User-Side Indoor Localization Using CSI Fingerprinting

Author(s):
Kazemi, Parham; Al-Tous, Hanan; Studer, Christoph; Tirkkonen, Olav

Publication date:
2022

Permanent link:
https://doi.org/10.3929/ethz-b-000580771

Originally published in:
https://doi.org/10.1109/SPAWC51304.2022.9833973

Funding acknowledgement:
813999 - Integrating wireless communication engineering and machine learning (EC)

User-Side Indoor Localization Using CSI Fingerprinting

Parham Kazemi[1], Hanan Al-Tous[1], Christoph Studer[2], and Olav Tirkkonen[1]

[1]*Department of Communications and Networking, Aalto University, Espoo, Finland*

[2]*Department of Information Technology and Electrical Engineering, ETH Zurich, Zurich Switzerland*
e-mail:{parham.kazemi, hanan.al-tous, olav.tirkkonen}@aalto.fi, studer@ethz.ch

***Abstract*—We consider a scalable User Equipment (UE)-side indoor localization framework that processes Channel State Information (CSI) from multiple Access Points (APs). We use CSI features that are resilient to synchronization errors and other hardware impairments. As a consequence our method does not require accurate network synchronization among APs. Increasing the number of APs considered by a UE profoundly improves fingerprint positioning, with the cost of increasing complexity and channel estimation time. In order to improve scalability of the framework to large networks consisting of multiple APs in many rooms, we train a multi-layer neural network that combines CSI features and unique AP identifiers of a subset of APs in range of a UE. We simulate UE-side localization using CSI obtained from a commercial raytracer. The considered method processing frequency selective CSI achieves an average positioning error of 60 cm, outperforming methods that process received signal strength information only. The mean localization accuracy loss compared to a non-scalable approach with perfect synchronization and CSI is 20 cm.**

***Index Terms*—Channel state information, user equipment (UE)-side indoor localization, fingerprinting, neural networks.**

I. Introduction

Indoor localization finds widespread use in our everyday lives and is critical to many industry applications, e.g., robot navigation and asset tracking [1]. Global Navigation Satellite Systems (GNSSs), the mainstream technology for outdoor navigation, do not typically work in indoor environments due to lacking Line-of-Sight (LoS) connectivity to a sufficient number of satellites. Similarly, distance-based localization techniques that leverage Time of-Arrival (ToA), Time-Difference-of-Arrival (TDoA), or Angle-of-Arrival (AoA) information perform poorly indoors, as the wireless signals typically have to propagate through walls [2].

Fingerprint-based localization methods are not restricted to LoS connectivity. Accordingly, they have gained significant attention recently [1], [3], [4]. In these methods, a database containing ground-truth locations and received signal strength indicators (RSSI) or Channel State Information (CSI) samples is first constructed. Then, measured RSSI or CSI samples are matched with the fingerprint database to generate a location estimate. RSSI-based methods typically perform poorly as compared to CSI-based approaches which capture fine-grained channel information at the subcarrier level. Indoor localization methods based on CSI samples have gradually become more and more accurate [1], [3], [4].

Localization at the User Equipment (UE)-side has the advantage that APs can transmit dedicated positioning reference signals which are used by all UEs to estimate CSI. Localization accuracy can be improved by processing information from multiple APs [5], [6]. If the network of APs is perfectly synchronized, the relative phase information in CSI from multiple APs carries relevant positioning information. Otherwise relative phases are irrelevant, and may be removed before fusion of multi-AP CSI. It is also crucial to observe that the delay- or frequency domain CSI estimated by a UE depends on the synchronization of the UE to the received signal from an AP.

Mass-market wireless transceivers are cost effective, and come with non-ideal analog radio frequency (RF) components. As a consequence, RF-induced phase noise (PN), carrier frequency offset (CFO), and I/Q-imbalance affect the received signals [7]. Estimation of impulse responses in Multiple-Input Multiple-Output Orthogonal Frequency-Division Multiplexing (MIMO-OFDM) systems in the presence of phase noise and I/Q imbalance was considered in [8]. A pre-processing method to clean CSI from phase fluctuations introduced by imperfect synchronization between the transmitter and the receiver was considered in [9]. A compressive sensing framework that recovers the channel delay components and their coefficients at high resolution under hardware imperfections (i.e., phase noise) was considered in [10]. This method improves the channel delay profile resolution, which, in turn, enables improved localization accuracy.

In this paper, we consider on-device CSI-based indoor positioning which is robust to timing mismatch and hardware impairments. For this, we apply CSI features that are resilient to hardware impairments. We provide a scalable approach for leveraging CSI for indoor positioning in a network with a large coverage area and many APs. Scalability is achieved by two methods. First, the training and localization samples of a UE at a given location consists of the CSI features of an active set of P APs with strongest received signal power, together with the unique identifiers of these APs. Second, the service area is divided into subareas. In each subarea the union of

the active sets of all sample UEs in the area consists of a limited set of APs. For each subarea, we train a multi-layer neural network to perform UE-side indoor positioning. We show simulation results with realistic channel vectors from a commercial raytracer.

II. System Model

We consider the propagation environment between multiple APs and U single-antenna UEs. Each AP is equipped with M antennas. The location of UE u is denoted by $\mathbf{x}_u \in \mathbb{R}^D$, where $D = 2$ or $D = 3$. Similarly, the location of AP b is $\mathbf{x}_b \in \mathbb{R}^D$. The UE is assumed to be synchronized to one AP. We do not assume network-level synchronization among APs.

A. Channel Model

The channel between an AP and a UE is assumed to consist of L multipath components, each with a path gain $a_l \in \mathbb{C}$ and delay $\tau_l \in [0, \tau_{\max}]$, where $\tau_{\max}$ denotes the maximum delay spread. The baseband continuous time Channel Impulse Response (CIR) between a UE and one antenna of an AP is thus

$$c(\tau) = \sum_{l=1}^{L} a_l \, \delta(\tau - \tau_l), \tag{1}$$

where $\delta(\cdot)$ is the Dirac delta function. The Channel Frequency Response (CFR) is $h(f) = \int e^{-j2\pi f \tau} c(\tau) d\tau$.

We assume that OFDM is used, and that the AP transmits pilot signals to the UE over N equi-spaced subcarriers with subcarrier spacing f_s. The subcarriers are indexed by $n = -\frac{N-1}{2}, \ldots, \frac{N-1}{2}$. With f_0 the carrier frequency, subcarrier n has center frequency $f_n = f_0 + n f_s$ and CFR

$$h_n \triangleq h(f_n) = \sum_{l=1}^{L} a_l \, e^{-j2\pi f_n \tau_l}. \tag{2}$$

The vector $\mathbf{h} = [h_1, \ldots, h_N]^T$ of subcarrier channels is the Discrete Fourier Transform (DFT) of the discrete time Tapped Delay Line (TDL) channel $\mathbf{c} = [c_1, \ldots, c_K]^T$ with $K < N$ delay taps, after extending $\mathbf{c}$ to length N by appending zeroes. Note that in the discrete time TDL, there would typically be more than L taps, K being proportional to $\tau_{\max}$.

Due to hardware imperfections, and synchronization errors, the CSI measurements are affected by various timing, phase, and magnitude distortions. The elements of the estimated CSI vector $\hat{\mathbf{h}} = [\hat{h}_1, \ldots, \hat{h}_N]^T \in \mathbb{C}^N$ are modeled as

$$\hat{h}_n = e^{-j\phi_n} h_n + z_n, \tag{3}$$

where $\phi_n = 2\pi n f_s \beta + \psi_n$ is a phase distortion term with time offset β, arising from synchronization error, and ψ_n a phase offset arising from other impairments. Estimation noise is modeled by $z_n \in \mathbb{C}$, assumed Gaussian and white, which also is assumed to capture the interference arising from CFO, I/Q imbalance and phase noise. The DFT of the estimated CFR gives rise to an estimated TDL $\hat{\mathbf{c}}$ with elements $\hat{c}_k$. We assume that that effective CIR does not exceed the Cyclic Prefix (CP) length. For more details on impairment modeling, see [11], [12].

B. CSI Feature

Since the estimated CSI is affected by phase distortion, it is not suitable to be directly used as a feature for localization [13], [14]. The absolute square of the estimated subcarrier channel can be written as

$$|\hat{h}_n|^2 = |h_n|^2 + \tilde{z}_n, \tag{4}$$

where $\tilde{z}_n = h_n^* e^{j\phi_n} z_n + z_n^* e^{-j\phi_n} h_n + |z_n|^2$. The phase distortion is transformed from a multiplicative factor to an additive term and the leading term is not affected by phase errors. The vector of absolute squares of subcarrier channels is equivalent to the Fourier transform of the CIR autocorrelation function

$$r(\zeta) = (c \circledast \underline{c}^*)(\zeta) = \sum_{k=1}^{L} \sum_{l=1}^{L} a_k a_l^* \delta(\zeta - (\tau_k - \tau_l)), \tag{5}$$

where $\circledast$ denotes a convolution, underlined quantities are time reversed, the superscript * denotes complex conjugate, and $\zeta \in [-\tau_{\max}, \tau_{\max}]$.

We shall use the discrete delay domain estimated TDL autocorrelation as a feature, i.e., the DFT of the vector of absolute squares of the subcarrier channels [14]:

$$r_\kappa = \sum_k \hat{c}_k \hat{c}_{k+\kappa-1}^*, \tag{6}$$

where $\kappa = 1, \ldots, K$ and K is the number of delay taps. As the delay spread of the channel is assumed to be well within the CP, we have $K << N$. The TDL autocorrelation thus has the same information as the vector of $|\hat{h}_n|^2$-values, but in a much lower dimension. This CSI feature is resilient to synchronization errors and hardware imperfections.

Considering all M antennas of AP b, the CSI features of UE u are denoted by:

$$\mathbf{R}_b^u = [\mathbf{r}_{1,b}^u, \ldots, \mathbf{r}_{M,b}^u] \in \mathbb{C}^{K \times M}, \tag{7}$$

where $\mathbf{r}_{m,b} \in \mathbb{C}^K$ is the autocorrelation vector of the estimated TDL of antenna m at AP b.

C. Benchmark CSI Features

A variety of features have been used for fingerprinting indoor positioning. RSSI is frequently used and can be easily obtained at the receiver. However, due to the complexity of real-world indoor channels, this feature is unable to provide high positioning accuracy. RSSI fingerprint based localization is insensitive to synchronization errors and phase distortion. We consider RSSI based localization as a benchmark approach to evaluate the localization accuracy in this paper.

As another benchmark feature we use the full estimated CSI. Although this feature can provide better accuracy than RSSI-based approaches, it requires complex algorithms to estimate and mitigate/cancel the effect of possible impairments, assumes that all UEs apply the same synchronization algorithm, and are able to perform error free synchronization. In most of the literature, CSI is assumed to be estimated accurately with perfect synchronizity for localization purposes.

The full complex-valued CIR was considered for fingerprinting in [15]. Here, we consider the magnitude of the discrete delay domain CIR as a CSI feature to benchmark performance.

III. Positioning Framework

For each of the CSI features discussed above, a Deep Neural Network (DNN) is trained to map the feature vector to the location of a UE. The UE receives the signal of several APs, and measures the RSSI for each AP. If RSSI-based positioning is used, a feature containing the measured signal strength of all APs in the indoor area is fed to the DNN.

For the benchmark full CSI feature, the UE estimates the CSI of all APs. A feature vector is generated based on the magnitude of the delay domain channel taps. Then, by stacking features of different APs into one high dimensional feature, the DNN takes in the large input feature and learns the mapping between the CSI feature vector and the UE location.

These methods scale poorly when the environment expands to cover a large area with a multitude of access points. Using the CSI from all APs, one can have accurate position estimation, but the dimension of the input feature as well as UE complexity and energy consumption will grow linearly with the number of APs. Also, a UE does not receive a high quality signal from all APs in a large indoor area.

We thus consider the input features of a UE to be collected from a UE-location specific active set of APs with the strongest received power. This enables scaling to larger areas without undue expansion of UE measurement requirements. A unique AP identifier is appended to each CSI feature vector so that the DNN can distinguish between different APs in the input. Here, for simplicity, we use AP location as an ID. To reduce DNN complexity we divide the service area into subareas and train a separate positioning DNN for each subarea. We assume that there is a classifier which is able to provide a coarse UE location estimate based on the CSI feature, i.e., to which room/subarea the UE belongs.

The features resilient to synchronization errors and hardware imperfections are then constructed as follows. The RSSIs of different APs are first measured at the UE. The P APs with largest received signal power are chosen to the active set, and their CSI is estimated. The autocorrelation feature vectors are then formed from the estimated delay domain channel taps. The location of the AP is concatenated as IDs to the corresponding CSI feature vector. Fig. 1 illustrates the architecture of the considered positioning system.

A. Feature Prepossessing

Since off-the-shelf machine learning frameworks are unable to process complex numbers, for further processing we convert the complex valued autocorrelation features to real-valued vectors stacking the real and imaginary parts together. Thus, the MK-dimensional matrix $\mathbf{R}_b^u$ of autocorrelations is converted to a $2MK$- dimensional feature vector. We assume that AP location is available and concatenate it to the received CSI feature, resulting in

$$\mathbf{f}_b^u = [\mathrm{Re}\{\mathrm{vec}(\mathbf{R}_b^u)\}^T, \mathrm{Im}\{\mathrm{vec}(\mathbf{R}_b^u)\}^T, \mathbf{x}_b^T]^T. \quad (8)$$

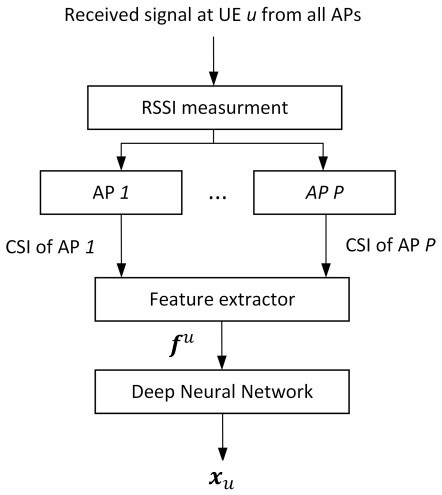

Fig. 1: Block diagram of the positioning system.

The main reason for adding the AP location is to identify the AP to the DNN. The final feature vector that is input to the DNN is constructed by stacking the $\mathbf{f}_b^u$ from the P best (strongest) APs in the subarea, i.e.,

$$\mathbf{f}^u = [\mathbf{f}_{1\star}^T, \ldots, \mathbf{f}_{P\star}]^T, \quad (9)$$

where $\mathbf{f}_{p\star}^T$ is the feature of the pth best AP.

B. Neural Network structure

For the positioning problem, we train a DNN to infer the location of a UE from the CSI feature. The structure of the DNN is depicted in Fig. 2. The DNN takes in a CSI feature and passes it through several fully connected layers until the output layer. In the figure, the number in each block represents the number of neurons in that layer. In all layers, the Rectified Linear Unit (ReLU) is used as the activation function, except the last layer where linear activation is used to generate the location of the UE.

In the training phase, the loss function of the predicted value against the ground truth value is computed. The Mean Squared Error (MSE) is considered as the loss function. The trainable parameters are then updated by back propagation. We use Batch Normalization (BN). All layers are initialized using the He initializer [16]. The training of the DNN takes place during an off-line phase with Adam optimizer. We assume that a dataset containing the CSI feature and corresponding ground truth location is available.

IV. Simulation Results

In our simulations, we consider 2D localization in an indoor office area with multiple rooms. The area is shown in Fig. 3.

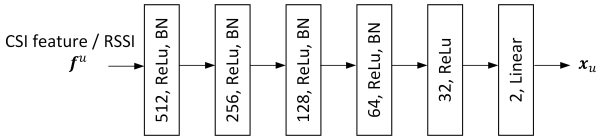

Fig. 2: DNN architecture for subarea/room localization.

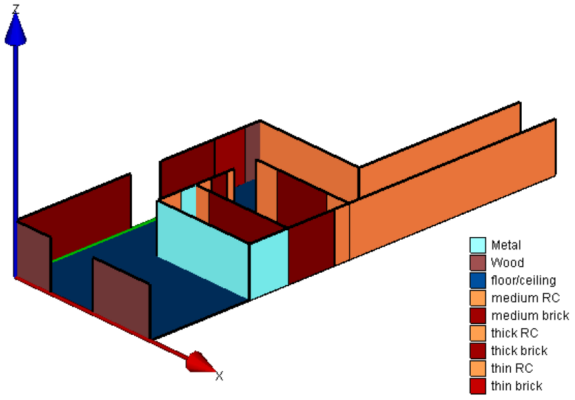

Fig. 3: 3D map of the floor plan. The color represents the construction material. The ceiling is removed for clarity.

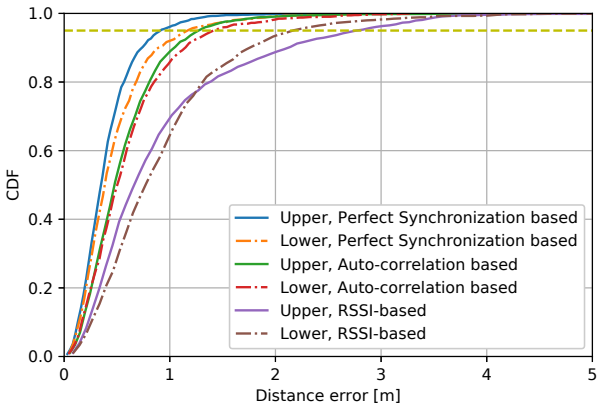

Fig. 4: CDF of the DE in [m] for different approaches. The CDF of the DE for the upper and lower areas are shown separately, since a different NN is trained for each area.

We use a range of different materials for the walls. We install 8 APs in the office area, each of which has 4 antennas. The carrier frequency is 2.4 GHz, the bandwidth is 20 MHz, and there are 1024 subcarriers. The AP transmit power is 27 dBm. The thermal noise power spectral density is -174 dBm/Hz, and the UE noise figure is 9 dB.

We generate a dataset of 17 thousand distinct locations with the corresponding CSI using the Wireless InSite channel simulator [17]. We split the dataset to 80% for training and 20% for testing.

We divide the floor plan into two subareas; the upper and lower subarea. The part of the building extending from the corridor to the room with metallic walls is in the upper area.

We consider the following positioning approaches:

- RSSI-based positioning: The received signal power from all 8 APs are measured and used as a feature to train a DNN for each subarea and then used for positioning.
- Perfect synchronization based positioning: We assume perfect network synchronization between all APs and all UEs. The magnitude of the delay domain CIR from all 8 APs is used as a feature to train a DNN for each subarea and then used for positioning.
- Autocorrelation based scalable positioning: for each UE we consider the CSI of an active set of the four strongest APs, i.e., $P = 4$. A feature vector as mentioned in Section III-A is constructed and used to train a DNN for each subarea and then used for positioning.

First we consider the statistics of the Distance Error (DE), i.e. the difference of the predicted and ground truth locations in the test set. Fig. 4 shows the CDF of DE obtained for different approaches for the test data in the upper and lower subareas. RSSI-based localization shows the worst performance, while perfect synchronization method is the best. The auto-correlation approach performs closer to the ideal perfect synchronization method than to the RSSI-method. Recall that the auto-correlation and RSSI based methods are robust to synchronization and phase error. The 95th percentile is indicated by the horizontal yellow dashed line.

Table I summarizes the accuracy of these approaches for the upper and lower subareas. We use the Mean Distance Error (MDE) and the 95% DE as the performance metrics to evaluate the accuracy of above discussed approaches. The gain of the autocorrelation based approach over the RSSI- based is striking, especially in the 95% accuracy measure. In addition, we evaluate the effect of wrong strongest AP estimation, by randomly permuting the order of AP features in the feature vector. We observe that the MDE is doubled.

It is instructive to analyze the positioning error as a function of location. In Fig. 5, the heat map of positioning error is depicted for perfect synchronization and auto-correlation based approaches for the test data, in the left and right half-figure, respectively. The location of the APs are also shown. The APs are installed on the ceiling and are inside the floor plan, but for readability they are depicted slightly outside of the area. The dashed white line indicates the split of the floor plan into two subareas. In most of the locations we have a good localization accuracy. The accuracy in the upper part of the corridor is somewhat degraded, since in this area only one AP has a strong received signal. The corridor is open ended, i.e., there are no reflections from the end. In this area, the ranging information available in the perfectly synchronized system from the weakly received APs in the the other rooms considerably improves performance as compared to the auto-correlator method. Also, in the closed room with partially metallic walls, we see some performance degradation.

TABLE I: Localization performance of fingerprinting positioning methods for lower and upper areas

Approach	MDE		95% DE	
	Upper	Lower	Upper	Lower
RSSI-based	0.93	0.92	2.82	2.10
Auto-correlation based scalable	0.56	0.60	1.25	1.42
Perfect synchronization based	0.40	0.46	0.95	1.17
Auto-corr - random permutation	1.2	1.4	-	-

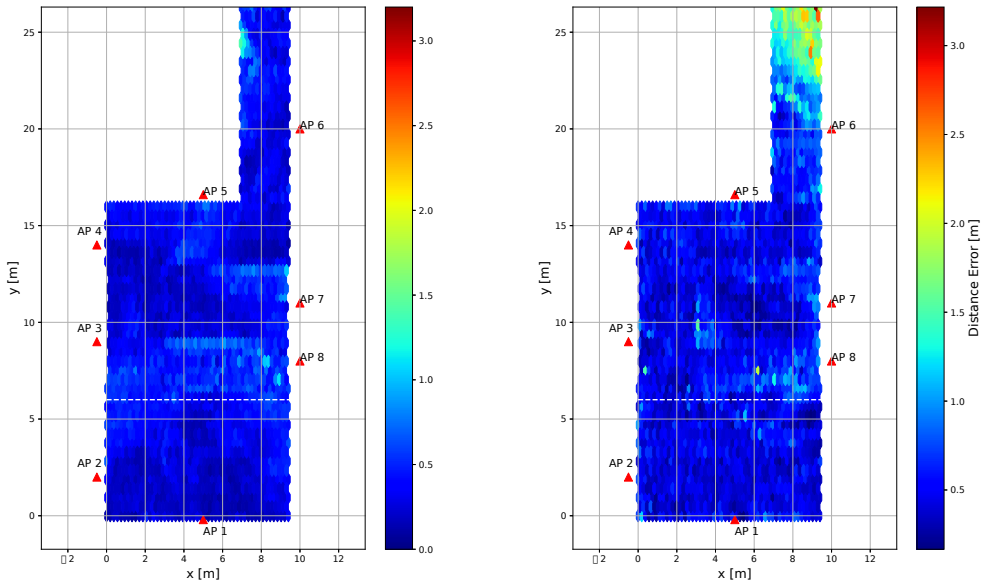

Fig. 5: Heat map of the distance error in [m] for (Left); perfect synchronization based approach (Right); scalable auto-correlation based approach. The white dashed line indicates the border of the two sub areas.

V. Conclusion

In this paper we have considered fingerprint based indoor localization. The estimated channel state information at the UE side was affected by synchronization errors and hardware impairments. A CSI feature resilient to such conditions was constructed based on the auto-correlation of the delay domain channel impulse response. We developed a scalable framework by considering the CSI features of the set of P strongest APs for each location with unique AP-identifiers. Simulation results showed that the localization accuracy loss compared to a non-scalable approach with perfect synchronization and CSI is only 20 cm.

Acknowledgment

This work was funded in part by the European Union under the framework of the project H2020-MSCA-ITN 813999 Windmill and the Academy of Finland (grant 319484).

References

[1] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu, "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2634–2657, 2020.

[2] F. Liu, J. Liu, Y. Yin, W. Wang, D. hai Hu, P. Chen, and Q. Niu, "Survey on WiFi-based indoor positioning techniques," *IET Commun.*, vol. 14, pp. 1372–1383, 2020.

[3] Q. Li, X. Liao, M. Liu, and S. Valaee, "Indoor localization based on CSI fingerprint by siamese convolution neural network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 12 168–12 173, 2021.

[4] C. Xiang, S. Zhang, S. Xu, X. Chen, S. Cao, G. C. Alexandropoulos, and V. K. N. Lau, "Robust sub-meter level indoor localization with a single WiFi access point—regression versus classification," *IEEE Access*, vol. 7, pp. 146 309–146 321, 2019.

[5] S. V. Prasad, E. Hossain, and V. K. Bhargava, "Machine learning methods for RSS-based user positioning in distributed massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8402–8417, 2018.

[6] M. Arnold, S. Dorner, S. Cammerer, and S. T. Brink, "On deep learning-based massive MIMO indoor user localization," in *Proc. IEEE Int. Workshop Sig. Processing Advances in Wireless Commun. (SPAWC)*, 2018, pp. 1–5.

[7] A. Mohammadian and C. Tellambura, "RF impairments in wireless transceivers: Phase noise, CFO, and IQ imbalance – a survey," *IEEE Access*, vol. 9, pp. 111 718–111 791, 2021.

[8] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "Reduced-complexity joint baseband compensation of phase noise and I/Q imbalance for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3450–3460, 2010.

[9] N. Tadayon, M. T. Rahman, S. Han, S. Valaee, and W. Yu, "Decimeter ranging with channel state information," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3453–3468, 2019.

[10] M. B. Khalilsarai, S. Stefanatos, G. Wunder, and G. Caire, "WiFi-based indoor localization via multi-band splicing and phase retrieval," in *Proc. IEEE Int. Workshop Sig. Processing Advances in Wireless Commun., (SPAWC)*, 2019, pp. 1–5.

[11] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[12] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving accurate CSI phases with commodity WiFi devices," in *Proc. IEEE Conf. Computer Commun., (INFOCOM)*, 2017, pp. 1–9.

[13] P. Ferrand, A. Decurninge, L. G. Ordoñez, and M. Guillaud, "Triplet-based wireless channel charting: Architecture and experiments," *IEEE J. Sel. Areas in Commun.*, vol. 39, no. 8, pp. 2361–2373, 2021.

[14] E. Gönültaş, E. Lei, J. Langerman, H. Huang, and C. Studer, "CSI-Based Multi-Antenna and Multi-Point Indoor Positioning Using Probability Fusion," *IEEE Trans. Wireless Commun.*, 2021.

[15] A. Sobehy, Éric Renault, and P. Muhlethaler, "CSI-MIMO: K-nearest neighbor applied to indoor localization," in *Proc. IEEE Int. Conf. Commun., (ICC)*, 2020, pp. 1–6.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 1026–1034.

[17] Remcom, "Wireless insite, https://www.remcom.com."