Absolute Positioning with Unsupervised Multipoint Channel Charting for 5G Networks

Jaakko Pihlajasalo, Mike Koivisto, Jukka Talvitie, Simo Ali-Löytty, Mikko Valkama
Tampere University
Tampere, Finland
{jaakko.pihlajasalo, mike.koivisto, jukka.talvitie, simo.ali-loytty, mikko.valkama}@tuni.fi

***Abstract*—The 5th generation mobile networks introduce large bandwidths with extended beamforming capabilities, which results in increased spatial selectivity of received channel state information. A channel chart is a map of the radio geometry that surrounds the base station and it can be generated in an unsupervised manner from the received channel state information without any knowledge of actual measurement locations. In this work, we generate channel charts for multiple base stations using multidimensional scaling and combine them for a better shape of the radio geometry. The combined chart cannot be directly applied for absolute positioning, but it can be extended. Extension is performed with affine and conformal mappings to the charts. The method for generating and combining the charts as well the method of extension to absolute positioning is explained. Evaluations are performed for two different scenarios, one of which is an open space scenario, and the other utilizes ray-tracing data. Finally, the charts are presented and analyzed together with positioning estimation results for the considered scenarios.**

I. Introduction

The 5th generation (5G) wireless networks continuously process vast amounts of information in order to meet the demanding communication requirements in terms of the data rate and capacity [1]. Along with a normal data transmission, user equipment (UE) devices are transmitting uplink pilot signals which facilitate uplink channel state information (CSI) estimation and subsequent data demodulation at the network-side. In addition to the general communications, the CSI observations can also be used to create a radio geometry map of the environment which in turn can be utilized for enhancing network functionalities by leveraging the available device position information [2]. Hence, channel charts (CC), which represent the positions of the UE devices through preserving the radio geometry of the environment [3], can be beneficial especially in the cases where the absolute device position is not readily available. In this work, we propose an unsupervised method for generating a CC which maintains the shape and size of the original map.

In the absence of measurement location information, the CCs should be generated by using only the obtained CSI measurements in an unsupervised manner. This way, functions can be automated and the use of laborious and sometimes expensive fingerprinting campaigns and the use of global navigation satellite systems (GNSS), can be avoided. In practice, the CCs can be generated for each base station (BS) separately, but if carried out this way, the shape of the charts can be distorted [3], [4]. Thus, measurements from multiple BSs can improve the charts significantly [5], if a proper measurement fusion method is considered in the distance matrix determination phase of the charting. Stemming from the aforementioned work, we propose a distance matrix combination method that relies on a weighted average based on the visibility condition of the UE devices. This method can be performed in an unsupervised manner and maintains the shape of the radio geometry relative to the BSs.

Even though the unsupervised positioning setting is attractive, it can also be challenging due to the dynamic nature of the channel, as the moving vehicles and people, weather conditions and multipath for buildings affect the channel. Due to this, the unsupervised model might need to be updated regularly. In addition, without certainty of UE positions related to the received CSI measurements, only using the data within an examined area can be challenging.

Even though the multipoint CCs presented in [5] have better shape, the charts are not in the correct scale and orientation, in comparison with the real coordinates. Additionally in [6], Siamese Neural Network is used for unsupervised as well as supervised and semisupervised CC, but absolute positioning is not accomplished for unsupervised CC. Therefore, the unsupervised charts in [5] and [6] are not suitable for absolute device positioning as such. In this work, we introduce an unsupervised method for extending the multipoint CC to absolute device positioning. This can be achieved with different transformations such as affine and conformal mapping. These mappings can be applied to the points of the CC in order to acquire a chart with better orientation and shape for absolute positioning. However, these transformations require knowledge of the BS positions in the CC to function. Therefore, we introduce a method for estimating the corresponding BS positions by utilizing the power of the CSI measurements.

This article is structured as follows. In Section II, the CC and a chart combination method are presented. In Section III,

This work was supported in part by the Finnish Funding Agency for Innovation (Business Finland) under the project 5G-FORCE, and in part by the Tampere University Graduate School.

the method for extension to absolute positioning is described. In Section IV, the simulations and performance of the proposed methods are presented. In Section V, we conclude the article.

II. Channel charting

The goal of channel charting is to create a mapping from a high dimensional CSI measurement $h \in \mathbb{C}^M$ to a coordinate $z \in \mathbb{R}^2$ in such a way that the local geometry is preserved. In this work, the considered multicarrier-multiantenna CSI measurement h has the dimensionality $M = B \cdot W$, where B is the number of antenna elements in the antenna array and W is the number of active subcarriers in the system. In the conducted simulations, carried out using extensive ray-tracing channel modeling [7], we employ an uniform circular array (UCA) with 20 antenna elements and 16 active subcarriers at a carrier frequency of 3.5 GHz. In the actual charting, the obtained CSI measurements are first used to compute the raw 2nd moment of the CSI vectors. These instantaneous matrices are then averaged over a few time-steps in order to cancel out the effects of fast fading, and the acquired averaged matrices are denoted as $\mathrm{H} \in \mathbb{C}^{m \times m}$. In [4], the CSI vectors were averaged over the subcarriers due to the assumed line of sight (LoS) conditions, whereas in this work, non line of sight (NLoS) measurements are also taken into account, and thus the averaging over the subcarriers is not carried out.

The obtained matrices H are used to compute the distance matrices for each BS, which are then used in multidimensional scaling (MDS). The pairwise distance matrix $\mathrm{D}^{(k)}$ of the kth BS is computed in a similar way as in [4] such as

$$\mathrm{D}_{i,j}^{(k)} = \left|\left| \frac{\mathrm{H}_i}{\|\mathrm{H}_i\|_F^\beta} - \frac{\mathrm{H}_j}{\|\mathrm{H}_j\|_F^\beta} \right|\right|_F, \text{ for } i,j = 1,...,N, \quad (1)$$

where H_i is the CSI matrix of the ith measurement, N denotes the number of available CSI matrices, and parameter $\beta = 1 + \frac{1}{2\sigma}$ with parameter σ that corresponds to the unknown path-loss exponent. We chose $\sigma = 8$ Finally, $\|\cdot\|_F$ denotes the Frobenius norm, which can be roughly interpreted as a power of the given CSI measurement.

The received distance matrices from each BS are combined by taking a weighted average of the matrices, where the weighting is based on the visibility of the users to the BS. We can assume that the distance estimates in matrices $D^{(k)}$ between LoS measurements are more accurate than NLoS measurements. Hence, the combined distance matrix will use only LoS measurements, when possible. In a real scenario it is not possible to directly say which measurements are LoS and which are NLoS. Thus, an estimation of the visibility is needed. In this work, simulations and numerical evaluations are only performed with perfect visibility knowledge as estimation of visibility is beyond the scope of this work. Some methods for visibility estimation are presented in [8].

The visibility information is utilized to combine the distance matrices from different BS. In the proposed solution visibility value of a pair of users visible to the BS is set to one, whereas if only one or neither of the users are visible, the value is zero. Thereafter, we set each distance with a zero visibility value to zero and sum the distances. Finally, we divide each distance by the number of distances used. These values can be represented in the following way.

$$\mathrm{D}_{i,j} = \begin{cases} \frac{\sum_{k=1}^{n_{BS}} v_{i,j}^{(k)} \mathrm{D}_{i,j}^{(k)}}{\sum_{k=1}^{n_{BS}} v_{i,j}^{(k)}} & \text{if} \quad \sum_{k=1}^{n_{BS}} v_{i,j}^{(k)} > 0 \\ 0 & \text{if} \quad \sum_{k=1}^{n_{BS}} v_{i,j}^{(k)} = 0 \end{cases}, \quad (2)$$

where $D_{i,j}^{(k)}$ is the distance value of the point-pair (i,j) in the distance matrix of the kth BS, and $v_{i,j}^{(k)}$ is the visibility value of the measurement pair (i,j) for kth BS. This is essentially the average of the distances of the visible measurements.

The combined distance matrix D is then used as an input to MDS. In this work, we will use stress function [9] and Sammon's mapping (SM) [10] as cost functions for evaluations. When the stress-based approach is utilized in MDS, the missing values, which are the off-diagonal zero valued elements, are ignored. In SM, the missing values are replaced with the average of the NLoS distances due to implementational issues.

To avoid computing the CC again for a new measurement, a recursive update for MDS can be beneficial. In applications, it would be beneficial that the rest of the CC is not changed due to the new measurement. Thus, we propose an estimation method for a new measurement to be located on the CC. We derive the update step for the "stress" cost function, but the derivation is similar for other cost functions. MDS minimizes the stress, which is a root of sum of squared residuals

$$\text{Stress}_D(\mathbf{z}_1,...,\mathbf{z}_N) = \left(\sum_{i \neq j = 1,...,N} (d_{ij} - \|\mathbf{z}_i - \mathbf{z}_j\|)^2 \right)^{1/2}, \quad (3)$$

where D is the distance matrix, d_{ij} is the (i,j)-th element of D, N is the number of points and $\mathbf{z}_i$ is the ith coordinate in the estimated map [9]. After minimizing this cost, we have estimated positions $\mathbf{z}_i$ for N points in the CC. The estimated position of a new measurement $\mathbf{z}_{N+1}$ can be estimated by minimizing

$$\text{Stress}_D(\mathbf{z}_{N+1}) = \left(\sum_{i = 1,...,N} (d_{i,N+1} - \|\mathbf{z}_i - \mathbf{z}_{N+1}\|)^2 \right)^{1/2}, \quad (4)$$

where we treat positions $\mathbf{z}_i$ solved with (3) as parameters, and only variables to be solved are the coordinates of the new measurement $\mathbf{z}_{N+1}$.

III. Extension to Absolute Positioning

The mapping obtained from the MDS is not in the real scale, shape and translation and thus it cannot be used for

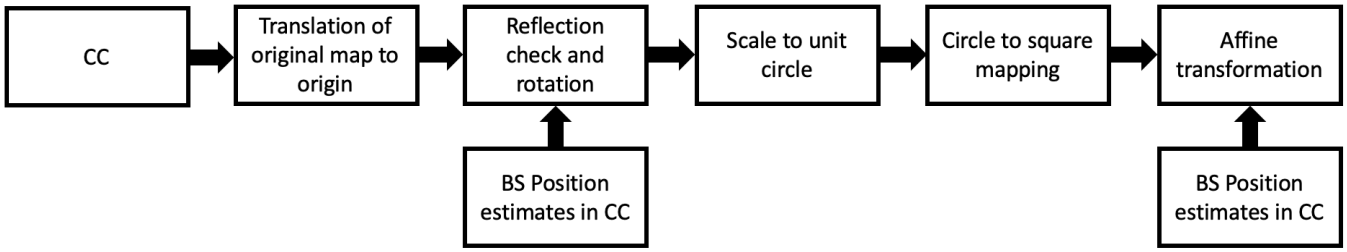


Fig. 1: A flowchart of the steps made for the CC extension to absolute scale.

absolute positioning. The resulting maps have some rotation, compared with real coordinates, they are in the wrong scale, and the coordinates are zero-centered. An affine transformation can be applied to rotate, scale and mirror the map to better represent the original map. The affine transformation needs some reference points from real coordinates. The order of the steps for the extension procedure can be seen in Fig. 1.

Since the learning of the CC is carried out in an unsupervised manner, no reference points are available. The real coordinates of the BSs are known, which can be used to estimate the transformations between the map and the CC. The coordinates of the BS in the CC are not known, but can be estimated using the power of the measurements. The estimated BS positions can then be used to apply necessary transformations to the CC.

Additionally, the chart can be circular, even though the measurement area itself can be rectangular, as in our scenarios. A conformal mapping, which is an angle-preserving transformation, from a circle to a square can be used to remove the roundness of the chart.

A. Estimation of BS position

The positions of the BSs can be estimated based on the power of the CSI measurements. We can assume that measurements with the highest power are close to the BS. The position of the BS in the CC can be estimated using measurements surrounding the BS. Locate some number of measurements with the highest power from the CC. In evaluations, we use 10 measurements with the highest powers. The BS position estimate can then be calculated as the average of the positions of such points on the chart.

Additionally, more sophisticated methods could be used to estimate the BS position. In [11] access points (AP) position was estimated using path loss models. BS position estimation could be replaced by a few measured reference positions, which would turn the problem into a semisupervised problem.

B. Affine and conformal transformations

The BS position estimates can be used to find the correct rotation between the CC and the known BS positions in the real coordinates. The rotation angle is found by calculating the angle between the segments from the origin to the real and estimated BS positions. First, the real map is translated to the origin, by subtracting the mean value of the BS positions from each BS position. Angles between each position and estimate pair can be calculated with the two-argument arctangent function by

$$\theta = \sphericalangle(\mathbf{x}, \mathbf{y}) = \text{atan2}(y_1 x_2 - y_2 x_1, y_1 x_1 + y_2 x_2), \quad (5)$$

where $\mathbf{x}$ is the two-dimensional coordinate of the BS in real coordinates that are translated to the origin, $\mathbf{y}$ is the estimated position of the BS in the CC, and subscript i corresponds to the ith coordinate of the position. The angles are calculated for each BS. The angle estimate $\bar{\theta}$ is the circular mean of angles θ given by

$$\bar{\theta} = \text{atan2}\left(\frac{1}{n}\sum_{j=1}^{n}\sin\theta_j, \frac{1}{n}\sum_{j=1}^{n}\cos\theta_j\right). \quad (6)$$

However, the resulting chart from the MDS can be reflected and the angle of rotation can thus be erroneous. Reflection can be checked, for example, by computing the variance of the rotations given by (5). If the variance is high, for example 1, the points can be reflected on either axis by choosing $z_i = -z_i$, where z_i is the ith coordinate of point $\mathbf{z}$ in the CC. Each point in the chart can then be rotated by multiplying each coordinate with a rotation matrix.

$$\mathbf{z}_{\text{rotated}} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \mathbf{z}, \quad (7)$$

where $\mathbf{z}$ is the two-dimensional coordinates of a point in the CC and $\mathbf{z}_{\text{rotated}}$ is the coordinates of a point that has been rotated.

The rotation is followed by the conformal mapping. In this work, conformal mapping is used to map a circle to a square, to remove some of the roundness of the charts. There are a few ways to map the circle to a square, which are explained in detail in [12]. We use the elliptical grid mapping, which is carried out on points inside a unit circle centered at the origin to a square that circumscribes the circle. Since points in the CC are not in a unit circle, the points in CC are scaled as such to avoid numerical problems with the square root. Scaling can be computed in polar coordinates, but it is equivalent to scaling by the maximum distance of a point to origin

$$\mathbf{y}_{\text{scaled}} = \frac{\mathbf{y}}{\max_i(\|\mathbf{y}_i\|_2)}, \quad (8)$$

where $\mathbf{y}_i$ is the ith point. This can be derived from the polar coordinate expression. Elliptical grid mapping is computed in the following way

$$x_1 = \frac{1}{2}\sqrt{2 + y_1^2 - y_2^2 + 2\sqrt{2}y_1} - \frac{1}{2}\sqrt{2 + y_1^2 - y_2^2 - 2\sqrt{2}y_1} \tag{9}$$

$$x_2 = \frac{1}{2}\sqrt{2 - y_1^2 + y_2^2 + 2\sqrt{2}y_2} - \frac{1}{2}\sqrt{2 - y_1^2 + y_2^2 - 2\sqrt{2}y_2}, \tag{10}$$

where y_1 and y_2 are the coordinates of a point in the CC and x_1 and x_2 are the coordinates of a point in the resulting chart. Now, the points need to be scaled and translated to correct positions and scale, which can be achieved by using an affine transformation. General affine transformation between the BS positions is of the form

$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 1 \end{bmatrix}, \tag{11}$$

where parameters a_1 to a_6 represent the unknown affine transformations between the points in the map $\mathbf{x}$ and the points in the chart $\mathbf{y}$. With the BS position estimates, the parameters a_1 to a_6 are the only unknowns, thus we can solve (11) for parameter vector $\mathbf{a}$

$$\mathbf{x} = \begin{bmatrix} y_1 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & y_1 & y_2 & 1 \end{bmatrix} \mathbf{a}. \tag{12}$$

A least squares solution of (12) can be solved with at least three BS positions $\mathbf{x}$ and $\mathbf{y}$, but more BS positions are better if the geometry of the points is not ideal. Each point in the CC is transformed using (11).

IV. Simulations and Results

In this work, we have simulated data in a part of the METIS Madrid grid [13] using an extensive ray-tracing channel modeling [7], proposed by the 3rd Generation Partnership Project (3GPP) in [14]. We consider two different scenarios each with 4 BSs. The first scenario is an open park area and the second scenario is a city block. We will call these scenarios LoS and NLoS scenarios respectively. Both scenarios have roads on the edges and the BSs are located in the intersections of these roads. In the simulations, the BSs are set 1.5m away from the corners of the buildings at the height of 10m and the UE are set in a uniform grid with 2.5m spacing at the height of 1.5m. The map of the Madrid grid and the area of the scenarios are illustrated in Fig. 2. The cost functions for the MDS were chosen to be SM for the LoS scenario and stress for the NLoS scenario. Both cost functions were tested for both scenarios, but these choices gave the best results. In Fig. 3a and Fig. 3d, the UE and BS positions are shown for the LoS and the NLoS scenarios, respectively. The UE positions are colored along the north-south direction. In Fig. 3b and Fig. 3e, the multipoint CCs of four BSs can be seen for both scenarios. The color gradients match the real maps, but the circular shape

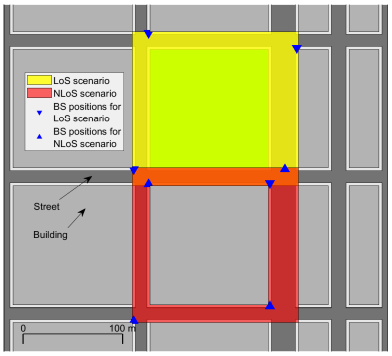
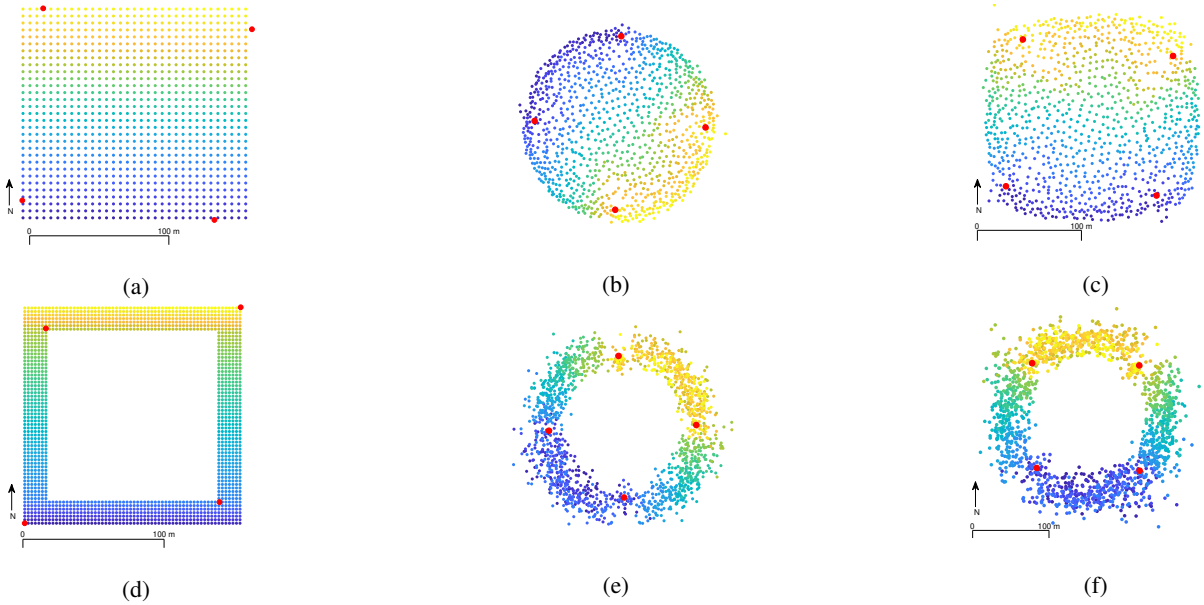


Fig. 2: Madrid grid with LoS and NLoS scenarios colored in yellow and red patches respectively. The green park area is part of the LoS scenario.

of the CC is not optimal since the goal of the chart is to represent the original topology. For the NLoS scenario, there is an empty space inside the points, which means that the CC is able to capture the topology of a scenario with UE around a building. After the extensions proposed and described in Section III, the resulting maps, shown in Fig. 3c and Fig. 3f, have comparable scale, shape and orientation with the real map data. The conformal mapping has removed some of the roundness to better represent the original map, especially in the LoS scenario.

In both scenarios, some of the BSs are located along the edges of the map. These BSs are not surrounded by users, and thus the BS position estimation will be biased away from the edge. This can be seen as the estimates being among the points in CC in Figures 3c and 3f, even though they should be on the outer and inner edges.

The positioning error between the CC and real positions is presented with a histogram. In Fig. 4 the histograms of 2D-errors are presented for the both scenarios. Mean errors in the LoS and NLoS scenarios are 19.7m and 32.0m, respectively. In Figs. 3c and 3f, the points are grouped on the edges of the CC, which likely causes the second peak for LoS scenario and the long tail for NLoS scenario error distributions. A better BS position estimation method, such as path loss models, could improve the accuracy. The positioning accuracy of Long-Term Evolution (LTE) networks, given in [15], is close to our results. In [6], the mean errors for a supervised LoS and NLoS scenario are 6.69m and 10.51m respectively. In addition, the mean errors for a semisupervised LoS and NLoS scenarios are 10.51m and 17.59m. Their unsupervised method gives good shape of the chart, but lacks accuracy information. Considering the unsupervised setting, our results compare fairly well to the supervised settings, and our method works as a proof-of-concept method. The results could possibly be improved by increasing the number of BSs in the area as well.



(a) (b) (c)

(d) (e) (f)

Fig. 3: Figures on top and bottom represent LoS and NLoS scenarios respectively. In each figure red points represent BSs and colored points are UE. On the left are maps of the UE and BSs in real coordinates. In the middle are the multipoint CCs of four UCA antennas in logical scale and on the right is the CCs in absolute scale after tranformations.

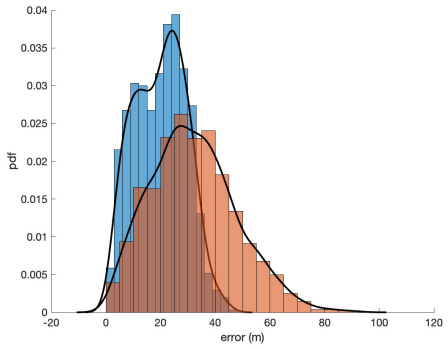


Fig. 4: Histograms of 2D-error in Los and NLoS scenarios colored with blue and red, respectively, and the estimated probability density functions in black.

V. Conclusions

A method for multipoint Channel Charting was presented for UCA antennas for two different scenarios. We proposed a method for weighting the distance matrices for the multipoint CC and cost functions for different visibility scenarios.

An unsupervised method for the extension of CC to absolute positioning is proposed and tested. Our results show that the extension works as a proof of concept, but the accuracy can be improved. Further research is needed for BS position estimation for better positioning accuracy.

References

[1] 5G-PPP, "5G empowering vertical industries," Feb. 2015. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2016/02/BROCHURE_5PPP_BAT2_PL.pdf

[2] M. Koivisto, A. Hakkarainen, M. Costa, P. Kela, K. Leppänen, and M. Valkama, "High-Efficiency Device Positioning and Location-Aware Communications in Dense 5g Networks," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 188–195, 2017.

[3] C. Studer, S. Medjkouh, E. Gönultas, T. Goldstein, and O. Tirkkonen, "Channel charting: Locating users within the radio environment using channel state information," *IEEE Access*, Jul. 2018.

[4] J. Pihlajasalo, S. Ali-Löytty, and M. Valkama, "Unsupervised Channel Charting with Channel State Information Using ULA and UCA Antenna Models in 5G Network," in *Proceedings of XXXV Finnish URSI Convention on Radio Science*. URSI, Oct. 2019.

[5] D. Junquan, S. Medjkouh, N. Malm, O. Tirkkonen, and C. Studer, "Multipoint channel charting for wireless networks," Oct. 2018, pp. 286–290.

[6] E. Lei, O. Castañeda, O. Tirkkonen, T. Goldstein, and C. Studer, "Siamese neural networks for wireless positioning and channel charting," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2019, pp. 200–207.

[7] METIS, "D1.4 Channel models," Feb. 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/METIS_D1.4_v3.pdf

[8] W. Xu, "Multi-antenna non-line-of-sight identification techniques for target localization in mobile ad-hoc networks," 01 2011.

[9] A. J. Izenman, *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[10] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers*, vol. C-18, no. 5, pp. 401–409, May 1969.

[11] H. Nurminen, J. Talvitie, S. Ali-Löytty, P. Müller, E. Lohan, R. Piché, and M. Renfors, "Statistical path loss parameter estimation and positioning using RSS measurements in indoor wireless networks," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nov 2012, pp. 1–9.

[12] C. Fong, "Analytical methods for squaring the disc," 2015.

[13] "Metis *D6.1 Simulation Guidelines*, ICT-317669.METIS/D6.1," Oct. 2013.

[14] "3GPP TR 38.901 v16.1.0 Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)," Dec. 2019.

[15] Spirent, "An Overview of LTE Positioning," Feb. 2012.