Network-side Localization via Semi-Supervised Multi-point Channel Charting

Junquan Deng[1], Olav Tirkkonen[2], Jianzhao Zhang[1], Xianlong Jiao[3] and Christoph Studer[4]

[1]*Sixty-third Research Institute, National University of Defense Technology, China, Emails: {jqdeng, jianzhao63s}@nudt.edu.cn*
[2]*Department of Communications and Networking, Aalto University, Finland, Email: olav.tirkkonen@aalto.fi*
[3]*College of Computer Science, Chongqing University, China, Email: xljiao@cqu.edu.cn*
[4]*Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland, Email: studer@ethz.ch*

***Abstract*—We consider the network-side mobile localization problem in future 5G and beyond wireless networks with distributed multi-antenna base stations (BSs). For this application, we propose a semi-supervised multi-point channel charting (SS-MPCC) framework, which consists of (i) collaborative collection of channel state information (CSI) and other side-information by distributed BSs; (ii) local CSI feature extraction and self-learning of a dissimilarity metric, and (iii) global graph construction and constrained manifold learning. We show that side-information from routine network operations, including timestamps, channel qualities, and a small set of labeled samples, can be exploited to construct a consistent global graph. The graph is then mapped to a 2D channel chart using constrained manifold learning for localization purposes. We evaluate the performance of SS-MPCC in a simulated urban outdoor scenario with realistic user motion. Our results show that SS-MPCC achieves a mean localization error of 5.6 m with only 10% of labeled CSI samples. SS-MPCC does not require accurate synchronization among multiple BSs and is promising for future cellular localization.**

***Index Terms*—MIMO CSI, channel charting, semi-supervised learning, network-side localization**

I. Introduction

Mobile location information acquisition is fundamental in building smart cites and intelligent transportation systems. However, how to realize seamless high-precision localization, especially on the network side, is a challenging problem in GPS-denied environments. Future 5G and beyond radio access networks (RANs) will be densely deployed with massive multiple-input multiple output (mMIMO) Base Stations (BSs) and high-frequency carrier frequencies in order to provide ubiquitous ultra-fast and reliable wireless services. In addition to communication, high-resolution sensing functions are envisioned to be integrated in such RANs, which opens up new opportunities for high-precision positioning and tracking [1]–[4]. Along these lines, 3GPP Rel-16 has specified positioning support for 5G New Radio (NR) [5]. However, current network-side NR positioning techniques are based on triangulation with Angle-of-Arrival (AOA) or trilateration with Time-Difference-of-Arrival (TDOA) measurements at multiple BSs, which require rigorous calibration or accurate synchronization among network elements. Moreover, the performance of such methods degrades in complex multi-path and obstructed environments. Fingerprinting techniques [6]–[8] with machine learning could be applied for cellular positioning in such challenging propagation environments. However, they require a large number of densely-sampled channel measurements, which scales only poorly to large areas and renders automation to dynamic environments challenging. As unlabeled samples are easy to collect, semi-supervised learning (SSL) methods [9] utilizing both labeled and unlabeled data are promising for building a large-scale automatic network-side positioning system. SSL has attracted increasing attention for indoor positioning applications based on Received Signal Strength Indicator fingerprints [10], [11]. However, utilizing mMIMO Channel State Information (CSI) [12], [13] in a networking communication environment for high-precision and seamless positioning and tracking remains a challenging problem.

Channel charting (CC), as put forward in [14], relies on using sporadically collected multi-antenna CSI from unknown locations in order to learn a chart that can be used for applications that do not require absolute position information (e.g., hand-over and user grouping). The principle of CC is to exploit the fact that high-dimensional CSI strongly depends on User Equipment (UE) location, which is low-dimensional. To equip CC with absolute positioning capabilities, a semisupervised autoencoder utilizing a subset CSI of samples with known spatial locations, as well as mobility side information, has been proposed in [15]; this approach does not require dense measurements. More recently, an unified CC architecture based on Sammon's mapping (SM) and Siamese networks has been presented in [16]; this approach can be used for both supervised and semisupervised positioning, and unsupervised relative localization. In [17], CC is performed for radio signals collected in a real environment from transmissions of a 64-antenna base station. A triplet neural network learns the mapping from CSI samples to CC-coordinates, using time stamp side information to classify pairs of samples to come from nearby or remote spatial locations. Unfortunately, in a semi-supervised setting with challenging non-Line-of-Sight (NLOS) channels and large-scale network coverage areas, the methods of [15], [16] are not capable of exploiting unlabeled data samples to improve

This work was supported in part by the National Science Foundation of China under grant 61901497 and 62072064, in part by Research Project of National University of Defense Technology under grant ZK 19-09 and in part by the Academy of Finland under grant 319484.

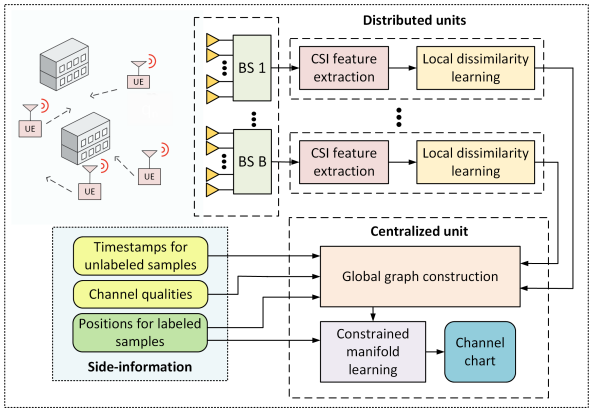


Fig. 1. Graph-based semi-supervised multi-point channel charting framework.

location estimates.

Multipoint channel charting (MPCC) [18], [19] is an unsupervised method that produces trustworthy channel charts by combining all CSI available at multiple BSs and by exploiting redundancy in multi-point CSI to combat the distortion occurring in single-point channel charting results. In this paper, we propose a semi-supervised multi-point channel charting (SS-MPCC) framework for large-scale network-side cellular localization. In SS-MPCC, multiple distributed multi-antenna BSs collaboratively collect wireless data from spatio-temporal mobile locations and learn an aligned channel chart in which historical and real-time mobile locations can be accurately matched and determined. We use timestamp information for the unlabeled samples, and location information of a few labeled samples to increase the smoothness and trustworthiness of the learned channel chart. Specifically, we use graph-based *constrained manifold learning* methods to calibrate the multi-point channel chart, so that points in the chart represent the true geographical mobile location. To demonstrate the effectiveness of SS-MPCC, we perform simulations in an urban outdoor multi-cell mmWave network. Using a simulator that models UE movements in a street grid, we show that a small set of CSI measurements is sufficient to perform accurate positioning within mobile networks.

II. SS-MPCC Framework

We now introduce the framework for semi-supervised multi-point channel charting. Our framework is depicted in Figure 1 and consists of a CSI and side-information collection procedure, CSI feature extraction and local dissimilarity matrix construction procedures at distributed units (DUs), as well as global graph construction and semi-supervised constrained manifold learning procedures performed at a centralized unit (CU).

A. CSI and Side-information Collection

We consider a typical 5G cellular network in which UEs move in the coverage area of multiple mMIMO BSs. During network operation, the UEs send pilot signals, e.g., Sounding Reference Signals (SRSs), to BSs for channel estimation. An estimated mMIMO channel vector at time t on a sub-carrier with frequency f is modelled as

$$\mathbf{h}_{t,f} = \sum\nolimits_{l=1}^{P} \alpha_{t,f}^{(l)} \mathbf{s}(\phi_t^{(l)}) + \mathbf{n}, \tag{1}$$

where P is the number of multi-path components, $\phi_t^{(l)}$ is the impinging direction-of-arrival (DOA) of the lth propagation path, and $\alpha_{t,f}^{(l)}$ is a random complex gain for the lth path which depends on the transmit power, path delays, reflection coefficients of the contributing scatters, and impacts of radio chain components. In addition, $\mathbf{n}$ is the estimation error and the vector $\mathbf{s}(\phi)$ represents the array steering vector (ASV). For an M-antenna uniform linear array with the plane-wave approximation, the ASV is

$$\mathbf{s}(\phi) = \left[1, e^{j\frac{2\pi}{\lambda} s \sin(\phi)}, \ldots, e^{j(M-1)\frac{2\pi}{\lambda} s \sin(\phi)}\right]^{\mathrm{T}}, \tag{2}$$

where λ is the carrier wavelength and s the antenna spacing.

The channel vector $\mathbf{h}_{t,f}$ changes rapidly when the UE moves in length scales of the wavelength λ. This is due to changes in the phases of the complex path gains of the multipath components. In comparison, the DOAs $\{\phi_{t,f}^{(l)}\}_{l=1}^{L}$ and the powers of $\{\alpha_t^{(l)}\}_{l=1}^{L}$ change more slowly and smoothly. To exploit this property, we use the frequency-domain covariance $\mathbf{C} = \mathbb{E}_f\left[\mathbf{h}_{t,f}\mathbf{h}_{t,f}^{\mathrm{H}}\right]$ as the CSI sample for channel charting, as it depends on the multipath DOAs and the powers. In practice, the covariance can be estimated across multiple coherence bandwidths, e.g., different OFDM sub-carriers.

In the considered system, there are two sets of CSI data, labeled and unlabeled samples. The first set consists of the CSI data with corresponding position coordinates which can be gathered in a dedicated site survey, whereas the second set is collected from multiple UEs on move. While collecting the unlabeled CSI samples, the timestamp and UE identifier are also recorded by the BSs, as in [15], [17]. We assume there are B BSs. The L different labeled CSI samples at BS b and the corresponding ground-truth location coordinates are denoted as $\mathcal{L}^{(b)} = \{\mathbf{C}_1^{(b)}, \ldots, \mathbf{C}_L^{(b)}\}$ and $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_L]$. The in total U unlabeled samples and corresponding timestamps collected at BS b then can be denoted as $\mathcal{U}^{(b)} = \{\mathbf{C}_{L+1}^{(b)}, \ldots, \mathbf{C}_{L+U}^{(b)}\}$ and $[t_1, \ldots, t_U]$. The corresponding unknown locations for the unlabeled samples are denoted by $\hat{\mathbf{P}} = [\mathbf{p}_{L+1}, \ldots, \mathbf{p}_{L+U}]$. It is straightforward to generalize the model to a situation with unlabeled samples gathered from multiple mobile UEs. Furthermore, the channel qualities, e.g., Reference Signal Receiving Powers (RSRPs), for CSI samples are also collected for data fusion purposes, as we discuss in Section III.

B. CSI Feature Extraction

We use the power angular profile (PAP) of multipath components (MPCs) as the CSI feature as in [18]. We use a Multiple Signal Classification (MUSIC) algorithm [20] to extract the PAP feature from a covariance CSI sample $\mathbf{C}$. The powers of multiple components are proportional to the eigenvalues of the covariance matrix. Noticing this, we first compute the eigen-decomposition of $\mathbf{C}$ and estimate the number $\hat{P}$ of prominent MPCs based on the descending

distribution of eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_M\}$. In practice, $\hat{P}$ must be smaller then the number M of BS antennas, and can be estimated as the minimum $\hat{P}$ satisfying $\frac{\lambda_1+\ldots+\lambda_{\hat{P}}}{\lambda_1+\ldots+\lambda_M} > \rho$, where ρ is a target power ratio, which is set as 0.8 in our experiments in Section IV. Then the MUSIC algorithm is applied to find $\hat{P}$ DOAs $\{\phi_1, \ldots, \phi_{\hat{P}}\}$ in its pseudo-spectrum. Finally, the PAP feature is represented as DOA−power pairs stacked in a vector as $\mathbf{f} = [\lambda_1, \ldots, \lambda_{\hat{P}}, \phi_1, \ldots, \phi_{\hat{P}}]$.

C. Local CSI Dissimilarity Learning

Given a set of CSI features $\{\mathbf{f}_n^{(b)}\}_{n=1}^{L+U}$ extracted from the labeled CSI samples $\mathcal{L}^{(b)}$ and unlabeled $\mathcal{U}^{(b)}$ at a local BS b, we need a method to measure their dissimilarity. As shown in Figure 2, each MPC corresponds to a virtual transmission (Tx) point [21] in the radio propagation environment. Spatially close mobile locations will produce similar virtual Tx points, as they share common scatterers and have similar MPCs. The coordinates of these virtual Tx points can be deduced from the PAP features. For a general multipath PAP feature $\mathbf{f} = [\lambda_1, \ldots, \lambda_{\hat{P}}, \phi_1, \ldots, \phi_{\hat{P}}]$, its $\hat{P}$ coordinates with respect to the local BS is given by

$$\mathcal{F}(\mathbf{f}) = [\mathbf{x}_1, \ldots, \mathbf{x}_{\hat{P}}], \tag{3}$$

where $\mathbf{x}_i = \rho(\lambda_i)\,[\lambda_i^{-\frac{1}{2}} \cos\phi_i, \lambda_i^{-\frac{1}{2}} \sin\phi_i]^{\mathrm{T}}$ is the ith virtual Tx point, and $\rho(\cdot)$ is a scaling function compensating the scattering loss. We then transform the PAP features $\{\mathbf{f}_n^{(b)}\}_{n=1}^{L+U}$ to a *point cloud*, which consists of all the virtual Tx points with respect to MPCs observed by the BS during CSI sample collection. As depicted in Figure 2, this point cloud exhibits clusters, which reflects the distributions of mobile UEs and the radio scatterers in the environments. The dissimilarity between two CSI samples depends on whether their PAP features have virtual points that belong to the same cluster. If two samples share virtual points in a same cluster, one should estimate their distance based on these similar virtual points. For this, take two features $\mathbf{f}_m^{(b)}$ and $\mathbf{f}_n^{(b)}$ from the feature set, assume that their virtual points are $[\mathbf{x}_{1,m}, \ldots, \mathbf{x}_{\hat{P},m}]$ and $[\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{\hat{Q},n}]$ of the form (3). Assume that there are a total of C clusters in the virtual Tx point cloud. After clustering, a virtual Tx point $\mathbf{x}$ would have a cluster label $l(\mathbf{x}) \in \{1, \ldots, C\}$. Denote the set of clusters by $\mathcal{C}_{m,n}$ where both $\mathbf{f}_m^{(b)}$ and $\mathbf{f}_n^{(b)}$ have at least one virtual point in each of them. If $\mathcal{C}_{m,n}$ is not empty, then we search a pair of $\{\mathbf{x}_{i,m}, \mathbf{x}_{j,n}\}$ from $[\mathbf{x}_{1,m}, \ldots, \mathbf{x}_{\hat{P},m}]$ and $[\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{\hat{Q},n}]$, such that $\max_{i,j}\{|\|\mathbf{x}_{i,m}\|_2 - \|\mathbf{x}_{j,n}\|_2|\}$ is minimal and $l(\mathbf{x}_{i,m}) = l(\mathbf{x}_{j,n})$. Then, the distance $\|\mathbf{x}_{i,m} - \mathbf{x}_{j,n}\|_2$ is used as the dissimilarity metric. If $\mathcal{C}_{m,n}$ is empty, then the distance $\|\mathbf{x}_{1,m} - \mathbf{x}_{1,n}\|_2$ between the points $\mathbf{x}_{1,m}$ and $\mathbf{x}_{1,n}$, which correspond to the strongest MPCs, is used as the dissimilarity metric.

Using above CSI feature extraction method and the self-learned dissimilarity metric, the local BS can construct a matrix $\mathbf{D}^{(b)} \in \mathbb{R}^{(L+U)\times(L+U)}$ for the CSI samples $\mathcal{L}^{(b)}$ and $\mathcal{U}^{(b)}$, whose element $D_{i,j}^{(b)}$ captures the dissimilarity between $\mathbf{C}_i^{(b)}$ and $\mathbf{C}_j^{(b)}$, for $i, j = 1, \ldots, L+U$.

The CSI feature extraction and dissimilarity learning procedures are performed at the DUs which reside at the network edge close to BSs. Then, the local dissimilarity matrix of all CSI samples and the corresponding timestamp and channel quantities are reported to a logical CU for multi-point channel charting. We thus avoid exchange of the high-dimensional channel covariances between DUs and CU. The CU then fuses the data collected from distributed BSs to form a channel chart via global graph construction and semi-supervised constrained manifold learning, as discussed in Section III.

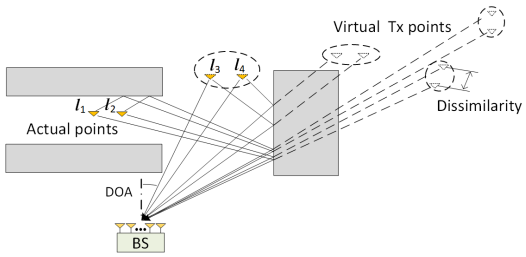


Fig. 2. Dissimilarity metric deduced from a virtual Tx point cloud.

III. Graph-based SS-MPCC

In [18], manifold learning algorithms, including Sammon's Mapping (SM), t-Distributed Stochastic Neighbor Embedding (t-SNE) and Laplacian Eigenmaps (LE), were considered for unsupervised MPCC, which can learn the relative position between unlabeled CSI samples. Here, we choose LE the baseline dimensional reduction algorithm, and adapt it for semi-supervised learning. Conventional SM and t-SNE can also be modified to utilize side information. For example, a parametric version of SM using Siamese networks has been considered in [16] for semi-supervised single-point channel charting. However, performing multi-point channel charting with such methods is an open research problem.

LE consists of neighborhood graph construction, weight matrix computation, and low-dimensional mapping procedures. We will show that incorporation of side information and merging of local dissimilarity matrices can be realized in the first two procedures. For application in SS-MPCC, we first construct a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the sets of vertices $\mathcal{V}$ and edges $\mathcal{E}$ represent the CSI samples collected from the $L+U$ locations, and the corresponding neighborhood relations.

The objective of SS-MPCC is to map the $L+U$ high-dimensional CSI samples into a low-dimensional space with a representational channel chart $\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_{L+U}]$, so that $\mathbf{z}_n$ should be close to the true mobile location $\mathbf{p}_n$, for all $n = 1, \ldots, L+U$. It can be addressed using *constrained manifold learning* methods [22], [23], which is prevalent in machine learning for exploring the underlying low-dimensional structure of the whole dataset with both labeled and unlabeled data. As the first L locations are known, we set $\sum_{n=1}^{L} \|\mathbf{z}_n - \mathbf{p}_n\|_2^2$ as a cost function which should be minimized. Furthermore, to preserve the neighborhood relationships of nodes in $\mathcal{V}$ deduced from CSI dissimilarities, we set $\sum_{m,n} w_{m,n}\|\mathbf{z}_m - \mathbf{z}_n\|_2^2$ as another cost function, where

$w_{m,n}$ is the edge weight for nodes m and n. The overall cost function is therefore given by

$$f(\mathbf{Z}) = \sum_{n=1}^{L} \|\mathbf{z}_n - \mathbf{p}_n\|_2^2 + \gamma \sum_{m=1,n=1}^{L+U} w_{m,n} \|\mathbf{z}_m - \mathbf{z}_n\|_2^2. \quad (4)$$

Here, $\gamma \geq 0$ is a tradeoff coefficient between two objectives. The first term in (4) represents the alignment error for labeled samples, while the second term preserves local smoothness of the channel chart. For $\gamma = 0$, only the anchor points will be used; for $\gamma \to \infty$, only CSI is used and anchors are ignored.

A. Global Graph Construction

To construct the global graph $\mathcal{G}$ based on the B local dissimilarity matrices $\{\mathbf{D}^{(b)}\}_{b=1,...,B}$, and side information $[t_1, ..., t_U]$ and $\mathbf{P} = [\mathbf{p}_1, ..., \mathbf{p}_L]$, we proceed as:

Step-1: Merge the local dissimilarity matrices to form a globally consistent one as in [18]. In practice, the signal from a sampled location $\mathbf{p}_n$ may not be received by a BS b. In this case, we set $\mathbf{f}_n^{(b)} = \varnothing$, and $D_{n,m}^{(b)}$, $D_{m,n}^{(b)}$ for all $m \neq n$ to be a large value, and set the channel quality $\gamma_n^{(b)}$ be zero. The cell-specific dissimilarities are then fused into a global multi-cell dissimilarity matrix $\mathbf{D}$ with elements

$$D_{m,n} = \left(\sum_{b=1}^{B} v_b \right)^{-1} \sum_{b=1}^{B} v_b D_{m,m}^{b}, \quad (5)$$

where v_b is a reliability weight. We consider using $v_b = [\min(\gamma_n^{(b)}, \gamma_m^{(b)})]^{\zeta}$, where $\{\gamma_j^{(b)}\}$ are channel qualities between locations $\{\mathbf{p}_j\}$ and BS b, estimated from RSRP or other measurements. In addition, The parameter ζ is a weighting exponent. MPCC is rather robust against changing the weighting exponents, and we set $\zeta = 2$ in this paper.

Step-2: Find k_e nearest neighbors based on the global multi-cell dissimilarity matrix $\mathbf{D}$ for each node m in $\mathcal{G}$. If node n is in the set of k_e nearest neighbors of node m, nodes m and n are connected in the graph.

Step-3: If the timestamps t_m, t_n of two nodes $m, n \in \{L+1, \ldots, L+U\}$ with unlabeled CSI samples satisfy $|t_m - t_n| < T_{th}$, connect nodes m and n in $\mathcal{G}$. Here T_{th} is the threshold for timestamps which is related to the movement speeds of UEs. It will be discussed in Section IV.

Step-4: Based on the neighborhood graph constructed above, compute the weight matrix $\mathbf{W}$ elements $\{w_{m,n}\}_{m,n=1}^{L+U}$. Denote $\mathcal{A} = \{1, \ldots, L\}$ and $\mathcal{B} = \{L+1, \ldots, L+U\}$. If the nodes m and n are connected, then the weight $w_{m,n}$ is

$$w_{m,n} = \begin{cases} \alpha e^{\frac{-D_{m,n}}{\theta_s^2}} + (1-\alpha) e^{\frac{-\|\mathbf{p}_m - \mathbf{p}_n\|_2}{\theta_d^2}}, & \text{if } n, m \in \mathcal{A}, \\ \alpha e^{\frac{-D_{m,n}}{\theta_s^2}} + (1-\alpha) e^{\frac{-|t_m - t_n|}{\theta_t^2}}, & \text{if } n, m \in \mathcal{B}, \\ e^{\frac{-D_{m,n}}{\theta_s^2}}, & \text{otherwise.} \end{cases} \quad (6)$$

Here, α is a tradeoff coefficient for mixed edge weights, and θ_s, θ_t and θ_d are the kernel parameters for CSI dissimilarities, timestamps and location coordinates respectively. The parameters α, θ_s, θ_t, and θ_d will be investigated in Section IV. Finally, if the nodes m and n are not connected, set $w_{m,n} = 0$. When k_e is much smaller than the number of nodes in the graph, the resulting graph would typically have a sparse weighted adjacency matrix $\mathbf{W}$.

It is straightforward to modify (6) to incorporate only position or timestamp information. For example, if timestamps are absent, then we just need to delete the second line in (6).

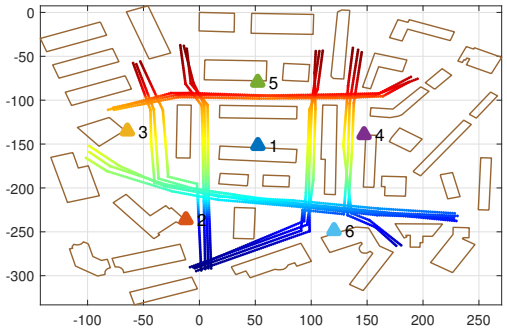


Fig. 3. Distributions of 6 BSs marked by numbers, the traces of UEs during CSI sample collection represented by dots. The mobile locations are color-mapped for visualization purposes.

B. Constrained Manifold Learning

Given the graph weight matrix $\mathbf{W}$ and the locations $\mathbf{P} = [\mathbf{p}_1, ..., \mathbf{p}_L]$ for the labeled samples, we can find a multi-point channel chart $\mathbf{Z}$, by minimizing the cost function in (4). Denote $\mathbf{Q} = [\mathbf{P}, \hat{\mathbf{P}}] \in \mathbb{R}^{(L+U)\times 2}$, and set the unknown locations $\hat{\mathbf{P}}$ to zeros. Then we rewrite $f(\mathbf{Z})$ as

$$\begin{aligned} f(\mathbf{Z}) &= \sum\nolimits_{n=1}^{L} \|\mathbf{z}_n - \mathbf{p}_n\|_2^2 + \gamma \mathrm{Tr}\left(\mathbf{Z}^{\mathrm{T}} \mathbf{\Lambda} \mathbf{Z} - \mathbf{Z}^{\mathrm{T}} \mathbf{W} \mathbf{Z}\right) \\ &= \mathrm{Tr}\left[(\mathbf{Z} - \mathbf{Q})^{\mathrm{T}} \mathbf{J} (\mathbf{Z} - \mathbf{Q}) + \gamma \mathbf{Z}^{\mathrm{T}} \mathbf{L} \mathbf{Z}\right], \end{aligned}$$

where $\mathbf{J} = \mathrm{diag}\left\{\underbrace{1, \ldots, 1}_{L}, \underbrace{0, , \ldots, 0}_{U}\right\}$, $\mathbf{\Lambda}$ is the degree matrix of $\mathcal{G}$ with the nth diagonal element $\Lambda_n = \sum_{m=1}^{L+U} w_{n,m}$, and $\mathbf{L} = \mathbf{\Lambda} - \mathbf{W}$ is the graph Laplacian of $\mathcal{G}$.

As in [22], by setting the derivative of objective function to be zero, the aligned channel chart can be estimated as

$$\mathbf{Z} = (\gamma \mathbf{L} + \mathbf{J})^{-1} \mathbf{J}\ \mathbf{Q}. \quad (7)$$

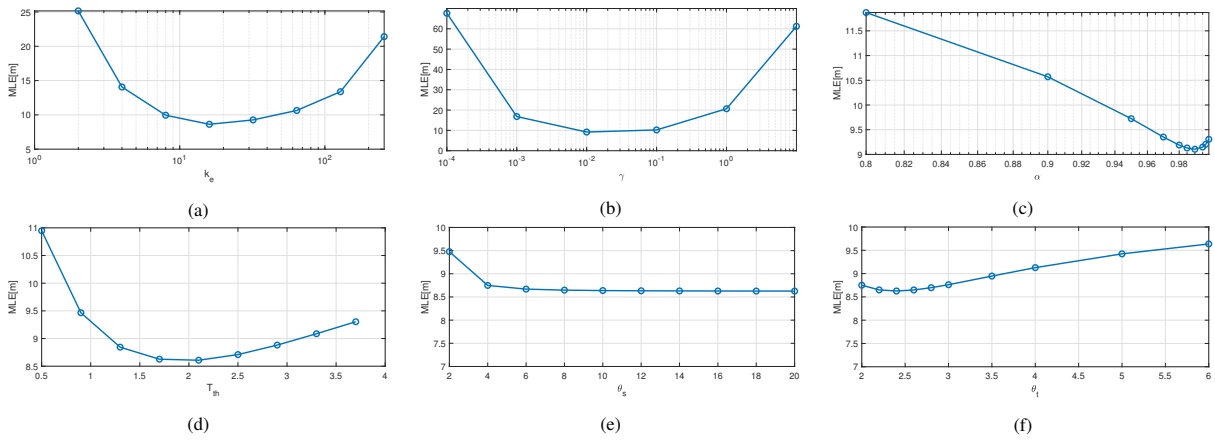
In (7), larger values of γ will increase the smoothness of the channel chart $\mathbf{Z}$, while γ should be close to 0 to map labeled samples to its ground-true positions. A suitable value of γ can be found experimentally as will be shown in IV.

IV. Simulation Results

We now demonstrate the efficacy of the proposed graph-based SS-MPCC for CSI-based localization with timestamps and a subset of CSI samples with marked positions.

A. Simulated Scenario and Evaluation Metrics

We consider an urban outdoor multi-cell mmWave network scenario as depicted in Figure 3. BSs are below rooftop, and signals will be reflected or blocked by the walls. A ray-tracing channel model is used to generate the multi-path channels. There are $B = 6$ mMIMO BSs equipped with ULAs, each has $M = 16$ elements. The antenna arrays of the BSs are oriented perpendicular to the building surfaces where the BSs



(a) (b) (c) (d) (e) (f)

Fig. 4. Effects of parameters on localization performance of graph-based SS-MPCC with both timestamp and location information; The default parameters are $k_e = 16$, $\gamma = L/3000$, $\alpha = 0.99$. $T_{th} = 2.4$, $\theta_s = 5$, $\theta_d = 5$, $\theta_t = 2$ and $L = 100$.

are mounted. UEs transmit reference signals with a fixed power of 23 dBm. The carrier frequency is 28 GHz, system bandwidth is 256 MHz with 128 OFDM subcarriers.

In the simulated scenario, UEs move along the roads with a average speed of 5 meters per second. We collect $U = 3000$ CSI samples from UE traces and the average distance between neighboring sampled locations is approximately 2 m, which is much larger the carrier's wavelength. The covariance matrix is estimated over 32 realizations of the channel vector $\mathbf{h}$ over multiple adjacent subcarriers. The channel quality is estimated as $\mathbb{E}\{\|\mathbf{h}\|_2^2\}$. For semi-supervised learning, additional L samples are generated randomly on the roads. To measure the performance of CSI-based localization methods, we use the following metrics:

1) Mean localization error: For semisupervised localization, a natural way for measuring performance is the mean localization error (MLE) for unlabeled samples:

$$MLE = \frac{1}{U}\sum_{n=1}^{U} \|\mathbf{z}_{L+n} - \mathbf{p}_{L+n}\|_2. \tag{8}$$

2) Kruskal's stress: In addition to the MLE, we use Kruskal's stress (KS) to measure how well the learned channel chart represents the true locations globally:

$$KS = \sqrt{\frac{\sum_{n,m}(\delta_{n,m} - \beta d_{n,m})^2}{\sum_{n,m}\delta_{n,m}^2}}. \tag{9}$$

Here, $\delta_{n,m} = \|\mathbf{p}_n - \mathbf{p}_m\|$, $d_{n,m} = \|\mathbf{z}_n - \mathbf{z}_m\|$, and $\beta = \sum_{n,m}\delta_{n,m}d_{n,m}/\sum_{n,m}\delta_{n,m}^2$. KS is in the range $[0, 1]$ and smaller values indicate better preservation of global geometry.

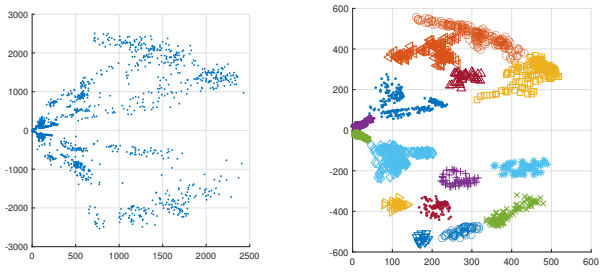
3) Trustworthiness and continuity: We use two standard metrics that characterize neighborhood preserving performance: *trustworthiness* (TW) and *continuity* (CT) [14]–[16]. The TW measures whether the mapping of high-dimensional samples to the low-dimensional space introduces false neighbors; the CT measures whether similar samples in high-dimensional space remain similar in low-dimensional space. TW and CT have values in $[0, 1]$ and larger values imply better preservation of neighborhood relationship. When evaluating TW and CT, a number K of neighbors should be specified.

B. Virtual Tx Point Cloud

Figure 5 depicts the virtual Tx point cloud extracted by BS 1. These points are related to 536 mobile locations inside the coverage area of BS 1, and exhibit clear clusters when compressed to a confined area via a nonlinear scaling function.

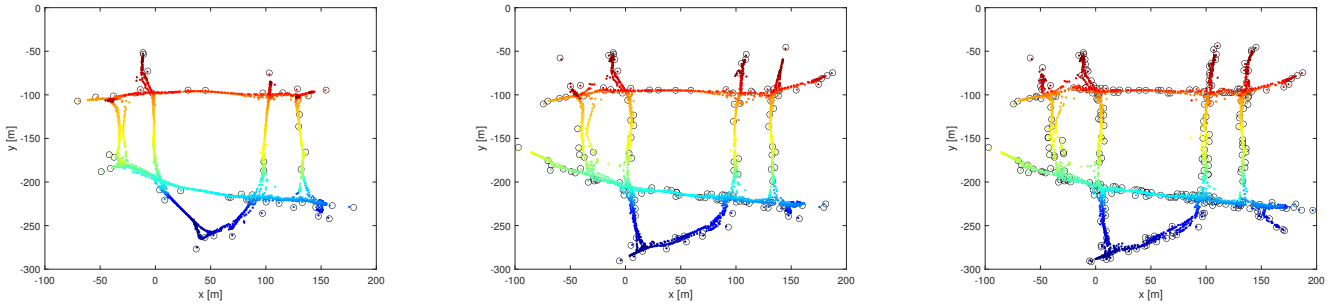
C. Investigations of Algorithm Parameters

There are seven parameters: the alignment-smoothness tradeoff coefficient γ, mixed edge weight tradeoff coefficient α, number of neighbors k_e, timestamp threshold T_{th}, and the heat kernel parameters θ_s,θ_d and θ_t. We first carried out experiments to study the impact of parameters k_e on the localization performance, as it largely determines the structure of the graph. As shown in Figure 4a, when other parameters are fixed, there is an optimal k_e. A very small k_e cannot well characterize the neighborhood relations among nodes. Here we set $k_e = 16$, which leads to the lowest MLE as in Figure 4a. T_{th} plays a similar role as k_e, we found that in our simulation scenario, $T_{th} = 2.4s$ is a good option as depicted in Figure 4d. Small kernel parameters will de-emphasis neighbor connections with larger dissimilarities, while large kernel parameters cannot find the fine-grained local structures. However, when the neighborhood graph is set, their impacts on MLE performance is limited, we set $\theta_s = \theta_d = 5$ and θ_t = 2 in the following simulations. The parameter γ has a profound impact on the learned channel chart as can be seen from Figure 4b. Values of γ larger than



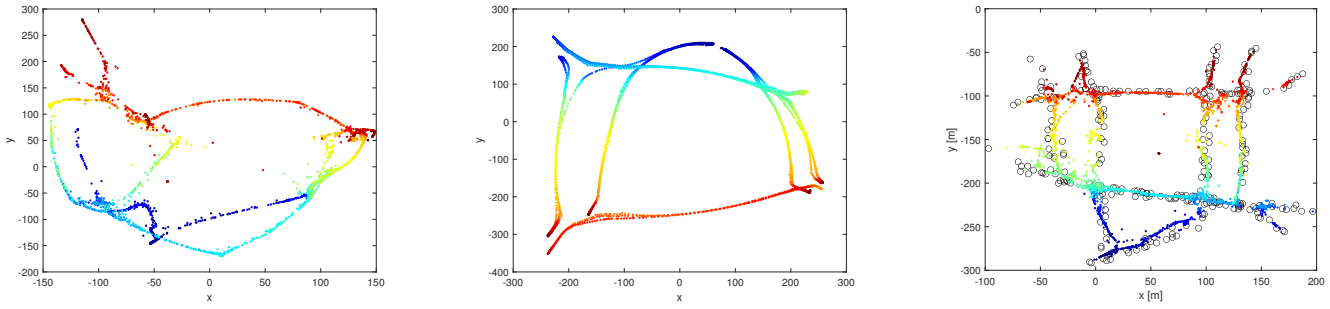
(a) (b)

Fig. 5. Point clouds of virtual Tx points extracted by BS 1, a) without scattering loss compensation; b) with scattering loss compensation and DBSCAN clustering.



(a) $L = 50$. (b) $L = 150$. (c) $L = 300$.

Fig. 6. Visualization of the localization performance with different numbers of labeled CSI samples, number of unlabeled samples is U = 3000. The MLEs are 12.6, 6.8 and 5.6 m for L =50, 150 and 300 labeled samples. The positions for labeled samples are marked by circles.



(a) unsupervised LE. (b) LE with timestamps. (c) LE with 300 labeled samples and no timestamp.

Fig. 7. Visualization of the channel charts learned by (a) unsupervised LE, (b) LE with timestamps and (c) LE with L=300 labeled samples and no timestamp. The parameters used for b) are $k_e = 16$, $\gamma = 0.1$, $\alpha = 0.99$. $T_{th} = 2.4$, $\theta_s = 5$, $\theta_t = 2$. The parameters used for c) are $k_e = 16$, $\gamma = 0.1$, $\alpha = 0.01$. $\theta_s = 5$, $\theta_d = 10$,

one leads to a shrunken chart as it helps to minimize the Graph Laplacian. It turns out $\gamma = \frac{L}{U}$ is a good choice in our simulations; for the mixed edge weight tradeoff coefficient α, we found that $1 - L/U$ worked well.

D. Performance Comparison

We now investigate the performance of the proposed SS-MPCC method for network-side localization. We compare SS-MPCC to unsupervised MPCC [18], and kNN method, which is widely used in CSI-based fingerprinting [6]. We will show that SS-MPCC is applicable to a variety of scenarios. It can work with different types and amounts of side-information.

First, as demonstrated by Figure 6, the proposed graph-based SS-MPCC method is capable of unveiling mobile devices' locations in the considered grid-like urban environment. With only $\frac{50}{3000} = 1.67\%$ of labeled CSI samples, the network topology is revealed, and the UEs can be located with a MLE as 12.6 m. When this ratio increases to 10 %, a MLE of 5.6 m is achieved. Note that such a positioning accuracy does not assume tight synchronization and/or high-precision calibration among multiple BSs, what they need to know is that a CSI sample is transmitted from a UE during the same time interval. This result can be understood by the fact that the timestamps and multi-point data fusion in SS-MPCC help in preserving the neighborhood relationships of the CSI data on each user trace and/or across multiple user traces, which is beneficial in learning the manifold of unlabeled CSI data. As a reward, a small set of labeled data is adequate for network-side positioning and tracking.

To gain more understanding on the efficacy and working mechanisms of SS-MPCC, we investigate its performance with only timestamps, and with only marked positions. For the sake of completeness, results of conventional unsupervised MPCC are also given, as illustrated in Figure 7 and summarized in Table I. The channel chart in Figure 7a generated by unsupervised MPCC is capable of preserving most of the local manifold structures, but with a twisted global structure, and interruptions at some points. By utilizing the timestamps, the chart in Figure 7b is smoother and the traces of UEs are revealed, with $TW = 0.9951$ and $CT = 0.9943$ for $K = 50$. With only partial position information, by modifying (6) and using proper parameters, an aligned channel chart is shown as in Figure 7c. Compared to Figure 7b, the TW and CT scores are smaller, but absolute positioning with a MLE = 16.7 m is enabled with 10% labeled CSI samples.

Finally, we compare SS-MPCC to kNN. kNN is only

TABLE I
PERFORMANCE COMPARISON

		kNN	MPCC	SS-MPCC		
Timestamps		×	×	✓	×	✓
L		300	×	×	300	300
MLE [m]		18.8	N/A	N/A	16.7	**5.6**
KS		.2547	.3593	.1938	.2540	**.0768**
TW	$K = 50$	.9699	.9629	**.9951**	.9819	.9950
	$K = 100$	.9689	.9529	.9898	.9814	**.9956**
CT	$K = 50$	.9641	.9711	.9943	.9735	**.9968**
	$K = 100$	.9591	.9645	.9912	.9719	**.9973**

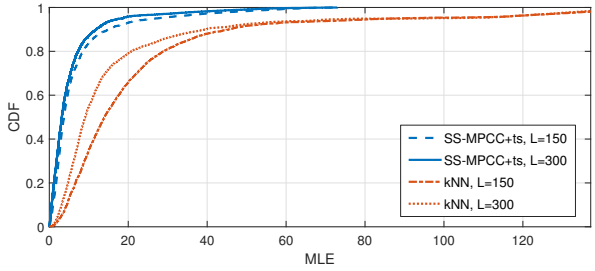


Fig. 8. Cumulative Distribution Function (CDF) of localization errors for different number (L) of labeled CSI samples.

applicable when there are labeled samples, it estimates the position of an unlabeled sample by computing the barycenter of k most similar labeled ones, and neglects the timestamp information. Here we set $k = 5$ for kNN, which yields optimal MLE performance. As shown in Table I, SS-MPCC performs much better than kNN for all metrics, especially when timestamps of unlabeled samples are exploited. From the CDFs of MLEs in Figure 8, we see that a portion of locations have large localization errors for kNN, which correspond to the positions without nearby labeled anchor points, especially at the network edges. Furthermore, as is shown in Figure 8, when the number of labeled samples drops from 300 to 150 by half, the performance of MPCC is only slightly affected, with mean MLE increased from 5.6 m to 6.8 m, while the MLE of kNN is increased from 18.8 m to 25.1 m.

V. Conclusions

We have shown that side-information, such as timestamps and a set of labeled samples, can be incorporated into a semi-supervised multi-point channel charting framework to support large-scale network-side localization and tracking in a distributed manner. Moreover, we have demonstrated that the multi-BS CSI feature fusion and incorporation of timestamps to increase manifold smoothness can be implemented in global graph construction based on graph Laplacian. Finally, through experiments in a simulated urban outdoor scenario, we have shown that SS-MPCC is able to perform large-scale network-side positioning for scenarios with realistic UE motion, even with a very small portion of labeled data. The proposed framework is scalable and automatic in the sense that it could be implemented for varying number of BSs, with spatially sparse labeled samples, and does not require accurate network synchronization. There are many opportunities for extensions of this work, including out-of-sample mapping to locate new samples on the channel chart for real-time positioning applications, and filtering algorithms for multi-target tracking.

References

[1] R. Mendrzik, F. Meyer, G. Bauch, and M. Win, "Localization, mapping, and synchronization in 5G millimeter wave massive MIMO systems," in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1–5.

[2] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3607–3644, Fourthquarter 2018.

[3] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 80–86, Dec 2017.

[4] C. Zhang, Y. Ueng, C. Studer, and A. Burg, "Artificial intelligence for 5G and beyond 5G: Implementations, algorithms, and optimizations," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 10, no. 2, pp. 149–163, June 2020.

[5] R. Keating, M. Säily, J. Hulkkonen, and J. Karjalainen, "Overview of positioning in 5G new radio," in *16th International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2019, pp. 320–324.

[6] Y. Chapre, A. Ignjatovic, A. Seneviratne, and S. Jha, "CSI-MIMO: Indoor Wi-Fi fingerprinting system," in *39th Annual IEEE Conference on Local Computer Networks*, 2014, pp. 202–209.

[7] X. Sun, X. Gao, G. Y. Li, and W. Han, "Single-site localization based on a new type of fingerprint for massive MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6134–6145, July 2018.

[8] J. Meng, A. Sharma, T. X. Tran, B. Balasubramanian, G. Jung, M. Hiltunen, and Y. Charlie Hu, "A study of network-side 5G user localization using angle-based fingerprints," in *IEEE International Symposium on Local and Metropolitan Area Networks*, July 2020, pp. 1–6.

[9] X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005.

[10] J. J. Pan, S. J. Pan, J. Yin, L. M. Ni, and Q. Yang, "Tracking mobile users in wireless networks via semi-supervised colocalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 587–600, March 2012.

[11] T. Pulkkinen, T. Roos, and P. Myllymäki, "Semi-supervised learning for WLAN positioning," in *Artificial Neural Networks and Machine Learning*. Springer Berlin Heidelberg, 2011, pp. 355–362.

[12] S. D. Bast, A. P. Guevara, and S. Pollin, "CSI-based positioning in massive MIMO systems using convolutional neural networks," in *IEEE 91st Vehicular Technology Conference*, May 2020, pp. 1–5.

[13] M. Gauger, M. Arnold, and S. ten Brink, "Massive MIMO channel measurements and achievable rates in a residential area," in *24th International ITG Workshop on Smart Antennas*, Feb 2020, pp. 1–6.

[14] 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, pp. 47 682–47 698, 2018.

[15] P. Huang, O. Castañeda, E. Gönültaş, S. Medjkouh, O. Tirkkonen, T. Goldstein, and C. Studer, "Improving channel charting with representation-constrained autoencoders," in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1–5.

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

[17] P. Ferrand, A. Decurninge, L. G. Ordoñez, and M. Guillaud, "Triplet-based wireless channel charting," 2020.

[18] J. Deng, S. Medjkouh, N. Malm, O. Tirkkonen, and C. Studer, "Multipoint channel charting for wireless networks," in *52nd Asilomar Conference on Signals, Systems, and Computers*, Oct 2018, pp. 286–290.

[19] C. Geng, H. Huang, and J. Langerman, "Multipoint channel charting with multiple-input multiple-output convolutional autoencoder," in *IEEE/ION Position, Location and Navigation Symposium (PLANS)*, April 2020, pp. 1022–1028.

[20] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, March 1986.

[21] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5613–5629, Dec 2019.

[22] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *J. Mach. Learn. Res.*, vol. 7, p. 2399–2434, Dec. 2006.

[23] M. Wang, W. Fu, S. Hao, D. Tao, and X. Wu, "Scalable semi-supervised learning by efficient anchor graph regularization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 7, pp. 1864–1877, July 2016.