IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 January 2024, accepted 18 January 2024, date of publication 24 January 2024, date of current version 5 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358274

RESEARCH ARTICLE

Selecting and Combining UWB Localization Algorithms: Insights and Recommendations From a Multi-Metric Benchmark

BEN VAN HERBRUGGEN[1], JONO VANHIE-VAN GERWEN[1], STIJN LUCHIE[1], YURI DURODIÉ[2], (Graduate Student Member, IEEE), BRAM VANDERBORGHT[2], MICHIEL AERNOUTS[3], ADRIAN MUNTEANU[4], (Member, IEEE), JARON FONTAINE[1], AND ELI DE POORTER[1]

[1]IDLab Research Group, Department of Information Technology, Ghent University–imec, 9052 Gent, Belgium
[2]Brussels Human Robotic Research Center (BruBotics), Vrije Universiteit Brussel–imec, 1050 Brussels, Belgium
[3]IDLab Research Group, Faculty of Applied Engineering, University of Antwerp–imec, 2020 Antwerpen, Belgium
[4]ETRO Department, Vrije Universiteit Brussel–imec, 1050 Brussel, Belgium

Corresponding author: Ben Van Herbruggen (ben.vanherbruggen@ugent.be)

This work was supported by the Fund for Scientific Research Flanders, Belgium, Fonds voor Wetenschappelijk Onderzoek (FWO)-Vlaanderen, FWO-Strategisch basisonderzoek (SB), under Grant 1SB7619N.

ABSTRACT Ultra-wideband (UWB) localization has emerged in GPS-denied environments as a crucial facilitator for diverse industries, including logistics, healthcare applications, and societal domains. Despite notable progress, UWB algorithms from the scientific literature are often evaluated in isolation in very specific conditions and hence difficult to compare. This paper introduces a novel benchmark platform designed to assess the performance of 11 prominent UWB accuracy-enhancing algorithms, both independently and in combination. A key feature of the platform is its incorporation of multiple diverse evaluation metrics, including mean average error, latency, and spatial error. We showcase the significance of adopting alternative metrics such as spatial error, which often, depending on the use case, offers greater relevance than the prevalently used mean average error. Furthermore, we show that ''more is better'' does not hold true when combining multiple accuracy-improving algorithms for UWB systems: combining multiple accuracy-improving algorithms reveals instances of diminishing returns and can even result in overall performance decline. Additionally, we caution against blind reliance on accuracy outcomes reported in the scientific literature when designing UWB systems for use cases that are different in terms of requirements or environment. Finally, we also provide algorithmic recommendations for distinct surroundings, exemplary applications, and usage contexts, assisting in driving efficient design in future UWB research and adoption.

INDEX TERMS Benchmark, filters, guidelines, localization, LOS/NLOS, software, UWB, ultra-wideband.

I. INTRODUCTION

Many industry 4.0 and consumer applications require spatial awareness of persons and internet of things (IoT) devices. To facilitate this spatial awareness, an application designer can choose different localization technologies depending on the requirements. With the absence of GPS signals indoors, different other technologies can be considered for indoor localization including Bluetooth Low Energy (BLE) [1], inertial measurement unit (IMU) [2], 5G [3], motion capturing (MOCAP) [4], and light detection and ranging (lidar) [5]. However, these technologies suffer from less accurate localization, provide no absolute positioning, require line-of-sight (LOS), and/or are not power efficient. Ultra-wideband (UWB) is a popular choice for high-accuracy indoor localization as the radio frequency (RF) technology has excellent indoor localization properties. Firstly, the high bandwidth enables a high time resolution and low transmission power [6]. Incoming UWB pulses can be timestamped with sub-nanosecond resolution, which enables

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed Kheir.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

the estimation of the position of mobile tags with accuracies that range from several centimeters (in optimal conditions) to around 30 cm accuracy in more realistic conditions [7]. In addition, this resolution causes good resilience against RF channel imperfections, such as multipath effects and fading. Thirdly, UWB does not cause interference with existing technologies in the same frequency bands, due to its low transmission power. The UWB technology is currently rapidly maturing with increased availability in consumer electronics, chip manufacturers, and custom system providers, while standardization and interoperability efforts are in progress [8]. University testbeds are deployed to investigate the compatibility and interoperability of different UWB transceiver platforms [9]. In literature, many scientific contributions are proposed to evaluate UWB systems in a plethora of use cases. For example, the FiRa consortium, a consortium of more than 100 members focusing on the development of UWB, defines 4 main use case categories with in total 37 unique applications for UWB technology [8]. In industrial environments, UWB can be used to track assets for which a software architecture is proposed in [10]. By using an indoor location system for asset tracking, supply chains can be optimized and labor costs can be reduced. Other use cases include athlete tracking in sports [11], [12], [13], radar [14], unmanned aerial vehicles (UAVs) [15], [16] and many more. Recent scientific research papers have proposed techniques to improve UWB indoor localization accuracy further, using techniques such as bias correction, machine learning based error correction, outlier removal, adaptive physical layer (PHY) settings, path filtering, etc. Many of these papers report promising performance gains, especially in non-line-of-sight (NLOS) conditions. The evaluation is done by comparing the approximated points with a ground truth. The ground truth points are the true locations of the device and need to be known with much higher precision than the estimated points. However, the results of these studies are difficult to interpret and compare: (i) the algorithms are evaluated in different setups (different environments, different mobility patterns, different hardware platforms, different ground truth methodologies, etc.), (ii) the systems use different localization techniques (e.g. two-way-ranging (TWR) versus time-difference-of-arrival (TDoA)), (iii) or uses different post-processing techniques (real-time position updates versus post-processing using filtering methods), (iv) using different evaluation metrics, etc. As such, **based on current research papers, it is practically impossible to separate the evaluation methodology's impact from the optimization techniques' actual gains.** Moreover, many proposed state-of-the-art improvement techniques focus on different levels of processing steps (e.g. ranging bias correction versus positioning error correction). However, improving the ranging error might not always lead to similar improvements in terms of positioning error. As such, it is not clear if combining optimizations at different phases would still improve the accuracy to the same degree as both individual techniques or to a lesser degree. At the moment, **there is no systematic investigation into the question of how individual optimization techniques perform when combined within the same localization system**.

The main contributions and innovations of this paper are:

1) Description of the design of a software platform to evaluate different off-the-shelf algorithms and post-processing pipelines for UWB positioning systems.
2) Different state-of-the-art software post-processing algorithms are objectively evaluated and clear guidelines on when to use them are formulated.
3) A multi-metric analysis from the perspective of secondary objectives (e.g. power consumption) rather than only accuracy metrics (e.g. mean absolute error).
4) The presented results are associated with five specific practical use cases, and recommendations have been developed for these use cases.
5) We provide dynamic open-source datasets with both TWR and TDoA collected data in a controlled environment with full LOS scenarios and scenarios including NLOS with mm-level accurate ground-truth. This enables performance analysis of future localization methods and a thorough comparison of their performance.

The remainder of this paper is structured as follows: Section II gives an overview of the basic UWB principles. Section III discusses the state-of-the-art related work for benchmarks of localization systems and UWB benchmark initiatives. This related work Section is followed by Section IV which provides an overview of the measurement setup. In Section V the used algorithms are discussed and in Section VI all the module combinations in the platform are evaluated. These results are put in perspective for different use cases in Section VII. Sections VIII and IX are discussing future work and are concluding this work, respectively.

II. BASIC PRINCIPLES

Different ways exist to estimate the mobile node's location in an indoor localization system: (i) received signal strength indicator (RSSI) values and/or the channel impulse response (CIR) information can be used to model or to allow fingerprinting within the environment using machine learning (ML) models [17], [18], [19], [20] (Fig. 1a), (ii) the angle which is estimated based on the phase information at an antenna array can be used in triangulation [21] (Fig. 1b), (iii) time information can be used in multilateration [22] (Fig. 1c-f) or (iv) hybrid approaches combining multiple localization techniques (e.g. Djosic et al.) [23] combines multilateration with fingerprinting.

Currently, in UWB localization time-based approaches are the most widespread due to their high accuracy, generalization to new environments (one of the drawbacks of ML fingerprinting), and rather simple hardware (compared to the more complex antenna arrays for AoA). For the

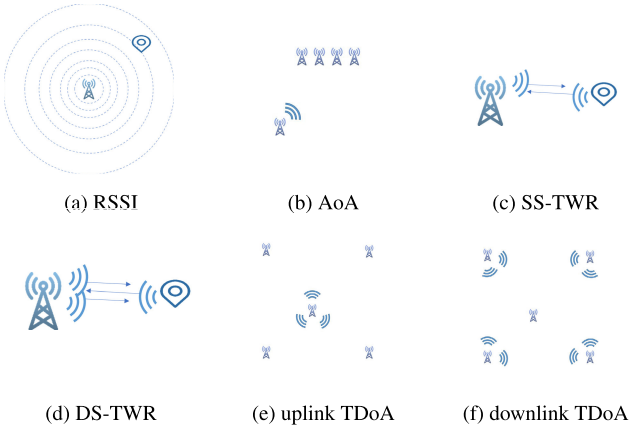


FIGURE 1. Different techniques exist to estimate positions based on UWB signals: based on (a) received signal strength indicator values, (b) angle-of-arrival, (c) single sided two-way-ranging, (d) double sided two-way-ranging, (e) uplink time-difference-of-arrival and (f) downlink time-difference-of-arrival.

time-based approach, two popular time-based techniques can be distinguished.

- In two-way-ranging (TWR) a sender and receiver are exchanging 2 or 3 UWB packets to measure the propagation time of the packet. Combining this with the propagation speed of the UWB signals results in a distance between the nodes. The simpler 2-packet version, single sided two-way-ranging (SS-TWR) (Fig. 1c) reduces the influence of clock offset between the nodes, but the clock drift error should still be compensated. This clock drift error compensation is inherently incorporated in the 3 packet schemes (double sided two-way-ranging (DS-TWR), Fig. 1d).
- A second time-based approach is time-difference-of-arrival (TDoA). Two variants of this technique exist, an uplink version (Fig. 1e) with lower power consumption for the tag and the downlink version (Fig. 1f) with high scalability to the number of tags and privacy for the tag's location. In TDoA, the difference in arrival time between the arriving packets from (downlink) or at (uplink) the anchors leads to a hyperbola in 2D-positioning or a hyperboloid in 3D-positioning of possible tag locations. The TDoA and distances are the base measurements that the localization engine will combine to a single position.

When the timing information from the anchors is collected, the localization server combines them into a position estimate. Many different methods exist to estimate the position and a use case-specific trade-off has to be made. A first example of applications that exploit UWB is the precise localization of robots [24]. These industrial robotic applications need highly accurate localization with low latency that is robust for changes in the environment which makes UWB an excellent choice. Another example is presented in [25], where the authors utilized a periodic extended Kalman filter (PEKF) to combine the range information from two anchors and one tag concerning the position of a rowing machine. UWB based localization systems are also used in other sports. A review study on the improvements of UWB indoor localization systems of the last 5 years discusses many possible use cases and corresponding localization algorithms [7]. Some of the localization algorithms exploit information from other sensors to enhance their performance. The technology is commercialized and a wide overview of different systems is used. Most of the current discussed systems use least-squares multilateration, Kalman filtering, or a particle filtering approach to estimate the position. These traditional approaches are still optimal for standard multilateration use cases, but different ML models are proposed to cover specific environments or to increase robustness [26], [27], [28].

There is a consensus among published UWB research regarding the performance degradation of UWB localization systems in NLOS conditions. However, it is worth noting that the presence of visual NLOS does not always impact performance to the same degree. Depending on the type of material, some visually opaque materials still allow signal penetration through obstacles, and as such there may still be a direct path available for the received signal. The study made in [29] extensively discusses the interpretation of NLOS, methods for simulating NLOS environments, and the objective determination of NLOS characteristics for comparing different environments. Their findings demonstrate that in the absence of multipath reflections, it is challenging to distinguish between line-of-sight (LOS) and visual NLOS.

III. RELATED WORK ON STANDARDIZED UWB BENCHMARKING

A plethora of literature works are reporting on localization systems but it is currently difficult to compare different optimizations as they are tested on different datasets in different environments. Surveys also tend to report the results of the developers which makes it difficult to select the best option. Therefore, there is a need for objective evaluation methods for localization systems. This is currently done by standards, localization competitions, and open-source datasets. In addition to these efforts, researchers already presented small benchmarks. The ISO/IEC 18305:2016 International Standard is one of the first initiatives to provide a standardized methodology for evaluating indoor localization systems and defines a complete framework for performing tests and evaluation of localization and tracking systems. The authors of [35] formulated many remarks on this standard. The most informative aspect for the user's questions lies in the quantiles of the errors, they specify how often the system gives wrong results. The authors suggest using the quantiles 0.5,0.75,0.9 and 0.95 to report the performance of a system. The standard specifies only metrics related to errors between positions and ground truth points but no other metrics related to the trajectory. The selection of the evaluation metric was also one of the incentives for the EVARILOS project [30]. This benchmark project was

TABLE 1. Overview of prior benchmarking systems and approaches. Despite the presence of multiple datasets, no comparisons have been made of UWB optimizations and combinations from scientific literature in the same conditions.

	Method	Metric	Ground truth	UWB	Guidelines	Robots	Different environments	Different algorithms	Transceiver
Heydariaan et al [29]	benchmark platform	boxplot		✓	✗	✗	✗	✗	DW1000
EVARILOS [30]	benchmark platform	custom formula	Not reported	✗	✓	✗	✓	✗	
EvAAL [31]	localization competition	75th percentile absolute error	MOCAP	✓	✓	✗	✗	✓	
IPSN microsoft [32]	localization competition	average location error	3-D laser scanners	✓	✓	✗	✓	✓	
Arjmandi et al [15]	open source dataset	MAE	Leica Nova MS60 Multistation	✓	✗	✓	✓	✗	P400
Stocker et al [27]	open source dataset	Mean Error & RMS	Leica TS11 Total Station	✓	✗	✗	✓	✗	DW1000
Queralta et al [33]	open source dataset	trends in graphs	MOCAP	✓	✗	✓	✗	✓	DW1000
LocURa4IoT [34]	testbed + open source dataset	trends in graphs	motorised rail	✓	✗	✗	✗	✗	DW1000
Shuh et al [9]	testbed + benchmark	packet reception rate	Not reported	✓	✓	✗	✗	✗	DW1000/DW3000
Our paper	**benchmarking platform + open source dataset**	**MAE and spatial error**	**MOCAP**	✓	✓	✓	✓	✓	**DW1000**

one of the first major RF localization projects focusing on benchmarking different RF localization systems. Based on different metrics such as accuracy, energy efficiency, latency, etc. a global score is calculated which is used for scoring different solutions. Secondary metrics such as mobility, repeatability, scalability, and interference profiles are also taken into account when evaluating a solution. Although EVARILOS is a relevant benchmark example, it does not include UWB. To overcome the difficulties in comparing localization solutions, different localization competitions are organized where the competitors work on the same data. The conference on Indoor Positioning and Indoor Navigation (IPIN) has a lot of experience in organizing localization competitions [31]. Over the years, different tracks have been organized, both held offline and online. The conference uses the evaluation platform EvAAL whose main goal is the evaluation of indoor localization systems in diverse scenarios and contexts but with a common metric. The solutions are evaluated with the 75th percentile of the mean absolute error, where the ground truth has been determined by using a mm-accurate MOCAP system. The conference had a specific UWB based track in an industrial environment. A second conference that organized localization competition is the Information Processing in Sensor Networks (IPSN) conference. This conference organized an indoor localization competition in collaboration with Microsoft for four consecutive years between 2014 and 2017. The lessons learned from this competition are bundled in [32]. UWB localization systems built on the DW1000 were scoring very well in the competition but required dedicated hardware installation with long deployment times. A second conclusion was that the different systems had lower accuracies during the competition than when tested in the lab of the researchers which expresses the need to have a unified evaluation environment for indoor localization systems. A last remark is that the localization error metric is influenced by the chosen points: if more easy points (without obstacles) are evaluated the error metrics will also be better.

The need for a unified dataset for evaluating different algorithms is also indicated in [15]. The authors have made a subset of drone-collected datasets available for UWB designers to test their algorithms on and to be able to validate the algorithms in detail. Environments available in this dataset are: indoor, field, building, bridge, and tunnel but the tag was differently mounted on the drone for the indoor scenarios than for outdoor scenarios making it more difficult to compare different environments. A range calibration is already performed before the publication of the dataset. With multilateration approaches, the mean absolute error is between 0.17 m (Field) and 2.02 m (Building). To collect the datasets four PulseON 440 UWB anchors are used and the ground truth is collected with the Leica Nova MS60 Multistation.

Currently, it is rather unknown how the ranging performance will be with different interoperable UWB transceivers. The results reported in [9] are one of the first who delved into interoperability, conducting experiments in a testbed environment where hardware from various UWB providers was installed and benchmarked their performance under interference from WiFi 6E communication. In their testbed, Decawaves DW1000 and DW3000 are interoperated with NXP transceivers, and all nodes are connected to a network of Raspberry Pis. The authors conclude that the coexistence between WiFi 6E and UWB is not trivial.

A different new UWB testbed to evaluate ranging information was presented at the IPIN conference 2018 [34]. The testbed has 20 fixed DW1000-based nodes and two mobile nodes placed at motorized rails for effective tests. The goal of the testbed is to investigate various physical properties such as propagation conditions, clock skew, signal reflections, etc. Two datasets are publicly available that are recorded in the testbed. The first dataset contains timestamp information from symmetric double sided two-way-ranging (SDS-TWR) and the authors compare SS-TWR, with and without clock skew compensation to DS-TWR. With compensation of the clock skew, similar accuracy can be reached while at the same time, one UWB packet can be saved per range. The second dataset focuses on a node moving from a hard NLOS environment to a LOS situation with the anchor. Mitigation techniques for NLOS errors can be further investigated here. To increase the reproducibility of the results, datasets are also published with their corresponding paper. For example, the authors of [33] published an open source dataset collected with a UAV, where the system has 3-6 anchors. A second extensive dataset with UWB ranging information for ML error mitigation is also available [27]. The number

of available datasets is increasing, but as environmental influences are very important and no common data format is available, it is difficult to compare them objectively.

An overview of the discussed benchmarking initiatives is given in Table 1. None of these show the same algorithms in different environments and the influence of a combination of different algorithms.

IV. MEASUREMENT SETUP

This section describes the setup that was used for a repeatable benchmark. During the design of the used setup, careful considerations are given to repeatability, reproducibility, and transparency in the results. While the data are collected in a realistic scenario, all the state-of-the-art algorithms are evaluated offline both in isolation and in combination with each other. The ground truth is collected with mm-level precision.

A. LAB INFRASTRUCTURE

The datasets collected to evaluate the benchmark platform were captured in a controlled and industrial lab environment of the IDLab research group at Ghent University (Fig. 2). The Industrial Internet of Things (IIoT) lab [36] contains a 240m^2 warehouse environment, representative of many Industry 4.0 use cases. To meet the benchmarking requirements, the lab's open space subsection was used to achieve repeatable and mm-accurate ground truth in the raw (not preprocessed and unfiltered) dataset. This open space of 7 × 11 × 2.5 m (L×W×H) is equipped with 8 Qualisys Miqus M3 MOCAP cameras [37], capable of tracking hundreds of passive infrared reflective motion capturing (MOCAP) markers with a quantified uncertainty in the millimeter range at speeds up to 340Hz.

To achieve a controllable NLOS influence in the environment without compromising the MOCAP setup, RF absorbing panels (−18 dB reflection loss at 6 GHz [38]) are added in the middle of the open space in the lab. The panels are placed in such a way that some of the links are attenuated, while also guaranteeing coverage from the MOCAP system to maintain an accurate ground truth (see Fig. 3). The placement of the absorbing panels results in both weak and strong NLOS. Weak NLOS is defined as links that are visually NLOS and RF absorbed but still permit the first path to be detected correctly [27]. In strong NLOS links, the first path is completely attenuated and the transceiver timestamps a reflected path. This controllable environment facilitates an objective comparison in various propagation scenarios. The introduction of absorbing panels allows for precise control over alterations to the environment and characterization of the differences between LOS and NLOS scenarios. The number of reflections is limited in the measurement area but metal surfaces are available close to the measurement area and still influence the signal propagation from tag to anchors. In the NLOS scenarios, the number of NLOS links to the anchors varies up to 8 blocked anchors, but for most points at least one link is LOS to permit the algorithms to detect unreliable links and mitigate these effects.

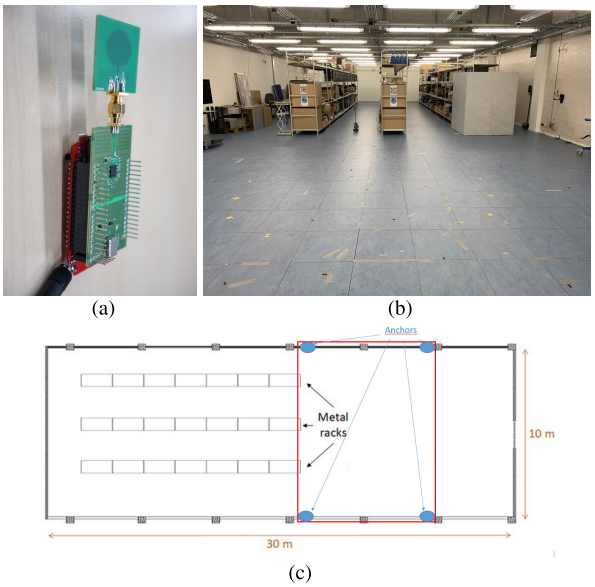


FIGURE 2. The measurements were performed in the open space area of the iiot-lab at Ghent University which simulates an industrial environment. (a) The UWB hardware platform for the data collection [6], (b) the iiot-lab environment, and (c) the ground plan of the environment.

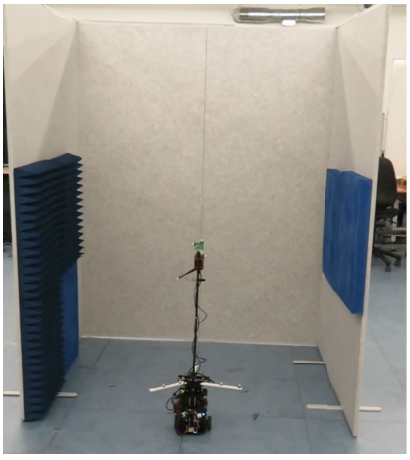


FIGURE 3. By adding RF absorbers, weak and strong NLOS links are created to evaluate the robustness of the different algorithms.

B. UWB

The data is collected with the Wi-Pos UWB system [6], a platform developed for research data collection with a wireless long-range Sub-GHz backbone combined with UWB ranging based on the DW1000 shown in Fig. 2. The UWB devices are configured with a center frequency of 6489.6 MHz, a bandwidth of 499.2 MHz (channel 5), and a preamble length of 1024 symbols at the suggested TX power (7.5 dB) from the DW1000 user manual [39]. The bitrate for the measurements is 850 kbps and a pulse repetition frequency (PRF) of 64 MHz is used. Eight UWB

anchors are placed in a cuboid setup of 8.0 × 10.8 × 2.2 m^3 in the previously mentioned lab infrastructure. Both uplink TDoA [22] and asymmetric double-sided TWR [40] are supported. While the first has good scalability to the number of anchors and low power consumption at the tag device, the latter does not require any form of clock synchronization between the anchor nodes. We collected TDoA and TWR data using two firmware versions. In the version without CIR, output for TWR includes timestamp, range, and anchors' addresses. For TDoA, the output includes a timestamp, tag address, anchors' addresses, superframe counter, and clock drift information. The system achieves a high update rate with limited UART communication. When CIR is collected, the same values are captured, along with additional values from registers: received power, first path power, first path amplitudes, peak path index and amplitude, preamble accumulation count, LDE threshold, CIR noise, and complex CIR samples between indexes 650 and 950. Transmitting this expanded information over UART slows down the system, resulting in a lower update rate. The reduced update rate, combined with the need for a higher throughput link between anchors and the localization engine, may not be justified compared to potential benefits in a system with a lower update rate and improved filtering.

C. ROBOTIC PLATFORM

All experiments were performed using a TurtleBot 3 Burger [41] robot base platform, with a height extension beam to place the UWB tag at a fixed height of 1 m. This places the tag at a more representative height for most use cases involving person and object localization, while also guaranteeing line-of-sight concerning the MOCAP cameras. This allows the robot platform to navigate based on the input from the MOCAP system, providing a mm-accurate ground truth at all times, even if the robot deviates more from the planned trajectory. Both the robot platform and the UWB tag are fitted with passive infrared reflective MOCAP markers, giving two Six Degrees of Freedom (6DoF) bodies that fully describe both the robot and UWB hardware in both position and rotation in 3D space at 100 Hz. These external measurements are timestamped and synchronized by the same system that records the robot platform metrics (from CPU usage to wheel encoders and position on the experiment trajectory) and are finally combined with the measurements of the actual system under test (SUT): the UWB tag and anchors. This approach satisfies the proposed repeatability, reproducibility, and transparency requirements by collecting all metrics with a unified time source. The same trajectory can be executed for different physical settings and changes in the environment which makes the experiment repeatable.

D. DATASETS

The robotic platform is set to capture different datasets on the same trajectory to compare the actual results.

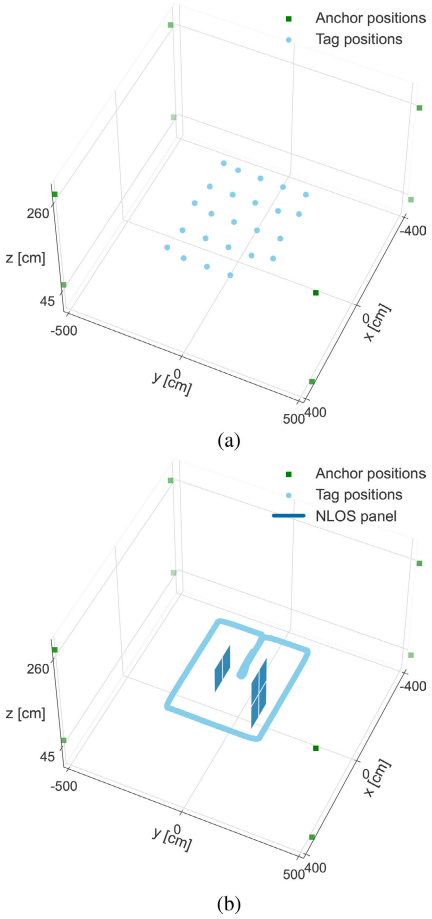


FIGURE 4. Both static (a) and dynamic (b) trajectories are collected in the iiot-lab. The dynamic scenario covers realistic sharp corners and areas with high numbers of NLOS links are created with the absorbing panels.

An overview of all captured and considered datasets is given in Table 2. The different datasets are selected carefully in a controlled environment to see the influence of the different algorithms for the following operational conditions: TWR vs TDoA, LOS vs NLOS, static vs dynamic, and lastly if the CIR information was collected or not. Due to the limited throughput from the nodes, the collection of the CIR will lower the update rate of the dataset. To evaluate the influence of static vs dynamic, two different movement patterns are executed in the lab. For static use cases, a 25-position grid was measured with a surface area of $16m^2$ (Fig. 4a). For the dynamic use cases the turtlebot drives from the origin along the y-axis to the point with x = −2. From this point, the robot rides a perfect square of 4 × 4 m, before returning to the origin (Fig. 4b). This dynamic pattern has been executed at low speed (0.1 m/s) and high speed (0.2 m/s). In total three different mobility patterns (static, low speed dynamic, and high speed dynamic) are evaluated. An overview of all the datasets and their settings is given in Table 2. All datasets have a mean absolute error (MAE) between 12.7 cm and 22.3 cm.

TABLE 2. A representative collection of datasets is used for software evaluation in the platform. Both LOS and NLOS environments are evaluated in different patterns with both TWR and TDoA techniques.

	TWR	TDoA	robot speed	NLOS	CIR collected	length [s/lap]	laps	Datapoints	UWB update rate [Hz]
1		✔	static		✔			4 552	9.9
2		✔	high		✔	109.8	4	5 326	9.9
3		✔	low		✔	197.6	4	9 989	9.9
4		✔	low			197.6	5	31 376	26.2
5		✔	high	✔	✔	110.4	4	5 114	9.9
6		✔	low	✔	✔	198.5	4	9 427	9.9
7		✔	high	✔		107.6	4	12 945	26.2
8		✔	low	✔		196.1	5	25 916	22.0
9	✔		static		✔			2 869	7.4
10	✔		high		✔	113.6	8	8 121	7.4
11	✔		low		✔	200.3	4	7 429	7.3
12	✔		high			110.0	4	9 580	17.2
13	✔		high	✔	✔	109.9	4	4 001	7.4
14	✔		low	✔	✔	197.0	1	1 857	7.4
15	✔		high	✔		110.8	4	5 336	9.5
16	✔		low	✔		197.8	4	20 033	17.5
Total	8	8	3 speeds	8	10		59	163 871	

E. EVALUATION METRICS

As previously mentioned, the user will assess the performance of the indoor localization system based on different metrics. In most practical applications, this will be the distance between the predicted location and the real location. We propose to build a scoring mechanism like presented in the EVARILOS benchmarking handbook [30], based on the user's desiderata. The metrics that we will use in the evaluation of the different datasets and post-processing scenarios are based on the requirements of the different use cases presented before. As a first metric, we will use the **absolute error (AE)**: the Euclidean distance between the ground-truth (MOCAP) position and the approximation position. While the MAE provides an average performance of the software combinations, we also include percentiles that provide additional information on the robustness of the chosen software combination by showing the distribution of the outliers. As a second metric, we use the **spatial error (SE)** which evaluates the path of the localization. Similar metrics have been proposed earlier [42]. Some software algorithms such as smoothing filters consider previously estimated points into account for improving the accuracy and thereby inherently introduce latency in the system. To calculate the time-compensated spatial error (SE), we determine the Euclidean distance from the calculated location to the closest point on the traveled trajectory. A third considered metric during the evaluation is the **power consumption** and corresponding battery lifetime of the devices. The fourth and last metric is the **cost** of the system for both deployment and execution. This mostly depends on the number of anchors, the need for wired synchronization between the anchors, and data throughput from anchors to the localization engine.

F. EVALUATION PROCEDURE

The benchmark platform is built upon Node-RED [43], a low-code programming tool for event-driven applications. The platform utilizes a standardized format to relay data between different algorithms. The specific algorithm's settings are centrally orchestrated to be able to run different versions of the same algorithm to evaluate the influence of certain settings. We will keep this benchmarking process manageable by only evaluating different combinations of algorithms and not differentiating between the settings of each specific algorithm.

V. EVALUATED ALGORITHMS

Many different optimizations can be added and combined to create a multi-dimensional analysis. In this section, all the evaluated state-of-the-art algorithms are introduced, together with their configuration parameters. Although the used parameters are mentioned for completeness, the parameters are defined based on previous work and are not optimized for these datasets in particular. This allows for validating the generalization of the results in different environments. Table 3 provides the available information on the results, including a reference to a more detailed description, the reported improvement, and how the evaluation was executed.

All state-of-the-art algorithms are listed in Table 4 with their specific input data, output data, and targeted UWB technique. As discussed, the algorithms will be evaluated both individually as well as when combined with each other. The implementations of the different evaluated modules were made compatible with each other. The possible combinations of algorithms are shown in Fig. 5. The bias correction and the machine learning can both be used on the TWR dataset after each other but only one positioning module can be

TABLE 3. State-of-the-art algorithms used for evaluation.

Paper	Description	Reported improvements	UWB	Static evaluation	Dynamic evaluation	New environment	NLOS	module
[22]	ML error mitigation	MAE: 29%	✓	✓	✗	✗	✓	machine learning TDoA
[22]	anchor selection TDoA	MAE: 9%	✓	✓	✗	✗	✓	anchor selection TDoA
[27]	received power correction	Not reported	✓	✓	✗	✓	✗	bias correction
[44]	received power correction	bias up to 10 cm	✓	✓	✗	✓	✗	bias correction
[28]	ML error mitigation	MAE ranging: 29%	✓	✓	✗	✗	✓	machine learning TWR
[45]	particle filter	MAE: 10%	✗	✓	✗	✗	✓	particle filter
[46]	Kalman filter	Not reported	✗	✗	✗	✗	✗	extended Kalman filter

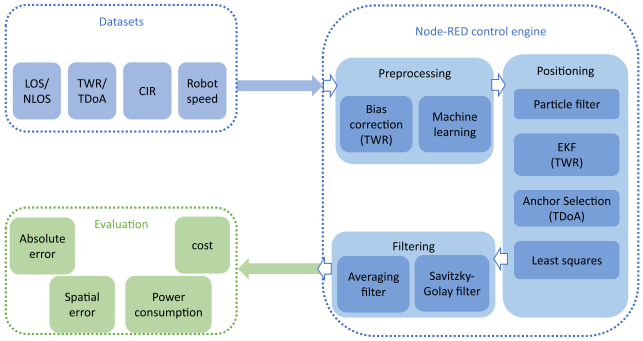


FIGURE 5. Overview of all the tested modules in the postprocessing chain. Different datasets are collected under different environmental parameters and are evaluated with different metrics.

combined to convert the timestamps into positions. In every combination of algorithms, one of the two filtering modules is used or the filtering is omitted. This means that algorithm combinations with 1 up to 4 modules are tested with in total 396 combinations of algorithms for the 16 considered datasets. The TWR datasets, three without CIR collection and five with CIR collection, have 18 and 42 combinations respectively. Similarly, the TDoA datasets, amounting to three with CIR collection and five with CIR collection, generate 9 and 21 combinations each.

A. BIAS CORRECTION (BC) (TWR)

The distance between the tag and anchor nodes is measured with asymmetric double sided two-way-ranging (ADS-TWR). Using a three-packet scheme not only cancels out the influence of clock offset but the influence of clock drift is minimized. In [44] (application note 11) the transceiver manufacturer mentions an offset on the ranging depending on the received power at the devices. This received power can be retrieved in two different ways. Firstly, an approximation is reported by the DW1000 but is only representative lower than −85 dBm. A second approach, also followed by the authors of [27], is to use the measured distance and Friis path loss formula to estimate the received signal strength. Once the received power is determined, a correction factor can be applied to the measured distance. During the first analysis of the datasets' received powers, almost 50% percent of the received powers was higher than −85 dBm and therefore we chose the second approach to estimate this received power for further analysis. For this benchmark, the relationship between received power level and bias as explained in the manufacturer's application note is used to correct the error based on the deducted received power. The same correction is done for LOS, weak NLOS, and strong NLOS links. The received power will be lower for NLOS links due to the attenuation on the direct path which will result in a different bias correction.

Parameters of this module: correction curve from the application note and Friis formula
Required input: distances

B. MACHINE LEARNING (ML)

Multipath propagation can cause the first paths to be wrongly detected, resulting in ranging errors. A popular approach in scientific literature is to use machine learning based approaches to correct the ranges. We evaluate one machine learning based approach for TWR and one for TDoA.

1) TWR

The authors of [28] proposed a machine learning based approach for mitigating the errors in a UWB system. In this paper, we implemented the same machine learning model, which was trained on data collected in the same environment as the conducted experiments. However, to pursue generalization the data was measured on random static points with random NLOS introduced and the evaluated experiments are not part of the training data. The points in the evaluated datasets are mostly unseen for the model, but the environment is the same.

2) TDoA

A similar approach can be applied for TDoA. The used model has been proposed in [22]. In this case, the convolutional neural network (CNN) requires 2 CIRs as input: one from each anchor that contributes to the TDoA. The CIRs from the links between the tag and anchors are provided to the machine learning model and allow the model to correct the TDoA when the obstruction is between the tag and one of the anchors. However, this approach cannot detect synchronization errors caused by NLOS effects between the anchor nodes.

Parameters of this module: the ML model architecture and the data trained on

Required input: distance (TWR)/ distance difference (TDOA), normalized CIR

3) LEAST-SQUARES (LS) MULTILATERATION

The timestamps collected at the anchors will resemble distances for TWR and distance differences for TDoA. This information is part of an optimization problem that can be solved with a least-squares solver.

a: TWR

Every anchor will report its distance sequentially to the localization engine (this can be the tag device or an edge node as well). Based on the measured distance, all possible points will lie on a circle (2D) or sphere (3D) with the anchor in a center. The localization engine will combine the last 4 received ranges into one position by solving the multilateration problem. This problem is defined as a minimization problem with the following cost function to minimize:

$$LE_{twr} = \sum_{k=0}^{4} \sqrt{(x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2} - d_k, \quad (1)$$

with x_k, y_k, z_k the coordinates of anchor k and d_k the distance measured with two-way-ranging (TWR). This localization error (LE) is minimized by executing an off-the-shelf least squares optimization solver based on [47].

b: TDOA

For time-difference-of-arrival (TDoA), the problem is analog to the problem in TWR except that the difference in timestamps will not result in spheres around the anchors but in hyperboloids. With one transmitted pulse, for every receiving anchor a timestamp is obtained. Equation 1 will be transformed to 2:

$$LE_{tdoa} = \sum_{k=0}^{n} \sum_{l=0}^{n} \sqrt{(x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2} - \sqrt{(x - x_l)^2 + (y - y_l)^2 + (z - z_l)^2} - tdoa_{k,l} * c, \quad (2)$$

with n the number of anchors in the system with correct reception, x_k, y_k, z_k the coordinates of anchor k, x_l, y_l, z_l the coordinates of anchor l and $tdoa_{k,l}$ the time difference measured after reception of the single UWB pulse at anchors k and l. This localization optimization problem is minimized by executing the same least squares optimization solver [47].

Parameters of this module: stop condition of the algorithms and initial guess for the optimization problem, coordinates anchors
Required input: distance (TWR)/ distance difference (TDOA)

4) PARTICLE FILTER (PF) POSITIONING

The localization engine can incorporate historical information (previous positions) in the estimation of the current position. A particle filter is a common method to do this. In this benchmark, a bootstrap particle filter similar to the one presented in [45] is used. The implementation consists of four main steps where the two techniques (TWR and TDoA) differ in the update step. Firstly, a set of particles has to be initialized a single time. We use a standard deviation of 75 cm for the normal distribution so that the initial estimate lies within a conservative and safe error margin compared to the expected NLOS error (between 20 and 60 cm). Subsequently, M particles are initialized, and a location ρ for each of them is randomly drawn from the distribution. Furthermore, each particle is assigned a speed v, direction θ, and weight ω. The motion model for the particles is chosen according to the robots' movement constraints. In this stage, all particles will have the same weight of $\frac{1}{M}$. In the second step, the new state of each particle is predicted. The new speed of each particle is randomly chosen from a normal distribution with the previous speed as the mean value and a standard deviation of 0.5 m/s within the initial bounds of 0 m/s and 1.5 m/s. Similarly, a new direction is determined. As a third step, the weights of all particles are updated and normalized. Finally, the particles are resampled based on their updated weights. particles with high weights will populate a new set of M particles more often than particles with lower weights. Additionally, a location estimate is computed in this step by taking the mean location of all particles in the new set.

a: TWR

The update of the weights of the TWR particle filter variant is done by comparing the list of distances. We look at the location of each particle and calculate the distances to the receiving anchors. Hence, a theoretical time of flight between the particle and each of the anchors can be calculated and subsequently, a similar list of time distances is acquired. With this information, we can determine the probability of a particle's predicted measurements matching the actual time measurements. The probability is calculated as a normal distribution with mean $\Delta\tau$ and standard deviation of 500 mm. The distribution to calculate the probability of the particles is $N((\Delta d\text{-}\Delta d_p),500)$ with Δd the measured distances and Δd_p the distance calculated based on the particles' position. The resulting value x_{TWR} is multiplied by the current weight of the particle [45].

b: TDOA

The weights of the TDoA particle filter variant are updated similarly. A theoretical time of flight between the particle's location and each of the anchors can be calculated and subsequently, a list of time differences $\Delta\tau_p$ is acquired. With this information, we can determine the probability of a particle's predicted measurements matching the actual time measurements. The probability is again calculated by comparing with a normal distribution with mean $\Delta\tau$ and standard deviation of 750 mm. The distribution to calculate the probability of the particles is $N((\Delta ddoa\text{-}\Delta ddoa_p),750)$ with $\Delta ddoa$ the measured distance difference of arrival

TABLE 4. Overview of the evaluated modules: their complexity and requirements.

abbreviation	module	complexity	latency	execution time	input	output
BC	bias correction	low	low	low	ranges + RX power	ranges
ML	machine learning	high	medium	high	ranges + CIR (TWR) TDoA + CIR (TDoA)	ranges + predicted error TDoA + predicted error
LS	least-squares multilateration	low	low	medium	ranges (TWR) TDoA (TDoA)	positions
LSAS	least-squares multilateration anchor selection	low	low	medium	TDoA	positions
PF	particle filter	high	low	high	ranges (TWR) / TDoA (TDoA)	positions
PF_{Input}	particle filter with machine learning input	high	low	high	ranges (TWR) + predicted error / TDoA (TDoA) + predicted error	positions
EKF	extended Kalman filter	medium	low	low	ranges	positions
AF	averaging filter	low	high	low	positions	positions
SG	Savitzky-Golay filter	medium	medium	medium	positions	positions

(ddoa) and $\Delta ddoa_p$ the ddoa calculated based on the particles' position. The resulting value x_{TDoA} is multiplied by the current weight of the particle [45].

Parameters of this module: weights, motion model, number of particles, noise parameters, coordinates of anchors
Required input: distance (TWR)/ distance difference (TDOA), ML predictions (optional)

5) EXTENDED KALMAN FILTER (EKF) (TWR)

An alternative to the particle filter that is computationally less demanding but still includes historical information and corresponding smoothing is the Kalman filter. The Kalman filter consists of two separate steps: the prediction step and the update step and tracks both the speed and position of the UWB tag therefore there are 4 parameters to predict. Each iteration of the Kalman filter starts with the prediction of position based on the current state parameters. Afterward, the extended Kalman filter (EKF) [46] updates this position based on the distance to the anchor node. The filter is an extension of the standard Kalman Filter that linearizes the model at each chosen discrete time point to apply the linear Kalman Filter equations. An EKF is often used in sensor fusion, combining multiple observation functions of different sensors to efficiently solve a non-linear state. This EKF only concerns UWB ranges so we get the following simplified implementation based on a distance measurement between tag and anchor at every observed time point. With the measurement matrix H:

$$H = \left[\frac{\bar{x} - x_1}{d} \frac{\bar{y} - y_1}{d} \frac{\bar{z} - z_1}{d} \right] \tag{3}$$

where $[x_1, y_1, z_1]$ are the coordinates of the anchor involved in the TWR measurement, and d is the calculated distance. This leads to a new state vector $\mathbf{x}$ at every observed point in time, giving the three-dimensional position estimate $[x, y, z]$ of the tag. The extended Kalman filter is initialized at [0,0,0] where the dynamic trajectories start.

Parameters of this module: noise parameters, coordinates of anchors
Required input: distances

6) ANCHOR SELECTION (LSAS) (TDOA)

Another approach towards improving accuracy in TDoA is to assume multiple anchor nodes are within range and to select only the subset of anchor nodes that are most reliable for calculating the position of a mobile tag. As the UWB time information is not flawless, in particular in harsh environments where many NLOS links are present, it might be better to use only the timestamps with the highest quality for approximating the position. NLOS links cause some signals to propagate longer on the reflected paths and the receiving timestamp will be too high. Based on the work of [22] an anchor pair selection procedure was implemented. As a first step, the position is calculated based on all available timestamp differences in the system. Based on this approximation, the positioning will be done again on the k-closest hyperboloids to this position thus limiting the influence of the shifted hyperboloids due to NLOS effects. Prior work from [22] showed a clear improvement in cases where clock drift might be an issue in the anchor nodes. In this benchmark campaign, utmost importance has been given to clock synchronization to make sure that errors originating from this source are as low as possible. For this reason, paths between the synchronizing anchor and all other anchors are kept LOS throughout the experiment.

Parameters of this module: number of anchors, coordinates of anchors
Required input: distance differences

7) POSITION SMOOTHENING FILTERING

The outcomes of the positioning engine often contain outlier positions due to NLOS effects. In case a moving trajectory is followed with frequent position updates, different types of filters can remove outliers by smoothening the trajectory, Applying these filters introduces processing latency in real-time positioning systems, as the filter needs to wait until future data points are available in order to retroactively correct prior positions. However, for many localization systems, it is sufficient to give an accurate localization after a few seconds to the end user for analysis. The latency for

correcting the positions is determined by the window size of the filter.

a: AVERAGING

To remove Gaussian noise that is still present in the system, a rolling averaging can be calculated before the output is going to the application. This filter is commonly used but is less suited for trajectories with sharp turns or abrupt deviations since these turns tend to be corrected into more gradual ones.

Parameters of this module: window size
Required input: positions

b: SAVITZKY-GOLAY (SG)

A Savitzky-Golay filter is a digital finite impulse response smoothing filter. The Savitzky-Golay filter will fit the trajectory to a polynomial of order N inside a window with a fixed length. The final trajectory will then be determined based on this polynomial for all data points. The size of the window will impact the performance of the filter: a larger window size will result in a smoother signal, while a smaller window size will be more responsive to sharp edges in the trajectory. The data points in the window include both historical and future points.

Parameters of this module: window and degree of the polynomial function
Required input: positions

VI. RESULTS

For every dataset, we first determine the so-called base accuracy. This base accuracy is the positioning error calculated by least squares multilateration without any extra modules added compared to the mm-accurate MOCAP system used as ground truth. Based on this base accuracy per dataset we will assess the performance and improvement of the different sets and combinations of modules. Without any improvement algorithms, the base scenarios have an MAE of 18.0 cm (± 10.2 cm) over all datasets. The spatial mean error for all datasets combined is 12.5 cm (± 9.7 cm). In the next part of this paper, we will discuss the influences of both environmental aspects and algorithm choices. To start with, the environmental influences (NLOS links, speed of the robot, …) are discussed. Subsequently, the consequences of the localization technique and localization method are studied. Finally, the influence of enabling different modules in the processing chain is addressed in detail. The different influences are shown in Fig. 6. Fig. 7 highlights the influences of different conditions which will be discussed in detail in the next parts of this section. For each of the cdf-curves, all relevant data is used: both TWR and TDoA data if the modules work for both technologies. For the influence of the ML, only the datasets with CIR information are used. For the bias correction (BC) graph, only TWR data is included as this module is only suited for distance correction. Both

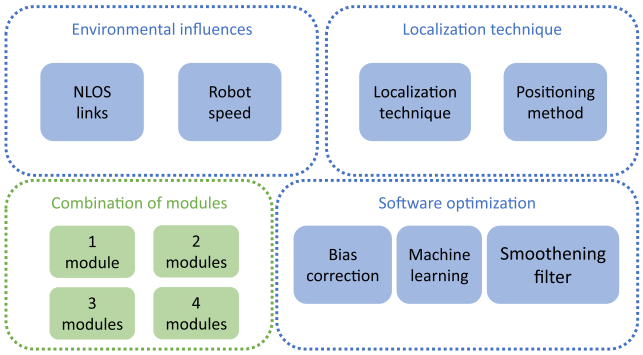


FIGURE 6. Selection of influences discussed in more detail in the results section are divided in three main categories: environment influences, influences of localization technique and positioning method and thirdly the extra software optimizations on top. The combination of modules and the number of used modules is also discussed in detail.

TABLE 5. Positioning results with NLOS links present.

	AE [cm]			SE [cm]		
	mean	std	75%	mean	std	75%
LOS	14.7	8.9	19.5	8.7	7.2	12.2
NLOS	15.6	9.8	20.7	9.4	8.1	13.4

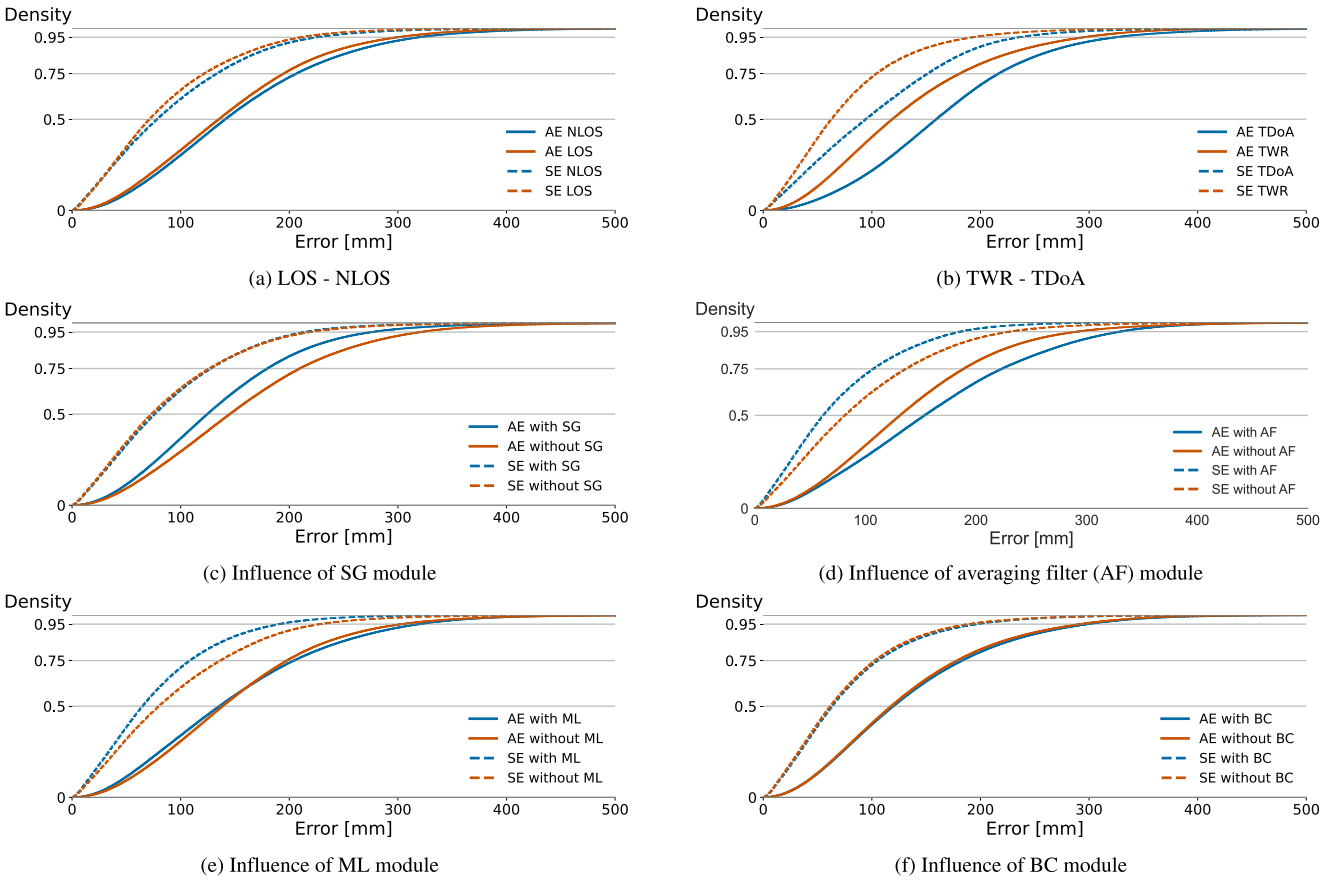
error metrics, the absolute error and spatial error which were introduced in Section IV-E are shown in the figure. The modules behave differently based on each evaluation metric for certain modules.

A. ALGORITHM RECOMMENDATIONS BASED ON ENVIRONMENTAL ASPECTS

1) INFLUENCE OF NLOS

The scenarios include NLOS by shielding several links between the tag on the robotic platform and the anchors. The NLOS shielding was designed to attenuate the direct path in such a way that a strongly attenuated signal might still arrive with an SNR in the noise floor, but may not be correctly detected. On average the ranging error increase is limited but the number of failed ranging attempts has increased. The average receive power for the NLOS scenarios is significantly lower. For all considered datasets we see an average accuracy drop of 0.9 cm in positioning accuracy. The particle filter and extended Kalman filter cope well with input highly varying input. Ideally, these positioning methods both need an estimation of the measurement error. In Table 5, both error metrics are given for LOS/NLOS situations for both TWR and TDoA.

Lesson learned 1: The presence of NLOS obstructions degrades the ranging measurements but has a significantly larger impact on TDoA than on TWR algorithms. In NLOS conditions, it is recommended to use either TWR, or to extend the TDoA with filtering methods to remove outliers.



(a) LOS - NLOS

(b) TWR - TDoA

(c) Influence of SG module

(d) Influence of averaging filter (AF) module

(e) Influence of ML module

(f) Influence of BC module

FIGURE 7. The influence of different software modules on multiple evaluation metrics: absolute error (AE) and spatial error (SE). The absolute error is the error between the estimated point and ground truth while the spatial error measures the distance between the estimated point and the traveled path. All relevant data (both TDoA and TWR) is used to make these graphs unless mentioned otherwise. (a) The presence of NLOS links results in higher errors. (b) TWR is significantly more accurate. (c) Adding a Savitzky-Golay-filter positively influences the absolute error but is similar in the performance of the spatial error. (d) The spatial error is lowered by applying the averaging filter. Still, the absolute error is higher due to the introduced latency. (e) Machine learning decreases the spatial error while the absolute error is both improving and diminishing on the data where CIR is collected. (f) Applying bias correction based on the receive level only has a limited influence on the TWR accuracy.

2) INFLUENCE OF MOBILITY PATTERN

Three different mobility patterns are investigated. Firstly, a static grid was determined to mimic low dynamic applications such as asset tracking (Fig. 4a). Hereby, the robot moves to a position and stays stationary. The second and third options include a square trajectory with slightly rounded corners run at low speed (0.1 m/s) and higher speed (0.2 m/s) respectively (Fig. 4b). The MAE for the static scenarios is lowest with 8.1 cm (± 5.3 cm) when TWR is used and 10.9 cm (± 7.3 cm) when using TDoA. The spatial error is practically the same as the absolute error for both techniques, which is expected in a static scenario. When the speed is low, the TWR accuracy is lower with average errors up to 12.2 cm (± 7.1 cm). When the robot is moving at high speed, the MAE is the highest (15.6 cm ± 9.2 cm). For the SE, in TWR the mean values are within a 0.5 cm interval, independent of the movement speed of the robot, with the highest error for the static scenario. TDoA, the high and low-speed scenarios show similar accuracy as the static scenario. The detailed characteristics of the mobility pattern when using both localization techniques can be found in Table 6.

Lessons learned 2: The AE increases with the increasing speed of the tag, with stationary tags having the smallest absolute errors. However, the SE of the devices remains stable, indicating that many algorithms do in fact calculate the position correctly but with additional latency. As such, the SE together with the latency might be a more meaningful metric combination for evaluating the accuracy of localization systems for mobile devices rather than using only the AE.

B. ALGORITHM RECOMMENDATION BASED ON LOCALIZATION TECHNIQUE

1) INFLUENCE OF LOCALIZATION TECHNIQUE

By executing the exact same trajectory and keeping the speed for different experiment runs fixed, but using different localization techniques, we can compare the behavior of the two-way-ranging and time-difference-of-arrival techniques. Two-way-ranging is more accurate, with a lower

TABLE 6. Positioning results for mobility patterns.

		AE [cm]			SE [cm]		
		mean	std	75%	mean	std	75%
TWR	high speed	15.6	9.2	21.3	7.6	6.0	10.3
	low speed	12.2	7.1	16.2	7.9	6.4	10.6
	static	8.1	5.3	10.4	8.0	5.3	10.3
TDoA	high speed	19.6	12.0	24.9	10.5	10.2	15.1
	low speed	16.6	8.9	21.0	10.7	8.3	15.2
	static	10.9	7.3	15.7	10.8	7.2	15.6

standard deviation than the time-difference-of-arrival variant. However, faster positioning update rates are achieved in time-difference-of-arrival schemes, as the number of packets required for localization is independent of the number of anchors used in the system. Instead, it is solely dependent on the number of transmissions by one or more tags. Every transmitted tag packet will result in a position in TDoA. The benchmark shows that TWR is much more robust in both absolute (13.5 ± 8.6 cm) and spatial (7.2 ± 6.2 cm) metrics compared to the higher absolute (18.6 ± 11.5 cm) and spatial (12.5 ± 10.8 cm) TDoA errors. In Table 7 both localization techniques are compared based on the four different metrics. During TWR schemes the tag clock keeps running, while in TDoA the tag enters sleep mode which reduces its power consumption. Some modules require only timing information, while others need to collect CIRs (e.g. for bias correction). The cost for the latter type is higher, as it requires a high throughput link between anchors and the central node or a sufficient amount of computing power at the anchor nodes. TDoA needs the synchronization of the anchors which also introduces an extra cost and increased complexity. In this benchmark, synchronization has been performed wirelessly, but similar conclusions can be achieved with wired synchronization. Wired synchronization increases deployment costs and is not possible for every deployment (e.g. due to safety restrictions).

Lessons learned 3: In our experiments, TWR is more accurate than TDoA due to lower constraints on clock synchronization between the anchor nodes. In contrast, the scalability and power consumption of TDoA are its biggest strengths. The choice between both techniques is dependent on the targeted application.

2) INFLUENCE OF POSITIONING METHOD

For both localization techniques, different methods are implemented that estimate the position. For both TWR and TDoA, a basic least-squares implementation without information on previous points is available. In addition, there is a particle filter and a second version of this particle filter which utilizes machine learning output as an indication of the noise in the system. An EKF is added for TWR and an anchor selection approach is available for TDoA. Firstly, the TWR results are discussed, and secondly the TDoA results.

a: TWR

For two-way-ranging the EKF is outperforming the other positioning methods (MAE: 11.1 ± 6.9 cm). The particle filter (MAE: 13.7 ± 8.0 cm) is still an improvement of the base least squares algorithm (MAE: 15.4 ± 9.3 cm). As the particle filter adds much more complexity to the system in comparison to EKF, the latter is the better choice in TWR indoor positioning systems. When using the output of the machine learning model as input for the particle filter to estimate the noise levels, the particle filter performance is reduced (15.1 ± 8.8 cm). For the spatial error, the particle filter with machine learning is the best choice with an error of 7.2 ± 5.4 cm. 0.3 cm better than its standard equivalent. The least-square approach is the worst positioning method in following the trajectory with an error of 8.3 cm.

b: TDOA

For time-difference-of-arrival, the particle filter with machine learning input is the best choice to use (MAE: 15.8 ± 10.7 cm), an improvement of 2 cm compared to the basic LS multilateration (MAE 17.7 ± 10.1 cm). The anchor selection method can be considered as it boosts the accuracy (MAE: 16.3 cm ± 10.4 cm). For the spatial error, the positioning methods show the same behavior. The detailed characteristics of the different positioning methods for both localization techniques can be found in Table 8.

Lessons learned 4: The basic multilateration methods are outperformed by EKF for TWR and PF for TDoA. These positioning methods require a sufficient update rate of the UWB system as they rely on previous positions to approximate the current position.

C. RECOMMENDATIONS PER SOFTWARE OPTIMIZATION.

1) INFLUENCE OF SMOOTHING FILTER

In this benchmark, we evaluated two different smoothing filters. An averaging filter (AF) has a positive impact on the spatial error (7.6 ± 5.9 cm compared to 10.4 ± 9.1 cm when no filter is applied), but the MAE of the system has increased from 15.2 to 16.5 cm. The main reason is that the average filter lags behind in position updates as it needs older positions to approximate the current position. As such, the faster the update rate of the tag, the lower the overall introduced latency will be, making the filter best suited for systems with high update rates and/or slow movements. However, the latency disadvantage is negated when the visual trajectory is only needed for non-real-time post-event analysis. The second considered filter is the Savitzky-Golay filter. This filter also introduces processing latency as it waits for future points to correct the position of the current point, but the estimated positions are closer to their true value. For real-time processing, the window can be kept rather low to limit the processing latency and take sharper corners into account. The SG filter increases the performance with 1.5 cm and 1.3 cm for the AE and SE respectively. Finally,

TABLE 7. Two-way-ranging is the most accurate technique but requires higher energy consumption. Depending on the data that was collected and the used technique, different combinations of modules are best.

technique	tag energy consumption			Deployment	update	LOS			NLOS			best	best
	#TX	#RX	power consumption	costs	rate	MAE	std	75th AE	MAE	std	75th AE	modules (AE)	modules (SE)
TWR without CIR	1.25	1	high	$	17.4	11.4 cm	7.5 cm	14.81 cm	12.6 cm	7.9 cm	16.8 cm	EKF SG	BC LS AF
TWR with CIR	1.25	1	high	$$	7.4	13.7 cm	8.3 cm	18.4 cm	16.4 cm	9.3 cm	22.2 cm	EKF ML	LS ML AF
TDoA without CIR	1	0	low	$$	25.7	17.1 cm	8.7 cm	20.7 cm	17.3 cm	11.0 cm	22.1 cm	PF SG	PF AF
TDoA with CIR	1	0	low	$$$	9.9	16.3 cm	9.7 cm	21.7 cm	17.5 cm	10.0 cm	22.5 cm	LS ML SG	LS ML AF

TABLE 8. Positioning methods for both localization techniques.

		AE [cm]			SE [cm]		
		mean	std	75%	mean	std	75%
TWR	EKF	11.1	6.9	14.7	7.6	4.9	10.0
	LS	15.4	9.3	20.5	8.3	7.5	11.3
	PF	13.7	8.0	18.3	7.5	6.0	10.3
	PF_{Input}	15.1	8.8	20.3	7.2	5.4	10.1
TDoA	LS	17.7	10.1	22.3	11.5	9.5	16.5
	LSAS	16.3	10.4	20.7	10.2	9.1	13.9
	PF	17.3	9.1	21.9	10.8	7.8	15.7
	PF_{Input}	15.8	10.7	21.3	8.2	7.6	11.7

TABLE 9. Accuracy of the system with and without smoothening filters. The evaluated AF filter has a larger window size than the evaluated SG filter, thereby mainly improving the SE rather than the MAE.

	AE [cm]			SE [cm]		
	mean	std	75%	mean	std	75%
No filter	15.2	9.9	19.8	10.4	9.1	14.6
AF	16.5	9.3	22.4	7.6	5.9	10.8
SG	13.7	8.4	18.0	9.1	7.3	12.8

it is possible to conclude that if no filter is used, the SE (10.4 ± 9.1 cm for all datasets) is higher than when filters are used while for the AE it depends on the chosen filter and introduced latency. The results with the two filters and without any positioning smoothing filter is summarized in Table 9. It is noteworthy that the traveled trajectory will also influence the performance of the filtering methods. For example, a circle will be smoothed by the AF to its center point. To limit the influence of this behavior, a realistic path was chosen with sharp corners and straight lines in between. Furthermore, the window of the filter is also kept small in this benchmark, to limit the influence in dynamic situations. The detailed characteristics for the positioning filters can be found in Section V-B7.

Lessons learned 5: Adding a smoothening filter increases positioning accuracy at the cost of additional latency. Smaller window sizes (such as the one in the evaluated SG module) mainly improve the MAE, whereas longer window sizes mainly improve the SE at the cost of higher latency. Ideally, the window size of a smoothening filter should be adapted dynamically, depending on the positioning update rate, the mobile tag speed, the sharpness of the trajectory, the maximum latency, and the preferred metric to optimize (AE versus SE).

2) INFLUENCE OF CIR-BASED MACHINE LEARNING

Many recent scientific papers [18], [27], and [28] employ ML-based error mitigation techniques that are based on the collected CIR values. 300 samples are gathered for each Channel Impulse Response (CIR), and within each sample, there are 2 bytes allocated for the real part and 2 bytes for the imaginary part. To this end, for each packet CIR 300*2*2 bytes of data need to be collected, which negatively impacts the maximum update rate of the positioning system. In addition, the post-processing chain needs more computational power for the inference of the pre-trained network. It is important to note that UWB error mitigation models are environment-specific and need to be retrained for optimal results [48]. For our evaluation, the neural network is trained on data that was captured in the same lab but doesn't originate from the evaluated datasets. For the evaluation of the machine learning error mitigation, only the datasets including CIRs are compared. Both AE and SE show a very clear improvement in performance in the TDoA datasets, with improvements of 2.6 cm and 2.4 cm respectively. When using TWR, the machine learning models could not improve the results and even degraded the localization accuracy. The added complexity is therefore often not worth implementation in TWR systems. The mean, standard deviation, and 75th percentile of the absolute en spatial error for the same scenarios with and without ML for both localization techniques is given in Table 10.

Lessons learned 6: Adding machine learning models can boost the performance of the localization, but most scientific papers neglect to report on the large negative impact of CIR-based error mitigation on the update rate and required energy consumption. Caution should be taken to activate CIR-based error mitigation only when sufficient NLOS effects are present in the system, and even then the reduced positioning update rate might negate some of the error mitigation benefits. Machine learning models should be retrained with data from the same environment for optimal performance.

3) INFLUENCE OF THE BIAS CORRECTION (TWR)

The Decawave application note suggests to use a received signal strength-based bias correction approach, as outlined

TABLE 10. Influence of machine learning on the results.

		AE			SE		
		mean	std	75%	mean	std	75%
TWR	with ML	14.6	8.9	19.7	7.6	5.7	10.3
	without ML	13.8	8.2	18.6	7.3	5.3	10.0
TDoA	with ML	15.7	10.0	21.2	8.5	7.3	11.9
	without ML	18.3	9.3	23.0	10.9	8.1	15.9

TABLE 11. Influence of received signal strength-based bias correction.

	AE			SE		
	mean	std	75%	mean	std	75%
TWR with BC	13.7	8.5	18.3	7.9	6.3	10.6
TWR without BC	13.4	8.3	17.9	7.6	6.1	10.3

in Section V-A. We notice a marginally higher MAE when the bias correction is applied (0.3 cm). For the spatial error, a marginally lower error of 0.4 cm could be observed. Although there is no significant influence on the positioning accuracy, we still saw an increase in the accuracy of the distance measurements of 2 cm. The error statistics of all combinations with and without bias correction are given in Table 11.

Lessons learned 7: The influence of received signal strength-based bias correction is minor. Despite its simplicity, such an approach needs specific calibration on the used hardware and antennas to result in an optimal performance improvement.

D. EFFECTS OF COMBINING MULTIPLE SOFTWARE ALGORITHMS

Although individual modules can positively or negatively impact evaluation metrics, it is unclear from the scientific literature to what extent these benefits are additive when combining multiple software modules. The influence of using only one basic positioning module compared to adding different modules can be seen in Fig. 8 for four different scenarios. For both TWR and TDoA, scenarios with and without NLOS are shown and all possible post-processing flows are considered and grouped based on the number of activated modules. When TWR is used to determine the position, at least one module is required to calculate the position from the distances. A maximum of four modules can approximate the final position. In TDoA, there is no bias correction variant and a maximum of three modules can be used sequentially. While in TDoA a lower error is observed when more modules are combined, the influence is rather limited in TWR cases, sometimes even degrading the MAE positioning accuracy. As such, the potential for improvements by combining multiple modules is larger for TDoA than for TWR. Moreover, the performance gains when combining multiple modules for TWR are higher in NLOS conditions

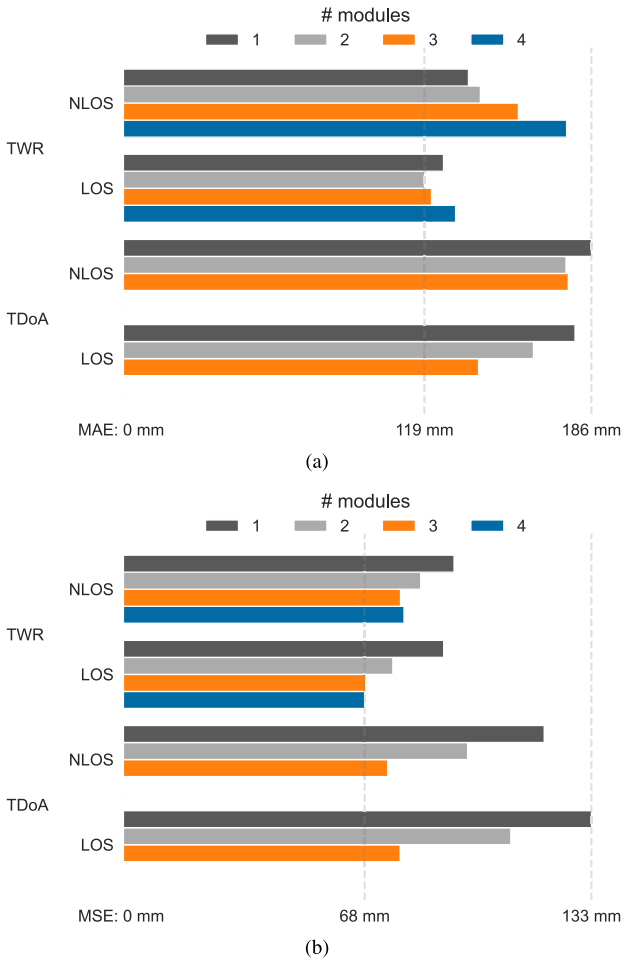


FIGURE 8. a) The mean absolute errors on the final positions for all scenarios where 1 up to 4 modules were applied on the UWB measurements. For TDoA it is beneficial to add more modules to the system as it will be more accurate on average. For TWR the system is already accurate and adding more modules will have a negligible or sometimes negative impact. b) Considering the spatial error, using more modules is improving the localization accuracy for both localization techniques.

(due to the lower base accuracy). In contrast, for TDoA even LOS conditions stand to benefit from combining multiple modules.

The relative improvement of all tested module combinations compared to the base post-processing pipeline (only theLS module to convert timestamps into positions) is shown in Fig. 9. The darkness of the bars shows the number of datasets that agree on the relative improvement, averaged out over the different mobility scenarios and both LOS and NLOS scenarios, thus representing diverse unknown conditions. For such diverse conditions, for TWR (Fig. 9a) several combinations result in worse performance for both metrics. It is noteworthy, that all combinations with the averaging filter increase the absolute error but decrease the spatial error. For TDoA (Fig. 9b), almost all modules and module combinations decrease the spatial error, but

several combinations increase the MAE. Notably, the anchor selection variant for TDoA seems to improve the performance sometimes but has a negative impact in multiple other scenarios.

Lessons learned 8: The relative benefits of multiple combined algorithmic modules can not easily be predicted based on the performance of individual modules, indicating hidden interactions between algorithmic improvements. For TWR, combining multiple modules can even degrade the performance, whereas due to its lower initial base accuracy, combining multiple modules more often positively benefits the accuracy of TDoA systems. Additionally, it is important to evaluate the combination with the suitable metric.

E. BENCHMARKS INSIGHTS DISCUSSION

From the previous results, we can derive some high-level conclusions on the benchmark process. Firstly, the results of the benchmark campaign are highly dependent on the used metric. Where some of the filters wouldn't be considered when using the MAE, for example, the averaging filter, it gives excellent results in the spatial domain and visually appealing curves. In addition, the targeted application will determine the requirements, a small low-power tag can have a low update rate and thus difficulty in using historical information. For high-update rate applications, historical information can be taken into account. The amount of historical data that can be taken into account (the last 100 ms, or the last second) depends on the maximal speed of the robot and the update rate of the UWB system. Taking more samples makes it easier to limit noise in the system, but introduces inaccuracies in itself.

Lessons learned 9: The reported accuracy from existing state-of-the-art scientific papers can not directly be relied upon for making UWB system design decisions for scenarios that are different in terms of environment, setup, or requirements. The need for benchmarking and unified testing datasets is high. There is a clear need for environment and experiment descriptions driven by the application domain.

VII. USE CASES BEST PRACTICES

Although every use case will differ in terms of requirements, in this section we will discuss requirements and guidelines for five typical generic use cases: asset tracking, crowd monitoring, geofencing, autonomous ground robots, and autonomous aerial robots. First, these use cases are introduced, and then we provide findings from the benchmark experiments that are specifically relevant to these datasets. The complete overview of five considered use cases is given in Fig. 10 and Table 12 which also summarizes the best module combinations to use.

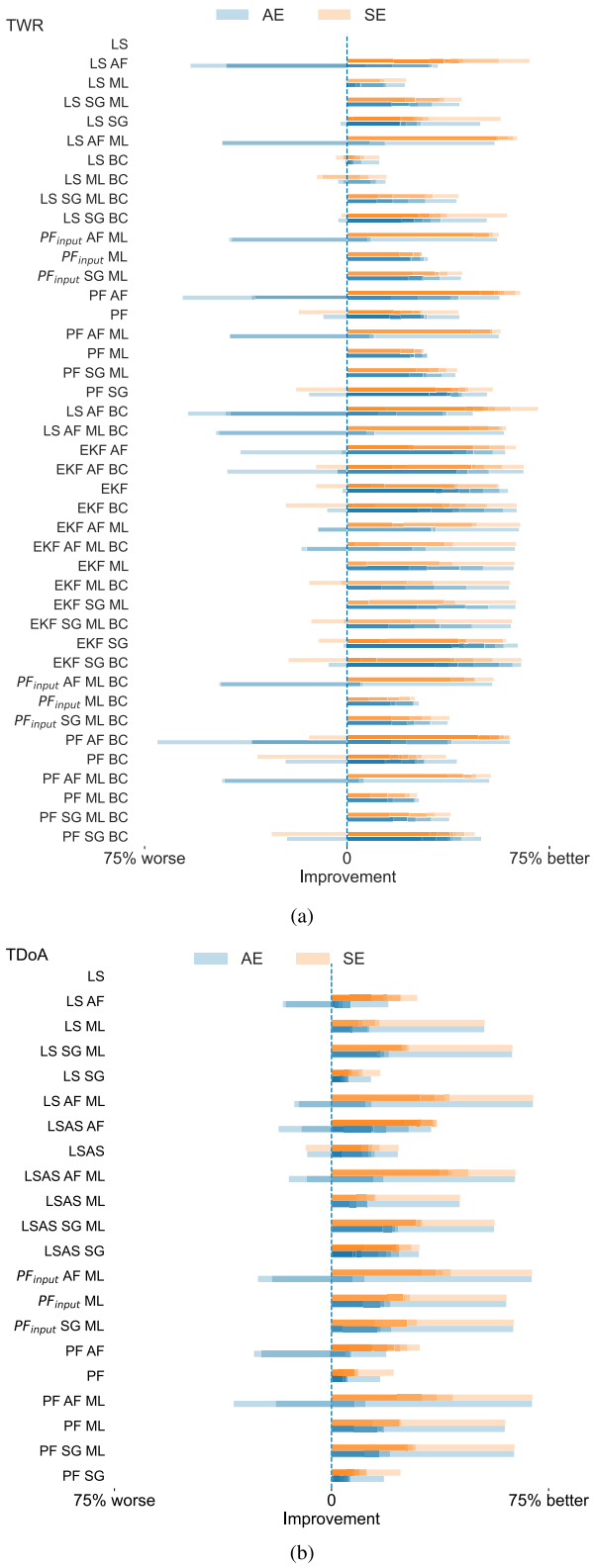


FIGURE 9. Relative increase of the performance compared to the base accuracy of the LS algorithm for different module combinations over all datasets of the MAE and SE for a) TWR and b) TDoA datasets.

A. ASSET TRACKING

Description A high number of tags with limited battery power is tracked. The number of UWB packets sent and received from the mobile nodes is limited and their energy consumption requires scheduling optimizations while going

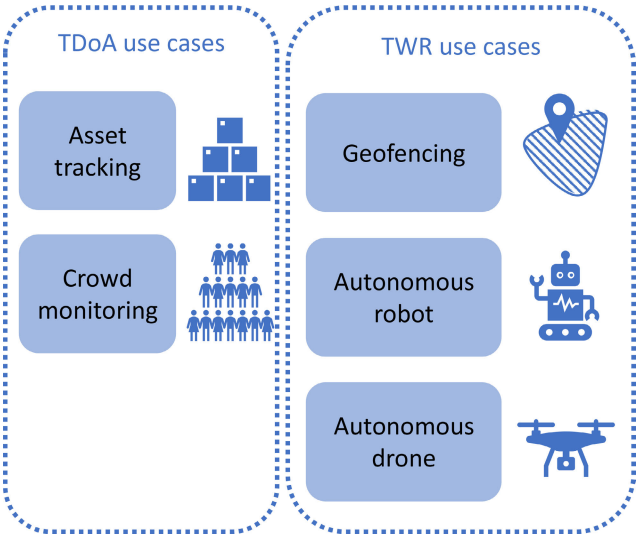


FIGURE 10. Different use cases have specific requirements and different localization techniques.

maximally in sleep or deep sleep mode. Assets are mostly static, but should also be tracked when moving around the localization zone.

Datasets The datasets that are representative of this use case use TDoA in LOS and NLOS environments. The static measurements are in particular relevant as asset tracking update rates are low.

Findings The asset tracking use case is typically power-constrained with lots of battery-powered tags to track. A position update will only be required once every hour/day. For this use case we suggest using (uplink) TDoA where the power consumption of the tag is limited to the transmission of a single UWB packet and therefore negligible to the power consumed during sleep mode. Machine learning can be added at the fixed anchor nodes to improve localization with the LS multilateration approach. Using a particle filter is not suggested as historical information will be of low value. To improve static localization, building on the results of the averaging filter, it could be possible to locate the nodes multiple times during a small time interval and take the average position. After this average position is calculated, the node can go back to sleep for hours.

B. CROWD MONITORING

Description Large crowds can be tracked when the required transmission time per position is low and many tags can transmit their packets each second.

Datasets The crowd monitoring use cases use TDoA for a high scalability. All used datasets are dynamic in nature with both NLOS and LOS at high and low speeds.

Findings Previously, it has been shown that TDoA is more scalable towards multiple users because it requires fewer UWB packets for a location update [49]. The movement of people demands faster update rates than the asset tracking and therefore positioning with historical information (particle filter) and filtering the positions (averaging filter or Savitzky-Golay filter with larger window sizes) are a good choice. Traditionally, a latency of a few seconds in processing delay for the SG is not a problem and the filter can be considered beneficial to use in this use case.

C. GEOFENCING

Description In geofencing applications a trigger event will be thrown based on the location of the tag and a predefined zone. This can be used to boost the efficiency of industrial processes or to make the environment safer. Precise location information is required for the tag to minimize both false positives and false negatives. The system should provide a natural responsiveness but the application is not time-sensitive.

Datasets Two datasets are selected for this use case, the static TWR and NLOS low-speed scenario. Both scenarios cover the use case requirements of high accuracy and low update rate to limit power consumption.

Findings High accuracy is required to determine if an object is inside the borders of the search area. For this purpose TWR is used. To reduce the number of false positives and false negatives filtering can be applied to remove noise from the positions, but the latency introduced is unfavorable for the response time.

D. AUTONOMOUS GROUND ROBOT

Description For an autonomous ground robot, accurate localization is needed so the robot can find its path throughout the localization zone.

Datasets The autonomous ground robot needs the high accuracy and stability of TWR. In all considered datasets CIR information is collected to improve positioning, and both low and high-speed scenarios are considered in LOS and NLOS environments.

Findings An autonomous ground robot needs its position accurately to operate safely in different scenarios and track its estimated path. The highly accurate TWR calculated with an EKF or PF and averaging filter with a small window size is recommended.

E. AUTONOMOUS AERIAL ROBOT

Description To support fast-moving autonomous aerial robots, a higher update rate is needed next to accurate localization. The collection of CIRs will not be feasible to maintain a high update rate. The ML based error mitigation can not be added in this use case and other positioning methods and filters need to be considered.

Datasets Two datasets are selected for the autonomous aerial robot use case: high speed TWR without CIR collection in both LOS and NLOS environments.

TABLE 12. The use cases that were highlighted in the paper and the proposed software optimizations based on the benchmark campaign.

use case	technology	base		best optimizations			
		AE	SE	best AE	AE [cm]	best SE	SE [cm]
asset tracking	TDoA	19.2	15.3	LS AF ML	5.6	LS AF ML	5.5
crowd monitoring	TDoA	19.1	12.4	LSAS SG ML	15.1	LSAS AF ML	6.9
geofencing	TWR	17.1	13.4	EKF AF ML	5.2	EKF AF ML	5.1
autonomous ground robot	TWR	17.8	10.7	EKF SG	9.5	LS AF ML	4.2
autonomous aerial robot	TWR	16.4	10.4	EKF SG	10.1	LS AF EC	5.3

Findings The drone variant is more demanding. The update rate needs to be high with as little latency as possible. To limit the latency, we suggest an EKF executed at the drone itself without filtering and without machine learning.

VIII. FUTURE WORK

This work has taken the first steps to compare different software post-processing steps into design knowledge for the indoor localization system designer, thus opening up new possibilities for innovative research and expansions. While utmost consideration is taken for designing the benchmark procedure, the methodology still possesses certain limitations. The main limitations are 1) the used trajectories, 2) the benchmark environment, 3) the firmware design choice, and 4) the set of algorithms tested. Although these limitations still exist, the outcomes found in this paper are still relevant. For the first limitation, the selection of the trajectories influences the results as different algorithms show different resilience against sharp corners in the trajectory. In this paper, only a single trajectory is used, albeit at two different robot speeds. In future experimentation, the influence could be investigated. The second limitation considers the environment. In this benchmark, the environment is changed by placing absorbers in the room to objectively compare LOS and NLOS inside the same room. The benchmark could be repeated in new, more complex environments. The third identified limitation is the firmware design choices which contain the choices made by the embedded developer inside the MAC and physical layer. For example, reply times between UWB messages influence the update rate and performance of the system. To generalize the results of this benchmark, different hardware systems and firmware implementations should be tested. Finally, in this paper, the algorithms were selected based on availability and relevance for general use cases. With the proposed benchmark approach and the availability of the dataset, new algorithmic combinations could be tested objectively for both TWR and TDoA.

In addition to these mentioned limitations, extensions of the current work are identified: the analysis can be extended toward existing open-source datasets (e.g. the IPIN competition dataset). This will increase the generalization of the results toward a self-adaptive dynamic localization system and seamless localization. To obtain this goal, the system will not only have to compare different post-processing strategies, but in addition, dynamically adapt itself for better performance. To do so the system can base its decision on the known environmental conditions (materials and attenuation in the indoor localization area), application requirements (update rate, latency, and scalability), and available infrastructure (computational power, power supply, etc.). A final open question remains how to design a confident single quality metric for indoor localization to make the modules able to weigh the most important information.

IX. CONCLUSION

When comparing UWB software components, it is often difficult to interpret the results correctly as the data and environment are different. In this paper, over 16 different datasets were gathered in the same industrial environment and each was evaluated with up to 42 different combinations of software post-processing components. The platform's uniqueness lies in its ability to simulate diverse conditions, encompassing various mobility scenarios and line-of-sight conditions. This work contributes to the advancement of UWB localization research by offering a standardized framework for testing and comparing algorithms. By exposing the algorithms to a wide array of scenarios, the platform better reflects real-world challenges, fostering robust algorithmic development.

The evaluation shows that TWR achieves higher performance than TDoA, while TDoA is beneficial in the scalability and power consumption of the tag node. Taking harsh environments into account, half of the datasets were collected with introduced NLOS links which decreases the accuracy of the localization systems for both TWR and TDoA. The use of an extended Kalman filter or a particle filter can increase robustness in these situations. Based on our analysis, we can draw the conclusion that making a carefully considered selection of software optimizations to implement in the system is frequently more advantageous than deploying all available optimizations. This decision should be tailored to the specific use case and localization technique. In our examination of five distinct UWB-targeted use cases, each with its unique set of requirements, we have presented the optimal combinations of software modules.

We evaluated the performance of different post-processing algorithms based on the absolute error (AE) and spatial error (SE). The AE is more oriented towards real-time

low latency systems where decisions are taken immediately and a trailing bias is also punished by the error metric. The SE checks an estimated position with the followed ground truth trajectory to have an error metric independent of latency and bias. This SE is in some use cases more helpful than the AE, e.g. for long-term analysis of movement patterns. Applications that don't require real-time updates for decision-making and allow a higher latency (>0.5 s) should prefer this metric over MAE. Finally, in this paper, we focused on the practical steps for performing an UWB benchmark and present our conclusions based on different metrics. The insights derived from our experiments offer valuable guidelines for algorithm selection in various real-world scenarios, fostering the broader adoption of UWB localization across industries and applications. We hope this work can help people rethink indoor localization optimization and foster an increased deployment of indoor localization systems in the real world.

REFERENCES

[1] K. Szyc, M. Nikodem, and M. Zdunek, "Bluetooth low energy indoor localization for large industrial areas and limited infrastructure," *Ad Hoc Netw.*, vol. 139, Feb. 2023, Art. no. 103024.

[2] L. Zhang, L. Huang, Q. Yi, X. Wang, D. Zhang, and G. Zhang, "Positioning method of pedestrian dead reckoning based on human activity recognition assistance," in *Proc. IEEE 12th Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2022, pp. 1–8.

[3] K. Gao, H. Wang, H. Lv, and W. Liu, "Toward 5G NR high-precision indoor positioning via channel frequency response: A new paradigm and dataset generation method," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2233–2247, Jul. 2022.

[4] M. Menolotto, D.-S. Komaris, S. Tedesco, B. O'Flynn, and M. Walsh, "Motion capture technology in industrial applications: A systematic review," *Sensors*, vol. 20, no. 19, p. 5687, Oct. 2020.

[5] Y. Xu, R. Zheng, S. Zhang, and M. Liu, "Confidence-rich localization and mapping based on particle filter for robotic exploration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2022, pp. 4471–4477.

[6] B. Van Herbruggen, B. Jooris, J. Rossey, M. Ridolfi, N. Macoir, Q. Van den Brande, S. Lemey, and E. De Poorter, "Wi-PoS: A low-cost, open source ultra-wideband (UWB) hardware platform with long range sub-GHz backbone," *Sensors*, vol. 19, no. 7, p. 1548, Mar. 2019.

[7] M. Elsanhoury, P. Mäkelä, J. Koljonen, P. Välisuo, A. Shamsuzzoha, T. Mantere, M. Elmusrati, and H. Kuusniemi, "Precision positioning for smart logistics using ultra-wideband technology-based indoor navigation: A review," *IEEE Access*, vol. 10, pp. 44413–44445, 2022.

[8] D. Coppens, A. Shahid, S. Lemey, B. Van Herbruggen, C. Marshall, and E. De Poorter, "An overview of UWB standards and organizations (IEEE 802.15.4, FiRa, Apple): Interoperability aspects and future research directions," *IEEE Access*, vol. 10, pp. 70219–70241, 2022.

[9] M. Schuh, H. Brunner, M. Stocker, M. Schuß, C. A. Boano, and K. Römer, "First steps in benchmarking the performance of heterogeneous ultra-wideband platforms," in *Proc. Workshop Benchmarking Cyber-Phys. Syst. Internet Things (CPS-IoTBench)*, May 2022, pp. 34–39.

[10] S. Shyam, S. Juliet, and K. Ezra, "A UWB based indoor asset tracking architecture for industry 4.0," in *Proc. 4th Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Jan. 2022, pp. 501–506.

[11] A. Bastida-Castillo, C. Gómez-Carmona, E. De la Cruz-Sánchez, X. Reche-Royo, S. Ibáñez, and J. Pino Ortega, "Accuracy and inter-unit reliability of ultra-wide-band tracking system in indoor exercise," *Appl. Sci.*, vol. 9, no. 5, p. 939, Mar. 2019.

[12] R. Vleugels, B. Van Herbruggen, J. Fontaine, and E. De Poorter, "Ultra-wideband indoor positioning and IMU-based activity recognition for ice hockey analytics," *Sensors*, vol. 21, no. 14, p. 4650, Jul. 2021.

[13] M. Rico-González, A. Los Arcos, F. Y. Nakamura, P. Gantois, and J. Pino-Ortega, "A comparison between UWB and GPS devices in the measurement of external load and collective tactical behaviour variables during a professional official match," *Int. J. Perform. Anal. Sport*, vol. 20, no. 6, pp. 994–1002, Nov. 2020.

[14] M. Cimdins, S. O. Schmidt, P. Bartmann, and H. Hellbrück, "Exploiting ultra-wideband channel impulse responses for device-free localization," *Sensors*, vol. 22, no. 16, p. 6255, Aug. 2022.

[15] Z. Arjmandi, J. Kang, K. Park, and G. Sohn, "Benchmark dataset of ultra-wideband radio based UAV positioning," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, 2020, pp. 1–8.

[16] J. Tiemann, F. Schweikowski, and C. Wietfeld, "Design of an UWB indoor-positioning system for UAV navigation in GNSS-denied environments," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, 2015, pp. 1–7.

[17] J. Fontaine, B. Van Herbruggen, A. Shahid, S. Kram, M. Stahlke, and E. De Poorter, "Ultra wideband (UWB) localization using active CIR-based fingerprinting," *IEEE Commun. Lett.*, vol. 27, no. 5, pp. 1322–1326, May 2023.

[18] S. Kram, M. Stahlke, T. Feigl, J. Seitz, and J. Thielecke, "UWB channel impulse responses for positioning in complex environments: A detailed feature analysis," *Sensors*, vol. 19, no. 24, p. 5547, Dec. 2019.

[19] G. Caso, M. Le, L. De Nardis, and M.-G. Di Benedetto, "Performance comparison of WiFi and UWB fingerprinting indoor positioning systems," *Technologies*, vol. 6, no. 1, p. 14, Jan. 2018.

[20] J. Luo and H. Gao, "Deep belief networks for fingerprinting indoor localization using ultrawideband technology," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 1, Jan. 2016, Art. no. 5840916.

[21] I. Dotlic, A. Connell, H. Ma, J. Clancy, and M. McLaughlin, "Angle of arrival estimation using decawave DW1000 integrated circuits," in *Proc. 14th Workshop Positioning, Navigat. Commun. (WPNC)*, 2017, pp. 1–6.

[22] B. Van Herbruggen, J. Fontaine, and E. De Poorter, "Anchor pair selection for error correction in time difference of arrival (TDoA) ultra wideband (UWB) positioning systems," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, 2021, pp. 1–8.

[23] S. Djosic, I. Stojanovic, M. Jovanovic, T. Nikolic, and G. L. Djordjevic, "Fingerprinting-assisted uwb-based localization technique for complex indoor environments," *Exp. Syst. Appl.*, vol. 167, Jul. 2021, Art. no. 114188.

[24] A. Ledergerber, M. Hamer, and R. D'Andrea, "A robot self-localization system using one-way ultra-wideband communication," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 3131–3137.

[25] L. R. Mendoza and K. O'Keefe, "Periodic extended Kalman filter to estimate rowing motion indoors using a wearable ultra-wideband ranging positioning system," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2021, pp. 1–8.

[26] C. Jiang, J. Shen, S. Chen, Y. Chen, D. Liu, and Y. Bo, "UWB NLOS/LOS classification using deep learning method," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2226–2230, Oct. 2020.

[27] M. Stocker, M. Gallacher, C. A. Boano, and K. Römer, "Performance of support vector regression in correcting UWB ranging measurements under LOS/NLOS conditions," in *Proc. Workshop Benchmarking Cyber Phys. Syst. Internet Things*, 2021, pp. 6–11.

[28] J. Fontaine, M. Ridolfi, B. Van Herbruggen, A. Shahid, and E. De Poorter, "Edge inference for UWB ranging error correction using autoencoders," *IEEE Access*, vol. 8, pp. 139143–139155, 2020.

[29] M. Heydariaan, H. Mohammadmoradi, and O. Gnawali, "Toward standard non-line-of-sight benchmarking of ultra-wideband radio-based localization," in *Proc. IEEE Workshop Benchmarking Cyber-Physical Netw. Syst. (CPSBench)*, Apr. 2018, pp. 19–24.

[30] T. Van Haute, E. De Poorter, J. Rossey, I. Moerman, V. Handziski, A. Behboodi, F. Lemic, A. Wolisz, N. Wirström, and T. Voigt, "The Evarilos benchmarking handbook: Evaluation of RF-based indoor localization solutions," in *Proc. 2nd Int. Workshop Measurement-based Experim. Res., Methodology Tools*, 2013, pp. 1–13.

[31] F. Potortì, S. Park, A. Jiménez Ruiz, P. Barsocchi, M. Girolami, A. Crivello, S. Lee, J. Lim, J. Torres-Sospedra, F. Seco, R. Montoliu, G. Mendoza-Silva, M. Pérez Rubio, C. Losada-Gutiérrez, F. Espinosa, and J. Macias-Guarasa, "Comparing the performance of indoor localization systems through the EvAAL framework," *Sensors*, vol. 17, no. 10, p. 2327, Oct. 2017.

[32] D. Lymberopoulos and J. Liu, "The Microsoft indoor localization competition: Experiences and lessons learned," *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 125–140, Sep. 2017.

[33] J. P. Queralta, C. M. Almansa, F. Schiano, D. Floreano, and T. Westerlund, "UWB-based system for UAV localization in GNSS-denied environments: Characterization and dataset," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2020, pp. 4521–4528.
[34] A. Van den Bossche, R. Dalcé, N. Gonzalez, and T. Val, "LocURa: A new localisation and UWB-based ranging testbed for the Internet of Things," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2018, pp. 1–6.
[35] F. Potortì, A. Crivello, P. Barsocchi, and F. Palumbo, "Evaluation of indoor localisation systems: Comments on the ISO/IEC 18305 standard," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2018, pp. 1–7.
[36] IDLab Ghent University. *Industrial IoT Lab*. Accessed: Jun. 6, 2022. [Online]. Available: https://www.ugent.be/ea/idlab/en/research/research-infrastructure/industrial-iot-lab.htm
[37] Qualisys AB. *Qualisys Motion Capture Systems*. Accessed: Dec. 19, 2022. [Online]. Available: https://www.qualisys.com/
[38] *3650—PU Foam Based Flat Absorbers*. Dordrecht, The Netherlands: H Shielding, 2016.
[39] Decawave, "Dw1000 user manual," Decawave Ltd., Dublin, Ireland, Tech. Rep., 2017.
[40] N. Macoir, J. Bauwens, B. Jooris, B. Van Herbruggen, J. Rossey, J. Hoebeke, and E. De Poorter, "UWB localization with battery-powered wireless backbone for drone-based inventory management," *Sensors*, vol. 19, no. 3, p. 467, Jan. 2019.
[41] Open Source Robotics Foundation. *Turtlebot3*. Accessed: Dec. 19, 2022-12-19. [Online]. Available: https://www.turtlebot.com/turtlebot3/
[42] A. Mathisen, S. K. Sørensen, A. Stisen, H. Blunck, and K. Grønbæk, "A comparative analysis of indoor WiFi Positioning at a large building complex," in *Proc. Int. Conf. Indoor Position. Indoor Navigation*, 2016, pp. 1–8.
[43] The OpenJS Foundation. *Node-Red, Low-code Programming for Event-Driven Applications*. Accessed: Dec. 19, 2022. [Online]. Available: https://nodered.org/
[44] Decawave, "Sources of error in dw1000 based two-way ranging (TWR) schemes," Decawave Ltd, Dublin, Ireland, Tech. Rep., 2014.
[45] M. Aernouts, N. BniLam, N. Podevijn, D. Plets, W. Joseph, R. Berkvens, and M. Weyn, "Combining TDoA and AoA with a particle filter in an outdoor LoRaWAN network," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, Apr. 2020, pp. 1060–1069.
[46] G. Mao, S. Drake, and B. D. Anderson, "Design of an extended Kalman filter for UAV localization," in *Proc. Inf., Decis. Control*, 2007, pp. 224–229.
[47] J. J. Moré, "The Levenberg–Marquardt algorithm: Implementation and theory," in *Proc. Numer. Anal. Biennial Conf. Held Dundee*. Cham, Switzerland: Springer, Jul. 2006, pp. 105–116.
[48] J. Fontaine, F. Che, A. Shahid, B. Van Herbruggen, Q. Z. Ahmed, W. Bin Abbas, and E. De Poorter, "Transfer learning for UWB error correction and (N)LOS classification in multiple environments," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 4085–4101, Feb. 2024.
[49] M. Ridolfi, S. Van De Velde, H. Steendam, and E. De Poorter, "Analysis of the scalability of UWB indoor localization solutions for high user densities," *Sensors*, vol. 18, no. 6, p. 1875, Jun. 2018.

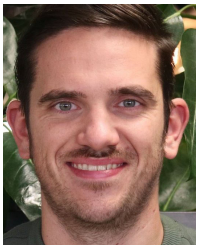
BEN VAN HERBRUGGEN was born in Antwerp, in 1995. He received the M.Sc. degree in electrical engineering from Ghent University, Belgium, in July 2018. He is currently pursuing the Ph.D. degree with the IDLab Research Group, Department of Information Technology (INTEC), Ghent University–imec. He is the author and coauthor of various publications on ultra-wideband localization solutions. His scientific work is focused on indoor localization systems based on ultra-wideband technology and the use of energy harvesters for wireless networking.

JONO VANHIE-VAN GERWEN received the M.Sc. degree in computer science engineering from Ghent University, Ghent, Belgium, in July 2009. Since August 2009, he has been a Research Engineer with the Department of Information Technology, Ghent University, where he is currently a Senior Research Engineer. His research activities have included performance analysis and benchmarking for wireless sensor networks, management, and optimization of wireless testbeds, together with enabling them for experimentally driven education. In 2018, he received a professional drone pilot license, and since he has been focused on indoor drone experimentation and automated operations.

STIJN LUCHIE was born in Veurne, in 1999. He received the M.Sc. degree in computer science engineering from Ghent University, Belgium, in September 2022. In October 2022, he joined as a Researcher with the IDLab Research Group, Department of Information Technology (INTEC), Ghent University. His primary area of scientific interest centers on ultra-wideband technology, with a specific focus on the self-calibration of anchor networks and the applications of UWB radar.

YURI DURODIÉ (Graduate Student Member, IEEE) received the Ph.D. degree from Vrije Universiteit Brussel, in April 2022, after five years of industry experience in system simulation. As part of the Brussels Human Robotic Research Center (BruBotics), his research is focused on enabling multi-agent collaboration with the use of ultra-wideband technology in combination with SLAM methods, in the absence of installed infrastructure.

BRAM VANDERBORGHT received the Ph.D. degree from Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 2007. He performed his research at the JRL Laboratory, AIST, Tsukuba, Japan, and the Postdoctoral Research with the Italian Institute of Technology, Genova, Italy. Since 2009, he has been a Professor with VUB. He had a European Research Council (ERC) starting grant and is currently coordinating three EU projects on smart and self-healing materials for soft robots. He is a Scientific Collaborator with imec, Brussels. His research interests include human–robot collaboration for applications for health and manufacturing, such as exoskeletons, prostheses, social robots, drones, and cobots.



MICHIEL AERNOUTS received the master's degree in applied engineering and the Ph.D. degree in IoT device localization with low power wide area networks from the University of Antwerp, in 2017 and 2022, respectively. During the Ph.D. degree, he worked with the IDLab-imec Research Group. After the Ph.D. degree, he became a Postdoctoral Researcher with imec, leading the Sustainable Intelligent Systems Team, and serving as a Research and Teaching Assistant. Since August 2023, he has been a Digital Forensics Engineer with the Belgian Federal Judicial Police.

ADRIAN MUNTEANU (Member, IEEE) received the M.Sc. degree in electronics and telecommunications from the Politehnica University of Bucharest, Bucharest, Romania, in 1994, the M.Sc. degree in biomedical engineering from the University of Patras, Patras, Greece, in 1996, and the Ph.D. degree (summa cum laude) in applied sciences from Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 2003.

From 2004 to 2010, he was a Postdoctoral Fellow with the Fund for Scientific Research-Flanders (FWO), Vlaanderen, Belgium. Since 2007, he has been a Professor with VUB. He is currently a Professor with the Electronics and Informatics (ETRO) Department, VUB. He is the author of more than 400 journal and conference publications, book chapters, and contributions to standards and holds eight patents in image and video coding. His research interests include image, video, and 3D graphics compression, 3D video, deep learning, distributed visual processing, error-resilient coding, and multimedia transmission over networks. He was a recipient of the 2004 BARCO-FWO prize for his Ph.D. work, and 15 other awards and scientific prizes at national and international conferences. He served as an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA and currently serves as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING.

JARON FONTAINE received the M.Sc. (Hons.) degree in information engineering technology from Ghent University, Ghent, Belgium, in 2017, and the Ph.D. degree in information engineering technology from the IDLab, Department of Information Technology (INTEC), Ghent University. His main research interests include machine learning (ML) techniques for wireless network applications (e.g., indoor localization systems using ultra-wideband, wireless technology recognition, and healthcare activity monitoring). Specifically, he focuses on optimizations of ML to run on edge and embedded devices and allowing smaller datasets to be used in new environments using transfer learning, semi-supervised learning, and data augmentation techniques. He has already published and coauthored numerous papers on these topics in journals and presented his work at conferences. He was a recipient of the FWO-SB Grant and funded by the Scientific Research Flanders (Belgium, FWO-Vlaanderen).

ELI DE POORTER is currently a Professor with the IDLab Research Group, imec and Ghent University. His team performs research on wireless communication technologies, such as (indoor) localization solutions, wireless IoT solutions, and machine learning for wireless systems. He performs both fundamental and applied research. For his fundamental research, he is currently the coordinator of several research projects (SBO, FWO, and GOA) and has more than 200 publications in international journals or the proceedings of international conferences. For his applied research, he collaborates with industry partners to transfer research results to industrial applications, as well as to solve challenging industrial research problems. He is also the Co-Founder of Lopos, a spin-off company (https://lopos.be) that offers privacy-aware UWB wearables for safety and social distancing.