DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

Benchmarking for the Indoor Localization of Autonomous Mobile Robots in Intralogistics

Benchmarking für die Indoor-Lokalisierung Autonomer Mobiler Roboter in der Intralogistik

Markus Knitt
Yousef Elgouhary
Jakob Schyga
Hendrik Rose
Philipp Braun
Jochen Kreutzfeldt

Institute for Technical Logistics
Hamburg University of Technology

This paper introduces a novel approach to benchmarking Indoor Localization Systems (ILS) for mobile robots in warehouse and manufacturing contexts. The study focuses on diverse localization technologies commonly used in mobile robotics and implements transparent and comparable performance metrics, an automated experimental procedure, as well as an intuitive performance visualization approach. Experiments were conducted using a custom-built robot equipped with various sensors, including LiDAR, Ultra-Wideband (UWB), and vision systems. A process for systematically analyzing the impact of environmental factors such as lighting, reflectivity, and obstacles on localization performance is proposed. The results provide insights into system robustness and accuracy under different conditions. The study enables more efficient experimental analysis of sensor fusion and optimization strategies for achieving optimal performance and offers a workflow to efficiently investigate sensor fusion concepts using real data.

[Keywords: Benchmarking, Localization, Robotics, Intralogistics]

In diesem Beitrag wird ein neuartiger Ansatz zum Benchmarking von Indoor-Lokalisierungssystemen (ILS) für mobile Roboter in Lager- und Produktionsumgebungen vorgestellt. Die Studie konzentriert sich auf verschiedene Lokalisierungstechnologien, die üblicherweise in der mobilen Robotik verwendet werden, und implementiert transparente und vergleichbare Leistungsmetriken, ein automatisiertes Experimentierverfahren und einen intuitiven Ansatz zur Leistungsvisualisierung. Die Experimente wurden mit einem speziell angefertigten Roboter durchgeführt, der mit verschiedenen Sensoren ausgestattet war, darunter LiDAR-, UWB- und Vision-Systeme. Es wird eine Methode vorgeschlagen, um die Auswirkungen von Umgebungsfaktoren wie Beleuchtung, Reflektivität und Hindernisse auf die Lokalisierungsleistung systematisch zu analysieren. Die Ergebnisse geben Aufschluss über die Robustheit und Genauigkeit des Systems unter verschiedenen Bedingungen. Die Studie ermöglicht eine effizientere experimentelle Analyse von Sensorfusions- und Optimierungsstrategien, um eine optimale Leistung zu erzielen, und bietet einen Arbeitsablauf für die effiziente Untersuchung von Sensorfusionskonzepten anhand realer Daten.

[Schlüsselwörter: Benchmarking, Lokalisierung, Robotik, Intralogistik]

1 INTRODUCTION

Mobile robots play a vital role in driving the ongoing transformation towards an interconnected, efficient, and flexible industry [1]. They serve as a key instrument for maintaining competitiveness in the face of rising customer demands, which are fueled by the growth of e-commerce, product individualization, and labor shortages [2]. The latest version of the *DHL Trend Radar*, published in 2022, identifies indoor mobile robots as the technology trend with the highest potential to disrupt the logistics industry [3]. This potential is expected to materialize within the next five years, as the technology of Autonomous Mobile Robots (AMRs) continues to advance, enabling their widespread deployment on a large scale. Unlike Automated Guided Vehicles (AGVs), which are attributed to basic line-following capabilities, AMRs possess advanced abilities regarding decision-making and real-time path planning [3]. The potential applications of mobile robots, including both AMRs and AGVs, in logistics are manifold, ranging from simple material transport to more complex tasks such as mobile manipulation, packaging, or palletizing [4].

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

Localization is a key capability of mobile robots, enabling them to navigate and operate effectively in complex environments [5]. The selection of a suitable Indoor Localization System (ILS) is paramount when developing a mobile robot to ensure its safe and reliable operation. As illustrated in Figure 1, for a localization system to be considered suitable, the system performance must satisfy the application's requirements [6]. A method for the derivation of location data requirements has already been proposed in the literature [7]. Hence, the present work focuses on benchmarking, which involves determining the performance of localization systems through empirical experiments. Hence, benchmarking serves as an indispensable tool, providing stakeholders with the necessary information to make informed decisions regarding system selection.

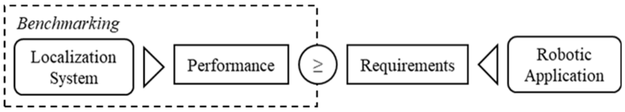

Figure 1: Selecting a localization system for a robotic application by matching system performance and requirements.

Diverse sensor technologies, such as Light Detection and Ranging (LiDAR), Inertial Measurement Units (IMUs), vision, or Ultra-Wideband (UWB), are commonly employed for localizing mobile robots [3]. The heterogeneity of these localization technologies introduces a wide range of influencing factors, such as radio interference, dynamic environments, and lighting conditions, that can potentially impact system performance [8]. Consequently, obtaining comparable and realistic benchmarking results for systems that are based on such diverse technologies poses a significant challenge [9]. This challenge is further amplified when the focus extends beyond a specific application or domain, encompassing a broader range of common influencing factors [10].

Consequently, existing benchmarking studies tend to focus on the comparison of similar technologies or specific algorithms, such as for vision-based localization [11] or LiDAR-based Simultaneous Localization and Mapping (SLAM) algorithms [12]. While such technology-driven benchmarking approaches are undoubtedly essential for advancing technology, the requirements and limitations of the real world are often disregarded, limiting their practical utility for the selection of localization systems.

This work presents a distinct approach by conducting application-driven benchmarking of commercially available localization solutions based on various technologies commonly used for mobile robots in warehouse and manufacturing scenarios. The primary objective is to provide a performance approach, conducting structured and automated comparisons between different systems based on comparable performance metrics. These facilitate the selection of the ideal localization system or technology for mobile robot applications. The key contributions of this work can be summarized as:

a) Elaboration of a testbed for automated execution of experiments

b) Intuitive data visualization for qualitative analysis of ILS' robustness

c) Use case analysis for different ILS influencing factors

The remaining paper is structured as follows. Section 2 provides an overview of related work. Next, the materials and methods for the benchmarking study are introduced in Section 3. Section 4 presents the results, which are subsequently discussed in Section 5. The paper closes by drawing conclusions and providing an outlook in Section 6.

2 Related Work

This section provides an overview on the topic of benchmarking of localization systems leading to the identification of the research gap addressed in this work.

As mentioned before, existing literature predominantly focuses on benchmarking studies that are centered around specific technologies. For example, Zuo et al. [13] conducted a comparative analysis of three widely used open-source SLAM algorithms for mobile robots using data from a LiDAR scanner, supplemented with IMU and odometry data. Besides position accuracy, the study evaluated orientation accuracy, processing power, and memory usage. Similarly, Ragot et al. [14] as well as Kasar et al. [11] conducted technology-driven benchmarking studies, comparing visual SLAM algorithms using image sequences from different camera types. Despite their valuable contributions to the scientific community, they primarily represent technology-driven benchmarking and offer limited practical utility for logistics stakeholders when it comes to selecting a localization system or technology.

However, other benchmarking studies place a stronger emphasis on the application aspects of mobile robotics. For example, Hofer et al. [15] evaluated four LiDAR scanners for mobile robot localization in the construction industry within three application-driven scenarios, using the same localization algorithm. Although the study does concentrate on a specific application domain, it maintains a primary focus on a component of a particular localization technology. Furthermore, although the study is insightful, repeatability and comparability across different scenarios and sensors could potentially be further enhanced by utilizing automated robot control [8].

In contrast, Karaagac et al. [16] and Crețu-Sîrcu et al. [17] conducted benchmarking of commercially available systems, with a specific focus on their application in industrial environments. While Karaagac et al. examined a Bluetooth and a UWB-based system across four different scenarios, addressing various system vulnerabilities, Crețu-

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

Sîrcu et al. compared a UWB-based system with an ultrasound-based system. These application-driven studies offer valuable insights into system performance within practical scenarios, contributing to market transparency by providing a deeper understanding of system capabilities. However, they are not centered on mobile robotics, but on industrial applications in general.

In conclusion, existing benchmarking studies of localization systems are predominantly focusing on certain technologies or even components. While some studies consider market-ready systems in industrial environments, they are not focused on the application of mobile robotics. Hence, there is a scarcity of application-driven benchmarking studies of localization systems specifically targeting the field of mobile robotics in warehouse and manufacturing scenarios. This work aims to address this research gap by implementing an application-driven benchmarking approach for the performance analysis of localization solutions in warehouse and manufacturing settings for mobile robots. By conducting experiments that reflect real-world applications and considering market-ready solutions, the goal is to offer valuable insights for stakeholders in system and technology selection.

3 Material and Methods

The systematic experimental analysis is crucial for understanding the impact of the application environment and the ILS configuration on the localization performance of ILS. This section describes the fully automated test setup and how the setup reflects a real-world intralogistics application and which ILS are implemented and tested. Next, the experiment process is presented, in which environmental disturbances are considered. Afterward, the automatic data collection pipeline is presented. Lastly, it is shown which ILS performance values are applied and why these are important for benchmarking ILS. The methodology for application-driven benchmarking of indoor localization systems, as presented by Schyga et al. [18], serves as the methodological foundation. It guides designing significant scenarios, specifying and executing experiments, and evaluating the performance for benchmarking in test halls.

3.1 Test Setup

This subsection describes which scenario is regarded for the application-driven test setup and how the setup reflects some of the influencing factors on ILS.

An application-driven scenario is defined by environmental influences and process influences [18]. The application scenario is a mobile robot inside a warehouse. This is a typical application and environment for mobile robots in the intralogistics domain, which is why this is chosen as a scenario. In the following the term *scenario* describes a set of environmental parameters, a set of process influences, and the configuration of the sensors and algorithms used for localization (cf. Figure 2). Process influences are the entity to be localized/tracked (ELT) itself, the motion, and the path. In our case, the ELT is a mobile robot that is used for typical logistical tasks, such as goods transportation, inventory management, or order picking. The robot operates with low speed (< 1 m/s), and low rotational speeds (< 90 deg/s) on a horizontal plane, driving through aisles of the warehouse or logistics area in straight lines and curves. There is no guidance using lines or markers. Since only one robot is used for the experiments, the process influences remain the same throughout the experimental series. Depending on the items that are stored inside the warehouse, the environment of warehouses can differ in the light reflectivity of the surfaces and the impact on electromagnetic wave propagation due to different material constants, such as the attenuation constant or reflection constant. How electromagnetic waves propagate in the environment influences the performance of radio-based ILS such as UWB systems. Metals for example have a high reflection coefficient while water and human tissue have a high attenuation coefficient for waves in the UWB spectrum.

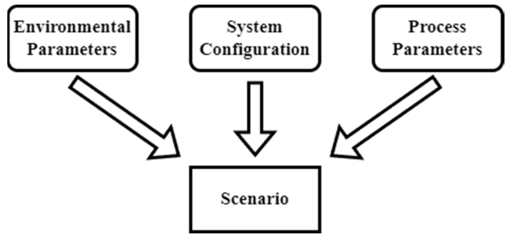

Figure 2: A scenario depends on the environmental parameters, the system configuration and the process parameters.

Typical for a warehouse are shelves in which items are stored. In the shelf slots, there can be items of different sizes stored or not items stored. For map-based localization, the static structures of the environment are relevant, which are the walls and the vertical beams of the shelves. Items that are stored inside the shelves should be removed before map creation or digitally removed during the clean-up process for the map. Lighting has also been shown to influence vision-based ILS. Warehouses can have different lighting due to the number and size of windows, the weather, and the number and brightness of artificial light sources. Dynamic obstacles, such as humans and other robots, show up on the laser scans provided by LiDAR sensors and also change the way electromagnetic waves propagate in space. Some warehouses are fully automized and there are no humans present and in other warehouses, humans and mobile robots are acting in the same environment simultaneously.

Figure 3 shows the test setup. The experiment area is in the center of the test setup with a geometrical size of 9 m × 9 m. To represent a typical structure of a shelf, the distance and dimensions of the vertical beams for a typical shelf were measured and cardboard boxes that roughly fit

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

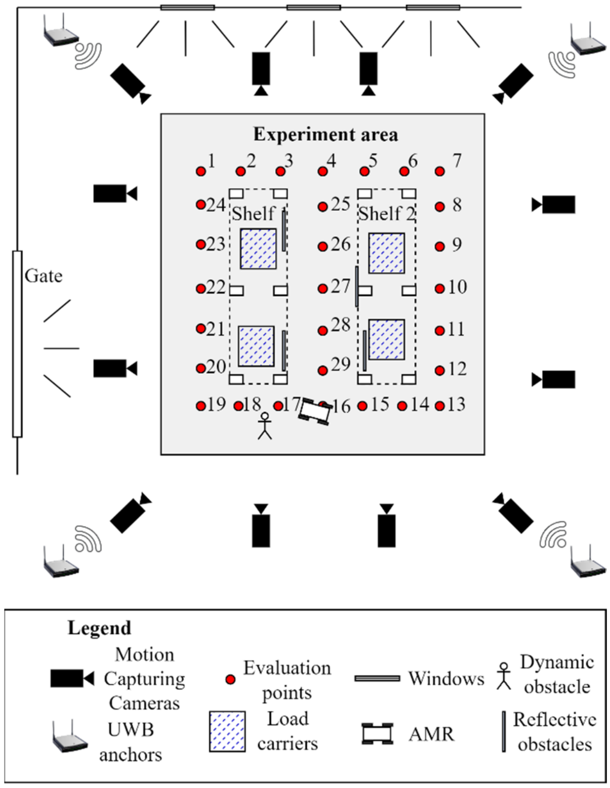

Figure 3: Experimental set up

the dimensions of the vertical beams were stacked on top of each other and placed according to the distance measured of an actual shelf. Items that are stored on the shelves are represented by metal carts. Stored items with high light reflection constants, like metal sheets or glass panes are placed sporadically in the test environment. The robot moves within the experiment area through a series of evaluation points. Evaluation points are the points on which the localization performance is later evaluated. The evaluation of the ILS performance is done using an optical motion-capturing system (*MoCap*) that consists of twelve cameras. These cameras are equipped with emitters that send out light in the infrared spectrum. The system can determine the pose of objects with millimeter accuracy by measuring the Time-of-Flight (TOF) to markers that get attached to the ELT. Additionally, the setup includes 4 UWB anchors. These anchors act as reference points and are responsible for transmitting UWB signals at precise intervals.

As described earlier, different factors can influence the performance of ILS. Throughout the environmental series, some factors that have shown an impact on some types of visual- and radio-based ILS were changed. The varied parameters were:

- Lighting: The lighting was varied by turning the ceiling lights on/off, opening/closing blinds for three large windows, and opening/closing the gate of the lab. The lighting is quantified by the illuminance, which was roughly determined by measuring the illuminance within the experiment area and tracking the maximum and minimum values. This choice of parameters represents warehouses with different lighting intensities.

- Reflectivity: The reflectivity in the environment is varied by adding metal sheets for some of the experiments. Reflective surfaces affect the laser scans that are measured with the LiDAR sensors. Reflective surfaces could be present in a warehouse or production environment, e.g., by glass panes or sheet metal that are being produced or stored or are part of machinery or the building itself.

- Dynamic human obstacles: Dynamic human obstacles were added for some experiments by having a person walk around the robot in circles while it was following its trajectory. This represents environments in which vehicles as well as humans operate.

- Metal obstacles: Metal load carriers are added for some experiments to the experiment area. This represents warehouses in which metal objects are stored.

The process influences (ELT, speed and acceleration settings, and the path) for this experimental series are the same for all experiments. The environmental influences that stay the same for all experiments are a room temperature between 20 °C and 23 °C, an atmospheric pressure of approx. 1013 hPa, relative humidity between 40% to 60 %, a low maximum distance to static structures (< 2 m), and a weak electromagnetic field (absolute magnetic flux density < 150 µT). These are typical values for many warehouses. Cold or dark warehouses or block storage areas are not represented by the experiments. Tropical climates, high-altitude atmospheric pressure, or areas with machines that produce strong electromagnetic fields are also not represented by the experiments. Another influencing factor that is not considered in this experimental series is the cross-interference of sensors by their light emitters which can occur when multiple entities in the same area are equipped with LiDAR sensors or depth cameras.

3.2 Sensors and Systems

An ILS consists of a set of sensors and an algorithm that processes the sensor data. In the following, the sensor technologies that were used in this study shall be briefly explained:

- *LiDAR:* LiDAR sensors are advanced detection and ranging devices that utilize laser technology to measure distances and create detailed representations of their surrounding environments. These sensors emit laser pulses and calculate the time it takes for the

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

pulses to reflect off objects, enabling them to accurately determine distances. In this study, three different LiDAR sensors are used:

a. *SICK microScan3*
b. *Velodyne Puck*
c. *SICK MRS1104*

- *Tracking Camera*: The underlying algorithm of the deployed *Intel Realsense T265* is a feature-based visual SLAM algorithm. The T265 uses sensor fusion techniques to combine the visual odometry and IMU data. By integrating these two sources of information, the camera achieves more robust and accurate tracking, overcoming the limitations of using either visual or inertial data alone.
- *Ultra-Wideband*: The system, consisting of four anchor nodes and a localization tag, the SICK LOCU system, utilizes UWB technology to determine the distances between anchor nodes and the localization tag once the localization tag calculates its position based on the distance measurements, it transmits the position information to a central processing unit or a control system. The data that the system provides is the 2D position of the UWB-Tag.

The chosen technologies represent a set of established, easily integratable solutions in the ILS market. The specific solutions are summarized in Table 1. To evaluate systems based on heterogeneous technologies, a custom-built mobile robot was equipped with a UWB-Tag, three LiDAR sensors, a tracking camera, and markers for an optical motion-capturing system (cf. Figure 4).

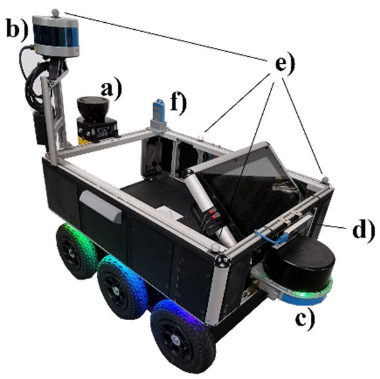

Figure 4: Sensor placement on our custom-built robot: a) SICK microScan3, b) Velodyne Puck, c) SICK MRS1104, d) Intel Realsense T265, e) Markers for Qualisys Motion Capturing System, f) UWB-Tag for SICK LOCU

Table 1: Overview of used sensors and sensor systems

System	**Technology**	**Data**
SICK microScan3	LiDAR	2D scan
Velodyne Puck	LiDAR	2D scan
SICK MRS1104	LiDAR	2D scan
Intel Realsense T265	Tracking-Camera	odometry
SICK LOCU	Ultra-Wideband	Position

All LiDAR sensors provide 3-dimensional point cloud data and 2-dimensional laser scan data. For this benchmarking study, the 2D-localization algorithm *AMCL (Adaptive Monte Carlo Localization)* is applied using laser scan data in combination with odometry data. Using the previously described sensors, four different ILS are compared. Implemented are six different AMCL-based ILS and one UWB-based ILS. AMCL uses a 2D laser scan of the environment and matches it to a pre-recorded map with a relatively low update rate. Odometry data can be used to update the position data more frequently between scan-to-map matches. The AMCL-based ILS differ in the source of the laser source. A tracking camera provides odometry data.

As a robotics framework, the Robot Operating System (ROS) was used. These systems were integrated into a testbed, enabling automated execution of experiments. The autonomous navigation of the robot was implemented using the *move_base* package. For obstacle avoidance, the laser scans of all three LiDAR sensors were combined into one laser scan. The implemented test bed allows a fully automated and reproducible test procedure for the performance analysis of the integrated ILS.

In the following, the benchmarking procedure is presented in a process-oriented manner. First, the definition of scenarios is presented. Next, the experiment specification is elaborated. After the presentation of the experiment execution, the performance evaluation will present how the experiment data was analyzed.

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

3.3 Experiments

Due to the large dimension of environmental influences and a large number of different systems and system configurations, it is required to conduct a large number of experiments to accurately reflect real-world relation to the performance of ILS. Automating the experiment procedure allows to create large data sets more efficiently. This section explains the automated experiment procedure. As shown in Figure 5, the experiment phase is divided into three stages: the setup process, the online/live phase, and the offline experiment.

Setup Process: Calibrating MoCap refers to the process of setting up and aligning the motion capture system to accurately track the movements of the robot. The calibration process ensures that the captured data is properly synchronized and accurately represents the real-world positions and orientations of tracked markers. Define trajectory refers to the path that a robot follows during the experiment. Additionally, the environment is changed scenario

Online/Live Phase: The robot moves on a predefined trajectory, and all sensor readings are recorded by saving the raw sensor data into a rosbag. This creates a valuable dataset that can be later used for various purposes, such as algorithm development, testing, and machine learning. Measure Environmental Parameters involves collecting data in the experiment area and defining all the parameters that affect the localization system. Following a given trajectory involves guiding robot movements along a predetermined path. Recording all sensor data in a rosbag: The data collected from all sensors on the robot during the execution of the trajectory is saved in a specific format (rosbag) that allows efficient storage and playback of data.

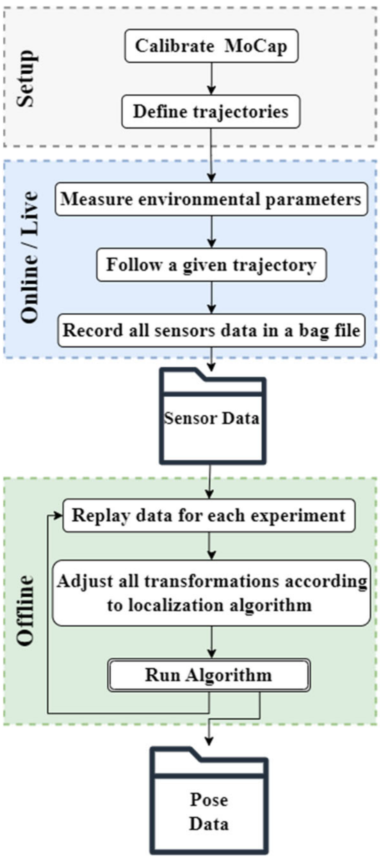

Figure 5: Flow diagram of the data collection process

Offline Experiment: This process involves adjusting all transformations based on the localization algorithm, replaying the recorded rosbag containing raw data, running multiple localization algorithms (Algorithm X), and saving the pose data for each algorithm along with the ground truth pose. Adjusting all transformations according to the localization algorithm: The robot's sensor data, such as camera images or laser scans, is processed through a localization algorithm. The localization algorithm estimates the robot's pose (position and orientation) relative to a global or local coordinate system, adjusting the transformations between sensor frames and the robot frame to align them with the chosen reference frame.

With the recording of raw sensor data, it is possible to replay the data and reconfigure the localization algorithms while keeping the input to the algorithm the same. This increases the reproducibility and the flexibility to benchmark sensor fusion approaches.

3.4 Performance Evaluation & Visualization

This subsection presents the evaluation of the pose data. All pose data for the different ILS is saved in a folder that corresponds to a specific experiment. A Python script checks which experiments were not evaluated and loops over the folders that were not evaluated yet. It loads the pose data from rosbags and loads the experiment configuration from a YAML file, downsampling the ground truth trajectory such that it has the same length as the ILS that is currently being evaluated. For each timestamp, the horizontal error and the orientation error are determined. For these error arrays the maximum, the mean, and the standard deviation is determined. Once the pose data for an ILS for a specific experiment is evaluated, the data is appended to a *data frame* object which is created using *pandas* - the Python data science library. Once the evaluation script has looped over all the experiments and the pose data within, the data frame is complete and the data frame is saved as a CSV file with the name *results.csv*. The data frame consists of 26 columns in which the following information is saved:

- Experiment-ID: date-timestamp at experiment start
- System Configuration: a short description of the ILS, scan source, odometry source, whether or not UWB-data is used

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

- Environmental parameters: a short description of the lighting situation, minimum and maximum illuminance values, Boolean values indicating whether or not dynamic obstacles, reflective surfaces, or metal carts were added to the experiment area for the respective experiment

- Trajectories: Complete trajectories for all ILS under test and the reference trajectory. The trajectories are defined by x, y, theta time series data

- Temporal error propagation: orientation error time series, position error time series

- Performance metrics: maximum position error, mean position error, standard deviation of position error and orientation error, maximum orientation error, mean orientation error.

Saving this information in a data frame allows for easier filtering, which offers more insights, e.g., using visualization techniques. The visualization script mainly utilizes methods from the Python library *matplotlib.*

4 Experimental Results

To demonstrate the potential of the experiment procedure and the benefits of adequate visualization techniques and meaningful performance metrics, example data from a small experimental series is presented. For each scenario (cf. Figure 2) one experiment was conducted. To show exemplary performance analysis of ILS in the logistics domain, a use case of data analysis in the following chapter. Since the focus lies on what information can be deducted from which type of plot and not on the analysis of the impact of each environmental factor, one scenario is shown. Example pose data for the experiment is presented in various ways. The plots that will be shown are merely an excerpt of the plots that are created within the visualization script. The complete visualization of the experiments is available on the corresponding GitLab project [19].

To increase the statistical significance of the results and to enable a meaningful analysis of the impact of the influencing factors, more experiments are to be conducted. Using the visualization evaluation script different plots are generated – 2D-trajectory plots, orientation-error-over-time-plots, and radar charts that allow visualizing robustness of an ILS towards different environmental influences.

Figure 6 shows the 2D trajectories of all ILS for a chosen example experiment. The robot's initial position is in the top-middle of the experiment area and ends in the lower-left corner. The bottom-right corner shows the environmental parameters of the experiment and the experiment ID. In this example scenario, no obstacles are added to the experiment area, and the illuminance during the experiment is between 600 to 1000 lux. For each system, the position error for different areas of the environment can be

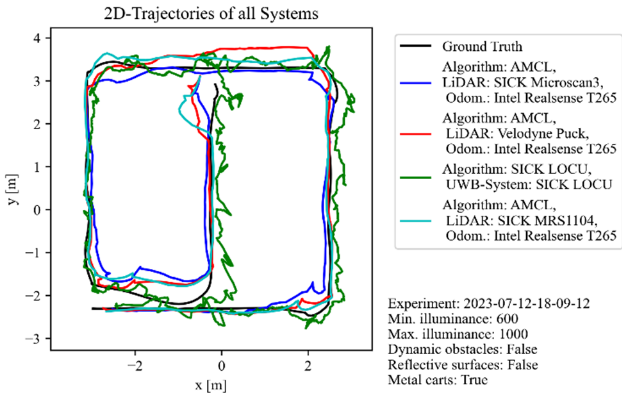

Figure 6: 2D-Trajectories show spatial influences on the position error.

qualitatively estimated by the viewer. The qualitative trajectory is then to be put into the context of the setup to identify the first implication of different environmental influences on the localization performance.

In the given example the result shows that the UWB-based ILS shows a smoother trajectory on the vertical section on the left side of the experiment area, a rougher trajectory on the right side of the experiment area, and the roughest curve between the shelves in the middle of the experiment area. This is possibly due to the lower distance towards the closest UWB anchor and due to the obstacles left and right of the middle vertical section of the trajectory. When planning the layout of the UWB anchors, the distance the maximum distance to the ELTs should be considered to achieve the desired performance. To investigate spatial dependencies on the localization performance of ILS in future experiments, live-tracking of the distances to the anchors should be implemented.

While the UWB-based localization seems to be influenced by the distances to the anchors, the position error of the AMCL-based localization systems is also dependent on the location within the experiment area. This can be seen by the larger distances of the ILS trajectories to the Truth trajectory on the bottom left part of the experiment area and the top horizontal section of the trajectory. These areas correspond to the areas with the highest illuminance values due to the proximity of the windows and the hangar gate (cf. Figure 2). The trajectory corresponding to the UWB-based ILS is less smooth than the trajectories that correspond to the AMCL-based ILS under certain environmental conditions.

Besides the position of the robot, determining the orientation of the robot is important for navigation and control. If there is an error in the robot's understanding of its orientation, it might misinterpret its position within a warehouse or the direction it needs to take. Analyzing the orientation error over time can help in noticing dependencies of dynamic environmental changes. Depending on the application and the accuracy it requires, the orientation error of the robot is required to remain within a certain range. Therefore, analyzing this error can help find the right set of sensors for mobile platforms. Figure 7 shows the orienta-

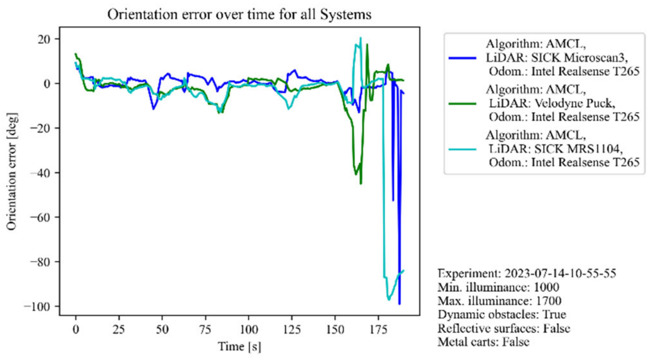

Figure 7: Orientation error time series

tion error of the ILS over time for a scenario with high illuminance. This example is chosen to demonstrate its potential regarding the intuitive understanding of the connection between environmental changes on the performance metrics of ILS. In the first 150 seconds of the experiment, the absolute value of the orientation error remains below 20 degrees for all systems. In the last 30 seconds of the experiment, the orientation error of all ILS shows large error spikes of up to 100 degrees. Toward the end of the trajectory, the robot approaches the hangar gate. In this specific experiment, the sun was shining directly through the windows of the hangar-gate which resulted in a relatively extreme lighting scenario for indoor applications. This most likely caused the large orientation error spikes. This type of plot becomes more valuable as environmental parameters such as the illuminance are tracked and saved over time opposed to noting the minimum and maximum values. By plotting the time series describing environmental parameters underneath the influences can be intuitively visualized. This feature will be implemented in the future.

Radar charts are a type of data visualization that can be visually appealing and informative for certain types of data. They are particularly useful for displaying multivariate data, where each variable is represented as a spoke emanating from a central point, forming a polygonal shape. A spoke refers to one of the straight lines that radiate outward from the center of the chart to the outer edge. Each spoke represents a different variable or dimension of data being plotted. The primary reason radar charts are visually interesting is due to their unique ability to convey patterns and relationships within multiple dimensions of data simultaneously. This feature allows us to study the impact of individual environmental parameters on any performance metric of ILS. Figure 8 shows an exemplary radar chart that shows the robustness of the ILS towards specific environmental influences. The robustness is shown by the magnitude of performance values such as the mean horizontal error or the

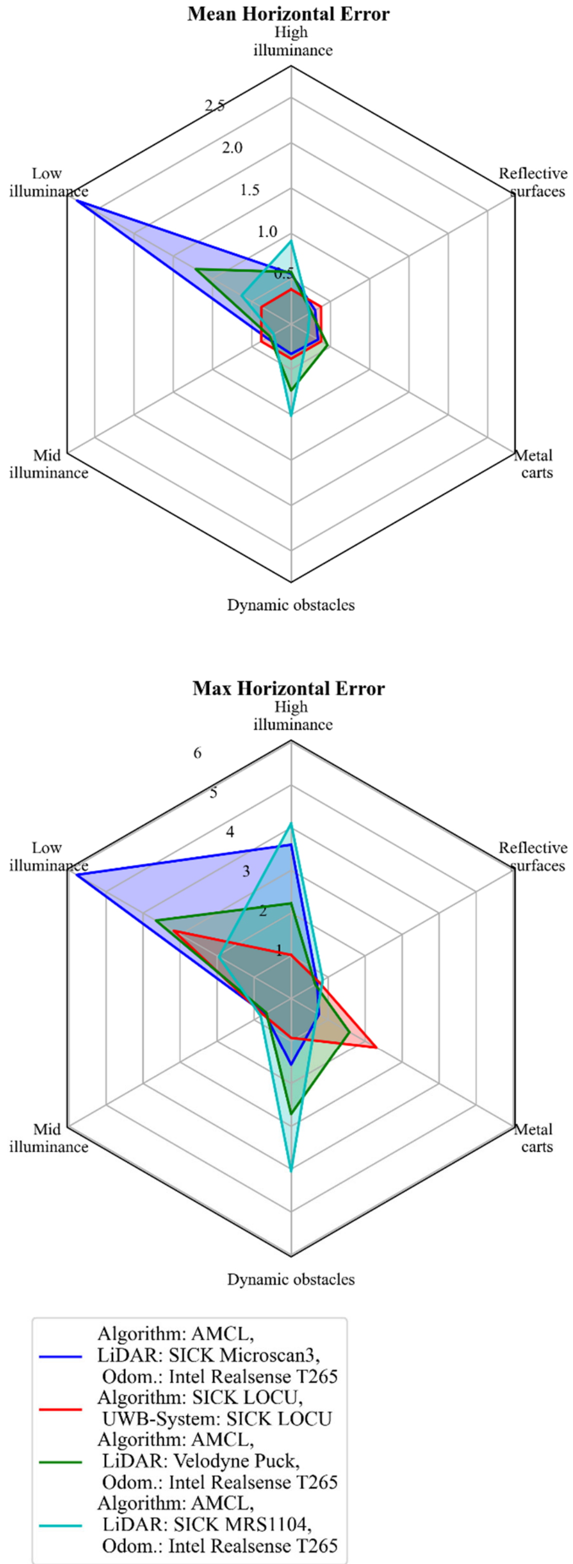

Figure 8: Radar charts are able to convey patterns and relationships within multiple dimensions of data simultaneously

maximum horizontal error. The horizontal error is a key performance metric as it directly impacts the size uncertainty space in which the ELT could be located. Depending

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

on the use case the maximum or mean value of a performance value is considered. For safety-critical use cases typically the maximum error is considered while for other use cases, the mean horizontal often in combination with the standard deviation is sufficient. The horizontal error for each ILS for each scenario category is plotted within one radar chart. The distance to the center of the radar chart indicates the magnitude of the performance value for the corresponding environmental influence.

In the case of the mean and maximum values for the horizontal error, a large radar chart of an ILS indicates a poor general performance of the ILS. A peak in the radar chart can indicate a dependency on an environmental influence. The UWB-based ILS shows even magnitudes for the different environmental influences for the mean horizontal error but the maximum horizontal error peaks when the environment has metal carts or low illuminance. The radar chart shows a larger maximum horizontal error for the UWB-based ILS when the illuminance is low. There should not be an influence of illuminance on the performance of UWB-based localization. This effect would most likely not be observed if more experiments were conducted. The radar chart corresponding to the mean horizontal error supports that. The result that the mean horizontal error is the highest for two of the AMCL-based ILS when the illuminance is low is counter-intuitive as well. This result is not due to the lighting but seems to be due to the repetition of similar patterns within the recorded map as can be seen by the trajectory shift of map-based ILS in Figure 9. The AMCL-based ILS compares a live laser scan

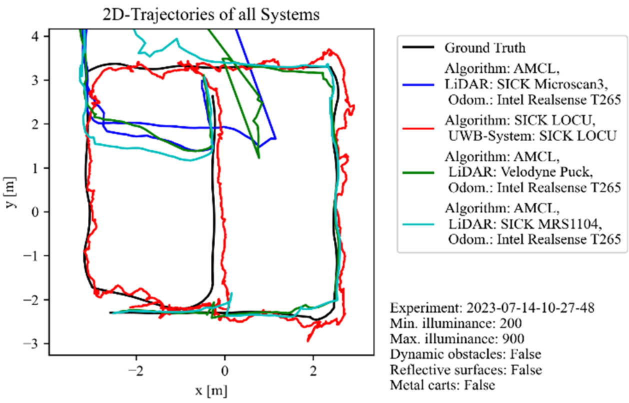

Figure 9: Shifted trajectories of map-based ILS due to repetitive patterns within the pre-recorded map

with a pre-recorded 2D map of the environment. This map includes the pillars of the logistics shelves. Since the pillars of the shelves form a pattern that is reoccurring within the map (cf. Figure 3), there can be instances where map-based ILS mistakenly calculates to be in a different row or aisle. Including reflective markers that act as reference points or a sensor fusion with the UWB-based ILS could alleviate the issue. Avoiding patterns in building structures would put additional requirements on the design of the warehouse and would be impractical. The occurrence of the trajectory shift appears to be random which results in a limited interpretability of the radar chart for small experimental series like the one in this study.

In summary, the utilization of different types of plots can greatly aid in the analysis of ILS performance for mobile robotics. Trajectory plots, for instance, offer valuable insights into the performance of various ILS in different spatial points. By revealing the strengths and weaknesses of ILS like UWB and vision-based systems concerning factors like distance and lighting conditions, these plots enable a comprehensive understanding of their spatial effectiveness. On a similar note, time series plots come into play when analyzing the effect of dynamic environmental changes. Additionally, the radar chart serves as a valuable tool for assessing overall system performance. By allowing for a comparison of diverse influences, this plot type aids in identifying the most crucial factors that impact a given application's success. In essence, the strategic utilization of these plot types offers an enhanced understanding of the intricate relationships and dependencies of ILS performance and the environment.

5 Discussion

The results show that the performance analysis of ILS can vary significantly under the impact of different environmental influences. Therefore, benchmarking needs to be performed to find the best-suited ILS for a specific application and its surrounding environment.

Potentials

The integration of radar charts introduces an element of intuitive comprehension in assessing the robustness of ILS. This innovation augments the transparency of performance metrics provided by system manufacturers and algorithm developers. These charts offer a clearer understanding of ILS performance, aiding in more informed decision-making.

The capability to simultaneously capture raw data and subsequently replay it opens avenues for the fusion of sensors in algorithms like AMCL. This facet increases flexibility during the exploration of novel algorithms and configurations, further enhancing the adaptability of ILS. Many algorithms require the same raw data such as laser scan data or odometry data. Likewise, different sensors are available to provide the same data. The offline execution of the localization algorithms enables the extensive configuration and combination of the raw data and the algorithms. Additionally, the simultaneous capture of the sensor data fosters the comparability of the localization data of the ILS under test.

Due to the large number of possible technologies, sensor models, and combinations of different sensors via sensor fusion, it becomes difficult to analyze a data set of that

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

set for any combination of environmental and process parameters. A recommender system would be highly beneficial for system developers as it increases transparency of the system performance for a given scenario allowing the system developer to make more calculated decisions. Performing a large number of experiments efficiently by leveraging the automation of the experiments allows to create a large data set that maps environmental influences and system configurations to ILS performance values. A data set like this can be used to train Machine Learning (ML) algorithms resulting in a model that can recommend an ILS and its configuration depending on parameters that describe the application environment and the process parameters. Such a model would simultaneously be able to predict the performance values of the ILS. A challenge to this goal would be the large number of data points needed to train ML models. Currently, the length of a data set l created by an experimental series consisting of *next* experiments and n_{ILS} different ILS is determined by $l = n_{exp}n_{ILS}$. This is because the environmental and process parameters are only measured once. By measuring the environmental and process parameters with a sample rate f_s the length of a data set is substantially increased for the same number of experiments and the same number of ILS under test. The length of the data set in that case would be determined by $l = \frac{n_{exp}n_{ILS}t}{f_s}$ where t denotes the duration of the experiments. This would also improve the quality of the data since the environmental and process parameters can be dynamic and multiple measurements can more accurately capture the current environmental and process influences that are applied in a specific moment.

Limitations

It is difficult to attribute the impact on ILS to specific environmental influences due to the limited statistical significance of the experiments. This limitation calls for a cautious interpretation of the results and emphasizes the need for a larger number of experiments. For UWB-based ILS, additional sensors are imperative to add orientation data, a critical component for typical robotic tasks such as navigation. Certain environmental factors, such as vast open spaces within logistics environments cannot be investigated in the current research facility due to a lack of space. The simultaneous collection of raw data is hampered by the physical limitations of sensor placement on the robot. Ideally, each sensor of the same type is placed in the same mounting point on the robot. To overcome this challenge, future experiments should be conducted multiple times while exchanging the sensors' placement on the robot. The results show that continuously measuring environmental parameters could increase the interpretability of the results.

Critical Reflection

State-of-the-art ILS performance is yet to be achieved, due to the absence of comprehensive sensor fusion techniques and little tuning of the used algorithms. This finding prompts further exploration of sensor fusion strategies and the enhancement of ILS parameters for future experiments. Benchmarking ILS proves to be a formidable undertaking, largely due to the time-intensive nature of conducting a sufficient number of experiments to attain statistically meaningful results. This inherent challenge underscores the need for efficient and pragmatic experimental designs in future benchmarking endeavors. The complex nature of capturing the multidimensional aspects of both the environment and system configuration remains an unresolved issue. Overcoming this challenge necessitates innovative approaches to capture the intricate interplay between these variables. Cost considerations emerge as a significant factor in the practical application of ILS. The study's focal point on data collection processes and visualization techniques underscores the commitment to enhancing reproducibility and transparency in benchmarking results. These efforts facilitate more informed comparisons and assessments within the robotics community. It's important to note that the study's scope was limited to analyzing absolute localization, while robotic systems encompass a broader range of capabilities. As the study indicates, relative localization through distance sensors and object detection via cameras are among the additional capabilities that warrant exploration in future investigations.

6 Conclusions & Outlook

This work introduces a novel approach to benchmarking indoor localization systems for mobile robots in warehouse and manufacturing scenarios. By conducting application-driven experiments, this study contributes to enhancing transparency and aiding stakeholders in making informed decisions regarding the selection of ideal localization systems or technologies.

The potential and limitations of various commercially available systems based on diverse technologies have been systematically explored and evaluated. The presented experimental results highlight key insights into the performance of different localization systems under varying environmental conditions. Notably, the study sheds light on the impact of environmental factors, such as lighting and distance to UWB anchors, on system performance. The innovative use of radar charts enhances the visualization of system robustness, facilitating a clearer understanding of how different systems respond to distinct environmental influences. Furthermore, it is demonstrated that spatial and temporal insights can be fostered by adequate data visualization techniques. The multitude of technologies and sensor combinations make analyzing data for different environmental and process scenarios challenging. A recommender system would aid developers by clarifying system performance and informing decisions. Automating experiments to gather extensive data on how environmental factors and system setups affect performance can help train

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

Machine Learning models. The main challenge is collecting enough data for effective model training. By continuously measuring parameters, the dataset can be enriched, leading to better insights into dynamic influences.

While the study presents valuable findings, certain limitations are acknowledged. The statistical significance of the experiments remains constrained, urging a cautious interpretation of results. The study also underscores the importance of sensor fusion strategies and optimization time frames to achieve optimal performance.

In reflection, this research paves the way for future investigations in sensor fusion techniques, efficient experimental designs, and advanced localization capabilities beyond absolute positioning. The commitment to transparency, data collection, and visualization techniques contributes to the reproducibility and comparability of benchmarking results within the robotics community. This allows robotics developers to make more justified decisions when selecting an ILS for an application. As the field of mobile robotics continues to evolve, the insights gained from this study will serve as a valuable resource for researchers, practitioners, and stakeholders seeking to navigate the complex landscape of Indoor Localization Systems.

Acknowledgments

This work is funded by the Federal Ministry of Education and Research (BMBF, FKZ: 01IS22047).

Federal Ministry of Education and Research

References

[1] M. A. Kamarul Bahrin, M. F. Othman, M. F. Othman, N. H. Nor Azli, and M. F. Talib, “Industry 4.0: A Review on Industrial Automation and Robotic,” *Jurnal Teknologi*, vol. 78, 6-13, 2016, doi: 10.11113/jt.v78.9285.

[2] M. Benz, M. Dold, S. Hochhäuser, and M. Jacob, “AGV System Localization: More Efficiency in Intralogistics,” SICK AG, 2022. Accessed: Jul. 17 2023. [Online]. Available: https://cdn.sick.com/media/docs/4/14/814/whitepaper_mobile_platforms_en_im0104814.pdf

[3] K. Dohrmann *et al.,* “DHL Logistics Trend Radar: Sixth Edition,” DHL, 2022. Accessed: Jul. 17 2023. [Online]. Available: https://www.dhl.com/global-en/home/insights-and-innovation/insights/logistics-trend-radar.html

[4] H. Unger, T. Markert, and E. Müller, “Evaluation of use cases of autonomous mobile robots in factory environments,” *Procedia Manufacturing*, vol. 17, pp. 254–261, 2018, doi: 10.1016/j.promfg.2018.10.044.

[5] P. Skrzypczyński, “Mobile Robot Localization: Where We Are and What Are the Challenges?,” in *Advances in Intelligent Systems and Computing Ser, Automation 2017: Innovations in Automation, Robotics and Measurement Techniques*, C. Zieliński and M. Kaliczyńska, Eds., Cham: Springer International Publishing, 2017, pp. 249–267.

[6] R. Mautz, “Indoor Positioning Technologies,” Habilitation, Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, 2012.

[7] J. Schyga, M. Knitt, J. Hinckeldeyn, and J. Kreutzfeldt, “Method for Specifying Location Data Requirements for Intralogistics Applications,” in *Proceedings of the Conference on Production and Logistic Systems (CPSL) 2023*.

[8] Sudeep Pasricha, “Overview of Indoor Navigation Techniques,” in *Position, Navigation, and Timing Technologies in the 21st Century*: John Wiley & Sons, Ltd, 2020, pp. 1141–1170.

[9] F. Potorti, A. Crivello, P. Barsocchi, and F. Palumbo, “Evaluation of Indoor Localisation Systems: Comments on the ISO/IEC 18305 Standard,” in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2018.

[10] F. Potortì, A. Crivello, and F. Palumbo, “The EvAAL Evaluation Framework and the IPIN Competitions,” in *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation*: Elsevier, 2019, pp. 209–224.

[11] A. Kasar, *Benchmarking and Comparing Popular Visual SLAM Algorithms*, 2018.

[12] H. Sier, Q. Li, X. Yu, J. Peña Queralta, Z. Zou, and T. Westerlund, “A Benchmark for Multi-Modal LiDAR SLAM with Ground Truth in GNSS-Denied Environments,” *Remote Sensing*, vol. 15, no. 13, p. 3314, 2023, doi: 10.3390/rs15133314.

[13] Q. Zou, Q. Sun, L. Chen, B. Nie, and Q. Li, “A Comparative Analysis of LiDAR SLAM-Based Indoor Navigation for Autonomous Vehicles,” *IEEE Trans. Intell. Transport. Syst.*, pp. 1–15, 2021, doi: 10.1109/TITS.2021.3063477.

[14] N. Ragot, R. Khemmar, A. Pokala, R. Rossi, and J.-Y. Ertaud, “Benchmark of Visual SLAM Algorithms: ORB-SLAM2 vs RTAB-Map,” in *2019 Eighth International Conference on Emerging Security Technologies (EST)*, 2019.

[15] Hofer, *Lokalisierung von mobilen Robotersystemen im Hochbau mittels LiDAR-Technologie: Abschlussbericht zur Kurzstudie.* [Online]. Available: https://epub.sub.uni-hamburg.de/epub/volltexte/2023/148646/

[16] Abdulkadir Karaagac, Jetmir Haxhibeqiri, Matteo Ridolfi, Wout Joseph, Ingrid Moerman, Jeroen Hoebeke, *et al.,* “Evaluation of accurate indoor localization systems in industrial environments: September 12-15, 2017, Limassol, Cyprus,” in pp. 1–8.

DOI: 10.2195/lj_proc_knitt_de_202310_01
URN: urn:nbn:de:0009-14-58098

[Online]. Available: http://ieeexplore.ieee.org/servlet/opac?punumber=8233358

[17] A. L. Crețu-Sîrcu *et al.,* "Evaluation and Comparison of Ultrasonic and UWB Technology for Indoor Localization in an Industrial Environment," *Sensors (Basel, Switzerland)*, vol. 22, no. 8, 2022, doi: 10.3390/s22082927.

[18] J. Schyga, J. Hinckeldeyn, and J. Kreutzfeldt, "Meaningful Test and Evaluation of Indoor Localization Systems in Semi-Controlled Environments," *Sensors*, vol. 22, no. 7, p. 2797, 2022, doi: 10.3390/s22072797.

[19] M. Knitt, J. Schyga, *Evaluation and Visualization Tools for Benchmarking ILS*, [Online]. Available: https://collaborating.tuhh.de/JakobSchyga/evaluation-and-visualization-tools-for-benchmarking-ils (accessed: Aug. 24 2023)

Markus Knitt M. Sc., Research Associate at the Institute for Technical Logistics, Hamburg University of Technology. In his master's studies in mechatronics engineering at the Hamburg University of Technology he specialized in robotics and intelligent systems. Since 2023 Markus Knitt has been working on the investigation and modelling of the behavior of Indoor Localization Systems in intralogistics.

Yousef Elgouhary B. Sc., Student Assistant at the Institute for Technical Logistics, Hamburg University of Technology. He studied mechatronics engineering at the October 6 University in Giza, Egypt. Since 2023, he has been pursuing a Master's degree in Geodesy and Geoinformatics at the HafenCity University Hamburg.

Jakob Schyga M. Sc., Research Associate at the Institute for Technical Logistics, Hamburg University of Technology. After finishing his master's studies in mechanical engineering at the Hamburg University of Technology. Since 2019 Jakob Schyga has been working on the investigation and application of Indoor Localization Systems in intralogistics.

Hendrik Rose M. Sc., Research Associate at the Institute of Technical Logistics at Hamburg University of Technology. Hendrik Rose completed his studies in industrial engineering with a focus on general engineering science and logistics at Hamburg University of Technology in 2020. Since 2023, he has been in charge of coordinating the robotics research at the Institute of Technical Logistics.

Philipp Braun M. Sc., Chief Engineer at the Institute of Technical Logistics at Hamburg University of Technology. Philipp Braun completed his studies in industrial engineering with a focus on information technologies and logistics at Hamburg University of Technology in 2020. Since 2023, he has been in charge of coordinating the logistics planning and simulation research at the Institute of Technical Logistics.

Prof. Dr.-Ing. Jochen Kreutzfeldt, Professor and Head of the Institute for Technical Logistics, Hamburg University of Technology. After studying mechanical engineering, Jochen Kreutzfeldt held various managerial positions at a company group specializing in automotive safety technology. Jochen Kreutzfeldt then took on a professorship for logistics at the Hamburg University of Applied Sciences and became head of the Institute for Product and Production Management. His current research focuses on warehouse and production planning as well as process optimization.

Address: Institute for Technical Logistics, Hamburg University of Technology, Theodor-Yorck-Strasse 8, 21079 Hamburg, Germany; Phone: +49 40 42878-2981, E-Mail: markus.knitt@tuhh.de