16th IPIN Competition 2026 - Track 7

5G Channel Impulse Responses with Sparse References

1 Introduction

Classical radio-frequency (RF) positioning in cluttered indoor environments is challenging, due to absorption, reflection, diffraction and scattering. However, modern algorithms, such as Fingerprinting [1], utilize those propagation effects contained in the received signals, to create a unique mapping from the radio signal to the position. Fingerprinting methods are widely used to estimate a rough position from narrow-band signals such as Wi-Fi or Bluetooth. However, with modern 5G new radio technologies, signals can be transmitted at higher bandwidths, enabling a much higher spatial resolution from which we can gain more precise position results by examining complex propagation conditions [2].

To leverage the benefits of the high spatial solution we can make use of the channel impulse response (CIR). Due to the high bandwidth, the reflections of the environment can be identified and exploit to create highly valuable fingerprints of the environment. Machine learning methods can utilize the high dimensional CIR to learn a end-to-end regression model for fingerprinting [3].

However, while channel state information (CSI) is easy to acquire, reference labels are not. To still be able to train a localization mode, methods like pretraining [4], [5], [6] or channel charting (CC) [7], [8], [9] have emerged.

2 Challenge Objectives

In this year's challenge, participants will receive CIRs from 8 distributed base stations (BSs) and are tasked to predict the user equipment (UE) position. To encourage solutions that leverage pretraining and/or channel charting techniques, we provide only 50 labeled reference samples. Therefore, the dataset presents several key challenges:

- **Multipath interference**: The recording environment exhibits significant multipath propagation
- **Signal degradation**: Certain areas experience partial or complete signal loss
- **Limited supervision**: The sparse reference data requires robust generalization strategies

Furthermore, due to high multipath interference and/or signal degradation, sometimes the CIR window is wrongly selected (and thus also has a wrong TOA), thus requiring robust algorithms capable of detecting and mitigating outlier measurements or exploiting temporal consistency through filtering techniques.

3 Environment and Measurement Setup

The dataset was recorded at the SNCF Technicentre in Bischheim, a railway maintenance facility featuring numerous large metallic obstacles including train compartments, access platforms, and maintenance equipment. These structures cause significant signal reflections and distortions, creating a challenging multipath-rich propagation environment. Figure 2 shows the recording environment.

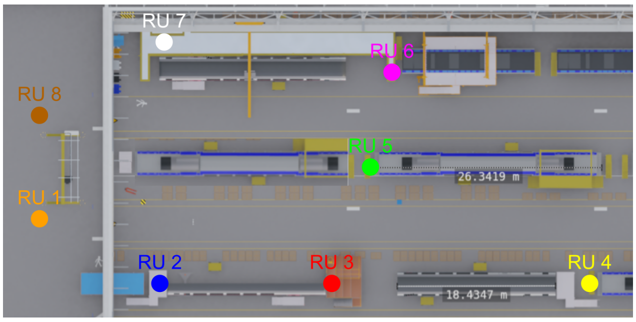

Figure 1: Map of the recording area. RU 4 is placed further away to simulate distant BS in larger deployments, while the rest are positioned directly around the recording area.

Figure 2: Image of the lower corridor of the recording environment. It is taken right of RU 1 looking in the direction of RU 4/5. There are many metallic objects, so there is lots of multipath interference in the data.

Figure 1 illustrates the spatial distribution of the 8 BSs deployed around the measurement area. BS 4 is intentionally positioned at a greater distance to emulate BSs in larger-scale deployments.

The UE is mounted on a mobile robot platform that operates at speeds up to $1.5\ \mathrm{m/s}$. It transmits 5G uplink Sounding Reference Signal (SRS) at a center frequency of $3.95\ \mathrm{GHz}$ and a bandwidth of $100\ \mathrm{MHz}$. Ground truth reference positions are captured using an iGPS system providing submillimeter accuracy at a sampling rate of $40\ \mathrm{Hz}$.

4 Dataset description

Datasets are provided as *HDF5* file. The file has a dataset called *ds* with the structure as described in Table 1. *CIR* contains the complex CIR set with the real/imaginary part in the first dimension, the different BSs in the second dimension, and the delay taps in the third dimension. CIRs are sampled at 122.880 MHz. The CIRs are not normalized and is therefore proportional to received signal strength. However, the UE might have adjusted the sending power, so be careful if using CIR's amplitude. *toa* contains the time of arrival (ToA) measurements in seconds for each CIR, relative to the first received CIR. Because CIRs may sometimes be missing due to signal blockage, resulting in incomplete datasets, we provide an *antenna_mask* that indicates for each antenna whether the data is valid (True) or invalid (False). The data timestamps (in seconds) are stored in *timestamp*. Note that the data is composed of different recordings, that you can differentiate by jumps in the timestamps. The *reference* field contains the 2D reference position. However, it will be mostly NaN with the exception of 50 samples. The training data can be downloaded at: `https://owncloud.fraunhofer.de/index.php/s/hDpvDXpqeJSqWWu`

Field Name	Shape	Data
CIR	(2, 8, 128)	float32
toa	(8)	float32
antenna_mask	(8)	bool
timestamp	()	float32
reference	(2)	float32

Table 1: Training Dataset Description

4.1 Testing

To validate your approaches, we plan to provide a similar dataset recorded in a different environment to the participants. Details will follow.

5 Competition Procedure

We will update this document and provide instructions on how to obtain the competition data and how to submit the final results before the competition.

5.1 Evaluation metrics

The Euclidean distance between estimated and true results (each 2D-positions) is the error metric, while the third quartile of the distribution of the errors is used as a performance metric.

6 Contact

For any questions or inquiries regarding this competition, please contact one of the track chairs:

- **Jonas Pirkl**: jonas.pirkl@iis.fraunhofer.de
- **Andreas Eidloth**: andreas.eidloth@iis.fraunhofer.de

To ensure a faster response, we recommend including both contacts in your email correspondence.

References

[1] M. Stahlke, T. Feigl, M. H. C. García, R. A. Stirling-Gallacher, J. Seitz, and C. Mutschler, “Transfer learning to adapt 5g ai-based fingerprint localization across environments,” in *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*, IEEE, 2022, pp. 1–5.

[2] A. Molisch, “Ultra-wide-band propagation channels,” *Proceedings of the IEEE*, vol. 97, pp. 353–371, Mar. 2009. DOI: 10.1109/JPROC.2008.2008836.

[3] A. Niitsoo, T. Edelhäußer, and C. Mutschler, “Convolutional neural networks for position estimation in tdoa-based locating systems,” in *Proc. 9th Intl. Conf. Indoor Positioning and Indoor Navigation, Nantes, France*, 2018, pp. 1–8.

[4] J. Ott, J. Pirkl, M. Stahlke, T. Feigl, and C. Mutschler, “Radio foundation models: Pre-training transformers for 5g-based indoor localization,” in *2024 14th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2024, pp. 1–6. DOI: 10.1109/IPIN62893.2024.10786154.

[5] J. Ott, M. Stahlke, T. Feigl, B. M. Eskofier, and C. Mutschler, *Simplicity is key: An unsupervised pretraining approach for sparse radio channels*, 2026. arXiv: 2505.13055 [eess.SP]. [Online]. Available: https://arxiv.org/abs/2505.13055.

[6] B. Guler, G. Geraci, and H. Jafarkhani, *A multi-task foundation model for wireless channel representation using contrastive and masked autoencoder learning*, 2025. arXiv: 2505.09160 [cs.LG]. [Online]. Available: https://arxiv.org/abs/2505.09160.

[7] C. Studer, S. Medjkouh, E. Gonultaş, T. Goldstein, and O. Tirkkonen, “Channel charting: Locating users within the radio environment using channel state information,” *IEEE Access*, vol. 6, no. 3, pp. 47 682–47 698, 2018.

[8] M. Stahlke, G. Yammine, T. Feigl, B. M. Eskofier, and C. Mutschler, “Velocity-based channel charting with spatial distribution map matching,” *IEEE J. of Indoor and Seamless Positioning and Navigation*, vol. 2, no. 1, pp. 230–239, 2024.

[9] P. Ferrand, A. Decurninge, L. G. Ordoñez, and M. Guillaud, “Triplet-based wireless channel charting,” in *IEEE Global Comms. Conf.*, Taipei, Taiwan, 2020, pp. 1–6.