DGK Ausschuss Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 993

Hossein Shoushtari

Data-Driven Inertial Localization Corrected by 5G Uplink-TDoA Sparse Positions

München 2025

Verlag der Bayerischen Akademie der Wissenschaften, München

ISSN 0065-5325 ISBN 978 3 7696 5385 4

DGK Ausschuss Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 993

Data-Driven Inertial Localization
Corrected by 5G Uplink-TDoA Sparse Positions

Dissertation
zur Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)
an der
HafenCity Universität Hamburg

vorgelegt von

M.Sc. Hossein Shoushtari

München 2025

Verlag der Bayerischen Akademie der Wissenschaften, München

ISSN 0065-5325 ISBN 978 3 7696 5385 4

Adresse des Ausschusses Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften:

Ausschuss Geodäsie (DGK) der Bayerischen Akademie der Wissenschaften
Alfons-Goppel-Straße 11 ● D – 80 539 München
Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100
e-mail post@dgk.badw.de ● http://www.dgk.badw.de

Vorsitzender:	Univ.-Prof. Dr. Jörg Pohlan, HafenCity Universität Hamburg
Erstgutachter:	Univ.-Prof. Dr.-Ing. Harald Sternberg, HafenCity Universität Hamburg
Zweitgutachter:	PD Dr. rer. nat. Lasse Klingbeil, Rheinische Friedrich-Wilhelms-Universität Bonn
Zusätzliche Professur:	Univ.-Prof. Dr.-Ing. Youness Dehbi, HafenCity Universität Hamburg

Einreichungsdatum: 30.04.2025
Verteidigungsdatum: 14.07.2025
Veröffentlichungsdatum: 14.10.2025

Diese Dissertation ist auf dem Server des Ausschusses Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften unter <http://dgk.badw.de/>
sowie in der Bibliothek der HafenCity Universität Hamburg unter
<https://repos.hcu-hamburg.de/handle/hcu/1158> digital publiziert

ISSN 0065-5325 ISBN 978 3 7696 5385 4

Abstract

Positioning in Global Navigation Satellite System (GNSS)-challenged areas for real-life applications, such as mobile location-aware services, augmented reality, and location history-based applications remains a long-awaited technology, particularly as a general or GNSS-comparable solution. The emergence of powerful 5G-enabled wearable devices, equipped with advanced multi-sensor systems, along with rapid developments in Artificial Intelligence (AI), presents opportunities for indoor pedestrian localization. These technologies support a wide range of use cases for visually impaired, firefighters, in hospitals or shopping as well as critical safety scenariossuch as contact tracing during the COVID-19 pandemic.

This research aims to develop a GNSS-comparable method to enable long-term pedestrian navigation for real-world applications. While data-driven, positioning methods, so called Inertial Localization (IL), using smartphone inertial sensors can provide relative displacement estimates over extended periods, they still suffer from drift over time, necessitating external corrections. Smartphones, however, can access sparse yet accurate 5G positioning data, which can be used to integrate corrections; thereby making long-term and reliable pedestrian navigation feasible.

To tackle open challenges, a data-driven inertial localization model using deep learning and a novel transformer network is proposed, Transformer Automatic Features Interaction (TAFI), which combines state-of-the-art machine learning with engineered physics-based features derived from domain expertise. A dataset of real 5G Uplink Time Difference of Arrival (TDoA) sparse positions, collected in cooperation with Fraunhofer Institut für Integrierte Schaltungen (IIS) 5G Bavaria Testbed, supports this approach and provides a better understanding of 5G positioning. The correction mechanism, in format of state estimation algorithms, automatically computes heading to correct the pose.

The method's performance was evaluated following the ISO 18305 standard. The results show that a GNSS-comparable with a 3-meter accuracy in 90% of the cases is feasible by integrating Inertial Measurement Unit (IMU) data and sparse but accurate 5G positions. The research finding establish a foundation for further research into data-driven IL using 5G positioning data.

Zusammenfassung

Die Positionierung in Global Navigation Satellite System (GNSS)-beeinträchtigten Gebieten für Anwendungenwie mobile ortsbezogene Dienste, Augmented Reality und ortsbezogene Anwendungen ist eine lang erwartete Technologie, insbesondere als allgemeine oder mit GNSS vergleichbare Lösung. Das Aufkommen leistungsfähiger 5G-fähiger tragbarer Geräte, die mit fortschrittlichen Multisensorsystemen ausgestattet sind, sowie die rasanten Entwicklungen im Bereich der künstlichen Intelligenz (KI) bieten Möglichkeiten für die Lokalisierung von Fußgängerinnen und Fußgängern in Gebäuden. Diese Technologien unterstützen eine breite Palette von Anwendungsfällen für Sehbeeinträchtigte, Feuerwehrleute, in Krankenhäusern oder beim Einkaufen sowie kritischen Sicherheitsszenarien, wie z. B. die Kontaktverfolgung während der COVID-19-Pandemie.

Ziel dieser Forschungsarbeit ist die Entwicklung einer mit dem GNSS vergleichbaren Methode, die eine langfristige Fußgänger:innennavigation für Anwendungen in Praxis ermöglicht. Zwar können datengesteuerte Positionierungsmethoden, so genannte Inertial Localization (IL), die Inertial Measurement Unit (IMU)-Sensoren von Smartphones verwenden, Schätzungen der relativen Verschiebung über längere Zeiträume liefern, doch leiden sie immer noch unter einer zeitlichen Drift, so dass externe Korrekturen erforderlich sind. Smartphones können jedoch auf spärliche, aber genaue 5G-Positionsdaten zugreifen, die zur Integration von Korrekturen verwendet werden können, wodurch eine langfristige und zuverlässige Fußgänger:innennavigation möglich wird.

Um offene Herausforderungen zu bewältigen, wird ein lernbasiertes Modell zur Trägheitslokalisierung vorgeschlagen, das Deep Learning und ein neuartiges Transformator-Netzwerk verwendet, TAFI, welches maschinelles Lernen mit physikalisch basierten Merkmalen kombiniert, die aus dem Fachwissen abgeleitet sind. Ein Datensatz von realen 5G Uplink Time Difference of Arrival (TDoA) sparse Positions, der in Zusammenarbeit mit dem Fraunhofer Institut für Integrierte Schaltungen (IIS) 5G Bavaria Testbed gesammelt wurde, ermöglicht ein besseres Verständnis der 5G-Positionierung und unterstützt den lernbasierten Ansatz. Der Korrekturmechanismus berechnet automatisch die Richtung und korrigiert die Pose.

Die Leistung der Methode wurde in Anlehnung an die Norm ISO 18305 bewertet. Die Ergebnisse zeigen, dass eine GNSS-ähnlicher Lösung mit einer Positionierung von 3-Meter-Genauigkeit in 90% der Fälle durch die Integration von IMU-Daten und spärlichen, aber genauen 5G-Positionen möglich ist. Die Forschungsergebnisse bilden die Grundlage für weitere Forschungen zur datengesteuerten IL unter Verwendung von 5G-Positionsdaten.

Acknowledgements

Firstly, I would like to thank my doctoral supervisor, Harald Sternberg, for his open and trusting support, especially during challenging times. Thank you for being my 'doktorvater'.

I would also like to thank Lasse Klingbeil for his positive comments on my research, his feedback, inspiration and discussions.

I would also like to express my sincere thanks to the Level 5 Indoor Navigation (L5IN) project team, starting with Jörg Müller-Lietzkow, without whom the acquisition of this project would not have been possible. I would also like to thank Dorian Harder, Cigdem Askar, Elena Falcini for their support with technical questions, investigations and corrections, as well as for their strengthening and encouraging conversations.

Thank you to my colleagues in the working group and the students for providing such a pleasant working environment, for your many helpful gestures, for your openness and for simply being nice. Special thanks to Firas Kassawat, Fickrie Muhammad, Annette Schneider, Clemens Semmlroth, Korvin Venzke and Esteban David.

Next thanks go to my parents, Morteza and Tahereh; my brother Mahdi; now truly considered as my sister, Sahar; and my wife, Saeideh, for your constant support, encouraging conversations and unwavering trust. This would not have been possible without you <3.

Finally, I would like to thank Thomas Willemsen from Hochschule Neubrandenburg, Maximilian Kasparek from Fraunhofer IIS, Michael Klingen from Breuer Nachrichtentechnik GmbH, and Dietmar Kohnenmergen from 1NCE for their positive and supportive collaboration. The world truly needs such constructive and inspiring individuals.

Hossein Shoushtari, 30.10.2025

Contents

1 Preface

This cumulative dissertation explores advancements in Inertial Localization (IL) through data-driven (learning-based) approaches. It further addresses the correction of relative trajectories derived from IL using sparse location data from 5G Uplink Time Difference of Arrival (TDoA) positioning, enabling real-life applications of long-term indoor localization with wearable devices such as smartphones. The thesis is based on seven publications, all of which underwent a peer-review process, except for the publication B2, which is a proof of concept used in publication B1. Notably, the three main objectives of this thesis are classified into objective groups A, B, and C (See Section 2.2), with the first publication in the objective groups, i.e. A1, B1 and C1, serving as the primary criterion for thesis evaluation at HafenCity University, Hamburg. The publications are summarized in Chapter 5, with Chapter 6 and 7 highlighting the key findings and numerical results. The author of this dissertation has made the primary contribution to all these publications, particularly by introducing the methodological advancements.

- Publication A1:
  Shoushtari, H., Harder, D., Kasparek, M., Müller-Lietzkow, J., Sternberg, H. 2023; Data-Driven Inertial Navigation assisted by UL-TDOA 5G Positioning, International technical Meeting (ITM), Institute of Navigation (ION), Long Beach, California. https://doi.org/10.33012/2023.18645
- Publication A2:
  Shoushtari, H., H., Kassawat, F., Harder, D., Venzke, K., Müller-Lietzkow, J., Sternberg, H., 2022; L5IN+: From an Analytical Platform to Optimization of Deep Inertial Odometry, International Conference on Indoor Positioning and Indoor Navigation (IPIN) , Beijing, China.
  https://ceur-ws.org/Vol-3248/paper24.pdf
- Publication B1:
  Shoushtari, H., Kassawat, F., Sternberg, H. 2024; Context aware transformer network and in-situ IMU calibration for accurate positioning, IPIN, Hongkong.
  https://doi.org/10.1109/IPIN62893.2024.10786130
- Publication B2:
  Shoushtari, H., Willemsen, T., Sternberg, H. 2023; Supervised Learning Regression for Sensor Calibration, DGON Inertial Sensors and Systems (ISS), Braunschweig, Germany.
  https://doi.org/10.1109/ISS58390.2023.10361922

- Publication C1:
  Shoushtari, H., Harder, D., Willemsen, T., Sternberg, H. 2023; Optimized Trajectory from Smartphone Sensors and 5G UL-TDoA Using Cluster Particle Filters (Original title in german: Optimierte Trajektorie aus Smartphone-Sensoren und 5G UL-TDoA mit Cluster Partikel Filter), Internationalen Ingenieurvermessung, Zürich, Hrsg.: Andreas Wieser, Wichmann Verlag, ISBN/EAN:9783879077342
- Publication C2:
  Shoushtari, H., Askar, C., Harder, D., Willemsen, T., Sternberg, H., 2021; 3D Indoor Localization using 5G-based Particle Filtering and CAD Plans, IPIN , Lloret de Mar, Spain.
  https://doi.org/10.1109/IPIN51156.2021.9662636
- Publication C3:
  Shoushtari, H., Willemsen, T., Sternberg, H., 2021; Many Ways Lead to the Goal-Possibilities of Autonomous and Infrastructure-Based Indoor Positioning, Electronics Special Issue Indoor Positioning Techniques.
  https://doi.org/10.3390/electronics10040397

2 Introduction

Autonomous navigation paradigms, from sensor-based to map-based ones, are highly relying on localization, which nowadays works as one of the key tasks in this domain. Inertial Localization (IL) methods, which utilize inertial sensors, specifically Inertial Measurement Unit (IMU) sensors to estimate position, have established themselves as standard tools in various indoor and outdoor environments (Thrun et al., 2005).
The idea of locating people and objects in indoor environments dates back to the 1990s when the Real Time Indoor Positioning System (IPS) was deployed in hospitals to track equipment like wheel chairs, and warehouses to track merchandise, etc. (Malik, 2013; Want et al., 1992). In the IPS domain, IL serves as a critical technique for positioning in Global Navigation Satellite System (GNSS)-challenged areas where their signals are unavailable. This enables applications such as pedestrian navigation, aerospace instances, autonomous navigation of vehicles and drones, and contact and asset tracking in different use cases (Sternberg & Fessele, 2009). For pedestrian indoor positioning, IMU sensors are particularly popular due to the widespread availability of (Micro Electro Mechanical System) MEMS-IMU sensors embedded in wearable devices such as smartphones, smartwatches, and smart glasses (Keller & Sternberg, 2013; Willemsen, 2016). Despite advancements in motion sensors within wearables, indoor localization has remained a long-anticipated technology. This is expected to change soon.

This chapter presents the motivation for the research, focusing on pedestrian indoor positioning using smartphones. It explores how machine learning and data-driven approaches have emerged as key solutions, while also highlighting the continued necessity for optimization and correction methods. In this context, 5G sparse positioning is considered a potential correction source. The main objectives of the thesis are then outlined from a broad perspective before being broken down into three specific objectives.

2.1 Motivation

Pedestrian Indoor Positioning with Smartphones
As of 2024, the number of smartphone users worldwide is projected to reach approximately 7.1 billion. In Germany, the smartphone penetration rate exceeds 80%, for instance (Statista, 2024). Smartphones, as multi-sensor systems, integrate a variety of preinstalled sensors, processors, and displays, making them a versatile platform for research and development (Sternberg, 2013). Their extensive adoption by a massive user

base further underscores their significance in technological innovation. In the context of pedestrian localization and navigation use cases, smartphones remain the device most commonly carried by individuals. Existing navigation applications, such as those developed by Google and Apple, already provide robust location-based services for outdoor navigation. However, IPS have gained increasing attention due to their potential to extend navigation capabilities to indoor environments due to the GNSS signals limitations.

IL Meets Machine Learning

The term inertial sensor in IL refers to the combination of a three-axis accelerometer and a three-axis gyroscope. Devices containing these sensors are commonly known as IMUs. IL methods can provide information about the pose (orientation and position) of any object to which they are rigidly attached using these inertial sensors. Integration of gyroscope measurements provides information about the sensor's orientation, while double integration of accelerometer measurements, after subtracting the Earth's gravity, provides information about the sensor's position. This approach is commonly referred to as Strapdown methods (Titterton et al., 2004).

Pedestrian Dead Reckoning (PDR) utilizes inertial measurements to detect steps, estimate stride length and heading, and ultimately determine position. Pedestrians can take advantage of this information along with estimated average stride length to track their relative position (Li et al., 2012). The extensive tuning required for PDR systems (e.g., step length estimation) represents a significant barrier to widespread adoption (Xiao et al., 2014). Researchers consider using more data to learn and adapt to multiple walking characteristics, and this is how IL meets machine learning domain.

Data-driven IL, particularly supervised learning approaches, presents promising opportunities addressing indoor positioning challenges related to varying device placement and changes in smartphone positioning during walking as common behaviors observed in real-life experiments (Chen, Lu, et al., 2018). Instead of relying on discrete step events, these algorithms use time windows and shift the windows during model training. This approach enables learning displacement patterns regardless of the device placement or the type of user movement. As an example of a real-life use case, consider a postal worker who interacts with their device in various scenarios: checking it, carrying it in a pocket or bag, answering a call, or placing it on a tray. Despite these dynamic conditions, the position estimation system is expected to maintain functionality and sustain submeter-level accuracy over an extended duration. In this context, data-driven IL methods have demonstrated significant advantages over Strapdown and PDR approaches, offering improved accuracy and sustained performance over extended durations with reduced error accumulation. (Yan et al., 2019).

Data-Driven IL Still Needs Support

Recently, data-driven IL has shown up to submeter-level positioning accuracy and durability for multiple minutes (Yan et al., 2019). However, considering a GNSS-comparable solution for indoor environments, regardless of the accuracy and durability of an IL method, some form of initialization pose and correction mechanism is required to start and then keep the trajectory drift within acceptable limits. Such a system must be capable of consistently and instantaneously delivering absolute position estimates, as opposed to relative ones. While the error accumulation in state-of-the-art MEMS sensors can be reduced, it cannot be entirely removed. Therefore, in the indoor positioning domain, as in other localization use cases outdoors, the localization task does not solely rely on IMU sensors. There are two main categories of real-time correction mechanisms (more details in 6.5-C. Correction of relative trajectories using sparse positional data). The first involves the use of maps, and the second relies on some form of position information to determine and use an absolute position. Using a map, through a so-called map-matching process, the relative trajectory can be laid down on walkable areas on the map. An algorithm could also request a short walk to compare the relative trajectory derived from IL with the map to address initialization tasks (Herath et al., 2022). By using absolute position information, the second group of algorithms could initialize and reinitialize the trajectory derived from IL, thereby enhancing accuracy and mitigating drift.

5G positioning

The fifth generation (5G) of cellular networks has been developed not only to revolutionize communication but also to offer advanced positioning capabilities. With the proliferation of powerful 5G-enabled smartphones, IPS can benefit significantly from the inclusion of 5G technology. In indoor environments, 5G holds the potential to act like the GNSS infrastructure and solutions, particularly due to the standardization and organizational support behind it and the widespread adoption of 5G-enabled devices, which is expected to provide most smartphone users, above 80% in Europe and North America, access to 5G-based positioning services within the next few years (Statista, 2024).

The Third Generation Partnership Project (3GPP), a global collaborative initiative for mobile communication standardization, has introduced several releases aimed at enhancing 5G positioning capabilities. Latest 3GPP standards require 5G systems to achieve positioning accuracy of 3 meters at 80% reliability in indoor environments and 10 meters at 80% reliability outdoors. Future releases, aim to push this boundary further, targeting sub-meter accuracy of up to 0.5 meters at 90% reliability or higher (3GPP, 2024). The Uplink Time Difference of Arrival (TDoA) positioning method is among the earliest 5G positioning techniques capable of addressing the limitations in accuracy definitions

However, despite its potential, 5G positioning faces practical limitations. For instance, deploying a network capable of providing sub-meter accuracy uniformly across all indoor spaces may not be financially viable. Additionally, 5G positioning shares limitations similar to GNSS in terms of coverage, particularly in environments with restricted signal penetration.

In the context of pedestrian IPS, several challenges emerge. Devices are inherently limited by battery capacity, making it impractical to continuously operate 5G networks at their maximum positioning frequency rate, as this would deplete device power within minutes. Consequently, usable absolute positioning data may only be available intermittently, potentially once every one to two minutes of walking. This phenomenon introduces the concept of sparse location data, where position data is provided intermittently based on network availability and configuration (Herath et al., 2021).

2.2 Objectives

To address the challenges and opportunities mentioned above, the overarching goal of this thesis is to develop and advance a solution for pedestrian IPS that is comparable to GNSS in terms of functionality and availability. This research aims to consider the practical constraints and requirements of real-life applications, ensuring the proposed system is robust, efficient, and user-centric (See Figure 2.1).

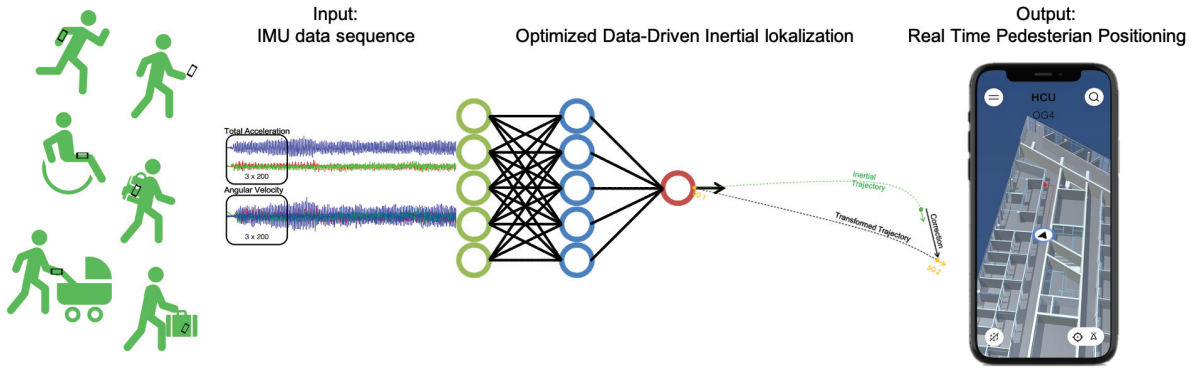


Figure 2.1: Real-life data-driven IPS supporting variations in smartphone placements.

This research focuses on pedestrian indoor positioning using smartphones. During the algorithm development phase, a broader range of wearable devices and diverse IMU sensors will be considered to enhance the accuracy and adaptability of IL methods. Leveraging the IMU sensors available in smartphones, their quality, and real 5G Uplink TDoA positioning data, this work aims to collect and utilize its own dataset to improve the state-of-the-art data-driven IPS and 5G Uplink TDoA positioning approaches in terms

of accuracy, reliability, availability, and durability. Sparse position updates derived from the absolute Uplink TDoA positions, will be employed for initialization and correction, while the investigation of map-based approaches may also be explored with respect to complexity. This research emphasizes extensive qualitative and quantitative evaluations of existing baselines and state-of-the-art IL methods across multiple benchmarks. The main objective of this thesis can be divided into three smaller objectives:

- **A:** Evaluation of smartphone- and 5G-based positional data (publication A1-A2),
- **B:** Improving and developing efficient data-driven IL approaches (publication B1-B2, partially in A2) and
- **C:** Design and development of map matching and correction approaches, compatible with the sparse location data (publication C1-C3).

A: Evaluation of smartphone- and 5G-based positional data
As a first step, evaluation and integrating smartphone sensors data with synchronized real 5G positioning data essential for advancing the next stages of development. The dataset collection will play also a pivotal role in identifying patterns, weaknesses, and potential areas for improvement in data-driven IL. These learning-based approaches rely heavily on data quality; unsynchronized or poor-quality datasets will inevitably lead to suboptimal model predictions. To better understand 5G positioning behavior, particularly in the absence of available datasets, a novel dataset should be collected, capturing raw 5G positioning signals and time measurements for Uplink TDoA under different line-Of-Sight (LOS) conditions. The evaluation of 5G positioning in indoor environments, including both narrow corridors and expansive spaces, must also be addressed. Furthermore, the development of simulation tools is critical for preparing to work with and process real data, ensuring continuity in algorithm development even when real 5G data is unavailable.

B: Improving and developing efficient data-driven IL approaches
Building upon the collected dataset and derived insights, this objective focuses on developing and enhancing data-driven IL techniques. The primary goal is to create a real-life compatible algorithm capable of handling diverse smartphone (also general wearables) placements and user behaviors. This includes scenarios such as producing step events during normal walking, handling activities like calling while walking, or accommodating situations where the phone is carried on a stroller (e.g., with no step events). Additionally, the algorithm aims to ensure accurate positioning and robust trajectory estimation over extended periods. Furthermore, the data-driven approach will integrate state-of-the-art machine learning methods, leveraging engineered features informed by physics and domain expertise. This combination of advanced machine learning and domain-specific feature

engineering is expected to improve algorithm novelty, performance and adaptability in real-world scenarios.

C: Design and development of map matching and correction approaches
This objective involves designing robust initialization and correction strategies that are compatible with sparse location data, addressing a critical challenge in real-life pedestrian IPS applications. Considering the capabilities and limitations of 5G positioning, as well as the goal of developing a GNSS-comparable IPS solution, this research aims to propose a realizable concept and solution for efficient initialization. The map matching strategy is expected to leverage sparse positional updates, and the map information, ensuring reliable system startup even with limited availability of absolute position data. Additionally, an automatic correction approach will be developed to optimize system accuracy over time. This should utilize sparse position points as infrequent correction anchors, minimizing their usage due to the constraints of sparse data availability and the battery limitations of user devices.

3 Theoretical Background

This chapter provides a brief overview of the most important and relevant theoretical knowledge related to data-driven inertial Inertial Localization (IL) and 5G positioning, which is essential for understanding the content of this thesis. This chapter encompasses three distinct topics: first, the inertial sensors and their calibration are explored, then the orientation as foundational input for IL; third, the theoretical background of deep learning models, with a particular emphasis on sequence models; and finally, an overview of 5G positioning. In line with the concept of a cumulative dissertation, the topics are presented at an intuition level that ensures a comprehensive understanding of the thesis content. However, for more detailed explanations, extensive references are included.

3.1 Coordinates, Frames and Inertial Sensors

Coordinate Systems

In the context of data-driven IL and 5G positioning, there are many coordinate systems and frames involved, each of them introduced and associated with a certain type of coordinates, either Cartesian, Polar coordinate system, etc. Therefore, different coordinates, and coordinate systems and frames are briefly defined, but they are well detailed in (Savage, 2000; Torge et al., 2023).

Cartesian coordinates are denoted as x, y, and z, representing the perpendicular distances of a point from the origin, projected onto the system's first, second, and third axes, respectively. Polar coordinates describe a point based on its distance L from the origin (called the *pole*) and the direction ϕ, which indicates the angle between the polar axis, a fixed ray from the pole, and the line connecting the point to the pole. Ellipsoidal or geodetic coordinates are defined by ellipsoidal latitude, longitude, and height. These coordinates are referenced to an ellipsoidal model of the Earth, making them particularly useful in Global Navigation Satellite System (GNSS) positioning.

Next, the frames are elaborated, inspired from (Klingbeil, Lasse, 2023). A frame is generally a realization of a coordinate system. The inertial frame (i-frame) is a non-rotating, space-fixed frame with its origin in the barycenter of the solar system or, as in this thesis, in the Earth's center of mass. The third axis is usually aligned in parallel to the Earth's mean rotation axis, the first axis is pointing towards the vernal equinox (which is by definition orthogonal to the third axis) and the second axis completes the right-handed coordinate system. The inertial frame is the reference for inertial sensors

used for trajectory estimation.

The terrestrial frame (e-frame) is situated in the center of mass of the Earth. It is an Earth Centered, Earth-Fixed (ECEF) frame, rotating synchronously with the Earth around its third axis. In general, the first axis points towards the Greenwich meridian in the equatorial plane and the second axis completes the right-handed system. Positions are expressed in terms of ellipsoidal coordinates. As examples, the World Geodetic System 1984 (WGS84) reference frame are mentioned, which both include the reference ellipsoid Geodetic Reference System 1980 (GRS80) with a minute difference in the flattening (Heinz, 2021).

Body frame (b-frame, also local or world frame) is attached to the moving platform. The origin is a point on the device or vehicle, e.g. the center of mass, the x-axis is pointing forward, the z-axis is pointing down, and the y-axis is completing the right-handed system (pointing to the right). It is different from Sensor Frame, a local system attached to the sensor, where raw sensor readings are given in this system. The variable relationship between the b-frame and sensor frame, as in real-life usage of wearable devices, cause challenges in position estimation, for instance using smartphone sensors (See Figure 3.1).

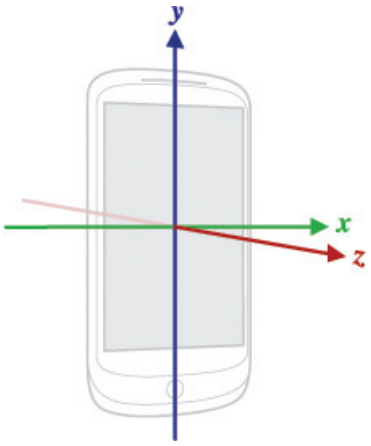


Figure 3.1: Android smartphone sensor coordinate system relative to a device (developer.android.com).

In Navigation Frame (n-frame) the origin is the same as for the body frame, the x-axis is pointing towards North, the z-axis is pointing down (parallel to gravity), and the y-axis is completing the right-handed system, pointing East. This system is also called North-East-Down (NED).

Inertial sensors

Inertial sensors build the core of most navigation systems, as they can provide information about the sensors motion. Two types of sensors are considered inertial; accelerometers

and gyroscopes (angular rate sensors). They measure the sensors angular rates and non-gravitational accelerations with respect to the inertial frame and can be used to determine both translational and rotational components of the trajectory.

Accelerometers measure the specific force acting on the sensor along its sensitive axis with respect to the inertial frame. The specific force is the nongravitational per unit mass, which is a dynamic acceleration. An accelerometer, standing motionless on the ground, measures $9.81\,\mathrm{m/s^2}$ or 1g antiparallel to gravity. Gyroscopes measure the angular rate of the sensor around its sensitive axis with respect to the inertial frame. A high-grade gyroscope lying motionless on the ground would still measure the Earths rotation rate, However, Micro Electro Mechanical System (MEMS) sensors are not sensitive enough to measure this. MEMS are tiny sensors, which can detect mechanical, magnetic, or even chemical changes and convert them into electrical information. These sensors are often combined as three-axis sensors in smartphones, such as the MPU-9250 (TDK InvenSense, Tokyo, Japan), and are called Inertial Sensor Assembly (ISA). The ISA sensors should be calibrated by the manufacturer periodically. A calibrated ISA, or an Inertial Measurement Unit (IMU) is the combination of three accelerometers and three gyroscopes perpendicular aligned to cover all spatial axis.

Calibration of inertial sensors

Calibration of inertial sensors is a critical task to ensure the accuracy and reliability of sensor systems. This process involves establishing a functional relationship that takes into account the interaction between the sensor data, considering sources of error such as bias, scale, orthogonality errors and temperature effects, and a known and error-free reference value. The aim is to match the measurements to references and estimate the calibration parameters for further use of the sensor values. In order to estimate these parameters and to assess the quality of the estimation, a series of repeated measurements are performed and therefore optimization methods need to be applied (Aggarwal et al., 2010).

The authors in (Poddar et al., 2017), summarize and review the works related to inertial sensor calibrations, where the methods are divided into two categories, those using high precision calibration equipment, and without equipment. In the first group, there is usually a rotating platform in which the exact rotation can be measured, such as works in (Willemsen, 2016) for smartphone sensors, and (Olivares et al., 2009) also using a total station. To consider the temperature, a thermal test setup should be used (Aggarwal et al., 2008). For the calibration of accelerometers without equipment, the methods rely on gravity as a stable physical calibration standard (Fong et al., 2008).

3.2 Attitude and heading reference systems

An Attitude and Heading Reference System (AHRS) uses an IMU consisting of MEMS, typically including tri-axis gyroscopes and accelerometers, and magnetic sensors, to measure Magnetic, Angular Rate and Gravity (MARG) values respectively. These measurements can then be used to derive an estimate of the object's attitude (Madgwick et al., 2011).

Gyroscopes have been successfully employed to determine system orientation in various applications. However, it is well known that they suffer from drift (Pandit et al., 1986; Van Dierendonck & Brown, 1969), which manifests itself as a time-variant low-frequency bias, leading to significant orientation errors over long-term operation (Han, 2009). In the case of low-cost gyroscopes, such as MEMS-based sensors found in smartphones, the drift rate can reach several degrees per second. Even for high-precision mechanical gyroscopes, accumulated errors remain a critical concern that cannot be overlooked.

Accelerometers and magnetometers, while can provide an absolute reference of orientation free from long-term drift effects, are susceptible to significant short-term noise, particularly in MEMS sensors. For accelerometers, this noise predominantly manifests itself as a time-variant low-frequency bias mixed with white noise, making it challenging to distinguish between dynamic acceleration and the inherent sensor noise. This issue is especially pronounced in MEMS accelerometers due to their small size and minimal mass, which increase their sensitivity to various noise sources (Mohd-Yasin et al., 2009). Magnetometers, on the other hand, are highly sensitive to environmental magnetic interference. Factors such as nearby ferromagnetic materials can significantly affect their readings, leading to inaccuracies in some applications. This environmental sensitivity requires hard and soft iron calibration and compensation techniques to ensure accurate measurements (Willemsen, 2016; Q. Zhang et al., 2016).

The task of an orientation filter or an AHRS is to compute a single estimate of orientation through the optimal fusion of gyroscope, accelerometer and optionally magnetometer measurements. The Kalman Filter (Kalman, 1960) has become the accepted basis for majority of filter algorithm to integrate gyroscopes values with some other augmenting sensors to provide a stable orientation solution. Parameter tuning and large computational load for implementation of Kalman-based solutions, provide a motivation for alternative approaches. The Madgwick Filter (Madgwick et al., 2011) is an orientation filter that is applicable to both IMUs and MARG sensor arrays. The filter employs a quaternion (Kuipers, 1999; Shoemake, 1985) representation of orientation.

A quaternion is a four-dimensional vector composed of a scalar and a vector component.

The scalar component represents the rotation angle, while the vector component defines the rotation axis(Jekeli, 2001). A quaternion can be expressed as:

$$\mathbf{q} = a + \boldsymbol{i}b + \boldsymbol{j}c + \boldsymbol{k}d \tag{3.1}$$

where, $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are the imaginary units that satisfy the quaternion multiplication rules. On the other hand, a, b, c, and d represent real numbers. A quaternion can also be expressed in Equation 3.2 (Diebel, 2006) as:

$$\mathbf{q} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^\top = \begin{bmatrix} q_0 & q_{1:3} \end{bmatrix}^\top \tag{3.2}$$

Quaternion multiplication combines rotations, which allows concatenating multiple rotations by multiplying their respective quaternions. One notable benefit of using quaternions over Euler angles (Diebel, 2006; Jekeli, 2001) is their computational efficiency when performing rotations and transformations. Quaternion operations, such as multiplication, conjugation, and normalization, can be carried out more efficiently compared to the operations involved in Euler angle computations. This efficiency comes from the quaternions compact representation and the use of quaternion-specific algebraic rules. Another advantage of using quaternions inrotation representation is that they avoid the problem of gimbal lock, which occurs when using Euler angles. Quaternions do not suffer from this issue and provide a smooth and continuous representation of rotations (Madgwick et al., 2011).

The estimation of the IMU's orientation quaternion can be achieved by combining fused MARG values and its optimization techniques whenever inertial measurements are acquired. Each quaternion represents the rotation required to transform an IMU's frame into another frame whenever a measurement is obtained (Jekeli, 2001).

Consider v as an acceleration or gyroscope measurement in either the sensor frame or the local (world) frame, represented as a quaternion:

$$\mathbf{v} = \begin{bmatrix} 0 & a_x & a_y & a_z \end{bmatrix}^\top \quad \text{or} \quad \begin{bmatrix} 0 & \omega_x & \omega_y & \omega_z \end{bmatrix}^\top \tag{3.3}$$

Equations 3.4 and 3.5 perform the described transformations from the sensor to the world frame and vice versa (Diebel, 2006):

$$\mathbf{v}_{world} = \mathbf{q}\mathbf{v}_{sensor}\mathbf{q}^{-1} \tag{3.4}$$

$$\mathbf{v}_{sensor} = \mathbf{q}^{-1}\mathbf{v}_{world}\mathbf{q} \tag{3.5}$$

3.3 Sequence deep learning models

Machine learning was initially defined in the 1950s as the field of study that gives computers the ability to learn without being explicitly programmed (Samuel, 1959). Machine learning is fundamentally data-driven, with its primary objective being the development of general-purpose methods to extract meaningful patterns from data, ideally with minimal domain-specific expertise. To achieve this, models are designed to represent the underlying data generation process. As paraphrased from (Burkov, 1997), a model is considered to have learned if its performance improves with data. The ultimate goal is to create models that generalize well to unseen data by optimizing their parameters to automatically identify patterns and structures (Deisenroth et al., 2020).

Supervised machine learning, commonly referred to as supervised learning, encompasses algorithms that learn to map input features (x) to corresponding labels (y), effectively establishing input-output functional relationships. While regression algorithms are designed to predict continuous numerical values from an infinite range of possibilities, the second major category of supervised learning algorithms, known as classification algorithms, focuses on predicting discrete categories from a finite set of possible outcomes (Ng, 2024b). In addition to supervised learning, machine learning encompasses several other paradigms, notably unsupervised learning (Hastie et al., 2009; Murphy, 2012), semi-supervised learning (Chapelle et al., 2010), and reinforcement learning (Sutton & Barto, 2012).

Regression learning serves as the fundamental core of many learning algorithms. By extending the concept of linear regression, artificial neural networks emerge (Da et al., 2016). A single neuron can be viewed as a linear regression unit, inspired by the structure and function of biological neural networks in animal brains, forming the basis of neural network concepts (Mcculloch & Pitts, 1943). Deep learning architecture, as in Figure 3.2, is a contemporary term for the neural network approach to artificial intelligence, which has been studied for over 70 years (Poggio et al., 2017). Furthermore, a regression problem can also be transformed into a classification problem by applying an activation function, such as the sigmoid, softmax, tanh, or Rectified Linear Unit (ReLU). These functions convert predicted values into probabilities within the range of 0 and 1 (Ng, 2024b). This transformation enables the model to estimate the likelihood of an instance belonging to a particular category, forming the basis of the logistic regression algorithm (Tolles & Meurer, 2016).

For regression learning algorithms, the closer the predicted results to the label data, the lower the loss value, ideally a value of zero. In mathematical optimization a loss or cost

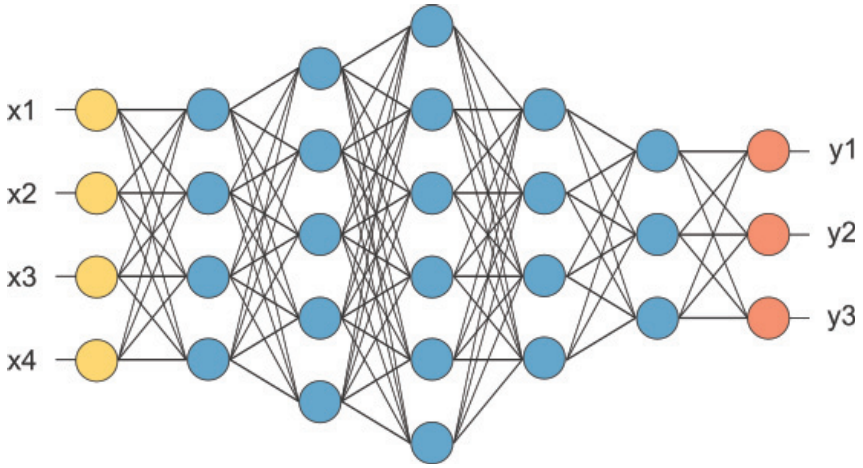


Figure 3.2: Illustration of a simple neural network architecture with an input layer in yellow, 5 hidden layers in blue and an output layer in orange (opennn.net).

function, also known as an error function, maps a event or variable values to a real number intuitively representing the "cost" of the event. The objective of machine learning is to optimize this cost function (Murphy, 2022). Given that learning algorithms often handle high-dimensional input and output vectors, the optimization process is typically expressed using numerical and iteration-based optimization techniques such as Gradient Descent (Deisenroth et al., 2020) or the first-order gradient-based optimization, so called Adam (Kingma & Ba, 2014). Optimization of a loss function through data and hidden layers, and thus, adjustments of parameters of each linear regression approximate model corresponding to a neuron, is done using forward and backward propagation in deep learning (Goodfellow et al., 2016). This method consists of two steps: forward propagation, predicting by letting the input data flow through the network, and backward propagation, where gradients of the loss function are calculated with respect to respective parameters (Deisenroth et al., 2020). This allows numerical optimization techniques to solicit updates to the parameters. Traditionally, matrices were only interpreted as inputting to neural networks. But in many real-world applications, sequence data is involved, and the ordering of elements is important.

Sequence models

Different neural network architectures have originally been designed for the specific challenges and structures of supervised learning data and tasks. Since Convolutional Neural Network (CNN)s are powerful tools used to process grid-like data, such as images, CNNs apply convolutional layers to learn the spatial hierarchies of features. This architecture has been very effective for computer vision, such as image recognition and classification (Murphy, 2022). Advanced models such as Residual Network (ResNet) have enabled the

training of very deep networks and improving performance in many tasks (K. He et al., 2015). Temporal Convolutional Network (TCN) is another recent proposed CNN architecture, which approximates recurrent architectures with dilated causal convolutions (Chung et al., 2014).

Recurrent Neural Network (RNN)s, on the other hand, are designed to handle the data by maintaining a form of state or memory that captures information from previous inputs. This capability makes RNNs well-suited for tasks where context and sequence are crucial, to build models for natural language, audio, and other sequence data. This type of model has demonstrated exceptional performance on temporal data, including Long Short-Term Memory (LSTM) networks, which address the vanishing gradient problem and enable long-range dependencies (Hochreiter & Schmidhuber, 1997); Gated Recurrent Unit (GRU)s, which simplify the architecture while maintaining performance efficiency (Ng, 2024a); and Bidirectional RNNs, which process information in both forward and backward directions (Schuster & Paliwal, 1997), improving contextual understanding in sequential data (Murphy, 2022).

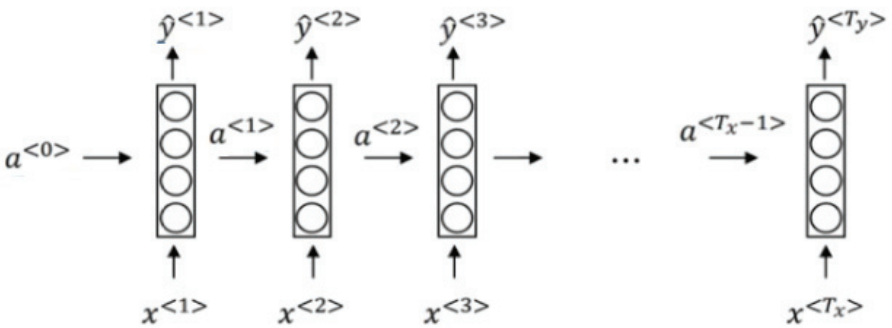


Figure 3.3: Illustration of a simple RNN architecture (Ng, 2024a)

The advantage of using RNN models for sequence tasks becomes evident when the lengths of inputs T_x and outputs T_y vary across different examples. Additionally, features learned across different positions of the sequence can be shared, which reduces the number of parameters in the model. Based on the relationship between input and output lengths, RNN architectures can be categorized into one-to-one, one-to-many, many-to-many, and many-to-one configurations (Murphy, 2022). There are an encoder and a decoder parts in some architectures. The encoder encodes the input sequence into one matrix and feed it to the decoder to generate the outputs. Encoder and decoder have different weight matrices (Cho et al., 2014). In an RNN architecture (see Figure 3.3), the current output $\hat{y}^{<t>}$ depends on the previous inputs and activations (Sutskever et al., 2014).

Transformers

Sequence models can be enhanced using an attention mechanism, which helps retain necessary information over long sequences. This, in turn, significantly increases the model's accuracy. The attention mechanism enables the model to determine where to focus given a sequence of inputs, using so-called contextual information. In other words, attention weights specify which inputs are required at each step to generate an output (Bahdanau et al., 2014). The attention mechanism was first introduced for machine translation but has been leveraged in several other fields, including computer vision, to create different novel architectures. A drawback to this algorithm is that it has a quadratic time or cost to run (Graves et al., 2014).

The Transformer that builds on the attention mechanism integrates the self-attention and multi-head attention mechanisms, and thus in a way enhances the traditional deep learning models by including some ideas from CNNs. This method allows for computations to be executed in parallel, which greatly increases efficiency by removing the sequential nature that had restricted progress in RNNs. Another important innovation in Transformers is positional encoding, which gives an index to input elements so that a non-sequential function can keep the sequence order of the inputs (Vaswani et al., 2017). This makes Transformers particularly effective for large-scale text processing tasks, as seen in models like BERT (Devlin et al., 2018) and GPT (Radford & Narasimhan, 2018), where contextual relationships and long-range dependencies are critical.

3.4 Positioning in 5G networks

Future indoor awareness requirements are expected to cause huge improvement in the communication network capacity, energy efficiency, and latencies of services as well as the number of connected devices. Therefore, the Fifth Generation New Radio (5G NR) network is the first cellular network which is not only designed with the sole focus on mobile phones, but also as basis for technologies such as Device-To-Device (D2D) or Internet of Things (IoT) devices such as smart meters, watches, or wearables (Kraeling & Brogioli, 2019; Wymeersch, 2017). The Third Generation Partnership Project (3GPP) is a global collaboration that sets the standards and protocols for mobile telecommunications.

The 3GPP has initiated the definition of 5G NR and its requirements have been structured under three main categories through Releases and Technical Specifications [1]:

- enhanced Mobile BroadBand (eMBB), whose aim is to provide wireless connectivity with very high bandwidth.
- massive Machine-Type Communication (mMTC), providing connectivity to a large number of IoT devices such as smart meters, watches, or wearables. mMTC requires very large cell and network capacity.
- Ultra-Reliable Low-Latency Communication (URLLC), targeted to provide low latency robust communication links for Vehicle-to-everything (V2X), remote surgery, and other safety-critical applications.

The 5G network includes other devices and use cases, referred to as industry verticals. Another significant difference with respect to previous cellular networks is the use of higher frequencies in the millimeter-Wave (mmWav) spectrum, starting at 24 GHz. 5G also brought important features such as massive Multiple Input, Multiple Output (MIMO), beamforming, cloud computing, and network virtualization (Shi & Jensen, 2011). All these features help increase the scalability and modularity of the network and to reach peak data rates of 20 Gbps with very high user density.

Measurement models

Release 16 specifies positioning signals and measurements for the 5G NR. They are summarized below, inspired from (García et al., 2020). Time of Arrival (ToA) measurement calculates the time lapse between the transmission and reception of a particular signal. Based on the propagation delay multiplied by the propagation speed of the signal (v_p) and considering the measurement error (e_X), the distance between the transmitter and receiver can be determined. Geometrically, the locus of points with a constant distance to a known point forms a sphere in 3D space. Assuming the position of either the transmitter or the receiver is known, the ToA measurement defines a sphere of location. This implies that the receiver is positioned on a sphere with a radius equal to the measured distance, centered at the transmitter, or vice versa. Using the equation of a sphere and incorporating measurement error, the observation equation is obtained.

A Round-Trip Time (RTT) measurement calculates the time lapse between the moment a signal is transmitted and the moment the response or acknowledgment to that signal is received. In the ideal case, the RTT measurement is twice the ToA measurement.

To eliminate the need for high-resolution time synchronization between the mobile ter-

[1] https://portal.3gpp.org (Accessed: 15 April 2025).

minal and base station, the Time Difference of Arrival (TDoA) between multiple pairs of antenna nodes can be measured. TDoA calculates the time lapse between the reception of two signals originating from different transmitters.

This is expressed in Equation (3.6), where $t_{\mathrm{rx}i}$ represents the time at which the signal from transmitter i is received:

$$\mathrm{TDoA} = t_{\mathrm{rx1}} - t_{\mathrm{rx2}} \tag{3.6}$$

The time difference measurement helps calculate the difference in distances between the receiver and each of the transmitters. As shown in Table 3.1, a hyperboloid equation is used for the observation model, where $\overline{(\mathrm{RxTx}_i)}$ denotes the distance between the receiver and transmitter i, and e_{TDoA} represents the error associated with the TDoA measurement.

The Phase of Arrival (PoA) measurement estimates the phase of the received signal and calculates the distance between the transmitter and receiver based on this phase measurement. If the transmitted signal is a sine wave with an initial phase of zero, the received phase ϕ_{rx} depends on the distance traveled by the signal, as given in Table 3.1, where k is a positive integer. The observation model can be derived by assuming that the propagation speed $v_p = c$ (the speed of light) and solving for the distance d, considering $\lambda = c/f$ as the signal wavelength.

An Angle of Arrival (AoA) measurement determines the angle at which a signal is received. Knowing the incoming angle of the received signal allows the receiver to estimate the transmitter's direction. The transmitter can be located at any point along the straight line defined by the angle α. Two classical methods are used to determine the AoA: by using the Phase Difference of Arrival (PDoA) or the geometric estimation-PDoA Measures the phase difference of signals received at two or more antennas. This method requires an antenna array with at least two elements at the receiver side and assumes the receiver is in the far field of the transmitter. Geometric estimation uses the known geometry of multiple antennas to infer the signal's arrival angle.

The Received Signal Strength (RSS) method estimates the distance between the transmitter and receiver based on the received power level. If the transmission power is known, the reception power can be used to estimate the signal's path loss. The distance D_{tra} can be computed using transmission equations such as Friis' equation (García et al., 2020). The error term e_{RSS} is considered in the observation model (See Table 3.1).

Table 3.1: Summary of Cellular-Based Measurements.

Measurement	**Model**
ToA	$(x-a)^2+(y-b)^2+(z-c)^2=(\mathrm{ToA}\cdot v_p)^2+e_{\mathrm{ToA}}$
RTT	$(x-a)^2+(y-b)^2+(z-c)^2=(\mathrm{RTT}/2\cdot v_p)^2+e_{\mathrm{RTT}}$
TDoA	$\overline{(\mathrm{RxTx}_2)}-\overline{(\mathrm{RxTx}_1)}=\mathrm{TDoA}\cdot v_p+e_{\mathrm{TDoA}}$
PoA	$(x-a)^2+(y-b)^2+(z-c)^2=\lambda\cdot(\phi_{\mathrm{rx}}/2\pi+k)$
AoA	$(x-x_1)^2+(y-y_1)^2+(z-z_1)^2=(l\cdot\sin\alpha)/\sin(\alpha+\beta)$
RSS	$\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}=D_{\mathrm{tra}}+e_{\mathrm{RSS}}$

While time-based measurements such as ToA, RTT, or TDoA require highly precise clocks, phase-based methods are primarily dependent on antenna design and phase ambiguity solutions. However, they can achieve millimeter-level accuracy using commercial off-the-shelf hardware[2]. AoA measurements, when combined with the aforementioned methods, enable single-anchor positioning, which is particularly advantageous for location based applications, in which they often fail to provide a financially viable solution. Among these, RSS measurements stand out due to their widespread availability and ease of acquisition. This accessibility makes RSS one of the most globally applicable sources for indoor positioning (see Section 4.2).

Positioning Technologies

Positioning measurements can be combined in various ways to calculate the coordinates (location) of a target object. Although the positioning measurements and parameters may be specific to cellular networks, the fundamental positioning algorithms remain consistent across different generations of cellular technology.

Proximity is the simplest positioning method, where the object's location is approximated by assuming it shares the same coordinates as the nearest antenna. More advanced approaches utilize multiple antennas and apply a weighted sum of their positions based on signal measurements. Triangulation determines the object's position by measuring angles from known reference points. In two-dimensional localization, at least two angle measurements from two known locations are sufficient to estimate the position. Trilateration, on the other hand, relies on distance measurements to compute the object's location. To obtain a 2D position, at least three reference points are required. Multilateration is based on TDoA observations. It calculates position by measuring the difference in distances between the object and multiple known reference points, and needs at least three measurements estimate a 2D position (García et al., 2020).

To standardize 5G positioning methods, 3GPP groups different positioning techniques

[2] https://www.koherent.io/ (Accessed: 7 October 2025).

based on whether they use Downlink or Uplink methods (García et al., 2020). However, this does not consider data-driven approaches, such as the widely used fingerprinting technique (Laska, 2023). Table 3.2 presents different compositions of 5G positioning technologies based on their underlying algorithms.

Table 3.2: Different compositions of 5G positioning technologies based on the methods.

Method	**Measurements**
Proximity	RSS, RTT
Triangulation	AoA (DL-AoD, UL-AoA)
Trilateration	ToA, RTT, RSS (NR-CID, Multi-RTT)
Multilateration	TDoA (DL-TDoA, UL-TDoA)

Positioning in 5G extends beyond Radio Access Technology (RAT)-based methods, incorporating alternative localization techniques such as GNSS, Bluetooth, and other RAT-independent approaches within the cellular network. Moreover, positioning can be executed either directly by the User Equipment (UE), in so called UE-based methods or managed centrally through the Location Management Function (LMF) within the 5G core network, so called LMF-based ones (Dwivedi et al., 2021).

4 Scientific Context

Indoor positioning solutions have gained considerable research attention in recent years. The launch of 5G networks and the worldwide growth of Artificial Intelligence (AI) have given a leverage to the problem of Indoor Localization. Many academic and industrial activities and events have occurred during this time, which has influenced the scientific context. Several milestones and events that shaped the trajectory for this dissertation are highlighted in Figure 4.1.

On the national level, the German Bundestag approved approximately 66 million euros in funding to support the research initiative of 5G technology, distributed by Bundesministerium für Verkehr und digitale Infrastruktur (BMVI). As part of this investment, six special grants were awarded in 2019 to research institutions and universities, with the aim of enabling digital transformation across social, economic, and administrative sectors, to contribute to long-term improvements in quality of life in Germany. The present dissertation was developed during the framework of one of these funded projects, Level 5 Indoor Navigation (L5IN), which began in 2019 (detailed in Section 4.4), coinciding with the authors research employment.

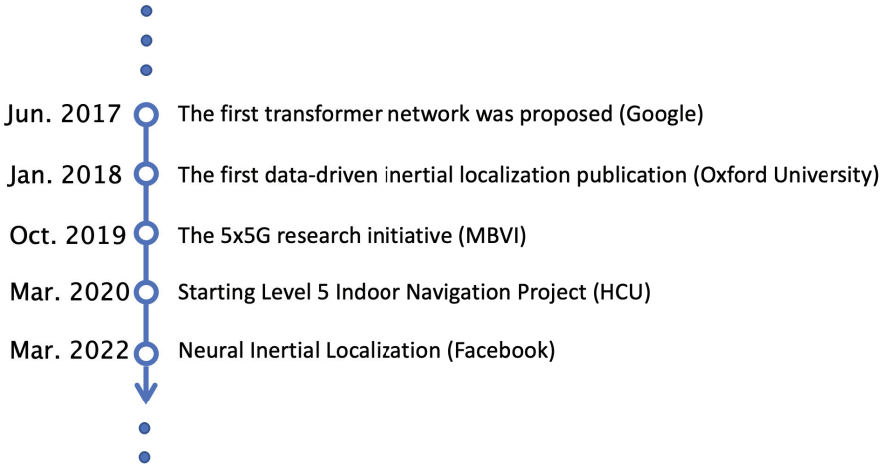


Figure 4.1: Some important events on the context of the thesis.

At the same period of time, on the global AI stage, the invention of Transformer networks, starting by researchers in Google with the landmark paper "Attention is all you need" (Vaswani et al., 2017), marked a turning point in machine learning domain. The highly active research domain came up with the first data-driven approaches in indoor positioning domain, as a usage of learning-based techniques to inertial sensor drift-reduction, launched

at University of Oxford (Chen, Lu, et al., 2018). These seminal works set the scientific tone that surrounded this thesis. The trend has been at its peak while companies such as Facebook have released the "neural inertial localization", using the transformer networks (Herath et al., 2022).

From 2020 to 2023, the COVID-19 pandemic has also highlighted the importance of pedestrian localization on public safety and mobility. But it also led to extended delays on research collaborations and event planning. For example, the organizers and steering committee of the 11^{th} Indoor Positioning and Indoor Navigation (IPIN) conference, among the most prestigious events in the field, had to cancel this year's conference with the pandemic having global repercussions.

Last but not least, contributing to the context of the thesis was the increasing of transformer network applications. While the paper "Improving Language Understanding by Generative Pre-Training" (Radford & Narasimhan, 2018) didn't gain much attention initially, its importance was made evident after OpenAI released ChatGPT in November 2022. This was reflective of AIs transformative potential, both its impact on the public and on the research community. It is worth to mention here, that the dissertation author utilized ChatGPT-4o during the writing process for proofreading.

This chapter aims to summarize the relevant state-of-the-art research in a concise and accessible format, with references to key works, starting with exploring different types of localization methods, followed by infrastructure-based positioning techniques. Next, the state estimation approaches, also known as correction or filtering approaches, are discussed. Lastly, L5IN project will be introduced.

4.1 Inertial localization

The use of inertial sensors to estimate the trajectory of a moving sensor, or a pedestrian or a platform carrying the sensor, is called Inertial Localization (IL). Although the general idea has been known for years (Savage, 2000), the data-driven approaches have only recently been introduced, and the leading benchmarks have named the problem differently, from inertial odometry (Chen, Lu, et al., 2018), to inertial navigation (Chen et al., 2020), to inertial localization (Herath et al., 2022). It should be noted that the terms odometry, navigation and localization have their standard definitions and differences, and they are different from the IL process, although these are taken into account.

Strapdown integration is among the first IL methods that can estimate the position by two time integrations of the dynamic acceleration (Titterton et al., 2004). Estimating

the dynamic acceleration, i.e. reducing gravity and other force effects such as Coriolis or centrifugal, is the challenging part, which may not be accurate with Micro Electro Mechanical System (MEMS) inertial sensors.

$$\mathbf{g} = \begin{bmatrix} 0 & 0 & 0 & 9.81 \end{bmatrix}^{\top} \tag{4.1}$$

Practical methods are used to overcome this challenge, focusing only on reducing the gravity values, such as estimating the Inertial Measurement Unit (IMU) sensor orientation, using Madgwick algorithms (Madgwick et al., 2011) and transforming the sensor values to the local frame (see Equation 3.4), where $9.81\,\mathrm{m/s^2}$ can be subtracted from the sensor z-axis. To check whether the quaternion vector computed an algorithms is the correct one, the inverse quaternion vector and the transfer function other way round (see Equation 3.5) can be used. The gravity vector in a quaternion way (See Equation 4.1) in sensor frame, should be the same or very close to the mean of the acceleration values in a time window, say 5 seconds for pedestrian motion. The mean vector has been shown to be very close to gravity using MEMS sensors, with respect to the sensor frame (Mizell, 2003). Although the process is not suitable for high-precision IMU sensors to accurately estimate the dynamic acceleration, it also works based on the definition of thresholds depending on the application, it can be used well for pedestrian localization purposes, therefore it is the basis of other IL methods (Xiao et al., 2014).

4.1.1 Step based methods

The combination of accelerometer and gyroscope is of particular interest to Pedestrian Dead Reckoning (PDR). The basis of PDR is the combination of a pedometer, a known or estimated stride length, and orientation. Several implementation examples and variants exist for realizing PDR. Notable approaches include multi-threshold step detection algorithms based on accelerometer and gyroscope signals (Li et al., 2012), the Madgwick filter for orientation estimation (Madgwick et al., 2011), and regression-based models for stride length estimation (Gu et al., 2019). Once a step is detected, PDR updates the position. As shown in Equation 4.2, the current 2D position $\mathbf{X}_i$ is computed based on the previous position $\mathbf{X}_{i-1}$, combined with the stride length T and the rotation matrix $\mathbf{R}_i$ build on the heading value.

$$\mathbf{X}_i = \mathbf{X}_{i-1} + \mathbf{R}_i \cdot T \tag{4.2}$$

PDRs are usually limited to the device placements, and works for instance very well in

foot mounted sensors, where Zero Velocity Update (ZUPT) are used in order to correct the sensor bias when the system stays stationary (Pinchin et al., 2012; Romanovas et al., 2013), as well as handheld positions such as those using smartphone sensors (Wang et al., 2020; Willemsen, 2016). Research works in step length estimation using regression models has been extended, and consideration of placement and accuracy limitation in PDR methods, established data-driven methods.

4.1.2 Data-driven methods

Data-driven methods have advanced significantly, employing learning techniques to enhance inertial tracking. The foundations of these approaches stem from early work in human activity recognition, such as (Zhao et al., 2017), where a time-window of IMU sensor values is used to classify activities like walking, running, and more. The concept of using time-windowed data, rather than relying on step-detection events, has also proven advantageous in approaches (Xiao et al., 2014).

Building on these ideas, the collection of the first inertial localization datasets enabled the development of early neural network-based approaches to localization (Chen, Lu, et al., 2018; Chen, Zhao, et al., 2018). The inertial sensor values, and accurate reference positioning values should be collected, in which a time window of the inertial sensor value in three axis each (say 1 seconds) considered as the input for the supervised learning, and displacement happened during the time window (velocity vector), tracked by the reference system, considered as the label data. In follow-up work, the Robust IMU Double Integration (RIDI) method (Yan et al., 2017) demonstrated how velocity vectors could be regressed by incorporating insights from activity recognition. This approach employed simpler machine learning models, such as Support Vector Machines (SVM), to reduce computational complexity (F. Zhang & O'Donnell, 2020). Notably, RIDI was the first open-source implementation in this area, helping to catalyze broader community engagement.

Subsequent developments led to more advanced architectures. Robust Neural Inertial Localization (RoNIN) (Herath et al., 2020) and Inertial Deep Orientation-Estimation and Localization (IDOL) further improved localization accuracy (Sun et al., 2021). Additionally, Neural Inertial Localization (NILoc) leverages Transformer models for map-matching, aligning trajectory points to road networks for accurate trajectory reconstruction (Herath et al., 2022). The Comprehensive Transformer-based Inertial Navigation (CTIN) model combines ResNet (Zhao et al., 2017)-based encoders with Transformer decoders to capture both spatial and temporal information (Rao et al., 2022).

4.2 Infrastructure-based positioning

Infrastructure-based positioning methods, as seen in section 3.4, are distinguished by the signal measurements they use and therefore the positioning algorithms. Different positioning techniques are used in a wide range of beacon technologies, mostly focused on Received Signal Strength (RSS) measurements, including beacon-based positioning as Wireless Local Area Network (WLAN), Bluetooth Low Energy (BLE), Ultra-Wideband (UWB), Radio Frequency Identification (RFID), etc. (Zafari et al., 2019) as well as cellular technologies such as those in 1G up to Long Term Evolution (LTE) and 5G (Del Peral-Rosado et al., 2018). The suitability of these varies depending on the use-case requirements. While certain beacon technologies, such as WLAN, have the advantage of being already densely deployed in most buildings to provide Internet connectivity, they have the disadvantage of needing constant adjustment to changes in the indoor environment (S. He & Chan, 2016). In this way, absolute position estimates are mostly obtained from approximate measurements such as RSS, in so-called fingerprinting approaches, whereas cellular positioning methods use timing-based or angle-based measurements, making them less sensitive to changes in the indoor environment (Blankenbach et al., 2017).

The problem of estimating the position of a given fingerprint naturally can also (similar to data-driven IL) fit the definition of a supervised machine learning problem, where the data is represented by the fingerprints labeled with the capture position. By collecting labeled fingerprints in the so-called offline phase, a model can be trained in a supervised environment that is able to predict the positions of unseen fingerprints in the online phase (Laska, 2023). It also combines well with IL methods and map constraints (Ebner et al., 2015; Klingbeil & Wark, 2008). Fingerprinting can be applied not only at the scale of a building with in a zero-effort crowdsourcing way (Rai et al., 2012), but also globally by collecting the signal information through mobile mapping systems or smartphone applications, using Global Navigation Satellite System (GNSS) as label data. For example, the Android Application Programming Interface (API) named as Fused Location Provider[1]. It combines input from GNSS, Bluetooth and WLAN to produce good indoor and outdoor location indicators and provide sparse position data with an accuracy of about 20 meters (Herath et al., 2021).

Positioning methods based on time-based measurements, such as Time of Arrival (ToA), Time Difference of Arrival (TDoA), but also angle-based measurements such as ToA, depend on the hardware and software of the transmitter or/and User Equipment (UE), making them not commonly used as fingerprinting approaches (Liu et al., 2007). While the synchronization limitation on the UE side should be the main challenge in realizing

[1]https://developers.google.com/location-context/fused-location-provider (Accessed: 15 April 2025).

these approaches on smartphones, 5G positioning, especially those Location Management Function (LMF) based solutions such as Uplink TDoA, make accurate positioning possible also for this kind of portable devices. The research is trying to implement angle-based methods as a 5G positioning technique (Dwivedi et al., 2021) similar to what has already been tested in UWB (Blankenbach & Norrdine, 2011).

4.3 State estimation algorithms

In the last two sections, it was shown that relative trajectory estimation using inertial sensors, but sick to a growing drift over time, and absolute positioning from infrastructure, but energy consuming, can be done. This combination is reminiscent of the idea of an Inertial Navigation System (INS), which uses the fusion of inertial sensors to continuously estimate the position, orientation and velocity of a moving object (Britting, 1971; Jekeli, 2001). The difference, however, is that INSs usually use GNSS which can provide absolute positions with a continuous frequency in case of availability, otherwise the INSs can keep navigation accurate for a short period of time relying only on IMU sensors. However, the challenge in indoor environments is that the availability of positions is very sparse, and the estimation of position needs to be done in real time (Herath et al., 2021). Therefore, accurate sparse position data and map constraints are typically used, respectively through Recursive Bayesian Estimatio (RBE) algorithms such as Kalman Filtering (KF) and Particle Filter (PF). While both of these approaches are well-known, they struggle with sparse observations, an issue that can be considered their research gap, or simply a scenario for which they were not designed.

4.3.1 Kalman filtering

Kalman filtering is used to estimate a state from motion model and control inputs as well as measurement models, using so called prediction and measurement models respectively (Kalman, 1960). It is assumed that system noise follow a Gaussian distribution, and both models are linear with respect to the state. If the models are non-linear, and need to be linearized using Jacobi matrices, then the algorithm is called Extended Kalman Filter (EKF). Authors in (Solin et al., 2018) uses EKF in context of data-driven IL, in which the landmarks positions such as stairs have been translated to sparse positions, and are used in the filter, for correction of the relative trajectory delivered from PDR algorithm, although heading correction mechanism is not explained or considered. Therefore, the challenge for an indoor INS method, where sensor heading measurement is lacking, remains open.

4.3.2 Particle filtering

In contrast to the correction mechanism using sparse location data, filtering with map constraints is better known. PF belongs to the Monte Carlo Methods, used to model highly non-linear problems with very noisy measurements (Wendel, 2011). PF is able to handle highly non-linear problems, where the uncertainties are represented by several estimation samples, so-called particles (Aggarwal et al., 2010). The goal in the implementation for map matching is to approximate the probability distribution function (PDF) of a position by a large number of weighted, independent particles (Gordon et al., 1993). Backtracking PF further consider the history of the trajectory to further improve position estimates (Widyawan et al., 2008). PF have the advantage that non-linear processes and observation models can be implemented without linearization (Willemsen, 2016). In notable studies, map-matching algorithms based on PF such as in (Davidson et al., 2010; Woodman & Harle, 2008) has been tested in a real-time and sequential manner, in which inertial sensor values could be also used (Yu et al., 2017).

4.4 Level 5 Indoor Navigation Project

In the scientific context and state-of-the-art Indoor Positioning System (IPS) approaches and technologies, and the increasing demand for location-based services, L5IN project has been defined and launched by HafenCity Universit (HCU), Hamburg. The main idea was that there is no global system that works well enough to answer the indoor applications, while 5G network infrastructure was seen as no extra infrastructure or reserved hardware, since of promised availability to the public in the near future.

The aim of the project was to develop and provide indoor positioning and navigation services within areas covered by a 5G network. To investigate this, a 5G positioning capability was integrated into a 5G non-standalone (NSA) network deployed at HCU. The key points covered were the experimental modular development system at HCU, which supports the integration of heterogeneous technologies such as smartphones and their sensors, UWB reference positioning system, a smartphone application representing the HCU 3D building information model and the 5G network (see Figure 4.2). PDR approaches and PF-based map matching were implemented in the smartphone application, and the UWB transmitters, which were able to use the Uplink TDoA positioning method as reference system using a separate receiver tag. The systems were synchronized using a single Network Time Protocol (NTP) server located in the HCU, which received the time signal via the installed GNSS antenna. The project was designed so that different elements and methods could be easily improved and replaced, thanks to its modularity.

In addition, a cloud-based data management and analysis platform was established to facilitate data analysis and evaluation.

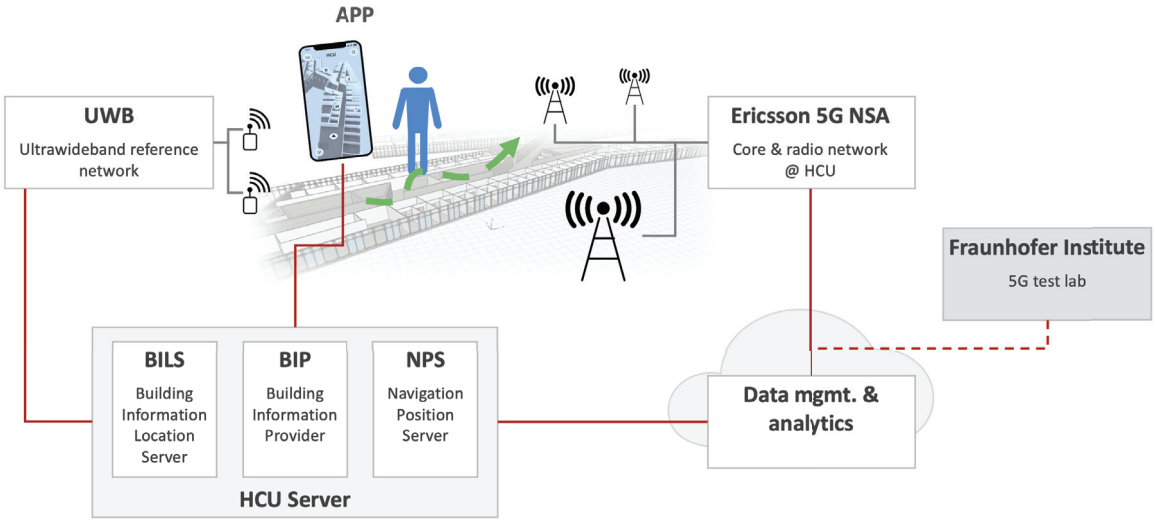


Figure 4.2: System overview of the L5IN project.

The project is working in partnership with the Fraunhofer Institut für Integrierte Schaltungen (IIS) to validate Release 16 positioning insights in a laboratory-scale experimental prototype, as Fraunhofer IIS operates a laboratory network for positioning solutions in a research environment (see Section 7.3).

5 Summaries of the Relevant Publications

This chapter briefly summarizes the publications that form the basis of this cumulative thesis, and illustrates the relationship between the publications.

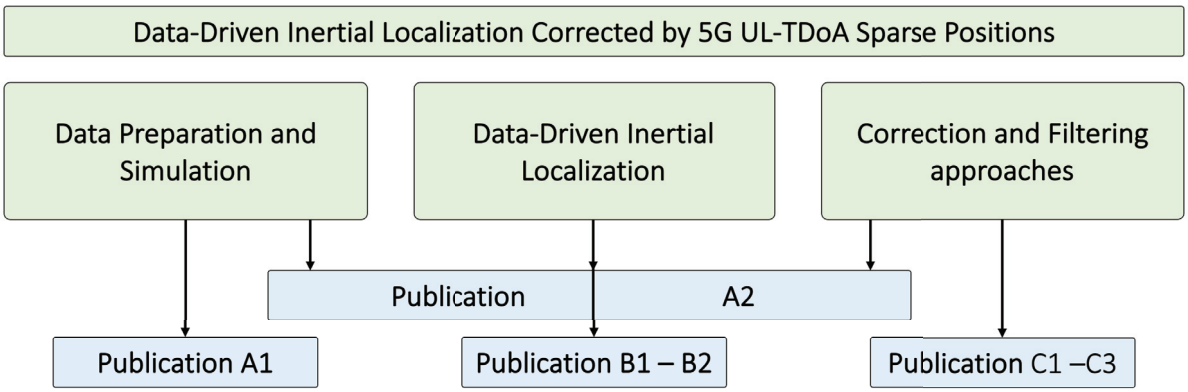


Figure 5.1: Content assignment of the relevant publications to clarify their respective contribution to the dissertation.

Connection between the publications

The primary objective of the publications is to develop a Global Navigation Satellite System (GNSS)-comparable solution for real-life Indoor Positioning System (IPS) suitable for smartphone-based pedestrian navigation use cases, considering the availability of 5G signal and positioning services. This objective is to be achieved by a data-driven Inertial Localization (IL) corrected with 5G sparse position data, from the 5G Uplink Time Difference of Arrival (TDoA) positioning method, for instance. In order to test the proposed solution, experiments and data collection are required, making a dataset. Then, once the data evaluated, two main approaches can be taken. Firstly, the IL can be made as suitable for real-life applications as possible, and as long-lasting and accurate as possible. Secondly, the necessity for correction should be minimised by automatic detection of the need for correction by sensor-based solutions, because this results in a more infrastructure-free solution, less energy consumption of the smartphone, and greater autonomy. In an ideal scenario, the sensor-based IL should be capable of triggering an alert when a correction is required. Overall, the present publications are divided into three main objectives, A, B, and C (see Section 2.2). Figure 5.1 classifies these publications according to their content and scientific context.

The connection of the publications can be viewed from a variety of perspectives, including the order in which the papers were published. The initial approach sought to identify a

comprehensive filtering method for integrating the measurements and solutions, concurrently with the endeavour to utilise unavailable real-world experimental data. Following the stabilisation of filter algorithms through the utilisation of simulated infrastructure data, a data preparation phase ensued, resulting in the formulation of final data-driven and filtering approaches.It is therefore recommended that readers peruse the publications in the order specified in this section.

Publication C3 (peer-reviewed)

- Shoushtari, H., Willemsen, T., Sternberg, H., 2021; Many Ways Lead to the GoalPossibilities of Autonomous and Infrastructure-Based Indoor Positioning, Electronics Special Issue Indoor Positioning Techniques.

This study provides a study of indoor positioning along with both autonomous and infrastructure-based methodologies. Two contrasting scenario studies, based on real experiments, were conducted to assess Pedestrian Dead Reckoning (PDR) and infrastructure-supported positioning, including the option of fifth-generation (5G) cellular networks. This work has been based on the early autonomous approaches, started to use Inertial Measurement Unit (IMU) embedded in smartphones for PDR. However, the findings revealed limitations in long-term reliability because of drift and error accumulation and user-device placement limitation of the last works has been addressed. A comprehensive literature review and a detailed story outline indicate that data-driven approaches have the potential to enhance the longevity of inertial localization, thereby introducing greater flexibility.

To counter these issues, it was proposed the use of 5G positioning based corrections. Two innovations were tested: Particle Filter (PF), in a form of a topological approach based on floor plans and routing graphs to improve positioning accuracy. The filter assigns weights to particles, hypothetical positions, based on how closely they align with the simulated 5G positioning data (reference points added noise), subsequently refining the estimate toward the most probable trajectory. Floor plans and routing graphs were also utilized within the particle filtering framework to further constrain the particles and improve accuracy by preventing improbable paths through obstacles or walls. Second, to use only a few of simulated 5G positions, to correct the step length, step heading and therfore the trajectory. The study also modeled 5G-based positioning scenarios and demonstrated better robustness in hybrid PDR with these Sparse position data. This hybrid approach observed significant accuracy enhancements as shown in early implementations surpassing only self-driven approaches.

Publication C2 (peer-reviewed)

- Shoushtari, H., Askar, C., Harder, D., Willemsen, T., Sternberg, H., 2021; 3D Indoor Localization using 5G-based Particle Filtering and CAD Plans, Indoor Positioning and Indoor Navigation (IPIN) , Lloret de Mar, Spain.

Building on the previous publication, the study proposes an innovative methodological framework incorporating simulated 5G locations, PF, smartphone sensor data and geospatial analyses. This research includes a critical innovation of an effective workflow converting CAD architectural plans into indoor maps and routing graphs, in the form of Geojson data, by application of basic geospatial analyses, such as, spatial joins, shortest paths and affine transformations. Providing maps and graphs for localization means providing environmental context for positioning algorithms and the application of geospatial analyses, makes the PF algorithm development flexible in terms of trying different scenarios for more development.

The study utilises particle filtering as a fundamental algorithm, wherein all particles originate from a shared initial point. The movements of these particles are calculated in three dimensions, incorporating covered length, heading information obtained from smartphone sensors through the implementation of the Madgwick filter, and the height change of each step. The application of the Madgwick filter represented a pioneering endeavour, enabling users to place the device with a high degree of flexibility. However, PDR shows its limitations in terms of needing steps events to update information, including height, which is not possible in an elevator, for instance. The application of Madgwick was the first try, that allowed a flexible device placement by the user. The integration of random noise accounting for real-world inaccuracy and particle position updates are performed iteratively for each particle, reducing any uncertainty associated with the control, association, and update phases of the filter. Afterward, different line of sight, rooms, walls, and routing graph weighting methods are applied to refine particle positions. As a result, particles remain promptly dispersed, and divergence is avoided by means of a resampling procedure.

Moreover, a 5G emulation tool was developed to simulate a realistic network environment, thereby generating realistic measurements such as signal transmission and reception times, as well as the angles and positions of the signal.This research has demonstrated the merits of correction using sparse location data, which significantly reduced drift from IMU-based or autonomous methods. The proposed methods were validated through experimental setups in a complex university building, resulting in significantly higher precision and accuracy of tracking when compared to the baseline, particularly when used in conjunction with other map-matching techniques.

Publication A2 (peer-reviewed)

- Shoushtari, H., Kassawat, F., Harder, D., Venzke, K., Müller-Lietzkow, J., Sternberg, H., 2022; L5IN+: From an Analytical Platform to Optimization of Deep Inertial Odometry, IPIN , Beijing, China.

Advancing investigations used in Level 5 Indoor Navigation (L5IN) project (see section 4), this research work presents L5IN+ analysis platform and web-based simulation tools developed to support data-driven IL (see 4.1 for more information on the difference in titles.) correction approaches for real-life indoor positioning applications. L5IN+ analysis platform include an online application called Positioning Lab (PosLab) and provides comprehensive evaluation tools and datasets, supporting the complete processfrom data collection and management to dataset creation and position estimation. This publication also includes initial explorations of data-driven methods. This included the use of a neural network architecture inspired by RoNIN-ResNet (Herath et al., 2020) to predict velocities from IMU sensor values. This network was specifically designed to cope with the drift issues due to noisy sensor readings and limited training data, by merging different quality of label data. Moreover, a correction method using Kalman Filtering (KF) was formulated and verified on simulated data to complement the absolute placements estimated from 5G measurements with the relative placements generated by the Residual Network (ResNet) model. Preliminary evaluation demonstrated promising improvements in positioning accuracy using this combined approach.

Publication A1 (peer-reviewed)

- Shoushtari, H., Harder, D., Kasparek, M., Müller-Lietzkow, J., Sternberg, H. 2023; Data-Driven Inertial Navigation assisted by UL-TDOA 5G Positioning, International technical Meeting (ITM), Institute of Navigation (ION), Long Beach, California.

This research work investigates the fusion of 5G Uplink TDoA positioning with data-driven IL operated by neural networks, focusing on dynamic positioning scenarios and settings, with a special emphasis on pedestrian navigation. With the proliferation of 5G capable smartphones, there are great opportunities to develop indoor navigation solutions, particularly by significantly enhancing positional accuracy under real-world conditions.

Extensive experiments were conducted within a test environment, covering diverse indoor scenarios that ranged from clear line-Of-Sight (LOS) conditions to highly obstructed Non-Line-Of-Sight (NLOS) environments. Absorber walls were strategically placed within the testbed to effectively simulate varying degrees of signal obstruction, representing the narrow corridors, which are challenging for 5G positioning coverage.

The experiments were tailored to enable some of the different 5G situations, such as a wallless open indoor space where there are no obstructions, and results were also provided with wall configurations that were specifically designed to represent hallways and less fictional indoor settings. A Qualisys optical system was used as reference system, tracking two additional targets, one attached to the top of the user's head to mimic normal human movement, and another on top of the smartphone device. In addition, fresh 5G raw signals were made available for positioning research for the first time in this study, which increased the real-world nature and novelty of the dataset.

Moreover, available data-driven IL banchmarks were taken into account, which used information from sensors on smartphones and advanced deep learning network architectures, including ResNet, Long Short-Term Memory (LSTM), and Temporal Convolutional Network (TCN). This was done to confirm the general agreement of the data gathered through the experiences. Among the different models used, ResNet shows more efficient learning speed and overall higher performance across various data types, indicating its potential for practical detection and cancellation of common sensor drifts found in IMU-based positioning. A key element of this work involved creating a large dataset called L5IN+, which contained a variety of data from smartphone sensor readings and Uplink TDoA 5G location data to positional estimates from Androids fused location provider. This dataset covered a wide range of pedestrian movements and multiple smartphone placements, greatly expanding the opportunity for training and validating deep learning models for inertial navigation.

Overall, the study highlights the merits and effectiveness of integrating data-driven inertial navigation methodologies with cutting-edge 5G positioning solutions, showcasing significant advances in navigation precision in intricate indoor settings. This comprehensive method represents a solid ground for the effective progress towards the goal of universally reliable indoor positioning systems.

Publication C1 (peer-reviewed)

- Shoushtari, H., Harder, D., Willemsen, T., Sternberg, H. 2023; Optimized Trajectory from Smartphone Sensors and 5G Uplink TDoA Using Cluster Particle Filters (Original title in german: Optimierte Trajektorie aus Smartphone-Sensoren und 5G Uplink TDoA mit Cluster Partikel Filter), Internationalen Ingenieur-vermessungskurs Zürich, Hrsg.: Andreas Wieser, Wichmann Verlag

This research work builds on the findings of previous publications by applying state-of-the-art filtering approaches to the L5IN+ dataset using real 5G-ULTDoA positions. The approaches include cluster-based PF, Weighted Least Square (WLS) and the previously introduced KF (in Publication A2). The central research question guiding this study

pertains to the efficacy of correcting a relative trajectory by sparse location data, an event that should not occur as frequently as few seconds. The PF method aims to mitigate heading drift by leveraging narrow corridors or/and topological information from the routing graph. Additionally, the KF and WLS methods are designed to rectify shifts. However, it is noteworthy that these methods have been evaluated individually in the study.

One key feature of the PF approaches is that they essentially split particles in disjoint spaces, two parallel corridors or two adjacent doors, for instance. The estimation of the position is achieved through the consideration of the preceding steps, a process referred to as the back-tracking solution (Harder et al., 2022). However, these weights will only be useful if the filter mergs to the appropriate space. But that doesnt always work out. Publication C1 recommends to define clusters, by applying mean-shift algorithm (Welbaum & Qiao, 2025). Clusters are incorporated, allowing weighting methods to be used on each particle along with identifying the winning cluster. Winning cluster may be detected with with respect to the distance of the clusters to the PDR trajectory.

In the KF, it is necessary to provide measurements at a pre-defined rate; although the filter has the capacity to manage measurement gaps using the movement model and other measurements, it is not possible to cover gaps of several seconds if only IMU readings and PDR are used as measurements and movement sources, which can be different using WLS approach. The WLS, in addition to the automatic detection of wining clusters approach, which can be used as a signal for the need of correction, has been considered as a replacement for the KF. The WLS has the advantage of being used as required, without the need for constant operation, and it has been demonstrated that it is equally capable as the KF in L5IN+ experiments. Specifically, the work shows that filter approaches can be employed to correct errors with the use of sparse location data and map representations.

Publication B2

- Shoushtari, H., Willemsen, T., Sternberg, H. 2023; Supervised Learning Regression for Sensor Calibration, DGON Inertial Sensors and Systems (ISS), Braunschweig, Germany.

This study explores the sensor calibration process in great detail, with a specific focus on supervised learning regression techniques for application to Micro Electro Mechanical System (MEMS) inertial sensors, specifically accelerometers. Inertial sensors are prone to systematic measurement errors, and therefore require appropriate calibration to improve accuracy. Conventional methods treat calibration as a pure mathematical optimization problem, commonly tackled using techniques such as Gauss-Markov adjustment. This study investigates sensor calibration as a supervised machine learning problem, as an

add-on process, employing regression methods to specifically circumvent the systematic sensor error that inherently complicates the calibration problem in MEMS.

The research develops a comparative model consisting of two calibration workflows, which are a traditional Gauss-Markov adjustment method and a supervised learning regression method. The Gauss-Markov method obtains the calibration parameters, such as offset, scale and orthogonality from an iterative adjustment process. It is a functional relationship with gravitational acceleration as the reference, using the measurement data obtained from controlled static conditions. The supervised learning regression method refers to the fact that the optimization problem can be treated as a supervised learning problem, where gradient descent is used to minimize the same cost function defined in the previous workflow. In supervised learning, the data is divided into parts that are used for training (Seen dataset) and testing (Unseen dataset), providing an assessment that more closely resembles real-world deployment, where the model must deal with data that it has not yet seen.

The experimental phase employs a multi-MEMS configuration comprising fifty MPU-9150 sensors, thereby facilitating extensive validation across multiple sensor units. In this experiment, the number of iterations in the Gauss-Markov model cannot be accurately matched to the number of training epochs; however, it demonstrates the efficacy of supervise learning approaches that incorporate calibration parameters.

Publication B1 (peer-reviewed)

- Shoushtari, H., Kassawat, F., Sternberg, H. 2024; Context aware transformer network and in-situ IMU calibration for accurate positioning. IPIN, Hongkong, China.

This research provides an extensive approach to inertial localization, specifically through the use supervised learning and Transformer networks to jointly calibrate the sensor, as the proved concept in the publication B2) and provide accurate IMU-based IPS . IL methods tend to exhibit very large drift and tracking errors in complex movements and different device placements, especially using MEMS sensors in wearables devices such as smartphones and smartwatches. To overcome these challenges, this paper proposes a novel Transformer model with physics-based features, by which context-aware enhancements can be applied to sensor data interpretation and localization performance.

The proposed model, referred to as the Transformer Automatic Features Interaction (TAFI), integrates raw IMU sensor data with physics-based engineered features, such as mean values, standard deviations, step length, and features derived through Simpson integration and the Madgwick algorithm. These context-enriched features enhance the models capability to accurately predict velocity vectors and improve the IL performance.

The Transformer-based architecture, augmented by convolutional layers, efficiently captures temporal dependencies and spatial relationships within IMU data sequences, offering robust and precise predictions even under diverse device placement conditions.

The TAFI architecture merges established domain-specific knowledge in inertial positioning with contemporary deep learning techniques. Initially, the Spatial Encoder extracts spatial embeddings from raw IMU data using convolutional layers and batch normalization, followed by Transformer encoders enhanced with convolutional layers to capture complex local and global patterns. The Contextual Component contributes significantly by embedding physics-based engineered feature into the learning process. These features undergo convolutional processing, batch normalization, and linear transformations, subsequently integrating contextual insights via Transformer encoders, further refining the understanding of sensor behaviors and interactions. The Temporal Decoder then synthesizes spatial, contextual, and temporal representations by creating temporal embeddings and employing positional encoding. Utilizing Transformer decoders, this module effectively leverages masked self-attention and multi-head attention mechanisms to model temporal dependencies thoroughly. Finally, the Output Module consolidates these complex and multi-dimensional representations into precise velocity predictions through linear transformations, Rectified Linear Unit (ReLU) activations, and dropout techniques to enhance robustness and prevent overfitting. Experimental evaluations were conducted using three distinct datasets: Robust IMU Double Integration (RIDI) (Yan et al., 2017), Robust Neural Inertial Localization (RoNIN) (Herath et al., 2020), and L5IN+. The TAFI model's performance was benchmarked against state-of-the-art models, including Ronin-ResNet (Herath et al., 2020) and Comprehensive Transformer-based Inertial Navigation (CTIN) (Rao et al., 2022).

Overall, the time-ordered publications demonstrate not only the progression of research aimed at addressing the defined questions, but also provide insights into how research can be conducted alongside a project, in this case the L5IN project. While maintaining a degree of independence, the research still relied on the necessary infrastructure that had to be installed. Some time constraints may have led to opportunities, as the lack of infrastructure in the early days was reflected in the development of simulation tools. Reproducing and building upon existing machine learning models prior to dataset collection proved to be an efficient strategy in optimizing the overall process. Finally, based on the experience gained and the knowledge accumulated throughout the previous process, non-trivial ideas emerged, leading to the design and implementation of a novel model.

From a technical perspective, several considerations arise regarding the publications pre-

sented together as one body of work. First, in the earlier publications, the capability to compute headings for multiple device placements was limited to the use of data-driven approaches such as IL. Later, however, it was shown that quaternion-based heading estimation can also support multiple placements, albeit with significantly lower accuracy insufficient for practical, real-world applications.

Second, the filtering approaches, including various forms of KF and PF, were discussed individually in the respective studies. In practice, however, these filters should operate simultaneously and in coordination. Achieving this requires additional engineering effort and adaptive adjustments to the methods, an aspect not fully addressed in the aforementioned publications.

Finally, the impact of the novelly introduced model should be regarded as a research contribution of this dissertation. In practical terms, while there is a clear performance improvement of data-driven methods compared to traditional IL techniques such as PDR, the choice between the proposed model and existing state-of-the-art methods should ultimately depend on the specific application requirements and the intended balance between accuracy, robustness, and computational cost.

6 Research Findings

In the previous chapter, the relevant publications forming the foundation of this dissertation were briefly summarized in a publication-wise manner. This chapter presents the major findings and insights derived from those publications, organized topic-wise in relation to the research objectives outlined in Section 2.2. Within topic "**A**," which addresses data preparation and evaluation, Section 6.1 answers the question: "*How can research and investigation into 5G positioning accuracy be conducted during the planning phase or in the absence of a positioning-capable cellular network?*" Meanwhile, Section 6.2 focuses on the availability and suitability of data sources, addressing the question: "*What properties should a dataset and data sources have, to effectively support machine learning and correction models?*". Within topic "**B**", concerning data-driven il, Section 6.3 introduces the novel data-driven Inertial Localization (IL) model, known as Transformer Automatic Features Interaction (TAFI). It explores *"whether it is feasible to combine physics-based domain expertise with learning-based approaches for accurate IL"*. Section 6.4 addresses the research question of *"whether supervised learning methods, as of TAFI, inherently consider the calibration of the Inertial Measurement Unit (IMU) sensor?"*. Finally, topic "**C**," deals with the filtering approaches and correction methods. Herein, Section 6.4 summarize the findings, thereby facilitating a better understanding of hybrid approaches using smartphone sensors and 5G infrastructures. The central research question guiding this study is *"How can data from smartphone sensors, absolute positioning provided by 5G networks, and environmental map data be optimally integrated?"*

This chapter seeks to address the aforementioned research questions by referring to the proposed solutions presented in the publications. Each section aims to frame detailed research questions, discuss the methodology used to address them, and present the corresponding findings. However, detailed numerical results will be discussed in the next chapter.

6.1 A. Design of a position generator as a 5G position simulator

The first contribution of this thesis, expressed in Publication A2, C2 and C3, is towards a solution of the availability problem in cellular positioning domain, data availability of smartphone sensors is less problematic in comparison. This problem becomes especially pronounced when data must be obtained from specific infrastructures that are often limited in accessibility during the research stage, and can slow down or even halt research

processes. For example, 5G or general cellular positioning has long been awaited, yet even at the time of writing this dissertation, few practical options are available.

While researchers in the field, as it has been elaborated in section 4.3.1, used discrete reference points (so called waypoints) or Quick-Response (QR) codes to evaluate the trajectory made by the smartphone sensors and remove the drift. From this the question arose, if the waypoints could be generated automatically, simulating the accurate positions from 5G networks, that are promised by the telcom industry. While the 5G standardization details and releases were published, the implementation in the real network was postponed. The question then arose whether a tool could simulate also cellular measurements. Additionally, the quality of cellular positioning directly depends on the design and deployment of the mobile network itself. Given the high costs of installing 5G infrastructure, it is important to determine the required accuracy levels for 5G positioning to ensure it meets the needs of specific use cases. To achieve this, how can one effectively experiment with use cases under varying scenarios, specifically by testing different levels of expected errors or standard deviations in coordinate measurements?

The simulator

A publicly available web-based simulation[1] tool has been developed to facilityte such experimentation and encourage further research in the above described context (See Figure 6.1), as the primary part of the PosLab introduced in Publication A2.

Figure 6.1: Interface of the web-based simulation tool.

The workflow starts with the calculation of a ground truth trajectory, which can be

[1] https://github.com/Hossein-Shoushtari/poslab (Accessed: 15 April 2025).

provided from a user-uploaded ground truth trajectory or it can be directly created in the application using chosen waypoints and sensor data (See Section 6.5 for details). Once the ground truth is in place, the simulation computes cellular network measurements (e.g., times, angles, and positions for sending and receiving a signal) while introducing custom systematic or/and white noise errors to emulate real-world scenarios. Given parameters such as positioning rate, expected error, and optionally simulated systematic errors, the simulation generates realistic measurement data expected from a cellular network like 5G, by applying the defined noise values to the ground truth trajectories positions at the defined positioning rate. The simulated outputs may then be tuned based on the number of simultaneous users and antenna placements. Therefore, the network quality can be fully customized to simulate different possible scenarios.

The systematic errors are calculated based on simulation settings (Figure 6.2). Such errors add further noise to localized segments of the trajectory. The user defines the number of intervals (the subsections of the trajectory prone to the systematic errors) and the duration of the errors. For example, a corridor without network coverage, where users pass through for 10 seconds, could have different error characteristics compared to other areas. The number of users also affects the timing measurements of the simulation. In particular, delays are imposed on the measurement if the system usage exceeds the limit defined by the network capacity.

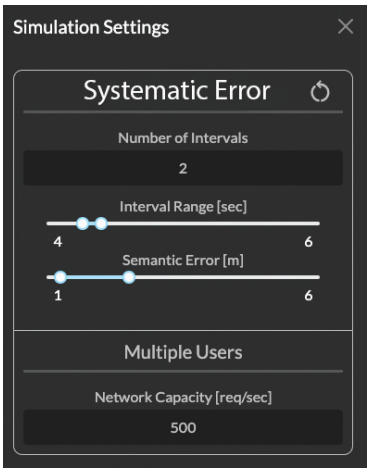


Figure 6.2: Simulation settings and the parameters.

The proposed simulation, supports loading map data in GeoJSON format, reference points, antenna positions, as well as Magnetic, Angular Rate and Gravity (MARG) data in CSV format. Users can create individual accounts and enhance their sensor datasets by integrating simulated cellular positioning measurements. Additionally, users have the capability to create map elements, such as points, lines, and polygons, during the simulation, which can then be exported. Ground truth trajectories, simulated measurements,

and user-generated drawings are available for download as a ZIP file. It should be noted that this simulation focuses mainly on positioning accuracy, with network details, such as number of connected users or bandwidth constraints being neglected at this stage. Nonetheless, this is only a temporary measure, a time delay is estimated, given a number of users and maximum amount of users for the network.

6.2 A. Data preparation and evaluation: Introducing the L5IN$^+$*Dataset*

Once data becomes available, the key considerations shift to the format and quality of the dataset itself, which are fundamental for any data-driven and correction approach. In learning-based methods, both the input data and its corresponding labels significantly affect the outcome. Advancing the methodology used in Level 5 Indoor Navigation (L5IN) project (see section 4), a contribution of this thesis is the introduction of the L5IN+ dataset in the A1 and A2 publications to provide different qualities and types of data needed to solve indoor positioning and 5G positioning challenges.

A key research question concerns the general characteristics of datasets required to build IL machine learning models. Effective datasets should include sensor measurements as input and displacement information for optimal labeling. To generalize and make the models reliable, these datasets should also account for variability in sensor placement and human behavior. The challenge lies in the determination of which behaviors should be considered to form a comprehensive dataset of real-world indoor pedestrian positioning applications.

Another important consideration is the required quality of the labeled data. These label data will not only be used to express the reference displacement in short time windows of a user movement, say 1 second, but also for evaluation purposes of the 5G positioning or correction model outputs. Timely and accessible correction positions will facilitate the mitigation of position drift, thereby extending the service life of Indoor Positioning System (IPS). In summary, the dataset is designed to support two main objectives, providing separate data sets for 1) machine learning and 2) correction models.

First, the L5IN+ dataset includes sensor data from a Huawei Mate 20 5G smartphone, serving as input for data-driven approaches. It consists of accelerations and angular velocities (from the 6-axis IMU sensor), magnetic field (from a 3-axis magnetometer), and more calculated values out of the sensors, as so called software-based sensors. The labeling data is obtained using the Qualisys optical motion capture system, which serves as a refer-

ence system and provides millimeter-level accuracy in human biomechanics applications, depending on the capture volume [2].

Second, the dataset contains 5G Uplink Time Difference of Arrival (TDoA) positioning data, including positioning measurements, results and raw channel impulse responses. It also contains position estimates from the Android fused location provider Application Programming Interface (API). This second set of measures is used to apply the correction methods. The reference data provided by the Qualisys optical system can still be used to evaluate this second set of measurements (for numerical results See Section 7.3).

Smartphone sensor data

The collection of the smartphone sensor data was done using an Android application developed in the L5IN project (see section 4). The application was enriched with a so-called researcher mode, which could collect the sensor data at the desired frequency, which was set to 200 Hz, the same as for other benchmarks (Chen, Zhao, et al., 2018; Herath et al., 2020). In the researcher mode, one could define a test route to be walked. Using the pre-defined route, researchers could also evaluate the quality of the navigation assistance, as well as the creation of waypoints with responsive timestamps (by clicking on the screen) in the absence of a reference positioning system. Figure 6.3 shows the researcher mode in the application.

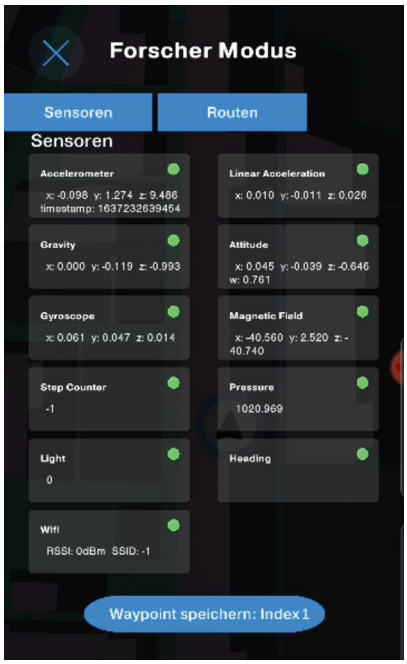


Figure 6.3: Interface of the researcher mode in L5IN application.

[2] https://www.qualisys.com/engineering/marine-vessels-and-structures (Accessed: 15 April 2025).

Qualisys data

The reference positions were measured using the Qualisys optical motion capture system, a high-precision 3D tracking solution based on camera-based detection of both passive and active markers. In the large L.I.N.K. hall, the system delivers sub-centimeter accuracy, and with optimized settings, it can achieve millimeter-level precision. The setup includes 28×Oqus 7+ series cameras arranged around the test area, providing comprehensive coverage. Additionally, two Miqus cameras are employed to capture video references. Both position and orientation data are recorded at a frequency of 100 Hz. Two active Qualisys tags are tracked: One tag attached to the smartphone, allows accurate tracking of the User Equipment (UE)'s movement and a second tag is attached to a cap worn by the subject to track the user regardless of the UE's placement. This configuration can not only help maintain a stable connection between that cameras and tags, but allows the machine learning models to be trained on both sets of labeling data in order to distinguish between user and device motion (See Figure 6.4).

Figure 6.4: Tags placement on user UE and a cap worn by the user.

5G Uplink TDoA positioning data

The data has been collected using a Huawei mate 20 5G, which supports the 5G New Radio (NR) band n77/n78. For the emulation of 5G Uplink TDoA in real time, the Fraunhofer Institut für Integrierte Schaltungen (IIS) FR1 Platform was used. This platform follows a Software Defined Radio (SDR) approach, enabling flexible and accurate signal generation and processing. The systems radio network consists of two main types of Transmission Reception Points (TRPs): Serving cells and positioning cells. Each cell

consists of a Universal Software Radio Peripheral (USRP) as RU and is connected to a server PC by a fiber optic 10 Gb/s network link for software-based signal processing. The serving cell is realized as a SDR 5G network and provides basic connectivity for the UEs. The neighbor cells operation is limited to the reception and processing of uplink signals that are transmitted by the UEs. Accordingly, the UEs are only connected to the serving cell, and no handovers to the neighbor cells occur.

Other considerations

Measurements and coordinates must share a common time reference and coordinate system. To ensure this, all input data was collected using a smartphone synchronized with a time server. The reference system was using another time server. To align reference values with the other measurements, time drift was calculated by identifying the best match and highest correlation between trajectories derived from 5G positions and the reference coordinates. Additionally, all coordinates were transformed into the the local coordinate system defined in the measurement hall, given the geo-referenced Global Navigation Satellite System (GNSS) points outside. Moreover, the dataset records different types of pedestrians and varying smartphone placements, such as handheld, pocket, backpack, and stroller. In one scenario the smartphone was mounted on a robotic crane, to mimic more non-step movements. This diversity can greatly augment training and validating inertial localization machine learning models. The retrieved movements also include step-based and non-step movements, with application as moving through an escalator, or strollers. These scenarios fall under the category of pedestrian movement and represent tasks that an IPS should be able to handle (for experience details See Section 7.3).

6.3 B. Development of a Transformer model for accurate inertial localization

This contribution of the dissertation primarily shapes publication B1, which is one of the main publications. Therefore, this section emphasizes the research questions and main findings to ensure comprehensive coverage and clarity of the content. To avoid repetition, further details can be found in the complete paper included in Appendix C.

During the time frame of this dissertation, data-driven approaches for accurate IL began to gain prominence. However, a challenge remained: To incorporate established physics-based domain knowledge rather than relying solely on "black-box" deep learning models. For instance, leveraging the periodic characteristics of acceleration values

allowed state-of-the-art physics-based methods to reliably detect step events. Similarly, integrating gyroscope data over short intervals provided accurate estimations of device orientation changes. Although using physics-based features, such as step count or orientation changes, may appear beneficial, the implementation details and how to effectively integrate this domain knowledge with the learning layers remain open questions. To integrate such established domain knowledge, the TAFI was developed. The proposed model TAFI in Publication B1, computes physics-based features and adds them as additional inputs to the model. Building on the original Transformer architecture (Vaswani et al., 2017), this network aims to estimate the velocity vectors by incorporating both the spatial relationships between the IMU sensor values and contextual information from features. The Transformer architecture combines these physics-based features with raw IMU data, avoiding redundancy by optimizing feature interactions. The raw sensor data and physics-based features are individually passed through numerous embedding layers and are then concatenated together.

Another advantage of TAFI is its ability to achieve high accuracy with a less complex architecture, i.e. few layers, compared with purely deep-learning Transformer models as (Herath et al., 2020; Rao et al., 2022). This illustrates how physics-based domain knowledge can be effectively structured to provide learning assistance. What is more, the lightweight nature of TAFI makes it appropriate for deployment on even resource-constrained devices including smartphones, though it has not yet been field tested in a smartphone setting. TAFI processes IMU input data in fixed time intervals of one second, and uses the label data from Qualisys optical system to predict and regress the velocity vector, i.e. displacements in x and y directions. Regardless of whether a persona makes a physical step or not, TAFI can estimate the displacements of pedestrian movements, even when no physical steps happened, such as using elevators or pushing a stroller on which the device is placed.

6.4 B. Validation of supervised learning localization models for IMU calibration

This contribution is primarily reflected in Publication B1, which constitutes one of the core components of this dissertation. Therefore, this section emphasizes the research questions and main findings to ensure comprehensive coverage and clarity of the content. To avoid repetition, further details can be found in the complete paper included in Appendix D.

Sensor Calibration is a standard task in measurement technology and allows the reduction of systematic measurement components in the measurement result. In the field of local-

ization and navigation, estimating the calibration parameters of IMU sensors, assuming they remain constant over their operational lifetime, is a common and valuable procedure, typically performed under the Gauss-Markov model. In contrast, as presented in Section 3.1, the calibration of IMU MEMS sensors is generally discouraged in the literature due to the temporal variability of their parameters. This raises the question of whether alternative estimation approaches can be employed. Although accelerometer calibration can be carried out without the need for additional equipment, the standard calibration of gyroscopes typically involves external hardware. Consequently, it is unreasonable to expect smartphone users to perform such procedures for applications like indoor navigation. Since supervised localization methods also establish functional mappings between labeled data and the measurement model, can supervised regression and learning-based localization approaches, such as TAFI, serve as in-situ calibration procedures during the training phase?

The proposed solution, demonstrated through a proof-of-concept calculation, incorporates sensor calibration during the training phase of a regression model. In an example involving accelerometer calibration, where a functional relationship is derived from the constant value of gravity, the results show that the machine learning-based estimation is significantly more accurate than conventional calibration approaches.

Indeed, the training epochs and numerical optimization in supervised learning are comparable to the iterative procedures applied on standard sensor calibration methods with the goal of minimizing an error or cost function. Since regression models are typically trained to optimize predictive performance on labeled data, rather than to explicitly estimate changes in sensor inputs as traditional calibration does, the calibration parameters are optimized indirectly as part of the training process. While supervised learning methods are not explicitly designed to estimate calibration parameters, they effectively consider these parameters to deliver accurate predictions, particularly beneficial for inertial sensors in smartphones.

6.5 C. Correction of relative trajectories using sparse positional data

One contribution of this thesis, expressed in Publication C1, C2 and C3, is towards the correction of relative trajectories, with a focus on sparse positional data, which is defined and expressed in Section 4.2. In short, to achieve a GNSS-comparable solution for indoor environments, and considering the high cost of infrastructure installation and the limitation of UE battery, the localization solution should be as autonomous as possible,

with as little infrastructure use and dependency as possible. Inertial sensors can provide relative trajectory accuracy up to within a few meters for few minutes, when data-driven approaches are utilized. Although long-term accuracy is critical for real-world applications and the longevity of inertial localization, these approaches still require some form of initialization and correction.

In this dissertation, the available sources of correction, in this case map data, routing graphs, building plans, and sparse location data from 5G positioning, are identified and assessed for their optimal application. While map data is typically autonomously available, continuous coverage from 5G positioning is practically infeasible. With an optimal 5G network design, accurate network-based positioning can expected approximately every one to two minutes of walking. However, the question arises how map information can support the correction procedure? Can autonomous approaches determine the need for correction from the infrastructure? In other words, do correction algorithms possess mechanisms to detect when autonomous localization fails and corrections must be triggered? Finally, how can corrections be carried out efficiently using sparse location data, applying an automatic re-initialization of the user's pose? The proposed solution can be divided into four key steps.

Step 1: Initialization

The pedestrian indoor positioning problem is, at its core, a challenge of estimating the user's pose, and floor level within a building. Considering a requirement for an instantaneous response from an IPS, infrastructure-based positioning methods such as 5G positioning can provide solutions for 3D position estimation or, alternatively, for 2D position along with floor identification. A challenge arises in instance estimating the user's heading at the beginning of the navigation or positioning process. Although magnetometer sensors can be used for this purpose, the accuracy of such sensors embedded in smartphones is insufficient or at least unreliable for precise heading estimation. Another potential approach is to calculate the initial heading ϕ using the arctangent formula:

$$\phi = \tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \tag{6.1}$$

where (x_1, y_1) and (x_2, y_2) represent the coordinates of the first two detected positions or steps. However, this method assumes the availability of highly accurate sparse positioning data, with sub-meter precision, at any location where the user intends to start walking or navigating. In practical scenarios, such an assumption is rarely met. Therefore, heading estimation using inertial sensors and sparse location data remains an open problem.

Nevertheless, the initialization and subsequent trajectory estimation of the user can still be addressed effectively.

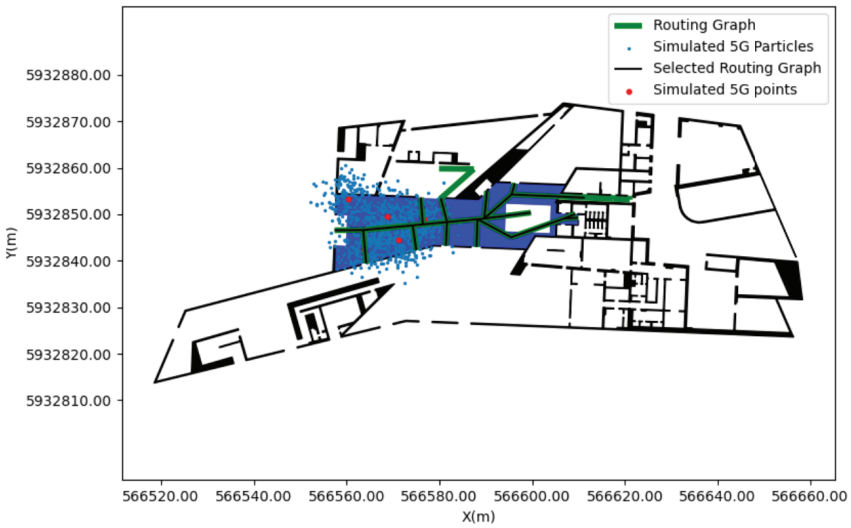


Figure 6.5: The initialization for PF using 5G sparse location data. The blue shape is the corridor or space that is determined as the one where the user located in (Publication C2).

As discussed in the literature review Section 4.3.2, Particle Filter (PF) is one of the filtering approaches that can play a key role in improving positioning accuracy through map matching techniques. During initialization, particles are generated and distributed according to the standard deviation of initialization position, which implies that there is no strict requirement for highly accurate sparse position data. Indeed, an initial position estimate with respective standard deviation of the position, and floor detection provided by the 5G network are sufficient. Based on this initial data, the spatial space (such as the room or corridor structure) and the routing graph can be identified (see Figure 6.5), in parallel with particle initialization. This essentially implies that the initialization process is completed and ready for further trajectory estimation.

Step 2: Correction with map data

The proposed particle filtering method for trajectory correction using map data comprises three main steps: propagation, weighting and resampling, as detailed in the publication C1 and C2. In short, particles are generated according to the initialization step, and then they are moved following the IL-provided movement model, considering a system noise. In the second step, weighting, the weight for particles is calculated due to its position concerning the environment or map constraints. Depending on where they ended up, particles crossing walls, in a physical sense, are weighted with zero, while particles closer to the routing graph get higher weights. In the resampling phase, particles with

zero weights are resampled uniformly at random with an initial weight of 1/N (where N denotes the total number of particles) per new particle. This equal weighting means that particles that are just introduced have no extra information for the first few steps. Resampling however consists of creating new particles around particles that have a high weight, as a wider interval has been formed around high weight particles and have a greater likelihood of reproducing new particles.

The two principal innovations introduced in this work are as follows. Firstly, the integration of geospatial analysis methods (e.g., calculation of centroids and computation of shortest distances) into the PF, which make the filtering process easier to develop and and enables the implementation of different scenarios. Second, it incorporates the mean-shift algorithm (Welbaum & Qiao, 2025) to allow the formation of a variable number of particle clusters. This allows for the calculation of separate weighted means in each cluster, especially useful when particles spatially separate. Mean-shift is an unsupervised machine learning clustering technique, technically a non-parametric algorithm that identifies clusters by iteratively shifting data points toward regions of higher data density. At its core, it uses Kernel Density Estimation, defined by:

$$\hat{f}(p) = \frac{1}{nh^d} \sum_{i=1}^{n} k\left(\frac{p - p_i}{h}\right) \tag{6.2}$$

where h represents the bandwidth parameter, n is the number of data points or particles, d is the dimensionality, and k is typically a Gaussian kernel function.

The clustering process involves repeatedly computing the mean shift vector, $m(p_i)$, for each particle p_i:

$$m(p_i) = \frac{\sum_{p_j \in N(p_i)} p_j - p_i}{|N(p_i)|} \tag{6.3}$$

Here, $N(p_i)$ is the neighborhood defined by the bandwidth h. Each point is shifted iteratively in the direction of $m(p_i)$, moving towards the local density maximum until convergence. Convergence occurs when the movements of the data points between iterations fall below a predefined threshold. At this stage, the data points have stabilized around the local maxima, or modes, of the density function. These modes become the final cluster centers, effectively grouping data points based on density. The selection of an appropriate bandwidth h is crucial, influencing the granularity and quality of clustering results. Methods such as Scott's rule, Silverman's rule, cross-validation, expert knowledge, and adaptive bandwidth approaches can help identify suitable bandwidth values to improve clustering outcomes (Chugh, 2024).

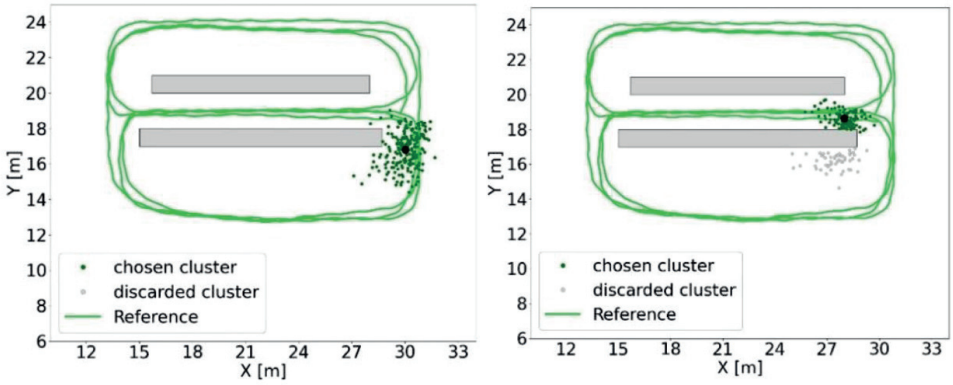


Figure 6.6: Example of particle clustering, where in left one cluster is, and in right two clusters are detected (Publication C1).

In the Figure 6.6 the clustering can be seen in two different steps, the clustering is useful to detect the correct side of the corridor, while the weighted mean calculation will only consider the correct cluster, otherwise it will estimate a position between the clusters, even with a low weight on the second particle cluster.

Step 3: Detection of Localization Failure

The identification of discrete clusters and their corresponding mean positions can offer valuable insights for detecting localization failures and determining when corrective mechanisms should be activated. Several conditions can be defined to facilitate this detection. For instance, the distance between the mean positions of clusters should not exceed a certain threshold, such as the width of a corridor, typically around 3 meters. Additionally, the standard deviation of the clusters is another critical parameter that must be carefully calibrated to effectively signal when corrections are necessary.

Furthermore, the calculated position obtained from IL , along with its cumulative accuracy, which should be maintained separately from the PF estimated position, can also indicate when infrastructure-based correction mechanisms need to be initiated. Specifically, if the IL accuracy is superior to the distance between the PF estimated position and the IL position, it clearly suggests that the trajectory has diverged and requires immediate correction.

Step 4: Correction with Sparse location Data

Once a correction event is triggered by the autonomous approach, the 5G network should respond by providing a position. The network will also verify whether it can adequately support this correction. Given that the 5G network is able to estimate a position and its

accuracy, if the accuracy of this estimated position is better than the distance between positions obtained from the IL or PF, the correction mechanism can be successfully applied. Otherwise, the autonomous localization method continues requesting corrections while simultaneously operating, resulting in progressively increasing errors in the user position estimation. This means, when accurate reference positions or waypoints (rather than potentially erroneous positions from the 5G network) are provided to the correction algorithm, the resulting trajectory can be considered as a ground-truth trajectory.

Details and how-to of the correction mechanism is shown in Figure 6.7. Initially, position #1 is known with high accuracy, marking the starting point of the trajectory. As the IMU-based trajectory progresses over time, it experiences drift, resulting in position and heading errors, leading to a drifted pose at position #2. The IMU-based trajectory provides not only heading information but also the entire sequence of trajectory positions P_d. Using sparse positional data, such as position #1 (the origin), the angular difference between the drifted pose and a known sparse measurement at position #2, denoted as diff $\phi_{d2,p2}$, can be computed, along with the associated error propagation values. Earlier positions along the drifted trajectory can also be evaluated for their respective deviations.

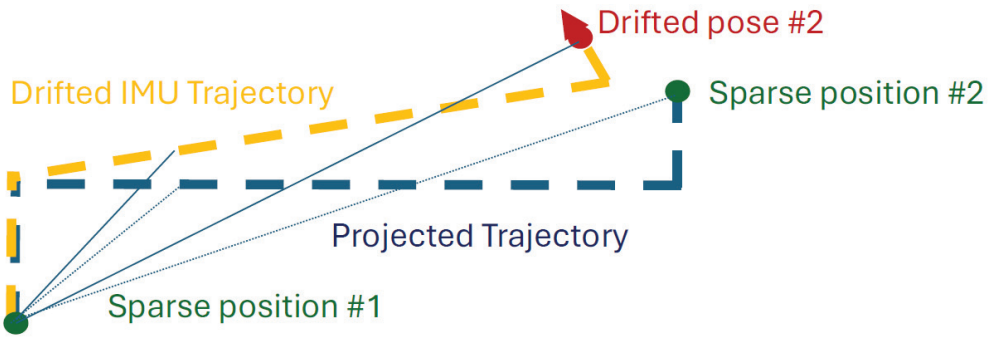


Figure 6.7: Symbolic representation of IMU drift correction using sparse position data.

By calculating the angular difference and the distance ratio between the origin and the drifted pose, the projected positions can be determined, as described in Equation (6.4). In essence, the method estimates the heading drift by comparing a sparse positional update with the drifted IMU-based trajectory. This difference is then propagated backward through all preceding trajectory points, incrementally refining the pose estimate. The final segment of the corrected trajectory yields an updated estimate of the heading and its associated quality.

$$P_t(\phi, L) = P_d \left(\text{diff } \phi_{d2,p2}, \frac{L_d \cdot L_{d1}}{L_2} \right) \tag{6.4}$$

where (ϕ, L) are the polar coordinates of the projected position P_t, and L_i represent the distances from the origin to various points along the drifted trajectory: L_{d1} is the distance to the drifted point 1, and L_2 is the distance to sparse position #2.

Overall, the investigations explored several potential approaches to address the lack of reliable GNSS measurements. These included the use of long-term motion estimation methods, the incorporation of additional information from map data, and the integration of infrastructure-based solutions as environmental sensors. The results demonstrate that combining data-driven IL, to maintain trajectory accuracy within acceptable error bounds for several minutes, with 5G uplink TDoA as an absolute positioning component, integrated through a multi-step correction mechanism, can result in drift-free user trajectory estimation. The next step involves numerically verifying these findings.

7 Evaluation

This chapter presents the numerical results and key findings of the proposed positioning approaches, covering both the 5G Uplink Time Difference of Arrival (TDoA) method and the data-driven Inertial Localization (IL) models, along with the correction filters applied. The main output of these algorithms are the estimated trajectories, and the central question is *how good these position estimates are?* To evaluate the quality of the positioning results, different assessment concepts will be discussed, including those aligned with the ISO/IEC 18305 standard. These concepts will help establish a framework for performance evaluation. The measurable metrics used to detail the main research question to specific ones will be introduced in the following section, followed by a detailed presentation of the results, including separate analysis on 5G Uplink TDoA positioning, data-driven IL and the correction methods through various plots and evaluation indicators.

In writing this chapter, particular care has been taken to avoid unnecessary repetition of content already published in related papers, in line with the principles of a cumulative dissertation. Nevertheless, additional details and results are provided here to ensure clarity and completeness, even if some overlap with earlier publications might appear necessary to maintain a coherent and comprehensive narrative.

7.1 Concepts and standardization

For an Indoor Positioning System (IPS) to come close to being comparable with and considered as a Indoor Global Navigation Satellite System (GNSS) it needs to provide the estimated positions with four key features: accuracy, smoothness, ubiquity, and immediacy (Angermann, 2019).

The most important is the property of accuracy, meaning correctness and precision, as they measures how close the estimated location is to the real location (ground truth). Several metrics are available to calculate the correctness of the models: Mean Absolute Erro (MAE), Root Mean Squared Error (RMSE), or Cumulative Distribution Function (CDF), while precision shows typically with standard deviation. But how accurate does an IPS need to be, particularly for guiding pedestrians indoors? This Thesis does not answer this question, as it needs statistical research, user studies, and collecting pedestrians feedback regarding acceptable error-margins according to the indoor navigation use case and the user experience. The other main property is smoothness. That is to say, how natural the estimated movement seems over time. A positioning system can exhibit

sudden jumps, erratic shifts or even imply that a person has walked right through a wall. These often arise as a result of limitations in sensors, availability of signals, or reliability of algorithms.

The third property is ubiquity. Mark Weiser and coworkers at Xerox PARC proposed the notion of ubiquitous computing in 1988, envisioning a future with technology seamlessly interwoven with the people's everyday lives (Weiser, 1999). The ubiquity of an indoor positioning system means that the system works anywhere and anytime for anyone, regardless of the matter which device is being used, its device placement (position) or environmental changes. Finally, the system must deliver instant position estimates and uses processing to provide accurate positioning without lag or latency. However, relative positioning methods, that rely heavily on user movement and history, do not provide instant positioning. To solve this problem, IPS can employ hybrid solutions that combine infrastructure-based systems with autonomous positioning approaches.

Guidelines on evaluation of IPS performance have emerged on the back of these principles, resulting in metrics definitions and the adoption of the International Standard: ISO/IEC 18305:2016(E)[1]. Any evaluation should provide a clear definition of the system under test, the entity being localized, and the context in which testing occurs, the standard states. In addition, experiments need to be repeatable, which means that the same results should be able to be obtained when the tests are performed several times in the same conditions. Tests must also be transferrable because the evaluation must not itself modify the test environment. Lastly, the standard necessitates the presence of a real-valued ground truth reference (at least one order and typically 3-5 times more accurate than the evaluated system) oriented to achieve meaningful evaluations.

Ground truth synchronization

It has been discussed in Section 6.2 that the ground truth trajectory can be measured using a Qualisys optical motion capture system or, as demonstrated in the Level 5 Indoor Navigation (L5IN) project, Ultra-Wideband (UWB) systems (see Section 4). Alternatively, the ground truth trajectory can be generated as described in Section 6.5, when the correction algorithm is feeded with reference points. However, the ground truth trajectory is more frequent and more accurate and does not necessarily share the same timestamps as the trajectory under evaluation. Therefore, for the purpose of evaluating a given trajectory, the reference trajectory is linearly interpolated to match the timestamps of the evaluated trajectory. This ensures that, for each timestamp of the estimated trajectory, whether estimated by the 5G network, IL, or filtering approaches, a corresponding ref-

[1] https://www.iso.org/standard/62090.html (Accessed: 15 April 2025).

erence position can be computed. Consequently, numerical evaluation metrics can be calculated in a one-to-one manner at each trajectory point.

7.2 Metrics

Building on the considerations discussed in the previous sections, this section introduces and details the selected evaluation metrics used, considering their relevance and importance to the pedestrian indoor positioning use case. Each metric is explained mathematically along with the meaning and performance information it provides. Let the ground truth coordinates at time step i be denoted by (x_i, y_i), and the corresponding estimated position by the localization algorithm be denoted by $(\hat{x}_i, \hat{y}_i)$. The following metrics are defined based on these values.

Mean Absolute Error (MAE) is used to measure the average positioning error across all time steps, providing an indication of the correctness of a trajectory. It treats all errors equally and expresses them in their real units. This is particularly useful when evaluating absolute positions, such as those determined by a 5G network. The MAE over n samples is defined as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |e_i| \tag{7.1}$$

where e_i is the Euclidean distance between the estimated and true positions at time step i. For the 2D case:

$$e_i = \sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2} \tag{7.2}$$

Absolute Trajectory Error (ATE) measures the correctness of a trajectory by calculating the RMSE across the entire evaluation trajectories. It indicates the position's correctness, when large errors are undesirable and should be emphasized. Absolute Trajectory Erro (ATE) focus is on the variance of the error, making it a useful accuracy indicator of positions delivered from IL, as they are later used in correction methods, where the variance can be incorporated into filtering processes.

$$\text{ATE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2} \tag{7.3}$$

Circular Error Probable (CEP) is a probabilistic key performance indicator, which describes the radius within which a given percentage of the positioning errors fall. It is presented via CDF graphs at different levels over all repetitions R, and defined as:

$$\text{CEP} = \min\{E : E \geq 0, |\{e_i : i = 1, 2, ..., N, \|e_i\| \leq E\}| \geq pN\} \tag{7.4}$$

where p is the desired probability level. The Circular Error Probable at 50% (CE50) and at 90% (CE90) are defined accordingly for $p = 0.50$ and $p = 0.90$. The metric can be used for both absolute and relative positioning methods, where as typical or average user experience are shown at 50%, and reliability guarantee shown at 90%, used generally in mission-and safety-critical, or regulatory environments, when outliers or large errors matter a lot. In general, CDF graphs and these indicators give an overview of the accuracy of a positioning system or algorithm.

Availability of the system is quantified by the loss rate of the expected positioning samples. The Loss Rate (LR) is defined as the mean of the individual loss rates over all repetitions R:

$$\text{LR} = \frac{1}{R}\sum_{r=1}^{R}\text{LR}_r \tag{7.5}$$

where the loss rate of a single repetition LR_r is given by:

$$\text{LR}_r = 1 - \frac{n_{\text{received},r}}{n_{\text{expected},r}} \tag{7.6}$$

Here, $n_{\text{received},r}$ denotes the number of samples received from the system under test during repetition r, and $n_{\text{expected},r}$ represents the expected number of samples, calculated as:

$$n_{\text{expected},r} = \Delta t_r \cdot f_s \tag{7.7}$$

where Δt_r is the duration of repetition r and f_s is the sample rate of the system. The availability metric is utilised to evaluate 5G networks and connectivity, as this is not a subject of uncertainty with IMU sensors.

Time Below Threshold and Distance Below Threshold are two additional performance indicators,introduced to assess the positioning error with respect to a predefined threshold x:

- **Time Below Threshold (TBx):** The fraction of total positioning time during which the positioning error is below x meters:

$$D_X(x) \equiv T(X \leq x) \tag{7.8}$$

- **Distance Below Threshold (DTx):** The fraction of the total traveled distance during which the positioning error is below x meters:

$$D_X(x) \equiv D(X \leq x) \tag{7.9}$$

These two indicators provide a practical interpretation of the system's reliability over fixed time and distance intervals. This metric is used to evaluate IL performance in terms of the longevity of the sensor-based methods and to account for drift over time, but this is not an issue of uncertainty in 5G positioning.

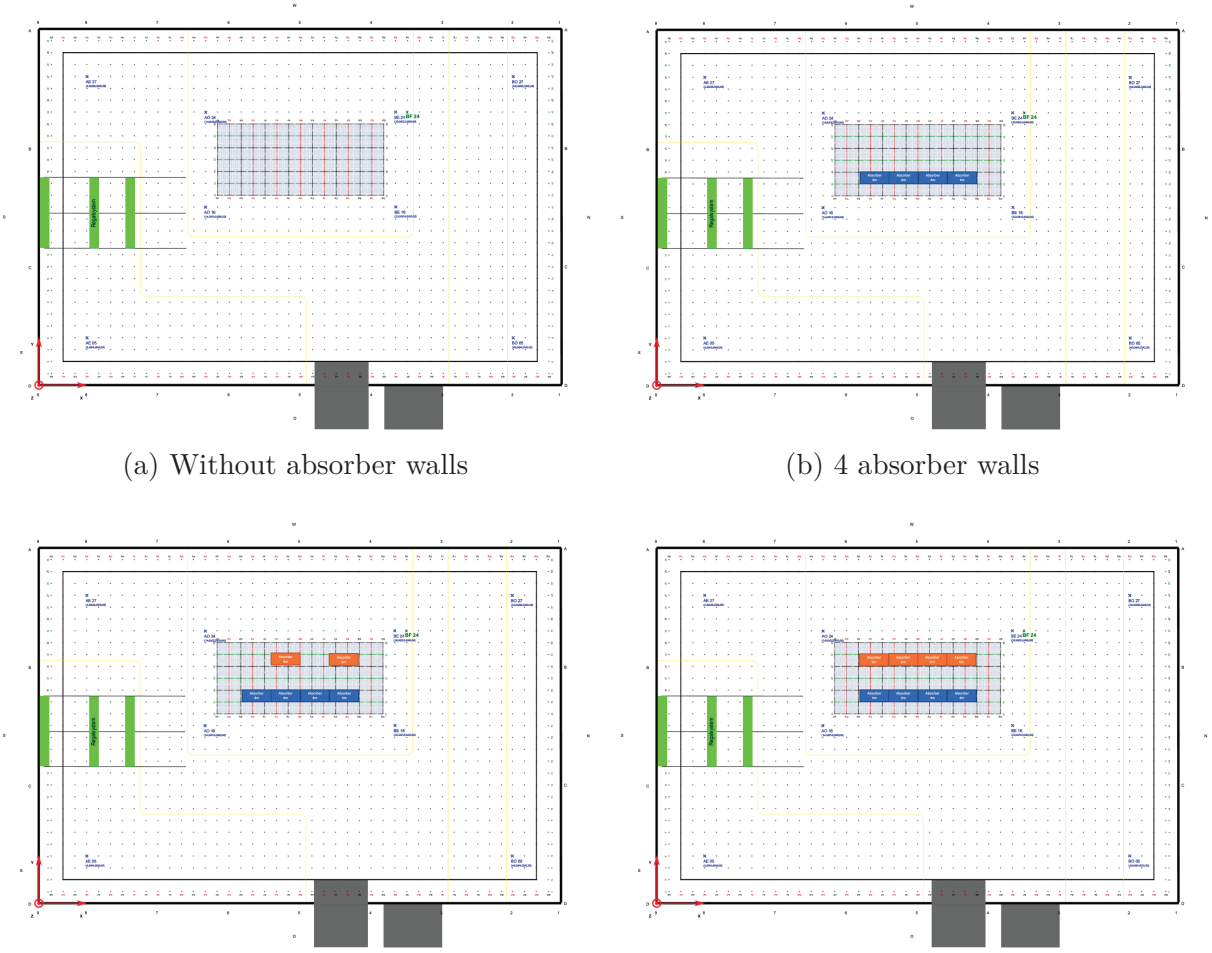
7.3 Experiments and results

The experiments were conducted at the L.I.N.K. test and application center at the Fraunhofer Institut für Integrierte Schaltungen (IIS) in Nuremberg, Germany, within a 30×44measurement hall, where smartphone sensor and 5G positioning data collection was performed using the Qualisys reference positioning system (for data sources, see Section 6.2). Section 7.3.1 focuses exclusively on the analysis of 5G positioning data. Section 7.3.2 presents experiments based on smartphone-derived datasets, while Section 7.3.3 discusses the outcomes of the correction and filtering procedures based on combined 5G and smartphone data.

7.3.1 5G UL-TDoA positioning

To assess the performance of 5G Uplink TDoA positioning, a variety of line-Of-Sight (LOS) and Non-Line-Of-Sight (NLOS) scenarios were tested. The desired measurement rate is 1 Hz, but this can only be expected in LOS. These scenarios were designed to reflect typical indoor environments; ranging from narrow corridors with limited or no signal coverage to open areas where signals from multiple antennas are clearly received. To

recreate these conditions, absorber walls were used to block signals from reaching the User Equipment (UE) in certain areas. Four different setup of absorber walls were considered, starting with no wall and increasing to eight absorber walls. For each setup, multiple sequences (or trajectories) were recorded, with different UE placements (further details are provided in Section 7.3.2) to capture a broad range of real-life conditions. In planning, the setups have been designed as shown in Figure 7.1, shown for better understand of the scenarios. Each trajectory was treated as a separate trial and repetition in evaluating the 5G positioning system. The numerical results of these experiments are presented below.



(a) Without absorber walls

(b) 4 absorber walls

(c) 6 absorber walls

(d) 8 absorber walls

Figure 7.1: Design layouts for corridor-like environments with varying absorber wall configurations: Orange walls represent 6 m high obstacles, while blue walls represent 4 m high ones.

The position of the antennas as well as the practical experiment with eight walls (see details below) can be seen in Figure 7.2. The red antenna is the primary transmitter, used as the reference for TDoA calculations and the blue antennas are the others. The antennas are installed at heights ranging from three to six metres.

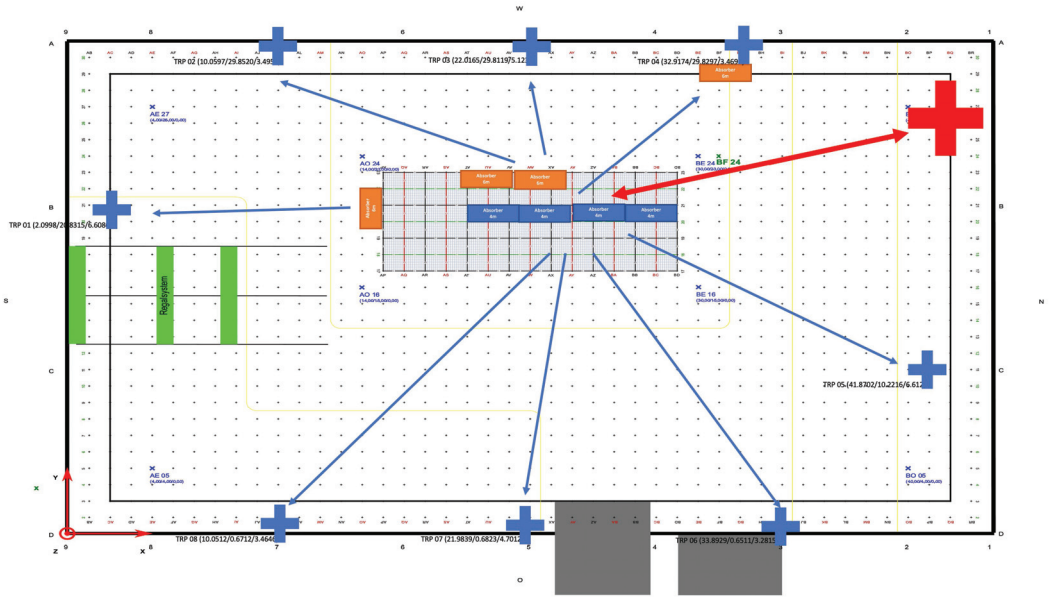


Figure 7.2: The position of the antennas

A Experimental setup with no absorber walls

In this experimental setup, 39 trajectories' data were collected, ranging from short trajectories of a few seconds to a 45 minute trajectory. Selected plots are showing in Figure 7.4. Metrics of correctness and precision (Table 7.1), named as MAE, standard deviation, minimum and maximum of errors were measured and the availability metrics, named as mean loss rate, its standard deviation and range for each repetition (Table 7.2) were also used for evaluation. Selected plots are showing in Figure 7.4. The grey area indicates where the reference system is unavailable, so evaluation is not possible.

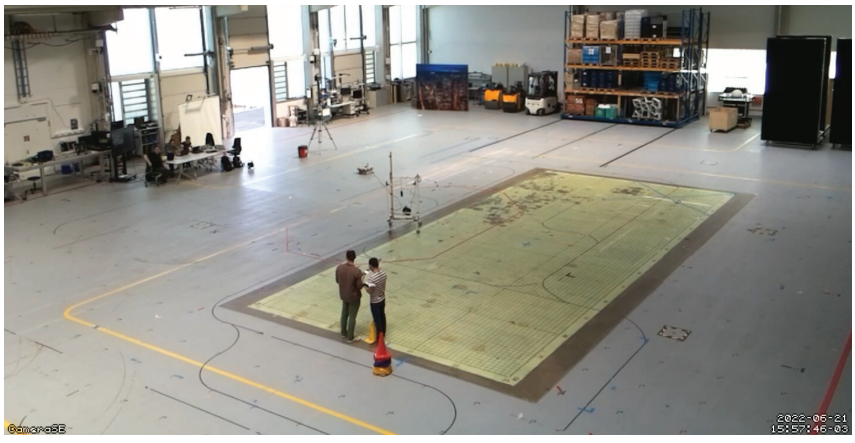
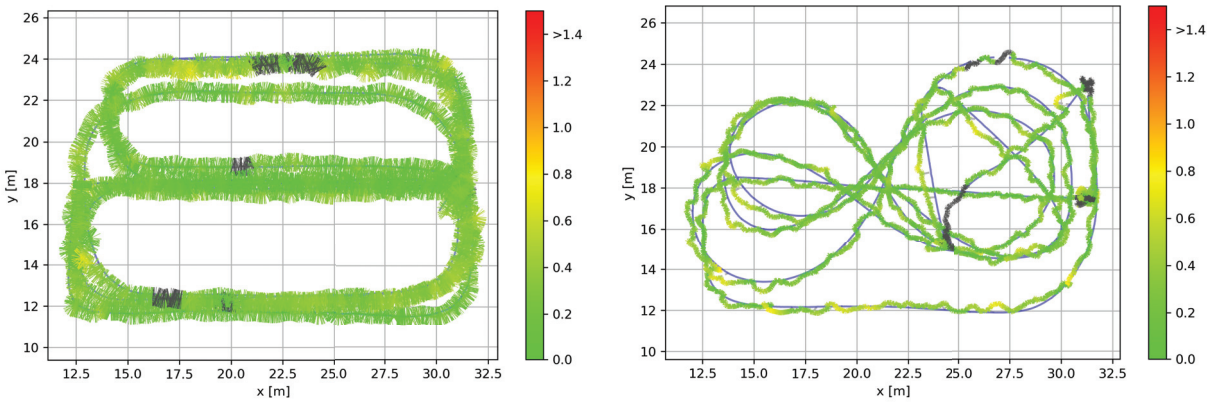


Figure 7.3: The measurement hall layout without absorber walls.

The analysis reveals a commendable overall performance, with an average MAE of 0.49 metres and a standard deviation of 0.62 metres when using all available 5G signals in LOS with a measurement rate of 1 Hz. Notably, several trajectories achieved sub-30 cm accuracy, underscoring the system's potential for high-precision applications.



(a) Trajectory 26 (b) Trajectory 37

Figure 7.4: Sample trajectories from experiments conducted without absorber walls, with error visualized using a color scale from green (low error) to red (high error). The grey area means no evaluation.

Certain trajectories exhibited significant maximum errors, reaching up to 33.97 meters, indicating occasional outliers that could appear in signal-based positioning methods. In a few cases, the reported standard deviation was lower than the MAE value, indicating that the precision does not necessarily reflect correctness. The consistency of this systematic error behavior can highlight the importance of addressing underlying biases in the positioning system to improve overall accuracy, but it is not the case here.

In terms of availability, the system maintained a high mean update rate of 99.80 Hz, with an average loss rate of 0.84%. While most trajectories demonstrated near-perfect availability, some experienced substantial packet loss, with one instance recording a 15.34% mean loss rate (see the availability metrics for trajectory 21). These findings in LOS conditions illustrates that, although the 5G positioning system generally offers reliable and accurate performance, there are instances of variability that warrant the combination with sensor-based positioning methods to ensure consistency across all operational scenarios.

Table 7.1: Correctness and precision in experiments with no absorber walls: MAE, Standard Deviation, and Min/Max Errors.

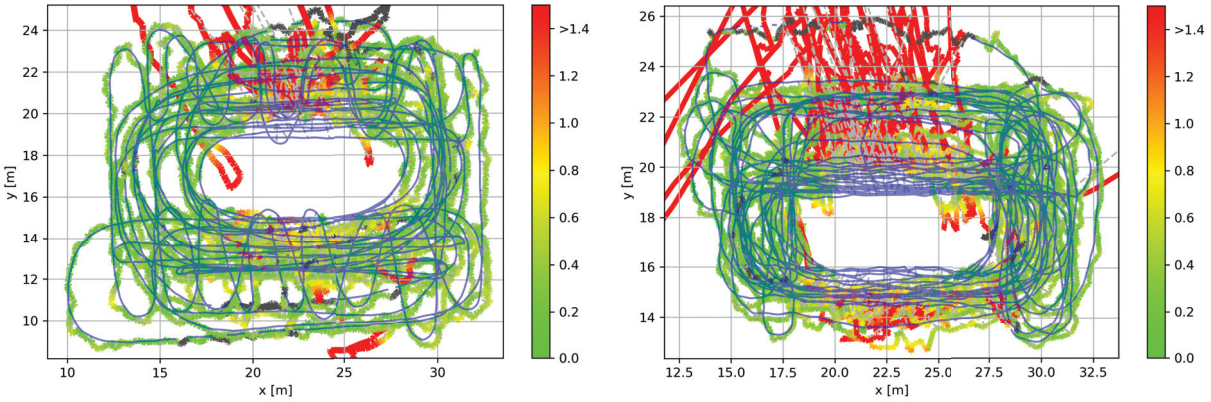
#	MAE [m]	Std. Dev. [m]	Min ... Max Error [m]
*A*1	0.74	0.36	0.01 ... 2.08
*A*2	0.70	0.62	0.01 ... 33.51
*A*3	0.66	0.32	0.01 ... 3.92
*A*4	0.72	0.29	0.01 ... 2.11
*A*5	0.69	0.26	0.03 ... 2.13
*A*6	0.84	0.99	0.02 ... 31.15
*A*7	0.98	1.02	0.02 ... 14.96
*A*8	0.68	0.32	0.01 ... 3.01
*A*9	1.42	0.90	0.02 ... 13.09
*A*10	0.71	0.40	0.01 ... 20.06
*A*11	0.70	0.43	0.01 ... 13.11
*A*12	0.85	0.76	0.01 ... 15.64
*A*13	0.51	0.38	0.00 ... 3.54
*A*14	0.87	1.40	0.01 ... 13.35
*A*15	0.53	0.18	0.01 ... 1.22
*A*16	0.48	0.28	0.01 ... 11.76
*A*17	0.42	0.17	0.01 ... 0.96
*A*18	0.47	0.17	0.02 ... 0.95
*A*19	0.46	0.17	0.02 ... 1.02
*A*20	0.41	0.16	0.00 ... 1.06
*A*21	0.38	0.71	0.00 ... 7.55
*A*22	0.30	0.90	0.00 ... 33.63
*A*23	0.22	0.39	0.00 ... 29.53
*A*24	0.27	0.39	0.00 ... 15.56
*A*25	0.23	0.68	0.00 ... 30.10
*A*26	0.23	0.12	0.00 ... 0.94
*A*27	0.25	0.38	0.00 ... 7.13
*A*28	0.21	0.11	0.00 ... 0.98
*A*29	0.25	0.40	0.00 ... 33.97
*A*30	0.24	0.17	0.00 ... 8.42
*A*31	0.21	0.10	0.00 ... 0.65
*A*32	0.16	0.30	0.00 ... 27.97
*A*33	0.27	0.16	0.00 ... 1.04
*A*34	0.25	0.28	0.00 ... 29.39
*A*35	0.22	0.34	0.00 ... 16.93
*A*36	0.23	0.38	0.00 ... 8.54
*A*37	0.24	0.15	0.00 ... 0.87
*A*38	0.25	0.24	0.00 ... 25.54
*A*39	0.21	0.13	0.00 ... 0.96
Ø	0.49	0.62	0.00 ... 33.97

Table 7.2: Availability in experiments with no absorber walls: Loss and Mean Rate.

#	Mean Loss Rate [%]	Std. Dev. [%]	Min ... Max [%]	Mean Rate [Hz]
*A*1	1.22	3.89	0.00 ... 99.99	99.73
*A*2	0.69	1.92	0.00 ... 99.99	99.85
*A*3	1.14	2.76	0.00 ... 99.99	99.75
*A*4	0.36	1.68	0.00 ... 99.99	99.94
*A*5	0.32	1.60	0.00 ... 99.99	99.94
*A*6	0.44	1.77	0.00 ... 99.99	99.91
*A*7	0.25	1.38	0.00 ... 99.99	99.94
*A*8	0.26	1.32	0.00 ... 99.99	99.95
*A*9	0.62	2.38	0.00 ... 99.99	99.87
*A*10	0.48	2.19	0.00 ... 99.99	99.91
*A*11	0.47	2.07	0.00 ... 99.99	99.91
*A*12	0.60	3.24	0.00 ... 99.99	99.94
*A*13	2.27	8.51	0.00 ... 99.99	99.78
*A*14	1.21	3.67	0.00 ... 99.99	99.79
*A*15	0.01	0.85	0.00 ... 99.99	99.99
*A*16	0.01	0.77	0.00 ... 99.99	99.99
*A*17	0.01	0.81	0.00 ... 99.99	99.99
*A*18	0.01	0.75	0.00 ... 99.99	99.99
*A*19	0.00	0.71	0.00 ... 99.99	100.00
*A*20	0.01	0.94	0.00 ... 99.99	99.99
*A*21	15.34	17.71	0.00 ... 73.82	96.11
*A*22	0.00	0.61	0.00 ... 81.01	99.99
*A*23	1.07	1.88	0.00 ... 79.75	99.69
*A*24	1.22	2.34	0.00 ... 86.03	99.64
*A*25	1.27	2.04	0.00 ... 77.50	99.57
*A*26	1.18	1.89	0.00 ... 72.53	99.58
*A*27	0.94	2.13	0.00 ... 77.93	99.76
*A*28	1.53	2.92	0.00 ... 93.25	99.57
*A*29	1.38	2.40	0.00 ... 69.10	99.59
*A*30	1.02	2.35	0.00 ... 80.29	99.71
*A*31	0.84	1.60	0.00 ... 90.62	99.71
*A*32	0.73	1.62	0.00 ... 81.46	99.75
*A*33	0.72	1.87	0.00 ... 76.99	99.81
*A*34	0.80	1.66	0.00 ... 83.25	99.75
*A*35	1.47	1.97	0.00 ... 77.33	99.48
*A*36	1.78	2.54	0.00 ... 66.59	99.38
*A*37	0.07	0.79	0.00 ... 82.05	99.98
*A*38	0.24	1.21	0.00 ... 77.73	99.95
*A*39	0.15	1.01	0.00 ... 82.66	99.97
Ø	0.84	3.74	0.00 ... 99.99	99.80

B Experimental setup with four absorber walls

In this experimental setup, four trajectories' data were collected, ranging from short trajectories of a few minutes to a 15 minute trajectory of walking. Selected plots are showing in Figure 7.5 and evaluation tables in terms of accuracy and availability are then in Table 7.3 and 7.4 respectively.



(a) Trajectory 1 (b) Trajectory 2

Figure 7.5: Sample trajectories in experiments with four absorber walls, with error visualized using a color scale from green (low error) to red (high error). The grey area means no evaluation.

The overall results and accuracy metrics, such as a MAE average of 0.42 meters with a standard deviation of 0.85 meters, are comparable to the LOS scenario. This is due to the frequent availability of at least three measurements to the transmitters required for 2D position estimation. However, short periods of signal-shadowing can still be observed in each trajectory (Figure 7.5), and is also evidenced by the low mean loss rate of 0.50%. This shows a typical signal disappearance in indoor areas, which can be bridged by other sensor-based methods.

Table 7.3: Correctness and precision in experiments with four absorber walls: MAE, Standard Deviation, and Min/Max Errors.

#	MAE [m]	Std. Dev. [m]	Min ... Max Error [m]
*B*1	0.40	0.77	0.00 ... 13.38
*B*2	0.56	1.21	0.00 ... 28.34
*B*3	0.29	0.35	0.00 ... 24.36
*B*4	0.38	0.57	0.00 ... 21.32
Ø	0.42	0.85	0.00 ... 28.34

Table 7.4: Availability in experiments with four absorber walls: Loss and Mean Rate.

#	Mean Loss Rate [%]	Std. Dev. [%]	Min ... Max [%]	Mean Rate [Hz]
$B1$	0.53	4.00	0.00 ... 84.24	99.93
$B2$	0.95	6.66	0.00 ... 71.35	99.94
$B3$	0.26	2.70	0.00 ... 77.63	99.96
$B4$	0.28	2.80	0.00 ... 66.89	99.97
Ø	0.50	4.34	0.00 ... 84.24	99.95

Γ Experimental setup with six absorber walls

In this experimental setup (See Figure 7.6), six trajectories' data were collected. Selected plots are showing in Figure 7.7 and evaluation tables in terms of accuracy and availability are then in Table 7.5 and 7.6 respectively.

Table 7.5: Correctness and precision in experiments with six absorber walls: MAE, Standard Deviation, and Min/Max Errors

#	MAE [m]	Std. Dev. [m]	Min ... Max Error [m]
Γ1	2.01	2.86	0.00 ... 20.42
Γ2	0.88	1.85	0.00 ... 23.88
Γ3	0.71	1.32	0.00 ... 12.48
Γ4	0.36	0.41	0.00 ... 4.60
Γ5	0.34	0.34	0.01 ... 4.00
Γ6	0.37	0.30	0.00 ... 3.39
Ø	1.08	2.10	0.00 ... 23.88

Table 7.6: Availability in experiments with six absorber walls: Loss and Mean Rate.

#	Mean Loss Rate [%]	Std. Dev. [%]	Min ... Max [%]	Mean Rate [Hz]
Γ1	3.86	15.12	0.00 ... 93.82	99.85
Γ2	1.81	10.69	0.00 ... 84.76	99.91
Γ3	0.25	2.52	0.00 ... 56.71	99.96
Γ4	0.22	1.30	0.00 ... 79.82	99.95
Γ5	6.99	12.21	0.00 ... 78.41	98.38
Γ6	18.27	16.55	0.00 ... 67.00	95.67
Ø	3.08	11.21	0.00 ... 93.82	99.51

The non-reliable signal behavior in NLOS conditions can be seen in Figure Figure 7.7. While in the example of Trajectory 1 (the worst performance) the position estimation was lost after a certain period, the shorter Trajectory 4 was able to maintain connectivity throughout. The error condition observed in Trajectory 4 (the best performance) could not

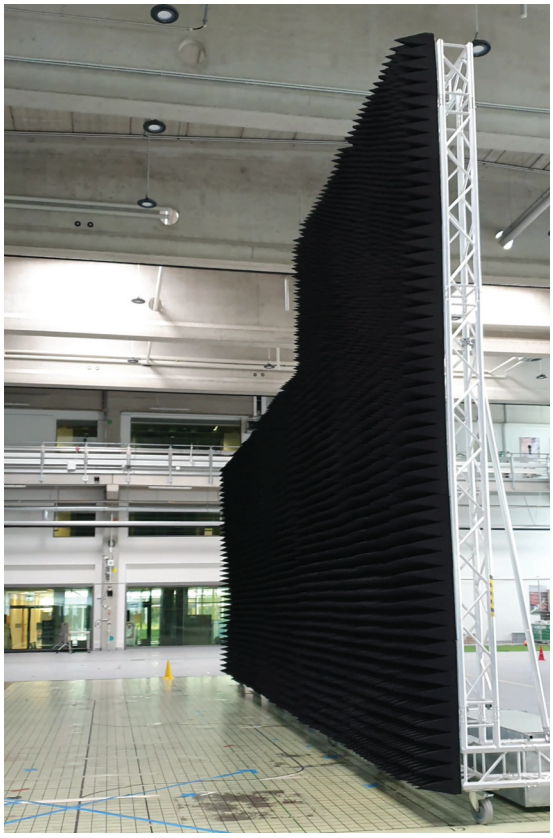
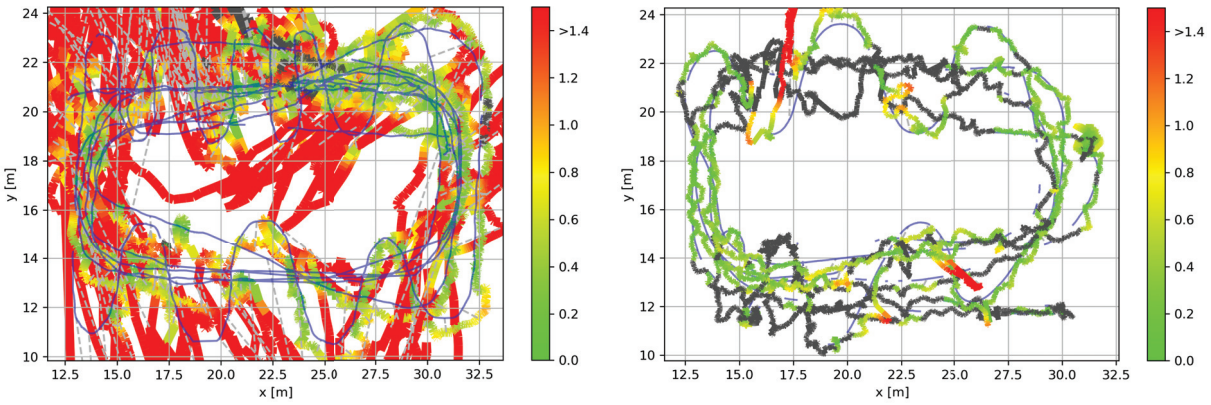


Figure 7.6: The measurement hall layout with six absorber walls.

be consistently reproduced, and therefore no definitive conclusion can be drawn, though the behavior remains noteworthy. As the number of absolute walls increases, accuracy and availability decrease, as expected, and this can also be seen in the metric values.



(a) Trajectory 1 (b) Trajectory 4

Figure 7.7: Sample trajectories in experiments with six absorber walls, with error visualized using a color scale from green (low error) to red (high error). The grey area means no evaluation.

Δ Experimental setup with eight absorber walls

In this experimental setup (See Figure 7.8), seven trajectories' data were collected. Selected plots are showing in Figure 7.9 and evaluation tables are in Table 7.7 and 7.8 respectively. In this scenario, there is no signal available within the simulated corridors, which are the most common indoor scenarios. Either in a corridor or in the very last part of the building, there are parts where there is no signal, but they are large enough for one to stay and navigate for a while. However, once the user enters an open area, 5G positioning becomes available again. It shows that the correction data source is only intermittently accessible.

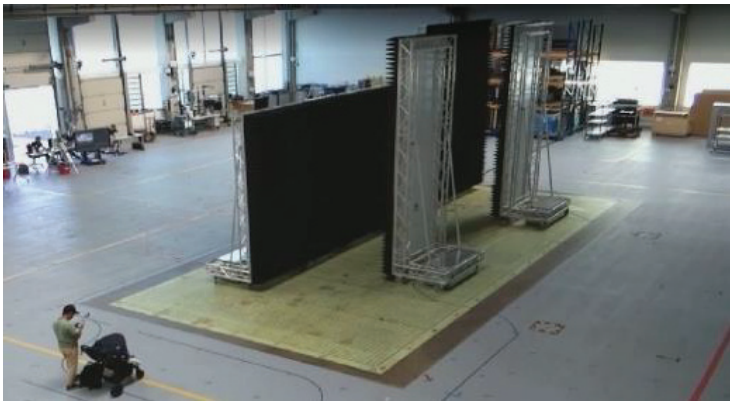
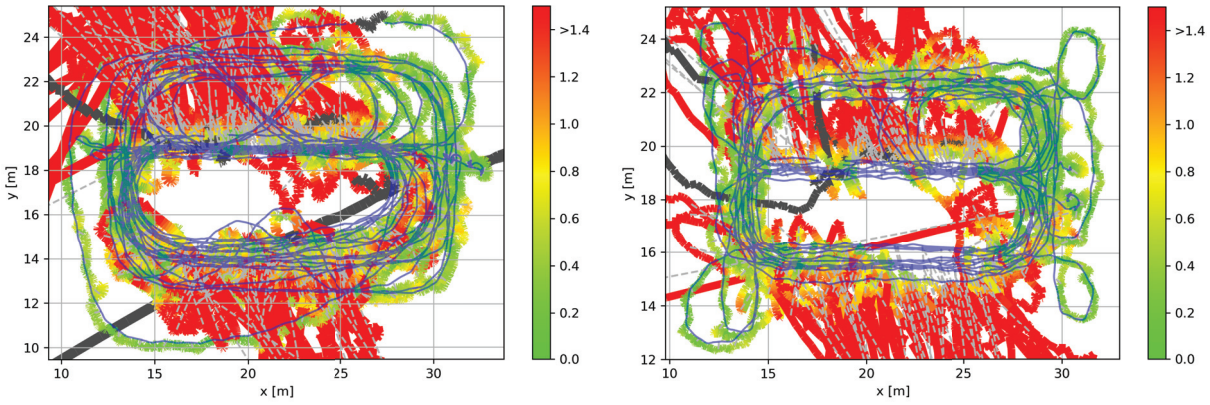


Figure 7.8: The measurement hall layout with six absorber walls.



(a) Trajectory 1 (b) Trajectory 2

Figure 7.9: Sample trajectories in experiments with eight absorber walls, with error visualized using a color scale from green (low error) to red (high error). The grey area means no evaluation.

Although the experimental evaluation reveals a moderate level of accuracy and high data availabilitywith an average MAE of 1.42 meters and an average loss rate of 2.67% across seven trajectoriesthe results should be interpreted with care. While the average loss

rate is lower than in the previous scenario, this improvement is largely attributed to a single trajectory (see Trajectory 6 in Table 7.5). It should also be noted that, due to the controlled laboratory conditions, the connection to the primary transmitter (serving as the reference for TDoA calculations) could not be lost, in order to prevent a system restart. Therefore, two-sided corridor configurations could not be tested.

Table 7.7: Correctness and precision in experiments with eight absorber walls: MAE, Standard Deviation, and Min/Max Errors.

#	MAE [m]	Std. Dev. [m]	Min ... Max Error [m]
Δ1	1.11	1.90	0.00 ... 26.92
Δ2	2.02	3.18	0.00 ... 29.93
Δ3	1.36	2.51	0.00 ... 30.17
Δ4	0.94	1.65	0.00 ... 25.22
Δ5	0.77	1.34	0.00 ... 27.42
Δ6	1.74	2.94	0.00 ... 32.70
Δ7	2.06	2.97	0.00 ... 29.46
Ø	1.42	2.48	0.00 ... 32.70

Table 7.8: Availability in experiments with eight absorber walls: Loss and Mean Rate.

#	Mean Loss Rate [%]	Std. Dev. [%]	Min ... Max [%]	Mean Rate [Hz]
Δ1	1.89	9.93	0.00 ... 84.48	99.88
Δ2	3.09	13.50	0.00 ... 85.42	99.90
Δ3	2.11	10.14	0.00 ... 83.21	99.89
Δ4	1.26	8.04	0.00 ... 82.76	99.93
Δ5	0.91	6.08	0.00 ... 81.13	99.92
Δ6	4.60	16.36	0.00 ... 92.97	99.79
Δ7	5.52	17.88	0.00 ... 92.63	99.78
Ø	2.67	12.26	0.00 ... 92.97	99.87

7.3.2 Data-driven inertial localization

The TAFI model, the proposed data-driven model for estimating user trajectories, using only inertial sensor values (see Section 6.3 for details), was assessed using various smartphone placements under a variety of experimental conditions. This involved a phone in a swing position, during a call, in the hand or in a pocket, inside a backpack or a bag, laid on a stroller, and carried by a robotic crane mimicking a moving walkway or escalator configuration (Figure 7.10). These diverse setups demonstrate the IL models ability to handle the variability of different positions or orientations of UEs, while maintaining reliable performance. While the dataset includes recordings from two users, where each tried to perform multiple walking patterns and speeds, expanding the dataset to include more users could lead to more robust and generalizable results. Whether the phone in hand gives rise to walking periodic acceleration signals or the phone is mounted on a stroller with no step-based motion, TAFI model should be able to estimate the displacement trajectory. The output, though, is only a relative trajectory, which drifts slightly over time. The objective is to maintain trajectory accuracy within an acceptable error margin... for as long a time as possible, and make the drift and need for correction minimized. The reliability is measured with two metrics: Time Below 1-Meter Threshold (TB1) and Distance Below 1-Meter Threshold (DT1). These metrics reflect the model's ability to maintain tracking accuracy, specifically how long and how far it can estimate motion without exceeding a 1-meter error in time and distance.

Details and configurations can be found in publication B1 (see Appendix C). The dataset is divided into seen datasamples from the L5IN+ dataset used during training or validationand unseen data, which the model encounters for the first time during position estimation. These subsets are randomly split in an 80/20 ratio, respectively.

Figure 7.11 shows one of the best performances of TAFI, on the Seen dataset. The model was able to predict the trajectory perfectly, with almost no drift over the entire trajectory for about 6 minutes. The evaluation was carried out across the entire dataset, as no significant difference between the model response in different placements has been recorded.

The model demonstrates strong performance, maintaining sub-meter accuracy for several minutes, which makes it well-suited for real-world applications. It also exhibits robustness across various device placements. Although TAFI incorporates the physics-based feature inside, combining the leaning approach and know-hows, but it does not perform optimally in all scenarios. Regular evaluation is required, and in some cases, corrective adjustments may be necessary.

(a) Stroller. (b) Handheld (c) Calling (d) Swinging

(e) Robotic crane (f) Backpack (g) Bag (h) Pocket

Figure 7.10: Visualizations for smartphone placements.

The performance on unseen data is remarkable, as it rather stays in 3 meter threshold for few minutes than 1 meter (See Figure 7.13). The detailed values are shown in Table 7.9 and 7.10 for seen and unseen datasets respectively.

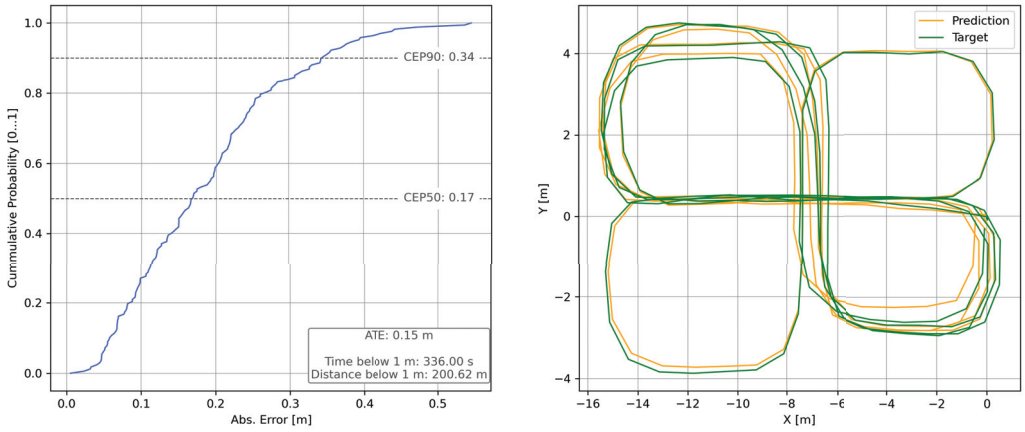
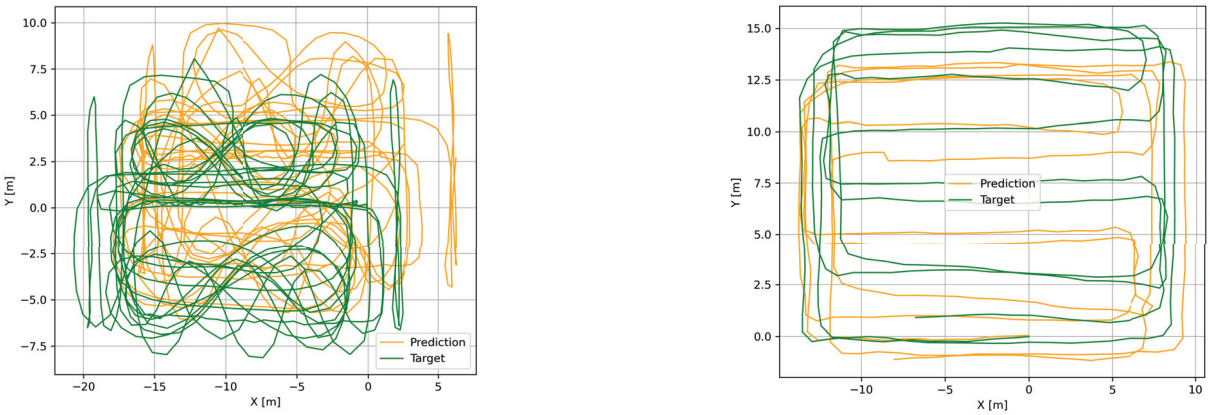
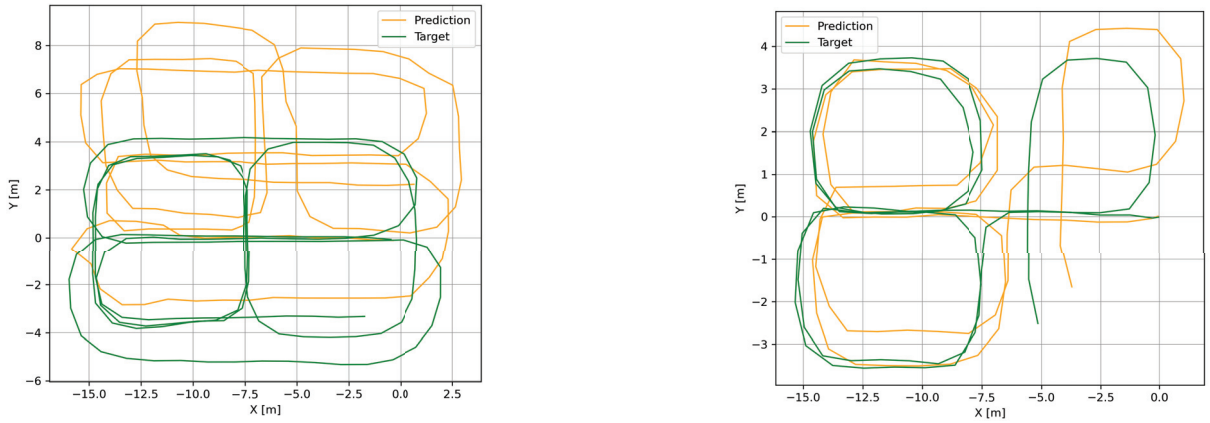


Figure 7.11: Trajectory 25 from seen dataset.



(a) Trajectory 6 (b) Trajectory 17

Figure 7.12: Sample trajectories from seen dataset.



(a) Trajectory 3 (b) Trajectory 4

Figure 7.13: Sample trajectories from unseen dataset.

Table 7.9: Accuracy and reliability on seen data in L5IN+ Dataset: ATE, Probabilistic Errors, and 1 m Thresholds.

#	ATE [m]	CEP50 [m]	CEP90 [m]	TB1 [s]	DB1 [m]
S1	0.54	0.56	1.11	642	372.68
S2	1.30	1.77	2.55	362	198.70
S3	0.77	1.09	1.58	170	71.48
S4	0.45	0.60	0.88	704	464.80
S5	0.51	0.64	1.20	246	162.67
S6	2.46	3.24	4.96	326	184.15
S7	0.36	0.43	0.77	704	417.60
S8	0.49	0.67	1.09	120	119.50
S9	0.34	0.48	0.71	336	187.94
S10	0.77	0.72	1.81	244	133.40
S11	0.40	0.54	0.81	480	297.19
S12	0.30	0.13	0.99	70	4.48
S13	0.38	0.46	0.76	304	152.35
S14	0.49	0.60	0.99	332	207.00
S15	1.60	1.95	3.33	28	9.93
S16	0.83	0.64	2.04	474	272.46
S17	1.39	1.62	2.82	108	28.61
S18	0.36	0.47	0.81	216	125.03
S19	0.82	1.11	1.58	84	30.25
S20	1.67	1.98	3.53	424	167.47
S21	0.23	0.29	0.50	272	159.16
S22	1.23	1.09	3.00	470	253.30
S23	0.92	1.09	1.92	104	43.52
S24–47	...	...	...	...	...
Ø	0.66	0.81	2.47	323.87	189.43

Table 7.10: Accuracy and reliability on unseen data in L5IN+ Dataset: ATE, Probabilistic Errors, and 1 m Thresholds.

#	ATE [m]	CEP50 [m]	CEP90 [m]	TB1 [s]	DB1 [m]
U1	0.77	1.09	1.85	170	71.48
U2	2.43	0.67	5.20	52	36.10
U3	2.56	0.63	5.19	56	36.69
U4	2.89	3.73	5.99	22	12.54
U5	0.69	0.80	1.51	66	39.47
U6	1.30	1.77	2.55	362	198.70
U7	0.54	0.56	1.11	642	372.00
U8–12	...	...	...	...	...
Ø	2.24	1.29	4.60	50.80	31.80

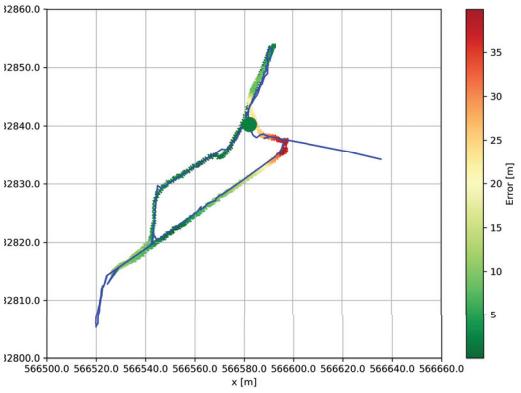
7.3.3 Correction and filtering

The experiments involving the four-step correction and filtering process, detailed in Section 6.5, have not yet been published. However, Publications C1, C2, and C3 collectively demonstrate the potential and effects of these methods, as discussed in Section 5. One of the main reasons for the lack of publication is the limitation of the test environment in L5IN+ dataset and the experiences in the measurement hall. Although absorber walls were used effectively to analyze 5G signal behavior, their limited number restricted the experimental setup's ability to evaluate map-matching techniques, such as long corridors connected to open areas. As a result, map-matching functions could not be thoroughly tested on the L5IN+ dataset. While sparse location correction could still be applied on its own, the full combined correction experiment was not feasible within this dataset.

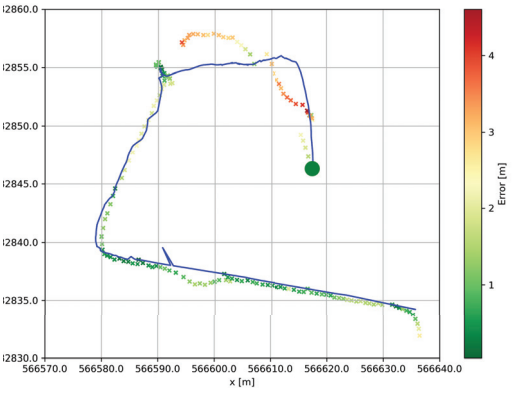
On the other hand, the L5IN project (see Section 4) setup provided a different opportunity. Although it lacked a millimeter-accurate reference system, it incorporated an UWB system using Uplink TDoA positioning method as a reference. With UWB positions were recorded at a frequency of 1 Hz and considered as reference trajectory, and used every 20 seconds in the the correction algorithm which was combined with a map-matching approach to shape the evaluation process. The data collection for these experiments was conducted using a dedicated smartphone application, by six members of the project team, independently and without direct involvement or supervision from the algorithms developer (the dissertation author). This setup was designed to provide an overview of the results and the practical effectiveness of the correction algorithm, helping to determine whether the proposed method could reliably achieve its intended objectives. Selected trajectories are illustrated in Figure 7.14, showcasing eight different correction events. The estimated trajectories are shown in blue. The distance between the reference trajectory, provided by the UWB, and the estimated trajectory, obtained from the filter, is defined as the positioning error, which is illustrated in the color bar.

In Trajectory 5, and even more clearly in Trajectory 16, it becomes evident how sparse position updates can effectively correct both the user's pose (position and heading). A similar observation can be made in Trajectory 17, where, despite an initial heading error of 180 degrees, the correction mechanism is able to realign the trajectory shortly afterward, once it has been called. However, certain challenges persist, particularly during initialization and navigation periods without correction signals. This is apparent in Trajectory 4, where the lack of correction data results in significant drift. In this case, the map-matching algorithm is unable to recover the trajectory once the user enters a signal-shadowed area. Trajectory 23 demonstrates how the correction mechanism can generate a smooth path by integrating IMU-based inertial localization with sparse position updates.

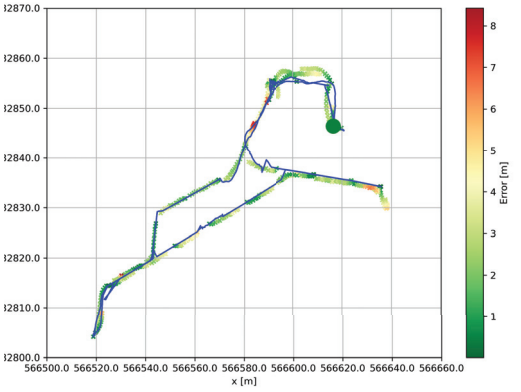
In contrast, relying solely on Uplink TDoA signals can cause unrealistic jumps in the trajectory, in addition to high energy consumption and limited coverage. The calculated metrics for all evaluated trajectories, including those mentioned above, are summarized in Table 7.11.



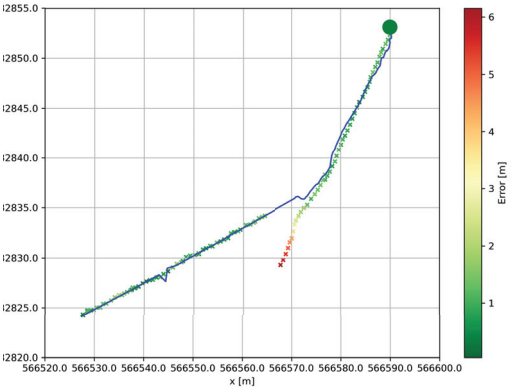
(a) Trajectory 4



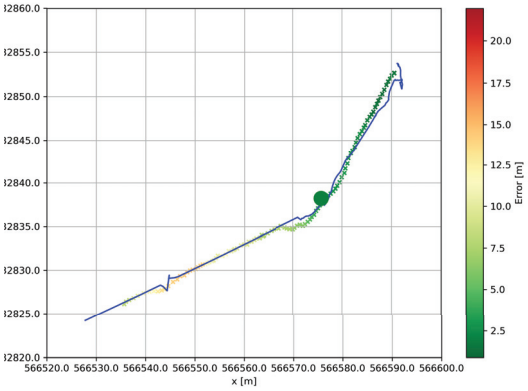
(b) Trajectory 5



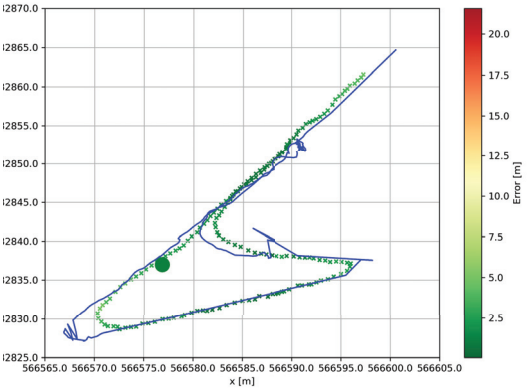
(c) Trajectory 9



(d) Trajectory 16



(e) Trajectory 17



(f) Trajectory 23

Figure 7.14: Selected visualization of corrected trajectories with green dot as start point, and error visualized using a color scale from green (low error) to red (high error). The estimated trajectories are shown in blue.

Overall, the results show that using sparse location data for correction can effectively support the trajectory, enabling seamless continuation of navigation. The achieved accuracyreflected in an erorr of less than 2 meters for 50% of the time, is adequate for typical pedestrian navigation scenarios. Considering the real-world use of smartphones and the semi-real-time calculation of corrections directly on the device, the proposed approach demonstrates both practicality and efficiency. As long as sparse location data remains available, the correction mechanism consistently performs as intended, underscoring the reliability and effectiveness of the method.

Table 7.11: Accuracy and precision of corrected random trajectories: MAE, Standard Deviation, and Min/Max Errors.

#	MAE [m]	Std. Dev. [m]	Min Error [m]	Max Error [m]	CEP50 [m]	CEP90 [m]
1	1.67	1.39	0.02	13.29	1.45	2.91
2	2.24	3.43	0.00	23.64	1.08	4.93
3	1.64	1.16	0.03	5.90	1.35	3.27
4	12.47	11.55	0.04	39.93	9.25	32.58
5	1.87	1.16	0.03	5.00	1.72	3.65
6	1.17	0.88	0.00	4.59	0.94	2.29
7	1.19	0.90	0.03	4.60	0.96	2.66
8	1.91	1.62	0.01	6.49	1.36	4.46
9	2.04	1.58	0.01	6.43	1.36	4.68
10	1.65	1.33	0.02	8.81	1.28	3.41
11	2.05	1.28	0.01	8.42	1.83	3.69
12	3.04	2.66	0.07	11.58	2.28	7.77
13	8.35	7.85	0.15	17.91	2.20	17.00
14	1.10	0.72	0.07	4.02	1.00	2.04
15	1.01	0.71	0.00	2.62	0.96	1.99
16	1.14	0.72	0.06	3.79	1.06	2.05
17	9.24	7.15	0.89	21.95	7.67	21.86
18	1.20	1.11	0.04	6.15	0.95	2.01
19	1.42	0.98	0.02	5.32	1.16	2.65
20	2.39	1.53	0.45	7.30	2.02	4.56
21	1.46	0.60	0.40	3.15	1.40	2.28
22	1.44	0.89	0.07	4.88	1.36	2.48
23	5.14	7.34	0.06	21.57	1.90	20.88
24	1.80	1.70	0.07	7.70	1.22	3.89
25	1.48	0.80	0.05	5.99	1.35	2.50
Ø	2.80	2.44	0.10	10.04	1.96	6.50

8 Conclusion and Outlook

Indoor Positioning System (IPS) often suffer from sensor drift over time, making a reliable indoor localization technology long-awaited. In recent years, hopes have turned toward Artificial Intelligence (AI) and the capabilities of 5G networks to deliver a Global Navigation Satellite System (GNSS)-comparable solution for indoor environments. Such advancements could enable a wide range of mobile location-based services, improve safety in industrial and public spaces, support augmented reality applications, and analysis based on the collection of location history.

A promising direction lies in the combination of data-driven methods, particularly recent breakthroughs using Transformer networks, and inertial sensors such as accelerometers and gyroscopes. These combinations have shown potential for accurately estimating relative positions over longer period of time. Simultaneously, 5G has emerged not only as a high-speed communication platform but also as a powerful enabler of real-time data transfer and potential positioning capabilities. Together, these technologies could mark a significant step forward in engineering geodesy and geoinformatics domains.

In one hand, existing data-driven models, along with their training datasets, offer a way to mitigate the sensor drift issue by leveraging deep neural networks. Considering the use case of pedestrian indoor navigation using smartphones, human motion, characterized by natural patterns measurable through physics-based models and inertial sensor data, can be effectively analyzed through the know-how obtained in the field and measurable movement based on the acceleration and rotation values. However, such insights have often been neglected in current data-driven solutions. This omission has sometimes led to unpredictable model behavior and reliance on deeply layered neural network architectures that make the learning process and predicted values difficult to interpret and explain.

On the other hand, the field of 5G positioning is still in a research phase and faces several challenges, such as dealing with multipath signals and the extraction of reliable measurements from raw signal data. While 5G positioning standards remain open, the actual implementation is largely in the hands of manufacturers, making the development process somewhat monopolized. As a result, progress has been limited, with only a few players controlling access to the technologies. For instance, positioning techniques like Uplink Time Difference of Arrival (TDoA), serving as a baseline for both 4G and 5G, have only recently become available from a single manufacturer. Other advanced 5G-specific positioning methods remain unimplemented and out of reach for many researchers.

Due to these limitations, collaboration with partner research institutes became neces-

sary. Before tackling technical research questions, organizational and logistical challenges had to be resolved. Through collaborative efforts, it has been possible to design real-world use cases, evaluate system performance in controlled environments, and publicly share everything from inputs to details and results. The investigations in this research primarily aim to move closer to a GNSS-comparable solution for indoor spaces by combining the most promising state-of-the-art researches and technologies. Initially, these approaches, data-driven IL, and 5G Uplink TDoA positioning evaluation, were studied independently. Later, they were integrated through filtering techniques to balance their respective strengths and compensate for individual shortcomings. The outcomes offer valuable insights and demonstrate practical potential for researchers, developers, and those interested in advancing 5G and AI technologies for human-centric applications, especially in pedestrian localization and navigation. From this broader investigation, aligned with the research objectives and key questions, three main aspects, explored through publications, can be highlighted:

- **Evaluation of 5G positioning in human centric applications**

  As part of the early national efforts to explore the capabilities of 5G networks across Germany, a public, web-based simulation tool was developed. This was followed by the collection, evaluation, and public release of 5G positioning datafrom raw signal measurements to estimated positions. The investigations revealed that 5G networks are indeed capable of providing useful signals in indoor environments. However, when focusing specifically on 5Gs positioning potential, several important lessons emerged that should guide research and development, and network design.

  One key takeaway is the importance of single-anchor positioning. This involves implementing and extracting Angle of Arrival (AoA) in combination with other ranging techniques such as Time of Arrival (ToA), and is crucial for both industrial and public applications of 5G positioning. Positioning should be considered an optional and value-added service, and must be able to compete financially with existing alternatives like Ultra-Wideband (UWB) beacon systems or even freely available Wireless Local Area Network (WLAN)-based positioning methods. This is why its particularly attractive to design private campus networks that not only support communication and real-time data transfer for a wide range of users, but also offer positioning as an additional feature, without requiring users to install multiple dedicated antennas solely for localization purposes.

  To be effective for pedestrian use cases, 5G positioning can just focus on delivering high-quality location data at key points within buildings, such as entrances and

floor transitions. Providing accurate position initialization in these areas is essential to enable seamless and reliable pedestrian navigation indoors. The rest can be covered with autonomous approaches and correction mechanisms. By delivering sparse yet strategically placed positioning updates, 5G can become a valuable enabler for the future of indoor navigation technologies.

- **Improving the data-driven inertial localization approaches**

  Building on recent advances in Transformer-based data-driven networks, this work contributes to the global effort to design and develop neural models that make inertial localization both long-lasting and suitable for near real-time applications on wearable devices such as smartphones and smartwatches. To be practical for these platforms, the models must be lightweight and efficient, ensuring compatibility with limited computational resources. This has led to the integration of physics-based features with raw sensor data in a compact, non-redundant format; an approach made possible by the capabilities of Transformer architectures; covering sensor calibration requirements. The proposed model, Transformer Automatic Features Interaction (TAFI), introduces this innovation by significantly reducing the complexity found in many state-of-the-art networks. While TAFI does not outperform existing benchmarks in terms of accuracy metrics, it offers a valuable trade-off between model size, interpretability, and performance.

  Evaluation of the model within a novel framework, built upon international standardization efforts, has demonstrated its potential. However, for broader adoption, the complexity of such models must be further reduced to allow seamless deployment across a wide range of smartphones. One of the key lessons from these experiments is the importance of model explainability. Understanding the causes of failure and being able to ensure consistent performance over extended periods are critical to making inertial localization reliable and robust for everyday use.

- **Improving the correction and filtering approaches**

  State estimation algorithms such as Kalman Filtering (KF) and Particle Filter (PF) have been widely used for decades in the field of mobile mapping and localization systems. These methods enable effective data fusion and integration of constraints such as map information to improve localization accuracy. In this research, however, the approach goes a step further by chaining multiple methods. It first ensuring reliable detection of when corrections are actually needed, and then applying them optimally using as few correction numbers as possible. These correction sources, often infrastructure-based, should inherently provide sparse location data and must be used strategically.

To the best of the author's knowledge, KF and PF remain among the most practical and compatible solutions for wearable devices. They have been thoroughly tested across a wide range of real-time experiments and have consistently demonstrated their effectiveness in live environments. Their practical reliability is well-established. A key lesson learned through this work is that more complex algorithms and tools should only be applied when necessary. The fact that a more powerful tool exists does not justify its use if a simpler solution already solves the problem efficiently. Compatibility and the intelligent chaining of well-established algorithms with state-of-the-art technologies should remain a central focus. While these classical algorithms may appear trivial in the current scientific context, their thoughtful integration with non-trivial ideas and modern techniques can lead to meaningful answers to previously unresolved research questions.

Although the effectiveness of the developed methods has already been demonstrated, the work on the GNSS-comparable IPS cannot be considered as complete, and several aspects are still to be resolved. Several challenges remain unresolved. This work primarily focuses on correction data derived from 5G infrastructure. However, as of now, 5G networks are not yet fully deployed as promised, and many of the positioning features are still not implemented in public networks. For such systems to be practical and scalable, access to accurate positioning data from publicly available infrastructure remains a critical requirement.

Regardless of these technological limitations, advancements in autonomous Inertial Localization (IL) using lightweight data-driven algorithms on smartphones and other wearable devices continue to open up new possibilities in future. Recent progress in explainable AI can contribute significantly in this space, especially in improving sensor calibration and enhancing the reliability of learning-based models. One promising direction is the 3D extension of data-driven IL; to explore whether altitude changes can be estimated accurately without relying on a barometric sensor. Initialization challenges, particularly with height and heading estimation using sparse location updates (typically within 2030 meters of accuracy), also remain as next challenges. next works may include exploring additional sensor integrations, such as cameras or Light Imaging, Detection and Ranging (LIDAR), in combination with 3D building models to address these limitations. The idea here is to develop a robust understanding of environmental materials using vision-based data acquisition methods. Such an understanding can also support 5G positioning and other signal-based localization techniques. 5G positioning offers great potential in terms of accuracy, where the implementation of methods such as AoA and Phase of Arrival (PoA) could be true game changers. The evaluations presented in this work will be further extended and shared with the research community to spark broader discussion

and collaboration.

Privacy and data protection are also important considerations. While pedestrian localization raises valid concerns regarding data security, major smartphone and wearable manufacturers, as well as social media platforms, are already collecting inertial sensor data in an industrial scale. These datasets can potentially be used to map environments or infer user location at a general level. This reality underlines the urgent need for clear data protection standards and secure frameworks that prevent unauthorized access to sensitive information. On the other hand, the societal benefits of IPS technologies cannot be overlooked. Applications tailored for specific user groups, such as indoor navigation for the visually impaired, localization support for firefighters operating in smoke-filled environments, or hospital systems designed to deliver timely assistance, highlight the value of developing targeted solutions rather than one-size-fits-all systems.

Overall, this thesis contributes to improving the efficiency of IMU-based localization systems, potentially benefiting any mobile platform equipped with such sensors. It demonstrates how cutting-edge AI and 5G technologies can be effectively integrated into the field of engineering geodesy and geoinformatics, helping to keep the field relevant, innovative, and aligned with modern technological developments.

List of Figures

List of Tables

List of Abbreviations

3GPP Third Generation Partnership Project

AHRS Attitude and Heading Reference System

AI Artificial Intelligence

AoA Angle of Arrival

API Application Programming Interface

ATE Absolute Trajectory Erro

BLE Bluetooth Low Energy

BMVI Bundesministerium für Verkehr und digitale Infrastruktur

CDF Cumulative Distribution Function

CEP Circular Error Probable

CNN Convolutional Neural Network

CTIN Comprehensive Transformer-based Inertial Navigation

CTIN Comprehensive Transformer-based Inertial Navigation

D2D Device-To-Device

ECEF Earth Centered, Earth-Fixed

EKF Extended Kalman Filter

eMBB enhanced Mobile BroadBand

GNSS Global Navigation Satellite System

GRS80 Geodetic Reference System 1980

GRU Gated Recurrent Unit

HCU HafenCity Universit

IDOL Inertial Deep Orientation-Estimation and Localization

IIS Institut für Integrierte Schaltungen

IL Inertial Localization

IMU Inertial Measurement Unit

INS Inertial Navigation System

ION Institute of Navigation

IoT Internet of Things

IPIN Indoor Positioning and Indoor Navigation

IPS Indoor Positioning System

ISA Inertial Sensor Assembly

ITM International technical Meeting

KF Kalman Filtering

L5IN Level 5 Indoor Navigation

LIDAR Light Imaging, Detection and Ranging

LMF Location Management Function

LOS line-Of-Sight

LR Loss Rate

LSTM Long Short-Term Memory

LTE Long Term Evolution

MAE Mean Absolute Erro

MARG Magnetic, Angular Rate and Gravity

MEMS Micro Electro Mechanical System

MIMO Multiple Input, Multiple Output

mMTC massive Machine-Type Communication

mmWav millimeter-Wave

NED North-East-Down

NILoc Neural Inertial Localization

NLOS Non-Line-Of-Sight

NR New Radio

NSA non-standalone

NTP Network Time Protocol

PDF probability distribution function

PDoA Phase Difference of Arrival

PDR Pedestrian Dead Reckoning

PF Particle Filter

PoA Phase of Arrival

PosLab Positioning Lab

QR Quick-Response

RAT Radio Access Technology

RBE Recursive Bayesian Estimatio

ReLU Rectified Linear Unit

ResNet Residual Network

RFID Radio Frequency Identification

RIDI Robust IMU Double Integration

RMSE Root Mean Squared Error

RNN Recurrent Neural Network

RoNIN Robust Neural Inertial Localization

RSS Received Signal Strength

RTT Round-Trip Time

SDR Software Defined Radio

TAFI Transformer Automatic Features Interaction

TCN Temporal Convolutional Network

TDoA Time Difference of Arrival

ToA Time of Arrival

TRPs Transmission Reception Points

UE User Equipment

URLLC Ultra-Reliable Low-Latency Communication

USRP Universal Software Radio Peripheral

UWB Ultra-Wideband

V2X Vehicle-to-everything

WGS84 World Geodetic System 1984

WLAN Wireless Local Area Network

WLS Weighted Least Square

ZUPT Zero Velocity Update

Bibliography

3GPP. (2024). *Ng radio access network (ng-ran); stage 2 functional specification of user equipment (ue) positioning in ng-ran* [Accessed on January 11, 2025]. Retrieved December 19, 2024, from https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3310

Aggarwal, P., Syed, Z., Niu, X., & El-Sheimy, N. (2008). A Standard Testing and Calibration Procedure for Low Cost MEMS Inertial Sensors and Units. *The Journal of Navigation*, *61*(2), 323–336. https://doi.org/10.1017/S0373463307004560

Aggarwal, P., Syed, Z., El-sheimy, N., & Noureldin, A. (2010). *MEMS-Based Integrated Navigation*. Norwood : Artech House.

Angermann, M. (2019). Seamless and Smooth Location Everywhere with the new FusedLocationProvider (Google I/O'19) - YouTube. Retrieved March 1, 2025, from https://www.youtube.com/watch?v=MEjFW_tLrFQ

Bahdanau, D., Cho, K. H., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1409.0473v7

Blankenbach, J., & Norrdine, A. (2011). Building information systems based on precise indoor positioning. *Journal of Location Based Services*, *5*(1), 22–37. https://doi.org/10.1080/17489725.2010.538016

Blankenbach, J., Sternberg, H., & Tilch, S. (2017). Indoor-Positionierung. Springer Spektrum, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47188-3{_}24

Britting, K. R. (1971). *Inertial navigation systems analysis*. Wiley-Interscience.

Burkov, A. (1997). The Hundred Page Machine Learning. *Computer*, *2005*(April), 414.

Chapelle, O., Schölkopf, B., & Zien, A. (2010). *Semi-supervised learning*. MIT Press.

Chen, C., Lu, X., Markham, A., & Trigoni, N. (2018). Ionet: Learning to cure the curse of drift in inertial odometry.

Chen, C., Zhao, P., Lu, C. X., Wang, W., Markham, A., & Trigoni, N. (2018). Oxiod: The dataset for deep inertial odometry.

Chen, C., Zhao, P., Lu, C. X., Wang, W., Markham, A., & Trigoni, N. (2020). Deep Learning based Pedestrian Inertial Navigation: Methods, Dataset and On-Device Inference. *IEEE Internet of Things Journal*, *7*(5), 4431–4441. https://doi.org/10.1109/JIOT.2020.2966773

Cho, K., van Merrienboer, B., Gulcehre, C., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *CoRR, abs/1406.1078*. http://arxiv.org/abs/1406.1078

Chugh, V. (2024). *Mean Shift Clustering: A Comprehensive Guide | DataCamp*. Retrieved March 1, 2025, from https://www.datacamp.com/tutorial/mean-shift-clustering

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. https://arxiv.org/abs/1412.3555v1

Da, S., Ivan, N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., & dos Reis Alves, S. F. (2016, January). *Artificial neural networks: A practical course*. Springer International Publishing. https://doi.org/10.1007/978-3-319-43162-8/COVER

Davidson, P., Collin, J., & Takala, J. (2010). Application of particle filters for indoor positioning using floor plans. *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service*, 1–4.

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020, April). *Mathematics for Machine Learning*. Cambridge University Press. https://doi.org/10.1017/9781108679930

Del Peral-Rosado, J. A., Raulefs, R., López-Salcedo, J. A., & Seco-Granados, G. (2018). Survey of Cellular Mobile Radio Localization Methods: From 1G to 5G. *IEEE Communications Surveys and Tutorials*, *20*(2), 1124–1148. https://doi.org/10.1109/COMST.2017.2785181

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 4171–4186. https://arxiv.org/abs/1810.04805v2

Diebel, J. (2006). *Representing attitude: Euler angles, unit quaternions, and rotation vectors* (Vol. 58). Matrix.

Dwivedi, S., Shreevastav, R., Munier, F., Nygren, J., Siomina, I., Lyazidi, Y., Shrestha, D., Lindmark, G., Ernstrom, P., Stare, E., Razavi, S. M., Muruganathan, S., Masini, G., Busin, A., & Gunnarsson, F. (2021). Positioning in 5G networks. *IEEE Communications Magazine*, *59*(11), 38–44. https://doi.org/10.1109/MCOM.011.2100091

Ebner, F., Fetzer, T., Deinzer, F., Koping, L., & Grzegorzek, M. (2015). Multi sensor 3D indoor localisation. *2015 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2015*. https://doi.org/10.1109/IPIN.2015.7346772

Fong, W. T., Ong, S. K., & Nee, A. Y. (2008). Methods for in-field user calibration of an inertial measurement unit without external equipment. *Measurement Science and Technology*, *19*(8), 085202. https://doi.org/10.1088/0957-0233/19/8/085202

García, A., Maier, S., & Philips, A. (2020). *Location-Based Services in Cellular Networks: from GSM to 5G NR*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE proceedings F (radar and signal processing)*, *140*(2), 107–113.

Graves, A., Wayne, G., & Danihelka, I. (2014). Neural Turing Machines. https://arxiv.org/abs/1410.5401v2

Gu, F., Khoshelham, K., Yu, C., & Shang, J. (2019). Accurate Step Length Estimation for Pedestrian Dead Reckoning Localization Using Stacked Autoencoders. *IEEE Transactions on Instrumentation and Measurement*, *68*(8), 2705–2713. https://doi.org/10.1109/TIM.2018.2871808

Han, J. W., S and. (2009). A novel method to integrate IMU and magnetometers in attitude and heading reference systems. *The Journal of Navigation, cambridge.org*. https://doi.org/10.1017/S0373463311000233

Harder, D., Shoushtari, H., & Sternberg, H. (2022). Real-Time Map Matching with a Backtracking Particle Filter Using Geospatial Analysis. *Sensors 2022, Vol. 22, Page 3289*, *22*(9), 3289. https://doi.org/10.3390/S22093289

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer New York. https://doi.org/10.1007/978-0-387-84858-7

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.

He, S., & Chan, S. H. (2016). Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. *IEEE Communications Surveys and Tutorials*, *18*(1), 466–490. https://doi.org/10.1109/COMST.2015.2464084

Heinz, K. (2021). *Grundlagen der physikalischen und mathematischen geodäsie*. https://doi.org/10.1007/978-3-662-62369-5

Herath, S., Caruso, D., Liu, C., Chen, Y., & Furukawa, Y. (2022). Neural inertial localization. https://arxiv.org/abs/2203.15851

Herath, S., Irandoust, S., Chen, B., Qian, Y., Kim, P., & Furukawa, Y. (2021). Fusion-dhl: Wifi, imu, and floorplan fusion for dense history of locations in indoor environments. *CoRR*, *abs/2105.08837*. https://arxiv.org/abs/2105.08837

Herath, S., Yan, H., & Furukawa, Y. (2020). RoNIN: Robust Neural Inertial Navigation in the Wild: Benchmark, Evaluations, & New Methods. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 3146–3152. https://doi.org/10.1109/ICRA40945.2020.9196860

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/NECO.1997.9.8.1735

Jekeli, C. (2001). *Inertial navigation systems with geodetic applications*. De Gruyter. https://doi.org/doi:10.1515/9783110800234

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Keller, F., & Sternberg, H. (2013). Multi-sensor platform for indoor mobile mapping: System calibration and using a total station for indoor applications. *Remote Sensing*, *5*(11), 5805–5824. https://doi.org/10.3390/rs5115805

Kingma, D. P., & Ba, J. L. (2014). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1412.6980v9

Klingbeil, L., & Wark, T. (2008). A wireless sensor network for real-time indoor localisation and motion monitoring. *Proceedings - 2008 International Conference on Information Processing in Sensor Networks, IPSN 2008*, 39–50. https://doi.org/10.1109/IPSN.2008.15

Klingbeil, Lasse. (2023, July). *Georeferencing of mobile mapping data* [Doctoral dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn]. https://hdl.handle.net/20.500.11811/10926

Kraeling, M., & Brogioli, M. C. (2019). Internet of Things. *Software Engineering for Embedded Systems*, 465–499. https://doi.org/10.1016/B978-0-12-809448-8.00013-8

Kuipers, J. (1999). *Quaternions and rotation sequences: A primer with applications to orbits, aerospace, and virtual reality*. Princeton University Press. https://books.google.de/books?id=_2sS4mC0p-EC

Laska, M. (2023). *Learned fingerprinting-based models for reliable localization in large buildings - RWTH AACHEN UNIVERSITY GIA - Deutsch* [doctoralThesis]. RWTH Achen University. https://doi.org/10.18154/RWTH-2023-07372

Li, F., Zhao, C., Ding, G., Gong, J., Liu, C., & Zhao, F. (2012). A reliable and accurate indoor localization method using phone inertial sensors. *UbiComp*. https://doi.org/10.1145/2370216.2370280

Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *37*(6), 1067–1080. https://doi.org/10.1109/TSMCC.2007.905750

Madgwick, S. O., Harrison, A. J., & Vaidyanathan, R. (2011). Estimation of IMU and MARG orientation using a gradient descent algorithm. *IEEE International Conference on Rehabilitation Robotics*. https://doi.org/10.1109/ICORR.2011.5975346

Malik, A. (2013). *Rtls for dummies*. John Wiley & Sons.

Mcculloch, W. S., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Journal of Symbolic Logic*, *9*(2), 49–50. https://doi.org/10.2307/2268029

Mizell, D. (2003). Using gravity to estimate accelerometer orientation, 252–252.

Mohd-Yasin, F., Nagel, D. J., & Korman, C. E. (2009). Noise in MEMS. *Measurement Science and Technology*, *21*(1), 012001. https://doi.org/10.1088/0957-0233/21/1/012001

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. The MIT Press.

Murphy, K. P. (2022). *Probabilistic Machine Learning: An Introduction*.

Ng, A. (2024a). Lecture note | neural networks and deep learning | coursera.

Ng, A. (2024b). Lecture note | supervised machine learning: Regression and classification | coursera.

Olivares, A., Olivares, G., Górriz, J. M., & Ramírez, J. (2009). High-efficiency low-cost accelerometer-aided gyroscope calibration. *Proceedings of the International Symposium on Test and Measurement*, *1*, 354–360. https://doi.org/10.1109/ICTM.2009.5412920

Pandit, S. M., Zhang, W., Pandit, S. M., & Zhang, W. (1986). Modeling random gyro drift rate by data dependent systems. *ITAES*, *22*(4), 455–459. https://doi.org/10.1109/TAES.1986.310781

Pinchin, J., Hide, C., & Moore, T. (2012). A particle filter approach to indoor navigation using a foot mounted inertial navigation system and heuristic heading information. *2012 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2012 - Conference Proceedings*. https://doi.org/10.1109/IPIN.2012.6418916

Poddar, S., Kumar, V., & Kumar, A. (2017). A Comprehensive Overview of Inertial Sensor Calibration Techniques. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, *139*(1). https://doi.org/10.1115/1.4034419

Poggio, T., Mhaskar, H., Rosasco, L., Miranda, B., & Liao, Q. (2017). Why and when can deep-but not shallow-networks avoid the curse of dimensionality: A review. *International Journal of Automation and Computing*, *14*(5), 503–519. https://doi.org/10.1007/S11633-017-1054-2

Radford, A., & Narasimhan, K. (2018). Improving Language Understanding by Generative Pre-Training.

Rai, A., Chintalapudi, K. K., Padmanabhan, V. N., & Sen, R. (2012). Zee: Zero-effort crowdsourcing for indoor localization. *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, 293–304. https://doi.org/10.1145/2348543.2348580

Rao, B., Kazemi, E., Ding, Y., Shila, D. M., Tucker, F. M., & Wang, L. (2022). CTIN: Robust Contextual Transformer Network for Inertial Navigation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(5), 5413–5421. https://doi.org/10.1609/AAAI.V36I5.20479

Romanovas, M., Goridko, V., Klingbeil, L., Bourouah, M., Al-Jawad, A., Traechtler, M., & Manoli, Y. (2013). Pedestrian Indoor Localization Using Foot Mounted Inertial Sensors in Combination with a Magnetometer, a Barometer and RFID. *Lecture Notes in Geoinformation and Cartography*, *0*(9783642342028), 151–172. https://doi.org/10.1007/978-3-642-34203-5{_}9

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/RD.33.0210

Savage, P. G. (2000). *Strapdown analytics*. Strapdown Associates.

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, *45*(11), 2673–2681. https://doi.org/10.1109/78.650093

Shi, Y., & Jensen, M. A. (2011). Improved radiometric identification of wireless devices using MIMO transmission. *IEEE Transactions on Information Forensics and Security*, *6*(4), 1346–1354. https://doi.org/10.1109/TIFS.2011.2162949

Shoemake, K. (1985). Animating rotation with quaternion curves. *ACM SIGGRAPH Computer Graphics*, *19*(3), 245–254. https://doi.org/10.1145/325165.325242

Solin, A., Cortes, S., Rahtu, E., & Kannala, J. (2018). Inertial odometry on handheld smartphones.

Statista. (2024). *Statista - das statistik-portal* [Accessed on January 9, 2025]. https://de.statista.com/

Sternberg, H. (2013). Engineering geodesy, and navigation lecture note [Accessed: 2020-03-01].

Sternberg, H., & Fessele, M. (2009). Indoor navigation with low-cost inertial navigation systems. *2009 6th Workshop on Positioning, Navigation and Communication*, 1–4. https://doi.org/10.1109/WPNC.2009.4907796

Sun, S., Melamed, D., & Kitani, K. (2021). IDOL: Inertial Deep Orientation-Estimation and Localization. *35th AAAI Conference on Artificial Intelligence, AAAI 2021*, *7*, 6128–6137. https://doi.org/10.1609/aaai.v35i7.16763

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *CoRR*, *abs/1409.3215*. http://arxiv.org/abs/1409.3215

Sutton, R. S., & Barto, A. G. (2012). *Reinforcement learning: An Introduction (Adaptive Computation and Machine Learning)* (Vol. 3). Bradford Books.

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. MIT Press.

Titterton, D., Weston, J., of Electrical Engineers, I., of Aeronautics, A. I., & Astronautics. (2004). *Strapdown inertial navigation technology*. Institution of Engineering; Technology.

Tolles, J., & Meurer, W. J. (2016). Logistic Regression: Relating Patient Characteristics to Outcomes. *JAMA*, *316*(5), 533–534. https://doi.org/10.1001/JAMA.2016.7653

Torge, W., Müller, J., & Pail, R. (2023, April). *Geodesy*. De Gruyter. https://doi.org/10.1515/9783110250008/PDF

Van Dierendonck, A. J., & Brown, R. G. (1969). Modeling Nonstationary Random Processes with an Application to Gyro Drift Rate. *IEEE Transactions on Aerospace and Electronic Systems*, *AES-5*(3), 423–428. https://doi.org/10.1109/TAES.1969.309844

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, ., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, *2017-December*, 5999–6009. https://arxiv.org/abs/1706.03762v7

Wang, X., Chen, G., Yang, M., & Jin, S. (2020). A Multi-Mode PDR Perception and Positioning System Assisted by Map Matching and Particle Filtering. *ISPRS International Journal of Geo-Information*, *9*(2), 93. https://doi.org/10.3390/ijgi9020093

Want, R., Hopper, A., Falcão, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems (TOIS)*, *10*(1), 91–102. https://doi.org/10.1145/128756.128759

Weiser, M. (1999). a Leading Computer Visionary, Dies at 46 - The New York Times. Retrieved March 1, 2025, from https://www.nytimes.com/1999/05/01/business/mark-weiser-a-leading-computer-visionary-dies-at-46.html

Welbaum, A., & Qiao, W. (2025). Mean shift-based clustering for misaligned functional data. *Computational Statistics & Data Analysis*, *206*, 108107. https://doi.org/10.1016/J.CSDA.2024.108107

Wendel, J. (2011). *Integrierte Navigationssysteme*. Oldenbourg Verlag, München. https://doi.org/10.1524/9783486705720

Widyawan, Klepal, M., & Beauregard, S. (2008). A Novel Backtracking Particle Filter for Pattern Matching Indoor Localization. *Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-Less Environments*, 79–84. https://doi.org/10.1145/1410012.1410031

Willemsen, T. (2016). *Fusionsalgorithmus zur autonomen positionsschätzung im gebäude, basierend auf mems-inertialsensoren im smartphone* [doctoralThesis]. HafenCity

Universität Hamburg. https://nbn-resolving.org/urn%3Anbn%3Ade%3Agbv%3A1373-opus-3156

Woodman, O., & Harle, R. (2008). Pedestrian localisation for indoor environments. *Proceedings of the 10th international conference on Ubiquitous computing*, 114–123.

Wymeersch, H. (2017). 5G Positioning Tutorial - YouTube. Retrieved March 1, 2025, from https://www.youtube.com/watch?v=VKF-Xgn0O6A

Xiao, Z., Wen, H., Markham, A., & Trigoni, N. (2014). Robust pedestrian dead reckoning (r-pdr) for arbitrary mobile device placement. *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 187–196. https://doi.org/10.1109/IPIN.2014.7275483

Yan, H., Herath, S., & Furukawa, Y. (2019). Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, and new methods.

Yan, H., Shan, Q., & Furukawa, Y. (2017). RIDI: robust IMU double integration. *CoRR*, *abs/1712.09004*.

Yu, C., Lan, H., Gu, F., Yu, F., & El-Sheimy, N. (2017). A map/ins/wi-fi integrated system for indoor location-based service applications. *Sensors*, *17*(6), 1272.

Zafari, F., Gkelias, A., & Leung, K. K. (2019). A Survey of Indoor Localization Systems and Technologies. *IEEE Communications Surveys and Tutorials*, *21*(3), 2568–2599. https://doi.org/10.1109/COMST.2019.2911558

Zhang, F., & O'Donnell, L. J. (2020). Support vector regression, 123–140. https://doi.org/10.1016/B978-0-12-815739-8.00007-9

Zhang, Q., Pang, H., & Wan, C. (2016). Magnetic interference compensation method for geomagnetic field vector measurement. *Measurement: Journal of the International Measurement Confederation*, *91*(May), 628–633. https://doi.org/10.1016/j.measurement.2016.05.081

Zhao, Y., Yang, R., Chevalier, G., Xu, X., & Zhang, Z. (2017). Deep Residual Bidir-LSTM for Human Activity Recognition Using Wearable Sensors. *Mathematical Problems in Engineering*, *2018*. https://doi.org/10.1155/2018/7316954

A Publication A1 - Peer reviewed

Shoushtari, H., Harder, D., Kasparek, M., Schäfer, M. Müller-Lietzkow, J., Sternberg, H. 2023; Data-Driven Inertial Navigation assisted by UL-TDOA 5G Positioning, International technical Meeting (ITM), Institute of Navigation (ION), Long Beach, California. https://doi.org/10.33012/2023.18645

Contribution of Co-Authors

The content structure of the paper was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Dorian Harder was involved in the data collection and evaluation section by applying the specified metrics. Maximilian Kasparek and Matthias Schäfer were responsible for the 5G positioning and reference data preparation. Harald Sternberg and Jörg Müller-Lietzkow oversaw and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 80 %. The specific contributions of the first author are outlined below:

Table A.1: Contribution to publication A1.

Involved in	**Estimated contribution**
Ideas and conceptual design	90%
Computation and results	70%
Analysis and interpretation	80%
Manuscript, figures and tables	80%
Total:	>80%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication A1 in Table A.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

B Publication A2 - Peer reviewed

Shoushtari, H., Kassawat, F., Harder, D., Venzke, K., Müller-Lietzkow, J., Sternberg, H., 2022; L5IN+: From an Analytical Platform to Optimization of Deep Inertial Odometry, International Conference on Indoor Positioning and Indoor Navigation (IPIN) , Beijing, China.
https://ceur-ws.org/Vol-3248/paper24.pdf

Contribution of Co-Authors

The content structure of the research work was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Firas Kassawat, as the first co-author, assisted in the model implementation and paper writing. Dorian Harder was involved in the data collection and simulation section. Korvin Venzke was responsible for website implementation tasks. Harald Sternberg and Jörg Müller-Lietzkow oversaw and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 85%. The specific contributions of the first author are outlined below:

Table B.1: Contribution to publication A2.

Involved in	**Estimated contribution**
Ideas and conceptual design	100%
Computation and results	80%
Analysis and interpretation	90%
Manuscript, figures and tables	80%
Total:	>85%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication A2 in Table B.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

C Publication B1 - Peer reviewed

Shoushtari, H., Kassawat, F., Sternberg, H. 2024; Context aware transformer network and in-situ IMU calibration for accurate positioning, IPIN, Hongkong.
https://doi.org/10.1109/IPIN62893.2024.10786130

Contribution of Co-Authors

The content structure was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Firas Kassawat, as the first co-author, assisted in the model implementation and paper writing. Harald Sternberg supervised and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 75%. The specific contributions of the first author are outlined below:

Table C.1: Contribution to publication B1.

Involved in	**Estimated contribution**
Ideas and conceptual design	90%
Computation and results	70%
Analysis and interpretation	80%
Manuscript, figures and tables	70%
Total:	>75%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication B1 in Table C.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

D Publication B2

Shoushtari, H., Willemsen, T., Sternberg, H. 2023; Supervised Learning Regression for Sensor Calibration, DGON Inertial Sensors and Systems (ISS), Braunschweig, Germany. https://doi.org/10.1109/ISS58390.2023.10361922

Contribution of Co-Authors

The content structure of this research work was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Thomas Willemsen, as the first co-author, assisted in the model implementation and paper writing. Harald Sternberg supervised and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 80%. The specific contributions of the first author are outlined below:

Table D.1: Contribution to publication B2.

Involved in	**Estimated contribution**
Ideas and conceptual design	100%
Computation and results	70%
Analysis and interpretation	90%
Manuscript, figures and tables	90%
Total:	>80%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication B2 in Table D.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

E Publication C1 - Peer reviewed

Shoushtari, H., Harder, D., Willemsen, T., Sternberg, H. 2023; Optimized Trajectory from Smartphone Sensors and 5G UL-TDoA Using Cluster Particle Filters (Original title in german: Optimierte Trajektorie aus Smartphone-Sensoren und 5G UL-TDoA mit Cluster Partikel Filter), Internationalen Ingenieurvermessung, Zürich, Hrsg.: Andreas Wieser, Wichmann Verlag, ISBN/EAN:9783879077342

Contribution of Co-Authors

The content structure of the paper was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Dorian Harder, as the second author, was involved in the data collection, implementation, and evaluation section, especially in regards to particle filter development. Thomas Willemsen and Harald Sternberg oversaw and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 75%. The specific contributions of the first author are outlined below:

Table E.1: Contribution to publication C1.

Involved in	**Estimated contribution**
Ideas and conceptual design	90%
Computation and results	70%
Analysis and interpretation	70%
Manuscript, figures and tables	80%
Total:	>75%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication C1 in Table E.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

F Publication C2 - Peer reviewed

Shoushtari, H., Askar, C., Harder, D., Willemsen, T., Sternberg, H., 2021; 3D Indoor Localization using 5G-based Particle Filtering and CAD Plans, IPIN , Lloret de Mar, Spain.
https://doi.org/10.1109/IPIN51156.2021.9662636

Contribution of Co-Authors

The content structure of this research work was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Cigdem Askar, the first co-author, was responsible for assisting with map extraction and paper writing. Dorian Harder was involved in the implementation section by executing the specified models. Harald Sternberg and Thomas Willemsen oversaw and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 85%. The specific contributions of the first author are outlined below:

Table F.1: Contribution to publication C2.

Involved in	**Estimated contribution**
Ideas and conceptual design	100%
Computation and results	90%
Analysis and interpretation	85%
Manuscript, figures and tables	80%
Total:	>85%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication C2 in Table F.1.

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg

G Publication C3 - Peer reviewed

Shoushtari, H., Willemsen, T., Sternberg, H., 2021; Many Ways Lead to the GoalPossibilities of Autonomous and Infrastructure-Based Indoor Positioning, Electronics Special Issue Indoor Positioning Techniques.
https://doi.org/10.3390/electronics10040397

Contribution of Co-Authors

The content structure of this paper was primarily developed by Hossein Shoushtari, who also played a major role in all aspects as the first author. Thomas Willemsen, as the first co-author, assisted in the model implementation and paper writing. Harald Sternberg supervised and proofread the project. Overall, the first authors contribution to the paper can be estimated to be over 80%. The specific contributions of the first author are outlined below:

Table G.1: Contribution to publication C3.

Involved in	**Estimated contribution**
Ideas and conceptual design	90%
Computation and results	70%
Analysis and interpretation	90%
Manuscript, figures and tables	90%
Total:	>80%

On behalf of the co-authors, I hereby confirm the correctness of the declaration of the contribution of Hossein Shoushtari for publication C3 in Table G.1:

Prof. Dr. -Ing. Harald Sternberg, HafenCity Universität Hamburg