

Ubiquitous Indoor Fine-Grained Positioning and Tracking: A Channel Response Perspective

Chenglong Li^{1,2,*}, Emmeric Tanghe², Sofie Pollin³ and Wout Joseph²

¹College of Electronic Science and Technology, National University of Defense Technology, 410073 Changsha, China

²WAVES group, Department of Information Technology, Ghent University-imec, 9052 Ghent, Belgium

³Wavecore, Department of Electrical Engineering, KU Leuven, 3001 Leuven, Belgium

Abstract

The future of location-aided applications is shaped by the ubiquity of Internet-of-Things devices. As an increasing amount of commercial off-the-shelf radio devices support channel response collection, it is possible to achieve fine-grained position estimation at a relatively low cost. In this paper, we focus on the channel response-based positioning and tracking for various applications. We first give an overview of the state of the art (SOTA) of channel response-enabled localization, which is further classified into two categories, *i.e.*, device-based and contact-free schemes. A taxonomy for these complementary approaches is provided concerning the involved techniques. Then, we present a micro-benchmark of channel response-based direct positioning and tracking for both device-based and contact-free schemes. Finally, some practical issues for real-world applications and future research opportunities are pointed out.

Keywords

Indoor localization, positioning and tracking, channel response, machine learning, multi-path propagation

1. Introduction

Location acquisition has obtained and continues to attract extensive attention due to the rapid proliferation of location-based services and the ubiquitous radio devices. Location awareness can help to automate and optimize the flows of many vertical applications [1, 2], such as automatic inventory management, smart logistics, human-machine interaction, intelligent factory, etc., as shown in Fig. 1. But for Global Navigation Satellite System-denied or indoor environments, accurate location acquisition is still challenging despite a range of radio frequency technologies and techniques that have been proposed and deployed during the past decades [3, 4].

Conventional localization solutions generally exploit the distance, angle, or Doppler information to estimate the location and trajectory. The localization accuracy highly depends on the corresponding ranging, angular, and velocity resolution. When relying on time-of-flight (ToF), the range resolution is directly related to the signal bandwidth as more bandwidth ensures

Proceedings of the Work-in-Progress Papers at the 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN-WiP 2023), September 25 - 28, 2023, Nuremberg, Germany


* Editors: Susanna Kaiser, Norbert Franke, and Christopher Mutschler

* Corresponding author.

✉ chenglong.li@nudt.edu.cn (C. Li); emmeric.tanghe@ugent.be (E. Tanghe); sofie.pollin@kuleuven.be (S. Pollin); wout.joseph@ugent.be (W. Joseph)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

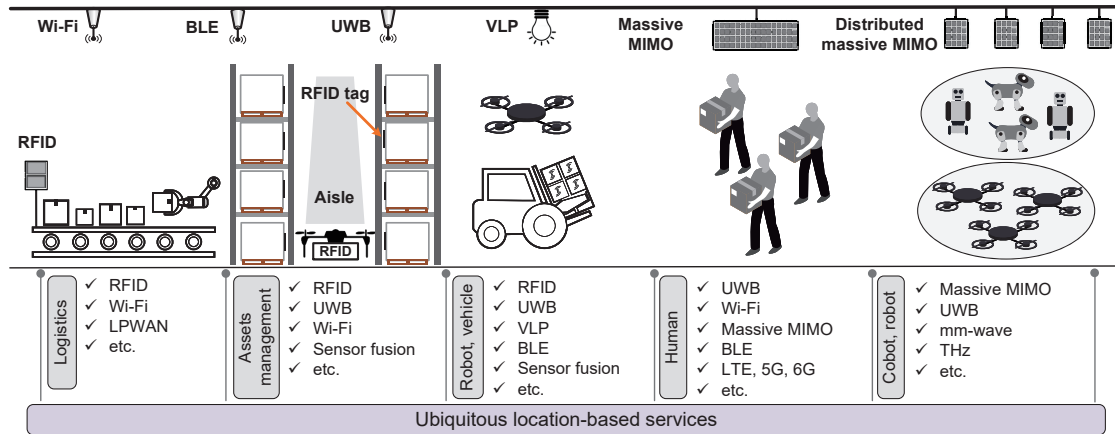


Figure 1: Pervasive radio devices enabling various location-aided applications (RFID: radio frequency identification. LPWAN: low-power wide-area network. UWB: ultra-wideband. VLP: visible light positioning. BLE: Bluetooth low energy. LTE: long-term evolution. 5G: fifth-generation. 6G: sixth-generation. mm-wave: millimeter wave. THz: Terahertz.)

finer-grained ToF and hence ranging. Angular resolution is increased when relying on a large antenna array with more antenna elements. However, it is not practical for bandwidth-limited systems, e.g., sub-6 GHz, due to the relatively large wavelength and thus large antenna sizes and inter-antenna spaces [5, 6]. One potential solution to this problem is to construct the virtual antenna array of a moving antenna, namely built upon the idea of synthetic aperture[7]. Finer velocity resolution is also achieved with a higher frequency as Doppler is inversely proportional to the carrier wavelength.

Channel response has a finer granularity than the conventional signal strength-based metrics, because it can capture small-scale spatial motions via the complex signal (amplitude and phase) changes in the temporal, frequency, and antenna domains. For example, in the case of one-way propagation, half-wavelength distance changes will cause π -rad phase shifts. There is an increasing community, implementing localization systems based on channel response that is available in advanced communication systems [8, 5, 9, 10, 7]. With the advances in radio hardware circuits and signal processing, many commodity radio devices support channel-sounding measurements, relying on channel state information (CSI) [8, 5], channel impulse response (CIR) [9, 10], or signal phase [7]. These physical-layer measurements enable accurate characterization of signal propagation. Compared with the received signal strength indicator (RSSI), channel response provides a much finer-grained metric to describe the smaller scale spatial variations. It makes the human-environment and machine-environment interactions more feasible, enabling various emerging applications.

This paper is dedicated to positioning and tracking for the pervasive location-based services. According to the authors' best knowledge, there is limited work that overviews the state of the art (SOTA) from the perspective of the channel response. We classify the SOTA approaches into learning-based and model-based methods regarding device-based and contact-free positioning and tracking. A further taxonomy is presented based on how these leading-edge algorithms perform with the channel response. We conclude this paper by discussing some practical

concerns and future research opportunities.

2. Channel Response-based Indoor Positioning and Tracking

From the perspective of the radio device, indoor positioning and tracking can in general be classified into two categories, namely device-based and contact-free (*a.k.a.* passive, device-free) schemes, as shown in Fig.2. Device-based positioning and tracking is a cooperative scheme with radio sensors attached to the users or agents that should be tracked. With the help of channel-sounding data, device-based solutions have boosted the positioning accuracy from meter-level to decimetre-level despite limited bandwidth and the limited number of antennas. By contrast, contact-free positioning and tracking aim for the localization of non-cooperative users, which relies on radar-like sensors integrated in the environment to facilitate passive sensing. Again, due to the available channel-sounding data from the commodity devices, it is feasible to achieve positioning and tracking similarly as done for the phased array radar. For instance, we can do range-Doppler-angle processing on the channel response.

Under assumption of a far-field MIMO channel, the channel transfer function at time t and frequency f (in equivalent base-band) can be given by

$$\mathbf{H}(t, f) = \sum_{l=0}^L \alpha_l \mathbf{c}_{\text{Rx}}(\boldsymbol{\Omega}_{\text{Rx},l}) \mathbf{c}_{\text{Tx}}^\top(\boldsymbol{\Omega}_{\text{Tx},l}) e^{-\mathcal{J}2\pi\left(\frac{d_l}{\lambda} + \nu_l t\right)}, \quad (1)$$

where α_l , d_l , ν_l denote the amplitude, distance, and Doppler shift of the l -th multi-path, respectively. λ is the wavelength and L the number of multi-path components. $c_i(\boldsymbol{\Omega}_{i,l})$, ($i = \text{Tx}, \text{Rx}$) is the steering vector at transceiver and $\boldsymbol{\Omega}_{i,l}$ the directional vector. $l = 0$ represents the line-of-sight (LoS) component. For device-based localization, if we can separate LoS from (1), the location of the users can be inferred because $\boldsymbol{\Omega}_{\text{Rx},0}$, $\boldsymbol{\Omega}_{\text{Tx},0}$, τ_0 , and ν_0 are the non-linear function of user's coordinates.

In the case of contact-free localization, the LoS component and the reflections or scatterings from the surroundings should be mitigated because the components associating with the target(s) are desired. Given several moving targets, the channel response of (1) can be rewritten as

$$\mathbf{H}(t, f) = \bar{\mathbf{H}}(\mathbf{f}) + \sum_{\mathbf{j} \in \mathcal{D}} \alpha_{\mathbf{j}} \mathbf{c}_{\text{Rx}}(\boldsymbol{\Omega}_{\text{Rx},\mathbf{j}}) \mathbf{c}_{\text{Tx}}^\top(\boldsymbol{\Omega}_{\text{Tx},\mathbf{j}}) e^{-\mathcal{J}2\pi\left(\frac{d_{\mathbf{j}}}{\lambda} + \nu_{\mathbf{j}} t\right)}, \quad (2)$$

where $\bar{\mathbf{H}}(\mathbf{f})$ is the sum of all static signals to be mitigated. \mathcal{D} denotes the set of dynamic targets. To obtain the locations of the multiple targets, the $\boldsymbol{\Omega}_{\text{Rx},\mathbf{j}}$, $\boldsymbol{\Omega}_{\text{Tx},\mathbf{j}}$, $\tau_{\mathbf{j}}$, and $\nu_{\mathbf{j}}$ should be associated with the corresponding moving target and used for contact-free positioning and tracking.

There are two popular technical routines for channel response-based localization. One is to estimate the distance, angle, or velocity-related metrics first, and then develop conventional positioning and tracking algorithms based on the estimated geometrical metrics. These distance- or angle-related metrics can be estimated from the CSI or CIR via super-resolution parameter estimation algorithms [11, 6], such as multiple signal classification (MUSIC), space alternating generalized expectation-maximization (SAGE), a maximum likelihood parameter estimation framework (RiMAX), etc. The second is to exploit the channel response measurements directly

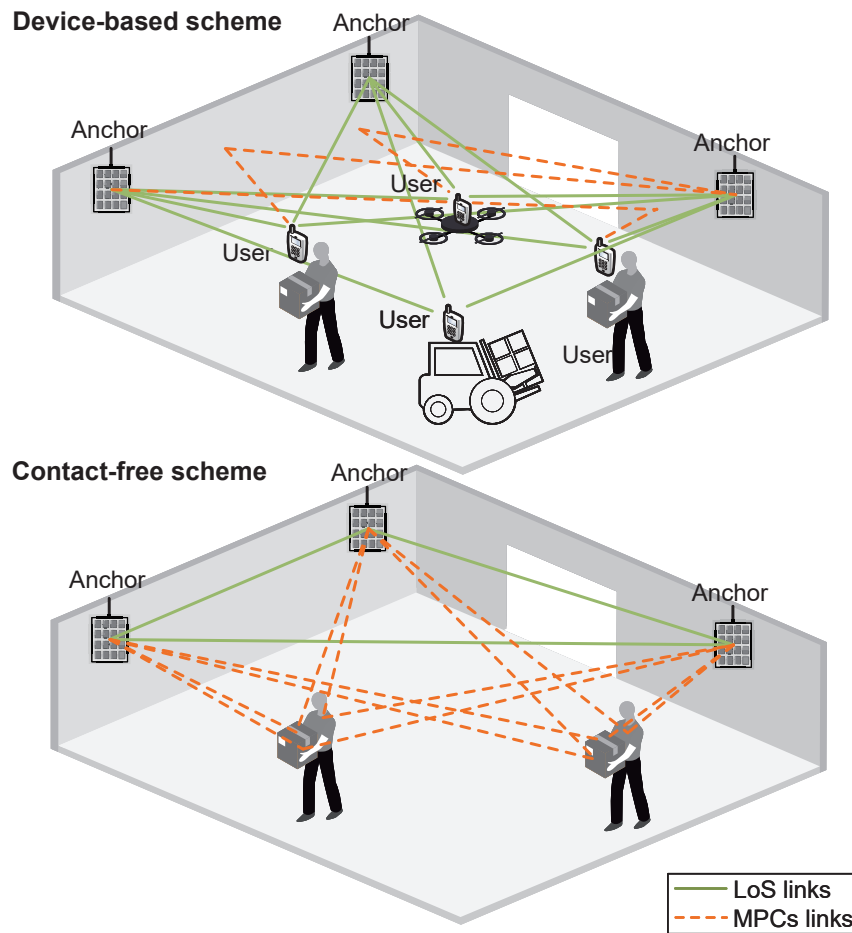


Figure 2: Device-based and contact-free indoor positioning and tracking.

for positioning and tracking purposes [12, 13], which evades the complex estimation of the intermediate metrics. The details of the taxonomy will be given in Section 3.

3. Taxonomy of State-of-the-Art

In this section, we will present the leading-edge channel response-based solutions regarding the device-based and contact-free schemes. The corresponding taxonomy of the channel response-enabling methods is shown in Fig. 3 and will be further elaborated below.

3.1. Device-based Positioning and Tracking

For indoor applications, the device-based schemes rely on sensors or tags attached to the targets, like robots, (manned or unmanned) vehicles, packages, etc. As shown in Fig. 2 (top), the users

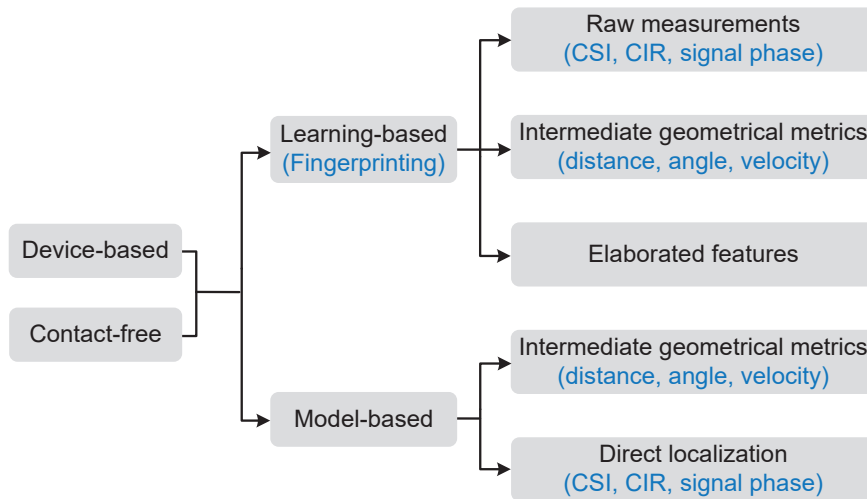


Figure 3: Taxonomy of the channel response-based positioning and tracking.

infer their locations relative to the anchors from the received signal which includes the LoS and multi-path components (MPCs) from the surroundings.

3.1.1. Learning-based Approaches

The learning-based (*a.k.a.*, data-driven) scheme in Fig. 3 generally establishes a non-linear mapping between the measurements and the targeted locations, which is intuitively easy to implement. Note that herein the fingerprinting localization approaches are also included within the learning-based framework because the off-line and on-line phases of fingerprinting are similar to the training and testing procedures, respectively.

Raw measurements: An intuitive positioning approach is to feed the raw channel measurements into a machine learning model directly to infer the coordinates of the users. Due to the pervasive access to Wi-Fi radios, there has been tremendous research focusing on CSI-based Wi-Fi localization in recent years. However, the measured channel response like CSI is a complex signal, and the widely used machine learning tools like neural networks generally are not capable of processing the complex values. Xiao *et al.* [8] pioneered CSI-based fingerprinting localization system (FIFS) using merely the amplitude. Comparatively, Wang *et al.* [14] proposed a phase-based fingerprinting system, PhaseFi. A deep neural network with three hidden layers was adopted to train the phase data. Another possible solution is to feed the amplitude and phase (or the real and imaginary components) separately to the different channels of the convolutional neural network (CNN). Moreover, the measured CSI from the commodity devices is distorted due to the imperfect synchronization and other system errors, which requires complicated pre-processing. In [5], De Bast *et al.* showed the CSI can also be calibrated via deep learning based on an encoder-decoder architecture.

Intermediate geometrical metrics: There are also many works on estimating the intermediate geometrical metrics, namely, angle, distance, or velocity, based on machine learning

tools. In this case, machine learning acts as a super-resolution channel parameters estimation algorithm. After this, the estimated angle or distance metrics are exploited for positioning or tracking purpose. In [15], Dai *et al.* proposed to use deep learning to derive angle-of-arrival (AoA) from a single snapshot of CSI, in which both classification- and regression-based models were investigated. Moreover, rather than estimating ToF from CIR directly, most of the learning-based research for ultra wide-band (UWB) localization focuses on the ToF or range correction [16].

Elaborated features: Even though learning from raw channel measurements is straightforward to implement, it is not easy to learn the location-related features unless a deep neural network is adopted. It might be not feasible for lightweight mobile edge computing. In [17], Wu *et al.* designed an angle-delay domain channel power matrix (ADCPM) for 3D massive MIMO localization. The ADCPM includes the multi-path power spectrum of delay, elevation, and azimuth in 4-D tensor, which achieved higher accuracy with reduced computational complexity and storage overhead. Another main drawback of learning-based approaches is the generalization ability for varied environments. To this end, another potential learning-based research direction is to elaborate the cross-domain features rather than using raw channel response.

3.1.2. Model-based Approaches

Learning-based solutions can achieve very high accuracy at the cost of labour-intensive data collection and labelling. Comparatively, the model-based approaches establish the explicit mathematical model from the channel response to the location, which are more physically interpretable. Furthermore, the model-based algorithms do not have the generalization problem as in the learning-based approaches, which are attractive from the perspective of real-world implementation.

Intermediate geometrical metrics: Within this framework, the distance-, angle-, and velocity-related metrics are estimated from the channel response via the super-resolution parameters estimation algorithms, *e.g.*, MUSIC, SAGE, RiMAX, etc. Then these intermediate geometrical metrics are used for positioning and tracking. Specifically, SpotFi was proposed based on commodity Wi-Fi devices [11]. The AoA of the multi-path propagation was estimated from the CSI first, then the AoA of a direct link was obtained using a filtering method. Due to the high-ranging resolution of UWB, multi-path assisted localization has attracted attention by exploiting the ToF estimates of both LoS and the specular multi-path reflections. Generally, given the prior knowledge of the floor plan or within the simultaneous localization and mapping framework [18], the fixed but unknown virtual anchors' locations can be obtained. Such this, positioning accuracy can be enhanced via the increasing spatial diversity of more (virtual) anchors.

Direct localization: Another promising solution is based on the channel response directly, as it depicts the rich propagation information and is also tightly location-related. This idea is especially suitable for bandwidth-limited systems, like ultra-high frequency RFID systems, sub-6 GHz Wi-Fi, BLE, etc. A promising accuracy can be achieved by exploiting the idea of synthetic aperture, namely the distributed antenna array or moving antenna. For example, in [7], Miesen *et al.* proposed to localize the RFID tag via a moving antenna on a robot arm. The moving antenna constructed a virtual antenna array, and the RFID tag was localized via a

matched filter.

3.2. Contact-Free Positioning and Tracking

Besides the device-based localization mentioned above, contact-free localization also plays an important role in our daily life. Contact-free localization exploits the signal reflected or scattered by the targets to estimate the locations, such as worker safety monitoring and accident alerts in the assembly area in the industry, as shown in Fig. 2 (bottom). This is more feasible as it does not require attaching any radio devices or sensors to the workers. Meanwhile, it also preserves personal privacy without the device-based individual identity.

While conventional contact-free methods primarily process the signal strength variations for coarse positioning, the channel response in commodity devices enables small-scale motion sensing like the prohibitively large phased array radar with a relatively low cost. Besides wide-band radio systems like mm-wave devices, narrow-band radio systems are also able to achieve accurate spatial sensing when deploying multiple antennas, or a motorized platform based on the idea of synthetic aperture radar. As shown in Fig. 3, there also are learning-based and model-based approaches generally.

3.2.1. Learning-based Approaches

Contact-free positioning is generally more complicated than the device-based scheme as it utilizes the reflection or scattering from the targets. The target-of-interest signal is hidden in the complex multi-path signal and has lower power than the LoS components and even the MPCs from the surroundings. Learning-based approaches can mitigate or relieve the interference via a data-driven model.

Raw measurements: Similar to the device-based scheme, learning from raw channel measurements for contact-free localization is relatively intuitive. The difference is the learning-based contract-free methods should mitigate the undesired background signals including the LoS propagation and the scattering from the surroundings. There are also mainly two categories of learning-based approaches: classification and regression. Classification-based methods are to compare the instantaneous measurements with the fingerprinting database or radio maps to infer the target's location. Regression-based methods are to infer the map from the raw channel measurements to the target's coordinates.

Intermediate geometrical metrics: Machine or deep learning tools are exploited to estimate the distance, angle, or velocity metrics based on the channel measurements in the cases of absent and present targets. In [9], Li *et al.* proposed a residual CNN-based reflected ToF estimation method via learning the differences between static and dynamic CIR or corresponding variance sequences. Then a particle filter was implemented for passive human tracking based on the ToF estimates. Moreover, it concluded that the variance-based metric is less domain-dependent than using the CIR directly.

Elaborated features: To advance cross-domain sensing and achieve (near) zero-effort sensing for new environments and system settings, a domain-independent feature, namely, the so-called *body-coordinate velocity profile*, was designed for the gesture classification [19]. However, according to the authors' best knowledge, there is no available literature on contact-

free positioning and tracking exploiting the elaborated domain-independent features. But it can be a promising direction concerning the scalability of the learning-based solutions.

3.2.2. Model-based Approaches

The taxonomy of model-based contact-free positioning and tracking is given as follows.

Intermediate geometrical metrics: One of the most popular approaches is to estimate the reflected angle, distance, or Doppler velocity from the target(s) using the super-resolution parameters estimation algorithms. In [20], Widar 2.0 exploited the conjugate multiplication between the adjacent antennas to mitigate the CSI measurement errors and static interference, and applied the SAGE algorithm to estimate the targeted MPCs. A binary optimization, together with a Hungarian algorithm, was used to associate the targeted MPCs along the trajectory. Instead of using CSI multiplication, Wu et al. [21] proposed a CSI-quotient model to estimate the targeted Doppler shift. The CSI-quotient model assumed the adjacent antennas possess the same static interference, which can obtain a higher targeted signal-to-noise ratio.

Direct localization: By using the intermediate geometrical metrics or the MPCs estimates for contact-free tracking, we not only obtain the targeted MPCs but also the cluttered components. It is necessary to associate the targeted MPCs along the time, namely evolving association. If there are multiple targets, associating the MPCs with the specific target, the so-called target association, is also required. Furthermore, if the radar-like sensors are distributed in the environment, we estimate the MPCs for each sensor. In this case, we also need to associate the MPCs of a specific target with all distributed sensors, that is sensor or link association. When taking this three-dimension association problem into account, the contact-free positioning and tracking will be computing complex and difficult to resolve [13]. Recently, there is work focusing on direct channel response-based contact-free localization without the consideration of MPCs association. Sakhnini *et al.* [22] established a radar-like tracking prototype based on a massive MIMO communication test bed. However, a single metallic cylinder target was considered and localized.

4. Micro-Benchmark: Channel Response-based Direct Positioning and Tracking

Channel response-based positioning and tracking definitely is a promising research field considering the pervasive applications, especially, when an increasing number of commodity radio devices provide the access to channel measurements. Among them, channel response-based direct localization obtains the location estimates through the integration of physical-layer channel information, which show its priority concerning low signal-to-noise ratio and multi-path effect [12, 13]. For device-based localization, instead of estimating the intermediate angle and distance metrics, we exploit the geometrical diversity of the multiple antenna. The user or target is localized via a matched filter built upon the idea of synthetic aperture. The user's

location \mathbf{P}_U can be estimated via

$$\begin{aligned}\hat{\mathbf{P}}_U &= \arg \max_{\mathbf{P}_U} \mathcal{C}(\mathbf{P}_U) \\ &= \arg \max_{\mathbf{P}_U} \left| \sum_{i=1}^{N_{\text{Tx}}} \sum_{j=1}^{N_{\text{Rx}}} \sum_{k=1}^{N_f} \tilde{\alpha}_{i,j,k} e^{\mathcal{J}(\phi_{i,j,k}(\mathbf{P}_{\text{UE}}) - \tilde{\phi}_{i,j,k})} \right|,\end{aligned}\quad (3)$$

where $\tilde{\alpha}_{(\cdot)}$, $\tilde{\phi}_{(\cdot)}$ denote the amplitude and the phase of the measured channel response. N_{Tx} , N_{Rx} , and N_f are the number of transceiver's antennas and sub-carriers, respectively. $\phi_{\cdot,\cdot,\cdot}(\mathbf{P}_{\text{UE}})$ is the phase calculated via the location candidates.

For contact-free localization, we can adopt a similar matched filter as in (3), but instead the $\tilde{\alpha}_{(\cdot)}$, $\tilde{\phi}_{(\cdot)}$ denote the amplitude and the phase of the targeted channel response after interference mitigation. Define the range-Doppler profile at a single antenna as $\hat{\mathbf{H}}_{(i,j)} = \mathbb{T}^H \mathbf{H}_{(i,j,\cdot,\cdot)} \mathbb{D}$, where \mathbb{T} and \mathbb{D} denote the discrete Fourier transform matrix, then we can obtain the targeted channel response $[\hat{\mathbf{H}}_{\text{Target}}]_{(i,j)} = \hat{\mathbf{H}}_{(i,j,\cdot,\cdot)}(\hat{n}_T, \hat{n}_D)$, where

$$(\hat{n}_T, \hat{n}_D) = \arg \max_{n_T, n_D} \mathbb{E} \left\{ (\mathbb{T}^H \mathbf{H}_{(i,j,\cdot,\cdot)} \mathbb{D}) \odot (\mathbb{T}^H \mathbf{H}_{(i,j,\cdot,\cdot)} \mathbb{D})^* \right\}, \quad (4)$$

and \odot denotes the Hadamard product. Note that the maximization in (4) indicates only single-target tracking is applicable through this method. For multi-target contact-free tracking, the targets association is needed. To handle this problem, compressive sensing and random finite set theory can be implemented, which is omitted here but refer to our latest work in [13] due to page limitation.

To track the user or target, we propose to track the changes of the matched filter in (3) through particle filter. For each location update, the K particles are weighted via the differences with the maximum of $\mathcal{C}(\mathbf{P}_U^{(i)})$, given as [13],

$$\begin{aligned}\hat{P}_w^{(i)} &= \exp \left(\frac{1}{2\sigma_c^2} \left(\min_{1 \leq i \leq K} \{ P_w^{(i)} \} - P_w^{(i)} \right) \right), \\ \text{s.t.} \quad P_w^{(i)} &= \left(\mathcal{C}_{\max} - \mathcal{C}(\mathbf{P}_U^{(i)}) \right)^2,\end{aligned}\quad (5)$$

where σ_c is the standard deviation of $\mathcal{C}(\mathbf{P}_U)$. $\mathcal{C}_{\max} = 1$ because $\mathcal{C}(\mathbf{P}_U)$ is normalized to $(0, 1]$.

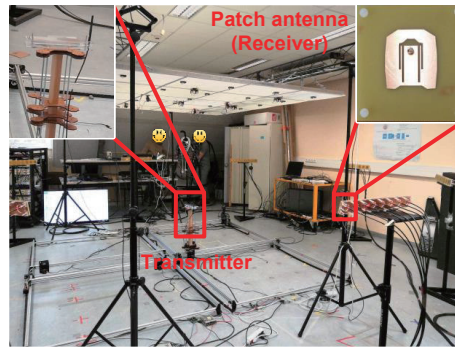
Fig. 4 shows the device-based and contact-free positioning and tracking results for a single user or pedestrian based on a sub-6 GHz distributed massive MIMO orthogonal frequency-division multiplexing (OFDM) communication test bed. The center frequency is 2.61 GHz and the bandwidth 18 MHz. The other parameter settings are given in Table 1. The transmitter is with a dipole antenna and the receiver with 64 patch antennas.

For both device-based and contact-free schemes, the impact of the number of antennas on accuracy is evaluated. The tracking accuracy is represented by the absolute distance errors between the tracking results and the ground truth. For the device-based scheme, the ground truth was obtained via the computerized numerical control X-Y table which guarantees a millimetre-level accuracy of the user equipment. For the contact-free scheme, the ground truth of the pedestrian was inferred via the active tracking. For the active tracking, we asked

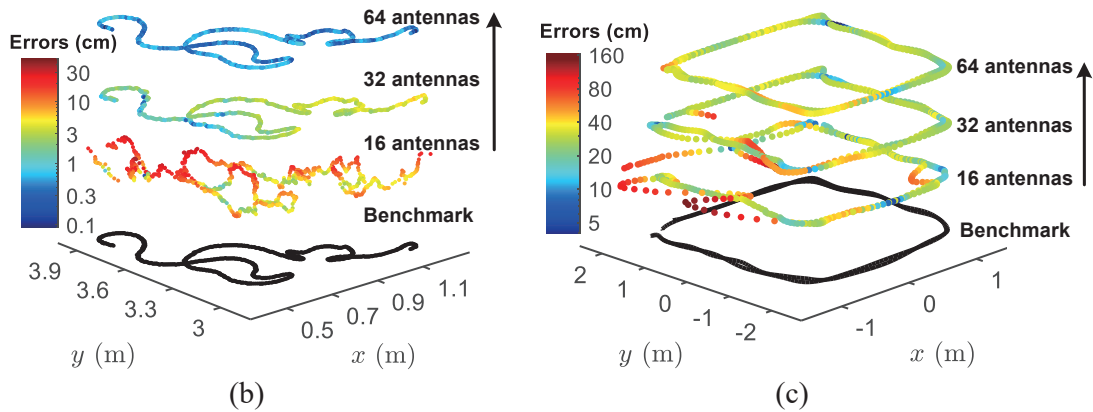
Table 1

Parameters settings of the massive MIMO-OFDM system

System parameters	Value
Transmitter power	15 dBm
Sampling rate	100 Hz
Wavelength	11.49 cm
# of Sub-carrier	100
# of antenna	64
Modulation	QPSK
Patch antenna size	7 cm×7 cm
Size of experimental space	6.5 m×10 m



(a)

**Figure 4:** Sub-6 GHz massive MIMO-OFDM system for positioning and tracking: (a) Measurement setup. (b) Device-based scheme. (c) Contact-free scheme.

each participant to place the antenna (another transmitter) on the top of the head to avoid the possible body shadowing effect. The active tracking ensures a centimetre-level accuracy that has been testified in [13], which is sufficient and applicable for pedestrian tracking given the size of human body. As expected, the accuracy increases as more antennas have been exploited as shown in Figs. 4(b)-(c). Note that we have not considered the antenna deployment

optimization herein but based on the sequential antenna selection. For device-based tracking, when 64 antennas are adopted, a (sub-) centimetre-level accuracy can be achieved. Fig. 4(c) shows the results of contact-free pedestrian tracking, in which a decimetre-level accuracy can be guaranteed under the assumption of cylindrical body model.

5. Discussions and Future Opportunities

Channel response makes the finer-grained positioning and tracking possible and applicable. However, there are still many challenges for real-world implementation. In the following subsections, we will discuss some open challenges and future research opportunities for channel response-based positioning and tracking.

5.1. Probabilistic Positioning and Tracking

Most available localization algorithms are based on such as the AoA and ToF estimates. However, these distance and angle metrics are only intermediate estimates from which the position is derived. Their estimation accuracy may be constrained by the resolution of the estimation algorithms. Meanwhile, the super-resolution algorithms generally require complex computations. One recently emerging solution for accuracy enhancement is based on soft information [23] instead of using a single-value distance or angle estimate. The ranging or angular likelihood is extracted from the channel response measurements. The likelihood is generally assumed as a Gaussian mixture model, which can be solved via the expectation-maximization algorithm or learned by deep probabilistic learning. On the other hand, soft-ranging or soft-angular information rarely appears in contact-free positioning and tracking, which can also be a promising research direction.

5.2. Simultaneous Localization and Mapping

Simultaneous localization and mapping is to construct the 2-D or 3-D map of an unknown environment and localize and track the targets simultaneously. To sense the surroundings, it requires a system that has a very high spatial resolution, such as UWB and mm-wave systems. With the environment knowledge, we can exploit the specular multi-path to enhance the localization performance with fewer anchors, so less human exposure as well [18]. However, obtaining the information on the complete surrounding is not necessary for positioning and tracking. Abundant knowledge will also increase the computing complexity and the inference latency. For example, if we use the specular MPCs to help localization, we only need the location of the planar reflectors (e.g., wall). On the other hand, MPCs exploitation requires complex association. The factor graph model and random finite set filtering are the two promising techniques for this issue.

5.3. Contact-Free Multi-Target Positioning and Tracking

Contact-free localization becomes a hot topic due to the advance in integrated communication and sensing, which achieves the sensing functionality by reusing the allocated spectrum,

hardware, and even signalling resources of the communication systems. First, background removal or ToI extraction is a big challenge for contact-free localization. The CSI-quotient model [21] can effectively mitigate the undesired signal using the CSI ratio of adjacent antennas. Second, according to the authors' best knowledge, most work on contact-free localization, especially for sub-6 GHz systems, concentrates on a single target. How to generalize to an unknown number of targets or multiple targets desires further investigation. The widely adopted MPCs-based localization requires complex MPCs association for multiple targets. The problem becomes more difficult if distributed sensors are adopted. Third, investigations on contact-free multiple static targets or hybrid static-dynamic targets localization are rare. It is challenging to use the Doppler shifts or channel variations for interference mitigation. One possible solution is to use high spatial resolution radio to generate the points cloud or image of the targets. Another method is to detect the vital sign (vibration) of the human (machine). However, how to evade the interference from the unexpected motion or dynamic targets needs further exploration.

5.4. Networked Positioning and Tracking

When in a cluttered environment, the signal is easy to be blocked by the metallic assets, which degrades the localizability and the positioning accuracy. On the other hand, for the sub-6 GHz bandwidth-limited systems, the angular resolution is limited even though a small antenna array can be installed. Therefore, improving the geometrical diversity of the anchors or access points can enhance the accuracy and coverage. During tracking, the local access points can share the channel information with the remote access points as a priori knowledge. Furthermore, for multi-target device-based tracking, the localized targets can also act as the new access points if they possess the transceiver functionality, namely, the networked positioning and tracking. Through information exchange, the field of sensing and the performance of each access point can be improved greatly.

5.5. Domain-Independent Positioning and Tracking

Even though the learning-based approaches are straightforward to implement and fairly easy to use, one of the main drawbacks of this scheme is that generalization for a new system setting (including hardware heterogeneity, antenna orientation, etc.) or a new environment is challenging. They are namely *domain dependent*. Training from scratch in a new domain or involving transfer learning can solve this problem, but it requires newly labelled dataset collection which adds effort more or less to practical implementations. To handle this problem, one solution is to establish an open-access dataset that is contributed by the whole community, like the well-known image recognition dataset ImageNet (<https://www.image-net.org/>). The open-access dataset should include the measurements at different sites, using different commodity devices, having different numbers of anchors, etc. The dataset provides a cross-domain platform for algorithm development and evaluation. It can also help to mitigate the performance gap from the well-controlled laboratory to practical implementation. Another potential learning-based research direction is to feed the less domain-dependent metrics to the model instead of the raw channel response. The domain-independent metrics generally require elaborated design, which should retain the rich channel information but be less dependent on the system setting

or environment.

6. Conclusion

Location acquisition has become an important content considering the ubiquitous location-aided vertical applications. Physical-layer channel information, i.e., channel response, in commodity devices allows us to better understand the signal propagation and develop the algorithms for positioning and tracking purposes. In this paper, the representative SOTA regarding device-based and contact-free schemes are classified into learning-based and model-based approaches. The corresponding pros and cons are discussed and analysed. This paper also presents of channel response-based direct positioning and tracking, which provides a micro-benchmark for both device-based and contact-free schemes. Furthermore, we will point out some limitations that the available work that have not solved well for channel response-enabled localization. Some open challenges and future opportunities are given as well.

Acknowledgments

This work is supported in part by the Excellence of Science (EOS) project MUlti-SERVICE WIREless NETWORKS (MUSE-WINET), in part by the imec.icon project InWareDrones, in part by the imec project UWB-IR, and in part by the Research Foundation Flanders (FWO) under Grant no. G098020N. The authors would like to thank Sibren De Bast and Robbert Beerten from KU Leuven for the help of the distributed massive MIMO experiments.

References

- [1] M. Wollschlaeger, T. Sauter, J. Jasperneite, The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0, *IEEE Industrial Electronics Magazine* 11 (2017) 17–27. doi:Mar.
- [2] C. Li, E. Tanghe, P. Suanet, D. Plets, J. Hoebeke, E. De Poorter, W. Joseph, ReLoc 2.0: UHF-RFID Relative Localization for Drone-Based Inventory Management, *IEEE Transactions on Instrumentation and Measurement* 70 (2021) 1–13.
- [3] F. Zafari, A. Gkelias, K. K. Leung, A Survey of Indoor Localization Systems and Technologies, *IEEE Communications Surveys & Tutorials* 21 (2019) 2568–2599.
- [4] Y. Li, Y. Zhuang, X. Hu, Z. Gao, J. Hu, L. Chen, Z. He, L. Pei, K. Chen, M. Wang, X. Niu, R. Chen, J. Thompson, F. M. Ghannouchi, N. El-Sheimy, Toward Location-Enabled IoT (LE-IoT): IoT Positioning Techniques, Error Sources, and Error Mitigation, *IEEE Internet of Things Journal* 8 (2021) 4035–4062.
- [5] S. De Bast, E. Vinogradov, S. Pollin, Expert-Knowledge-Based Data-Driven Approach for Distributed Localization in Cell-Free Massive MIMO Networks, *IEEE Access* 10 (2022) 56427–56439.
- [6] C. Li, S. De Bast, E. Tanghe, S. Pollin, W. Joseph, Toward Fine-Grained Indoor Localization Based on Massive MIMO-OFDM System: Experiment and Analysis, *IEEE Sensors Journal* 22 (2022) 5318–5328.

- [7] R. Miesen, F. Kirsch, M. Vossiek, UHF RFID Localization Based on Synthetic Apertures, *IEEE Transactions on Automation Science and Engineering* 10 (2013) 807–815.
- [8] J. Xiao, K. Wu, Y. Yi, L. M. Ni, FIFS: Fine-Grained Indoor Fingerprinting System, in: 2012 21st International Conference on Computer Communications and Networks (ICCCN), 2012, pp. 1–7.
- [9] C. Li, E. Tanghe, J. Fontaine, L. Martens, J. Romme, G. Singh, E. De Poorter, W. Joseph, Multistatic UWB Radar-Based Passive Human Tracking Using COTS Devices, *IEEE Antennas and Wireless Propagation Letters* 21 (2022) 695–699.
- [10] T. Altstidl, S. Kram, O. Herrmann, M. Stahlke, T. Feigl, C. Mutschler, Accuracy-aware compression of channel impulse responses using deep learning, in: 2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2021, pp. 1–8. doi:10.1109/IPIN51156.2021.9662545.
- [11] M. Kotaru, K. Joshi, D. Bharadia, S. Katti, SpotFi: Decimeter Level Localization Using WiFi, in: Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM), 2015, p. 269–282.
- [12] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, M. Coulon, Direct Localization for Massive MIMO, *IEEE Transactions on Signal Processing* 65 (2017) 2475–2487. doi:10.1109/TSP.2017.2666779.
- [13] C. Li, S. De Bast, Y. Miao, E. Tanghe, S. Pollin, W. Joseph, Contact-Free Multitarget Tracking Using Distributed Massive MIMO-OFDM Communication System: Prototype and Analysis, *IEEE Internet of Things Journal* 10 (2023) 9220–9233. doi:10.1109/JIOT.2023.3234041.
- [14] X. Wang, L. Gao, S. Mao, CSI Phase Fingerprinting for Indoor Localization With a Deep Learning Approach, *IEEE Internet of Things Journal* 3 (2016) 1113–1123.
- [15] Z. Dai, Y. He, V. Tran, N. Trigoni, A. Markham, DeepAoANet: Learning Angle of Arrival From Software Defined Radios With Deep Neural Networks, *IEEE Access* 10 (2022) 3164–3176.
- [16] H. Wymeersch, S. Marano, W. M. Gifford, M. Z. Win, A Machine Learning Approach to Ranging Error Mitigation for UWB Localization, *IEEE Transactions on Communications* 60 (2012) 1719–1728. doi:10.1109/TCOMM.2012.021719.
- [17] C. Wu, X. Yi, W. Wang, L. You, Q. Huang, X. Gao, Q. Liu, Learning to Localize: A 3D CNN Approach to User Positioning in Massive MIMO-OFDM Systems, *IEEE Transactions on Wireless Communications* 20 (2021) 4556–4570.
- [18] K. Witrisal, P. Meissner, E. Leitinger, Y. Shen, C. Gustafson, F. Tufvesson, K. Haneda, D. Dardari, A. F. Molisch, A. Conti, M. Z. Win, High-Accuracy Localization for Assisted Living: 5G systems will turn multipath channels from foe to friend, *IEEE Signal Processing Magazine* 33 (2016) 59–70.
- [19] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, Z. Yang, Widar3.0: Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021) 1–18.
- [20] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, Y. Liu, Widar2.0: Passive Human Tracking with a Single Wi-Fi Link, in: Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys), 2018, p. 350–361.
- [21] D. Wu, Y. Zeng, R. Gao, S. Li, Y. Li, R. C. Shah, H. Lu, D. Zhang, WiTraj: Robust Indoor Motion Tracking with WiFi Signals, *IEEE Transactions on Mobile Computing* (2021) 1–17.

- [22] A. Sakhini, S. De Bast, M. Guenach, A. Bourdoux, H. Sahli, S. Pollin, Near-Field Coherent Radar Sensing Using a Massive MIMO Communication Testbed, *IEEE Transactions on Wireless Communications* 21 (2022) 6256–6270.
- [23] A. Conti, F. Morselli, Z. Liu, S. Bartoletti, S. Mazuelas, W. C. Lindsey, M. Z. Win, Location Awareness in Beyond 5G Networks, *IEEE Communications Magazine* 59 (2021) 22–27.