min-Max (W_3) algorithm with the IEEE 802.11ah macro model in Figures 5.6 and 5.7, respectively.

The time to retrieve a location estimate should be taken into account when dealing with real-time and time-critical localization applications. Therefore, I compared the computation times for the different weight functions in the E-Min-Max algorithm. The time needed to calculate a position estimate for a single message on a computer with 16 GB RAM using a single core @ 1.8 GHz is shown in Figure 5.8. While it only takes 3 ms for the Min-Max algorithm with W_0 (5.13) to compute an estimate, the computation time for the weighted variants of the E-Min-Max algorithm varies from 43 ms to 83 ms. As can be observed in Equations (5.13) to (5.19) in Section 5.3, some weight functions require more complex operations, e.g., the square of the Manhattan distance in W_6.

Table 5.3: Mean location estimation errors for every combination of path loss model and weight function of the E-Min-Max algorithm. The optimal weight function for each path loss model is indicated in bold.

Path loss model	W_0	W_1	W_2	W_3	W_4	W_5	W_6
Urban Okumura-Hata	776	746	772	**733**	758	742	742
Suburban Okumura-Hata	**700**	800	938	731	786	788	733
Rural Okumura-Hata	**722**	1174	1687	806	1105	1021	821
Urban COST-231 Hata	789	743	758	**742**	756	751	752
Suburban COST-231 Hata	773	747	776	**731**	758	740	739
IEEE 802.11ah macro	763	757	804	**729**	765	743	736
IEEE 802.11ah pico	825	741	**733**	746	747	743	750
COST-231 Walfisch-Ikegami	896	790	**761**	789	**761**	766	766

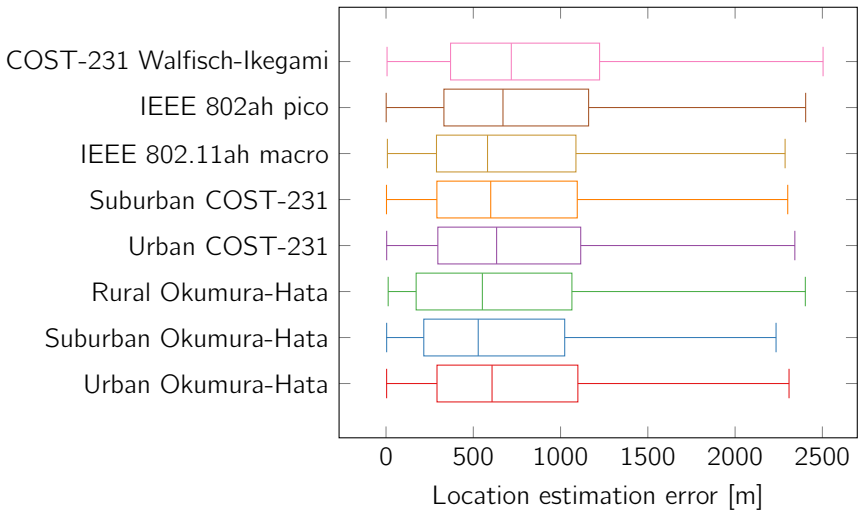


Figure 5.4: Location estimation errors of the Min-Max algorithm (without weights, i.e., W_0) for all path loss models.

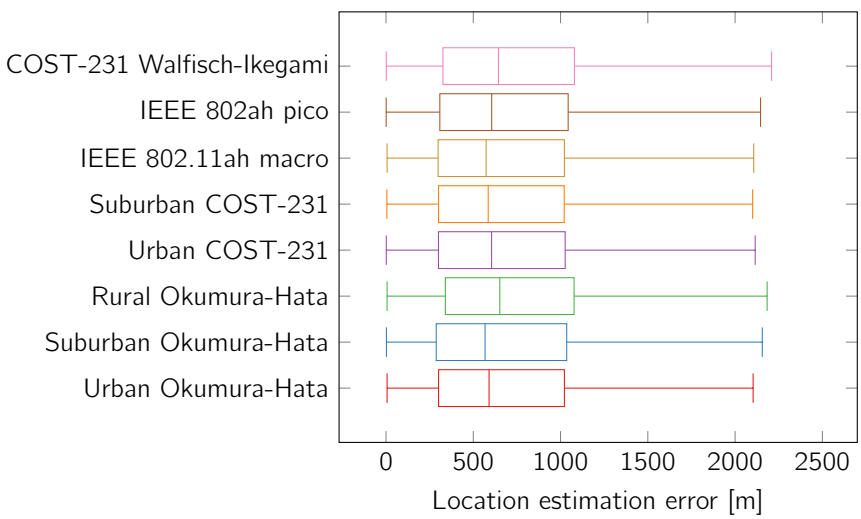


Figure 5.5: Location estimation errors of the E-Min-Max algorithm with weight function W_3 for all path loss models.

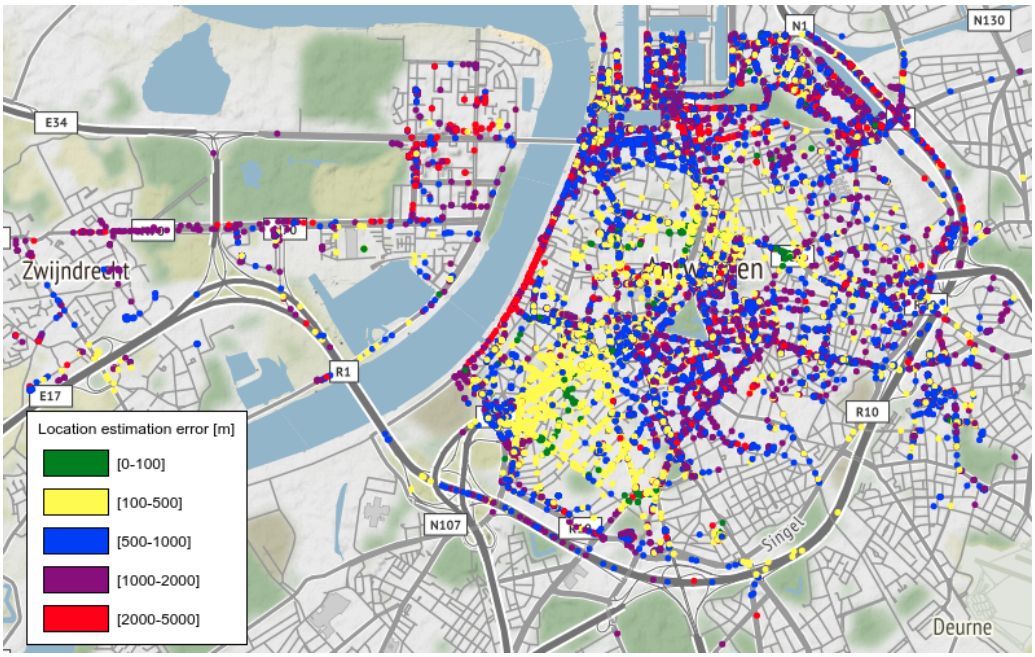


Figure 5.6: Location estimation errors for the Min-Max algorithm (without weights, i.e., W_0) and suburban Okumura-Hata path loss model, plotted on the GPS coordinate of every message in the data set.

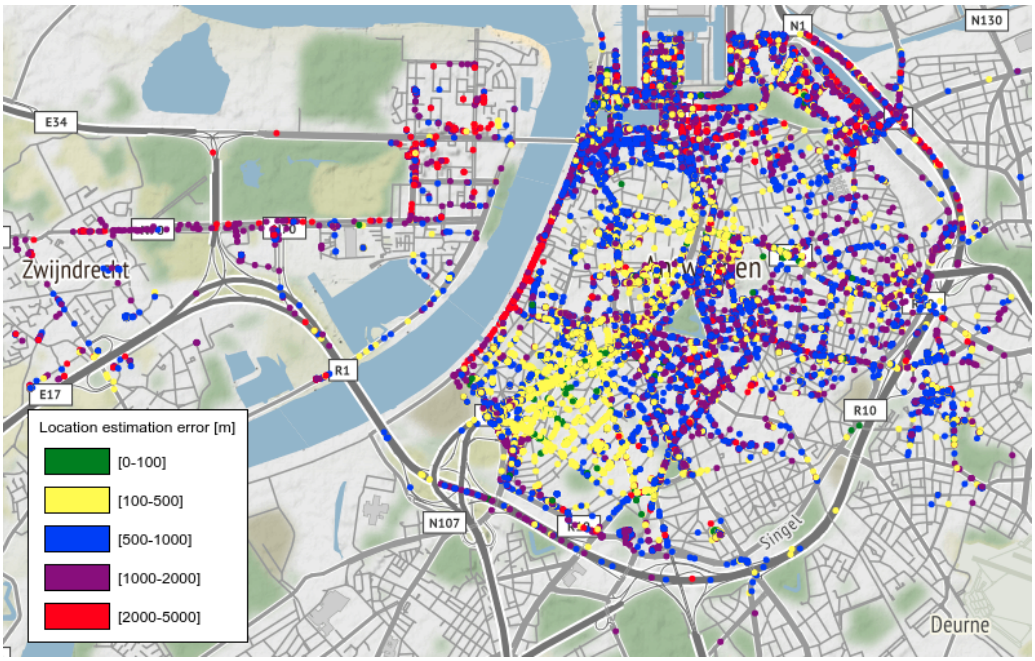


Figure 5.7: Location estimation errors for the E-Min-Max algorithm with weight function W_3 and the IEEE802.11ah macro model, plotted on the GPS coordinate of every message in the data set.

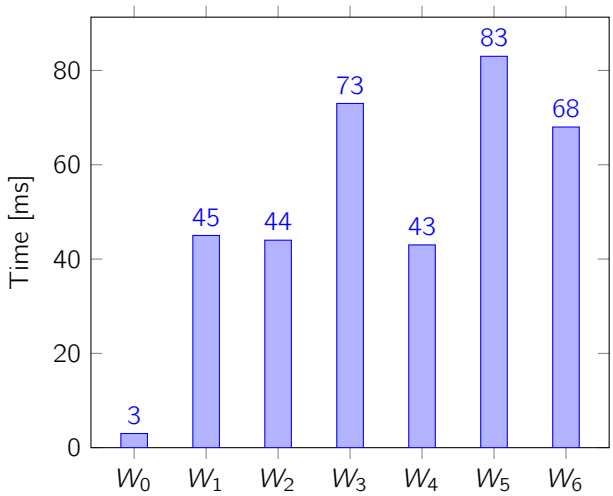


Figure 5.8: Average time needed to compute a single location estimate using different weight functions of the E-Min-Max algorithm, based on all 55,259 messages in the data set.

5.5 Discussion

Balancing location accuracy and computational performance is important when designing a localization application in an IoT context. In some applications, location accuracy is the most important aspect, e.g., the localization of a smartphone in a specific room. Outdoor applications with a need for high localization accuracy should use GNSS-based solutions. In contrast, other IoT applications, such as large-scale asset tracking, only require a rough location estimate and focus more on extreme battery lifetime. Such applications are realized by exploiting the advantages of LPWAN, such as low power consumption, long-range communication and indoor and outdoor coverage. The work in this chapter provides a benchmark for established RSS-based localization algorithms in LoRaWAN, for both fingerprint-based and range-based approaches.

5.5.1 Fingerprint-based localization

In total, ten pattern-matching ML algorithms are evaluated on a data set with 130,430 samples. Given the GPS coordinates as ground truth reference data, the objective was to assess the ability of each regression algorithm to accurately locate a mobile transmitter, using the RSS values to multiple gateways. Moreover, the computational performance of each algorithm was studied. In order to optimize the input data for the regression algorithms, I changed the representation of the RSS data, performed scaling and a PCA analysis on the data set as preprocessing steps. Afterwards, I evaluated every algorithm in terms of location estimation error, total computation time and R^2 score.

The optimal representation of signal strength values is determined for every algorithm being evaluated. Except for SVR, all algorithms benefit from transforming the RSS values to an exponential or so-called powed representation. These results are expected, given the

logarithmic nature of signal strength values. Thus, changing these RSS representations often leads to an increased localization performance.

The use of kernels in SVR is time consuming. Since the fit time complexity of SVR is more than quadratic with the amount of samples, the algorithm is unable to scale to data sets with more than a couple of 10,000 samples, resulting in significantly higher computation times. Moreover, taking into account the mean localization error of over 1 km, SVR is not a good choice to evaluate with our large data set.

Given their simplicity, the six linear regression algorithms all yield estimation errors around 800 m. However, out of all fingerprint-based algorithms, this category of algorithms computes a location estimate in the least amount of time. This can be important for fingerprint-based applications where a location estimate should be generated or updated in near real-time.

Finally, the weighted *k*NN and Random Forest algorithms achieve the highest localization accuracy. In both algorithms, a mean location estimation error of around 340 m is achieved when the RSS values are transformed to the powed representation. The Random Forest ensemble technique successfully avoids overfitting by averaging the predictions of multiple decision trees, while still being able to compute the results faster than the *k*NN algorithm, due to the bagging technique. Therefore, out of all ML algorithms, the Random Forest algorithm provides the best overall localization performance.

5.5.2 Range-based localization

Similar to the fingerprint-based experiments, I evaluated both the accuracy and computational performance of a range-based localization approach. I defined three gateway selection strategies, adapted the weight functions of the established E-Min-Max algorithm and evaluated the use of several outdoor path loss models.

The selection and number of participating gateways has a significant impact on the resulting localization accuracy. Therefore, three gateway selection strategies were developed. In a first strategy, all receiving gateways in the LoRaWAN network were used. However, in 58% of all 130,430 messages, only one or two gateways received the message. Regardless of the path loss model used, this consequently lead to location estimation errors above 1 km. With less than three receiving gateways, the RSS range-based algorithm boils down to an RSS proximity algorithm. The results of this strategy are in line with the RSS proximity results presented in [79]. With the aim to increase the localization accuracy, I developed a second gateway selection strategy. In this strategy, I only consider messages with three or more receiving gateways. For the most accurate path loss model, this strategy leads to a mean location estimation error of 869 m. However, when examining the estimation errors of this approach, a significant number of outliers were detected, because some distant gateways were still be able to receive the LoRaWAN signals. In order to limit the number of these gateways which are far away from the area under consideration, I defined a bounding box for gateways in a third strategy. In total, 55,259 out of 130,430 messages of the public data set satisfied the requirement of being received by three or more LoRaWAN gateways located within the bounding box. This strategy leads to an overall increase in localization

accuracy and thus is used to evaluate the performance of the outdoor path loss models and the E-Min-Max algorithm.

The purpose of the extended version of the Min-Max algorithm is to provide a weight to every vertex of the resulting rectangle, indicating how close the mobile transmitter is located to a specific vertex. The weight functions depend on how well the distance from the transmitter to a gateway is estimated. When investigating the performance of the E-Min-Max algorithm using an urban path loss model, shorter distances are estimated more accurately, when compared to long distances to gateways that are located farther away. Therefore, for all urban path loss models, the weighted E-Min-Max algorithm performs better than the non-weighted Min-Max algorithm. However, when using more rural path loss models, more free space path loss is considered during the translation from RSS values to distances and less building interference and multipath propagation effects are taken into account. Therefore, if a rural path loss model is used in the E-Min-Max algorithm, the distance estimates from a mobile device to nearby gateways are less accurate, whereas distances to gateways located farther away are estimated more accurately. Consequently, the resulting weights shift the original location estimate from the Min-Max algorithm towards the wrong vertices, leading to a decrease in localization accuracy.

The suburban and rural variants of the Okumura-Hata model exhibit the behavior mentioned above, as can be observed in Table 5.3. Therefore, these models might seem to outperform all other path loss models when using the non-weighted (W_0) Min-Max algorithm. For the suburban Okumura-Hata model, the mean and median location errors are 700 m and 528 m, respectively. However, when using different weight functions, the location estimation errors increase significantly for the suburban and rural Okumura-Hata models. Thus, it is of utmost importance to find the optimal combination of path loss model and weight function for the E-Min-Max algorithm.

In contrast to the findings in [117, 155], weight functions W_5 and W_6 of the E-Min-Max algorithm do not result in higher localization accuracy, despite the additional computational complexity.

In the box plots of Figures 5.4 and 5.5, it becomes clear that there is no path loss model that outperforms the other ones significantly. Therefore, I investigated the spatial distribution of the location estimation errors for the best performing combinations of path loss models and weight functions of the Min-Max and E-Min-Max algorithms. As can be observed in Figures 5.6 and 5.7, most location estimation errors below a threshold of 500 m originate from messages sent from within the city center of Antwerp, where the building density is the highest. Furthermore, a significant number of outliers up to 5 km can be observed from messages collected near water and more open areas. Due to this heterogeneity of the environment, it is a challenging task to choose a path loss model. In order to avoid the use of a single existing path loss model in a large-scale environment, a custom path loss model can be created for the application environment. However, this eliminates the advantage of a fast deployment in a range-based localization approach.

For the COST-231 Walfisch-Ikegami model, W_2 and W_4 are the best performing weight functions in terms of localization accuracy. This path loss model yields the most varying location estimation errors because model parameters such as street width and inclination angle vary from one location to another. Given the heterogeneous environment in our setup,

the COST-231 Walfisch-Ikegami model is not feasible to evaluate the average localization accuracy of a large-scale environment.

In four out of eight path loss models, weight function W_3 outperforms the other weight functions in terms of localization accuracy. However, the computational complexity of each location estimation algorithm should be taken into account as well. Based on a single LoRaWAN uplink message, a location estimate can be calculated within 3 ms when using the non-weighted Min-Max algorithm. Adding weights to this algorithm can increase location accuracy, but requires more time to compute the increased number of mathematical expressions, depending on the weight function. For example, it takes 73 ms to compute a location estimate when using weight function W_3.

In [117, 155], weight functions with the Manhattan distance (i.e., W_3, W_5 and W_6) were originally introduced to reduce the computational complexity of calculating the Euclidean distance in a two-dimensional space. However, both the Euclidean and Manhattan distances are converted to calculate the distance from point A to B on the surface of the Earth, modeled as a WGS-84 ellipsoid. To this end, the Euclidean distance can be replaced by the Karney distance, while the Manhattan distance needs to be calculated by computing the Karney distance twice, i.e., one time on the latitude line from point A and another time on the longitude line to point B. As a result, the Manhattan distance on a WGS-84 ellipsoid actually requires more computation time than the Karney distance. Therefore, weight functions W_3, W_5 and W_6 require more time to compute an estimate, as shown in Figure 5.8.

5.5.3 Comparing fingerprinting to ranging

Since I developed a range-based approach to locate a mobile device in a city-scale environment, as well as several fingerprint-based algorithms using the same data set, I am now able to compare these two RSS-based localization approaches to each other.

In the range-based approach, the E-Min-Max algorithm has been investigated using different path loss models and weight functions. The significant number of outliers due to the heterogeneous environment, fading and shadowing effects lead to a mean location estimation error of 700 m. In contrast, a fingerprint-based approach includes the effects on the signal strength in the training phase, leading to 50% more accurate results. For example, the Random Forest ensemble technique yields a mean localization error of 340 m.

When implementing a fingerprint-based approach in a localization application, several considerations should be taken into account. Building a large outdoor LPWAN training database takes a considerable amount of time, effort and cost. In the case of unlicensed LPWAN, messages need to be collected with respect to the uplink duty cycle regulations. Ideally, the messages in a fingerprinting database should be uniformly spread across the application environment. All of this leads to an increased deployment time, whereas in a range-based approach, the localization system can be deployed as soon as an appropriate path loss model has been chosen. Despite the maintenance cost, crowd-sourced fingerprinting initiatives such as The Things Network (TTN) assist in reducing the cost and time to build a large outdoor LPWAN database [156].

Another limitation of fingerprinting is the fact that mobile devices need to be localized in a predefined area, because training data is required for the pattern matching technique. Contrarily, a range-based localization system can be deployed anywhere, as long as the area is covered by at least three gateways. Unlike in a fingerprinting approach, the sometimes confidential gateway locations need to be known in a range-based approach. Finally, both localization approaches can be implemented on the server-side to preserve the minimal energy consumption of LPWAN communication. In Table 5.4, the benefits and limitations of both range-based and fingerprint-based localization approaches are summarized.

Table 5.4: Overview of the evaluated RSS-based localization algorithms, comparing the characteristics of both range-based and fingerprint-based localization.

	Range-based localization	**Fingerprinting-based localization**
Evaluated algorithms	Min-Max, E-Min-Max	ML algorithms: Linear regression, SVR, kNN, Random Forest
Accuracy (mean error)	700 m	340 m
Deployment time	Low	High
Cost to deploy in new environment	Medium	High
Maintenance cost	Low	High
Size of environment	Unlimited	Limited to training area
Number of RX gateways required	3 or more	1 or more
Gateway locations required	Yes	No
Power consumption	Low (unaffected)	Low (unaffected)

5.6 Conclusion

In light of the extreme battery lifetime requirements in some IoT use cases, GNSS-based solutions are not feasible. Therefore, I used an energy-efficient and long-range LPWAN network to estimate the location of a mobile device. Since simulations can lead to over-optimistic results, I chose to conduct this research in a real-world city-scale environment using a publicly available data set. I evaluated two RSS-based approaches in terms of accuracy and computational performance: fingerprint-based and range-based localization.

Outdoor RSS-based fingerprinting localization can be challenging, given the dynamic environment of a city. It takes time, effort and cost to create and maintain a training database. Nonetheless, a fingerprint-based approach results in a significantly higher localization accuracy when compared to range-based approaches. The Random Forest ML algorithm leads to an average location estimation error of 340 m and the shortest computation time due to the bagging technique. Interestingly, the location accuracy is equivalent to the accuracy of the Sigfox fingerprinting experiments conducted in the same environment in Chapter 4.

In a range-based approach, distance estimations obtained from RSS measurements are strongly influenced by various wireless propagation effects, such as shadowing and fading due to multipath propagation, leading to a mean accuracy of around 700 m. However, in opposition to fingerprint-based solutions, range-based localization algorithms can be implemented in applications where a fast deployment or near real-time location updates are of utmost importance. In this approach, a path loss model and location estimation algorithm should be chosen carefully, keeping the building density and type of environment in mind. However, given our dynamic test environment with both urban and rural areas, the E-Min-Max algorithm in combination with any path loss model yields a significant number of outliers. To reduce the outliers, several gateway selection strategies have been evaluated, successfully increasing the localization accuracy.

Given the comparative study on RSS-based LoRaWAN localization approaches in this work, it should be straightforward to choose the most suitable approach to design an IoT application, considering the trade-off between localization accuracy, deployment cost and computational performance.

Chapter 6

Opportunities with LoRa 2.4 GHz

RECENTLY, Semtech has released a range of LoRa transceivers which operate at the globally available 2.4 GHz frequency band, on top of the existing sub-GHz, km-range offer [50]. As stated in Chapter 2, this enables hardware manufacturers to design region-independent LoRa modules. Moreover, the 2.4 GHz transceivers promise an ultra-long communication range while withstanding heavy interference in this widely used band.

LoRa at 2.4 GHz has the potential to become an interesting solution for a variety of applications that involve indoor localization, such as warehouse management, but also for outdoor localization applications, such as construction site monitoring and livestock tracking. Moreover, the adoption of this technology can add flexibility to applications that require consistent asset tracking in both indoor and harsh outdoor environments, e.g., smart ports.

The goal of this chapter is to study the inverse relationship between the maximum communication range and the corresponding data rate of LoRa in the 2.4 GHz ISM band. Therefore, I first provide a mathematical description of the physical layer of LoRa in the 2.4 GHz band. Secondly, I investigate the maximum communication range of this technology in three different scenarios. Free space, indoor and urban path loss models are used to simulate the propagation of the 2.4 GHz LoRa modulated signal at different spreading factors and bandwidths. Additionally, I investigate the corresponding data rates. The results show that due to the configurable bandwidth and lower data rates, LoRa outperforms other technologies in the 2.4 GHz band in terms of communication range. In addition, both communication and localization applications deployed in private LoRa networks can benefit from the increased bandwidth and localization accuracy of this system when compared to public sub-GHz networks.

The remainder of this chapter is structured as follows. A mathematical background of LoRa at 2.4 GHz is provided in Section 6.1. Subsequently, three path loss models are presented in Section 6.2 in order to estimate the maximum communication range and corresponding data rate in a free space and in indoor and urban environments. The results of these estimations are shown in Section 6.3 and compared to other technologies operating in the 2.4 GHz band in Section 6.4, which also discusses the impact on the application potential of LoRa at 2.4 GHz. Finally, general conclusions are drawn in Section 6.5.

This chapter is based on:

Thomas Janssen, Noori BniLam, Michiel Aernouts, Rafael Berkvens, & Maarten Weyn. *LoRa 2.4 GHz Communication Link and Range*. Sensors, vol. 20, no. 16, p. 44366, Aug. 2020, https://doi.org/10.3390/s20164366.

6.1 LoRa in the 2.4 GHz Band

The physical layer of LoRa is a proprietary and closed source. Therefore, there are no official references or protocol specifications for the transmitted Radio Frequency (RF) signal [157]. Accordingly, several research groups have been working to provide an understanding of the LoRa modulation scheme in the sub-GHz frequency band. Vangelista [158], for instance, has provided a mathematical model, called Frequency Shift Chirp Modulation (FSCM), that describes the LoRa modulation process. The same model has been adopted by Bernier et al. [48]. On the other hand, Knight [157], Robyns et al. [159] and Ghanaatian et al. [160] have provided a model called Chirp Spread Spectrum (CSS) modulation based on the reverse engineering of LoRa's physical layer. Even though the formulation of the LoRa modulation scheme in the literature has been provided for the sub-GHz frequency band, the basic response of the modulation is expected to be the same for LoRa signals in the 2.4 GHz frequency band. Therefore, in this section, I modify the available physical layer models of sub-GHz LoRa to make them suitable for use in the 2.4 GHz frequency band.

Assume $x_s(k)$ is the transmitted LoRa sample; then, the received sampled signal $x_r(k)$ with index k can be expressed as

$$x_r(k) = a_r x_s(k-\tau) e^{i2\pi\Delta f k} + \omega(k), \tag{6.1}$$

where $a_r < 1$ is the received signal amplitude, τ is the time delay of the sample $x_s(k)$ at the receiver, Δf is the frequency offset between the transmitter and the receiver, and $\omega(k)$ is the identically independently distributed (i.i.d.) complex-valued Gaussian noise with zero-mean and variance σ^2; i.e., $CN(0,\sigma^2)$. The time and frequency synchronization are beyond the scope of this work. Therefore, in the following, I will consider a simplified version of Eq. (6.1), shown in Eq. (6.2).

$$x_r(k) = a_r x_s(k) + \omega(k) \tag{6.2}$$

The LoRa standard linear upchirp—also called a base chirp—can be expressed as [159, 160]

$$x_s(k) = e^{i2\pi\left(\frac{\mathrm{BW}}{2K}k^2 + f_\circ k\right)}, \tag{6.3}$$

where BW is the operational bandwidth of the LoRa signal in the 2.4 GHz frequency band (as shown in Table 6.1) and $K = 2^{\mathrm{SF}}/\mathrm{BW}$ is the symbol duration, with SF representing the spreading factor (also shown in Table 6.1). Finally, $f_\circ$ is the initial frequency, which can be expressed as

$$f_\circ = s\frac{\mathrm{BW}}{2^{\mathrm{SF}}}, \tag{6.4}$$

where $s \in \{0, 1 \ldots 2^{\mathrm{SF}}\}$ is the transmitted data symbol. Setting $s = 0$ results in an upchirp, in which the frequency continuously increases during the symbol duration K.
One can also present Eq. (6.3) as

$$x_s(k) = W_K^{\frac{\mathrm{BW}}{2}k^2 + Kf_\circ k}, \; W_K = e^{i2\pi/K}. \tag{6.5}$$

Table 6.1: Parameters used for path loss modeling.

Model Parameter	**Symbol**	**Value**	**Unit**
Frequency	f	2.4	GHz
Spreading factor	SF	5–12	-
Bandwidth	BW	203/406/812/1625	kHz
Code rate	R_C	4/5	-
Transmission power	P_{TX}	12.5	dBm
Transmitter antenna gain	G_{TX}	2	dBi
Transmitter cable losses	L_{TX}	2	dB
Fading margin	L_m	0	dB
Receiver antenna gain	G_{RX}	2	dBi
Receiver cable losses	L_{RX}	2	dB
Base station height	h_b	20	m
Mobile station height	h_m	2	m

The model in Eq. (6.5) is the linearly cyclically shifted version of a base Zadoff–Chu (ZC) sequence [161]. The ZC sequence possesses a unique autocorrelation property, in which the periodic autocorrelation is orthogonal (i.e., equal to zero) for all shifted replicas [162]. Therefore, the LoRa communication protocol uses this unique property to impose a random multiple access technique. Accordingly, an efficient utilization of the unlicensed spectrum can be obtained. The correlation between the received signal and the base chirp leads to

$$\begin{aligned} z(k) &= \frac{1}{K}\sum_{p=0}^{K-1} x(k+p)x_s^*(k)_{\text{mod}\,K} \\ &= \frac{1}{K}\sum_{p=0}^{K-1} \left(a_r x_s(k+p)+\omega(k+p)\right)x_s^*(k), \\ &= \begin{cases} a_r E_s + \nu_\omega & \text{for}\, p=0 \\ \nu_\omega & \text{for}\, p\neq 0 \end{cases} \end{aligned} \tag{6.6}$$

where E_s is the energy of the symbol x_s. Furthermore, ν_ω is the correlation between complex noise and the base chirp, which can be expressed as

$$\nu_\omega = \frac{1}{K}\sum_{p=0}^{K-1} x_s^*(k)\omega(k+p), \tag{6.7}$$

in which $\nu_\omega \sim CN(0,\sigma^2/K)$.

Figure 6.1 presents two received LoRa signals that constitute eight preamble (upchirp) symbols at 2.4 GHz with a bandwidth equal to 812 kHz. The short signal was transmitted at a SF of 9, while the longer signal was transmitted at a SF of 10. Figure 6.1a,b shows the combined received signals in the time domain and in the spectrogram (i.e., time and frequency) domain, respectively. Figure 6.1c,d shows the cross-correlation functions (6.6) when the received signals have been cross-correlated with base chirps of the SF equal to 10 and 9, respectively. It is clear that the two signals can be distinguished correctly, even though they interfere with each other. The unique orthogonality property of the ZC sequence

allows the LoRa communication system to provide a multiple access technique in the 2.4 GHz frequency band.

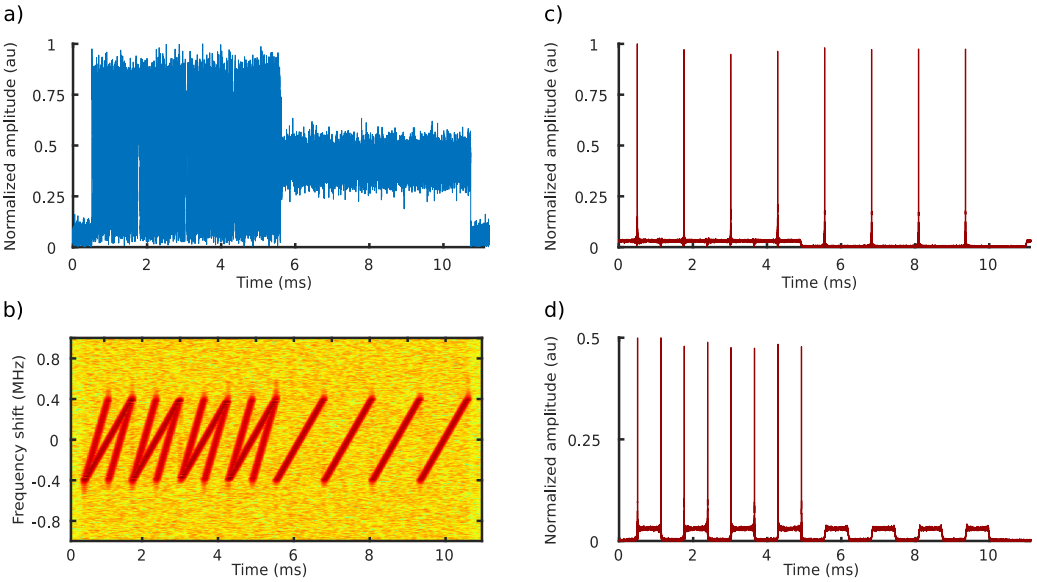


*Figure 6.1: Two received Long Range (LoRa) signals constitute eight preamble upchirp symbols at 2.4 GHz with a bandwidth equal to 812 kHz. The SF of the short signal, which ended after approximately 5 ms, is equal to 9, while the SF of the long-duration signal is equal to 10. Figures (***a***) and (***b***) represent the combined received signals in the time domain and in the spectrogram (i.e., time and frequency) domain, respectively. Figures (***c***) and (***d***) are the cross-correlation functions* (6.6) *when the received signals have been cross-correlated with base chirps of the SFequal to 10 and 9, respectively.*

6.2 Path Loss Modeling

In order to obtain the maximum communication range of a LoRa signal at 2.4 GHz (further denoted as d), one needs to find the maximum link budget for which the signal can be received properly; i.e., at the receiver sensitivity P_{RX}. This receiver sensitivity depends on two key factors: the used SF and bandwidth (BW). While the SF can range from 5 to 12, the possible bandwidths of LoRa at 2.4 GHz are 203, 406, 812 and 1625 kHz. Furthermore, the combination of a certain SF and BW results in a certain data rate, along with the receiver sensitivity, as shown in Table 6.2. The raw data rate R_b, expressed in kbit/s, can be calculated as

$$R_b = \frac{SF * BW}{2^{SF}}, \tag{6.8}$$

with SF and BW as defined in Table 6.1. As an example, a LoRa signal transmitted with an SF of 8 and a BW of 406 kHz results in a receiver sensitivity of −116 dBm and a data rate of 12.69 kbit/s. The receiver sensitivities and data rates used in this work originate from the datasheet of the Semtech SX1280 LoRa module [51].

Table 6.2: Receiver sensitivities (P_{RX} in dBm) and corresponding data rates (R_D in kbit/s) of the SX1280 LoRa module for every combination of spreading factor (SF) and bandwidth (BW).

	BW (kHz)							
	203		**406**		**812**		**1625**	
SF	P_{RX}	R_D	P_{RX}	R_D	P_{RX}	R_D	P_{RX}	R_D
5	−109	31.72	−107	63.44	−105	126.88	−99	253.91
6	−111	19.03	−110	38.06	−108	76.13	−103	152.34
7	−115	11.1	−113	22.2	−112	44.41	−106	88.87
8	−118	6.34	−116	12.69	−115	25.38	−109	50.78
9	−121	3.57	−119	7.14	−117	14.27	−111	28.56
10	−124	1.98	−122	3.96	−120	7.93	−114	15.87
11	−127	1.09	−125	2.18	−123	4.36	−117	8.73
12	−130	0.595	−128	1.19	−126	2.38	−120	4.76

The total link budget of a wireless communication signal propagating from the transmitter to receiver can be represented as [163]

$$P_{RX} = P_{TX} + G_{TX} - L_{TX} - L_p(d) + G_{RX} - L_{RX}, \tag{6.9}$$

where P_{RX} is the received power in dBm, P_{TX} is the transmission power in dBm, G_{TX} is the antenna gain at the transmitter in dBi, L_{TX} is the cable loss in dB, $L_p(d)$ is the path loss in dB in terms of the function from the distance d, G_{RX} is the antenna gain at the receiver in dBi and L_{RX} is the cable loss at the receiver in dB. Except for the path loss, all these parameters are set to typical values which are commonly used when simulating a wireless communication link between two dipole antennas [163]. The values of these parameters are summarized in Table 6.1.

The path loss $L_p(d)$ is defined as the propagation loss caused by the signal traveling from the transmitter to receiver over a distance d. The goal in this research is to maximize d while still being able to successfully receive the LoRa-modulated signal at the receiver. For the sake of simplicity, the simulated LoRa signal contains eight preamble symbols and no payload bytes.

Depending on the environment, the path loss should be modeled differently. Therefore, in the next three subsections, I discuss indoor, outdoor (urban) and free space path loss models to translate the propagation loss into a distance between the transmitter and receiver. The parameters required by these models are also summarized in Table 6.1. No fade margin is taken into account; this is further discussed in Section 6.4.

6.2.1 Free Space Environment

The first scenario can be described as a free space environment in which there is a LoS (i.e., the primary Fresnel zone must be at least 60% clear.) between the TX and RX locations. In this case, I can use the widely used Free Space Path Loss (FSPL) model to evaluate the maximum communication range. This model calculates the loss between two isotropic

radiators in free space, without considering any obstacles, reflections or interference. The model solely relies on the frequency and distance between the transmitter and receiver to calculate the path loss:

$$L_{p,LoS}(d) = 32.44 + 20\log_{10}(f) + 20\log_{10}(d), \tag{6.10}$$

where f is in MHz and d is in km. The combination of a given SF and BW yields a certain sensitivity P_{RX}. Consequently, given the maximum path loss obtained from Eq. (6.9), one can calculate the maximum distance as

$$d = 10^{(L_{p,LoS}(d) - 32.44 - 20\log_{10}(f))/20}. \tag{6.11}$$

6.2.2 Indoor Environment

In the second scenario, I evaluate the maximum communication range in an indoor environment. To this end, I slightly adapt a heuristic algorithm that was developed based on real measurements in an office-like environment [164]. The path loss model is based on the Indoor Dominant Path (IDP) model, which focuses on the dominant path between the Transmit (TX) and Receive (RX) location. In general, the total path loss is the sum of the distance loss, accumulated wall loss and interaction loss and can be calculated as

$$L_{p,in}(d) = L_{p_0}(d_0) + 10 * n * \log_{10}(\frac{d}{d_0}) + \sum_i L_{W_i} + \sum_j L_{B_j}, \tag{6.12}$$

where $L_{p_0}(d_0)$ represents the path loss at a distance d_0 and n is the path loss exponent. The accumulated wall loss is the sum of losses L_{W_i} caused by each wall along the dominant path. Finally, the interaction loss is the sum of losses L_{B_j} caused by all directional changes of the propagating signal.

Given the semi-empirical nature of this path loss model, some parameters need to be set to commonly used values in order to provide a generally applicable model that can predict ranges in other indoor environments. Therefore, $L_{p_0}(d_0)$ is set to 40 dB at a distance $d_0 = 1$ m, as suggested in [164]. The path loss exponent is set to $n = 5$, which is generally used for obstructed paths inside buildings [165]. For the accumulated wall and interaction loss, values of 6 and 3 dB have been taken into account respectively, as found specifically for the office-like environment in [164]. Consequently, the path loss model can be simplified to

$$L_{p,in}(d) = 40 + 5 * 10 * \log_{10}(d) + 6 + 3. \tag{6.13}$$

Thus, the range can be empirically estimated based on the path loss:

$$d = 10^{(L_{p,in}(d) - 49)/50}. \tag{6.14}$$

6.2.3 Urban Environment

An urban path loss model is used in the third scenario to evaluate the range of LoRa at 2.4 GHz in an outdoor city-scale environment. The Okumura-Hata Urban Path Loss model

is an empirical model that is often used in sub-GHz wireless communication systems, as shown in Chapter 5. While the COST-231 urban model extended its use up to 2 GHz, the Electronic Communication Committee (ECC) modified the original Okumura-Hata model to work with frequencies up to (and beyond) 3 GHz in the ECC-33 model [166, 167]. Therefore, the ECC-33 model is suitable to evaluate the maximum communication range of LoRa at 2.4 GHz in an urban environment. The path loss equation for this model is given by

$$L_{p,urban}(d) = A_{fs} + A_{bm} + G_b + G_m, \tag{6.15}$$

where A_{fs} is the free space attenuation, A_{bm} is the basic median path loss, G_b is the base station height gain factor and G_r is the receiver height gain factor, which can be calculated as

$$\begin{aligned} A_{fs} &= 92.4 + 20\log_{10}(d) + 20\log_{10}(f), && (6.16)\\ A_{bm} &= 20.41 + 9.83\log_{10}(d) + 7.894 log_{10}(f) \\ &\quad + 9.56[\log_{10}(f)]^2, && (6.17)\\ G_b &= \log_{10}(\frac{h_b}{200})\{13.958 + 5.8[\log_{10}(d)]^2\}, \text{ and} && (6.18)\\ G_m &= [42.57 + 13.7\log_{10} f][\log_{10}(h_m) - 0.585] && (6.19) \end{aligned}$$

for medium-sized urban environments. Given the complexity of this set of equations, I extract the maximum range by iterating over values of d from 1 m to 10 km and solving the optimization problem given a certain path loss $L_{p,urban}(d)$.

6.3 Range Versus Data Rate: Results

Figures 6.2 to 6.4 show the maximum communication range and corresponding data rate at each combination of SF and bandwidth for the free space and indoor and urban environments, respectively. In all cases, the highest possible data rate decreases in a logarithmic way when the communication range between the transmitter and receiver increases.

Using the FSPL model, it is found that a 2.4 GHz LoRa signal can travel up to 133 km in free space and still be received properly. Obviously, this is only a theoretical range and cannot be realized in real-world environments.

The performance of a more realistic indoor path loss model has been visualized in Figure 6.3. When transmitting with a spreading factor of 12 and the lowest bandwidth (i.e., 203 kHz), the path loss equals 142.5 dB. Consequently, a maximum communication range of 74 m can be achieved. Furthermore, the highest possible data rate at that range becomes 0.595 kbit/s. At the other extreme, the highest achievable data rate of 253.91 kbit/s is possible at a range of up to 18 m.

Finally, the communication range of the urban ECC-33 path loss model varies from 3 m at the highest achievable data rate of 253.91 kbit/s to 443 m at the lowest possible data rate of 0.595 kbit/s.

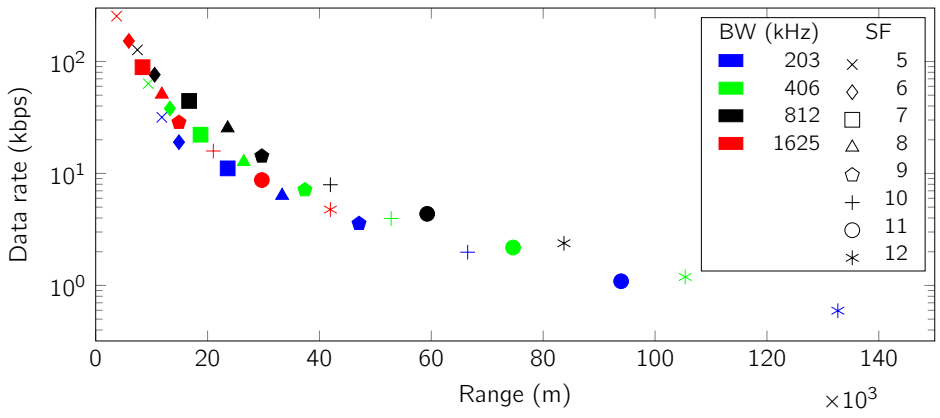


Figure 6.2: Communication range and data rate for every combination of spreading factor (SF) and bandwidth (BW) in a free space line of sight (LoS) environment.

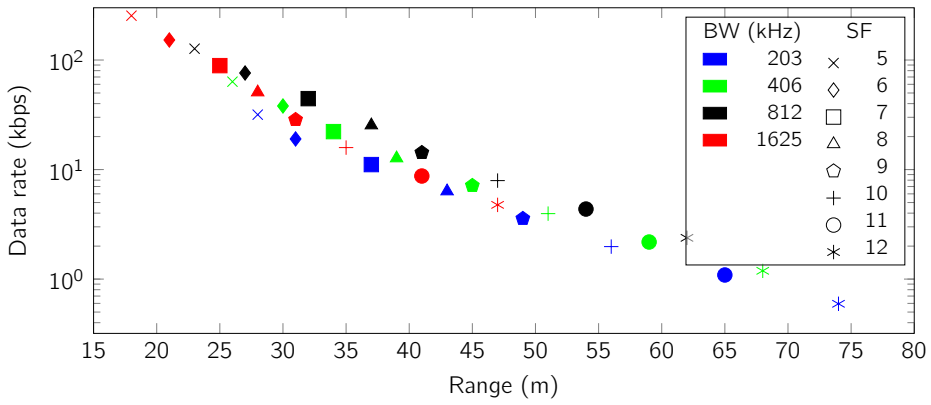


Figure 6.3: Communication range and data rate for every combination of spreading factor (SF) and bandwidth (BW) in an indoor environment.

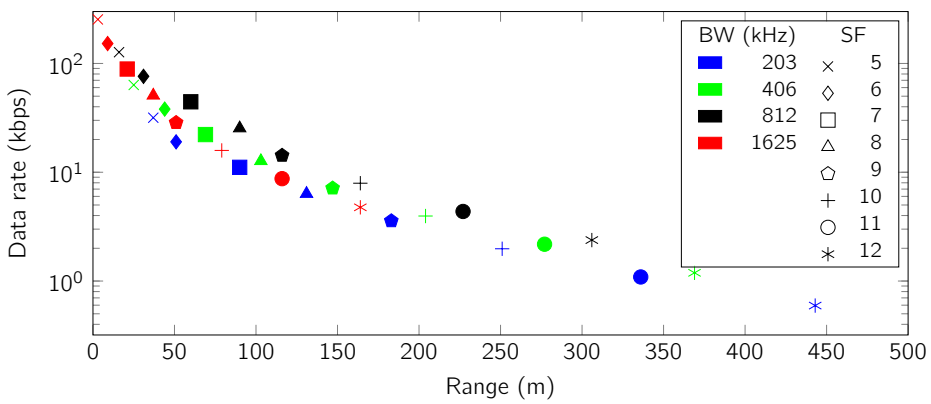


Figure 6.4: Communication range and data rate for every combination of spreading factor (SF) and bandwidth (BW) in an urban environment.

6.4 Discussion

I investigated the maximum communication range of LoRa in the 2.4 GHz band, which is defined as the maximum distance between a transmitter and receiver at which a LoRa-modulated message can be received properly. My investigations included three environments: LoS free space, NLoS indoor and urban outdoor. In all three scenarios, I maximized the range by reducing the bandwidth and increasing the spreading factor.

Based on the total link budget of the wireless communication system, including receiver sensitivity, antenna gains and cable losses, I was able to estimate the range of LoRa at 2.4 GHz. It is important to note that I did not include a fade margin in the link budget calculations. The fade margin can be defined as the level of received power in excess of that required for a specified minimum level of system performance. The reason for excluding this loss parameter in Eq. (6.9) is the high variability of fade margin in different scenarios. For instance, a 5 dB fade margin decreases the maximum urban range from 443 m to 277 m, while a 10 dB fade margin further decreases the range to 164 m. Thus, this should be taken into account when analyzing the results. Nonetheless, the most significant factor in a link budget is the path loss.

The free space line-of-sight scenario resulted in a theoretical maximum range of 133 km when transmitting at the highest SF and using the lowest bandwidth. In reality, the signal will always have to cope with obstacles, multipath propagation effects and interference with other signals. Therefore, these ranges will never be achieved in a real-world environment. Nevertheless, the results of the Free Space Path Loss model are useful as a benchmark as they enable comparison between different frequencies and technologies. For instance, the maximum range of LoRa at 868 MHz calculated with the FSPL model equals 921 km, which is almost seven times the range of LoRa at 2.4 GHz.

For the indoor range estimation of 2.4 GHz LoRa, I evaluated an indoor path loss model. In order to provide the highest possible accuracy, I adopted and slightly modified a model based on real-world measurements, taking into account both wall and interaction loss [164]. The estimated range varies from 18 m to 74 m, depending on the SF and BW. It should be noted that a path loss exponent of 5 was chosen, simulating an obstructed indoor environment. However, in an indoor LoS scenario, the range might therefore be increased.

Finally, the maximum communication range in an urban environment was found to be 443 m. This is similar to the results of the outdoor experiments with the SX1280 transceiver carried out by Wolf et al. [57]. They found that the ToF ranging feature failed for ranges over about 500 m. Thus, this validates my range estimations of the ECC-33 path loss model.

Besides the communication range, I investigated the data rates for all combinations of SFs and BWs and consequently associated this information with the highest achievable range. The highest possible data rate of LoRa at 2.4 GHz equals 253.91 kbit/s, which is almost seven times higher than the maximum data rate of LoRa at 868 MHz. This data rate can be achieved if the distance between the transmitter and receiver is not greater than 3739 m, 18 m and 3 m in a free space, indoor and urban environment, respectively.

Some significant differences in terms of range arise when comparing LoRa to other technologies operating in the 2.4 GHz band. As mentioned earlier, the range of the latest Bluetooth

standard equals 50 m and 165 m in an indoor and outdoor environment, respectively. Moreover, the maximum range of 2.4 GHz Wi-Fi networks typically varies around 100 m. Thus, the outdoor range of LoRa is almost three times larger than the outdoor range of BLE 5 and more than four times larger compared to typical IEEE 802.11 networks. This is mainly due to the lower bandwidth and data rates used in LoRa, as well as the robustness of the LoRa-modulated signal. These numbers clearly indicate the significant difference in intended applications between LoRa (such as long-range communication and localization) and Wi-Fi and Bluetooth (such as video and audio streaming).

Since LoRa modulation at 2.4 GHz has a higher bandwidth than sub-GHz LoRa modulation, the rising edge of a signal pulse can be determined more accurately. Therefore, I expect that time-based localization methods for this technology will result in lower estimation errors. However, the results in this chapter show that it is not possible to achieve the same long communication ranges as sub-GHz LoRa and other sub-GHz LPWANs. Therefore, more LoRa receivers have to be deployed to cover wide areas, which makes it a less feasible solution to build large public networks. On the other hand, 2.4 GHz LoRa is an interesting option for both communication and localization in privately deployed networks that are purposed for asset tracking and monitoring in large warehouses, construction sites, farms, etc.

6.5 Conclusion

With the expansion of LoRa to the globally available 2.4 GHz ISM band, hardware manufacturers are able to design a uniform LoRa module which functions independently of the region of deployment. However, as a consequence of moving to a higher frequency, the range of LoRa is reduced when compared to the sub-GHz range of several kilometers. In this chapter, I have provided a technology overview of LoRa operating in the 2.4 GHz band. By calculating the link budget of a system operating in this band, the range of a LoRa modulated signal is estimated in a free space and in indoor and outdoor scenarios. When compared to other technologies operating in the 2.4 GHz band, LoRa outperforms them in terms of communication range, due to the configurable SF and bandwidth. Thus, when configuring a LoRa channel at 2.4 GHz, a trade-off between range and data rate should be considered. Moreover, this trade-off leads to more flexible applications, such as the localization of assets in a private LoRaWAN network, which will be investigated in future work.

Chapter 7

RSS-based NB-IoT Localization

WITH hundreds of networks being deployed in a short amount of time, NB-IoT is stealing the LPWAN market shares from its competitors. The deep indoor coverage of the cellular IoT technology guarantees the energy-efficient communication link between a device and a nearby base station.

As discussed in Chapter 2, 3GPP Release 14 introduced OTDoA in the NB-IoT specification. However, this advanced positioning feature is barely implemented as it requires expensive and complex infrastructure upgrades, such as eNB synchronization. Hence, there is a need for alternative methods to locate NB-IoT devices.

To the best of the author's knowledge, no RSS-based outdoor localization experiments with NB-IoT have been conducted at the time of writing. When discussing the localization accuracy of a system, simulations are likely to lead to over-optimistic expectations of typical localization performance. Hence, instead of simulating the experiments, I face the challenges of working with real measurements of a public NB-IoT network in a real-world city-scale environment. In contrast to GNSS-based solutions, the position of a mobile device equipped with a small battery can be estimated in indoor and outdoor environments for several years. In this chapter, I investigate the accuracy of NB-IoT localization by evaluating three outdoor RSS-based algorithms: proximity, ranging and optimized fingerprinting. Additionally, I address the practical issues and limitations of dealing with real-world measurements.

During the measurement campaign, I discovered a mobility issue in NB-IoT. In contrast to other LPWAN and cellular technologies which use multiple gateways or cells to locate a device, only a single base station cell can be used for RSS-based localization in NB-IoT. When a UE transmits an uplink message with the request to get the signal strengths to nearby base stations, only the currently serving cell responds. As the number and placement of participating base stations significantly impacts the localization accuracy [62], it is highly preferred to receive a response from multiple nearby base stations. Hence, I evaluated the behavior of different NB-IoT chipsets, such as the Altair ALT1250, Fibocom MA510-GL, Mediatek MT2503, Quectel BC95, Sequans NB01Q, u-blox Sara N2 and NB-IoT chips from Huawei, Intel, and Qualcomm. While some of them should be able to receive a response from multiple neighboring cells, only the serving cell responded when connected to the NB-IoT networks of all Belgian providers. In an article from Rohde & Schwarz, the authors try to address this issue in the design of NB-IoT: "*To reduce the complexity and cost of NB-IoT*

chipsets, 3GPP has limited the mobility to idle mode only with cell reselection measurements. This means that neighbor cell measurements can only occur in idle mode, and once the NB-IoT device is in connected mode, it stops searching for neighbored cells." [168]. Therefore, I address this limitation in the current NB-IoT hardware and software by studying the mobility of the cellular-based 3GPP standard in a localization context. Experimental results show that the lack of handover support leads to increased cell reselection time and poor cell sector reliability, which in turn results in reduced localization performance. Fortunately, from 3GPP Release 16 onward, a UE may optionally report about strongest neighbors [36], and chipset manufacturers have started upgrading their firmware accordingly.

The proposed RSS-based methods, along with the mobility issue described above, will never reach GNSS-level location accuracy. Nonetheless, the resulting accuracies often satisfy most of the IoT use case requirements. LPWAN localization is further encouraged by the fact that in privacy sensitive applications, it is not desired to know the exact location (i.e., within a few meters) of a person or a package. For example in logistics, a rough estimate of where your package is located is usually sufficient. Meanwhile, the exact location of the driver is not exposed.

The contributions in this chapter are summarized as follows:

- I provide first-hand results of the outdoor localization accuracy of three RSS-based algorithms (proximity, ranging and fingerprinting) in a public NB-IoT network, based on the signal strength to a single cell.
- I optimize the fingerprinting algorithm to further increase the localization accuracy.
- I address the mobility issues of NB-IoT in a localization context. More specifically, I investigate the reason why only the serving cell can be used to locate a transmitter and I evaluate the cell sector reliability, as well as the cell reselection time.

The remainder of this chapter is structured as follows. Section 7.1 describes the outdoor localization experiments. The data sets used for both localization and cell reselection evaluation are described first. Second, three optimized RSS-based localization algorithms are proposed, namely proximity, range-based and fingerprint-based localization. The section ends with a description of the setup to evaluate the cell reselection in NB-IoT. The results in terms of localization accuracy and mobility performance are shown in Section 7.2. Finally, the main observations regarding localization accuracy and mobility of NB-IoT are discussed in Section 7.3 and summarized in Section 7.4.

This chapter is based on:

Thomas Janssen, Maarten Weym, & Rafael Berkvens, *A Primer on Real-world RSS-based Outdoor NB-IoT Localization*, in *2020 International Conference on Localization and GNSS (ICL-GNSS)*, pp. 1–6, Jun. 2020, https://doi.org/10.1109/ICL-GNSS49876.2020.9115578.

Thomas Janssen, Rafael Berkvens, & Maarten Weyn. *RSS-Based Localization and Mobility Evaluation Using a Single NB-IoT Cell*. Sensors, vol. 20, no. 21, p. 6172, Oct. 2020, https://doi.org/10.3390/s20216172.

7.1 RSS-based Localization Experiments

This section describes the steps taken during the real-world localization experiments. It starts with an explanation on how data is collected during the measurement campaign. Subsequently, three RSS-based methodologies to locate a mobile UE are presented. First, a simple proximity algorithm is presented. Second, I modify a range-based algorithm, which uses a path loss model in combination with the sector information of the received eNB. Third, I create a fingerprinting database and optimize the hyperparameters of a kNN algorithm. Finally, I describe the setup to evaluate the mobility of NB-IoT.

7.1.1 Measurement Campaign

Two sets of data are collected for this research. The first data set contains NB-IoT messages collected during an outdoor measurement campaign in the city of Antwerp, Belgium. This data set is used to evaluate the performance of RSS-based localization algorithms in an urban environment. The second data set is smaller and is used to evaluate the mobility of NB-IoT.

All NB-IoT messages are collected by sending uplink messages from a u-blox Sara N211 System on Chip (SoC) over a Release 13 NB-IoT network to a backend. Each message contains the Cell ID of and the RSS value to the responding eNB. By connecting a GPS module to the UE, the current location (i.e., latitude, longitude, altitude) is added to the message. The GPS coordinate is used as a ground truth reference location. Therefore, in the following, I define the location estimation error as the geographical distance from the estimated location to the location provided by the GPS receiver. A picture of the battery-equipped UE with GPS is shown in Figure 7.1.

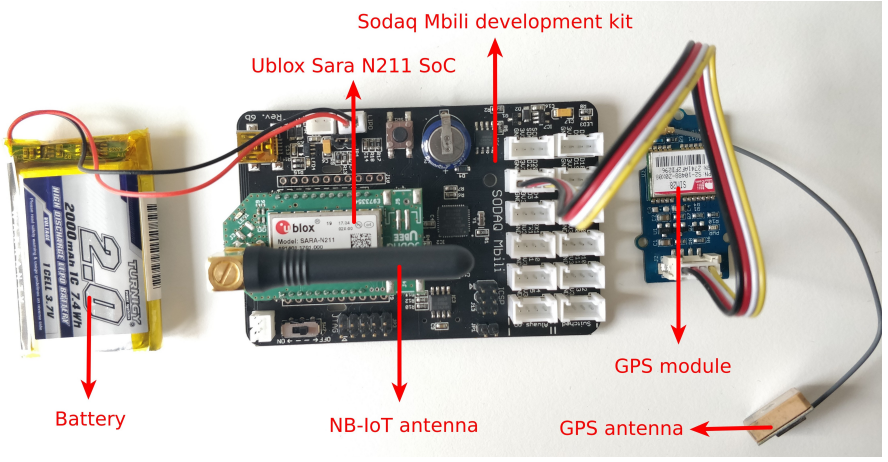


Figure 7.1: Composition of the User Equipment (UE) used to collect NB-IoT messages.

Since a colleague collected a large outdoor publicly available Sigfox and LoRaWAN data set in the city of Antwerp [32], I decided to perform the NB-IoT measurements in the

same environment. In this way, I am able to compare the localization performance of the three most popular LPWAN technologies available on the market. The heterogeneous zone of interest covers an area of 53 km^2 and is fully covered by a public NB-IoT Release 13 network, consisting of 83 eNBs. During the measurement campaign, I collected 1307 NB-IoT messages within the predefined zone, as shown in Figure 7.2.

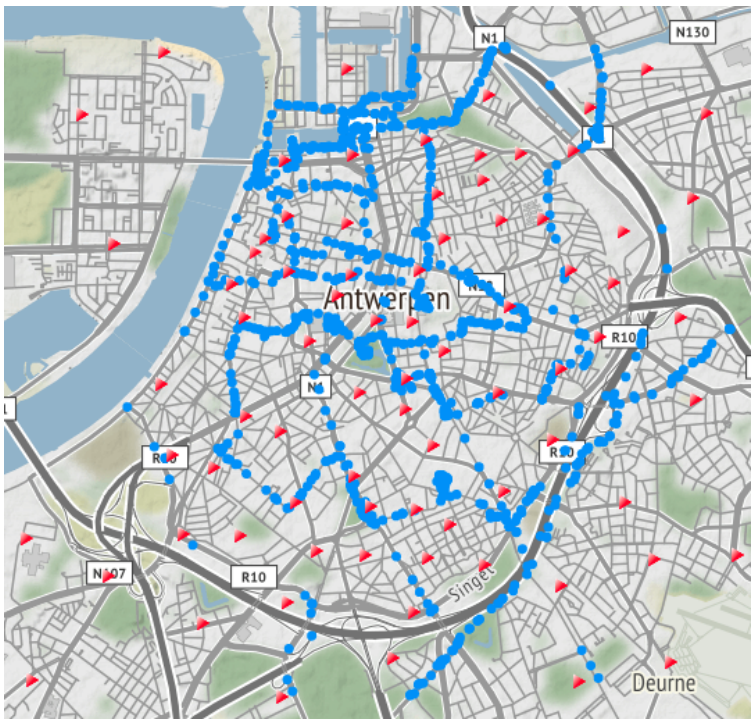


Figure 7.2: Map of the 53 km^2 test environment in the city of Antwerp, Belgium, used to study the localization accuracy of NB-IoT. The 1307 blue dots represent the GPS locations of the transmitter where a message was sent, while the 83 red triangles indicate the eNB locations.

A smaller data set is collected in northern Antwerp. In order to evaluate the mobility issues addressed earlier, I want to analyze the performance of cell selection and reselection in NB-IoT. The eNB centered around the track of my measurement campaign consists of three directional antennas, with each cell covering a sector of 120°. Note that other configurations such as four antennas separated 90° exist. To evaluate the handover between two antennas or cells, I collected 95 messages by moving counterclockwise around the eNB at a constant speed of 7 km/h and an update rate of 5 s. Figure 7.3 depicts a map of the uplink messages, together with the cell sector boundaries.

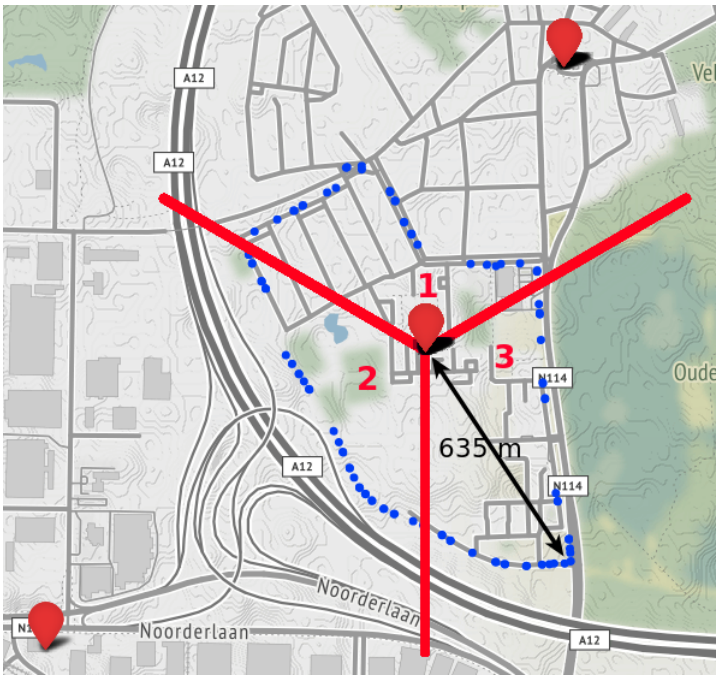


Figure 7.3: Map of the measurement campaign around a single eNB with three cell sectors, used to study the mobility. The 95 blue circles represent the GPS locations of each uplink transmission. The sector boundaries of each cell of the eNB are shown in red.

7.1.2 Proximity Localization

In a proximity-based localization algorithm, the location of the eNB with the strongest link to a mobile UE is used as the estimated location of that UE. This basic yet efficient RSS-based algorithm can already satisfy the requirements of several localization or tracking applications, such as logistics use cases and asset tracking. However, the accuracy highly depends on the base station density and the type of environment.

7.1.3 Range-based Localization

In RSS-based ranging approaches, the RSS from a UE to a specific eNB is translated into a distance by means of a propagation channel model or path loss model. Most path loss models can be customized by configuring environment-specific parameters. Given the urban nature of my large-scale environment, I set the height of the mobile station h_{MS} to 2 m and the height of the base station h_{BS} to 27 m. The carrier frequency f varies between 800 and 900 MHz and the transmission power is set to 14 dBm.

In this work, I evaluate the performance of four urban path loss models. First, the Hata model is an empirical path loss model based on the widely used model of Okumura [107]. The European Cooperation in Science and Technology (COST) received funding to extend the

Hata model, resulting in the COST-231 model [107]. Last, two empirical outdoor path loss models based on the 3GPP Spatial Channel Model (SCM) were proposed by the IEEE TGah group, which standardizes IEEE 802.11ah. Although the original models were devised for LTE, they have been transformed for use in sub-GHz frequency bands [152]. The first model targets a macro LoS deployment scenario (further referred to as *AH-macro*), while a second 3GPP model targets a pico deployment scenario (further referred to as *AH-pico*). The mathematical expressions of all these models are introduced in the LoRaWAN experiments in Section 5.3.1.

Given the mobility issues addressed earlier and in Section 2.2.3, it is not possible to retrieve information about multiple neighboring cells in NB-IoT. In other words, only the antenna of the serving eNB can be used to estimate the location of a UE. As a consequence, well known ranging localization approaches such as triangulation or multilateration cannot be applied. Therefore, I propose a novel algorithm that estimates the position of a UE based on the combination of the serving cell and sector information. First, I estimate the distance d between the UE and eNB using one of the four previously presented path loss models. This results in a circle around the eNB with a radius equal to d, representing all possible UE locations. With the aim to further improve the localization accuracy, I take into account the azimuth of the directional antenna of the serving cell, which covers a sector of 120°. Thus, the antennas provide a spatial filter, eliminating location estimates located outside a given sector. As I assume the location of the UE on the resulting arc to be uniformly distributed, the final location estimate is set to the center of the arc, as illustrated in Figure 7.4.

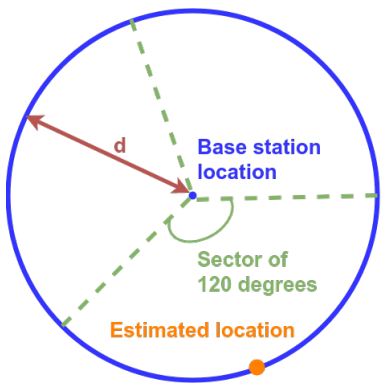


Figure 7.4: Visualization of the RSS ranging algorithm with sector information. The RSS from a single NB-IoT base station antenna is translated to a distance d. The orange dot indicates the resulting location estimate.

7.1.4 Fingerprint-based Localization

Fingerprinting localization is well established in indoor localization applications. The reason for this is that the first phase, i.e., the building of a training database, requires a significant amount of time and effort. However, recent crowd-sourced database initiatives can tackle this problem. Similar to the work in Chapter 4 for Sigfox and Chapter 5 for LoRaWAN, I optimize a ML-based fingerprinting algorithm for localization of NB-IoT devices in an outdoor environment.

As in many ML-based approaches, the localization data set described in Section 7.1.1 is split into 70% training messages, 15% validation messages and 15% test messages. The actual fingerprinting technique is split into two phases, as visualized in Figure 7.5. The first and offline phase consists of collecting training messages into a fingerprint database. Each entry of the fingerprint database consists of the RSS value to the serving eNB and its Cell Global Identity (CGI) (i.e, the features), as well as the GPS coordinate of the ground-truth location (i.e., the target). Consequently, in the second and online phase, a validation fingerprint is being matched to the earlier collected fingerprints, stored in the training database. Using the *k*NN algorithm, I calculate the distance between the validation sample and each training sample in signal space. Afterwards, the location of the UE is estimated by computing the centroid of the ground truth locations of the *k* nearest neighbors.

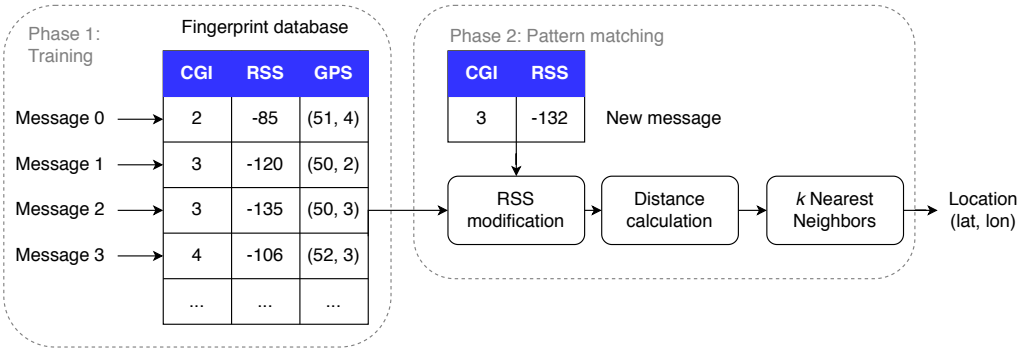


Figure 7.5: Procedure of the fingerprint-based kNN algorithm.

Similar to the Sigfox and LoRaWAN fingerprinting experiments, I also optimize several *k*NN-based algorithms in this work. In brief, I iterate over four different RSS representations, 31 distance metrics and several values of *k*. After the most optimal parameter configuration is found, I validate my results using the messages in the test data set.

7.1.5 NB-IoT Mobility Evaluation Setup

In traditional RSS-based localization algorithms, the position of the UE is determined by combining the estimated distances to multiple base stations in a multilateration algorithm. For example in LTE, the location of a UE is estimated based on timing information to multiple eNBs, leading to an increased localization accuracy. Similarly, an OTDoA-capable UE is able to request information from neighboring cells, due to the presence of a location server in the OTDoA-capable NB-IoT network. However, I discovered that only a single (serving) cell of an eNB can be used for RSS-based localization in NB-IoT. Therefore, in this chapter, I study the mobility of NB-IoT in a localization context. In order to assess the performance of the cell reselection process, I collected messages at a constant velocity and update rate around a single eNB with three cells, as discussed in Section 7.1.1 and shown in Figure 7.3. The azimuths of the three directional antennas are separated 120 degrees, pointing to 0°, 120° and 240°. The trajectory around the eNB was chosen with the objective to force a handover between each cell of the eNB. Consequently, I am able to evaluate two mobility parameters: The *sector reliability*, i.e., whether the UE is located inside the correct

sector of the serving cell, and the *cell reselection time*, i.e., how long it takes to switch to another base station antenna.

7.2 Results

In this section, I present the preliminary results of the RSS-based localization algorithms in terms of location estimation error. Subsequently, the mobility of NB-IoT is evaluated in terms of sector reliability and cell reselection time.

7.2.1 RSS-Based Localization Accuracy

The proximity algorithm results in a mean location estimation error of 340 m and a median of 294 m. In many applications, these rough location estimates can be sufficient, especially when the cost and complexity of the localization needs to be reduced to a minimum.

To optimize the ranging algorithm for my test environment, several urban path loss models have been evaluated. As shown in Table 7.1, all path loss models yield similar location estimation errors. The 3GPP AH-macro model yields the smallest mean and median localization error of 320 m and 259 m, respectively. Although the ranging algorithms perform only slightly better than the proximity algorithm, one needs to keep in mind that only a single eNB cell sector is used to estimate the position of the mobile node.

Table 7.1: Location estimation errors for different urban path loss models of an RSS ranging algorithm in a public NB-IoT network in the city center of Antwerp, Belgium.

Path loss model	Mean error [m]	Median error [m]	95th percentile [m]
Hata	325	272	783
COST-231	327	276	780
3GPP AH-macro	**320**	**259**	**790**
3GPP AH-pico	330	279	776

Finally, a fingerprint-based kNN algorithm is evaluated and optimized. As suggested in [124], different RSS representations and distance metrics are evaluated, along with a parameter sweep of k, i.e., the number of neighbors considered during the matching phase. Table 7.2 shows the location estimation errors and optimal value of k of the NB-IoT fingerprinting algorithm for each RSS representation and distance. In general, the estimation errors vary from 184 m to 207 m. In contrast to previous research, the lineal (i.e., positive and normalized) RSS representations yield the smallest location estimation error. In combination with the Pearson χ^2 distance metric, the smallest location estimation error of 184 m is achieved for the validation set. After iterating over different values of k, it appears that considering the ground truth reference locations of the $k = 2$ nearest neighbors results in the highest localization accuracy.

The location estimation errors from the validation data set are validated by running the algorithm with the optimal parameter set again with unbiased data from the test data set. This yields a mean and median location estimation error of 204 m and 132 m, respectively. Moreover, 95% of all measurements result in a localization error under 679 m.

The location estimation errors for the outdoor proximity, ranging and fingerprinting algorithms are summarized in box plots in Fig. 7.6.

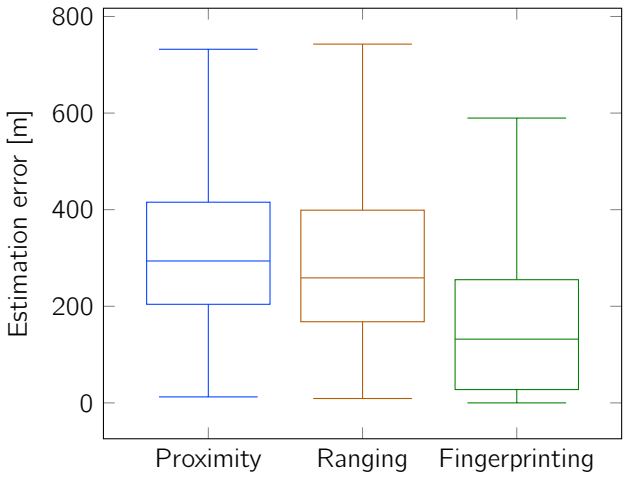


Figure 7.6: Location estimation errors for every outdoor RSS-based localization algorithm when using a single NB-IoT base station antenna. The left box plot shows the proximity algorithm results. The box plot in the middle shows the errors of the ranging algorithm with the 3GPP AH-maco path loss model. The right box plot shows the estimation errors of the validation fingerprinting data set using the optimal parameter configuration.

7.2.2 Mobility Evaluation

Figure 7.7 visualizes the connections between every transmission location and the currently serving cell, as well as the respective cell sector boundaries. Starting at the north side and moving counterclockwise, the first UE positions (i.e., red dots) are located within sector 1, which is the sector of the serving cell. When the UE moves outside sector 1, the cell reselection takes place and in most cases the adjacent cell corresponding to sector 2 becomes the serving cell. After leaving sector 2, the UE remains connected to the cell related to sector 2 and never connects to the cell covering the area of sector 3. Instead, I observe serving cells from other eNBs for short amounts of time (i.e., cell sectors 4 and 5). Thus, only two out of three cells of the center eNB are used.

Table 7.2: Results of the kNN-based fingerprinting algorithm with a single NB-IoT base station antenna, showing the mean location estimation errors and optimal values of k for every RSS representation and distance metric.

Distance measure	Positive RSS		Normalized RSS		Exponential RSS		Powed RSS	
	k	Error [m]	k	Error [m]	k	Error [m]	k	Error [m]
Euclidean	1	201	1	201	1	205	1	205
Manhattan	1	201	1	201	1	205	1	205
Minkowski-3	1	201	1	201	1	205	1	205
Minkowski-4	1	201	1	201	1	205	1	205
Minkowski-5	1	201	1	201	1	205	1	205
Gower	1	201	1	201	1	205	1	205
Lorentzian	1	201	1	201	1	205	1	205
Intersection	1	201	1	201	1	205	1	205
Chebyshev	2	196	2	194	2	198	2	198
Sørensen	2	202	2	202	1	205	2	202
Soergel	2	202	2	202	1	205	2	202
Czekanowski	2	202	2	202	1	205	2	202
Motyka	2	202	2	202	1	205	2	202
Kulczynski	2	185	2	185	1	205	2	185
Canberra	2	202	2	202	1	207	2	202
Divergence	2	202	2	202	1	207	2	202
Clark	2	202	2	202	1	207	2	202
Wavehedges	2	202	2	202	1	207	2	202
Jaccard	2	202	2	202	1	190	2	202
Dice	2	202	2	202	1	190	2	202
Hellinger	1	204	1	204	1	205	1	205
Matusita	1	204	1	204	1	205	1	205
Squared χ^2	1	204	1	204	1	205	1	205
Probabilistic Symmetric χ^2	1	204	1	204	1	205	1	205
Squared Chord	1	204	1	204	1	205	1	205
Squared Euclidean	4	185	1	201	1	205	1	205
Average(L_1, L_∞)	4	185	1	201	1	205	1	205
Pearson χ^2	2	**184**	2	**184**	1	207	2	185
Neyman χ^2	4	185	4	185	1	205	4	185
Additive Symmetric	2	185	2	185	1	205	4	185
Kumar-Johnson	2	185	2	185	1	205	4	185

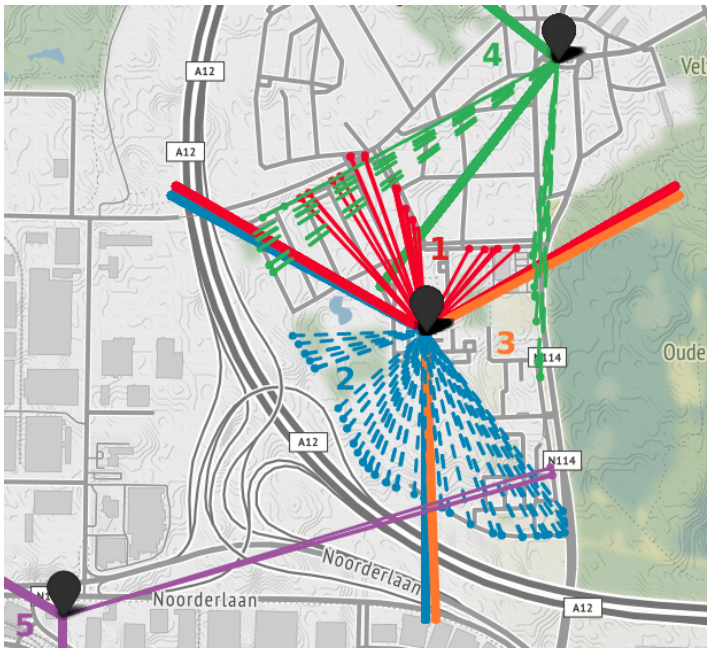


Figure 7.7: Visualization of NB-IoT cell reselection when moving at a constant speed around an eNB with three cell sectors. The color and pattern of the line connecting the GPS location of the UE to the eNB indicates the currently serving cell. The cell sector boundaries are indicated with thicker lines.

The time to select another serving cell varies from 51 s to 104 s, with a mean cell reselection time of 73 s. Within these periods, it is not possible to send NB-IoT messages. Furthermore, the UE remains connected to the cell covering sector 2 for more than four minutes after leaving that sector, without ever connecting to the cell corresponding to sector 3.

7.3 Discussion

While OTDoA in NB-IoT networks promises high-level accuracy, the technology is still in its infancy. As synchronizing eNBs in an NB-IoT network is a challenging and costly task, few operators already deployed this feature in their networks. Moreover, a location server is required in an OTDoA-capable network, which performs the actual localization. Furthermore, a highly accurate clock required in the UE leads to additional device complexity, which is often not desired.

In this chapter, RSS-based localization has been studied using real-world NB-IoT measurements in a city-scale environment. Three RSS-based GNSS-less algorithms are evaluated in terms of localization accuracy. The mean location estimation error varies from 340 m with a basic proximity algorithm to 204 m with an optimized fingerprinting approach. Set side by side, my research in Chapter 4 using multiple Sigfox base stations in the same environment resulted in a mean location estimation error of 340 m in the most optimal fingerprinting

approach. Moreover, I studied RSS-based localization with LoRaWAN in Chapter 5, which coincidentally lead to an accuracy of 340 m as well, using another fingerprinting approach and multiple base stations. Even though there are more NB-IoT base stations than Sigfox and LoRaWAN base stations located in the test environment, only a single base station antenna was used per measurement in the NB-IoT experiments. Contrarily, in the LoRaWAN experiments, up to 10 gateways were considered to estimate the location of the transmitting device. In the Sigfox experiments, the maximum number of receiving gateways increases to 48. From this point of view, the ability to reach multiple NB-IoT base stations promises to increase the localization accuracy significantly, outperforming other LPWAN technologies.

While the proximity algorithm results in a mean accuracy of 340 m, the localization accuracy of the RSS-based ranging algorithm increases only slightly to 320 m. This result can be explained by discussing three limitations of the proposed ranging algorithm. First, estimating the path loss in an urban but heterogeneous environment is not straightforward. Therefore, finding an optimal path loss model is a challenging task. Even if the path loss model estimates all distances very well, the average localization error highly depends on the sector size. This is due to the fact that all locations of UEs sending from within a certain cell sector and distance d to the eNB are estimated at the same centered location of that sector. A second limitation in the ranging algorithm is the fact that only a single directional cell antenna can be used to estimate the position of a UE. This is in contrast to Sigfox and LoRaWAN, where multiple receiving gateways yield more accurate results by performing multilateration. This argument indicates the growing need for multiple receiving cells in NB-IoT. Finally, the last limitation involves the issue of cell sector reliability, which is why I carried out a separate analysis of the mobility issues in an NB-IoT localization context.

Fingerprint-based localization approaches all share one major disadvantage: creating and maintaining a large-scale outdoor fingerprint database requires a significant amount of time and effort. For this reason, outdoor fingerprinting is not that popular as when compared to outdoor ranging or indoor fingerprinting. However, this issue is being tackled by crowdsourcing initiatives such as *TheThingsNetwork* [156], which enable efficient worldwide collection of training data through sensors, chipsets and even smartphones. Besides, fingerprint-based approaches have several benefits over others. In contrast to ranging, the locations of the base stations do not have to be known. Furthermore, when collecting training data, the particularities of the signal strength are stored in the fingerprint database. For this reason, the proposed fingerprinting algorithm consistently yields the highest localization accuracy.

In order to evaluate the localization accuracy of an RSS-based fingerprinting approach, the parameters of a kNN algorithm are optimized. While in the Sigfox experiments, the exponential RSS representation led to the highest localization accuracy, the best results in this chapter were obtained using the lineal (i.e., positive and normalized) RSS representation. Furthermore, the optimal parameter set can be expressed by the Pearson χ^2 distance and $k = 2$. Since the 1307 measurements in the fingerprinting data set are spread over an area of 53 km^2, the distance between each measurement in the training set is quite large, which in turn leads to a smaller value of nearest neighbors k. Finally, the optimal parameter configuration is validated by running the kNN algorithm with unseen data, leading to a mean and median location estimation error of 204 m and 132 m, respectively.

During the collection of real-world measurement data for the localization study, I encountered some mobility issues in the NB-IoT protocol. Therefore, I highlight these issues in this work

and analyze the impact on the localization performance. One major problem I faced was the inability to use information of neighboring cells in my localization algorithms. While in LTE, this feature has been implemented for years, only the serving cell can be used for RSS-based localization in NB-IoT. Some studies speculate that this limitation might be on purpose, in order to increase scalability and decrease signaling overhead and power consumption [60, 61]. Nonetheless, multiple neighboring cell reports can definitely increase the accuracy of RSS-based localization. Fortunately, after the publication of this work, more and more chipset manufacturers started implementing the scanning for neighboring cells, which is an optional feature in NB-IoT. To further support this, De Nardis et al. used a network scanner to demonstrate that the average positioning error drops below 100 m when scanning for multiple nearby eNBs [63].

The results of the mobility analysis in Figure 7.7 demonstrate the cell reselection process of NB-IoT. When moving at a constant speed of 7 km/h, the cell reselection process took 73 s on average. Important to note is that this cell reselection time depends on both the velocity of the UE and cell reselection parameters configured by the network operator. Decreasing the update rate of the cell reselection increases battery lifetime but is unfavorable from a localization perspective, as this introduces latency and reduced location update rates. Additionally, I investigated the cell sector reliability. After each cell reselection process, most of the messages are located within the sector of the new serving cell. However, it is worth discussing two peculiarities that can be observed from Figure 7.7. First, the UE remains connected to the same serving cell after moving outside the serving cell sector. Second, serving cells from eNBs located further away are reported, especially after losing the connection to the currently serving cell. The first issue might arise when the connection to the currently serving cell remains stronger than the connection to an adjacent cell, e.g., due to multipath or interference effects. The second issue can be caused by side and back lobes of directional antennas, which seem to establish a better connection for a short amount of time. Furthermore, despite being located in the serving cell sector, the UE might select a stronger cell from a nearby eNB if there is no more LoS to the current eNB. Thus, in line with the ideas of Moon et al. [65], it can be concluded that the performance of the cell reselection process in the NB-IoT technology should be improved.

7.4 Conclusion

The cellular-based NB-IoT technology connects a plethora of devices to the IoT. Consequently, there is an urgent need to localize such devices in an energy-efficient way. While OTDoA is introduced in 3GPP Release 14, few operators deployed this localization feature in their networks, given the cost and complexity to synchronize eNBs. Therefore, I investigated RSS-based localization approaches, providing an energy-efficient solution at low cost and low complexity.

Simulations of localization experiments often result in over-optimistic localization performance. Therefore, I insisted to perform outdoor localization experiments using real-world measurements. In this way, I can paint a more realistic picture of the localization performance. Additionally, I face the practical issues and constraints of working with real-world data.

This chapter provided the firsthand results of RSS-based localization experiments using a public NB-IoT network. Experiments carried out in a large-scale urban environment led to mean location estimation errors ranging from 340 m in a basic proximity algorithm to 204 m in an optimized fingerprinting algorithm. While a fingerprinting approach usually requires more time and effort, the location accuracy can thus be increased. Eventually, crowd-sourced fingerprinting can solve city-scale collection of training data.

Traditional multilateration approaches could not be applied, since many UE firmwares only report the serving cell rather than all neighboring cells within range. In order to address this issue, I studied the mobility of NB-IoT in a localization context. Due to unpredictable antenna patterns and environmental influences, a UE might connect to a cell while not located inside the corresponding cell sector. This poor cell sector reliability and the long cell reselection times indicated the need for proper handover support and observation of multiple cells in NB-IoT networks. After addressing and publishing these findings, I am happy to notice that chipset manufacturers such as Nordic Semiconductor have started updating their firmware to include multi-cell discovery for localization purposes [169]. Future research may therefore investigate the increase in location accuracy when considering multiple cells in a multilateration approach.

Chapter 8

From Ground to Space: A Survey on Energy-efficient Large-scale Positioning: LPWAN vs. GNSS vs. LEO-PNT

As an alternative to GNSSs, all previous chapters described how terrestrial LPWANs can be used to estimate the location of a mobile transmitter. Advantages of these technologies are the optimized energy consumption profiles for IoT use cases and the ability to provide location updates in both indoor and outdoor environments. On the other side of the coin, the positioning accuracy is rather limited when compared to GNSS and the coverage is bound to the range of the often nationwide terrestrial networks.

Despite all the limitations of GNSS described in Chapter 1, GNSS technologies and techniques have continued to evolve [170]. While Assisted GNSS (A-GNSS) receivers were introduced in the market years ago, novel techniques are only being adopted very recently. For example, snapshot processing techniques (S-GNSS) enable cloud processing by transmitting raw observables over a terrestrial connectivity link to the cloud, successfully reducing the UE complexity and energy consumption [171].

Recently, a myriad of companies started deploying LEO satellite constellations in the race towards the constant global coverage on Earth for the emerging market of satellite IoT [172, 173]. On the one hand, big tech companies like SpaceX, OneWeb, and Amazon are deploying hundreds and even thousands of satellites to provide worldwide broadband Internet. On the other hand, smaller companies such as Kineis, Lacuna Space, Wyld and Hiber focus on very low energy satellite communication and positioning of mobile end devices, which are especially of our interest. As illustrated in Figure 8.1, LEO satellites are around 20 times closer to Earth compared to GNSS satellites. Therefore, the oftentimes stronger LEO satellite signals enable PNT applications in GNSS-denied environments.

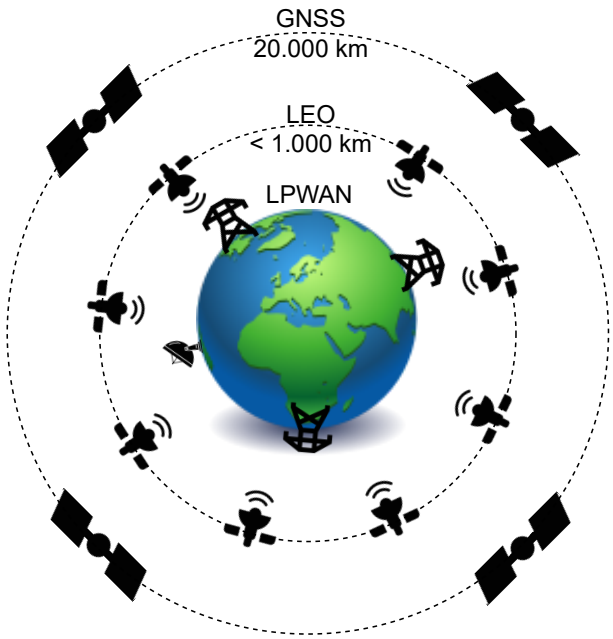


Figure 8.1: Three categories of technologies enabling large-scale positioning.

This chapter comprises a survey carried out for ESA, in which I investigate what energy-efficient large-scale positioning techniques are available today and how they perform when compared to each other. I tackle this by making the following contributions:

- I provide an overview of state-of-the-art, energy-efficient and large-scale positioning techniques using LPWAN, LEO and GNSS technologies. To the best of my knowledge, I am the first to combine these in a single comprehensive survey.
- I compare the performance of each positioning technique in terms of 16 dimensions and visualize them in a performance matrix.
- For each positioning technique, I evaluate the interoperability and the possibility to integrate multiple techniques in a single satellite IoT device.
- Through example IoT positioning use cases, I discuss a set of important trade-offs to consider during their design.

The survey in this chapter is structured in the following way. I first introduce state-of-the art positioning techniques leveraging LPWAN, GNSS and LEO systems in Section 8.1. Their performance is evaluated in Section 8.2. Using a defined set of dimensions, I am able to compare these techniques and to create a performance matrix. Section 8.3 discusses the trade-offs to be made when designing a location-enabled IoT use case. Finally, Section 8.4 summarizes the main conclusions and discusses remaining challenges.

This chapter is based on:

Thomas Janssen, Axel Koppert, Rafael Berkvens, & Maarten Weyn. *A Survey on IoT Positioning leveraging LPWAN, GNSS and LEO-PNT (submitted on July 30, 2021)*. IEEE Internet of Things Journal, pp. 1–25, 2023.

8.1 State-of-the-art positioning techniques

When discussing positioning systems, it is important to distinguish between positioning technologies and techniques. Within the scope of this work, a positioning technology is characterized by wireless connectivity combining a network of end devices, gateways and/or satellites, which enables positioning. In contrast, a positioning technique refers to a certain method or algorithm to estimate the position of a device or UE, independent of the used technology. Therefore, multiple positioning techniques can be applied using the same technology. LPWAN, GNSS and LEO technologies constitute three large-scale categories of positioning technologies, as illustrated in Figure 8.1. A high-level overview of state-of-the-art positioning technologies and techniques in each of these categories discussed in this work is shown in Figure 8.2. Their core concepts are briefly described in the following subsections. In order to provide a comprehensive survey, please note that the state-of-the-art LPWAN positioning techniques of Chapter 2 are briefly reintroduced.

Alternative surveys exists, although they specifically target LPWAN [21], GNSS [174] or LEO [172, 175, 176] technologies. Besides, many works focus on communication rather than localization. To the best of our knowledge, no other work in literature has investigated and compared this wide range of IoT positioning solutions, taking into account their energy efficiency.

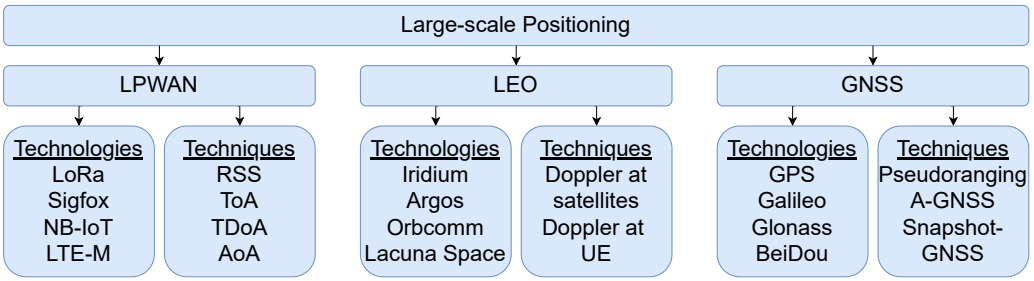


Figure 8.2: Overview of large-scale positioning technologies and techniques discussed in this work.

8.1.1 LPWAN

Terrestrial LPWANs are designed for long-range and low-power communication of small messages [34]. In more recent years, the networks of IoT transceivers and ground stations are also used as a means to provide a localization solution. LoRaWAN and NB-IoT are by far the most prominent LPWAN technologies available on the market [114, 9]. Where LoRaWAN provides operational flexibility and the choice for a private or public network, NB-IoT can be easily deployed on top of existing cellular infrastructure. Other LPWAN technologies include Sigfox and LTE-M, which offer battery lifetimes of several years as well.

As in most LPWAN positioning techniques, an RSS-based approach determines the location of a mobile IoT device through uplink communication. When a UE transmits a message, the RSS is measured at nearby gateways. This information is sent as metadata along with the

payload to the cloud, where the data processing and location estimation steps are performed. An RSS ranging technique uses a path loss model to translate the signal strength into a distance to a certain gateway. The position estimate can subsequently be calculated using various algorithms, such as Least Squares or Min-Max. Another RSS-based technique is fingerprinting, in which training RSS data with ground-truth information is collected in the area of interest, and a new fingerprint is matched to this training database to locate the mobile transmitter. While this technique incorporates multipath effects and environmental influences, it requires a lot of effort, cost and time to create a large fingerprinting database. The accuracy of these positioning techniques highly depends on the number of receiving gateways, as well as the accuracy of the path loss model in a given environment, as discussed in Chapter 4.

Another popular technique to estimate the location of an LPWAN-enabled device is time-based ranging. In traditional ToA approaches, the absolute time for a signal to travel from transmitter to receiver is measured. Multiplying by the speed of light yields the distance between the UE and the gateway. If enough gateways received the signal of a mobile transmitter, a multilateration algorithm is used to estimate the location of the transmitter. However, to avoid the need for synchronization between the UE and gateways, TDoA has become more popular. In this technique, the distance between the target and reference points is calculated based on the difference of arrival times at these reference points [177]. Geometrically, this leads to a hyperbola. With at least four gateways, the final location estimate can then be calculated as the intersection of the hyperbolas. It should be noted that when discussing TDoA, one mostly refers to uplink TDoA approaches. However for NB-IoT, the 3GPP has defined OTDoA in Release 14. Despite the limited number of networks currently supporting this feature, the first OTDoA experiments show promising positioning improvements [91].

In combination with RSS- or time-based techniques, the AoA of an LPWAN signal can be determined using an antenna array at the gateway side and a triangulation algorithm [93]. Finally, it was demonstrated that an increased positioning accuracy can be achieved using a combination of TDoA and AoA via sensor fusion in a particle filter [96].

8.1.2 LEO

While there are hundreds of LEO satellite constellations in orbit or to be launched, they are designed with different objectives in mind. For example, Iridium and Globalstar provide a voice service, while SpaceX, Amazon and OneWeb aim to deliver global broadband Internet [178]. Similarly, Telesat aims to deliver secure broadband connectivity. The Argos system is designed for Earth observation purposes [179], similar to the Sentinel satellites of the EU Copernicus programme [180]. Omnispace focuses on the integration of their satellite network with a terrestrial NB-IoT network, while Hiber, Wyld and Lacuna Space aim to achieve this using a network of LoRa gateways and LEO satellites [172]. Although all of these examples may not be primarily designed for positioning purposes, the satellite IoT market has made a myriad of companies to shift focus towards the monitoring and locating of remote IoT devices leveraging LEO satellites. For instance, Satelles is developing a service which provides a true ranging signal similar to GNSS, leveraging Iridium satellites. Two companies are currently testing a system targeting the autonomous driving market. Xona

Space Systems is developing a standalone LEO-PNT system using a dedicated constellation of 300 cubesats [181]. The company aims to deliver a reliable and resilient PNT service that is ten times more accurate compared to GNSS [182]. Similarly, Geespace is developing a 240-satellite constellation that will feature combined Precise Point Positioning (PPP) and Real-Time Kinematics (RTK) services, aiming to provide centimeter-accurate precise positioning and connectivity for automaker Geely [183]. While many of the positioning solutions are still in a research or testing phase, some industry leaders already provide early access to a commercial localization service. The interesting part of this type of positioning approach is the fact that most LEO satellites support two-way communication via ground stations, enabling to transmit a location estimate from the UE to the cloud.

In order to provide a positioning service leveraging LEO satellite signals, most currently available solutions exploit the Doppler effect. For instance, the Argos system operated by CLS and Kinéis provides satellite telemetry services for scientific and environmental applications. Through precise Doppler measurements with the Argos constellation, end users are provided with a location estimate, along with an indication of the estimation accuracy [179, 184]. By sending multiple uplink messages, a single receiving satellite performs a Doppler measurement. The time and frequency observation of the received signal is forwarded via a ground station to a solver, which estimates the user position using either a Least Squares algorithm, or a more advanced Extended Kalman Filter (EKF) [185]. Several improvements to Doppler positioning are being investigated, such as only transmitting during a satellite pass using forecasting software.

Due to the increasing number of LEO constellations provided by different operators, it has become a challenging task to provide a universal positioning technique. However, the Doppler positioning technique can also be performed by the UE, rather than by the satellites. Exploiting SoOP from LEO satellite constellations is one of the most recent developments and is referred to as instantaneous Doppler positioning. This approach has the potential to leverage mega-constellations for zero-cost worldwide access to space signals using Software Defined Radios (SDRs), removing the need for specific indoor infrastructure [26]. Farhangian et al. designed a LEO satellite receiver to perform local Doppler measurements using downlink signals from multiple LEO constellations in an opportunistic way [186]. The feasibility of this approach was demonstrated using Iridium NEXT, GlobalStar, Orbcomm and Starlink satellites in both simulations and experimental setups [187, 188, 189, 190]. Moreover, the fusion of mixed SoOP has been proven beneficial in weak signal environments as well [191]. Finally, the Doppler measurements can be used as assistance data in A-GNSS (see Section 8.1.3), as well as in Inertial Navigation Systems (INSs) [192].

8.1.3 GNSS

When it comes to GNSS-based localization solutions, the trend in the last decade was to manufacture multi-constellation GNSS receivers, e.g., combining the American GPS, European Galileo, Russian GLONASS and Chinese BeiDou satellite constellations in a single chipset. In this way, both global coverage and availability are extended. However, innovations to increase the energy efficiency lie in the used GNSS technique. In general, GNSS techniques relevant to IoT use cases can be further categorized into conventional observable-based GNSS, A-GNSS, and S-GNSS techniques.

Conventional GNSS receivers attempt a continuous signal tracking, which yields pseudorange, Doppler and phase observations. The tracking stage is preceded by an acquisition stage, in which the satellite signals are detected and the tracking loops are initialized. GNSS positioning is based on ToA, as the time to travel from the satellite to the receiver is used to calculate the distance, i.e., pseudorange, between them. Provided that satellite orbit and clock information is known, and at least four satellites are in view, the receiver can determine its position based on the pseudoranges. While Doppler-only positioning could be performed with similar principles as for LEO positioning, it is rarely applied due to the low accuracy of several kilometers. The reason for this is that GNSS satellites orbit the Earth at significant lower velocities than LEO satellites. A common way to cope with the high energy consumption of conventional receivers is duty cycling, i.e., periodically waking up to receive GNSS signals and going back to low-power sleep modes. However, this technique does not meet the energy requirements of IoT use cases. Therefore, a significant amount of research is devoted to novel energy-efficient GNSS techniques.

Obtaining a first GNSS fix on the UE can consume a considerable amount of time and energy. Therefore, several techniques exist to reduce the Time To First Fix (TTFF), and consequently, the energy consumption. In order to compute a first fix, the satellite signals have to be acquired and the ephemeris data containing information on the satellite orbits and clock needs to be decoded from the satellite navigation message. The acquisition requires multiple correlations for different time (i.e., delay of the ranging code modulated on the carrier) and frequency (i.e., carrier Doppler) offsets. The more a-priori information is available, the narrower the search space and the more efficient the acquisition processing becomes. The principle of A-GNSS has been developed in order to provide such assistance data from an external source to the GNSS receiver, with the aim to reduce the TTFF [193]. The assistance data can be a rough location and time estimate of a terrestrial network, as well as ephemeris data, which can be valid for up to a few weeks. Providing ephemeris data makes decoding it from the GNSS signal obsolete. For example, LPWAN can provide this information in an energy-efficient way. Moreover, this connection with a terrestrial network rises the opportunity to communicate the GNSS location to the cloud. Furthermore, if there is no possibility to connect to a terrestrial network, a GNSS receiver can reduce the TTFF by predicting the ephemeris data, based on previously calculated location, time and orbital parameters.

Snapshot processing techniques constitute a third set of energy-saving GNSS techniques. The main idea of these cloud processing techniques is to only sample a short portion of the received satellite signal (referred to as a snapshot), digitize the samples and transmit them via a connectivity link to the cloud, where the data is processed and the location is calculated [171]. By performing the most power-hungry functions in the cloud, the overall energy consumption is drastically reduced. The connectivity link can be provided though ground stations or LEO satellites. Depending on the length of the snapshot and the limitations of this link, a trade-off needs to be made between how many processing is performed on the device and how many data is sent to the cloud [194]. Furthermore, a S-GNSS receiver requires some adaptations from standard GNSS processing to derive a Position, Velocity and Time (PVT) solution. A basic block diagram for snapshot processing is shown in Figure 8.3. Even with small snapshot lengths, the frequency and code phase can be detected. To calculate the pseudoranges and solve some ambiguities, a rough estimate of the current location and time is often required. This information can be sent to the receiver via an LPWAN

connection, together with the ephemeris data. Using the latter, the current position and time can be calculated. Finally, it has been proven that even without a rough position and time estimate, meter-level accuracy can be achieved [195].

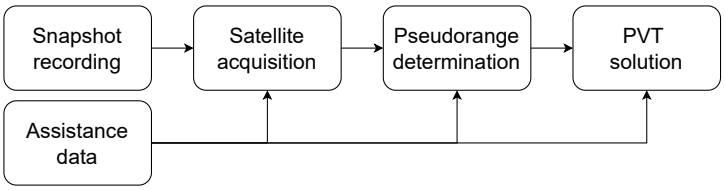


Figure 8.3: Block diagram of snapshot processing using GNSS. Note that all steps except the snapshot recording might be outsourced from the UE to the cloud.

8.1.4 Additional sensors

The aforementioned technologies and techniques can be complemented with sensors providing more accurate or contextual information of the device location. Inertial Measurement Units (IMUs), e.g., a combination of accelerometer, gyroscope and magnetometer can be used for dead reckoning in GNSS-denied environments, or to save energy if the transmitter has not moved since its last position update. Barometers are often used to estimate heights, e.g., to determine the floor level in an indoor positioning use case. Finally, Wi-Fi scanning, Near Field Communication (NFC) tags and BLE ranging can yield better positioning performance in indoor or urban environments. As the goal of this survey is to assess the performance of large-scale positioning techniques, i.e., tens of squared kilometers, these additional sensors fall outside of the scope of this work.

8.2 Positioning performance evaluation

In this section, I compare the current state-of-the-art of the aforementioned positioning techniques using LPWAN, GNSS and LEO technologies. The matrix in Table 8.1 shows an overview of the performance comparison in terms of 16 dimensions. This qualitative matrix enables the relative comparison between localization approaches. More context and a more detailed discussion on the performance of each positioning technique with respect to the dimensions is provided in the following subsections.

8.2.1 Hardware availability

The first dimension indicates how accessible the hardware of a technology is, and if there are Commercial Off-The-Shelf (COTS) chipsets available.

Table 8.1: Qualitative performance comparison matrix of LPWAN, GNSS and LEO positioning techniques in terms of 16 dimensions. A score of 1 (red) is highly limiting, while a score of 5 (green) is highly beneficial. A minus sign (gray) denotes not applicable.

	LPWAN			LEO			GNSS			
	RSS	T(D)oA	AoA	LEO + LPWAN	Doppler at satellites	Doppler at UE	Snapshot GNSS	GNSS + LEO	A-GNSS	GNSS
Hardware availability	5	5	3	3	4	3	4	3	5	5
Network accessibility	5	5	5	4	3	4	5	4	5	5
Energy consumption profile	5	5	5	3	3	3	3	2	2	1
Localization accuracy	1	2	3	3	3	4	4	5	5	5
Ubiquity of coverage	3	2	3	5	4	4	4	5	4	4
Scalability	5	4	4	5	4	4	4	5	4	5
TTFF	5	5	5	3	2	2	4	2	4	2
Data rate & BW	-	-	-	-	-	-	-	-	-	-
Interoperability	5	5	5	5	5	5	5	5	5	5
Communication of observables	5	5	5	5	5	5	5	5	5	1
Index of technology readiness and maturity	5	5	4	2	3	2	3	4	5	5
Standardized or proprietary	-	-	-	-	-	-	-	-	-	-
UE cost	5	5	5	3	2	4	4	2	3	4
UE complexity	5	5	5	4	4	3	4	3	3	3
Location update rate	4	4	4	3	2	4	4	2	5	5
Local or remote processing	-	-	-	-	-	-	-	-	-	-

Table 8.2: Non-exhaustive list of commercially available positioning chips and modules. Positioning techniques marked with an asterisk (*) denote that the chip or module does not natively implement the technique but provides support (i.e., required hardware/software) for it.

Category	Technology	Chipset / module name	Positioning technique
LPWAN	NB-IoT	u-blox SARA-N3/R5 series	RSS*, OTDoA*, AoA*
		Nordic Semiconductor nRF9160	RSS*, OTDoA*, AoA*
		Qualcomm 212 LTE modem	RSS*, OTDoA*, AoA*
		Quectel BC660K-GL	RSS*, OTDoA*, AoA*
	LoRa	Semtech SX1276	RSS*, OTDoA*, AoA*
		Semtech LR1110 & LR1120	RSS*, OTDoA*, AoA*. LoRa geolocation
		Microchip RN2483	RSS*, OTDoA*, AoA*
	Sigfox	Sigfox TD1207R	RSS*
GNSS	GPS, Galileo, GLONASS, BeiDou	u-blox MAX M10S	Pseudoranging, A-GNSS
		Quectel LC79D	Pseudoranging, A-GNSS
		Baseband Technologies snapshot GNSS receiver	A-GNSS. snapshot GNSS, cloud processing
		Syntony SoftSpot IoT	A-GNSS. snapshot GNSS, cloud processing
		Semtech LR1110	A-GNSS, cloud processing
		Maxim MAX27690 RF frontend	Snapshot GNSS, cloud processing
LEO	Iridium	Jackson Labs PNT-62xx STL receiver	Instantaneous Doppler positioning
	Argos	ARTIC R2 chipset	Doppler positioning at satellites
		Arribada Horizon ARTIC R2 development kit	A-GPS, INS
	Orbcomm	Orbcomm OG2-M modem	Instantaneous Doppler positioning
	Globalstar, GPS	Globalstar SPOT Trace	Doppler positioning at satellites
	LoRa, GNSS, Wi-Fi	Semtech LR1110	Doppler, GNSS and Wi-Fi scanning
		Miromico FMLR-LR1110-X-STL0Z module	Doppler, GNSS and Wi-Fi scanning

8.2.1.1 LPWAN

Since the rise of the IoT, LPWAN devices are becoming highly available to both industrial and commercial users. End devices are so commonly integrated in our society that they have become ubiquitous. A few examples include smart meters, temperature and humidity sensors. LoRaWAN and NB-IoT, two of the most popular LPWAN, each provide several UEs in a different way. While Semtech is the major manufacturer of LoRa chips, some manufacturers have a license to produce them (e.g., Microchip) or collaborate with Semtech (e.g., ST Microelectronics). Alternatively, manufacturers may develop a LoRa module based on a chip from Semtech. In contrast, any manufacturer is allowed to produce NB-IoT-enabled chipsets and modules, provided that the corresponding 3GPP standard is followed.

The first rows of Table 8.2 provide an overview of commonly used LPWAN chipsets and modules. Common NB-IoT manufacturers include U-blox, Nordic Semiconductors, Qualcomm and Quectel. Among the LoRa chips, the LR1110 chip from Semtech integrates LoRa with GNSS and Wi-Fi, providing a geolocation service through the 'LoRa Edge' platform. Finally, the company behind Sigfox provides LPWAN modules that work together with GNSS and accelerometers to provide a low-power localization service.

In general, one can conclude that LPWAN chipsets and modules are highly available. As RSS- and timing-based localization techniques generally do not depend on the manufacturer or type of UE, they are given a score of 5 in Table 8.1. Note however that UEs must support advanced localization techniques. For example, OTDoA requires accurate timestamps. Furthermore, besides a transmitter, AoA-based techniques require an antenna array at the receiver side to determine the angle of the incoming signal. Although antenna arrays are widespread, most LPWAN gateways are only equipped with a single antenna. Therefore, it is often not possible to deploy AoA in a public LPWAN network.

8.2.1.2 GNSS

Conventional observable-based GNSS and A-GNSS features are implemented in nearly all recent GNSS receivers and smartphones. Moreover, industry-leading companies such as u-blox provide an assistance service along with their multi-constellation GNSS chipsets. In contrast, S-GNSS receivers can be less complex and expensive as some traditional building blocks are not required in this type of receivers. On the one hand, the building blocks of S-GNSS receivers are widely available and consist of an RF frontend and a storage element, in order to digitize and store an incoming signal for processing at a convenient time in the cloud. An example is the Maxim MAX2769 GNSS-specific frontend. On the other hand, only few snapshot receivers are commercially available, e.g., the Baseband Technologies S-GNSS receiver. Furthermore, the aforementioned LR1110 chipset from Semtech is a LoRa chip which enables passive Wi-Fi and GNSS scanning [196]. The device captures a short portion of the satellite signal, extracts pseudoranges and aggregates them into a NAV message, which can be sent to the cloud for position estimation.

8.2.1.3 LEO

The category of LEO-based positioning techniques is the most recent category, and therefore, chipsets and modules of these techniques are not as ubiquitous as LPWAN or GNSS devices. For example, both Jackson Labs and Orolia do provide Iridium-enabled devices but these do not support actual location estimation yet. In contrast, commercial positioning hardware is available for the Argos system. The ARTIC R2 chipset, for example, is compatible with the Argos-2, Argos-3, and the upcoming Argos-4 system. An open source reference design is provided, along with all technical details of the chip. Moreover, an Arduino library and multiple development kits are widely available. Furthermore, the KIM1 module provided by CLS and certified by Kinéis and CNES offers a more finished product, requiring less development. A shield board is also available to ease integration. Other companies providing LEO hardware include Orbcomm and Lacuna Space, as listed in Table 8.2. Finally, the feasibility of instantaneous Doppler-based positioning using LEO signals has mostly been demonstrated based on SDR implementations, rather than tailored end products [197].

Recent hardware modules support the combination of LEO with LPWAN and GNSS. Kinéis is partnering with Bouygues Telecom to integrate the Argos system with the LoRaWAN standard. Hiber, Lacuna Space and Wyld are competing companies, also providing a combination of a LEO constellation and a LoRa network. Orbcomm has designed a 'dual-mode' platform, in which their LEO constellation is combined with a cellular network. Similarly, Intellian is manufacturing the user terminals for OneWeb, aiming to deliver commercial communications services to remote regions and industrial sectors. Globalstar provides devices combining LEO and GPS satellites to provide near real-time positioning in areas without terrestrial networks. While some of these companies are still developing and evaluating their solutions, some of them already offer commercially available hardware, as listed in Table 8.2.

8.2.2 Network accessibility

A second dimension indicates how accessible a network of gateways or satellites is, e.g., for commercial, personal, or industrial use. Table 8.3 lists currently available networks which are used for positioning. Additionally, I discuss any restrictions or limitations on the usage of these networks.

Table 8.3: Overview of available positioning networks and a non-exhaustive list of providers.

Category	**Network technology**	**Network provider(s)**
LPWAN	NB-IoT	Orange, Vodafone, T-mobile, China Mobile, Telia
	LoRa	The Things Network, Actility, private network operators
	Sigfox	Sigfox (UnaBiz), Engie M2M, HELIOT, WND
GNSS	GPS	US Air Force
	Galileo	European GNSS Agency
	GLONASS	Russian Federation
	BeiDou	China National Space Administration (CNSA)
LEO	Iridium	Iridium
	Argos	CLS, Kinéis
	Orbcomm	Orbcomm
	Globalstar	Globalstar
	LoRa	Lacuna Space

8.2.2.1 LPWAN

Since the emergence of the IoT, the number of low-power long-range networks worldwide has been growing rapidly. LPWAN technologies are deployed in various ways. Currently, 148 public and private LoRaWAN network operators are active in 162 countries [8]. Similarly, as of September 2022, 167 operators are actively investing in NB-IoT technology, of which 124 have commercially launched NB-IoT networks in 80 countries [198]. Sigfox networks are operated nationwide, either by Sigfox or a partnering telecom provider. All of these networks are accessible for commercial, industrial, and personal use. Roaming between these networks has been a hurdle, but recent initiatives aim to tackle the problem and accelerate LPWAN roaming worldwide. For instance, full LoRaWAN roaming is available in 27 countries around the world as well as via the satellite network of Lacuna Space.

8.2.2.2 GNSS

GNSS networks are highly accessible. While there are signals dedicated to certain user groups (e.g., military or public authorities), everybody can use most signals from the different constellations free of charge. The system providers publish all required information to exploit the open services. The plethora of multi-constellation GNSS receivers allow the end user to use satellites from multiple constellations simultaneously.

8.2.2.3 LEO

In general, LEO satellite networks are not as accessible as when compared to GNSS constellations. First, most LEO positioning providers, such as Argos and Lacuna Space, require a paid subscription to use their Doppler positioning service. Second, many LEO constellations are not finished yet and only a small number of often region bound beta testers can participate in the program. When passively performing Doppler measurements on the UE using SoOP from multiple constellations, however, the network accessibility increases.

8.2.3 Energy consumption profile

As the aim of this survey is to provide energy-efficient positioning techniques for the IoT, the energy consumption profile is one of the most critical dimensions. This section covers various energy-related parameters, ranging from overall UE energy consumption, over battery lifetime, to the availability of different energy profiles (e.g., sleep modes, idle mode, cold/warm/hot start). It is very important to highlight that, even though I provide numerical results originating from data sheets, simulations and experiments, the overall energy consumption highly depends on a plethora of parameters, which may significantly differ based on the used hardware, the use case and the environment. Examples of such parameters are the location update rate, transmission power, payload size and sleep mechanisms.

8.2.3.1 LPWAN

Most LPWAN localization systems work through the 'localization by communication' concept, i.e., by sending an uplink message. Therefore, the energy consumption of positioning techniques such as RSS, TDoA and AoA equals the energy consumption of this uplink communication using a certain LPWAN technology. Several recent studies have analyzed, simulated and demonstrated the ultra-low power consumption of LPWAN technologies. Singh et al. provide an analysis of the actual energy consumption profiles of Sigfox, NB-IoT and LoRaWAN [35]. The analysis shows that a LoRa transmitter consumes 37.05 mJ to transmit a 5-byte uplink message and has an average sleep current of 81 µA at 3.7 V, while NB-IoT transmission consumes 63.48 mJ, with a deep sleep current of 0.10 µA at 3.7 V. However, the overall energy consumption can vary significantly depending on the configuration parameters such as payload size, SF, update rate and sleep modes. For different update rates, the resulting estimated battery lifetimes are shown in Figure 8.4.

The energy consumption profiles of LPWAN technologies show a peak in current consumption during message transmission and in the idle period, which highlights the need for sleep modes. Examples are the eDRX and Power Saving Mode (PSM) of NB-IoT, as shown in Figure 8.5. While a Quectel BG96 NB-IoT module consumes 623.7 mW during transmission at 23 dBm, these modes consume only 3.63 mW and 10 µW, respectively [199]. Furthermore, the SF or LoRa provides the flexibility to tune the balance between energy consumption, data rate and communication range, depending on the application requirements [200].

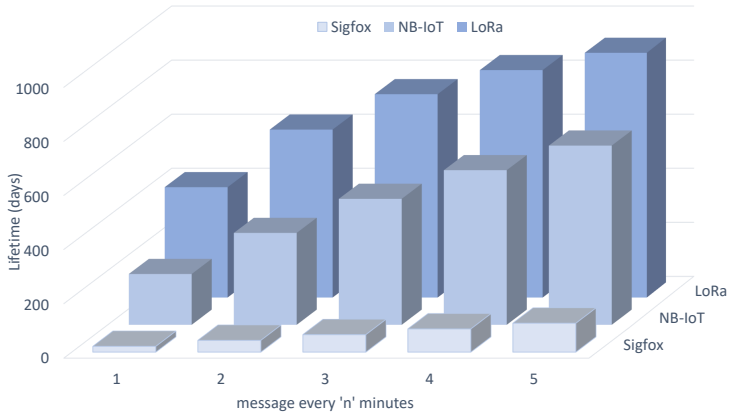


Figure 8.4: Battery lifetime for different LPWAN technologies and uplink message update rates, using a 5-byte payload size and a 2500 mAh battery [35].

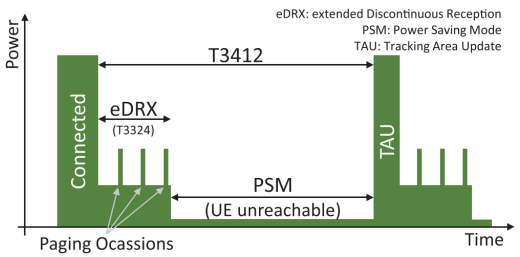


Figure 8.5: NB-IoT transmission cycle with sleep mechanisms [199].

8.2.3.2 GNSS

In high contrast to LPWAN, GNSSs originally were not designed with low energy consumption in mind. The GNSS technology report of 2020 lists several market segments and typical state-of-the-art receiver specifications [201]. In the IoT market segment, a typical receiver consumes 17 mA during signal acquisition and 0.5–8 mA during tracking, using a power supply of 1.4–4.3 V. The feasibility of adding a GNSS receiver to a LoRaWAN tracking device in terms of location accuracy, battery lifetime and location update rate is analyzed in [5]. The study shows that a GNSS receiver should only be omitted if a location error of more than 100 meters is acceptable and the energy budget is extremely constrained, provided that the LoRa SF is configured correctly. Furthermore, the battery lifetime of LoRaWAN trackers is estimated, depending on different application requirements. When tracking an animal with 48 location updates per day, a SF equal to 9 and a minimum battery lifetime of 3 years, the battery of the IoT device would last 4688 days without GNSS receiver and 2446 days with GNSS receiver. However, when tracking an animal using the same location update rate but with a desired 10 year battery lifetime, a GNSS receiver can no longer be used.

An empirical study on energy consumption of GNSS chipsets in smartphones, which also have energy constraints, demonstrates that a smartphone with a dual-frequency GNSS chipset consumes on average 28% and 37% more power compared to a single frequency GNSS smartphone, in indoor and outdoor environments, respectively [202]. Figure 8.6 shows the

energy consumption profile for the two smartphone GNSS receivers. Using a location update rate of 1 s, the mean energy consumption of the single- and dual-frequency receivers equals 232 mJ and 318 mJ, respectively. Due to this difference, the battery of the smartphone with single-frequency GNSS receiver lasts 10 hours longer.

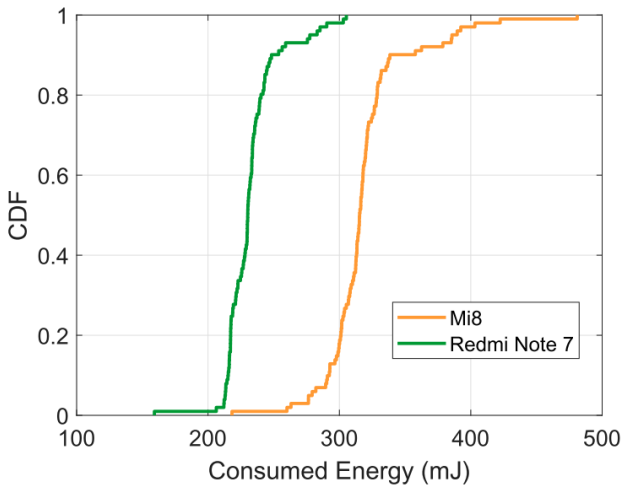


Figure 8.6: Energy consumption of a single (Redmi Node 7, green) and dual (Mi 8, orange) frequency receiver in two smartphone models [202].

During initial signal acquisition, a GNSS receiver consumes more energy than the subsequent tracking mode. Hence, the TTFF has a significant impact on the overall energy consumption. This is especially true for IoT applications with low update rates. In this case, virtually every localization attempt can be regarded as a first fix. Therefore, several energy-saving GNSS techniques are focusing on TTFF reduction, as discussed in Section 8.1.3. While the TTFF evaluation is discussed in more detail in Section 8.2.7, the focus here is on the energy consumption profiles of these novel techniques.

The first and most widely adopted energy-saving technique in GNSS receivers is duty cycling. By putting the receiver in sleep mode between location updates, the total energy consumption can be reduced significantly, especially in IoT use cases where a location update is only required every few hours, days, weeks or even months.

The widely adopted A-GNSS approach ensures all data needed to compute a location is present in the UE, successfully omitting power-hungry satellite communication to retrieve e.g., coarse location, time or ephemeris data. Several GNSS manufacturers provide a platform or service to download this data and send it to a UE, e.g., the AssistNow platform of u-blox. Furthermore, the integration of an assistance network and a GNSS receiver in an all-in-one SoC leads to a lower overall power consumption. A SoC integrating GNSS and NB-IoT consumes 50 mW for receiving and 1610 mW for transmitting, while the always-on-block consumes 15 µW and the sleep current is smaller than 10 µA at 3.8 V [203]. When using a 300 mAh battery, this results in a lifetime of 306 days for a daily uplink message, while the lifetime significantly decreases to only 15 days when an hourly location update is required.

Snapshot processing and cloud computing are two emerging techniques to reduce the energy

consumption of a GNSS receiver. They are especially of interest in case the GNSS receiver is connected to an LPWAN transceiver, as the latter is able to transmit snapshot data to a processing center for subsequent outsourced position calculation. Taking a snapshot of up to 25 ms with a cloud GNSS receiver is an order of magnitude more energy-efficient than a conventional A-GNSS receiver [194]. The snapshot receiver of Baseband Technologies lasts for 18 days to 1 year depending on the snapshot length, while a conventional receiver would only last for 2 hours on the same 10 mAh battery [204]. Finally, u-blox recently introduced their 'CloudLocate' service, offering a S-GNSS approach in which the receiver acquires a snapshot of a few seconds, performs some preprocessing steps such as the extraction of code phases, sends this information to the cloud and turns itself off. Designed for use cases with battery-operated devices with large power autonomy and internet connectivity, this approach performs well in terms of energy consumption, successfully filling the gap between traditional (A-)GNSS and GNSS-less positioning, as shown in Figure 8.7. According to u-blox, the additional power demand constitutes only 10% of the total UE power consumption.

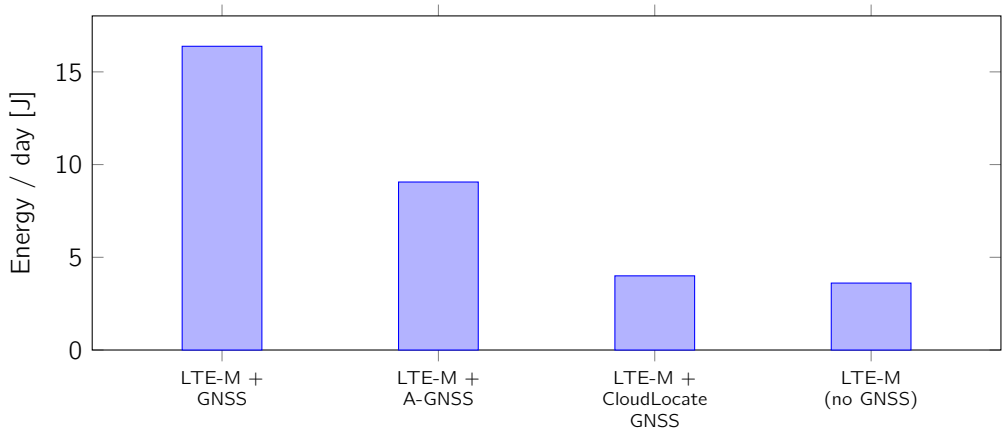


Figure 8.7: Total daily energy consumption comparison of different GNSS techniques, for a UE combining a u-blox M10 chip with LTE-M connection, with 6 location updates per day [205].

A white paper of the European GNSS Agency (GSA) describes the relative amount of energy saved with the aforementioned techniques, compared to a standard single-frequency GNSS receiver [206]. A-GNSS can be up to 10 times more energy-efficient, while snapshot processing and cloud computing can be 2-25 times more energy-efficient. For the latter category, higher energy efficiency is achieved when more location processing functionality is outsourced to the cloud. Finally, Figure 8.8 shows the relationship between the energy efficiency of each technique and the connectivity requirements of the terrestrial network.

8.2.3.3 LEO

While constellations such as Starlink are designed for broadband mobile Internet access, other LEO constellations are designed for low-power communication with terrestrial IoT devices. Despite many studies evaluating and improving the accuracy of LEO-based positioning systems, little attention has been paid to their energy consumption profile. Therefore, I now provide an overview of energy characteristics as specified in data sheets. Thus, it is im-

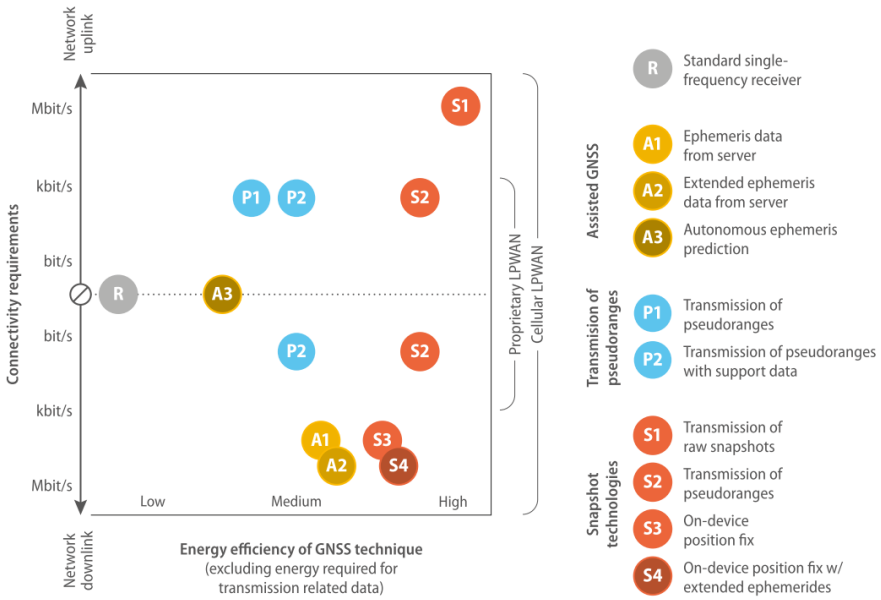


Figure 8.8: Relationship between connectivity requirements and energy efficiency of different GNSS techniques [206].

portant to keep in mind that the actual energy consumption of these systems is not widely evaluated yet.

The Argos system was introduced with the aim to serve environmental applications, including wildlife tracking and oceanography. The so-called Platform Terminal Transmitters (PTTs) are designed to have an autonomy of multiple years. An Argos transceiver has a typical transmission power of 500 mW, but this setting can be configured in a range from 250 mW to 2 W, which is equivalent to 24–33 dBm. A popular chipset is the ARTIC R2, which supports bidirectional communication and is compatible with the Argos-2, Argos-3 and the future Argos-4 system. On average, the chipset consumes 15–20 mA when receiving, 350 mA when transmitting and has an extremely low sleep current of less than 1 µA, at a supply voltage of 1.8 V or 3.3 V [207]. Arribada provides an Argos module integrating an ARTIC R2 transmitter and a GPS receiver and claims to achieve a 20 µA sleep current and 5 years of autonomy. Additionally, the module provides support for hybridization with a cellular or LoRaWAN daughter board [208].

Lacuna Space uses the LR1110 'all-in-one' chip from Semtech to perform Doppler positioning using LEO satellites. Even though no details about the actual power consumption of this localization technique are public, the energy consumption profile of the LR1110 for different modes is shown in Table 8.4, using an operating voltage around 3.3 V. Obviously, the power consumption depends on the bandwidth, SF and transmit power. Furthermore, the GNSS scanner typically needs to scan for 1–2 s, depending on the assistance data, and leads to a power consumption of 8.5 µWh for GPS. For the on-board passive Wi-Fi scanner, it takes 65–75 ms to scan 3 Wi-Fi channels and capture 6 MAC addresses, consuming 0.5–0.7 µWh.

Table 8.4: Energy consumption profile of the Semtech LR1110 chip [209].

Mode	**System component**	**Current consumption (@ 3.3 V)**
Receive (SF12, 125 kHz)	LoRa	5.7 mA
	Wi-Fi scan	3-11 mA
	GNSS scan	5-10 mA
Transmit (868 MHz, 14 dBm)	LoRa module	28 mA
Transmit (868 MHz, 22 dBm)	LoRa module	118 mA
Sleep (no RTC)	All	1.6 µA
Power down	All	0.8 µA

8.2.4 Positioning accuracy

In this section, I discuss the average, 3D RMS and 95th percentile of the difference between the estimated location and the ground-truth location. The lower this location estimation error, the higher the positioning accuracy of a technique.

8.2.4.1 LPWAN

Leveraging LPWAN communication to locate a device has been a popular research topic in recent years. While some approaches only aim to provide location awareness, i.e., a rough location estimate, other approaches try to improve the localization accuracy in order to attract more IoT use cases.

In general, RSS-based localization algorithms perform the worst in terms of positioning accuracy. Reasons for this are the high number of multipath effects (shadowing, reflections etc.) and signal interference. Moreover, the environment plays a significant role when judging the accuracy of RSS-based LPWAN localization. Although some studies categorize different environments into urban, sub-urban and rural areas, these terms are not clearly defined and ambiguous. This consequently leads to inaccuracies when applying signal propagation models in RSS ranging algorithms. Other popular algorithms range from simple proximity estimation to advanced Machine Learning and Neural Network-based fingerprinting. While the former has a typical localization error of several hundreds of meters to a few kilometers, the latter is able to locate a transmitter with a mean location error below 500 m. A benchmark of RSS-based ranging and fingerprinting algorithms using LoRaWAN is detailed in Chapter 5. The *k*NN and Random Forest algorithms yield the most accurate fingerprint-based results, while it was found that changing the path loss model in range-based approaches does not significantly impact the final location accuracy.

Timing-based approaches generally are more accurate as when compared to RSS-based algorithms. TDoA experiments in a public LoRa network resulted in a median and maximal location error of 150 m and 350 m, respectively [79]. However, TDoA requires accurate synchronization between gateways and is therefore not feasible in some LPWAN technologies.

For instance, applying TDoA in the UNB technology of Sigfox is not feasible as the accuracy is directly proportional to the bandwidth (see Section 8.2.8). Furthermore, at least four nearby gateways need to receive an uplink message in order to estimate the location of the transmitter. For these reasons, it is not always possible to provide a TDoA location estimate. Lastly, the first OTDoA experiments in a laboratory environment report an RMS positioning accuracy of 48.5 m and 65.5 m in normal and extended coverage, respectively [91].

AoA systems using LPWAN signals are proven to accurately estimate the angle of arrival, with an error below 5 degrees, 80% of the time [93]. The combination of AoA and RSS in NB-IoT is very welcome, as a lot of NB-IoT modules and networks only report the currently serving cell, instead of all nearby base stations (see Chapter 7). Furthermore, the combination of AoA and TDoA (also referred to as TDAoA) yields more accurate results, with a mean localization error of 159 m in a NLoS environment [96].

With the aim to compare localization algorithms and LPWAN technologies in a fair way, together with a colleague, I have investigated the accuracy of several LPWAN localization algorithms in the same urban environment in the city of Antwerp, Belgium (see Chapters 4, 5, 7 and [96]). The results are summarized in Figure 8.9. It is important to mention the difference in base station number and density. For example, while only a single NB-IoT base station is reported per measurement, some Sigfox messages were received by more than 40 gateways. However, the high density of the cellular NB-IoT network compensates for this fact, resulting in mean location estimation errors between 204 m and 340 m when applying fingerprinting and proximity algorithms, respectively. Despite several outlier detection algorithms, the outliers of LPWAN localization algorithms remain significant. Finally, the accuracy can be further improved by combining different techniques. Examples include applying Artificial Intelligence (AI) for optimized fingerprinting, estimating heights based on altimeters, and implementing road mapping filters.

8.2.4.2 GNSS

In high contrast to LPWAN localization techniques, GNSS techniques achieve much higher accuracies, up to several orders of magnitude. Important to note is that I evaluate the accuracy of Standard Positioning Service (SPS) GNSS receivers, as these are most common in IoT tracking devices. Hence, advanced positioning techniques such as RTK and PPP fall outside the scope of this discussion.

The GNSS technology report of 2020 lists a horizontal positioning accuracy of 5–10 m with a dual-frequency GNSS receiver, and a typical accuracy of 15–30 m for a single-frequency receiver [201]. These are typical accuracies however, and thus cannot always be guaranteed, i.e., in complex propagation environments such as high-speed moving trains, dense urban scenarios, tunnels and multi-story car parks [210]. Furthermore, pseudorange measurements are affected by various effects, such as errors in the satellite clock and orbit information, errors in the ionospheric or tropospheric models and multipath effects. The Dilution Of Precision (DOP) accounts for the error propagation and provides an indication for the accuracy of a location estimate.

The accuracy of consumer-grade GPS and Assisted GPS (A-GPS) receivers has been eval-

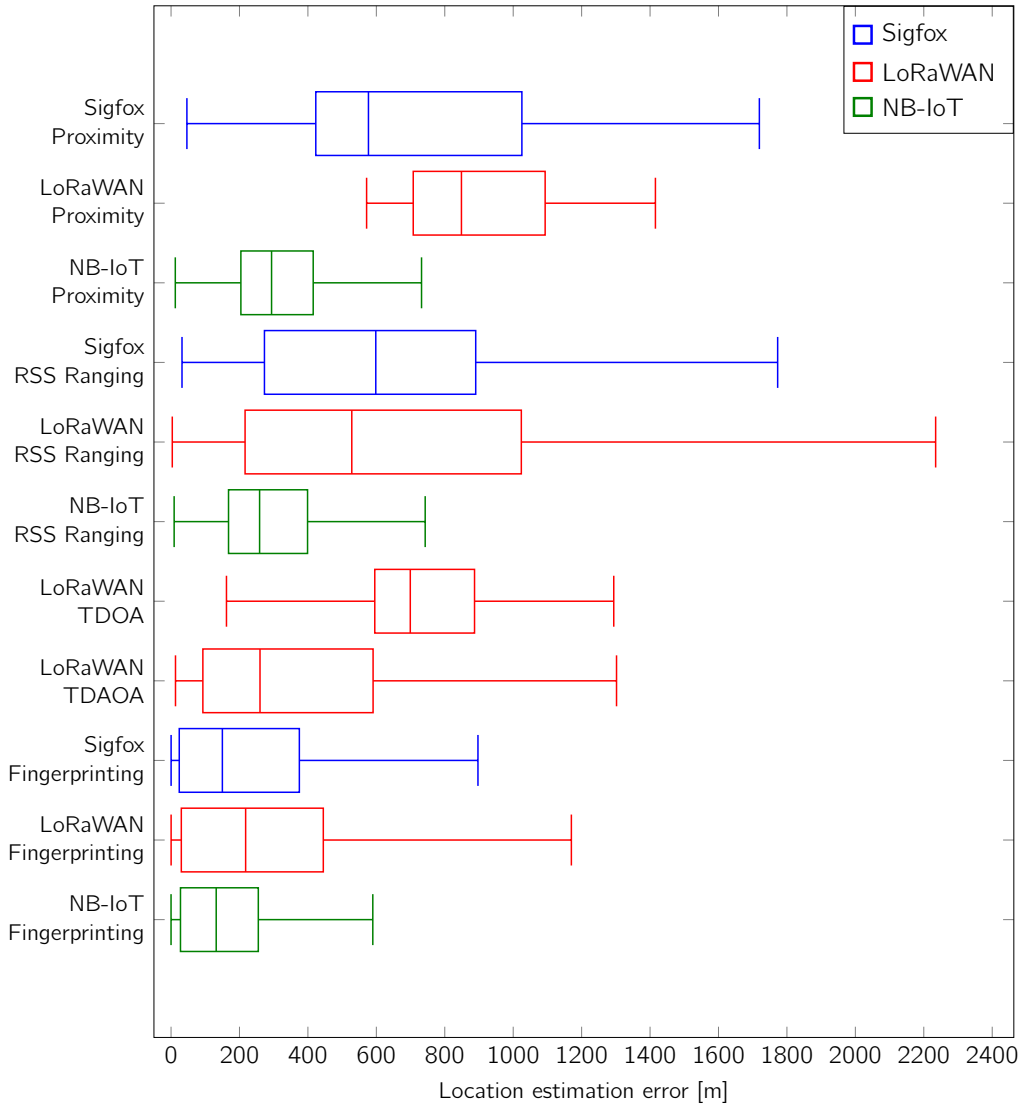


Figure 8.9: Positioning accuracy of LPWAN localization algorithms, compiled from Chapters 4, 5, 7 and [96].

uated in [211]. No remarkable differences were observed two minutes after the first fix. Nonetheless, in both approaches, the accuracy of the first fix is generally lower, due to the fact that a receiver initially has a fix on only 3-5 satellites. It was shown that with 95% probability, the first fix was accurate within 28–77 m, depending on manufacturer and type of the GNSS receiver.

When no terrestrial communication link is available, a GNSS receiver can autonomously predict ephemeris data. However, due to satellite orbit perturbation and environmental factors, ephemeris data is subject to change. This leads to a decrease in orbit prediction accuracy over time, which in turn leads to a reduced positioning accuracy.

S-GNSS and cloud processing techniques enable low energy positioning in return for a reduced sensitivity and accuracy [206]. Given the low energy consumption requirement, most S-GNSS receivers are using a single frequency, resulting in less accurate positioning, as mentioned before. Nonetheless, modern coarse-time navigation algorithms achieve an accuracy of a few meters from a one-shot position solution, depending on the length of the acquired signal and the accuracy of the code phase measurements. It was demonstrated that taking a 2 ms snapshot yields an accuracy below ± 1 m in north and east directions despite a reduced precision, as shown in Figure 8.10 for various snapshot lengths [212]. With a snapshot length of 1 s, the CloudLocate GNSS solution from u-blox achieves a median accuracy within 6 m [205]. Finally, it was demonstrated that a cloud GNSS sensor in an outdoor environment may offer the same horizontal accuracy as a conventional GNSS receiver, while consuming less energy. This relationship is visualized for different carrier-to-noise densities (C/N_0) in Figure 8.11 [194].

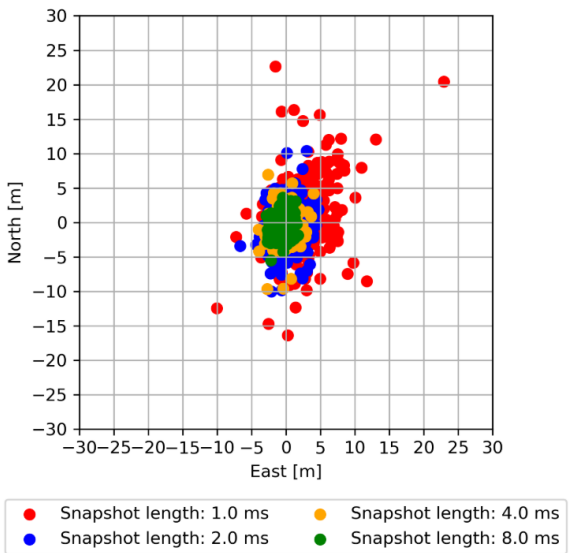


Figure 8.10: Time free pseudorange positioning results for various snapshot lengths, using GPS, Galileo and BeiDou signals [212].

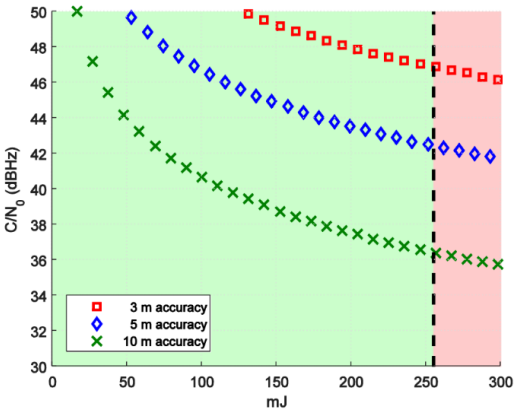


Figure 8.11: Expected energy consumption and accuracy of a cloud-based GNSS sensor under different C/N_0 environments [194].

8.2.4.3 LEO

The accuracy of positioning techniques using LEO satellites is studied to some extent in state-of-the-art literature. More specifically, the accuracy of Doppler positioning systems is being evaluated in a limited amount of test scenarios. Therefore, accuracy numbers may vary significantly.

Tan et al. proposed an instantaneous Doppler positioning solution using SoOP of Iridium NEXT satellites [187]. The inputs of the positioning system are the observed Doppler shifts and a precise orbit model. The SGP4 simplified perturbations model uses periodically updated Two Line Element (TLE) data obtained from the North American Aerospace Defense Command (NORAD). The proposed solution requires at least 4 LEO satellites in view and at least 6 Doppler measurements at different moments in time. The location of a static receiver was estimated in an open sky with a total of 7 satellites in view for 30 minutes. Using 25 different Doppler measurements in a Least Squares algorithm, the mean error in the east-direction is significantly higher than the mean error in the north-direction, as shown in Figure 8.12. The Doppler positioning was further improved by height aiding, resulting in a largest mean error of 46 m and 24 m in the east- and north-direction, respectively (also shown in Figure 8.12). Kalman filtering improves the accuracy even further, achieving a 2D position error of 22 m (1σ) for a static receiver in an open sky. This algorithm was also tested in a dense forest, which lead to a location estimation error of 108 m. Finally, it was concluded that satellites with high elevations and equally spaced velocity directions are more suitable for improving the horizontal positioning accuracy, while the subtracks of satellites at lower elevations should be used to restrict the vertical positioning errors.

Operating the Argos constellation, Kinéis claims to provide a native Doppler positioning accuracy of 150 m with the latest Argos satellites [208]. Optionally, users can combine this service with GNSS, in order to increase the accuracy if desired. More recently, the system can determine the position of a mobile transmitter using a single satellite. However, the resulting accuracy can vary from several hundreds of meters to several kilometers, depending on the number of transmissions during a single satellite pass. A complete and open manual

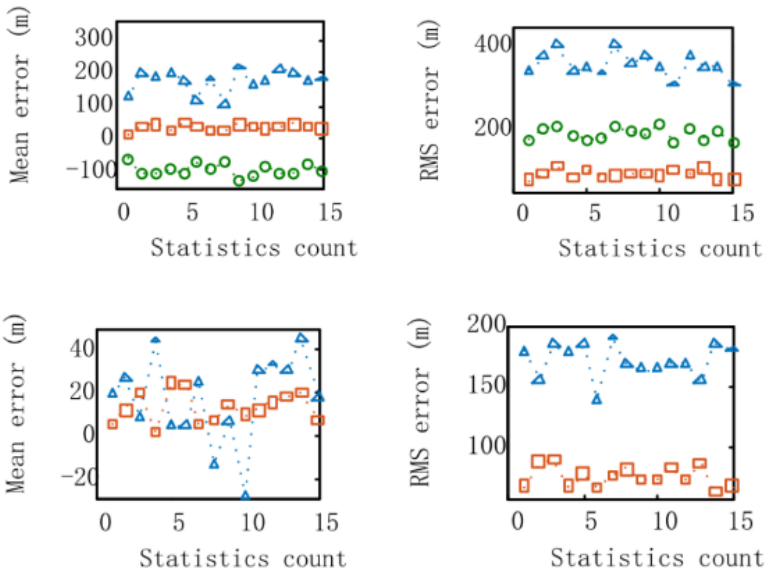


Figure 8.12: Mean and RMS errors of east (blue), north (red) and up (green) directions. Top: instantaneous Doppler positioning, Bottom: height aiding [187].

of the Argos positioning algorithms is available online [213]. Before 2011, a nonlinear Least Squares (LS) algorithm was used, which required at least 4 messages to get information about the accuracy of the position estimation. Since 2011, a new algorithm utilizing an EKF was introduced, which was proven more accurate and reliable [185]. Moreover, an error estimate can be provided through 7 so-called location classes, even with a single message per satellite pass. A location class is defined by the estimated positioning error of a measurement, as well as the number of messages sent during a satellite pass. A comparison of both LS and EKF algorithms in terms of location error and number of messages is shown in Figure 8.13. The reason for the biggest outliers is the incorrect choice between nominal and mirror locations. Moreover, both algorithms underestimate the positioning error due to non-normally distributed errors and changing frequency measurement noise, e.g., due to temperature changes. Furthermore, Lopez et al. developed an EKF algorithm that is able to switch between multiple motion models based on behavior (e.g., winter sleep versus hunt in animal tracking use cases), achieving higher accuracy.

Although Orbcomm does not provide an actual positioning service, their satellites can be used to perform Doppler-based positioning. Due to the lack of accurate LEO products and incomplete constellations, researchers found a significant gap between expected accuracy through simulations, and measured accuracy in real-life experiments: While simulation results show an 11 m accuracy with 25 LEO satellites over a period of four minutes, only 2 Orbcomm satellites are visible to the UE in reality, resulting in an accuracy of 360 m over a period of one minute [197]. When using the Starlink mega-constellation, the authors obtained an accuracy of 33.5 m and improved it to 7.7 m by adding an altimeter [190]. Thus, the accuracy of LEO positioning systems will increase as more constellations will be completed in the future.

Using the Iridium constellation of 66 LEO satellites, Satelles provides a globally operational PNT solution complementary to GNSS. The Satelles Time and Location (STL) service

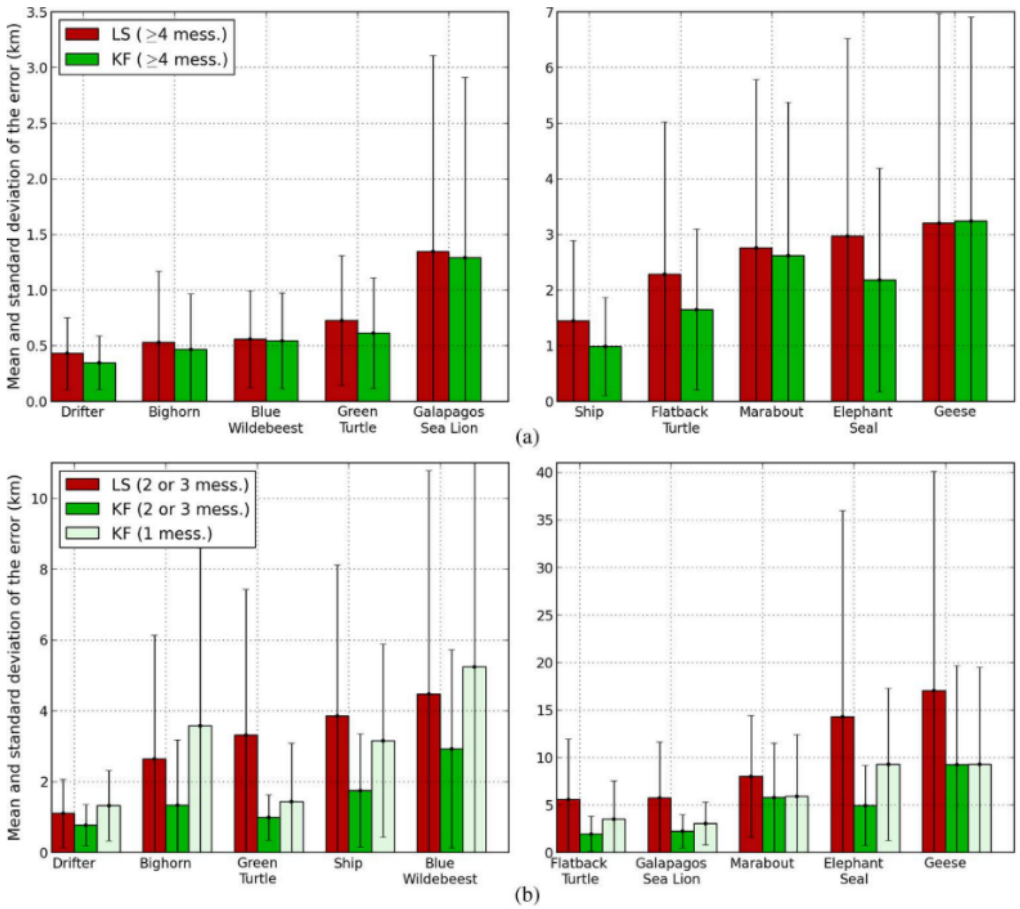


Figure 8.13: Mean and standard deviation errors for the Argos LS and EKF algorithms in different tracking uses cases (a) with at least 4 messages (b) with less then 4 messages [185].

mostly focuses on accurate timing requirements. While the STL data sheet does not specify an estimated position accuracy, a test performed by Satelles demonstrated an accuracy of around 20 m in deep indoor environments [29].

Reid et al. observed an interesting trend in positioning accuracy, depicted in Figure 8.14. By investing in new positioning infrastructure, the accuracy improves by an order of magnitude every 30 years. Following this trend implies that decimeter-level performance will be achieved by the mid-2020s [30]. Reid, co-founder and CTO of Xona Space Systems, predicts that LEO-based positioning could provide more than 10 times better accuracy and 100 times better interference mitigation compared to legacy GNSS. This enables applications such as autonomous driving, which require < 30 cm accuracy and very high reliability. Therefore, Xona Space Systems aims to provide Pulsar, a navigation system based on 300 dedicated LEO-PNT satellites, which should be operational in orbit by 2026. The system will employ signal ranging, similar to how GNSS works [181]. However, very few of the design parameters are made publicly available. Similarly, the Geely Technology Group is developing the GeeSpace system targeting the automotive industry. While the group aims

to offer centimeter-level positioning accuracy leveraging enhanced GNSS-based technology, the exact design parameters are also not yet available in the open literature [183]. Other players on the market include DDK Positioning and Trustpoint. They are in the process of designing PNT systems independent of GNSS, aiming to provide accurate, reliable, and secure LEO-PNT services.

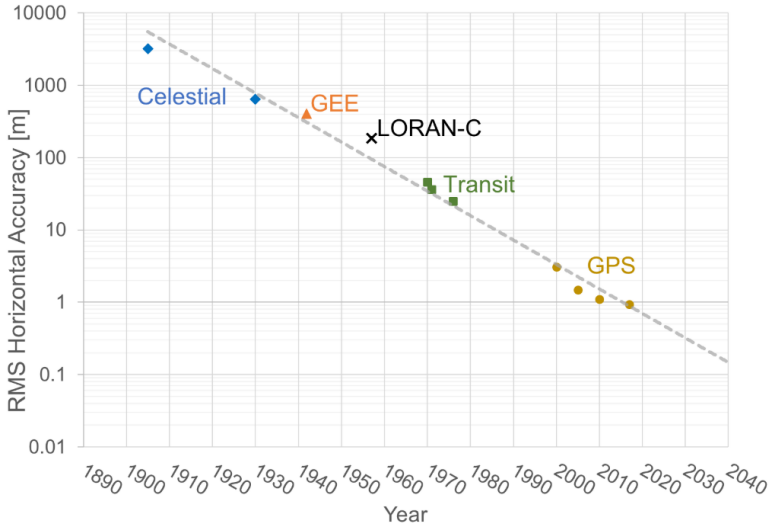


Figure 8.14: Trend in location accuracy of various technologies over the last century [30].

8.2.4.4 Combined approaches

It is common to combine or integrate the aforementioned standalone solutions, often resulting in an increased accuracy.

A first example is the solution provided by Lacuna Space, which comes in many flavors depending on customer demands. When accurate positioning is required, GNSS satellites are used to locate the UE and LEO satellites are only used to communicate the estimated position to the customer, if no LoRaWAN network is available. If a LoRaWAN network is available, the UE can also be located using the aforementioned LoRa Edge geolocation solution, which is less accurate but saves more energy. In addition, if no terrestrial network is available and battery lifetime is important, Lacuna Space offers a Doppler-based positioning service, along with their LEO satellite communication. Thus, the trade-off between accuracy and energy consumption becomes clear. For the native Doppler positioning system, Lacuna Space itself currently reports an initial accuracy of a few kilometers, which will be improved with more satellite passes. To overcome this issue, Wi-Fi and GNSS scanning are integrated [196]. Wi-Fi scanning is used for indoor positioning where satellite signals do not reach and has a typical accuracy of 30 m. Experiments from Irnas show that the passive GNSS scanner of the Semtech LR1110 leads to an average accuracy of around 30 m after 4.5 s, despite some outliers of around 100 m [214].

Aiming to track whales in oceans for a long time, LEO Doppler measurements and FastLoc

GPS experiments were carried out in [179]. While the former is used to communicate the observables and provide a rough location estimate, the latter is used to determine a more accurate GPS location. Hence, the feasibility of combining LEO and MEO satellites was demonstrated.

Finally, it was shown that LEO constellations can be used as a backup positioning system for GNSS. It was experimentally demonstrated that LEO Doppler measurements can reduce the position error of INS from 31.7 m to 8.8 m, 30 seconds after GNSS signals became unavailable [189]. The authors elaborated on this by evaluating different satellite propagation models and comparing three navigation frameworks. The combination of a LEO-aided INS Simultaneous Tracking And Navigation (STAN) framework and a two-body model with second gravitational zonal coefficient J_2 results in a 3D-Root Mean Square Error (RMSE) and final position error of 5.3 m and 5.4 m, respectively [192].

8.2.5 Ubiquity of coverage

The ubiquity of coverage indicates the availability of a positioning system on Earth and is measured in a quantitative (e.g., 90% worldwide) or qualitative (e.g., deep indoor) way.

8.2.5.1 LPWAN

The coverage of LPWAN technologies has been studied extensively, both in simulation and real-life environments. With the aim to evaluate which technology provides the best coverage for IoT devices, a simulation study for Sigfox, LoRa and NB-IoT was carried out in a 7800 km^2 area [215]. The results show that NB-IoT provides the best coverage, even in deep indoor environments, with a Maximum Coupling Loss (MCL) of 164 dB. Moreover, the cellular technology was proven to have the smallest outage probability when the inter-site distance is equal, followed by Sigfox and LoRa, respectively. A similar study revealed that NB-IoT outperforms LoRa in terms of coverage, in both urban and rural environments. The main reason for this is the directivity of NB-IoT antennas, which provide a better coverage for devices farther away from the eNodeB but near the main beam [114]. Besides the excellent performance in outdoor environments, extensive measurement campaigns confirm the deep indoor coverage of NB-IoT provided through existing LTE infrastructure [216].

In general, LPWAN technologies are able to provide excellent coverage in environments where terrestrial-based infrastructure is installed. Due to the presence of the communication signal, localization of the transmitter becomes possible within the same range. Despite the rapidly increasing number of mobile IoT networks, in some countries and especially in the continent of Africa, there is no cellular LPWAN connectivity yet, as shown in Figure 8.15. Furthermore, positioning algorithms such as TDoA require a minimum number of receiving gateways. Therefore, not all LPWAN positioning techniques can be applied in any environment with coverage.

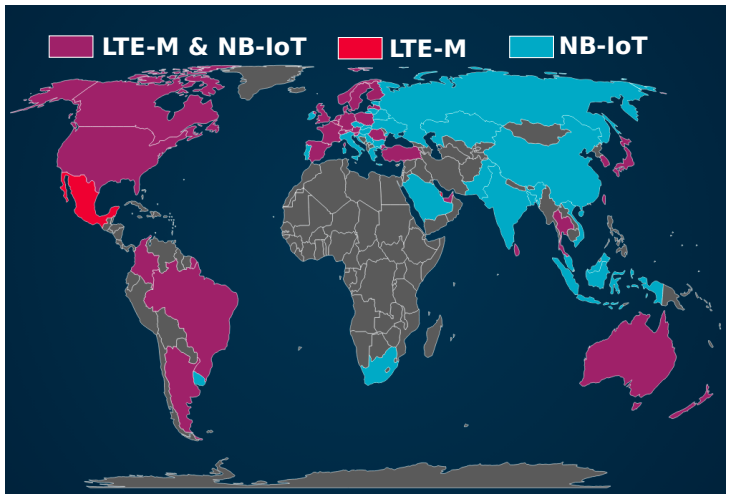


Figure 8.15: Coverage of cellular IoT networks worldwide [217].

8.2.5.2 GNSS

In opposition to LPWAN transmitters, GNSS receivers usually benefit from global coverage, due to the combination of multiple (both global and regional) constellations. Figure 8.16 shows the evolution of the total number and distribution of satellites across each global GNSS constellation [201].

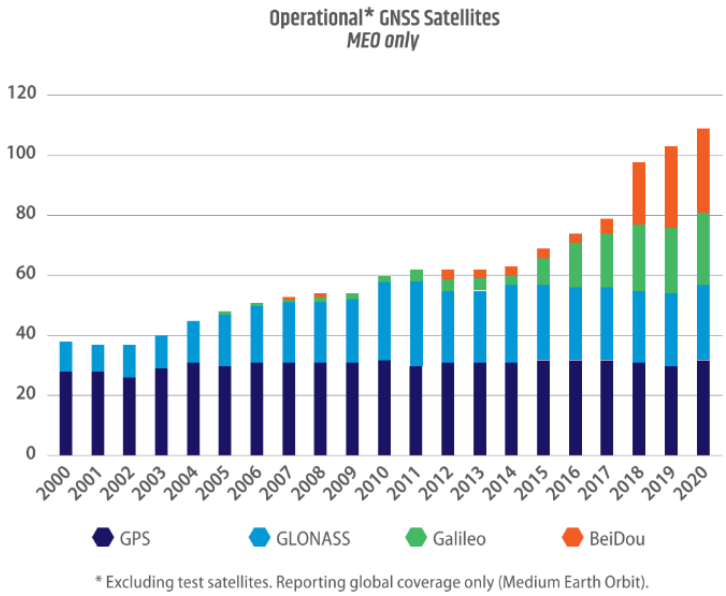


Figure 8.16: Evolution of number and distribution of operational GNSS satellites [201].

Despite their ubiquitous worldwide coverage, GNSSs have two main limitations in terms

of coverage: signal blockage and multipath effects. First, GNSS signals do not penetrate well through walls, reducing the number of satellites in view, making positioning difficult or even impossible in indoor and underground environments. Second, NLoS propagation causes multipath effects, which impact the observation quality and thus the positioning performance, especially in dense urban environments.

Aside from faster position fixes, A-GNSS improves the performance in difficult indoor and urban environments thanks to the increased receiver sensitivity [206]. However, A-GNSS and S-GNSS techniques can only be applied in areas where a communication network is available.

8.2.5.3 LEO

In comparison to GNSS satellites, satellites in lower Earth orbits provide a smaller coverage, as the smaller distance to Earth decreases the size of the satellite footprint. More specifically, LEO satellites are generally placed at altitudes below 1000 km, which is around 20 times smaller than the altitude of GNSS satellites, leading to a significantly smaller footprint. For example, Figure 8.17 shows the Iridium NEXT constellation which consists of satellites at an altitude of 780 km. The resulting satellite footprint has a diameter around 3000 km, whereas a GPS satellite covers an area of 12,000 km in diameter [178]. Thus, many more satellites are required to cover the Earth in LEO than in MEO. In order to provide global coverage with only a few LEO satellites, most satellites follow a polar orbit, i.e., flying over the North and South poles. As the Earth rotates in the meantime, a single satellite will eventually map out the entire globe without blind spots. Therefore, there is a higher satellite coverage near the poles than at the equator. Although a single revolution around the Earth only takes about 100 minutes, there is no permanent coverage everywhere on Earth. To solve this problem, several companies such as Argos, OneWeb and Starlink are developing constellations of tens, hundreds and even thousands of satellites. With more satellites in view, more accurate positions can be computed. Additionally, a higher location update rate can be achieved, as discussed in Section 8.2.15.

Due to the closer distance to Earth, LEO satellite signals experience less path loss and deliver more robust signals than GNSS signals, making them more suitable in difficult to reach environments. A perfect example of the use of LEO satellites can be found in the Automatic Identification System (AIS) for the tracking and monitoring of vessels. It was found that when using LEO satellites instead of GNSS satellites, a stronger signal was obtained at the AIS receiver and coverage was extended [219]. Moreover, LEO satellite signals are able to penetrate better in indoor environments. Tan et al. claim that LEO satellite SoOP can work in severe environments such as when rushing to deal with an emergency, fire control inside deep buildings and in combat [187]. In contrast, companies providing LEO services experience limited or precluded coverage in indoor environments. To solve this coverage issue, Lacuna Space is experimenting with the integration of Wi-Fi scanning for indoor positioning. Once an indoor position is determined, it can be sent to a LEO satellite via an outdoor gateway. A final strategy is to wait for certain coverage conditions, such as the delayed transmission of Argos messages when a whale equipped with a UE surfaces after a dive [179].

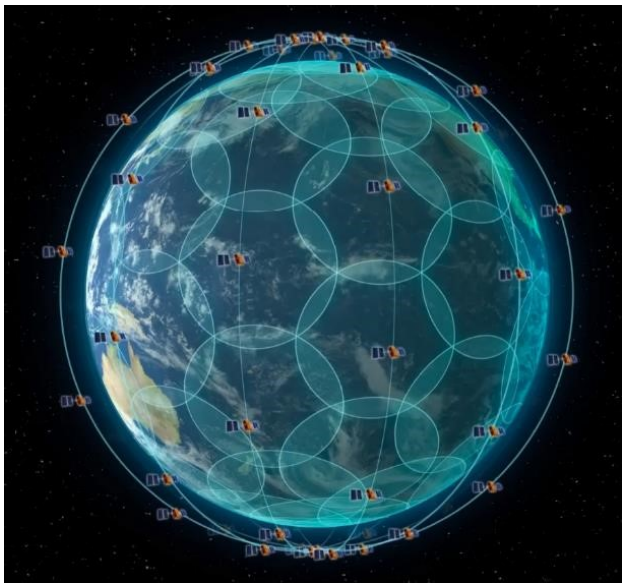


Figure 8.17: The 66 LEO satellites of the Iridium NEXT constellation with their footprints on Earth [218].

8.2.5.4 Combined approaches

Aiming to track battery-constrained IoT devices, LEO satellites can successfully extend the coverage of a terrestrial NB-IoT network [69]. The integration of LEO and LPWAN becomes more popular, as this combination of technologies provides truly global and deep indoor coverage. Companies integrating LEO and LPWAN solutions include Lacuna Space (LoRa + LEO), Hiber (LoRa + LEO), Wyld (LoRa + LEO), OmniSpace (NB-IoT + LEO) and Orbcomm (cellular + LEO). It should be noted that a roaming agreement needs to be in place with a public or private network where the device operates, unless it concerns an open public network such as TTN.

Finally, the combination of GNSS and LEO constellations leverages augmentation of GNSS for navigation. Moreover, LEO satellites can serve as a full standalone backup. Globalstar, for instance, provides UEs which are able to communicate with LEO and locate with GPS satellites, in order to enable near real-time tracking in e.g., mountainous areas where no terrestrial network is available.

8.2.6 Scalability

In 2021, more than 12 billion devices were connected to the Internet of Things, excluding computers, laptops, smartphones and tablets [1]. With the explosive growth in mind, modern communication and localization infrastructure requires a scalable design. In this section, I focus on the scalability of state-of-the-art positioning solutions towards billions of devices.

8.2.6.1 LPWAN

The scalability of LPWAN localization techniques depends on the technology and required network infrastructure. A capacity study of the Sigfox, LoRa and NB-IoT technologies is conducted in [220]. Capacity experiments were carried out in a 8000 km^2 dense urban environment in Denmark, with a varying number of IoT devices per person. The results show that NB-IoT outperforms Sigfox and LoRa, with an uplink failure probability below 4% in the 95th percentile, with 10 devices per person. Reasons for this are the superior coverage of NB-IoT, the use of link adaptations and a licensed band, which in turn leads to less duty cycle violations and interference. Moreover, NB-IoT benefits from the scalability of existing LTE infrastructure, currently supporting up to 200k devices per cell with respect to Quality of Service (QoS) [221]. This number will only increase with upcoming 5G networks. In contrast, despite LoRa gateways serving up to 50k LoRa end devices [34], the scalability of LoRa networks has been questioned in literature. Main drawbacks are the higher chance of collisions due to the use of unlicensed bands and the longer airtime at higher SFs [222]. Moreover, the use of different carrier frequencies in different regions (e.g., 868 MHz in EU vs. 915 MHz in US) makes it a challenging task to track an asset worldwide. As discussed in Chapter 6, Semtech addressed this issue by introducing LoRa at the worldwide available 2.4 GHz frequency band.

Aside from the underlying technologies, the hardware in both end devices and gateways impacts the scalability of a certain positioning technique. While any LoRa network can be used for RSS-based ranging or fingerprinting localization, implementing TDoA-based ranging requires accurate clocks in the surrounding gateways, as well as synchronization between them. Similarly, AoA-based approaches require gateways equipped with a more complex antenna array. Deploying such features in a large-scale network is a costly and time-consuming task for network operators. Therefore, TDoA and AoA are considered less scalable than RSS in the performance comparison matrix of Table 8.1.

8.2.6.2 GNSS

GNSS in general is considered highly scalable, as it is a broadcast system which can be used by an infinite number of users. In the past decades, GNSS systems have been scaled already, which is why no recent literature is questioning the scalability of the positioning system anymore. Hardware is inexpensive, widely available and produced in large numbers. The system is used by millions of users on a daily basis. Moreover, the integration of A-GNSS in our smartphones has become indispensable. Obviously, the scalability of A-GNSS and cloud processing techniques also depends on the scalability of the used communication network.

8.2.6.3 LEO

Due to relatively inexpensive nanosatellites or cubesats and decades of satellite technology advancements, a plethora of companies are currently deploying LEO constellations on a very large scale. In fact, one of the main reasons for the high interest in large-scale deployments is

the capacity of LEO constellations. The more satellites in a constellation, the more satellites are in view and the higher the capacity of the system [178]. However, the high number of satellites requires a scalable constellation design. A framework to identify an optimal design for a constellation of cubesats and for different use cases is developed in [172].

With only 8 operational LEO satellites, the Argos system is currently serving 22,000 active transmitters per month, spread over 100 countries. In addition, more than 60 ground stations worldwide and 2 data processing centers ensure the scalable delivery of a location estimate. The upcoming Kinéis constellation of 25 nanosatellites will further increase the system capacity [223]. Finally, Lacuna Space aims to provide a near real-time service with 240 satellites in orbit. The first step is to launch 24 cubesats by 2023. From the launch of their commercial service onward, the constellation will continuously expand. In addition, better revisit times and more capacity will be provided as per market demand.

8.2.7 TTFF

The TTFF is of high priority in GNSS, as it significantly impacts the duration for which the components needs to be powered and thus energy consumption. For other systems and use cases, this dimension might be of lower priority. Even though TTFF is mostly used in GNSS terminology, I use the term here in a wider context to indicate the time it takes to obtain a first position estimate. This parameter is not only highly related to the energy consumption profile (see Section 8.2.3), but also important in low-latency and near real-time tracking applications.

8.2.7.1 LPWAN

In LPWAN positioning techniques, the TTFF can be seen as the sum of the time it takes to send an uplink message with observables, the Time of Flight, the time for nearby gateways to receive the messages and add metadata (e.g., timestamps for TDoA), and the time to send the message from the gateways via the operator's network and a localization server to the end user. In this summation, the wireless transmission obviously is the most time consuming, introducing latency in the communication and hence also in location updates. In LoRa technology, the SF determines the transmission speed. For instance, sending a LoRaWAN packet using SF8 takes twice as long as when using SF7. Duty cycle regulations however limit the amount of airtime in unlicensed bands. For example, using the EU 868 MHz band limitations, the maximum payload size and 250 kHz bandwidth, the maximum airtime equals 3608.6 ms [156]. While NB-IoT transmitters do not have to cope with these limitations, the time it takes to get location information depends on various parameters and sources of latency, including the time to wake up the UE from PSM and synchronize with the network, payload size, subcarrier spacing, multi-tone capability, the chosen resource unit, the number of repetitions and the efficiency of the location estimation algorithm. This results in an absolute minimum transmission time of 1 ms, and a worst-case latency of 40,960 ms [224]. In general, empirical experiments teach us that LPWAN location estimates can be produced within a few seconds after transmission.

8.2.7.2 GNSS

There are three requirements for a first fix in GNSS positioning: signal acquisition, availability of ephemeris, and availability of a precise time of week. The TTFF of a conventional GNSS receiver can be split into the receiver warm-up time, the acquisition time, the settling time for code and carrier tracking, the navigation ephemeris read time, the time to retrieve the system time reference and the time to compute the navigation solution [225]. Information available to the receiver at the start-up will influence the time spent at these different stages, especially acquisition and navigation ephemeris read time. The need to update this information depends on how good the information can be maintained, i.e., the validity of the ephemeris or almanac, and accuracy of the real-time clock. When there is no information available (i.e., cold start) the receiver has to go through all the stages mentioned before and has to search the full frequency-code delay search space during acquisition. This can take up to several minutes. When there is coarse information on the position, the time, the frequency and the satellite positions (e.g., based on the almanac), the search space can be constrained leading to a decreased TTFF of around 30 s (i.e., warm start). In case of accurate knowledge of all the factors, the TTFF can be reduced to 1 s (i.e., hot start). The TTFF is also influenced by the environment, especially during a cold start. Weak signals or signal blockage lead to a longer TTFF because of possible data bit errors, which extends the navigation data read time considerably [193]. The usual estimations from chipset manufacturers for the TTFF are 30 s for a cold start and 1 s for a hot start.

Several techniques have been developed to reduce the TTFF. Through an external communication link, an A-GNSS receiver retrieves aiding information such as ephemeris, almanac, satellites status, precise network timing, and a coarse approximate position based on e.g., network cell ID. Several chip manufacturers, commercial service providers and scientific organizations have implemented A-GNSS services and platforms to ease the provision of aiding information.

Alternatively, a receiver can implement self-assistance. This technique is similar to conventional A-GNSS services, but aiding information is computed at the UE based on information received in the past. A UE can compute ephemeris with a validity interval up to several weeks which is much more accurate than conventional almanacs. In order to compute orbits predictions, the UE has to integrate satellite motions by modeling all forces acting on GNSS satellites [226]. Although this process can be power consuming, self-assistance can reduce the TTFF by a factor of up to 5 times [227]. More recent studies show how neural networks are able to more accurately predict orbits with less computational effort [228].

Coarse-time positioning circumvents the need for decoding the time of week, maintaining the time by an accurate receiver clock or receiving time assistance from A-GNSS (better than 1 ms) and thus reduces the time by the waiting time for decoding the time of week in case of coarse (seconds to minutes) time knowledge [193]. Using A-GNSS or self-assistance together with coarse-time positioning, the TTFF can be reduced by constraining the search space during acquisition, which reduces the time and improves the sensitivity, by eliminating the need to decode satellite ephemeris and the need to wait for decoding the satellite time of week.

Finally, cloud-based S-GNSS techniques do not require any of the aforementioned infor-

mation at the receiver side. Observables are sent to the cloud, where the latest aiding information is widely available.

8.2.7.3 LEO

In satellite IoT positioning applications, the TTFF can be defined as the sum of the following terms:

1. The difference in time between the request of a position update and the passing of a satellite above the UE,
2. The time for a UE to reach a satellite (transmission time),
3. The time to relay the message from the satellite to a ground station,
4. The time to calculate a position in a processing center.

While the latter term is negligible when using enough processing power, the other three terms can have a significant impact on the total TTFF.

For the first term, the availability of LEO satellites plays an important role to assess the overall latency of the positioning system. Because of the relatively small footprint of LEO satellites (see Section 8.2.5), a UE often has to wait to transmit an uplink message until a satellite passes. Therefore, a first Doppler location estimate can only be produced when a satellite is in view. For the Iridium constellation of 66 satellites, each satellite orbits the Earth about every 100 minutes. This subsequently results in an average satellite revisit time of around 9 minutes, provided that the UE did not move. Thus, in the worst-case scenario, the satellite revisit time of an Iridium transmitter can increase up to 9 minutes [229]. Obviously, the satellite revisit time reduces by adding more LEO satellites to the constellation, which is an ongoing task in many constellations. For more information about the satellite revisit time, I refer to Section 8.2.15.

The second term refers to the total transmission time through space. LEO satellites are closer to Earth and therefore introduce a lower Round-Trip Time (RTT) as when compared to GNSS satellites. Even for the more uppermost LEO satellites at 2000 km, the RTT is only 13.3 ms [229]. According to Samsung, a LEO constellation below 1580 km has the potential to be faster than Earth-bound fiber optic networks [178].

The third term accounts for the time needed by the satellite to pass over a ground station and forward data such as Doppler measurements to it. In the case of Lacuna Space, this typically takes a few minutes, with a maximum delay of 12 hours. A latency of a few minutes will be guaranteed soon, when more ground stations are deployed around the world.

In summary, the TTFF of LEO positioning systems can vary significantly, depending on the number of satellites and ground stations worldwide. In the near future, this number is expected to rise in a rapid fashion, achieving a latency of a few minutes.

8.2.8 Data rate & bandwidth

Table 8.5 lists the bandwidth and data rate for each technology considered in this work. In this section, I discuss how these parameters influence the positioning performance. However, a lower or higher bandwidth and data rate are not necessarily beneficial or disadvantageous for the overall positioning performance or power efficiency, hence the gray colored (meaning not applicable) row in Table 8.1.

Table 8.5: Bandwidth and data rates for each positioning technology.

Category	Technology	Bandwidth	Data rate
LPWAN	Sigfox	0.1 kHz (UL), 0.6 kHz (DL)	100 bps (UL), 100 bps (DL)
	LoRa	125/250/500 kHz	300 bps - 50 kbps
	NB-IoT	200 kHz	200 kbps
GNSS	GPS	2.046 MHz (L1), 20.46 MHz (L5)	50 bps (L1), 50 bps (L5)
	Galileo	32.0 MHz (E1B/C), 24.0 MHz (E5a), 24.0 MHz (E5b)	125 bps, 25 bps, 125 bps
	GLONASS	1.022 MHz (G1), 1.022 MHz (G2)	50 bps (G1), 50 bps (G2)
	BeiDou	4.092 MHz (B1), 24.00 MHz (B2)	50 bps (B1), 50 bps (B2)
LEO	Iridium	31.5 kHz	4.8 kbps (current), <512 kbps (NEXT)
	Argos	110 kHz	4.8 kbps (UL), 124 bps (UL, VLD-A4), 400 bps (DL)
	Orbcomm	15 kHz	4.8 kbps
	Lacuna Space	*Unknown*	*Up to 20 50-byte uplinks per satellite pass*

8.2.8.1 LPWAN

Out of all LPWAN technologies, the UNB technology of Sigfox has the smallest bandwidth of only 100 Hz uplink and 600 Hz downlink. Such narrow-band signals are not feasible to perform TDoA positioning. Together with a data rate of 100 bit/s, only 12 bytes can be sent in a single Sigfox uplink message. Note that this is just enough to communicate a traditional GNSS position.

While the bandwidth and data rate are fixed in many LPWAN technologies, they can be configured in LoRa. By increasing the SF from 7 to 12, the data rate decreases and a longer airtime is required. On the upside, the signal becomes more robust against interference, leading to an extended communication range. Several tools exist to calculate the airtime depending on the number of input bytes and chosen configuration (i.e., SF, bandwidth and

region) [156]. The SF can be configured statically on the UE with the aim to optimize data rate, airtime and energy consumption. Alternatively, the SF can be chosen dynamically by the network through an ADR mechanism, which takes into account the Signal-to-Noise Ratio (SNR) and the number of gateways that received the most recent uplinks.

As specified in 3GPP Release 13, NB-IoT occupies a frequency bandwidth of 200 kHz, which corresponds to a single LTE resource block. Within a licensed frequency band, NB-IoT can be deployed 'in-band', in the guard band or as a standalone operation. In high contrast to Sigfox and LoRa, the maximum data rate of NB-IoT is 200 kbit/s and an unlimited number of messages can be sent, with a maximum payload length of 1600 bytes per message [34]. This subsequently enables the faster and unlimited communication of observables of e.g., S-GNSS receivers to the cloud.

8.2.8.2 GNSS

Depending on the chosen GNSS technique, there are different requirements for data download to the IoT device and upload to the cloud. Data to be downloaded include:

- GNSS almanac (i.e., coarse satellite orbit and clock information), used to improve the acquisition performance by constraining the frequency search space. The almanac is decoded from the GNSS navigation messages or received via external means.
- GNSS ephemeris (i.e., precise satellite orbit, bias and clock information), required to compute a position. Ephemeris can be directly obtained via the signal in space or via external means as well.
- Coarse position and fine time, either derived from a terrestrial network or obtained via an assistance service in the cloud.

There exist a variety of assistance data services, which differ in the volume of the data set and its validity period, which in turn determines the frequency at which updates have to be sent to the UE. Most services providing ephemeris data reach a validity period of several days with a few kB of assistance data [230, 209].

In the opposite direction, data to be transferred from the receiver to the cloud can include a time-stamped position, a full set of pseudorange and Doppler observations or a GNSS signal snapshot. For transmitting a time-stamped position, a minimum payload size of 20 B is required and can be extended with information on DOP and velocity. As an example for the size of an observation set, the Semtech LR1110 receiver transmits pseudorange and Doppler observations using 34 bits per satellite and a small header. For 20 satellites this would be equivalent to around 85 B.

The size of a GNSS snapshot should be based on a trade-off between the snapshot length, the sampling frequency and the number of quantization levels. The higher the sampling frequency the better the resolution of the code phase and thus the better the resulting positioning accuracy. However, a complete code length of the replica is favorable in order to ease the reconstruction of the full pseudorange. The snapshot length L as well as the

sampling frequency f_s define the data size S. For a complex signal the size of the data expressed in bit is defined by:

$$S = 2 \cdot Q \cdot L \cdot f_s, \tag{8.1}$$

where Q represents the quantization. For example, a 20 ms snapshot using a sampling frequency of 20 MHz and 8-bit quantization results in a complex signal (I and Q) of 800 kB. If the sampling frequency is reduced to 4 MHz, only 160 kB needs to be communicated. The relationship between snapshot length, sampling frequency and snapshot size is illustrated by Figure 8.18 for an 8-bit quantization.

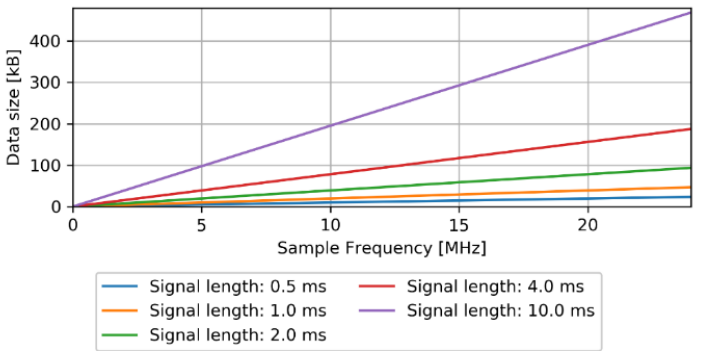


Figure 8.18: The relationship between snapshot length, sampling frequency and snapshot size for an 8-bit quantization [212].

8.2.8.3 LEO

Depending on the duration of a LEO satellite pass over a UE, a minimum data rate should be achieved, otherwise both uplink and downlink data can be lost. For example, in the cases of Lacuna Space and Argos, a passing satellite can only be reached for a duration of around two and ten minutes, respectively.

Argos uplink messages are sent at a data rate of 4.8 kbit/s. Each message contains a preliminary synchronization sequence, the total message length, transmitter identification number, user data and a checksum. Since the 3rd generation of the Argos system, two-way communication is supported. Downlinks sent at 400 bit/s are used for acknowledgement (ACK) of uplink messages, as well as for sending data such as timing information and satellite ephemeris data, in case the UE wants to locate itself or assist a GNSS receiver. Due to the introduction of ACK messages, redundant messages are no longer required. Moreover, the two-way communication enables customers to send commands to the UE. Such downlink messages may contain up to 128 bits by 8-bit increments. As an example, a command can be sent to the UE to increase the transmission frequency for a certain amount of time if more location updates are desired.

Apart from the aforementioned data rates, Argos has also developed a Very Low Data rate standard for the Argos-4 system (VLD-A4). This uplink standard has been designed for

very low-power transmitters (e.g., wildlife trackers) that transmit very small uplink messages. VLD-A4 has a modulated bit rate of 200 bit/s, which corresponds to a user bit rate of 124 bit/s. The message structure is shown in Figure 8.19. For a very short message containing only the 28-bit ID, the total transmission time equals 515 ms. A maximum of 56 bits of user data can be appended to this message, resulting in a transmission time of 965 ms. The message repetition period is minimum 30 seconds and can be configured by CLS according to the application and the geographical position of the UE. The VLD-A4 standard is also integrated in the ARTIC-R2 chipset. Finally, while the current bandwidth of Argos-3 equals 110 kHz, this number will be multiplied with a factor 8 in the future Argos-4 system of Kinéis [223].

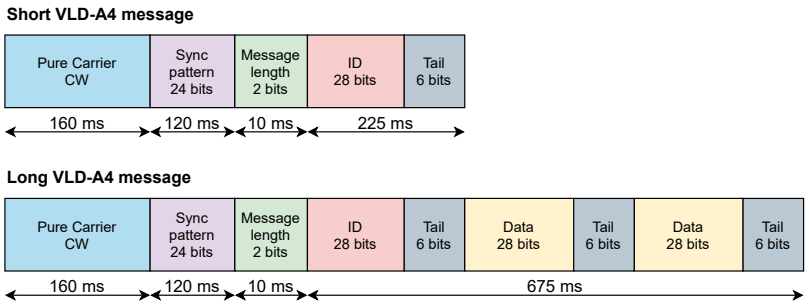


Figure 8.19: Message structure of the Argos VLD-A4 transmission standard.

Due to the way transmission works over a LEO satellite network, rather than throughput or data rate, it can be more useful to discuss message size and the number of messages per satellite pass. In the case of Lacuna Space, the maximum message payload size, i.e., excluding header data, varies between 45 and 125 bytes. The Long Range Frequency Hopping Spread Spectrum (LR-FHSS) modulation introduced for satellite IoT allows larger payloads at high power and low data rates. In theory, a UE can transmit 20 messages of 50 bytes (or equivalent longer messages) during a single satellite pass of 2 minutes. However, as in terrestrial LoRa networks, local operating conditions apply, limiting the airtime in several regions. This means that in Europe, for example, only 2 messages of 50 bytes or at most one 125-byte message can be sent per satellite pass.

8.2.9 Interoperability

The degree to which technologies can co-exist or even co-operate on the same UE without interfering with each other is referred to as the interoperability. In this section, I elaborate on the opportunities of combining LPWAN, GNSS and LEO technologies into a single solution, aiming to achieve a better overall positioning performance. A summary is shown in Figure 8.20.

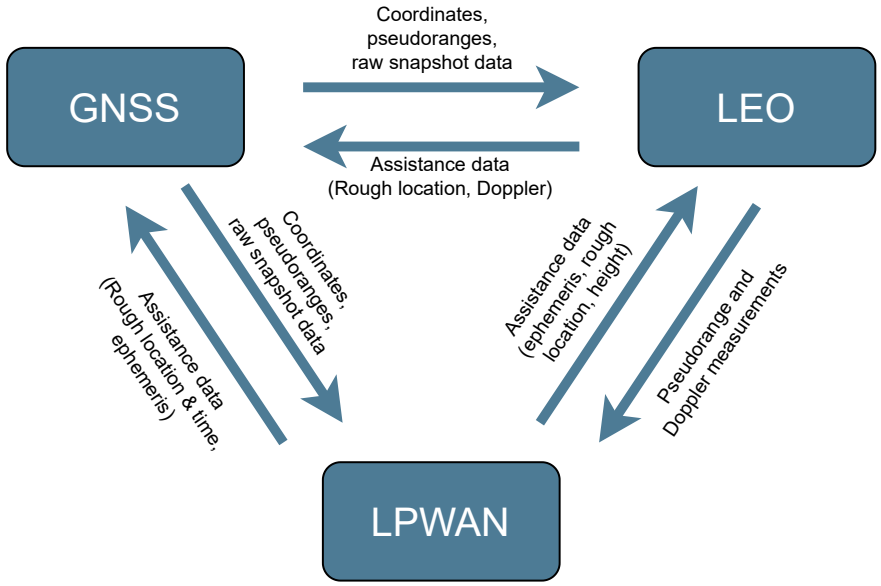


Figure 8.20: Interoperability of LPWAN, GNSS and LEO positioning systems.

8.2.9.1 LPWAN

A multiple Radio Access Technology (multi-RAT) UE can communicate with multiple low-power terrestrial networks. By implementing such a multimodal communication architecture, a UE is able to switch to an optimal localization method, depending on the context and constraints of the active wireless technology, as demonstrated in [72]. For example, a UE is accurately located on a construction site using TDoA in a private LoRa network, while during transport, RSS-based localization in a public Sigfox network is used to save energy.

8.2.9.2 GNSS

Augmenting GNSS positioning with LPWAN or LEO communication technologies creates a myriad of opportunities for energy-efficient positioning in an IoT context. A UE becomes capable to communicate a location estimate to the end user. Moreover, raw observables of S-GNSS receivers can be communicated to the cloud in order to save energy. In the opposite communication direction, assistance data can be provided to the GNSS receiver through a terrestrial communication link or via a LEO satellite network, successfully reducing the TTFF. Due to its effectiveness and efficiency, many LPWAN chip manufacturers provide built-in A-GNSS support. The combination of GNSS and LEO satellite communication is demonstrated by GlobalStar, which integrates LEO and GPS satellites for near real-time tracking in mountainous areas where no terrestrial network is available.

8.2.9.3 LEO

Terrestrial LPWANs can aid LEO-based positioning systems by communicating assistance data, such as a rough location estimate and orbital parameters of LEO satellites, as well as to exchange data from pseudorange and Doppler measurements. The integration of LPWANs for satellite IoT is being investigated by Lacuna Space, Fleet Space, OmniSpace, Hiber and many more companies. By doing so, the coverage of terrestrial LPWANs is extended by leveraging constellations of LEO satellites, resulting in a myriad of opportunities in the satellite IoT market. Potential use cases include global energy-efficient asset tracking and monitoring. Finally, LEO Doppler positioning can serve as a fallback solution for GNSS positioning in case the UE is located in a harsh environment where GNSS signals do not reach.

8.2.10 Communication of observables

A next dimension indicates whether there is a possibility to communicate the estimated position to the end user, or the need to communicate observables to a remote location processing system.

8.2.10.1 LPWAN

The approach in almost all LPWAN positioning techniques is to communicate one or more observables over the operator's network to a certain backend server. Observables in these uplink communications typically include RSS measurements, accurate timestamps, sensor readings etc. These features are either collected as nearby gateway metadata or message payload data and are forwarded over the Internet to a server acting as a localization engine, where the actual position of the UE is calculated. Finally, the location estimate is often visualized on a map in the end user application.

8.2.10.2 GNSS

In essence, observable-based GNSS techniques do not provide a means to communicate observables or a position estimate to an end user. Due to this self-localization, a GNSS tracker is often equipped with additional LPWAN or LEO communication hardware, as discussed in Section 8.2.9. Prominent examples of commercial services combining local and cloud processing while optimizing the amount of data to be exchanged are u-blox 'CloudLocate' [205] and Semtech 'LoRa Cloud' [231]. In the future, it is expected to see an increasing trend towards cloud processing and the integration of LPWAN and LEO in GNSS.

8.2.10.3 LEO

In contrast to GNSS, native positioning using LEO satellites leverages the communication link between the satellites and the ground stations to forward the location estimate to an end user application. For example in the Argos system, Doppler measurement data is obtained through the uplink connection to the satellite, relayed to receiving stations on Earth and forwarded to Global Processing Centers (GPCs), where the location is calculated and distributed to end user applications. Alternatively, this data can be stored in the memory of Argos chipsets for later retrieval, successfully saving energy and increasing battery lifetime. Similarly, satellites from Lacuna Space send frequency and timing information to a solver at the Lacuna Space backend, where the geolocalization is performed.

8.2.11 Index of technology readiness and maturity

Both academic and industrial research introduce novel positioning technologies and improve existing algorithms to push the state-of-the-art forward. However, every technology or algorithm faces challenges in terms of design, implementation, production and large-scale adoption. These challenges are included in the index of technology readiness and maturity, also referred to as the Technology Readiness Level (TRL).

8.2.11.1 LPWAN

Communication through LPWAN has been widely adopted, with many studies evaluating the performance and applications demonstrating the possibilities [232, 14]. Among the first major long-range energy-efficient networks, Sigfox and LoRa(WAN) have proven their excellent communication performance for more than a decade. Since 3GPP Release 13 in 2016, the NB-IoT standard has contributed significantly in the cellular IoT market. However, localization with LPWAN is not as mature as LPWAN communication. In the early years after the release of the first LPWAN technologies, mostly academic research was devoted to RSS-based and timing-based positioning algorithms with these novel technologies, in order to remove the need for a GNSS receiver in low-power IoT applications. Nowadays, several industry leaders provide a cloud positioning solution, e.g., the LoRa geolocation service of Semtech. While the standardization of e-Cell-ID and OTDoA in NB-IoT is finalized, the OTDoA feature is not widely available yet, as it requires an upgrade of the NB-IoT base station network as well as new network components, which is a challenging and costly task for network operators.

8.2.11.2 GNSS

Since the development of GPS by the US Department of Defense in 1967, GNSS technologies have evolved significantly. GNSS receivers nowadays are highly available and inexpensive, and the satellite constellations are publicly accessible at no charge. As duty cycling and A-GNSS are highly integrated in modern GNSS chipsets, these GNSS techniques are considered

highly mature. Snapshot and cloud processing algorithms, contrarily, are currently gaining popularity and completing the breakthrough in the IoT market. The fusion of these GNSS techniques with LPWAN technologies results in ubiquitous, energy-efficient and accurate positioning applications. As the transformation to cloud processing is currently ongoing, these positioning techniques are not considered highly mature yet, but it is expected that they will be widely adopted in the near future.

8.2.11.3 LEO

Similar to LPWAN, communication using LEO satellites can be considered mature, while satellite IoT localization is still in its infancy. A myriad of companies already provide satellite communication for decades. For instance, mobile communication via Iridium satellites is provided since 1998. Furthermore, newer LEO constellations such as Starlink aim to provide global high-speed Internet communication. Only a handful of companies have the primary objective to provide native positioning and navigation for the satellite IoT market. The Argos system can be considered as the most mature among the currently existing LEO positioning solutions. Several Argos chipsets and development kits are commercially available, and the Doppler positioning algorithm has been evaluated and optimized a few times. Nevertheless, as in many cases, the LEO constellation is not completed yet. Therefore, there is no 24/7 coverage anywhere on Earth yet. However, this will change rapidly as thousands of LEO satellite launches are planned in the coming years.

8.2.12 Standardized or proprietary

This section briefly specifies whether a protocol or technology has been standardized or made proprietary. This dimension does not influence the positioning performance, hence the corresponding gray row (meaning not applicable) in Table 8.1.

8.2.12.1 LPWAN

In the category of LPWAN technologies, there is a clear distinction between standardized and proprietary technologies. On the one hand, cellular technologies such as NB-IoT and LTE-M are based on 3GPP standards. The specification of these licensed technologies was frozen in Release 13 and is continuously updated in next releases, adding new features such as e-Cell-ID and OTDoA positioning in Release 14 [233]. On the other hand, several proprietary LPWAN technologies have arisen. Sigfox creates its own devices which are basically a 'black box' for end users. In the case of LoRa, only the signal modulation has been made proprietary, while LoRa end devices are licensed by Semtech and can be developed by other manufacturers. Moreover, LoRa standards are created and improved by the LoRa Alliance.

8.2.12.2 GNSS

As a mature positioning technique, GNSSs have been standardized in different specifications such as the GPS SPS performance standard. In order to cope with messages originating from multiple constellations on a single UE, the International GNSS Service (IGS) introduced a Receiver Independent Exchange Format (RINEX) for raw satellite navigation data [234]. While A-GNSS is a well-established technique described in several standards [235], the concept of S-GNSS has not been fully standardized yet. Ad-hoc methods of digitized data formats do not encourage interoperability and therefore need standardization. To this end, the Institute of Navigation (ION) aims to develop a specification for standardized metadata and formats. Adoption of this standard both by the data collection hardware and the SDR receiver would enable an SDR to process data from multiple sources seamlessly [236].

8.2.12.3 LEO

The rather immature positioning techniques such as Doppler positioning using LEO satellites have not been standardized yet. With the aim to speed up the development and integration with other services, several industry leaders decided to move towards an open-source system. For instance, the reference design and specification of Argos transceivers is fully open-source, as well as the antenna reference designs. Lacuna Space will also open-source the design of their devices, helping customers with the integration of other sensors into their system.

8.2.13 UE cost

The UE cost is defined as the total cost to use a certain positioning technique, including the cost for hardware, network access and positioning services. This section is highly related to Section 8.2.6, as the overall cost evolves when scaling up to billions of devices.

8.2.13.1 LPWAN

In 2020, the average terminal cost for Sigfox, LoRa and NB-IoT was $2-3 [9]. These ultra-low device costs are due to the massive number of IoT devices worldwide and the rapidly increasing interest to connect nearly everything to the Internet. Additionally, in the case of cellular IoT, a Subscriber Identity Module (SIM) or e-SIM is required in order to register to the network. The subscription cost varies per region and operator. As an example, the average cost of a data SIM card from a Belgian NB-IoT network operator equals €3 and the message transmission fee is in the order of €1/MB per SIM. Similarly, LoRaWAN end users need to pay a subscription cost, which is determined by the network operator and may include services such as the LoRa Geolocation Application Programming Interface (API). The subscription allows the end user to send a limited number of messages over the LoRa network each month. If this number is exceeded, an additional cost will be charged, depending on the policies of the operator. Alternatively, some LoRa operators such as TTN offer a connectivity service with fair-use policy free of charge. Lastly, AoA-based positioning

does not impact the overall UE cost, as the antenna arrays are placed at the gateway side, which can be low cost units as well [95].

8.2.13.2 GNSS

The price of a state-of-the-art GNSS receiver varies significantly depending on its capabilities. As an example, the u-blox MAX-M10S is a single-frequency multi-constellation GNSS module designed for low-power IoT applications and costs $21 when buying less than 10 units. Support for A-GNSS is often provided at almost no extra cost, as it is commonly integrated in nearly all modern GNSS chipsets. Moreover, it was shown that by integrating a GNSS receiver with cellular IoT connectivity on a single chip, the cost of the Bill Of Materials (BOM) can be reduced significantly [203]. Although GNSS hardware is more expensive than LPWAN hardware, no additional service cost needs to be paid in order to use the satellite broadcast systems. Furthermore, GSA and Ubiscale aim to deliver a Galileo service that integrates NB-IoT in the 'Galileo-of-Things' project, enabling lower chipset cost [201, 237].

Snapshot processing GNSS receivers can be less expensive than an ordinary GNSS receiver, as the acquisition of GNSS signals only requires an RF frontend [195]. Therefore, an RTL-SDR dongle of a few euros can be sufficient, while more optimized GNSS-specific frontends and SDRs can be rather expensive [238]. Finally, it is worth mentioning that the actual position computation in the cloud also comes at a cost. In an analysis of cloud GNSS approaches, the economic cost of the required cloud resources was studied. For a cloud GNSS sensor sending raw snapshot samples of a 15 ms duration every hour to a reserved server, the annual cost per sensor was estimated to be $1.46 [194].

8.2.13.3 LEO

Similar to the cost of LPWAN usage, the cost of satellite IoT solutions is the sum of the hardware cost, a subscription cost and a message cost. The hardware cost varies depending on and the level of optimization. For example, a multipurpose low-cost Very High Frequency (VHF) dipole antenna and an inexpensive RTL-SDR dongle can be used to sample LEO signals in a receiver-sided Doppler positioning approach [197]. However, if a more robust device or reliable positioning service is required, an off-the-shelf solution is recommended. For example, the Argos Receiver Transmitter with Integrated Control (ARTIC) R2 chipset costs €47, while the 'plug-n-play' KIM-1 module costs €50. Due to the all-in-one design, the Semtech LR1110 chip only costs €8.84. Furthermore, the subscription and message costs depend on the operator's business model. For example, Orbcomm charges a monthly message fee of $7 to get access to the OG2-M network on top of the $1.35 per kB sent. Lacuna Space only charges $5 per month per UE, but this excludes the LoRa geolocation service. Being a non-for-profit system, Argos offers unlimited usage of the satellite network for €63 per month. Finally, it is important to note that these subscription costs are quite expensive when compared to LPWAN subscription costs. Therefore, it is more cost-effective to exchange data over a terrestrial link, if available.

8.2.14 UE complexity

In order to determine a position, a set of hardware components such as chipsets, sensors and antennas are required. All of these increase the computational complexity of the UE, which is covered in this section.

8.2.14.1 LPWAN

LPWAN end devices are designed with low complexity in mind. First, this can be observed in the network architecture. LoRa networks, for instance, use a simple star topology, instead of more complex meshing topologies. In NB-IoT, a UE only communicates with a single serving cell, eliminating the need to scan for multiple nearby cells if a good connection is established. Second, RSS, TDoA and AoA positioning approaches do not increase the complexity of the UE, as they only require more complexity at the gateway side. Hence, zero location processing is performed on the device itself. Lastly, a multi-modal LPWAN localization approach does increase the overall UE complexity as intelligent radio switching mechanisms need to be implemented and executed by the device [19].

8.2.14.2 GNSS

When comparing single- to dual-frequency GNSS receivers, the complexity of the latter more than doubles. This is due to the fact that the UE antenna must support both frequencies, a second saw filter needs to be implemented and the signal needs to be recombined again to send down the coax to the receiver [239].

When adding LPWAN or LEO connectivity to GNSS-enabled UEs, the complexity changes depending on the used technique. For instance, adding A-GNSS features requires incorporating information into the GNSS tracking loops. In contrast, S-GNSS receiver designs are usually simpler due to the absence of signal processing blocks. In this case, the reduction in complexity due to the remote processing capability is more significant. However, increasing the sampling frequency of a S-GNSS receiver increases the pseudorange accuracy after the acquisition in exchange for more computational burden. The trade-off between computational complexity and collaborative GNSS hybridization is further described in [240].

Several techniques are being investigated to further decrease GNSS complexity. The Accurate GNSS Positioning for Low-power and Low-cost Objects (APOLLO) project aims to provide a Galileo-based location solution using a 100% software GNSS receiver. By getting rid of chipset constraints, the goal is to reduce device complexity by a factor 10 [201]. Furthermore, a Compressed Sensing (CS) technique requires a smaller number of samples, reducing the amount of memory needed [241].

8.2.14.3 LEO

Similar to LPWAN, position calculation in LEO satellite systems often occurs in the cloud, successfully reducing the UE complexity. Many Doppler positioning systems use an uplink communication approach, where the final UE position is estimated in ground processing centers. For example, Lacuna Space uses standard LoRa devices such as the Semtech LR1110 with a slightly different modulation and antenna to communicate with their satellites. Moreover, a GNSS receiver in LEO systems could be used to steer the LEO clock, reducing the onboard clock requirement and complexity [178]. In high contrast, capturing SoOP from multiple LEO constellations with an SDR at the UE side requires local signal processing, adding computational complexity.

8.2.15 Location update rate

Where some positioning applications need real-time and fast location updates (e.g., in vehicle-to-vehicle communication), others only require a location update once a day. Independent of the application, I here discuss the maximum achievable location update rate for every technology.

8.2.15.1 LPWAN

In unlicensed frequency bands, the amount of airtime is regulated in order to reduce interference from nearby communications and to maintain a fair use policy. For example, Sigfox and LoRa transmitters should respect the 1% duty cycle limitation in the EU 868 MHz band. In the case of Sigfox, this results in the transmission of maximum 140 uplink messages per day. LoRa operators (e.g., TTN) can implement even stricter limitations. In opposition to unlicensed LPWAN technologies, cellular IoT technologies such as NB-IoT do not face these duty cycle limitations, enabling the possibility to produce fast location updates, e.g., every second if needed.

8.2.15.2 GNSS

One major benefit of traditional GNSS receivers is their ability to produce fast location updates locally. Once a receiver has acquired a first fix and goes into tracking mode, it can usually produce a position estimation at an update rate of 1 Hz or higher. However, in order to communicate the position update to a remote user, the system depends on the location update rate of the LPWAN or LEO communication channel. Therefore, the delay between the production of the GNSS coordinates and the reception of the location update by the remote user should be taken into account in (near) real-time tracking applications. Similarly, the raw observables sent to the cloud in a S-GNSS approach are only valid for a limited amount of time.

8.2.15.3 LEO

The location update rate of LEO positioning systems is determined by the number of satellite passes per day, and thus by the number of satellites in orbit. The satellite revisit time is defined as the period during which a UE at a given location has to wait until the next satellite passes to transmit a message. Due to the near-polar orbit of many LEO satellites, the satellite revisit time shortens with latitude. For example, an Argos satellite is able to receive messages from a UE at the poles 14 times per day, while this number decreases when the UE moves towards the equator. In the future constellation of 25 nanosatellites, Kinéis expects the average satellite revisit time to drop below 15 minutes [223]. A similar simulation carried out by Lacuna Space shows that 240 LEO satellites are sufficient to provide connectivity every five minutes. When using SoOP from multiple LEO constellations, more frequent location updates can be calculated locally. For instance, when the Starlink constellation of almost 12,000 is finished, a median number of 100 satellites will be in view at any latitude [178]. Finally, it should be noted that combining LEO and LPWAN communication enables a higher overall location update rate.

8.2.16 Local or remote processing

A final dimension of this survey indicates whether the localization algorithm or other processing steps are performed locally on the UE or remotely in the cloud.

8.2.16.1 LPWAN

When applying localization algorithms such as RSS ranging or TDoA in terrestrial networks, oftentimes the UE is only required to send uplink messages with network statistics or sensor data. All further processing is performed on a localization server in the cloud, which in turn forwards a location estimate to the end user application. The involved processing steps may include determining the range to each gateway leveraging path loss models, RSS data or accurate timestamps, combining this data with AoA and sensor data in a sensor fusion algorithm, and applying a multilateration algorithm to determine a final coordinate.

8.2.16.2 GNSS

In high contrast, traditional GNSS receivers perform all processing steps locally, ranging from full signal acquisition to pseudorange generation and coordinate production. However, given the communication and low-energy requirements of IoT as, there is an increasing interest towards remote processing of GNSS signals. Recent GNSS receivers do not only integrate LPWAN or LEO communication, but also enable cloud processing. As described in Section 8.1.3, S-GNSS receivers capture small portions of a signal, digitize them and send them over a communication channel to the cloud for further processing. As an early adopter of this technique, Ubiscale provides a solution which shifts power-draining GPS (and Wi-Fi) processings to the cloud to minimize size, power consumption and cost of trackers [237].

8.2.16.3 LEO

Location processing in LEO systems can be performed either locally or remotely, depending on the architecture of the positioning approach. When the satellites perform Doppler measurements based on uplink transmissions from the UE, the data is forwarded via a ground station to a processing station on Earth, where the location of the UE is calculated. On the other side, when a UE scans for all available LEO satellite signals, an algorithm on the UE determines the position of the UE itself. While the former approach requires less on-device processing and thus lower UE complexity, the latter does not require uplink communication to the satellites. In practice, most systems such as the one of Argos and Lacuna Space work using the first approach (i.e., by sending precise frequency and timing information from Doppler measurements at the side of the satellites to a geolocation solver on Earth), as this approach eases the communication of the estimated location to a remote end user.

8.3 Trade-offs during use case design

The performance matrix as shown in Table 8.1 and discussed in the previous sections can serve as a methodology to identify an optimal positioning solution when designing an IoT positioning application. This section interprets the results of the performance matrix and discusses important design trade-offs by providing example use cases.

Most prominent is the trade-off between positioning accuracy and energy consumption. The vast majority of other trade-offs directly or indirectly impacts this trade-off. Hence, an application designer should choose a certain positioning technique primarily based on the position accuracy and energy requirements. Consider a construction company aiming to monitor valuable equipment such as heavy machinery and cranes across several sites. Because the battery lifetime of the trackers on the equipment is of utmost importance, LPWAN positioning may be the best choice in this case. As the sites are far away from each other, the positioning accuracy of several hundreds of meters is acceptable. However, if a higher accuracy is required, other techniques should be used. Similarly, a wildlife tracker may perform LEO Doppler positioning or snapshot GNSS rather than traditional GNSS to save energy. In high contrast, the battery of a pet tracker can be replaced more frequently, allowing for meter-level GNSS accuracy. Furthermore, S-GNSS techniques provide great flexibility. If a higher positioning accuracy is desired, the snapshot size is increased and more observables are communicated to the cloud, at the expense of additional energy consumption.

In a myriad of positioning use cases, it is beneficial in terms of UE cost, computational complexity and energy consumption to perform the location processing in the cloud. Moreover, the final location estimate often needs to be available to remote end users. For these reasons, the cloud processing paradigm has been established in the LPWAN and LEO markets and is gaining popularity in mass-market GNSS receivers. Notwithstanding the foregoing, there also is a growing interest towards local processing of LEO satellite SoOP, as this approach does not require communication, and signals from multiple LEO constellations can be used, improving the coverage in time and space, as well as the location update rate. In the wildlife tracking use case, for example, location and sensor data can be processed and stored locally. Afterwards, the logs can be retrieved manually or requested occasionally to

save on communication energy.

The choice for a certain communication technology comes along with its data constraints, such as data volume, bandwidth and data rate. Unlicensed terrestrial networks need to comply with regional duty cycle regulations and need to deal with limited payload sizes, data rates and location update rates. Therefore, when many observables (e.g., GNSS snapshots) need to be communicated to a cloud solver, a cellular technology such as NB-IoT is preferred. However, if smaller payloads need to be transmitted sporadically (e.g., weekly ephemeris data in A-GNSS), a more flexible LoRa network can be used, possibly in combination with a LEO constellation to extend the application coverage.

The high interoperability of LPWAN, GNSS and LEO solutions enables interesting novel use cases such as energy-efficient global tracking, managing natural resources, improving food production and optimizing global infrastructure. In particular, the compelling multimodal aspect offers the ability to intelligently change positioning strategies. For example, when the previously mentioned construction company is transporting equipment across cities or even countries, a network of terrestrial base stations or LEO satellites is used in order to save energy. However, when an asset reaches its destination site, the accuracy becomes more important. Based on a proximity detection algorithm, the UE might now switch to a more accurate A-GNSS approach, with the last known location and time provided by the communication network. Finally, switching between technologies also depends on their availability. For instance, the wildlife tracker may switch from LPWAN to LEO communication if no terrestrial network is available.

Figure 8.21 summarizes the key dimensions of every positioning technique discussed in this study. In this figure, it becomes clear that the traditional observable-based GNSS technique has one main limitation: the system was not designed with low energy consumption in mind. Hence, techniques such as A-GNSS and S-GNSS aim to solve this issue, leveraging LPWAN or LEO satellite communication networks [173]. As a more energy-efficient alternative, these networks can be used for positioning purposes as well, in exchange for positioning accuracy.

While LPWAN hardware is widely available and networks are easily accessible, LEO satellite networks for both communication and localization have not reached this point yet. Currently, there is no LEO constellation available which covers each place on Earth permanently (i.e., 24/7). As more and more LEO satellites will be launched into orbit in the coming years, satellite revisit times will shorten, increasing the location update rate, as well as the overall coverage. The most promising feature of LEO satellite networks is the ability to communicate a location estimate in areas where no terrestrial networks are available, such as remote mountainous, desert or jungle regions. Moreover, ESA is planning to add a layer of LEO-PNT satellites to complement the Galileo system [31, 242]. Hence, I believe LEO constellations have a promising future, combining the benefits of LPWAN and GNSS positioning solutions.

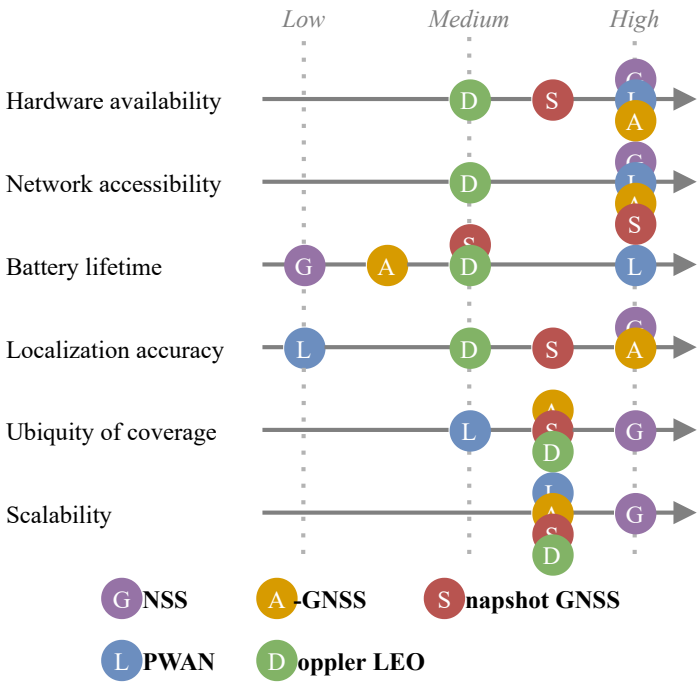


Figure 8.21: Summary of the positioning performance analysis.

8.4 Conclusion & Future Work

In this survey for ESA, I have provided a performance analysis of state-of-the-art, large-scale and energy-efficient positioning techniques for the IoT. Based on 16 dimensions, I composed a performance comparison matrix, which can be used by application designers to determine the most optimal positioning solution. In most cases, this solution will consist of a combination of positioning techniques and communication technologies, which emphasizes the high interoperability between LPWAN, GNSS and LEO systems. Through example use cases, important design trade-offs were discussed, in which the location accuracy and energy requirements play a decisive role.

While techniques such as LPWAN positioning and A-GNSS are widely adopted, others have remaining challenges to be investigated in the future. First, dedicated S-GNSS receivers are not widely available, as well as opportunistic LEO positioning hardware. As LEO constellations continue to expand, commercial positioning services will become available for the general public after the commercialization of this hardware. Second, based on the performance comparison matrix in this survey, some practical guidelines can help manufacturers with the integration of LPWAN, GNSS and LEO technologies, e.g., how to combine technologies and which configurations (sleep modes, snapshot lengths etc.) to use. Related research challenges involve the optimization of intelligent switching between positioning and communication technologies, potentially based on energy and accuracy requirements. Third, more accurate orbit products can improve LEO positioning, as the current accuracy makes

up a large part of the localization error budget. Moreover, the propagation of LEO signals in indoor environments to enable indoor positioning should be further investigated. Finally, advanced yet low-complex compression algorithms could lower the data constraints of communication networks for efficient data transfer to the cloud. These efforts along with this state-of-the-art survey should deliver a better understanding of current challenges to enable ubiquitous, energy-efficient and large-scale positioning solutions in the skyrocketing market of satellite IoT.

Chapter 9

Conclusion

THE IoT has caused the modern society to connect billions of smart devices with the aim to improve a myriad of applications and mankind in general. The challenge is to keep track of this plethora of devices. Therefore, we need communication and localization technologies. In times of a pandemic and chip shortage, it has become even more clear how much we rely on such technology to 'stay connected'. Moreover, the rapidly increasing energy prices have raised the question: How can we locate IoT devices on a large scale and in an energy-efficient way to enable location-aware sustainable IoT applications?

The objective of this thesis was to identify novel, energy-efficient and large-scale technologies and techniques that can be used for positioning purposes, as well as to characterize their positioning performance. A substantial part of my research focused on localization using LPWAN. For each of the three most prominent LPWAN technologies, I conducted RSS-based localization experiments. By implementing fingerprint- and range-based techniques, I significantly improved the localization accuracy of Sigfox and LoRaWAN, compared to the state-of-the-art. In the case of NB-IoT, I provided the first RSS-based localization results based on experiments conducted in the same real-world environment, allowing fair performance comparison. Closer to the end of my 4-year journey, I replaced the terrestrial LPWAN gateways with LEO satellites and explored the wonderful world of satellite IoT. In a survey for ESA, I analyzed the state-of-the-art of novel energy-efficient positioning techniques. In this survey, I characterized and compared the positioning performance of LPWAN, GNSS and LEO-PNT.

In this final chapter, Section 9.1 summarizes the main findings of this thesis according to the contributions listed in Section 1.2. With the aim to continue innovations in the location-enabled (satellite) IoT market, Section 9.2 describes future research directions.

9.1 Main findings

This section summarizes the main findings and conclusions of this thesis, based on the four contributions listed in Section 1.2.

Contribution 1: Improve RSS-based localization with Sigfox and LoRaWAN

Chapters 3 and 4 provide insights in how to leverage the Sigfox LPWAN technology for localization purposes. The UNB nature of Sigfox does not allow timing-based localization methods. Therefore, I analyzed and improved various RSS-based techniques to locate a mobile Sigfox transmitter with minimal energy consumption.

In Chapter 3, I explored the feasibility of combining Sigfox and Wi-Fi. I came up with a 3-step approach. First, the mobile device scans for BSSIDs and RSS values of nearby Wi-Fi access points. Second, due to Sigfox payload limitations, only two BSSIDs are sent over a Sigfox network to a backend. Third, the backend contacts a database of access points with their locations and estimates the location of the mobile device. By optimizing the considered RSS data, this approach resulted in a median estimation error of 39 m, which is acceptable in many mobile IoT applications. However, a passive Wi-Fi receiver is required in this approach, which slightly increases the overall UE complexity and energy consumption.

Chapter 4 employs a Joint Communication And Sensing (JCNS) approach, aiming to use the Sigfox communication signal for localization purposes as well, eliminating the need for additional hardware. Consequently, this approach has the huge benefit that no additional localization energy is required, as the location is calculated in the cloud. This is in opposition to the approach in Chapter 3 or in GNSS. However, on the other side of the coin, the localization accuracy is poor. In order to improve the accuracy, I optimized a fingerprinting algorithm, successfully reducing the mean location estimation error from 689 m to 340 m.

The goal of Chapter 5 is to provide a benchmark of RSS-based localization algorithms in a public LoRaWAN network and to make a fair performance comparison. On the one hand, 10 ML algorithms were evaluated in a fingerprinting approach. On the other hand, I evaluated different path loss models in an RSS ranging approach. Despite the need for a maintained training database, the results show that a fingerprint-based approach leads to 50% more accurate position estimates compared to range-based approaches, with estimation errors of 340 m and 700 m, respectively. Furthermore, I demonstrated that the kNN and Random Forest algorithms perform best in terms of location accuracy.

Finally, Chapter 6 describes the potential of LoRa localization utilizing the worldwide available 2.4 GHz band. Through mathematical analysis, I investigated the maximum communication range of this technology. The results show a maximum range of 133 km in free space, 74 m in an indoor office-like environment, and 443 m in an outdoor urban context. Due to the configurable bandwidth and lower data rates, LoRa outperforms other technologies in the 2.4GHz band in terms of communication range. In addition, both communication and localization applications deployed in private LoRa networks can benefit from the increased

bandwidth and localization accuracy of this system when compared to public sub-GHz networks.

Contribution 2: Provide first insights in RSS-based localization with NB-IoT

Cellular IoT, especially Narrowband-IoT (NB-IoT), is stealing LPWAN market shares from its competitors. Located under the umbrella of 5G mobile network technologies, NB-IoT guarantees deep indoor coverage and energy-efficient communication due to a dense network of existing LTE base stations. While OTDoA positioning was introduced in 3GPP Release 14, network operators refuse or hesitate to implement it because of the increased cost and synchronization complexity. Therefore, I investigated RSS-based localization techniques for NB-IoT. Because simulations of localization experiments often result in over-optimistic localization performance, I insisted to perform outdoor localization experiments using real-world measurements. In this way, I can paint a more realistic picture of the localization performance.

Chapter 7 provides the firsthand results of RSS-based localization experiments using a public NB-IoT network. The experiments I carried out in a large-scale urban environment yield mean location estimation errors ranging from 340 m in a basic proximity algorithm to 204 m in an optimized fingerprinting algorithm. While a fingerprinting approach usually requires more time and effort, the location accuracy can thus be increased. Eventually, crowd-sourced fingerprinting can solve city-scale collection of training data.

The discrepancy between what the NB-IoT standard prescribes and how chipset manufacturers and network operators choose to implement it, has caused several issues. By working with real-world data, I faced these practical issues. The most prominent example is the fact that for many UEs, it was not possible to scan for multiple nearby cells if there is a decent link quality to the currently serving cell. After addressing and publishing these findings, I am happy to notice that chipset manufacturers such as Nordic Semiconductor have started updating their firmware to include optional multi-cell discovery for localization purposes.

In summary, I conducted LPWAN localization experiments in the same real-world city-scale environment, enabling fair comparison between Sigfox, LoRaWAN and NB-IoT. The mean localization accuracy is in the order of a few hundred meters. While this may seem very inaccurate, this level of accuracy is often sufficient in many large-scale, location-aware IoT applications. For instance, imagine tracking assets on a truck which is driving on a highway at 120 km/h. This means the truck moves 333 m every 10 s. So, with an LPWAN location update rate which is typically lower, a location estimation error of 340 m is negligible in this non-realtime scenario. If a more accurate position of the asset is required when it arrives at its destination, one can switch from localization based on LPWAN to GNSS.

Contribution 3: Explore the possibility for PNT services leveraging LEO satellites

The idea of putting satellites into orbits that are approximately 20 times closer to Earth is not new. In the *New Space* era however, LEO satellites have been rediscovered. LEO constellations come in different sizes and provide a wide range of services, such as broadband communication and Earth observation. However, up to now, there is no fully operational LEO constellation designed for PNT services, and only a few operators provide an exploratory PNT service. Therefore, I investigated several positioning techniques leveraging LEO satellites and analyzed their performance and application potential.

Most currently available solutions exploit the Doppler effect of the fast moving LEO satellites to provide a positioning service. The Doppler measurements can take place either at the satellites or at the UE. Argos and Lacuna Space, for instance, apply the former approach, obtaining a positioning accuracy of 150 m. This approach makes sense given the uplink nature of satellite communication. Moreover, the approach does not impact the UE complexity and energy consumption. This is not the case when measuring the Doppler frequency at the UE. However, such approach can leverage any SoOP from any (mega-)constellation for PVT computation. This leads to varying accuracies up to 8 m, depending on the number of satellites in view and additional sensors used. In turn, the limitation is that a location estimate cannot be communicated to a remote user.

The most recent development involves building a constellation specifically designed to provide PNT services from LEO satellites. In other words, GNSS-like ranging signals are broadcast from LEO satellites to Earth. Both Xona Space Systems and GeeSpace believe such a system might reach sub-meter accuracy, enabling applications such as autonomous driving. Furthermore, a dedicated LEO-PNT system can function independent of GNSS, aiming to provide accurate, reliable, and secure PNT services.

Contribution 4: Compare energy-efficient positioning technologies and techniques

The goal of this final and overarching contribution was to identify and compare energy-efficient alternatives to traditional GNSS. Therefore, I combined the knowledge I gathered from the terrestrial LPWAN localization experiments and the survey for ESA on space-based positioning technologies and techniques.

Despite standalone GNSS services reaching meter-level accuracies and therefore being widely adopted, GNSS technology exhibits several shortcomings. Because GNSS satellites move in MEO and given their low transmission power, GNSS signals often cannot reach indoor or dense urban environments. Multipath errors degrade the accuracy drastically. Furthermore, GNSSs were not designed with low energy consumption in mind. In contrast, energy consumption is of utmost importance in mobile IoT applications. Therefore, more energy-efficient positioning alternatives are gaining increasing popularity. In this context, the question raises whether the high GNSS availability and accuracy are required by the application or if the requirements could be fulfilled by one of these alternatives. A final limitation of tra-

ditional GNSS approaches is the local PVT calculation, requiring an additional connectivity link to communicate the computed location to a remote user.

Several industries have shown an increasing demand for alternative and innovative positioning solutions that are more suited in an IoT context. Hence, I conducted a survey on state-of-the-art, large-scale and energy-efficient positioning techniques for IoT applications. More specifically, I analyzed the performance of terrestrial-based LPWAN techniques, novel GNSS solutions, and innovative positioning techniques leveraging LEO satellite constellations. I created and discussed a performance comparison matrix based on 16 key dimensions including energy consumption, positioning accuracy, coverage and scalability.

Attempts to rectify the aforementioned shortcomings of GNSS are currently taking place. Through a connectivity link, remote location processing techniques are gaining more attention. For example, S-GNSS techniques enable cloud processing by sending raw observables via a terrestrial network to the cloud. This improves both the UE complexity and energy consumption.

The balance between positioning accuracy and energy consumption is thus only one of the numerous trade-offs use case designers have to consider. Moreover, the application requirements vary significantly, as well as the positioning techniques that are feasible in a certain environment. Through the performance comparison matrix, I hope the survey gave more insights into what technology or technique can best be applied in a certain scenario.

ESA is planning to design a multi-layer 'system of systems', including PNT satellites in LEO and MEO, as well as terrestrial networks [31, 242]. This supports the conclusion of my analysis, highlighting the importance of interoperability. The survey shows that there is no 'one-size-fits-all' technology. Cooperation between LPWAN, GNSS, LEO satellites and additional sensors is required to enable energy-efficient communication and positioning applications in the emerging market of satellite IoT.

9.2 Future work

The previous section summarized the outcomes of the research I conducted in the past four years. This final section lists a number of potential improvements to my work in the short term, as well as future research directions in the long term.

The results of the LPWAN localization experiments with Sigfox, LoRaWAN and NB-IoT show a significant number of outliers, especially for the RSS-based ranging algorithms. These outliers are caused by the heterogeneous environment, fading and shadowing effects. Due to this heterogeneity of the environment, it is a challenging task to choose an appropriate path loss model. In order to avoid the use of a single existing path loss model in a large-scale environment, a custom path loss model can be created for the application environment. However, this eliminates the advantage of a fast deployment in a range-based localization approach. While I introduced three gateway selection strategies in Chapter 5, more advanced outlier detection algorithms could be implemented. For example, an AI-based algorithm may flag a measurement when the location estimation error exceeds the 95th percentile during the training phase. Then, when a similar measurement is observed during the operational

phase, the algorithm detects this and may discard the measurement.

Instead of just discarding a measurement or providing a location estimate, the end user can also be informed about the estimation accuracy of the location, i.e., the localization reliability. By doing so, the end user has more valuable information of each measurement and can judge whether or not the location estimate is considered valid for the targeted application. Furthermore, the area of uncertainty in overlapping circles in a multilateration algorithm (or squares in Min-Max) can serve as a metric for localization reliability.

In the first years after the introduction of NB-IoT, the cellular technology had to deal with teething problems. A major issue from a localization perspective was the fact that a UE could only request the RSS from the serving cell. After publishing this issue described in Chapter 7, more and more NB-IoT chipset manufacturers have started implementing optional multi-cell discovery for positioning purposes. Hence, the logical next step is to verify how this feature improves the localization accuracy.

While the current work on LPWAN mainly focused on positioning, future research may involve navigation. By implementing a navigation framework based on a particle filter, the location of a mobile IoT device can be updated with an increased accuracy. This framework can potentially add road snapping features and integrate sensor data from an IMU to further improve the accuracy.

With new technology, the overall UE energy consumption will only decrease, while the energy harvesting capability will only increase. Hence, a positioning solution without batteries might become a reality, paving the road towards a sustainable IoT. The ultimate goal is to create a system that is able to track any IoT device, anywhere on Earth, anytime, without batteries.

A major trend in the IoT landscape is the extension of terrestrial LPWAN communication to satellite-based networks. For example, Semtech has designed LoRa chipsets that are able to communicate with both terrestrial gateways and LEO satellites. These chipsets are being used by start-ups such as Lacuna Space, Wyld and Hiber. Likewise, cellular LPWAN technologies such as NB-IoT (under the umbrella of 5G technologies) are complementing LEO satellite services. By doing so, the coverage of LPWAN is extending to remote regions without terrestrial infrastructure such as deserts, mountains, oceans and jungles.

In line with this trend, support for NTNs is included in 3GPP Releases 17 and 18, with the aim to provide worldwide access to 5G, 6G and beyond services via satellite networks. Mobile devices using 5G New Radio (NR) are generally considered to be more energy-efficient compared to previous generations of mobile technology. This is because 5G uses advanced signal processing and multiple antenna techniques such as beamforming to transmit data more efficiently, which reduces the amount of energy required for data transmission. For this reason, the ongoing hybridization efforts for 5G and LEO will enable even more innovative energy-aware communication and localization applications.

The rapidly increasing number of LEO satellites in orbit has a positive impact from a positioning perspective. When there are always enough satellites in view at any place on Earth, a UE will not have to wait for a satellite pass to transmit its messages. This will lower the overall UE complexity and communication delay in near-realtime applications. Moreover, the location update rate can be increased if desired. That being said, the large number of

LEO satellites with a relatively short lifetime produce a vast amount of space debris. Thus, the removal or potential reuse of space debris poses a major challenge in the near future.

Navigating on Earth with sub-meter level accuracy. Observing live high-quality satellite footage of any place on Earth. Sending SOS messages from your smartphone to satellites in case of an emergency. Organizing commercial spaceflights for space tourism and rideshare missions. Exploring Mars with rovers and drones. Looking billions of years back in space and time to learn about the history of our universe through the James Webb Space Telescope (JWST). All of these astonishing achievements would have been impossible without efficient communication and positioning systems. The space exploration era has only just begun and is a given for decades ahead.

Bibliography

[1] M. Hasan, "Number of connected IoT devices growing 18% to 14.4 billion globally," 2022, available online: https://iot-analytics.com/number-connected-iot-devices/ (accessed on 17 July 2022).

[2] B. Jovanovic, "Internet of Things statistics for 2022," 2022, available online: https://dataprot.net/statistics/iot-statistics/ (accessed on 17 July 2022).

[3] J. Gante, L. Sousa, and G. Falcao, "Dethroning GPS: Low-Power Accurate 5G Positioning Systems Using Machine Learning," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 10, no. 2, pp. 240–252, jun 2020.

[4] K. Chen, G. Tan, J. Cao, M. Lu, and X. Fan, "Modeling and Improving the Energy Performance of GPS Receivers for Location Services," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 4512–4523, apr 2020.

[5] M. Aernouts, "Localization with Low Power Wide Area Networks," PhD thesis, Universiteit Antwerpen, 2022.

[6] O. Osechas, F. Fohlmeister, T. Dautermann, and M. Felux, "Impact of GNSS-Band Radio Interference on Operational Avionics," *Navigation*, vol. 69, no. 2, jun 2022.

[7] Sigfox S.A., "Sigfox," 2022, available online: https://www.sigfox.com/en (accessed on 23 March 2022).

[8] LoRa Alliance, "LoRaWAN: Wide Area Networks for IoT," 2022, available online: https://www.lora-alliance.org/ (accessed on 23 March 2022).

[9] E. M. Migabo, K. D. Djouani, and A. M. Kurien, "The Narrowband Internet of Things (NB-IoT) Resources Management Performance State of Art, Challenges, and Opportunities," *IEEE Access*, vol. 8, pp. 97 658–97 675, 2020.

[10] K. Mikhaylov, J. Petäjäjärvi, K. Mikhaylov, M. Hämäläinen, and J. Iinatti, "Evaluation of LoRa LPWAN technology for remote health and wellbeing monitoring Evaluation of LoRa LPWAN Technology for Remote Health and Wellbeing Monitoring," *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, no. March, pp. 1–5, 3 2016.

[11] O. O. Ajayi, A. B. Bagula, H. C. Maluleke, Z. Gaffoor, N. Jovanovic, and K. C. Pietersen, "WaterNet: A Network for Monitoring and Assessing Water Quality for Drinking and Irrigation Purposes," *IEEE Access*, vol. 10, pp. 48 318–48 337, 2022.

[12] R. K. Singh, M. H. Rahmani, M. Weyn, and R. Berkvens, "Joint Communication and Sensing: A Proof of Concept and Datasets for Greenhouse Monitoring Using LoRaWAN," *Sensors*, vol. 22, no. 4, p. 1326, feb 2022.

[13] A. Nazir, H. Mosleh, M. Takruri, A.-h. Jallad, and H. Alhebsi, "Early Fire Detection: A New Indoor Laboratory Dataset and Data Distribution Analysis," *Fire*, vol. 5, no. 1, p. 11, jan 2022.

[14] M. Pointl and D. Fuchs-Hanusch, "Assessing the Potential of LPWAN Communication Technologies for Near Real-Time Leak Detection in Water Distribution Systems," *Sensors*, vol. 21, no. 1, p. 293, 2021.

[15] L. Kirkpatrick, I. Hererra Olivares, A. Massawe, C. Sabuni, H. Leirs, R. Berkvens, and M. Weyn, "Proxlogs: Miniaturised proximity loggers for monitoring association behaviour in small mammals," *bioRxiv*, 2021.

[16] A. Cilfone and L. Davoli, "LoRaFarM: a LoRaWAN-Based Smart Farming Modular IoT Architecture," *MDPI Sensors*, 2020.

[17] I. F. Priyanta, F. Golatowski, T. Schulz, and D. Timmermann, "Evaluation of LoRa Technology for Vehicle and Asset Tracking in Smart Harbors," *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1, pp. 4221–4228, 2019.

[18] H. Wang, Y. Liu, Y. Wei, Y. He, K. F. Tsang, L. L. Lai, and C. S. Lai, "LP-INDEX: Explore the Best Practice of LPWAN Technologies in Smart City," in *2020 IEEE International Smart Cities Conference (ISC2)*. IEEE, sep 2020, pp. 1–5.

[19] K. Mikhaylov, M. Stusek, P. Masek, V. Petrov, J. Petajajarvi, S. Andreev, J. Pokorny, J. Hosek, A. Pouttu, and Y. Koucheryavy, "Multi-RAT LPWAN in Smart Cities: Trial of LoRaWAN and NB-IoT Integration," in *2018 IEEE International Conference on Communications (ICC)*, vol. 2018-May. IEEE, may 2018, pp. 1–6.

[20] B. Buurman, J. Kamruzzaman, G. Karmakar, and S. Islam, "Low-Power Wide-Area Networks : Design Goals , Architecture , Suitability to Use Cases and Research Challenges," *IEEE Access*, vol. 8, pp. 17 179–17 220, 2020.

[21] Y. Li, Y. Zhuang, X. Hu, Z. Gao, J. Hu, L. Chen, Z. He, L. Pei, K. Chen, M. Wang, X. Niu, R. Chen, J. Thompson, F. M. Ghannouchi, and N. El-Sheimy, "Toward Location-Enabled IoT (LE-IoT): IoT Positioning Techniques, Error Sources, and Error Mitigation," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4035–4062, mar 2021.

[22] P. Phalaagae, A. M. Zungeru, B. Sigweni, J. M. Chuma, and T. Semong, "Applications and Communication Technologies in IoT Sensor Networks," in *Green Internet of Things Sensor Networks*. Cham: Springer International Publishing, 2020, pp. 9–23.

[23] R. C. Shit, S. Sharma, D. Puthal, and A. Y. Zomaya, "Location of Things (LoT): A Review and Taxonomy of Sensors Localization in IoT Infrastructure," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2028–2061, 2018.

[24] Semtech, "Satellite IoT Connectivity Service Integrated With LoRaWAN," 2021, available online: https://www.semtech.com/company/press/semtech-and-echostar-mobile-to-test-satellite-iot-connectivity-service-integrated-with-lorawan (accessed on 17 July 2022).

[25] P. Wang, J. Zhang, X. Zhang, Z. Yan, B. G. Evans, and W. Wang, "Convergence of Satellite and Terrestrial Networks: A Comprehensive Survey," *IEEE Access*, vol. 8, pp. 5550–5588, 2020.

[26] R. M. Ferre, E. S. Lohan, H. Kuusniemi, J. Praks, S. Kaasalainen, C. Pinell, and M. Elsanhoury, "Is LEO-Based Positioning with Mega-Constellations the Answer for Future Equal Access Localization?" *IEEE Communications Magazine*, vol. 60, no. 6, pp. 40–46, jun 2022.

[27] E. Kulu, "Satellite Constellations - NewSpace Index," 2022, available online: https://www.newspace.im/ (accessed on 19 July 2022).

[28] Wikipedia, "2022 in Spaceflight," 2022, available online: https://en.wikipedia.org/wiki/2022_in_spaceflight (accessed on July 19 2022).

[29] B. Y. D. Lawrence, H. S. Cobb, G. Gutt, M. Connor, T. G. R. Reid, T. Walter, and D. Whelan, "Navigation from LEO: Current Capability and Future Promise," Tech. Rep. July, 2017. [Online]. Available: https://web.stanford.edu/group/scpnt/gpslab/pubs/papers/Lawrence_GPSWorld_July2017.pdf

[30] T. G. Reid, B. Chan, A. Goel, K. Gunning, B. Manning, J. Martin, A. Neish, A. Perkins, and P. Tarantino, "Satellite Navigation for the Age of Autonomy," *2020 IEEE/ION Position, Location and Navigation Symposium, PLANS 2020*, pp. 342–352, apr 2020.

[31] European Space Agency, "New navigation missions for enhanced satnav and Earth mapping," 2022, available online: https://www.esa.int/Applications/Navigation/New_navigation_missions_for_enhanced_satnav_and_Earth_mapping (accessed on 17 July 2022).

[32] M. Aernouts, R. Berkvens, K. Van Vlaenderen, and M. Weyn, "Sigfox and LoRaWAN Datasets for Fingerprint Localization in Large Urban and Rural Areas," *Data*, vol. 3, no. 2, 4 2018.

[33] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low Power Wide Area Networks: An Overview," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.

[34] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," *ICT Express*, vol. 5, no. 1, pp. 1–7, 2019.

[35] R. K. Singh, P. P. Puluckul, R. Berkvens, and M. Weyn, "Energy Consumption Analysis of LPWAN Technologies and Lifetime Estimation for IoT Application," *Sensors*, vol. 20, no. 17, p. 4794, aug 2020.

[36] E. Rastogi, N. Saxena, A. Roy, and D. R. Shin, "Narrowband Internet of Things: A comprehensive study," *Computer Networks*, 2020.

[37] E. Pasqua, "5 things to know about the LPWAN market in 2021," 2021, available online: https://iot-analytics.com/5-things-to-know-lpwan-market/ (accessed on 19 July 2022).

[38] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 60–67, oct 2016.

[39] Engie M2M, "Enabling your IoT solution," 2022, available online: https://corporate.engie.be/ (accessed on 23 March 2022).

[40] I. Lunden, "Sigfox, the French IoT startup that had raised more than $300M, files for bankruptcy protection as it seeks a buyer," 2022. [Online]. Available: https://techcrunch.com/2022/01/27/sigfox-the-french-iot-startup-that-had-raised-more-than-300m-files-for-bankruptcy-protection-as-it-seeks-a-buyer/

[41] A.-F. Pelé, "Singapore's UnaBiz Takes Over Sigfox," apr 2022. [Online]. Available: https://www.eetimes.eu/singapores-unabiz-takes-over-sigfox/

[42] H. Bong, "Sigfox is now a UnaBiz technology," 2022, available online: https://www.linkedin.com/posts/henribong_lpwan-unitedlpwanworld-sigfox-activity-6949631719444516864-xn5U/ (accessed on 17 July 2022).

[43] S. Sendra, L. García, J. Lloret, I. Bosch, and R. Vega-Rodríguez, "LoRaWAN Network for Fire Monitoring in Rural Environments," *Electronics*, vol. 9, no. 3, p. 531, mar 2020.

[44] A. Augustin, J. Yi, T. Clausen, and W. Townsley, "A Study of LoRa: Long Range & Low Power Networks for the Internet of Things," *Sensors*, vol. 16, no. 9, p. 1466, 9 2016.

[45] P. Gkotsiopoulos, D. Zorbas, and C. Douligeris, "Performance Determinants in LoRa Networks: A Literature Review," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1721–1758, 2021.

[46] W. Ayoub, A. E. Samhat, F. Nouvel, M. Mroue, J.-c. Prevotet, J. C. Prévotet, J.-c. Prevotet, and J. C. Prévotet, "Internet of Mobile Things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and Supported Mobility," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1561–1581, 2019.

[47] A. Picard, J.-C. Lapayre, F. Hanna, and R. Muthada Pottayya, "COMMA: A new LoRaWAN communication optimisation mechanism for mobility adaptation of IoT," *IET Wireless Sensor Systems*, vol. 11, no. 3, pp. 120–130, jun 2021.

[48] C. Bernier, F. Dehmas, and N. Deparis, "Low Complexity LoRa Frame Synchronization for Ultra-Low Power Software-Defined Radios," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3140–3152, May 2020.

[49] E. Sisinni, P. Ferrari, D. Fernandes Carvalho, S. Rinaldi, P. Marco, A. Flammini, and A. Depari, "LoRaWAN Range Extender for Industrial IoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5607–5616, aug 2020.

[50] Semtech, "Wireless RF: LoRa 2.4 GHz," 2022, available online: https://www.semtech.com/products/wireless-rf/lora-24ghz (accessed on 31 May 2022).

[51] Semtech, "Semtech SX128x Long Range datasheet," p. 143, 2019, available online: https://www.semtech.com/products/wireless-rf/lora-24ghz/sx1280 (accessed on 30 March 2022).

[52] M. Aernouts, N. BniLam, N. Podevijn, D. Plets, W. Joseph, R. Berkvens, and M. Weyn, "Combining TDoA and AoA with a Particle Filter in an Outdoor LoRaWAN Network," in *IEEE/ION Position, Location and Navigation Symposium (PLANS)*. St-Louis, Missouri: IEEE, apr 2020, pp. 1060–1069.

[53] A. Mahanti, N. Carlsson, C. Williamson, and M. Arlitt, "Ambient interference effects in Wi-Fi networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6091 LNCS. Springer, Berlin, Heidelberg, 2010, pp. 160–173.

[54] L. Polak and J. Milos, "Performance analysis of LoRa in the 2.4 GHz ISM band: coexistence issues with Wi-Fi," *Telecommunication Systems*, vol. 74, no. 3, pp. 299–309, July 2020.

[55] H. Karvonen, C. Pomalaza-Ráez, K. Mikhaylov, M. Hämäläinen, and J. Iinatti, "Experimental Performance Evaluation of BLE 4 Versus BLE 5 in Indoors and Outdoors Scenarios," in *Advances in Body Area Networks*. Springer, 2019, pp. 235–251.

[56] Semtech, "Application Note: Ranging with the SX1280 Transceiver," pp. 1–30, 2017, available online: https://www.semtech.com/products/wireless-rf/lora-24ghz/sx1280 (accessed on 30 March 2022).

[57] F. Wolf, K. Le Deroff, S. de Rivaz, N. Deparis, F. Dehmas, and J.-P. Cances, "Benchmarking of Narrowband LPWA Physical Layer Ranging Technologies," in *2019 16th Workshop on Positioning, Navigation and Communications (WPNC)*. IEEE, oct 2019, pp. 1–6.

[58] M. Kanj, V. Savaux, and M. Le Guen, "A Tutorial on NB-IoT Physical Layer Design," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2408–2446, 2020.

[59] A. Lombardo, S. Parrino, G. Peruzzi, and A. Pozzebon, "LoRaWAN Versus NB-IoT: Transmission Performance Analysis Within Critical Environments," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1068–1081, jan 2022.

[60] K. Radnosrati, C. Fritsche, F. Gunnarsson, F. Gustafsson, and G. Hendeby, "Localization in 3GPP LTE based on one RTT and one TDOA observation," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2020.

[61] S. Popli, R. K. Jha, and S. Jain, "A Survey on Energy Efficient Narrowband Internet of Things (NBIoT): Architecture, Application and Challenges," *IEEE Access*, vol. 7, pp. 16 739–16 776, 2019.

[62] F. Tong, Y. Sun, and S. He, "On Positioning Performance for the Narrow-Band Internet of Things: How Participating eNBs Impact?" *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 423–433, 1 2019.

[63] L. De Nardis, G. Caso, O. Alay, U. Ali, M. Neri, A. Brunstrom, and M.-G. Di Benedetto, "Positioning by fingerprinting with multiple cells in NB-IoT networks,"

in *2022 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE, jun 2022, pp. 01–07.

[64] Y. Lalle, L. C. Fourati, M. Fourati, and J. P. Barraca, "A Comparative Study of LoRaWAN, SigFox, and NB-IoT for Smart Water Grid," in *2019 Global Information Infrastructure and Networking Symposium (GIIS)*. IEEE, 12 2019, pp. 1–6.

[65] Y. Moon, S. Ha, M. Park, D. Lee, and J. Jeong, "A Methodology of NB-IoT Mobility Optimization," in *2018 Global Internet of Things Summit (GIoTS)*. IEEE, 6 2018, pp. 1–5.

[66] S. Cluzel, L. Franck, J. Radzik, S. Cazalens, M. Dervin, C. Baudoin, and D. Dragomirescu, "3GPP NB-IOT coverage extension using LEO satellites," in *IEEE Vehicular Technology Conference*, vol. 2018-June. IEEE, jul 2018, pp. 1–5.

[67] O. Liberg, S. E. Lowenmark, S. Euler, B. Hofstrom, T. Khan, X. Lin, and J. Sedin, "Narrowband Internet of Things for Non-Terrestrial Networks," *IEEE Communications Standards Magazine*, vol. 4, no. 4, pp. 49–55, dec 2020.

[68] S. K. Routray, A. Javali, A. Sahoo, K. P. Sharmila, and S. Anand, "Military Applications of Satellite Based IoT," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, aug 2020, pp. 122–127.

[69] O. Kodheli, N. Maturo, S. Chatzinotas, S. Andrenacci, and F. Zimmer, "NB-IoT via LEO satellites: An efficient resource allocation strategy for uplink data transmission," *IEEE Internet of Things Journal*, vol. 4662, 2021.

[70] G. Sciddurlo, A. Petrosino, M. Quadrini, C. Roseti, D. Striccoli, F. Zampognaro, M. Luglio, S. Perticaroli, A. Mosca, F. Lombardi, I. Micheli, A. Ornatelli, V. Schena, A. Di Mezza, A. Mattioni, D. Morbidelli, G. Boggia, and G. Piro, "Looking at NB-IoT over LEO Satellite Systems: Design and Evaluation of a Service-Oriented Solution," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[71] 3GPP, "NTN & Satellite in Rel-17 & 18," 2021, available online: https://www.3gpp.org/news-events/partners-news/2254-ntn_rel17 (accessed on 17 July 2022).

[72] M. Aernouts, F. Lemic, B. Moons, J. Famaey, J. Hoebeke, M. Weyn, and R. Berkvens, "A Multimodal Localization Framework Design for IoT Applications," *Sensors 2020, Vol. 20, Page 4622*, vol. 20, no. 16, p. 4622, 8 2020.

[73] Moons, Bart, "Multi-Level Interoperability in Heterogeneous Low Power Wide Area Networks," Ph.D. dissertation, Ghent University, 2022.

[74] M. Marcelic, B. Sandric, J. Jelenic, and M. Jurcevic, "Determining location in LPWAN using multilateration," in *2019 2nd International Colloquium on Smart Grid Metrology, SMAGRIMET 2019 - Proceedings*. IEEE, may 2019.

[75] A. Bensky, *Wireless Positioning Technologies and Applications, Second Edition*, 2nd ed. Artech House, 2016.

[76] B. Jin, X. Xu, T. Zhang, B. Jin, X. Xu, and T. Zhang, "Robust Time-Difference-of-Arrival (TDOA) Localization Using Weighted Least Squares with Cone Tangent Plane Constraint," *Sensors*, vol. 18, no. 3, p. 778, 3 2018.

[77] D. Bissett, "Analysing TDoA Localisation in LoRa Networks," Master's thesis, TU Delft, 2018.

[78] N. Podevijn, D. Plets, J. Trogh, L. Martens, P. Suanet, K. Hendrikse, and W. Joseph, "TDoA-Based Outdoor Positioning with Tracking Algorithm in a Public LoRa Network," *Wireless Communications and Mobile Computing*, vol. 2018, pp. 1–9, 5 2018.

[79] D. Plets, N. Podevijn, J. Trogh, L. Martens, and W. Joseph, "Experimental Performance Evaluation of Outdoor TDoA and RSS Positioning in a Public LoRa Network," *IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation*, no. September, pp. 24–27, 2018.

[80] W. S. Jeon, S. B. Seo, and D. G. Jeong, "Effective Frequency Hopping Pattern for ToA Estimation in NB-IoT Random Access," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10 150–10 154, 10 2018.

[81] J. A. del Peral-Rosado, J. A. Lopez-Salcedo, and G. Seco-Granados, "Impact of frequency-hopping NB-IoT positioning in 4G and future 5G networks," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 5 2017, pp. 815–820.

[82] H. Sallouha, A. Chiumento, and S. Pollin, "Localization in Long-range Ultra Narrow Band IoT Networks using RSSI," in *2017 IEEE International Conference on Communications (ICC)*. Paris, France: IEEE, 2017, p. 6.

[83] X. Lin, J. Bergman, F. Gunnarsson, O. Liberg, S. M. Razavi, H. S. Razaghi, H. Rydn, and Y. Sui, "Positioning for the Internet of Things: A 3GPP Perspective," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 179–185, 12 2017.

[84] A. Hoglund, X. Lin, O. Liberg, A. Behravan, E. A. Yavuz, M. Van Der Zee, Y. Sui, T. Tirronen, A. Ratilainen, and D. Eriksson, "Overview of 3GPP Release 14 Enhanced NB-IoT," *IEEE Network*, vol. 31, no. 6, pp. 16–22, 11 2017.

[85] 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 and Tutorials*, vol. 20, no. 2, pp. 1124–1148, apr 2018.

[86] K. Radnosrati, G. Hendeby, C. Fritsche, F. Gunnarsson, and F. Gustafsson, "Performance of OTDOA positioning in narrowband IoT systems," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 10 2017, pp. 1–7.

[87] S. Hu, A. Berg, X. Li, and F. Rusek, "Improving the Performance of OTDOA Based Positioning in NB-IoT Systems," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, vol. 2018-Janua. IEEE, 12 2017, pp. 1–7.

[88] I. Sobron, I. Landa, I. Eizmendi, and M. Velez, "Adaptive TDOA Estimation for Positioning in NB-IoT," in *2019 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*. IEEE, 10 2019, pp. 149–152.

[89] G. Pan, T. Wang, X. Jiang, and S. Zhang, "Deep Learning based OTDOA Positioning for NB-IoT Communication Systems," 4 2020.

[90] A. Berg, "OTDOA-Based Positioning in NB-IoT," Ph.D. dissertation, Lund University, 2017.

[91] M. Salomon, S. Lippuner, M. Korb, and Q. Huang, "Implementation and performance evaluation of cellular NB-IoT OTDOA positioning," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 4 2020, pp. 1365–1371.

[92] S. Sedighi, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "Localization Performance of 1-Bit Passive Radars in NB-IOT Applications," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. Guadeloupe, French West Indies: IEEE, dec 2019, pp. 156–160.

[93] N. BniLam, D. Joosens, M. Aernouts, J. Steckel, and M. Weyn, "LoRay: AoA Estimation System for Long Range Communication Network," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2020.

[94] N. BniLam, "Angle of Arrival Estimation for Low Power and Long Range Communication Networks," Ph.D. dissertation, University of Antwerp, 2021.

[95] N. Bnilam, D. Joosens, J. Steckel, and M. Weyn, "Low Cost AoA Unit for IoT Applications," in *13th European Conference on Antennas and Propagation, EuCAP 2019*, 2019.

[96] M. Aernouts, N. BniLam, R. Berkvens, and M. Weyn, "TDAoA: A combination of TDoA and AoA localization with LoRaWAN," *Internet of Things*, vol. 11, p. 100236, sep 2020.

[97] L. Wielandner, E. Leitinger, and K. Witrisal, "RSS-Based Cooperative Localization and Orientation Estimation Exploiting Antenna Directivity," *IEEE Access*, vol. 9, pp. 53 046–53 060, 2021.

[98] S. Gezici, Zhi Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, jul 2005.

[99] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A Survey of Various Propagation Models for Mobile Communication," in *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, 6 2003, pp. 51–82.

[100] F. Wolf, "Multi-Channel Ranging System for the Localization of Wireless Connected Objects in LPWAN: From Modeling to Field Trials," Ph.D. dissertation, University of Limoges, 2020.

[101] A. Anjali, "NB-IoT for Localization and Target Detection," in *Principles and Applications of Narrowband Internet of Things (NBIoT)*, 2021, pp. 105–126.

[102] M. Bolić, M. Rostamian, and P. M. Djurić, "Proximity Detection with RFID: A Step Toward the Internet of Things," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 70–76, 2015.

[103] M. Aernouts, B. Bellekens, R. Berkvens, and M. Weyn, "A Comparison of Signal Strength Localization Methods with Sigfox," *2018 15th Workshop on Positioning, Navigation and Communications, WPNC 2018*, 2018.

[104] P. F. e. Silva, V. Kaseva, and E. S. Lohan, "Wireless Positioning in IoT: A Look at Current and Future Trends," *Sensors (Basel, Switzerland)*, vol. 18, no. 8, 7 2018.

[105] T. Chrysikos, G. Georgopoulos, and S. Kotsopoulos, "Site-specific validation of ITU indoor path loss model at 2.4 GHz," *2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks and Workshops, WOWMOM 2009*, no. 978, 2009.

[106] International Telecommunication Union, "ITU-R propagation models," Tech. Rep., 2008. [Online]. Available: https://www.itu.int/rec/R-REC-P

[107] Y. Singh, "Comparison of Okumura, Hata and COST-231 Models on the Basis of Path Loss and Signal Strength," Tech. Rep. 11, 2012.

[108] A. Bhuvaneshwari, R. Hemalatha, and T. Satyasavithri, "Semi Deterministic Hybrid Model for Path Loss Prediction Improvement," *Procedia Computer Science*, vol. 92, pp. 336–344, 2016.

[109] B. Bellekens, L. Tian, P. Boer, M. Weyn, and J. Famaey, "Outdoor IEEE 802 . 11ah Range Characterization Using Validated Propagation Models," *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, 2017.

[110] A. ElNashar, *Design, Deployment and Performance of 4G-LTE Networks: Practical Approach*. Wiley, 2014.

[111] M. Stusek, D. Moltchanov, P. Masek, K. Mikhaylov, O. Zeman, M. Roubicek, Y. Koucheryavy, and J. Hosek, "Accuracy Assessment and Cross-Validation of LP-WAN Propagation Models in Urban Scenarios," *IEEE Access*, vol. 8, pp. 154 625–154 636, 2020.

[112] G. Caso, O. Alay, L. De Nardis, A. Brunstrom, M. Neri, and M.-G. Di Benedetto, "Empirical Models for NB-IoT Path Loss in an Urban Scenario," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13 774–13 788, sep 2021.

[113] Y. Dalveren and A. Kara, "Performance Evaluation of Empirical Path Loss Models for a Linear Wireless Sensor Network Deployment in Suburban and Rural Environments," *Hittite Journal of Science & Engineering*, vol. 7, no. 4, pp. 313–320, dec 2020.

[114] L. E. Ribeiro, D. W. Tokikawa, J. L. Rebelatto, and G. Brante, "Comparison between LoRa and NB-IoT coverage in urban and rural Southern Brazil regions," *Annals of Telecommunications*, vol. 75, no. 11-12, pp. 755–766, dec 2020.

[115] S. Demetri, M. Zúñiga, G. P. Picco, F. Kuipers, L. Bruzzone, and T. Telkamp, "Automated estimation of link quality for LoRa," in *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*. New York, NY, USA: ACM, apr 2019, pp. 145–156.

[116] Q. Jin, M. Huang, and X. Kuang, "Min-Max Wireless Sensor Network Topology Optimization Localization Algorithm Based on RSSI," *International Journal of Computer and Communication Engineering*, vol. 4, no. 1, pp. 9–12, 2015.

[117] J. J. Robles, J. S. Pola, and R. Lehnert, "Extended Min-Max algorithm for position estimation in sensor networks," in *Proceedings of the 2012 9th Workshop on Positioning, Navigation and Communication (WPNC)*, 2012, pp. 47–52.

[118] G. P. Telles, O. K. Rayel, and G. L. Moritz, " Weighted-Centroid Localization using LoRaWAN Network on Large Outdoor Areas ," *Internet Technology Letters*, no. March, pp. 5–10, 2022.

[119] H. Linka, M. Rademacher, O. Aliu, and K. Jonas, "Path Loss Models for Low-Power Wide-Area Networks: Experimental Results using LoRa," in *VDE ITG-Fachbericht Mobilkommunikation*, 2018.

[120] H. Kwasme and S. Ekin, "RSSI-Based Localization Using LoRaWAN Technology," *IEEE Access*, vol. 7, pp. 99 856–99 866, 7 2019.

[121] F. J. Aranda, F. Parralejo, F. J. Álvarez, and J. A. Paredes, "Performance analysis of fingerprinting indoor positioning methods with BLE," *Expert Systems with Applications*, p. 117095, apr 2022.

[122] S. He and S. H. G. Chan, "Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[123] R. Faragher and R. Harle, "Location Fingerprinting With Bluetooth Low Energy Beacons," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2418–2428, 11 2015.

[124] J. Torres-Sospedra, R. Montoliu, S. Trilles, Ó. Belmonte, and J. Huerta, "Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9263–9278, dec 2015.

[125] L. Zhang, S. Valaee, Y. B. Xu, L. Ma, and F. Vedadi, "Graph-Based Semi-Supervised Learning for Indoor Localization Using Crowdsourced Data," *Applied Sciences*, vol. 7, no. 5, p. 467, apr 2017.

[126] P. Sapiezynski, R. Gatej, A. Mislove, and S. Lehmann, "Opportunities and Challenges in Crowdsourced Wardriving," in *Proceedings of the 2015 ACM Conference on Internet Measurement Conference - IMC '15*. Tokyo, Japan: ACM Press, 2015, pp. 267–273.

[127] F. Alhomayani and M. H. Mahoor, "Deep learning methods for fingerprint-based indoor positioning: a review," *Journal of Location Based Services*, vol. 14, no. 3, pp. 129–200, jul 2020.

[128] Q. Song, S. Guo, X. Liu, and Y. Yang, "CSI Amplitude Fingerprinting-Based NB-IoT Indoor Localization," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1494–1504, 6 2018.

[129] G. G. Anagnostopoulos and A. Kalousis, "ProxyFAUG: Proximity-based Fingerprint Augmentation," in *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, nov 2021, pp. 1–7.

[130] ——, "A reproducible analysis of rssi fingerprinting for outdoor localization using sigfox: Preprocessing and hyperparameter tuning," in *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, sep 2019, pp. 1–8.

[131] ——, "A Reproducible Comparison of RSSI Fingerprinting Localization Methods Using LoRaWAN," in *2019 16th Workshop on Positioning, Navigation and Communications (WPNC)*. IEEE, oct 2019, pp. 1–6.

[132] I. Daramouskas, V. Kapoulas, and M. Paraskevas, "Using Neural Networks for RSSI Location Estimation in LoRa Networks," *10th International Conference on Information, Intelligence, Systems and Applications, IISA 2019*, jul 2019.

[133] Z. A. Pandangan and M. C. R. Talampas, "Hybrid LoRaWAN Localization using Ensemble Learning," in *2020 Global Internet of Things Summit (GIoTS)*. IEEE, jun 2020, pp. 1–6.

[134] J. Purohit, X. Wang, S. Mao, X. Sun, and C. Yang, "Fingerprinting-based Indoor and Outdoor Localization with LoRa and Deep Learning," *2020 IEEE Global Communications Conference, GLOBECOM 2020 - Proceedings*, dec 2020.

[135] H. Sallouha, A. Chiumento, S. Rajendran, and S. Pollin, "Localization in Ultra Narrow Band IoT Networks: Design Guidelines and Tradeoffs," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9375–9385, dec 2019.

[136] Y. Etiabi, M. Jouhari, and E. M. Amhoud, "Spreading Factor and RSSI for Localization in LoRa Networks: A Deep Reinforcement Learning Approach," 2022.

[137] F. Lemic, V. Handziski, M. Aernouts, T. Janssen, R. Berkvens, A. Wolisz, and J. Famaey, "Regression-based Estimation of Individual Errors in Fingerprinting Localization," *IEEE Access*, pp. 1–1, 2019.

[138] G. G. Anagnostopoulos and A. Kalousis, "Analysing the Data-Driven Approach of Dynamically Estimating Positioning Accuracy," in *ICC 2021 - IEEE International Conference on Communications*. IEEE, jun 2021, pp. 1–7.

[139] ——, "Can I Trust This Location Estimate? Reproducibly Benchmarking the Methods of Dynamic Accuracy Estimation of Localization," *Sensors*, vol. 22, no. 3, p. 1088, jan 2022.

[140] C. C. Robusto, "The Cosine-Haversine Formula," *The American Mathematical Monthly*, vol. 64, no. 1, p. 38, jan 1957.

[141] Combain, "Combain Positioning Solutions," 2022, available online: https://combain.com/ (accessed on 23 March 2022).

[142] UnwiredLabs, "LocationAPI: Reliable mobile mapping," 2022, available online: http://locationapi.org/ (accessed on 23 March 2022).

[143] Mozilla, "Mozilla Location Service," 2022, available online: https://location.services.mozilla.com/ (accessed on 23 March 2022).

[144] Navizon, "Navizon: WiFi and Cell-ID location database with Global coverage," 2015, available online: https://www.navizon.com/ (accessed on 23 March 2022).

[145] Bobzilla, Arkasha, and Uhtu, "WiGLE," 2022, available online: https://wigle.net/ (accessed on 23 March 2022).

[146] Business Wire, "Make Your Smart Phone Pay For Itself by Running Navizon," 2017, available online: http://www.businesswire.com/news/home/20100128005942/en/Smart-Phone-Pay-Running-Navizon (accessed on 3 July 2017).

[147] UnwiredLabs, "The Unwired Blog: Going from "Accurate" to "Really accurate"," 2017, available online: https://unwiredlabs.com/blog/improved-location-accuracy/ (accessed on 23 March 2022).

[148] S.-h. Cha, "Comprehensive Survey on Distance / Similarity Measures between Probability Density Functions," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 1, no. 4, pp. 300–307, 2007.

[149] D. Tran and T. Nguyen, "Localization In Wireless Sensor Networks Based on Support Vector Machines," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 7, pp. 981–994, 7 2008.

[150] M. Awad and R. Khanna, "Support Vector Regression," in *Efficient Learning Machines*. Berkeley, CA: Apress, 2015, pp. 67–80.

[151] Y. Wang, C. Xiu, X. Zhang, and D. Yang, "WiFi Indoor Localization with CSI Fingerprinting-Based Random Forest," *Sensors 2018, Vol. 18, Page 2869*, vol. 18, no. 9, p. 2869, 8 2018.

[152] A. Hazmi, J. Rinne, and M. Valkama, "Feasibility study of IEEE 802.11ah radio technology for IoT and M2M use cases," in *2012 IEEE Globecom Workshops, GC Wkshps 2012*, 2012, pp. 1687–1692.

[153] Minseok, Jeong and Bomson, Lee, "Comparison between path-loss prediction models for wireless telecommunication system design," in *IEEE Antennas and Propagation Society International Symposium. 2001 Digest. Held in conjunction with: USNC/URSI National Radio Science Meeting (Cat. No.01CH37229)*, vol. 2. IEEE, 2001, pp. 186–189.

[154] C. F. F. Karney, "Algorithms for geodesics," *Journal of Geodesy*, vol. 87, no. 1, pp. 43–55, 1 2013.

[155] Shaoguo Xie, Yanjun Hu, and Yi Wang, "An improved E-Min-Max localization algorithm in wireless sensor networks," in *2014 IEEE International Conference on Consumer Electronics - China*. IEEE, 4 2014, pp. 1–4.

[156] The Things Industries, "The Things Network," 2022. [Online]. Available: https://www.thethingsnetwork.org/

[157] M. Knight and B. Seeber, "Decoding lora: Realizing a modern lpwan with sdr," *Proceedings of the GNU Radio Conference*, vol. 1, no. 1, 2016. [Online]. Available: https://pubs.gnuradio.org/index.php/grcon/article/view/8

[158] L. Vangelista, "Frequency Shift Chirp Modulation: The LoRa Modulation," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1818–1821, 2017.

[159] P. Robyns, P. Quax, W. Lamotte, and W. Thenaers, "A multi-channel software decoder for the LoRa modulation scheme," *IoTBDS 2018 - Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security*, vol. 2018-March, no. July, pp. 41–51, 2018.

[160] R. Ghanaatian, O. Afisiadis, M. Cotting, and A. Burg, "Lora Digital Receiver Analysis and Implementation," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 5, pp. 1498–1502, 2019.

[161] B. Popovic, "Generalized chirp-like polyphase sequences with optimum correlation properties," *IEEE Transactions on Information Theory*, vol. 38, no. 4, pp. 1406–1409, jul 1992.

[162] M. Hua, M. Wang, K. W. Yang, and K. J. Zou, "Analysis of the frequency offset effect on Zadoff-Chu sequence timing performance," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4024–4039, 2014.

[163] F. Heereman, W. Joseph, E. Tanghe, D. Plets, L. Verloock, and L. Martens, "Path loss model and prediction of range, power and throughput for 802.11n in large conference rooms," *AEU - International Journal of Electronics and Communications*, vol. 66, no. 7, pp. 561–568, 2012.

[164] D. Plets, W. Joseph, K. Vanhecke, E. Tanghe, and L. Martens, "Simple Indoor Path Loss Prediction Algorithm and Validation in Living Lab Setting," *Wireless Personal Communications*, vol. 68, no. 3, pp. 535–552, feb 2013.

[165] N. A. M. Razali, M. H. Habaebi, N. F. Zulkurnain, M. R. Islam, and A. Zyoud, "The distribution of path loss exponent in 3D indoor environment," *International Journal of Applied Engineering Research*, vol. 12, no. 18, pp. 7154–7161, 2017.

[166] M. Mollel and M. Kisangiri, "Comparison of Empirical Propagation Path Loss Models for Mobile Communication," *Computer Engineering and Intelligent Systems*, vol. 5, no. 9, pp. 1–11, 2014. [Online]. Available: http://iiste.org/Journals/index.php/CEIS/article/view/15435

[167] M. F. Iskander, "Channel Characterization and Propagation Models for LTE Path Loss Prediction In Urban and Suburban Ghana," *International Journal of Wireless & Mobile Networks (IJWMN)*, vol. 11, no. 6, pp. 23–32, 2019.

[168] Rohde & Schwarz, "NB-IoT: scanners are a must for coverage measurements (part 2)," 2019, available online: https://blog.mobile-network-testing.com/market-technology-trends/evolving-technologies/nb-iot-measurements-part-2/ (accessed on 31 March 2022).

[169] Nordic Semiconductor, "Multicell Location," 2022, available online: https://developer.nordicsemi.com/nRF_Connect_SDK/doc/latest/nrf/libraries/networking/multicell_location.html (accessed on 17 July 2022).

[170] J. Morton, F. Van Diggelen, J. Spilker, B. Parkinson, S. Lo, and G. Grace, *Position, Navigation, and Timing Technologies in the 21st Century*, Y. T. J. Morton, F. Diggelen, J. J. Spilker, B. W. Parkinson, S. Lo, and G. Gao, Eds. Wiley, dec 2020.

[171] A. Minetto, F. Dovis, A. Vesco, M. Garcia-Fernandez, À. López-Cruces, J. L. Trigo, M. Molina, A. Pérez-Conesa, J. Gáñez-Fernández, G. Seco-Granados, and J. A. López-Salcedo, "A Testbed for GNSS-Based Positioning and Navigation Technologies in Smart Cities: The HANSEL Project," *Smart Cities*, vol. 3, no. 4, pp. 1219–1241, oct 2020.

[172] A. Kak and I. F. Akyildiz, "Designing Large-Scale Constellations for the Internet of Space Things with CubeSats," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1749–1768, 2021.

[173] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, Wide-Area, and Non-Terrestrial IoT: A Survey on 5G Advances and the Road Toward 6G," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1117–1174, 2022.

[174] G. W. Hein, "Status, perspectives and trends of satellite navigation," *Satellite Navigation*, vol. 1, no. 1, pp. 1–12, 2020.

[175] S. Liu, Z. Gao, Y. Wu, D. W. Kwan Ng, X. Gao, K. K. Wong, S. Chatzinotas, and B. Ottersten, "LEO Satellite Constellations for 5G and Beyond: How Will They Reshape Vertical Domains?" *IEEE Communications Magazine*, vol. 59, no. 7, pp. 30–36, 2021.

[176] F. S. Prol, R. M. Ferre, Z. Saleem, P. Valisuo, C. Pinell, E. S. Lohan, M. Elsanhoury, M. Elmusrati, S. Islam, K. Celikbilek, K. Selvan, J. Yliaho, K. Rutledge, A. Ojala, L. Ferranti, J. Praks, M. Z. H. Bhuiyan, S. Kaasalainen, and H. Kuusniemi, "Position, Navigation, and Timing (PNT) Through Low Earth Orbit (LEO) Satellites: A Survey on Current Status, Challenges, and Opportunities," *IEEE Access*, vol. 10, pp. 83 971–84 002, 2022.

[177] N. Podevijn, D. Plets, J. Trogh, A. Karaagac, J. Haxhibcqiri, J. Hoebeke, L. Martens, P. Suanet, and W. Joseph, "Performance Comparison of RSS Algorithms for Indoor Localization in Large Open Environments," in *IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation*. IEEE, 11 2018.

[178] T. G. Reid, A. M. Neish, T. F. Walter, and P. K. Enge, "Leveraging commercial broadband LEO constellations for navigation," in *29th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2016*, vol. 4. Institute of Navigation, 9 2016, pp. 2300–2314.

[179] L. M. Irvine, M. H. Winsor, T. M. Follett, B. R. Mate, and D. M. Palacios, "An at-sea assessment of Argos location accuracy for three species of large whales, and the

effect of deep-diving behavior on location error," *Animal Biotelemetry*, vol. 8, no. 1, pp. 1–17, 2020.

[180] EUSPA, "EUSPA EO and GNSS Market Report," EUSPA, Tech. Rep., 2022. [Online]. Available: https://www.euspa.europa.eu/2022-market-report

[181] G. Schrock, "Low Earth Orbit Satellites for Resilient Positioning, Navigation, and Timing," 2022. [Online]. Available: https://gogeomatics.ca/leo-pnt-xona

[182] J. Rainbow, "Xona to test GPS-alternative demo satellite with customer," 2022. [Online]. Available: https://spacenews.com/xona-to-test-gps-alternative-demo-satellite-with-customer/

[183] K. Çelikbilek, Z. Saleem, R. Morales Ferre, J. Praks, and E. S. Lohan, "Survey on Optimization Methods for LEO-Satellite-Based Networks with Applications in Future Autonomous Transportation," *Sensors*, vol. 22, no. 4, p. 1421, feb 2022.

[184] P. W. Robinson, D. P. Costa, D. E. Crocker, J. P. Gallo-Reynoso, C. D. Champagne, M. A. Fowler, C. Goetsch, K. T. Goetz, J. L. Hassrick, L. A. Hückstädt, C. E. Kuhn, J. L. Maresh, S. M. Maxwell, B. I. McDonald, S. H. Peterson, S. E. Simmons, N. M. Teutschel, S. Villegas-Amtmann, and K. Yoda, "Foraging Behavior and Success of a Mesopelagic Predator in the Northeast Pacific Ocean: Insights from a Data-Rich Species, the Northern Elephant Seal," *PLoS ONE*, vol. 7, no. 5, p. e36728, 5 2012.

[185] R. Lopez, J. P. Malardé, F. Royer, and P. Gaspar, "Improving Argos Doppler Location using Multiple-model Kalman Filtering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4744–4755, 2014.

[186] F. Farhangian and R. Landry, "Multi-Constellation Software-Defined Receiver for Doppler Positioning with LEO Satellites," *Sensors*, vol. 20, no. 20, p. 5866, 10 2020.

[187] Z. Tan, H. Qin, L. Cong, and C. Zhao, "New Method for Positioning Using IRIDIUM Satellite Signals of Opportunity," *IEEE Access*, vol. 7, pp. 83 412–83 423, 2019.

[188] H. Benzerrouk, Q. Nguyen, F. Xiaoxing, A. Amrhar, H. Rasaee, and R. J. Landry, "LEO satellites Based Doppler Positioning Using Distributed nonlinear Estimation," *IFAC-PapersOnLine*, vol. 52, no. 12, pp. 496–501, 2019.

[189] J. J. Morales, J. Khalife, A. A. Abdallah, C. T. Ardito, and Z. M. Kassas, "Inertial Navigation System Aiding with Orbcomm LEO Satellite Doppler Measurements," in *Proceedings of the 31st International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2018*, 10 2018, pp. 2718–2725.

[190] J. Khalife, M. Neinavaie, and Z. Z. Kassas, "The First Carrier Phase Tracking and Positioning Results With Starlink LEO Satellite Signals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 2, pp. 1487–1491, apr 2022.

[191] Z. Tan, H. Qin, L. Cong, and C. Zhao, "Positioning Using IRIDIUM Satellite Signals of Opportunity in Weak Signal Environment," *Electronics*, vol. 9, no. 1, p. 37, 12 2019.

[192] J. J. Morales, J. Khalife, U. S. Cruz, and Z. M. Kassas, "Orbit modeling for simultaneous tracking and navigation using LEO satellite signals," in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2019*. Institute of Navigation, 9 2019, pp. 2090–2099.

[193] F. Van Diggelen, *A-GPS: Assisted GPS, GNSS, and SBAS*, 1st ed. Artech House, 2009.

[194] V. Lucas-Sabola, G. Seco-Granados, J. Lopez-Salcedo, J. Garcia-Molina, and M. Crisci, "Efficiency Analysis of Cloud GNSS Signal Processing for IoT Applications," in *30th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2017*, vol. 6, no. 1, 11 2017, pp. 3843–3852.

[195] I. Fernández-Hernández and K. Borre, "Snapshot positioning without initial information," *GPS Solutions*, vol. 20, no. 4, pp. 605–616, 10 2016.

[196] J. Knapp and R. Fuller, "Challenge Accepted: Multi-modal Localization with the LoRa EdgeTM LR1110 for GNSS-Challenged Environments," in *Proceedings of the 34th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2021*. Institute of Navigation, oct 2021, pp. 2221–2241. [Online]. Available: https://www.ion.org/publications/abstract.cfm?articleID=18073

[197] J. J. Khalife and Z. M. Kassas, "Receiver Design for Doppler Positioning with Leo Satellites," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2019-May, pp. 5506–5510, 2019.

[198] Global mobile Supplier Association (GSA), "NB-IoT & LTE-M Executive Summary," 2022. [Online]. Available: https://gsacom.com/technology/narrow-band-iot-m2m/

[199] P. Jorke, R. Falkenberg, and C. Wietfeld, "Power Consumption Analysis of NB-IoT and eMTC in Challenging Smart City Environments," in *2018 IEEE Globecom Workshops (GC Wkshps)*. Abu Dhabi: IEEE, 12 2018, pp. 1–6.

[200] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the Limits of LoRaWAN," *IEEE Communications Magazine*, 2017.

[201] European GNSS Agency, "GNSS User Technology Report 2020," Tech. Rep. 3, 2020.

[202] B. Karki and M. Won, "Characterizing Power Consumption of Dual-Frequency GNSS of Smartphone," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*. IEEE, 12 2020, pp. 1–6.

[203] J. Lee, J. Han, C.-L. Lo, J. Lee, W. Kim, S. Kim, B. Kang, J. Han, S. Jung, T. Nomiyama, J. Lee, T. B. Cho, and I. Kang, "NB-IoT and GNSS All-In-One System-On-Chip Integrating RF Transceiver, 23-dBm CMOS Power Amplifier, Power Management Unit, and Clock Management System for Low Cost Solution," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 12, pp. 3400–3413, 12 2020.

[204] Baseband Technologies, "Snapshot positioning," Tech. Rep., 2019. [Online]. Available: http://www.basebandtech.com/wp-content/uploads/Snapshot-Receiver.pdf

[205] Ublox, "CloudLocate Services Overview," Tech. Rep., 2021.

[206] European GNSS Agency, "Power-efficient positioning for The Internet of Things," Tech. Rep., 2020. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/a57514a5-e1c2-11ea-ad25-01aa75ed71a1/language-en

[207] CLS, "The Argos Chipset: Less power = longer buoy lifetimes," Tech. Rep., 2014. [Online]. Available: https://library.wmo.int/pmb_ged/dbcp-td_51_en/presentations/DBCP-30-STW-6-Woodward-Argos-Chipset.pdf

[208] Kinéis, "Get started with Space IoT," Tech. Rep., 2020. [Online]. Available: https://www.kineis.com/wp-content/uploads/2020/11/200526-Webinar-presentation-EN.pdf

[209] Semtech, "Semtech LR1110 datasheet," 2020. [Online]. Available: https://www.semtech.com/products/wireless-rf/lora-transceivers/lr1110

[210] S. Spinsante and C. Stallo, "Hybridized-GNSS Approaches to Train Positioning: Challenges and Open Issues on Uncertainty," *Sensors*, vol. 20, no. 7, p. 1885, 3 2020.

[211] M. Lehtinen, A. Happonen, and J. Ikonen, "Accuracy and time to first fix using consumer-grade GPS receivers," *SoftCom 2008: 16th International Conference on Software, Telecommuncations and Computer Networks*, pp. 334–340, 2008.

[212] P. Dumitraschkewitz, "GNSS snapshot techniques for quality of service monitoring," Master's thesis, Graz University of Technology, 2020. [Online]. Available: https://diglib.tugraz.at/download.php?id=609a6a8325545&location=browse

[213] Argos, "ARGOS User's Manual," 2016. [Online]. Available: https://www.argos-system.org/manual/

[214] Irnas, "LR1110 chip: one solution for LoRa and GNSS tracking," 2020. [Online]. Available: https://www.irnas.eu/lr1110-chip-one-solution-for-lora-and-gnss-tracking/

[215] M. Lauridsen, H. Nguyen, B. Vejlgaard, I. Z. Kovacs, P. Mogensen, and M. Sorensen, "Coverage Comparison of GPRS, NB-IoT, LoRa, and SigFox in a 7800 km² Area," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, vol. 2017-June. IEEE, 6 2017, pp. 1–5.

[216] K. Kousias, G. Caso, O. Alay, A. Brunstrom, L. D. Nardis, M.-g. D. Benedetto, and M. Neri, "Coverage and Deployment Analysis of Narrowband Internet of Things in the Wild," *IEEE Communications Magazine*, vol. 58, no. 9, pp. 39–45, 9 2020.

[217] GSMA, "Mobile IoT Deployment Map," 2021. [Online]. Available: https://www.gsma.com/iot/deployment-map/

[218] Iridium, "Iridium network," 2022, available online: https://www.iridium.com/ (accessed on 18 June 2022).

[219] E. Husni, "Design of Automatic Identification System (AIS) Receiver for Low Earth Orbit (LEO) Satellite," *International Review on Modelling and Simulations (IREMOS)*, vol. 9, no. 6, p. 435, 12 2016.

[220] B. Vejlgaard, M. Lauridsen, H. Nguyen, I. Z. Kovács, P. Mogensen, and M. Sørensen, "Coverage and Capacity Analysis of Sigfox, LoRa, GPRS, and NB-IoT," in *Vehicular Technology Conference*. IEEE, 2017.

[221] P. Masek, M. Stusek, K. Zeman, J. Hosek, K. Mikhaylov, S. Andreev, Y. Koucheryavy, O. Zeman, J. Votapek, and M. Roubicek, "Tailoring NB-IoT for Mass Market Applications: A Mobile Operator's Perspective," in *2018 IEEE Globecom Workshops (GC Wkshps)*. Abu Dhabi: IEEE, 12 2018, pp. 1–7.

[222] B. Reynders, Q. Wang, P. Tuset-Peiro, X. Vilajosana, and S. Pollin, "Improving Reliability and Scalability of LoRaWANs Through Lightweight Scheduling," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1830–1842, 6 2018.

[223] CLS Telemetry, "The future of Argos," 2021. [Online]. Available: https://www.cls-telemetry.com/argos-solutions/the-future-of-argos-argos4ng/

[224] L. Feltrin, G. Tsoukaneri, M. Condoluci, C. Buratti, T. Mahmoodi, M. Dohler, and R. Verdone, "Narrowband IoT: A survey on downlink and uplink perspectives," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 78–86, 2 2019.

[225] M. Paonni, M. Anghileri, S. Wallner, J. A. Avila-Rodriguez, and B. Eissfeller, "Performance assessment of GNSS signals in terms of time to first fix for cold, warm and hot start," *Institute of Navigation - International Technical Meeting 2010, ITM 2010*, vol. 2, no. January, pp. 1221–1236, 2010.

[226] M. Seppänen, J. Ala-Luhtala, R. Piché, S. Martikainen, and S. Ali-Löytty, "Autonomous Prediction of GPS and GLONASS Satellite Orbits," *Navigation*, vol. 59, no. 2, pp. 119–134, 6 2012.

[227] M. Lytvyn, A. Kemetinger, and P. Berglez, "How can an orbit prediction module speed up the TTFF and help to authenticate the position?" in *2012 6th ESA Workshop on Satellite Navigation Technologies (Navitec 2012) & European Workshop on GNSS Signals and Signal Processing*. IEEE, 12 2012, pp. 1–6.

[228] J. Pihlajasalo, H. Leppäkoski, S. Ali-Löytty, and R. Piché, "Improvement of GPS and BeiDou extended orbit predictions with CNNs," *2018 European Navigation Conference, ENC 2018*, pp. 54–59, 2018.

[229] X. Yang, "Low Earth Orbit (LEO) Mega Constellations - Satellite and Terrestrial Integrated Communication Networks," Ph.D. dissertation, University of Surrey, 2018.

[230] RX Networks, "Predicted GNSS," 2021. [Online]. Available: https://rxnetworks.com/location.io#!P-GNSS

[231] Semtech, "LoRa Cloud Geolocation," 2022, available online: https://www.loracloud.com/ (accessed on 18 June 2022).

[232] H. Zhu, K. F. Tsang, Y. Liu, Y. Wei, H. Wang, C. K. Wu, and W. H. Wan, "Index of Low-Power Wide Area Networks: A Ranking Solution toward Best Practice," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 139–144, 2021.

[233] J. Dian and R. Vahidnia, "LTE IoT Technology Enhancements and Case Studies," *IEEE Consumer Electronics Magazine*, pp. 1–1, 2020.

[234] O. Montenbruck, P. Steigenberger, L. Prange, Z. Deng, Q. Zhao, F. Perosanz, I. Romero, C. Noll, A. Stürze, G. Weber, R. Schmid, K. MacLeod, and S. Schaer, "The Multi-GNSS Experiment (MGEX) of the International GNSS Service (IGS) - Achievements, prospects and challenges," *Advances in Space Research*, vol. 59, no. 7, pp. 1671–1697, 2017.

[235] The Radio Technical Commission for Maritime Services, "RTCM standards," Arlington, 2021. [Online]. Available: https://www.rtcm.org/

[236] Institute of Navigation, "GNSS Software Defined Receiver Metadata Standard," 2021. [Online]. Available: https://sdr.ion.org/

[237] Ubiscale, "Ubiscale technology," 2022, available online: https://ubiscale.com/ (accessed on 10 June 2022).

[238] J. Curran, C. Fernandez-Prades, A. Morrison, and M. Bavaro, "Innovation: The continued evolution of the GNSS software-defined radio," *GPS World*, 2018. [Online]. Available: https://www.gpsworld.com/innovation-the-continued-evolution-of-the-gnss-software-defined-radio/

[239] European GNSS Agency, "GNSS User Technology Report," Tech. Rep. 2, 2018.

[240] A. Minetto, G. Falco, and F. Dovis, "On the Trade-Off between Computational Complexity and Collaborative GNSS Hybridization," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, vol. 2019-Septe. IEEE, 9 2019, pp. 1–5.

[241] J. C. Bermúdez Ordoñez, R. M. Arnaldo Valdés, and F. Gómez Comendador, "Energy Efficient GNSS Signal Acquisition Using Singular Value Decomposition (SVD)," *Sensors (Basel, Switzerland)*, vol. 18, no. 5, 2018.

[242] European Space Agency, "ESA plans for low-orbiting navigation satellites," 2022, available online: https://www.esa.int/Applications/Navigation/ESA_plans_for_low-orbiting_navigation_satellites (accessed on 30 October 2022).