Urban Navigation with LTE using a Large Antenna Array and Machine Learning

Whiton, Russ; Chen, Junshi; Johansson, Tobias; Tufvesson, Fredrik

Published in:
2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)

DOI:
10.1109/VTC2022-Spring54318.2022.9860844

2022

Urban Navigation with LTE using a Large Antenna Array and Machine Learning

Russ Whiton*†, Junshi Chen*‡, Tobias Johansson†, and Fredrik Tufvesson*
*Department of Electrical and Information Technology, Lund University, Lund, Sweden
Email:{russell.whiton, junshi.chen, fredrik.tufvesson}@eit.lth.se
‡Terranet AB, Lund, Sweden
†Volvo Car Corporation, SE-405 31 Gothenburg, Sweden
Email:tobias.johansson@volvocars.com

***Abstract*—Channel fingerprinting entails associating a point in space with measured properties of a received wireless signal. If the propagation environment for that point in space remains reasonably static with time, then a receiver with no knowledge of its own position experiencing a similar channel in the future might reasonably infer proximity to the original surveyed point. In this article, measurements of downlink LTE Common Reference Symbols from one sector of an eNodeB are used to generate channel fingerprints for a passenger vehicle driving through a dense urban environment without line-of-sight to the transmitter. Channel estimates in the global azimuthal-delay domain are used to create a navigation solution with meter-level accuracy around a city block.**

***Index Terms*—Channel Fingerprinting, Terrestrial Navigation, Software-Defined Radio, Convolutional Neural Networks.**

I. Introduction

Wireless signals are used for positioning on a massive scale, with end-users for such systems numbering in the billions. The expansion of Global Navigational Satellite System (GNSS) constellations and commodification of receivers means that any modern mass-market electronic consumer device including phones, watches or vehicles can tap into a network of high-accuracy positioning reference signals wherever the receiver antenna has an un-attenuated view of the sky [1]. The availability restriction of clean sky views, paired with other limitations like susceptibility to spoofing and jamming, create a desire for alternative wireless positioning solutions based on terrestrial transmitters. 4G Long-Term Evolution (LTE) commercial base stations provide a rich basis for opportunistic positioning [2].

Localization methods aside from trilateration (used in GNSS) are triangulation, proximity, and fingerprint matching, all of which have attracted attention in both industry and academia. The focus of this paper is on fingerprint matching using statistical tools, in which we seek to find a mapping from the received wireless signal to a position in space through domain knowledge about electromagnetic wave propagation and signal processing [3]. This offers advantages in that no knowledge of transmitter locations is required, it allows for capturing of channel behavior in complicated propagation environments that are tough to model parametrically, and position estimates are provided in a reference frame of the user's choice.

The simplest feature available to the user by most commercial wireless receivers is received signal strength, but the low dimensionality of this parameter precludes anything close to a bijective mapping in a large outdoor environment, and the combination of large- and small-scale fading weakens the correlation even further in heavy multipath environments where alternatives to GNSS are most desirable. To benefit from the full richness of the wireless channel information, it is desirable to use an observation model that includes the geometrical information embedded in multipath propagation [4]. Channel State Information (CSI) is a raw form of measurement data, a complex-value for each subcarrier of the LTE signal per antenna. With a sufficiently large data set and network architecture, end-to-end learning from this information should be able to extract the relevant features. However, feature engineering allows for more computationally tractable and memory-efficient network design, portability to different hardware and software platforms, intuition about what the network is "learning" and the ability to augment the network with external information to attain better results with limited training data.

The problem of channel representation for outdoor fingerprinting with a single sector of a commercial base station is an interesting benchmark. In [5], 11 channel representations were investigated and 75-meter median error was achieved. In [6], 19-meter median accuracy was achieved with CSI vectors. [7] considers the problem of tracking with an angular-delay representation of the channel. Our research builds on these ideas, and contributions are as follows:

- We apply and extend [7], which suggests using an azimuthal-delay representation of the received signal as input to a Convolutional Neural Network (CNN). We demonstrate that the choice of a *global* azimuthal-delay domain allows for effective fingerprinting not only with CNNs, but through a non-parametric k-Nearest Neighbors (kNN) estimator.
- Downlink channel measurements are visualized in a world reference frame to motivate the choice of a global azimuth-delay domain representation.
- Results for the two estimators (CNN and kNN) are shown for both raw data and with a Kalman filter applied.

We demonstrate that without line-of-sight, a meter-level navigation solution can be generated with signals of limited bandwidth and no input regarding transmitter location. These results are predicated on having known User Equipment (UE) orientation and clock synchronization.

II. Measurement Setup

Cell-Specific Reference Symbols (CRS) in LTE are signals that are known to the UE in advance as a function of physical cell identity, with 504 possible sequences distributed across resource blocks to facilitate down-link channel estimation [8]. What makes CRS appealing for channel estimation also makes them appealing for opportunistic localization, in that they are open, frequently transmitted, known to the UE *a priori* and of higher bandwidth than the Primary and Secondary Synchronization Sequences.

A system was built for receiving and logging CRS symbols from commercial LTE base stations operating at 2.6 GHz, based on a National Instruments USRP-2953R Software-Defined Radio. Sampling of CRS symbols was done at a rate of 30.72 MHz, or approximately 32.56 ns per sample and logged onto a laptop hard drive. The USRP-2953R also controlled a 128-element switched Stacked Uniform Circular Array consisting of 4 vertically-stacked rings of 16 dual-polarized antenna elements each, which was mounted on the roof of a passenger vehicle. A rubidium standard, disciplined by GPS, was used as a frequency reference and a GPS receiver provided an absolute time reference for correlation with localization ground truth data.

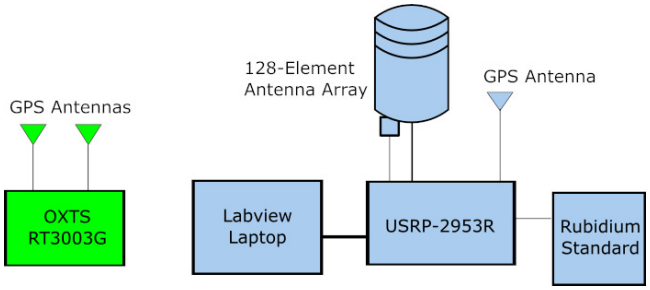

Fig. 1. Block diagram of key measurement system components.

A high-end localization reference system [9] with a survey-grade GNSS receiver and high-grade Inertial Measurement Unit (IMU) was used to provide estimates of vehicle and antenna position and orientation, through Post-Processed Real-Time Kinematics. A block diagram of key system components is shown in Figure 1, and a photo of the test vehicle with all components mounted is shown in Figure 2.

Four laps of around 400 meters each were driven around one block in central Lund, Sweden for an aggregate distance of 1628 meters over 28 minutes and 17 seconds at an average speed of 0.96 m/s (driving speed was limited to 1.5 m/s to avoid exceeding the channel coherence time[1]). The driving

[1]This is a limitation of the measurement system, which switches at the rate of CRS symbol transmission. Sampling antennas in parallel would negate limitations in speed.

Fig. 2. Photo of array antenna mounted on Volvo V90 Passenger Vehicle.

area spanned 103 meters East-West and 108 meters North-South. Two laps were driven in each direction, as depicted in Figure 3.

Fig. 3. Visualization of Ground Truth position for one of four laps driven in Lund, Sweden (55.71°N, 13.19°E). The approximate position of the Base Station (eNodeB) is displayed as well. Image created with Google Earth.

The surrounding buildings consist of four- to five-story residential or mixed-use buildings. The receive antenna array never had line-of-sight to the transmitting eNodeB, as the line-of-sight trajectory was always obstructed by multi-story buildings.

A sweep among all antenna elements (a channel snapshot) was completed every 75 LTE subframes (75 ms), and a total of 22,626 75-ms snapshots were generated over the drive trajectory.

III. Data Representation and Estimators

A. Channel and Position Representation

For every snapshot at time index i, a channel estimate in the UE frame, Y_i^{UE}, is associated with a localization vector x_i. Channel estimation is done using the Expectation-Maximization-based RIMAX algorithm [10], and parameterized into a variable and discrete number L_i of multipath components, e.g., $Y_i^{UE} \triangleq [\tau_l, \alpha_l, \theta_l^{UE}, \phi_l^{UE}]$ where l spans

from 1 to L_i. τ_l represents component delay[2], α_l represents component signal strength, and θ_l^{UE}, ϕ_l^{UE} represent component azimuth- and elevation-of-arrival in the UE frame, respectively. Doppler was not estimated, under the assumption that reflecting objects which contribute significant energy to the channel impulse response were static and that the ego-vehicle was moving slowly.

In [7], simulated channel data consisting of several multi-path clusters based on a COST 2100 channel model realization are processed into the azimuth-delay domain, motivated as a sparse channel representation that maps to position $x_i \in \mathbb{R}^{2\times 1}$ in a manner described as approaching bijective for most practical propagation scenarios. For this to hold true for a static transmitter and a mobile UE unconstrained in both position and orientation, a more complete parameterization of x_i is necessary. A complete 6 degree-of-freedom representation is $x_i \in \mathbb{R}^{6\times 1} = [e_i, n_i, u_i, \gamma_i, \lambda_i, \eta_i]^T$, including three-dimensional position e_i, n_i, u_i (East-North-Up Cartesian Frame) and three-dimensional orientation[3] $\gamma_i, \lambda_i, \eta_i$ (yaw, pitch, and roll [11]). We extend [7] through our use of external orientation information $\gamma_i, \lambda_i, \eta_i$ to rotate the angles-of-arrival of multipath components from the UE coordinate system $y_{l,i}^{UE}$ to express them in an East-North-Up coordinate system $y_{l,i}^{ENU}$, expressed in Cartesian form in Equation 1.

$$y_{l,i}^{ENU} = y_{l,i}^{UE} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\eta_i) & sin(\eta_i) \\ 0 & -sin(\eta_i) & cos(\eta_i) \end{bmatrix} \begin{bmatrix} cos(\lambda_i) & 0 & -sin(\lambda_i) \\ 0 & 1 & 0 \\ sin(\lambda_i) & 0 & cos(\lambda_i) \end{bmatrix} \begin{bmatrix} cos(\gamma_i) & sin(\gamma_i) & 0 \\ -sin(\gamma_i) & cos(\gamma_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The components $y_{l,i}^{ENU}$ constitute the global angular representation of the channel, $Y_i^{ENU} \triangleq [\tau_l, \alpha_l, \theta_l^{ENU}, \phi_l^{ENU}]$. The footprints were discretized into a matrix of azimuthal-delay bins with M delay bins and N azimuthal bins, $\mathbf{H}_i \in \mathbb{R}^{M\times N}$ with signal power α_l contributing to the power for a given azimuthal-delay bin. Elevation values are not considered in the formulation of $\mathbf{H}_i$, but extension to three dimensions, e.g., [12] is possible. The *M* azimuthal bins can be interpreted as a discretization of the full span of compass directions. We ignore the altitude u_i component of the position because the route does not have variable elevation for the same East and North coordinates[4], leaving us with the same two-dimensional localization estimate $x_i \in \mathbb{R}^{2\times 1} = [e_i, n_i]^T$ that now accounts for freedom in orientation. A desirable extension which we leave for future work is to add orientation (primarily heading) as an estimable parameter.

[2]Delay measurements (equivalently expressed in meters) were biased, but not time-varying based on a visual inspection of delay domain data over the full test drive. Built into this representation is effectively absolute clock synchronization between transmitter and receiver, a parameter which is typically estimated on a per-epoch basis for time-of-arrival systems built on low-cost crystal oscillators on the UE side. CNNs are effective if the shape of information features is invariant to the location in the input space, but absolute time synchronization provides a large amount of information about proximity to the base station, even if the mapping is highly non-linear for an urban canyon scenario. We leave the investigation of sensitivity to time synchronization for future work.

[3]Non-isotropic UE antenna pattern and variable propagation environment (vegetation or dynamic objects such as traffic) will also preclude a bijective representation. Effective channel footprinting entails finding representations of Y_i that are both sparse and robust to these types of effects.

TABLE I
CNN ESTIMATOR DESCRIPTION

Parameter	Value
Input Layer Dimensions	69 x 71
Delay Range (meters / ns)	0 to 700 / 0 to 2333
Azimuthal Range (degrees)	-180 to 180
Delay Resolution (meters / ns)	10.1 / 33.8
Azimuthal Resolution (degrees)	5.1
CAP Layer 1 Filters / Weights	16 / 272
CAP Layer 2 Filters / Weights	32 / 8224
CAP Layer 3 Filters / Weights	64 / 32832
Dense Layer 1 Weights	524416
Kernel Size (All conv. layers)	3
Pool Size (All pooling layers)	2x2
Output Layer Shape	2

B. Data Aggregation

Training on individual snapshots provides both higher-frequency position estimates at run-time and a larger training data set, but averaging over a longer time interval allows for more energy to be aggregated in each estimate. Using discretized multipath components as the basis for populating the matrix $\mathbf{H}_i$ results in very few non-zero angle-delay bins for an individual snapshot if the number of components L_i is low, resulting in most subsequent estimator operations being multiplication of zeros. Therefore, energy contributions from snapshots are summed, 20 at a time, into snapshot groups (indexed hereafter with k) to produce position estimates every 1.5 seconds, e.g., $\mathbf{G}_k = \sum_{i=1+20(k-1)}^{20+20(k-1)} \mathbf{H}_i$.

In a multi-sensor navigation system, unbiased and low-rate estimates of absolute position are complementary to high-rate sensors such as IMUs or wheel odometry with unbounded long-term bias [13], so for this application it is reasonable to produce position estimates at a lower rate.

C. Estimators

1) Convolutional Neural Network Estimator: The CNN estimator consists of three Convolution-Activation-Pooling (CAP) layers, a flattening layer, a fully-connected layer and an output layer estimating position e_k, n_k for each snapshot group k. Relevant parameters of the CNN estimator are summarized in Table I.

Moderate changes to parameters including number of layers, filters, kernel size and learning rate did not significantly impact performance and are not included in this text. Network optimizing is left to future work.

[4]Some driving scenarios could benefit from altitude estimates, such as parking garages, freeway over- and under-passes, stacked highways, etc.

TABLE II
KNN ESTIMATOR DESCRIPTION

Parameter	Value
Input Dimensions	4899 x 1
Delay Range (meters / ns)	0 to 700 / 0 to 2333
Azimuthal Range (degrees)	-180 to 180
Delay Resolution (meters / ns)	10.1 / 33.8
Azimuthal Resolution (degrees)	5.1
Neighbors n used	5
Neighbor weights	Uniform

2) k-Nearest Neighbors Estimator: A kNN estimator was used, with the intention of establishing a performance benchmark for the CNN using the same underlying input data. Unlike the CNN estimator, there are no learned parameters. The two-dimensional azimuthal-delay representation from the CNN is flattened, and the same angular-frequency bins are used without any convolution operation that considers adjacent values within Y_i^{UE}. The estimator is described in Table II.

D. Kalman Filtering

For each estimator, a Kalman Filter employing a constant velocity model is used to filter the series of independently predicted points from the kNN and CNN estimators. The state vector includes two-dimensional positions and velocities, e.g., $x_k = [e_k, n_k, e'_k, n'_k]^T$. The filter is initialized at $x_{k=1}$ with position estimates from the estimator predictions after the first snapshot group, velocities of zero (e.g., $\hat{e}'_{k=1} = \hat{n}'_{k=1} = 0$) and a covariance matrix $P_{k=1} = diag([9, 9, 1, 1])$.

IV. RESULTS

A. Channel Footprints

Azimuthal-delay domain spectra for four geographically proximate snapshot groups (spread within 2 meters), one on each lap, are shown in Figure 4. A visual inspection of these channel footprints shows visibly similar clusters, even with position offsets on the order of tens of wavelengths and heading differences of 180 degrees, and dynamic environmental changes such as size and location of other vehicles. This supports the choice of global azimuthal-delay representation of $\mathbf{H}_i$.

It is also clear that significant energy is contained in clusters that have path differences on the order of several hundred meters (several hundred nanoseconds), and that the channel estimation shows "smeared" angle-of-arrival estimates where clusters appear to occupy a broad range of azimuthal values.

B. Estimator Performance

Laps 1-3 were used as a training set in the case of the CNN and as the underlying search space in the case of the kNN estimator. Lap 4 was used for testing. The positioning performance for Lap 4 with each estimator is shown as a time series and error distribution function in Figure 5 and as a map in Figure 6. Performance metrics are listed in Table III. The kNN estimator is more prone to extreme errors, but still manages to produce a navigation solution that traces the trajectory of UE in a counter-clockwise manner around the block. Each estimator shows similar directional performance, with the exception of high East-West error for the kNN, possibly caused by a single point. This represents a performance improvement as compared with [5], [6], but different environmental conditions, survey areas and densities preclude such comparisons being considered comprehensive.

The CNN appears to be a superior estimator, likely owing to the ability to better integrate larger quantities of data in the individual estimates rather than discarding all but a few points, some of which may contain large errors.

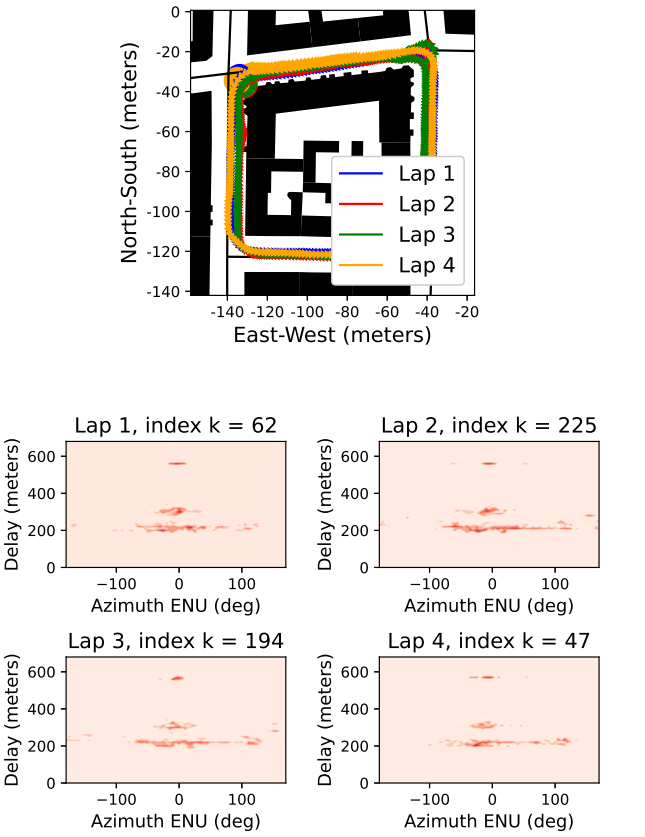

Fig. 4. Azimuthal (East-North) - Delay spectra for one snapshot group on each of four laps around Central Lund, Sweden. e_0, n_0 corresponds to 55.7107°N, 13.1898°E, and a Universal Transverse Mercator projection is applied to generate a two-dimensional representation of position. The points are taken from the Northwest corner of the route.

TABLE III
HORIZONTAL ERROR VALUES FOR ESTIMATORS

Estimator	Avg. Error (m)	Error Std. (m)	Max Error (m)
CNN	6.2	4.9	35.9
kNN	10.4	17.3	108.8
CNN (Filtered)	5.7	2.4	12.1
kNN (Filtered)	7.2	4.5	25.9

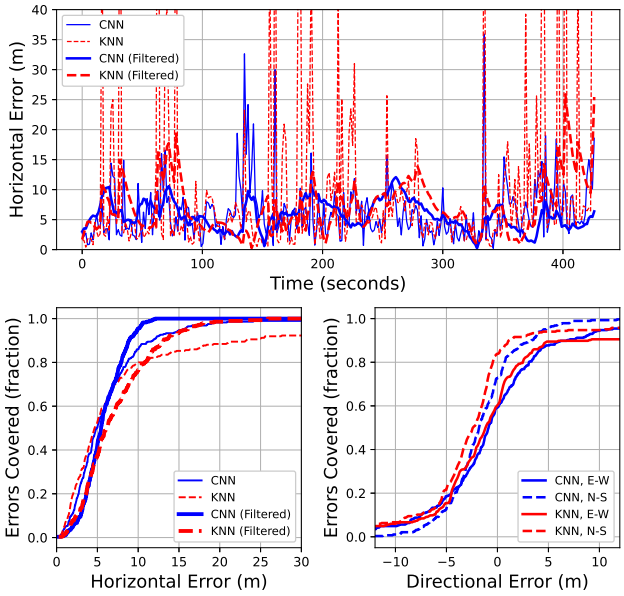

Fig. 5. Time series and cumulative distribution functions of absolute and directional position errors for Lap 4.

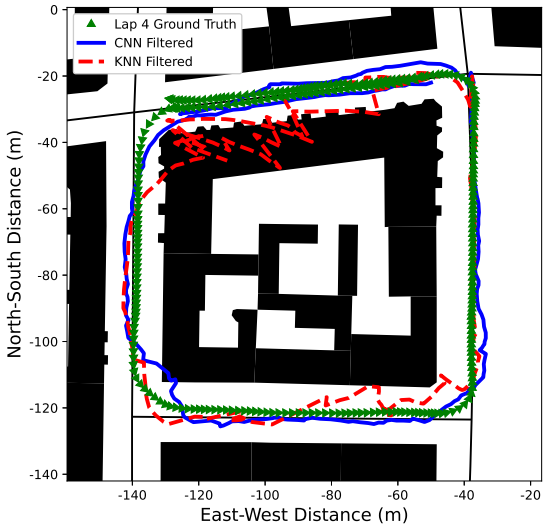

Fig. 6. Ground Truth vehicle trajectory, as well as Filtered Position Estimates from kNN and CNN.

V. Conclusions and Further Work

The results show that with a single sector of a single eNodeB in non-line-of-sight conditions, meter-level estimations of absolute position can be generated based on limited drive data over the same road segment through an azimuthal-delay representation of the wireless channel. They show that the statistics of the wireless channel can be used to create a navigation solution even in non line-of-sight conditions. The estimates are augmented by external information regarding UE orientation and presume effective clock synchronization with the transmitter.

In future work, it would be interesting to alter the CNN architecture including input and output parameters, investigate sensitivity to timing errors, and to integrate multiple sectors and base stations.

Acknowledgment

This work was financed by the Swedish Innovation Agency VINNOVA through the MIMO-PAD Project (Reference number 2018-05000).

The authors would like to thank Martin Nilsson for his help in setting up the measurement system.

References

[1] N. Joubert, T. G. Reid, and F. Noble, "Developments in Modern GNSS and Its Impact on Autonomous Vehicle Architectures," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, pp. 2029–2036.

[2] Z. Kassas, J. Khalife, K. Shamaei, and J. Morales, "I Hear, Therefore I Know Where I Am: Compensating for GNSS Limitations with Cellular Signals," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 111–124, Sep. 2017.

[3] D. Burghal, A. T. Ravi, V. Rao, A. A. Alghafis, and A. F. Molisch, "A Comprehensive Survey of Machine Learning Based Localization with Wireless Signals," 2020. [Online]. Available: https://arxiv.org/pdf/2012.11171.pdf

[4] K. Witrisal, P. Meissner, E. Leitinger, Y. Shen, C. Gustafson, F. Tufvesson, K. Haneda, D. Dardari, A. F. Molisch, A. Conti *et al.*, "High-Accuracy Localization for Assisted Living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 59–70, 2016.

[5] X. Ye, X. Yin, X. Cai, A. P. Yuste, and H. Xu, "Neural-Network-Assisted UE Localization Using Radio-Channel Fingerprints in LTE Networks," *Ieee Access*, vol. 5, pp. 12 071–12 087, 2017.

[6] H. Zhang, Z. Zhang, S. Zhang, S. Xu, and S. Cao, "Fingerprint-based Localization using Commercial LTE Signals: A Field-Trial Study," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. IEEE, 2019, pp. 1–5.

[7] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive mimo fingerprint-based positioning," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017, pp. 1–6.

[8] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-advanced for mobile broadband*. Academic press, 2013.

[9] Oxford Technical Solutions Ltd, "Rt3000 v3," 2020, online; accessed 06-December-2021. [Online]. Available: https://www.oxts.com/products/rt3000/

[10] A. Richter, "Estimation of Radio Channel Parameters, Models and Algorithms," Ph.D. dissertation, Technische Universität Ilmenau, Germany, 2005.

[11] International Organization for Standardization, "Road vehicles — Vehicle dynamics and road-holding ability — Vocabulary," International Organization for Standardization, Geneva, CH, Standard, dec 2011.

[12] C. Wu, X. Yi, W. Wang, L. You, Q. Huang, X. Gao, and Q. Liu, "Learning to localize: A 3D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, 2021.

[13] P. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, Second Edition*. Artech House, 03 2013.