Statistical Evaluation and Synthetic Generation of Ultra-Wideband Distance Measurements for Indoor Positioning Systems

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY-SA 4.0

SUBMISSION DATE / POSTED DATE

08-10-2021 / 15-10-2021

CITATION

Schwarzbach, Paul; Weber, Richard; Michler, Oliver (2021): Statistical Evaluation and Synthetic Generation of Ultra-Wideband Distance Measurements for Indoor Positioning Systems. TechRxiv. Preprint.
<https://doi.org/10.36227/techrxiv.16773967.v1>

DOI

[10.36227/techrxiv.16773967.v1](https://doi.org/10.36227/techrxiv.16773967.v1)

Statistical Evaluation and Synthetic Generation of Ultra-Wideband Distance Measurements for Indoor Positioning Systems

Paul Schwarzbach^{ib}, Richard Weber^{ib}, and Oliver Michler^{ib}

Abstract—Wireless positioning systems are a key enabler for a variety of applications. In indoor environments, Global Navigation Satellite Systems are not available, hence technological alternatives are required. Ultra-Wide Band (UWB) radios provide high-precision distance measurements, however are still prone to ranging errors due to the environment. For localization system design and testing, this leads to cost-intensive field trials in which localization methods are validated. As an alternative, publicly available UWB datasets exist, which however are also limited to only certain environmental conditions. In order to provide a more general evaluation basis, we present and discuss a procedure which allows the synthetic generation of UWB distance measurements with respect to theoretical error sources and parameterization based on a conducted measurement campaign. This contribution builds an understanding of commonly present errors for wireless distance measurements and their effects on ranging residual distributions in the distance domain. In addition, we discuss the influences of non-line-of-sight and multipath conditions. Based on this work, distance measurements with respect to the multipath richness of the propagation environment can be generated, helping to flexibly evaluate positioning methods for various emerging applications for telematics systems in the consumer, industrial, or transportation sector. These application fields are ranging from unmanned aerial vehicles, in-house parking systems, and intelligent freight wagons to connected aircraft cabins.

Index Terms—Indoor Positioning System (IPS), Real-time Location System (RTLS), Sampling, Sensor Model, Ranging Simulation, Ultra-Wide Band (UWB), Wireless Sensor Networks (WSN)

I. INTRODUCTION

LOCATION-BASED services (LBS) are rapidly growing in both quantity and importance. Driven by the digitization, automation, as well as the accompanying interconnection of people, devices and processes, LBS find their way into many application fields. Especially in traffic telematics, where the current position of vehicles and mobile devices is a fundamental prerequisite for monitoring and controlling tasks, LBS are in demand. Global Navigation Satellite Systems (GNSS) are still the primary source for location-related information for many applications. However, in GNSS denied areas, so-called Indoor Positioning Systems (IPS) [1] are gaining more attention as a technological basis for LBS.

This paragraph of the first footnote will contain the date on which you submitted your paper for review.

P.S. and O.M. are with the Institute of Traffic Telematics, Technische Universität Dresden, 01069 Dresden, Germany (e-mail: paul.schwarzbach@tu-dresden.de, oliver.michler@tu-dresden.de).

R.W. is with ZigPos GmbH, Räcknitzhöhe 35a, 01217 Dresden, Germany (e-mail: richard.weber@zigpos.com).

With the emerge of the Internet of Things (IoT), the market penetration of connected and smart devices is rapidly increasing [2]. For IoT associated tasks, wireless systems generally provide cheap, retrofitable, and energy efficient technologies. In addition to the informational interconnection of people, devices, and processes, spatial relations between transceivers can be obtained by analyzing wireless signal characteristics. Next to the already mentioned application fields, wirelessly-aided social distancing and contact tracing has also received research attention with the rise of COVID-19 [3].

In order to evaluate the localization capacities of wireless transmission systems, extensive, costly, and time-consuming field trials have to be performed, intended to ensure application proximity of the given data as well as to neglect specific hardware influences or effects of the propagation environment. Especially for applications with limited access like transportation or industrial environments (e.g. aircraft cabins or railway domain), this quickly becomes the bottleneck for validating localization systems.

Therefore, an efficient simulation of the input data based on an assessment of observable types of errors as well as their magnitude can drastically increase the value chain of IPS. Therefore, this paper discusses and quantifies observable error types for wireless IPS, based on a real-world measurement campaign using a commercial Real-time Location System (RTLS) based on Ultra-Wideband (UWB). Thereupon, a distance domain, sampling-based ranging simulation is presented. While this work focuses on UWB ranging, the underlying effects can also be extended to different technologies and radio-based distance measurement principles.

A. High-precision IPS based on Ultra-Wideband

For high-precision localization applications, a go-to technology is UWB, based on IEEE 802.15.4a. Spatial information between devices using UWB are often derived by utilizing time of flight measurements [4]. When compared to other wireless location-enabling technologies, UWB offers a variety of advantages, due to available bandwidths of at least 500 MHz in frequency bands ranging from 3.1 GHz to 10.6 GHz, as originally specified by the FCC. [5]. This allows robustness against signal multipath interferences and jamming [6].

Frequent application fields for UWB-based positioning systems include unmanned aerial vehicle tracking [7], navigation for automated guided vehicle [8], health care body wearable sensors [9] or logistics respectively industrial fields [10], where UWB provides a reasonable trade-off between accuracy, costs, and energy consumption.

For the implementation and validation of IPS-related work, such as measurement methods, localization approaches, data fusion concepts, and scope of application, a realistic and application-oriented assessment of suitable quality and performance metrics is essential. However, this requires corresponding data as a necessary prerequisite, where data acquisition approaches can be distinguished between:

1) Self-contained data surveying and available datasets:

The first category uses real-world data obtained in certain scenarios. At first, self-contained data surveying describes the collection of data, which is the predominant approach for assessing localization approaches. However, the traceability, further use by other researchers and the comparison of results is limited. In addition, a suitable and often costly reference system is required, especially in dynamic scenarios.

In order to overcome the drawbacks of self-surveying, another approach is to use available datasets. For UWB IPS, a variety of these exist in literature, e.g. [11]–[13]. On the downside, a dataset is never fully comprehensive with regards to possible applications, environments, associated interferences, or error influences (e.g. noise, multipath errors, constellation).

2) Simulation and synthetic generation of data:

Supplementary to utilizing real-world data, another validation method is the simulation of observation data. Generally, the synthetic generation of measurement data provides a variety of advantages, especially for early-stage validation. Among others, these include: time-, work-, and staff efficiency, the highest degree of flexibility in terms of geometric constellation, equipment, and ultimately errors types caused by environmental

influences. However, the proximity to reality of the generated data is crucial. Concurrently, a reasonable computational complexity, which does not hinder the advantages of this approach, needs to be maintained. Another advantage is the capability to evaluate the performance under ideal conditions. To further boost the realism of synthetic data, so-called semi-real data, which is based on both theoretical stipulations and surveyed quantities, can be exploited.

B. Contribution and Structure of this Paper

The main contribution of this paper is the presentation of a semi-empirical and lightweight simulation approach for wireless distance measurements for IPS. This includes a statistical evaluation and error modeling for UWB distance measurements in a demanding indoor scenarios and a simulation procedure, which individually considers error types.

In order to achieve this, we discuss measurement uncertainties as well as radio propagation effects and other influencing characteristics, which all lead to measurement deviations. Our contribution uses these theoretical fundamentals, combined with an empirical measurement campaign and a statistical evaluation of the appearing measurement residuals. The presented measurement campaign and used hardware provide insight on propagation and ranging characteristics in a challenging indoor environment using state-of-the-art UWB modules.

Previous works in the field of stochastic UWB multipath fitting apply the usage of skew-t [14] and gamma [15] distributions or Gaussian mixture models [16]. In contrary, our work suggests the fitting of a log normal distribution for these error components, which can straight forwardly be derived from a Maximum Likelihood (ML) estimation given a dataset, while allowing a precise modeling of the right-skewed shape of the non-line-of-sight (NLOS) and multipath error components.

Subsequently, a derived simulation approach is presented, allowing customized parameterization with respect to utilized hardware, technologies, and propagation environments. In order to achieve this, a probability-based simulation approach, where each individual error influence can be simulated and tuned, is proposed. This provides a low-complexity implementation with manageable parameter settings, while covering the impacts of potential error sources within the measurement process. While this work focuses on the UWB technology, general propagation effects are also transferable to different radio-based ranging approaches. Concluding, this allows a custom parameter tuning of noisiness, multipath occurrence, and error magnitude and therefore provides a comprehensive basis for IPS validation for various application fields.

The paper is structured as follows: After the introduction (Sec. I), we discuss measurement uncertainties, possible error influences for radio-based ranging and suitable accuracy metrics for wireless positioning systems in Sec. II. In Sec. III we present our measurement setup and the surveyed UWB dataset. Additionally, we perform a statistical evaluation of the distance errors. Consequently, Sec. IV describes the synthetic generation of wireless distance measurements with respect to error influences and the obtained measurement residuals. Finally, our contribution ends with conclusions in Sec. V.

II. RADIO-BASED DISTANCE MEASUREMENTS

A. Accuracy and Measurement Uncertainties

The quality of the observations is a measure for the sensory perception of reality and has a direct influence on the performance metrics of IPS, as observations are used as input data. The most important parameters of the observation quality are accuracy and availability.

In order to quantitatively evaluate the observation data, a suitable quality measure needs to be defined. Generally, the measure accuracy describes the degree of conformity between a reference value $\mathcal{F}^*$ and observed, measured or estimated values $\mathcal{F}$, whereby further terms such as correctness and precision are often derived from accuracy [17]. In contrast, the deviation of the two describes the error $\mathcal{R}$. Unlike, errors in position domain, which are typically expressed utilizing the root mean square error (RMSE) or the mean absolute error (MAE) [18], the distance measurement error (also referred to as measurement residual) is expressed as $\mathcal{R} = \mathcal{F} - \mathcal{F}^*$ [19].

Random variables are suitable to describe and evaluate measurement data. They characterize a physically or logically defined quantity, to which a certain measurement value r can be assigned as an unpredictable randomly observed result by executing an arbitrarily repeatable random experiment. The result is called observation or realization of the random variable S in the form of real numbers within a certain interval, where a sufficient number of realizations as representative sample allows a conclusion on the population of the observations.

The irregularities within a measurement data set are mainly responsible for the occurring inaccuracies and uncertainties. Those include grossly erroneous or contaminated observations as well as hidden disturbance effects. Such effects rarely remain constant and vary depending on time and situation. According to [20], the resulting inaccuracies can essentially be attributed to two types of errors:

- Random errors (measurement noise) characterize the criterion precision of the accuracy. Such errors are residual deviations which cannot be described in terms of amount or direction due to inadequacies of the measuring method or the nature of the phenomenon, e.g. noise of measurement devices.
- Systematic errors (bias, offset) characterize the criterion trueness of the accuracy. Such errors are deviations (e.g. constant measurement offsets), which have functionally describable causes.

According to the Central Limit Theorem, random measurement uncertainties are composed of a large number of so-called elementary errors, which is the basis for modeling a Cumulative Distribution Function (CDF) and for the application of corresponding statistical methods. The widely accepted model for the statistical distribution of measured values is the Gaussian normal distribution. For example, L measurements $r_i \in S$ with $i = 1, 2, \dots, L$ were performed and assigned to the measurement vector $\mathbf{r} = (r_1 \ r_2 \ \dots \ r_L)^T$. Assuming normally distributed measured values, these can be described by two parameters, the arithmetic mean

$$E(S) \approx \bar{r} = \frac{1}{L} \sum_i r_i \quad (1)$$

as ML estimation [21] of the usually unknown true value r^* and the empirical measurement variance

$$\begin{aligned} \text{Var}(S) \approx s^2 &= \frac{1}{L-1} \sum_i (r_i - \bar{r})^2 \\ &= \frac{1}{L-1} (\mathbf{r} - \bar{r} \mathbb{1}_L)^T (\mathbf{r} - \bar{r} \mathbb{1}_L) \end{aligned} \quad (2)$$

where $\mathbb{1}_L$ denotes the single-column vector of length L obtained from the ML-estimation of the theoretical variance σ^2 . From the summary observation of apparently irregular single phenomena, the observations $r_i \sim N(r^*, \sigma^2)$, regularities for the mass phenomenon in question are determined. If such regularities are found, a statement of accuracy for the single phenomenon can be deduced.

Systematic measurement deviations, on the other hand, are typically caused by unknown influences. Hence, even with a theoretically infinite number of observations, the resulting estimated value $\bar{r}$ still differs from the true value r^* . In the context of wireless ranging, typical systematic distance measurement errors are often associated with multipath propagation. By carefully analyzing the measurement process, such effects can be controlled and at least partially eliminated.

Lastly, gross errors (outliers) represent serious deviations between the observed and the true values as a result of process respectively hardware defects or target confusions.

B. Sources of Errors

According to [22], the measurement errors treated in Sec. II-A can be divided into model errors (system dynamics and concentric propagation), calculation errors (memory limitation and resolution of data types), as well as intrinsic and extrinsic setup errors. Intrinsic setup errors refer to hardware-related influencing factors such as various signal strength fluctuations, circuit tolerances, hardware calibration during production, the accuracy of clocks or clock generators such as quartz oscillators, the hardware-related constant measurement offset, and the stability of the power supply. Typically, intrinsic error are empirically quantifiable and given by the manufacturer.

Extrinsic setup errors on the other hand result from antenna orientation and the underlying radio transmission channel. Among others, the following characteristic effects typically have a negative impact on the signal propagation:

- Reflection: Signal waves are reflected at obstacles according to the law of reflection, which often leads to multipath propagation and possibly polarization changes.
- Diffraction: Certain obstacles like edges or gratings lead to wave deflection and thus to a changed signal propagation direction.
- Shadowing: Due to the obstruction of the direct path between transmitter and receiver, shadowing can lead to a variety of effects. Generally, signals are attenuated or completely blocked off. In addition, time delays during the penetration of objects occur, if their expansion is significantly larger than the signals bandwidth.

Further effects include refraction, penetration through objects, scattering due to rough surface reflections, frequency shifts caused by relative movements between transmitter and receiver, as well as interferences and superpositions, causing signal fading.

Subsequently, we want to emphasize that both intrinsic and extrinsic error types are highly dependent on both the applied hardware and the examined measurement environment. Whereby in the context of wireless ranging the latter includes both the geometric propagation environment (e.g. presence of obstacles) and the electromagnetic environment (e.g. presence of radio interferences). The impacts of the respective influences and their potential effects on the conducted simulation approach will further be discussed in Sec. IV.

C. Ranging Simulation

The basis for synthetically generating spatial relations between communication devices using the physical properties of the transmitted signals, generally is the modeling of the signal propagation itself. For this task, the main trade-off is between the accuracy and computational complexity of the propagation model. Modeling and simulating wireless transmission systems is typically achieved by one of the following approaches:

1) *Numerical approaches*: numerically solve Maxwell's equations [23], generally leading to the highest computational complexity among the mentioned approaches. Also numerically solving the electromagnetic field in demanding real-world environments is a challenging task.

2) *Ray tracing approaches*: describe possible propagation paths of an emitted signal, obtained from a Maxwell high-frequency approximation, where rays are used to describe several propagation phenomena [24]. Ray tracing methods provide a comparably accurate simulation for signal propagation, but they also require a variety of additional information about the propagation environment, such as an environmental model consisting of a geometry and an electromagnetic model. Furthermore, physical layer properties as well as specific antenna need to be included [25].

3) *Empirical approaches*: are derived from the evaluation of available real measurement data. Generally, the validity of the obtained findings is limited to the conditions of the underlying measurements, inter alia including utilized hardware, geometric constellations, weather conditions, and propagation environments. However, empirical approaches can be enhanced by taking underlying physical effects into account, leading semi-empirical approaches.

Depending on the applied measurement principle for obtaining spatial information between radio devices as well as the targeted use case, abstraction level, and complexity, different simulation approaches can be beneficial. As the conducted works aims at providing a lightweight, yet flexible and adaptable ranging data generation, a semi-empirical approach is presented. In accordance with the aforementioned propagation phenomena and the resulting signal reception types, the conducted and sampling-based approach only uses the individual and distance-dependent occurrence probabilities for each propagation type. Lastly, the corresponding error magnitudes can then be obtained from underlying measurements.

III. EMPIRICAL DATA SURVEY

A. Hardware Setup

The presented measurement campaign was performed using the *ZigPos* UWB RTLS [26], which is a standardized localization systems for various types of LBS applications. It consists of three main components: A small-form-factor barebone computer kit gateway for network organization, several stationary anchors used as fixed points for localization and mobile tags or badges (cf. Fig.1), whose location is of interest. On the one hand, a 2.4 GHz transceiver, which is mainly used for network organization and low energy data exchange within the network. On the other hand, the UWB front end as mentioned above, which is used for high accuracy localization tasks.

For the latter, the associated hardware is based on the IEEE 802.15.4a compliant *DecaWave* DW1000 IC transceiver [27], which applies Two-Way-Ranging. The measurements were conducted at a center frequency of 6.5 GHz using a bandwidth of 0.5 GHz. The applied measurement rate was set to 10 Hz.

Fig. 1: *ZigPos* RTLS UWB Badge

B. Measurement Setup

The measurement campaign was carried out in an indoor environment, more specifically in a bus depot located in Dresden, Germany. As depicted in Fig. 2, this scenario provides a common indoor positioning environment for transportation applications. A RTLS in this environment can serve as a key enabler for an intelligent operations and fleet management. In addition, the bus depot represents an heterogeneous environment with regards to radio-based propagation, as it contains different visibility conditions (line-of-sight (LOS), obstructed line-of-sight, NLOS) as well as different materials, like concrete, (partially metallized) glass, or bus chassis.

To realize a significant and holistic assessment of the environment and to enhance the following simulation procedure, a statistically representative number of approximately 100.000 distance measurements in spatially varying constellations were surveyed. This reduces hardware-related factors influencing the measurement accuracy and thus simplifies both observation and analysis of significant effects in the underlying radio transmission channel. Additionally, spatially correlated effects, like poor reception conditions for specific constellations, are evenly represented.

For this task, a set of seven stationary anchors were installed, aiming to uniformly cover the area of investigation. The measuring points for the badge, which are also shown

Fig. 2: Traffic environment for real-world UWB data collection: (a) and (b) show a parking hall of the bus depot, which represents a typical, challenging, and multipath-rich environment, including obstacles and shadowing objects, while also indicating a direct signal path (green) and a reflected, indirect path (red). In addition, (c) depicts the locations of the stationary anchors (green) and the quasi-static measuring points of the mobile badge (red) for the presented measurement campaign.

in Fig. 2, are varying throughout the entire available area. The reference distances were determined by the Euclidean distances between the positions of the mobile node and the anchors. The distances between the anchors and badge positions range up to around 100 m. This generates as many different phenomena in the transmission channel as possible, allowing a holistic representative data basis for statistical evaluation.

C. Observation Data

Regarding the conducted measurement campaign, the observation data includes measurement disturbances as well as valid measurements with Gaussian errors, multipath errors, and outliers. Collectively, these are presented in Fig. 3a, which depicts the measurement errors over the true distances. The color intensity indicates the frequency of the errors.

The measured Gaussian errors (in the area of the yellow horizontal lines) are quasi mean-free and characteristic for UWB measurements with respect to their magnitude. The outliers shown in Fig. 3a, which are located in the area below and far above the yellow straight line, represent measurement outliers, which from experience are due to noise signals detected by the UWB transceiver and due to their rarity are not considered within the subsequent statistical evaluation.

The measured multipath errors (in the area above the yellow straight line) are all positive and depend on the structure of the obstacle objects, the building geometry and the respective node positions, but less on the used bandwidth or the resulting measurement noise [28]. The errors range up to about 20 m, with smaller observed values for shorter distances.

Furthermore, Fig. 3b shows that the multipath error fraction remains relatively constant for all distances, but increases with respect to all valid observations (black, blue, and green lines). The relative share of measurement disturbances (yellow line) increases as well, which is mainly due to the correlation of distance and probability for blocking all potential paths. Moreover, we refrained from testing the stochastic independence of temporally successive measurements, since the few locally observed dependencies do not justify disproportionately complicating the modeling by including such effects.

D. Statistical Residual Evaluation

Propagation paths for empirical modeling of distance measurements can be distinguished as follows [29]–[31]:

- measurement of the direct path,
- measurement of the shadowed direct path in the NLOS channel, and
- measurement of multipath due to blockage of the direct path in the NLOS channel.

In this work, the first two cases are modeled by a normal distribution respectively a log normal distribution, which may not necessarily be mean-free. This is especially true as a result of signal propagation delays arising from possible shadowing in the second case.

As shown in the Graphical Abstract, the resulting Gaussian probability density (yellow line) for the presented measurement data, whose errors are summarized as an adequate simplification over all distances, can be generated with the measurement variance $\sigma_{hw}^2 = (0.33 \text{ m})^2$. This is determined from the ML estimation given in (2). However, before applying this estimation, the corresponding error type needs to be isolated from the mixed residual distribution in order to avoid bias due to multipath errors and outliers. Given the UWB-based ranging principle at hand, negative measurement residuals are not caused by the propagation environment, but rather by measurement noise as discussed in Sec. II-B. Due to the symmetry of the Gaussian distribution, which is commonly used to describe noise, σ_{hw}^2 can be estimated from the negative residuals and then also be applied to the positive residuals.

Subsequently, measurement residuals caused by NLOS and multipath propagation can be isolated from the dataset by subtracting the fraction of the normally distributed noise errors given the just estimated variance. A distinction between NLOS measurements and outliers is achieved by empirically setting the residual interval to $[-3\sigma_{hw}; 20 \text{ m}]$. The lower bound considers general noise effects also within the NLOS channel and the upper bound is set due to the properties of UWB. Fig. 3 reveals the distance dependency of NLOS measurements, however as UWB uses comparably low transmission powers, the maximum range for UWB signals is spatially

Fig. 3: Evaluation of real sensor data and model development: (a) shows the observed measurement errors and their relative frequency over the true distance (effective information content up to 70 m), which are separated by error components in (b).

limited, especially in complex indoor and NLOS scenarios as additional reflection attenuation occurs. Therefore, residuals higher than the threshold of 20 m are considered as outliers.

After isolating the NLOS and multipath error components, the subsequent modeling of the error R is performed using a right-skewed log normal distribution [32], where the skewness of the log normal distribution depends on the diversity of the NLOS channel.

$$\mathcal{LN}(R; \bar{R}_{mp}, \sigma_{mp}^2) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{mp}R} \exp\left(-\frac{(\log R - \bar{R}_{mp})^2}{2\sigma_{mp}^2}\right) & \text{if } R > 0 \\ 0 & \text{else.} \end{cases} \quad (3)$$

The distribution parameters $\bar{R}_{mp}$ and σ_{mp} are obtained based on the ML estimation in (1) and (2) and the substitution of the measured errors R_i by $\log R_i$. This yields the probability density for the present measured data with $\bar{R}_{mp} = 0.8$ m and $\sigma_{mp} = 1.07$ m, where [28] mentions comparable values. Finally, the sum of the Gaussian and log normal distributions results in the overall model as a mixed distribution. In addition, there is the possibility to check the parameter quality by means of Chi-square, Anderson-Darling or Cramér-von Mises distribution tests as well as to perform the modeling alternatively by histogram or kernel density estimators.

IV. GENERATION OF DISTANCE MEASUREMENTS

The simulation of a distance measurements can be formulated as [29]:

$$r = \begin{cases} d + \epsilon & \text{if } p > \frac{d}{d_{\max}} \\ \emptyset & \text{else,} \end{cases} \quad (4)$$

based on the reference distance $d = \|X_a - X_m\|_2$, which describes the Euclidean distance between the simulated position of the anchor X_a and the mobile object X_m . Furthermore, the sporadic measurement perturbations shown in Fig. 3b with probability $p \sim U(0, 1)$ linearly dependent on d and the

maximum possible measurement range $d_{\max} = 100$ m. Both are empirically modeled as a Bernoulli experiment, where $d_{\max}$ is based on the surveyed data in Sec. III.

For the generation of the error variable ϵ , two possibilities exist in general. In the simplest case, r can be simulated according to [33] by bootstrapping, i.e. sampling $\epsilon \sim P(R^{[1:L]})$ from the L real measurement errors $R^{[1:L]}$. Their occurrence is defined as equally likely or distance dependent. Although the number and distribution of the given measured data are shown to be suitable for this purpose, this approach is not pursued further, since the simulation would otherwise be limited exclusively to the measured sensor-related errors.

The error variable favored within the simulation procedure

$$\epsilon = \begin{cases} \epsilon_{mp} & \text{if } p_\epsilon > 0.8 - 0.3 \frac{d}{d_{\max}}, \epsilon_{mp} < d \\ \epsilon_{out} & \text{if } p_\epsilon < p_{out} \\ \epsilon_{hw} & \text{else} \end{cases} \quad (5)$$

contains the sensor-related errors in form of the white Gaussian variables $\epsilon_{hw} \sim \mathcal{WN}(0, \sigma^2)$ and the variables $\epsilon_{out} \sim \mathcal{U}(-d, d_{\max} - d)$ for uniformly distributed outliers. Here, the measurement variance σ^2 as well as the outlier probability p_{out} are modeled variably with the claim to perform the evaluation as generally valid as possible, i.e. independent of a concrete radio technology and its characteristic measurement noise.

Furthermore, (5) consists of the environmental error variable ϵ_{mp} . This simulates a multipath instead of the direct path and thus represents the much more limiting variable. Thereby, according to the real sensor data visualized in Fig. 3a, longer multipath paths are empirically prevented, especially for shorter d . Thus, for $d \rightarrow 0$ no multipath results and for $d \rightarrow d_{\max}$ a multipath results in about 50% of all cases, with [28] giving similar values in this respect. The occurrence of multipath paths linearly dependent on d and $d_{\max}$ as well as the outliers is modeled with probability $p_\epsilon \sim \mathcal{U}(0, 1)$ again as a Bernoulli experiment. The final parameter setting was determined based on the presented statistical evaluation.

Fig. 4: Synthetic generation of observation data. (a) shows the simulated measurement errors as a function of anchor to badge distance, whose relative frequency separated by the individual error components is depicted in (b).

In addition, partial anchor- or direction-dependent biases of the multipath probability could be included, which would simulate local stochastic dependencies of temporally successive measurements and thus slightly increase the error potential.

The error variable $\epsilon_{\text{mp}} \sim \mathcal{LN}(\bar{R}_{\text{mp}}, \sigma_{\text{mp}}^2)$, which follows the log normal distribution as given in (3) with model parameters $\bar{R}_{\text{mp}}$ and σ_{mp} , is given by

$$\epsilon_{\text{mp}} = \exp(\bar{R}_{\text{mp}} + \sigma_{\text{mp}} \epsilon_{\text{mpWN}}), \quad (6)$$

with $\epsilon_{\text{mpWN}} \sim \mathcal{WN}(0, 1)$, which is comparatively close to reality based on the model development performed for the traffic environment at hand. Sampling of ϵ_{mpWN} can be implemented according to [34] and [35] based on $\mathcal{U}(0, 1)$ by the Box-Muller method, the Marsaglia-Polar method, the Rule of Twelve, or the Ziggurat algorithm. Alternatively, the inversion method (with the simulation lemma realized by series expansion), which is applicable for arbitrarily distributed samples, as well as the rejection method, which is also called acceptance-rejection method, are suitable. Furthermore, it is conceivable to consider only positive values as resulting valid distance values r in (4) for plausibility reasons.

To underline the presented synthetic distance measurement generation approach, a simulation example is provided, aiming to emulate the surveyed data set. Fig. 4a shows an exemplary simulation result in the form of errors of semi-real distances, the number of which decreases over distance as a result of simulated measurement disturbances. Here, the relative frequency of the generated error components shown in Fig. 4b is comparable to that shown in Fig. 3b.

V. CONCLUSIONS

High-precision applications are rapidly growing, allowing a widespread of LBS based on a capable RTLS. Especially in indoor environments, a go-to technology for these tasks is UWB, as it provides comparably accurate distance measurements even in demanding scenarios. For RTLS realization, excessive data-based evaluation is typically required. However, the acquisition of this data is time-consuming and expensive.

Therefore, a common approach is the use of synthetically generated data, where simulation methods differ in their accuracy and complexity. For lightweight implementation, this paper presented a semi-empirical simulation for wireless distance measurements based on surveyed real-world data.

The synthetic data generation is based on the visibility conditions in the propagation environment, which are represented by individual occurrence probabilities for line-of-sight and non-line-of-sight propagation, as well as outliers. The corresponding error magnitudes, e.g. maximum distance, Gaussian noise, or right-skewness of the multipath distribution, can be obtained from the presented measurement campaign. For different scenarios or propagation environments, e.g. for dense multipath environments, the maximum transmission range might reduce, while the share of multipath propagation increases. Given the presented approach, such scenarios can easily be adapted by simply parameter tuning.

Summarizing, the proposed simulation approach provides a flexible parameter tuning for different scenarios and therefore enables cost-efficient as well as scalable localization method evaluation, while contributing synthetic and application-oriented data.

ACKNOWLEDGMENT

The authors would like to thank ZigPos GmbH Dresden.

REFERENCES

- [1] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.
- [2] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, "Internet of things (iot) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5g-iot scenarios," *IEEE Access*, vol. 8, pp. 23 022–23 040, 2020.
- [3] P. Schwarzbach, J. Engelbrecht, A. Michler, M. Schultz, and O. Michler, "Evaluation of technology-supported distance measuring to ensure safe aircraft boarding during COVID-19 pandemic," *Sustainability*, vol. 12, no. 20, p. 8724, Oct. 2020. [Online]. Available: <https://doi.org/10.3390/su12208724>

- [4] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrami, M. Al-Ammar, and H. Al-Khalifa, "Ultra wideband indoor positioning technologies: Analysis and recent advances," *Sensors*, vol. 16, no. 5, p. 707, May 2016. [Online]. Available: <https://doi.org/10.3390/s16050707>
- [5] IEEE, "Ieee standard for local and metropolitan area networks—part 15.4: Low-rate wireless personal area networks (lr-wpans)," *IEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006)*, pp. 1–314, 2011.
- [6] M. Segura, V. Mut, and C. Sisterna, "Ultra wideband indoor navigation system," *IET Radar, Sonar Navigation*, vol. 6, no. 5, pp. 402–411, 2012.
- [7] W. You, F. Li, L. Liao, and M. Huang, "Data fusion of uwb and imu based on unscented kalman filter for indoor localization of quadrotor uav," *IEEE Access*, vol. 8, pp. 64 971–64 981, 2020.
- [8] D. Shi, H. Mi, E. G. Collins, and J. Wu, "An indoor low-cost and high-accuracy localization approach for agvs," *IEEE Access*, vol. 8, pp. 50 085–50 090, 2020.
- [9] T. Otim, A. Bahillo, L. E. Díez, P. Lopez-Iturri, and F. Falcone, "Towards sub-meter level uwb indoor localization using body wearable sensors," *IEEE Access*, vol. 8, pp. 178 886–178 899, 2020.
- [10] M. Delamare, R. Bouteau, X. Savatier, and N. Iriart, "Static and dynamic evaluation of an UWB localization system for industrial applications," *Sci*, vol. 2, no. 2, p. 23, Apr. 2020. [Online]. Available: <https://doi.org/10.3390/sci2020023>
- [11] 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," 2020.
- [12] K. Minne, "Experimental evaluation of uwb indoor positioning for indoor track cycling," 2019. [Online]. Available: <https://data.mendeley.com/datasets/fkhfjfsprk/1>
- [13] V. B. C. J. E. J. A. García-Naya, "Nlos identification and mitigation using low-cost uwb devices," 2019. [Online]. Available: <https://dx.doi.org/10.21227/swz9-y281>
- [14] H. Nurminen, T. Ardesiri, R. Piché, and F. Gustafsson, "Robust inference for state-space models with skewed measurement noise," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1898–1902, 2015.
- [15] A. R. Jiménez Ruiz and F. Seco Granja, "Comparing ubisense, bespoon, and decawave uwb location systems: Indoor performance analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 2106–2117, 2017.
- [16] A. Prorok, L. Gonon, and A. Martinoli, "Online model estimation of ultra-wideband tdoa measurements for mobile robot localization," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 807–814.
- [17] ISO Central Secretary, "ISO 5725-1:1994(E) - Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions," International Organization for Standardization, Geneva, CH, Standard, 1994.
- [18] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [19] ISO Central Secretary, "ISO 3534-1:2006(E) - Statistics - Vocabulary and symbols - Part 1: General statistical terms and terms used in probability," International Organization for Standardization, Geneva, CH, Standard, 2006.
- [20] J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, ser. A series of books in physics. University Science Books, 1997.
- [21] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.
- [22] A. Savvides, W. L. Garber, R. L. Moses, and M. B. Srivastava, "An analysis of error inducing parameters in multihop sensor node localization," *IEEE Transactions on Mobile Computing*, vol. 4, no. 6, pp. 567–577, 2005.
- [23] F. Gustrau, *EM Modeling of Antennas and RF Components for Wireless Communication Systems*. Berlin, Heidelberg: Springer-Verlag, 2009.
- [24] M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge University Press, 1999.
- [25] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.
- [26] ZigPos, "Zigpos documentation," <https://products.zigpos.com/>, 2020.
- [27] Decawave, "Decawave software and documentation," <https://www.decawave.com/software/>, 2020.
- [28] N. A. Alsindi, B. Alavi, and K. Pahlavan, "Measurement and modeling of ultrawideband toa-based ranging in indoor multipath environments," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1046–1058, 2009.
- [29] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Transactions on wireless communications*, vol. 5, no. 3, pp. 672–681, 2006.
- [30] O. Michler, R. Weber, and G. Förster, "Model-based and empirical performance analyses for passenger positioning algorithms in a specific bus cabin environment," in *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2015, pp. 200–208.
- [31] R. Weber, U. Gosda, O. Michler, and J. Ringel, "Wsn-based passenger localization in severe nlos environments using sdpr," in *Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on*. IEEE, 2013, pp. 1–7.
- [32] A. Prorok, P. Tomé, and A. Martinoli, "Accommodation of nlos for ultra-wideband tdoa localization in single- and multi-robot systems," in *2011 International Conference on Indoor Positioning and Indoor Navigation*, 2011, pp. 1–9.
- [33] B. Efron, "Bootstrap methods: another look at the jackknife," in *Breakthroughs in statistics*. Springer, 1992, pp. 569–593.
- [34] L. Devroye, *Non-Uniform Random Variate Generation*. Springer, 1986.
- [35] D. E. Knuth, *The art of computer programming: sorting and searching*. Pearson Education, 1998, vol. 3.

Paul Schwarzbach was born in Dresden in 1990. He received his Diploma in Transport Engineering from Technische Universität Dresden, Germany in 2016. Since 2016, he is a Research Assistant and Ph.D. student at Technische Universität Dresden at the Institute of Traffic Telematics, where his research interest focuses on radio-based and hybrid localization systems as well as state estimation for vehicular applications.

Richard Weber was born in Annaberg-Buchholz, Germany in 1986. He received his Diploma in Transport Engineering from Technische Universität Dresden in 2011. Since then, he worked as a Research Assistant at Technische Universität Dresden, Germany, where he received his Ph.D. degree in 2020. Ever since, he is a senior software engineer at the IT venture ZigPos GmbH, Dresden, Germany focusing on RTLS and IoT systems, with special emphasis on UWB.

Oliver Michler was born in Dresden, Germany in 1967. He received his Diploma in Electrical Engineering from Technische Universität Dresden, Germany in 1993. From 1993 to 1997, he was a Research Assistant with the Technische Universität Dresden and received his Ph.D. in 1999. Until 2008, he was a project manager at Video Audio Design GmbH, Research Associate at Fraunhofer Institute of Transportation and Infrastructure Systems and a Professor for Signal Processing and Electronic Measurement Techniques at University of Applied Sciences Dresden. Ever since, he became a full Professor for Transport Systems Information Technology at Technische Universität Dresden, where he also became the Director of the Institute of Traffic Telematics in 2019. In addition, he is a scientific advisory board member of several international conferences, organizations and start-up businesses.