Kazemi, Parham; Al-Tous, Hanan; Studer, Christoph; Tirkkonen, Olav

SNR Prediction in Cellular Systems based on Channel Charting

Published in:
2020 8th International Conference on Communications and Networking, ComNet2020 - Proceedings

DOI:
10.1109/ComNet47917.2020.9306087

Published: 27/10/2020

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Kazemi, P., Al-Tous, H., Studer, C., & Tirkkonen, O. (2020). SNR Prediction in Cellular Systems based on Channel Charting. In *2020 8th International Conference on Communications and Networking, ComNet2020 - Proceedings* Article 9306087 IEEE. https://doi.org/10.1109/ComNet47917.2020.9306087

SNR Prediction in Cellular Systems based on Channel Charting

Parham Kazemi*, Hanan Al-Tous*, Christoph Studer†, Olav Tirkkonen*
* Department of Communications and Networking, Aalto University, Espoo, Finland
† School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA
Email: {parham.kazemi, hanan.al-tous, olav.tirkkonen}@aalto.fi, studer@cornell.edu

***Abstract*—We consider a machine learning algorithm to predict the Signal-to-Noise-Ratio (SNR) of a user transmission at a neighboring base station in a massive MIMO (mMIMO) cellular system. This information is needed for Handover (HO) decisions for mobile users. For SNR prediction, only uplink channel characteristics of users, measured in a serving cell, are used. Measuring the signal quality from the downlink signals of neighboring Base Stations (BSs) at the User Equipment (UE) becomes increasingly problematic in forthcoming mMIMO Millimeter-Wave (mmWave) 5G cellular systems, due to the high degree of directivity required from transmissions, and vulnerability of mmWave signals to blocking. Channel Charting (CC) is a machine learning technique for creating a radio map based on radio measurements only, which can be used for radio-resource-management problems. A CC is a two-dimensional representation of the space of received radio signals. Here, we learn an annotation of the CC in terms of neighboring BS signal qualities. Such an annotated CC can be used by a BS serving a UE to first localize the UE in the CC, and then to predict the signal quality from neighboring BSs. Each BS first constructs a CC from a number of samples, determining similarity of radio signals transmitted from different locations in the network based on covariance matrices. Then, the BS learns a continuous function for predicting the vector of neighboring BS SNRs as a function of a 2D coordinate in the chart. The considered algorithm provides information for handover decisions without UE assistance. UE-power consuming neighbor measurements are not needed, and the protocol overhead for HO is reduced.**

***Index Terms*—massive MIMO, mmWave, channel charting, handover, SNR prediction.**

I. Introduction

Beyond Fifth-Generation (B5G) networks are expected to offer important connectivity advantages: energy savings, higher system capacity, reduced latency, and higher data rates. Massive Multiple-Input-Multiple-Output (mMIMO), Millimeter-Wave (mmWave) and Ultra Cell Densification (UCD) are key technologies for successful development of B5G networks, needed to meet unprecedented speeds, near-wireline latencies and ubiquitous connectivity with diverse Quality-of-Service (QoS) requirements [1]–[3].

Large amounts of spectrum are available at mmWave frequencies above 30 GHz, which can be utilized to support data rates of multiple Gb/s. A key challenge in developing systems in mmWave bands is the potential for rapid channel dynamics. Radio propagation in mmWave bands suffers from high path loss, reduced scattering which in turns reduces the available diversity, and increased effect of blockage. Diffraction is reduced in mmWave bands, and accordingly small changes in the position or orientation of the handset relative to objects in the environment can cause large variations in the channel quality [3]. MmWave Base Stations (BSs) have to use directional antennas to increase the Signal-to-Noise Ratio (SNR), even to reach a radius up to 200 meters. Hence, hundreds of BSs will be needed to cover large spaces, and User Equipment (UE) will have to frequently perform Handover (HO) procedures. Successful deployment of mmWave cellular systems thus requires effective HO procedures [1].

Modeling, measuring and predicting the radio channel characteristics of mmWave systems are currently active research areas [4]. Using physical location information is a promising approach to decrease HO signaling overhead in mMIMO dense networks [5], [6], subject to taking care of the user location privacy.

B5G systems encounter fundamental challenges from having to cope with resource-constrained devices to managing the underlying heterogeneous networking and computing infrastructures. Network management and resource allocation of B5G using traditional algorithms becomes complicated, inflexible, and expensive. Motivated by the burgeoning progress of Artificial Intelligence (AI) and its breakthroughs in a variety of domains, the communication research community is currently seeking solutions from Machine Learning (ML) for intelligent controls on the Physical (PHY) and Medium Access Control (MAC) layers of future networks. B5G networks are expected to be intelligent enough to adapt to very dynamic topologies, intensive computation and storage applications, and diverse QoS requirements for ultra-high efficiency and resiliency purposes [7], [8].

To efficiently manage an mMIMO network, and to perform cognitive networking tasks, the network state which includes the spatial distribution and trajectories of the UEs, neighborhood relationships among the UEs, and handover boundaries among neighboring cells needs to be estimated. A novel Channel Charting (CC) framework is proposed for mMIMO systems in [9], exploiting the massive amounts of Channel State Information (CSI) available at the BSs. In CC unsupervised machine learning techniques are used to create a radio map of the cell served by a BS, which preserves the neighborhood relations of UEs, using features that characterize the large scale fading effects of the channel. The obtained CC can be used for local Radio Resource Management (RRM) in the cell. In [10], an extension to Multi-Point CC (MPCC)

978-1-7281-5320-9/20/$31.00 ©2020 IEEE

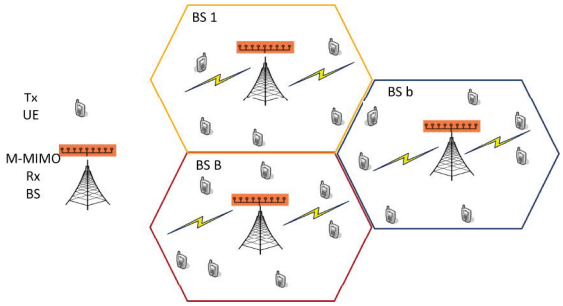

Fig. 1: Multipoint mMIMO system.

framework was discussed, to support advanced multi-cell RRM and to accurately map cell edge UEs. For improved charting performance, features are extracted and clustered based on advanced signal processing and ML techniques. In [11], the MPCC framework was generalized to incorporate new data points related to UEs at new locations, to an existing CC/MPCC, and to estimate the radio features related to a new location in the chart.

The level of channel variability in mmWave frequencies has widespread implications for virtually every aspect of cellular system design. In this paper, we consider Network Centric Handover (NCH) based on CC in a mmWave mMIMO system. To understand whether it is time to handover a user to a neighboring cell, the serving BS predicts the average SNR at neighboring BSs, based on radio signals received at the serving BS.

To the best of our knowledge, this is the first time that SNR prediction of neighboring BSs is exploited based on relative locations, such as CC. SNR prediction based on CC is an attractive approach since neither the physical location information nor the downlink channel measurement at the UE terminal are needed to predict the SNR of a neighboring BS. HO algorithms can be designed based on the predicted neighboring BS SNR, reducing power consumed at the UE, and signaling overhead.

The remainder of this paper is organized as follows. In Section II, the system model is introduced. In Section III and Section IV, the CC based NCH approach and SNR prediction of neighboring cells are presented, respectively. Numerical results are discussed in Section V. Finally, conclusions are drawn in Section VI.

II. System Model

The mMIMO mmWave cellular system under consideration is schematically shown in Figure 1. Each BS $b = 1, \ldots, B$ has M antenna elements and each UE $k = 1, \ldots, K$ has a single antenna element. For simplicity, we assume that BSs are equipped with Uniform Linear Arrays (ULAs). The channel vector of UE k received at BS b in one coherence bandwidth can then be modeled as [12]

$$\boldsymbol{h}_{k,b} = \sum_{l=1}^{L_{k,b}} \beta_{b,k}^{(l)} \boldsymbol{a}\big(\phi_{b,k}^{(l)}\big), \tag{1}$$

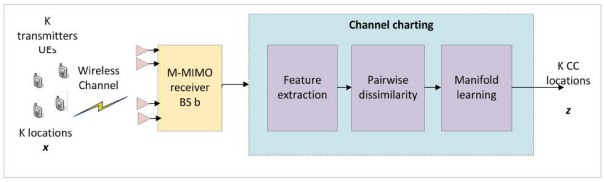

Fig. 2: The principle of channel charting.

where $L_{k,b}$ is the number of multi-path components for the wireless channel, $\phi_{b,k}^{(l)}$ is the direction of arrival of the lth path and $\beta_{b,k}^{(l)}$ the complex valued channel gain of the lth path, and $\boldsymbol{a}$ is BS array steering vector. For ULA, it can be modeled as

$$\boldsymbol{a}(\phi) = [1, \mathrm{e}^{i\frac{2\pi}{\mu} d \sin(\phi)}, \ldots, \mathrm{e}^{i\frac{2\pi}{\mu}(M-1)d\sin(\phi)}]^T, \tag{2}$$

where μ is the carrier wavelength, and d is the antenna spacing. The average uplink SNR $\gamma_{b,k}$ at BS b is computed as:

$$\gamma_{b,k} = \frac{P_0}{\sigma_0^2} \mathbb{E}\left[||\boldsymbol{h}_{b,k}||^2\right], \tag{3}$$

where P_0 is the transmit power of UE k, σ_0^2 is the noise power at BS b and $\mathbb{E}$ is the expectation operator.

The covariance matrix of UE k at BS b is computed as:

$$\boldsymbol{R}_{b,k} = \mathbb{E}[\boldsymbol{h}_{b,k}\boldsymbol{h}_{b,k}^H] = \boldsymbol{A}_{b,k}\boldsymbol{S}_{b,k}\boldsymbol{A}_{b,k}^H, \tag{4}$$

where the array steering vectors are collected to the matrix $\boldsymbol{A}_{b,k} = \left[\boldsymbol{a}\big(\phi_{b,k}^{(1)}\big), \ldots, \boldsymbol{a}\big(\phi_{b,k}^{(L_{k,b})}\big)\right]$, and the multi-path expected power components to the diagonal matrix $\boldsymbol{S}_{b,k} = \mathrm{diag}\left(\mathbb{E}\left[|\beta_{b,k}^{(1)}|^2\right], \ldots, \mathbb{E}\left[|\beta_{b,k}^{(L_{b,k})}|^2\right]\right)$.

A. Channel Charting Principle

Channel charting is based on the assumption that there is a continuous mapping from the spatial location $\boldsymbol{p}_k$ of UE k to the covariance matrix $\boldsymbol{R}_{b,k}$, given as [9], [10]:

$$\mathcal{H}_b : \mathbb{R}^s \to \mathbb{C}^{M\times M}; \mathcal{H}_b(\boldsymbol{p}_k) = \boldsymbol{R}_{b,k}\ . \tag{5}$$

Here s is the spatial dimension, which is either 2 or 3. A block diagram representing CC at BS b is shown in Figure 2. Using the estimated covariance CSI $\{\boldsymbol{R}_{b,k}\}_{k=1}^K$ collected at BS b form K unknown UE spatial locations $\{\boldsymbol{p}_k\}_{k=1}^K$, the CC finds a low dimension channel chart $\{\boldsymbol{z}_k\}_{k=1}^K$, such that

$$\|\boldsymbol{z}_k - \boldsymbol{z}_m\| \approx \alpha \|\boldsymbol{p}_k - \boldsymbol{p}_m\|, \text{ for } k, m \in \{1, \cdots, K\}, \tag{6}$$

where α is a scaling factor. Note that neither the UEs spatial locations $\boldsymbol{p}_k$ nor the locations of the BSs are needed; the CC is constructed in a *fully unsupervised* manner, solely based on the covariance CSI $\{\boldsymbol{R}_{b,k}\}_{k=1}^K$ at BS b.

A feature vector for UE k at BS b is constructed from the multi-path components of the mmWave channel as [10]:

$$\boldsymbol{f}_{b,k} = \left[\lambda_{b,k}^{(1)}, \ldots, \lambda_{b,k}^{(L_{b,k})}, \phi_{b,k}^{(1)}, \ldots, \phi_{b,k}^{(L_{k,b})}\right], \tag{7}$$

where $\lambda_{b,k}^{(l)} = \mathbb{E}\left[|\beta_{b,k}^{(l)}|^2\right]$ is the lth multi-path component power, and $\phi_{b,k}^{(l)}$ its angle of arrival. The multipath component parameters are estimated from the covariance matrix $\boldsymbol{R}_{b,k}$

using, e.g., the Multiple Signal Classification (MUSIC) algorithm [13].

Next, multipath components are clustered. For this, the Density-Based-Spatial-Clustering-of-Applications-with-Noise (DBSCAN) algorithm [14] can be used. The dissimilarity between two UEs (k, m) is based on identifying multi-path components in their feature vectors that are similar. In [10], carefully crafted radio features are used, and a data-driven CSI dissimilarity measure is developed, which are suitable for mm-Wave channels. The dissimilarity matrix has the elements $[\boldsymbol{D}_b]_{k,m} = d_b(\boldsymbol{f}_{b,k}, \boldsymbol{f}_{b,m})$ for $k, m = 1, \ldots, K$.

The channel chart is then obtained by applying a manifold learning algorithm to the pairwise dissimilarity matrix of the covariance CSI samples. The feature space is dimensionally reduced to a s-dimensional charting space;

$$\mathcal{C}_s : \boldsymbol{f}_k \mapsto \boldsymbol{z}_k \ , \tag{8}$$

where $\boldsymbol{z}_k \in \mathbb{R}^s$ represents the location of UE k in the chart, at the time of measuring the channel. Note that $\boldsymbol{z}_k$ may not have anything to do with a true location, the charting principle attempts to recover relative distances, and neighborhood information of UEs.

Different techniques such as principle component analysis, Sammon's Mapping, Laplacian eigenmaps, t-Distributed Stochastic Neighbor Embedding (t-SNE) and autoencoders are considered for channel charting in [9], [10], [15]. The quality of the generated CC is evaluated by measuring how well the points in the CC preserve the characteristics of the spatial geometry of the true UE locations using the Continuity (CT), Trustworthiness (TW) and Kruskal-Stress (KS) measures.

To use a channel chart for RRM functionalities, new UEs can be added to an existing CC based on their radio frequency CSI (i.e., covariance matrix) as in [11]. Then, from a CSI measurement of a new UE, possible CSI states can be predicted, by comparing to the CSI of nearby positions in the chart.

B. Handover Procedure

The HO process is a core element of cellular networks to support user mobility (see [16] for HO in GSM/CDMA and [3], [17], [18] in LTE/NR systems). Modeling the HO performance for 5G/ B5G is an active research area [19]–[22]. Novel techniques are needed to reduce signaling overhead, link failure probability and unnecessary HOs in mmWave and heterogeneous dense future networks.

HO is the process of changing the BS serving a mobile UE such that service continuity is guaranteed. Ideally, a UE is always served by the best BS. A simple rule for selecting the best BS is based on the average Received Signal Strength (RSS) level, i.e., the UE changes its association if another BS provides a higher RSS than the serving BS, which may happen when the user moves away from the serving BS towards another BS.

In general, HO is described by three phases: initiation, preparation, and execution. In the initiation phase, the UE reports reference signals measurements from neighboring BSs to the serving BS. For instance, the signal measurement report in 4G-LTE includes Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ) measurements (see [17]). In the preparation phase, signaling is exchanged between the serving BS and the target BS, and the target cell admission control procedures are ran. Once certain HO criteria are met, the user releases the serving BS channels and attempts to access the target BS using the Random Access Channel (RACH). After successful reception of an acknowledgment, the UE sends a confirmation message to notify the network that HO is executed. A HO procedure involves signaling overhead for both the UE, serving BS, target BS, and the core network. This is overhead from the perspective of payload data, which interrupts the data flow and decreases user plane throughput. The frequency at which such interruptions happen is a function of the relative values of the BS density and user velocity. The HO delay, measured from the beginning of the initiation phase to the end of the execution phase, can be significant. Mobility has a direct impact on user experience, since data transmission may subject to interruption due to high signaling overhead and the change of the serving BS [17], [23], or a radio link failure may interrupt the whole session. Therefore, at high velocities and/or dense cellular environments, it is desirable to decrease the frequency of handovers. This motivates investigating network-centric handover approaches to reduce HO delay, and signaling overhead.

C. Handover Procedure in mmWave Systems

Path loss at mmWave bands can be overcome by the use of massive MIMO systems. In mmWave cellular systems, the BS transmits multiple narrow beams towards the UE. Beamformed transmissions over up to 64 beams are allowed in 5G-NR for mmWave frequencies. In this architecture, in addition to inter-cellular handovers, beam switching handovers have to be considered [24]. Intra-cellular beam switching occurs when the target beam is selected among beams belong to the serving BS. Inter-cellular beam switching happens when the target beam is selected from a different BS [3], [17]. Hierarchical beamforming over multiple phases is assumed in 5G-NR where the initial beam acquisition is performed with wide beams [25].

When a UE moves around, it may experience frequent switching between neighboring beams although it stays in the same BS. Each beam acts like a small cell BS, which has its own physical and logical channels. The beams within the cell share the cell identity which is unique within the network, while each beam has a beam identity unique within the BS.

In the literature, two types of beam handover policies are considered; network controlled UE assisted beam switching, and UE controlled beam switching. In the former, beam switching is decided and initiated by the BS based on measurement reports provided by the UE. In the latter, the UE selects the target beam based on measurements and performs a beam switching procedure [24].

In this work, we consider fully network controlled handover. The UE uplink pilot signal is received at the serving BS and used to predict SNRs at neighboring BSs for handover

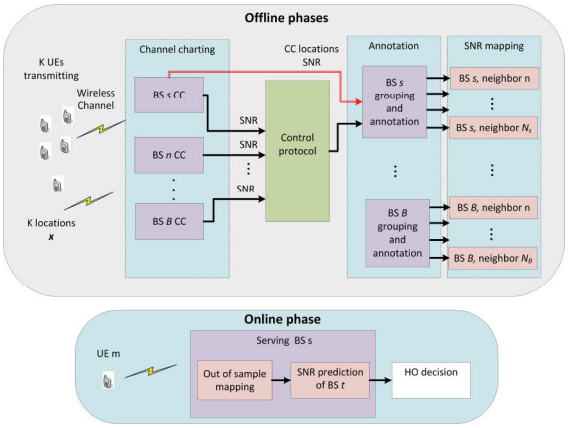

Fig. 3: The principle of Channel Charting based Network Controlled handover.

management. For simplicity, we assume that the UE has one transmit antenna, and we do not consider beam management. Inter-beam and intra-beam handover based on CC will be considered in future work.

III. CC Based Network Centric Handover

CC-based-Network-Centric-Handover (CCNCH) is based on the large scale radio features of a UE, measured at the serving BS. The serving BS has a CC, constructed offline, where the CC locations are annotated with measured SNR values of the neighboring BSs, and a prediction algorithm to predict the SNR of a neighboring BS. During online operation, the annotated CC, and the SNR prediction algorithm are used for making HO decisions for the current population of users. CCNCH is a distributed algorithm that is implemented at each BS. It is illustrated in Figure 3.

The input of CCNCH training is a set of measurements from UEs, collected in a short enough time scale such that the UE is in one location from the point of view of large scale channel characteristics. All BSs in the network that are able to detect the UE, measure the UE channels, and construct a received SNR from the UE. These measurements represent samples from the continuous function of radio signals from the s-dimensional spatial coverage area of the network to radio feature space.

Offline training consists of three phases: channel charting, annotation, and training of SNR prediction model.

- Channel charting phase: Each BS s considers the large scale radio features $\{\boldsymbol{R}_{s,k}\}_{k=1}^{K_s}$ of the set of UEs $\mathcal{K}_s$ which have a sufficient received signal quality at BS s. From this information, the BS constructs a single-cell channel chart $\mathcal{C}_s$ as explained in Section II-A. Note that $\mathcal{K}_s$ does not only have UEs for which s is the serving cell. The coverage area of $\mathcal{K}_s$ extends to cells neighboring s.
- Annotation phase: BS s gets the SNRs $\gamma_{t,k}$ measured at neighboring BS t for UEs $k \in \mathcal{K}_s^{(t)}$ in the subset

$$\mathcal{K}_s^{(t)} = \mathcal{K}_s \bigcap \mathcal{K}_t \tag{9}$$

of UEs that both BS s and t have information about. BS s collects this information from each neighboring BS t in the set $\mathcal{T}_s$ of neighbors of s, which consists of possible target cells for handovers initiated in s. The channel chart of BS s is then divided to groups, using knowledge of the own received SNR $\gamma_{s,k}$, in addition to the received SNR from neighboring BSs. There is a group $\mathcal{G}_s^{(t)}$ for each target BS $t \in \mathcal{T}_s$, as well as group $\mathcal{G}_s^{(0)}$. If $k \in \mathcal{G}_s^{(t)}$, it indicates that if k were served by s, knowing the SNR as received at t would be needed for handover decisions. If $k \in \mathcal{G}_s^{(0)}$, HO would not be actual for the user, i.e., the SNR at BS s is much stronger than at any neighboring BS. Note that the groups $\mathcal{G}_s^{(t)}$ for two different neighbors may overlap. For each group $\mathcal{G}_s^{(t)}$, an annotated channel chart $\mathcal{C}_s^{(t)}$ is created, where the charting coordinates $\mathbf{z}_k$ for $k \in \mathcal{G}_s^{(t)}$ are annotated with $\gamma_{t,k}$ and $\gamma_{s,k}$.
- SNR mapping: BS s uses the annotated CC $\mathcal{C}_s^{(t)}$ and a supervised learning algorithm to develop a function $g_s^{(t)}(\mathbf{z}, \gamma_s)$ to predict the SNR γ_t of the target BS t in this group of UE locations. The SNR prediction algorithm is explained in the next section.

In the online phase, the developed SNR mapping functions will be used by BS s to predict the neighbor-cell SNRs of a mobile user m that it serves. The prediction is solely based on received signal measurements at BS s. First, the BS estimates the large scale radio features $\boldsymbol{R}_{s,m}$ based on transmissions of the user. Using out-of-sample extension of CC [11], UE m is mapped to a CC location $\boldsymbol{z}_m$. Given the CC location $\boldsymbol{z}_m$ and received SNR $\gamma_{s,m}$, BS s determines which groups the user belongs to. For each possible target cell t, the BS estimates the SNR by applying $g_s^{(t)}(\mathbf{z}_m, \gamma_{s,m})$. A HO procedure can then be implemented based on the difference between SNRs $\gamma_{t,m}$ and $\gamma_{s,m}$ of the target and serving BSs. Other handover parameters such as the Time-to-Trigger (TTT) can be applied to avoid ping-pong events or early/late handover events.

IV. Neighbor Cell SNR Prediction

Based on the annotated channel chart $\mathcal{C}_s^{(t)}$, source cell BS s should find the function $g_s^{(t)}(\cdot)$ that predicts the SNR $\hat{\gamma}_t$ at target cell t for a transmission at any CC location $\mathbf{z}_m$ and received SNR $\gamma_{s,m}$ in the source cell.

We shall use machine learning methods to find this function. The training set contains information of UEs $k \in \mathcal{G}_s^{(t)}$. The inputs are channel chart coordinates $\mathbf{z}_k$ and serving cell SNRs $\gamma_{s,k}$, while the target cell SNRs $\hat{\gamma}_{t,k}$ are outputs. We consider dB-valued SNRs both at input and output, to have the dynamic range of the variables under control, and take the prediction Mean Squared Error (MSE) as the cost function to be minimized.

We consider three learning algorithms for prediction; Gaussian Process Regression (GPR), a Support Vector Machine (SVM), and Neural Networks (NN).

The use of Gaussian processes in machine learning has been comprehensively discussed in [26]. A Gaussian process is a stochastic process with a finite number of random variables

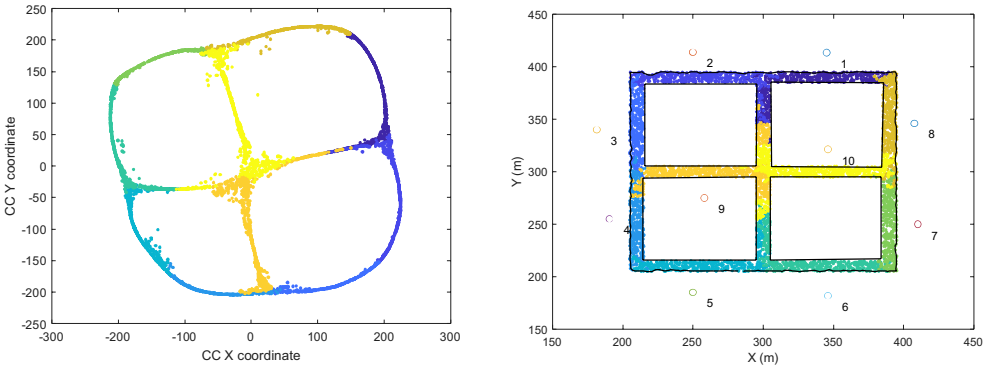

Fig. 4: Channel charting for a 10 BS network. Left: CC. Right: Ground truth spatial location in Manhattan Grid. Each color is associated to UEs that are served by a BS. BS locations indicated by numeric labeled circles.

following a joint Gaussian distribution. Given a training sample set, a Gaussian process generalizes these samples to a continuous function where each sample is considered to have a normal distribution. The key idea is that the correlation between the function values at different points depends on the input values, and this dependency can be determined through a covariance function, or a kernel. Taking the most probable value for each possible input value, one gets a GPR.

SVM is another well known kernel method for regression, which is able to learn the function between input and output by using kernel functions to maximize the margin between classes when transforming them to a higher dimensional space [27].

We compare GP and SVM to a NN based regression function [28]. For this, we consider a fully connected neural network, with three real-valued inputs, a number of hidden layers, and one real-valued output, providing the SNR prediction. For the learning process, forward and backward propagation phases are applied. First, weights are initialized randomly. In the forward phase, the input is fed to the network through input neurons, and is propagated across the hidden layers until the output layer. The error between the predicted output and the given output from the input data is calculated. Then, in the backward phase, based on the error, we use the Levenberg-Marquardt method to adjust the weights and biases so that the output MSE is minimized. We split the data set randomly into three sets; 80% is used for training, 10% for validation and 10% for testing. To avoid overfitting, during training, the performance is tested against the validation set. Once the validation error is larger than the training error for six consecutive iterations, we backtrack to the weights that provided the smallest validation error. Prediction results are then provided for the testing set.

The prediction accuracy of a NN depends on the weight initialization; depending on initialization, the NN may converge to different local optima. To mitigate this, we use several random initializations to generate multiple NN predictors. The NN that provides the best performance on the validation set is selected.

To investigate the role of initialization we adopt a systematic approach to determine the average performance as a function of the number of random initializations. We set the maximum number of initializations $N_{\max}$, generate the corresponding NNs and evaluate the performance using the MSE of the validation set. The average performance of a set $\mathcal{I}_n$ consisting of n random initializations for $1 < n < N_{\max}/2$ is evaluated by considering G sets each of size n. For each set the NN with best performance is selected, then averaged over G sets.

V. Simulation

An urban outdoor multi-cell mmWave scenario is considered as discussed in [10]. The system parameters are shown in Table I. A ray tracing channel model is used to generate multi-path channels. The UE locations are randomly generated on the streets of a Manhattan grid. The CSI of the UEs are estimated at multiple BSs.

First, to get a gist of CCNCH in a multicellular environment, we consider a scenario with 10 BSs, 6 streets and 5000 UEs. A channel chart based on Laplacian eigenmaps is generated as shown in Figure 4. Each UE in this figure is colored based on its best BS. For comparison, the ground truth location of UEs and BSs are also shown.

For each of the BSs $s = 1, \ldots, 10$, we construct CC groups $\mathcal{G}_s^{(t)}$ for handover target cells $t \in \mathcal{T}_s$ so that $\mathcal{G}_s^{(t)}$ consists of the UEs for which the two best BSs are s and t in any order. The SNR mapping function to predict the target BS t in each group $\mathcal{G}_s^{(t)}$ $s = 1, \ldots, 10$ and $t \in \mathcal{T}_s$ is then created. NN, GPR and SVM predictors are used.

For GPR, we select the exponential kernel, since it has the best performance compared to other kernels functions for the data set. A NN with three hidden layers with 10 neurons each is used.

TABLE I: Simulation parameters [10].

Parameter	Value	Parameter	Value
Carrier frequency	28GHz	Bandwidth	256MHz
UE Tx power	23dBm	BS noise power	-86dBm

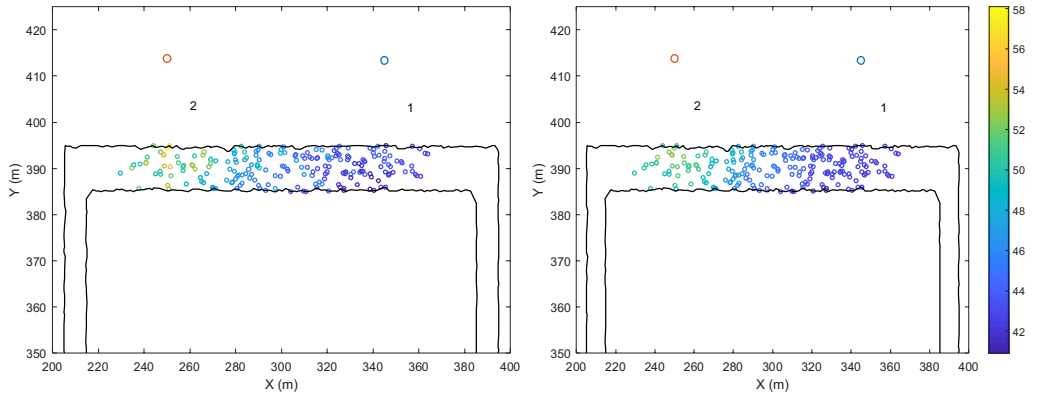

Fig. 5: Target BS SNR as function of ground truth location. BS 1 & 2 locations marked. Left: Ground truth SNR. Right: Predicted SNR.

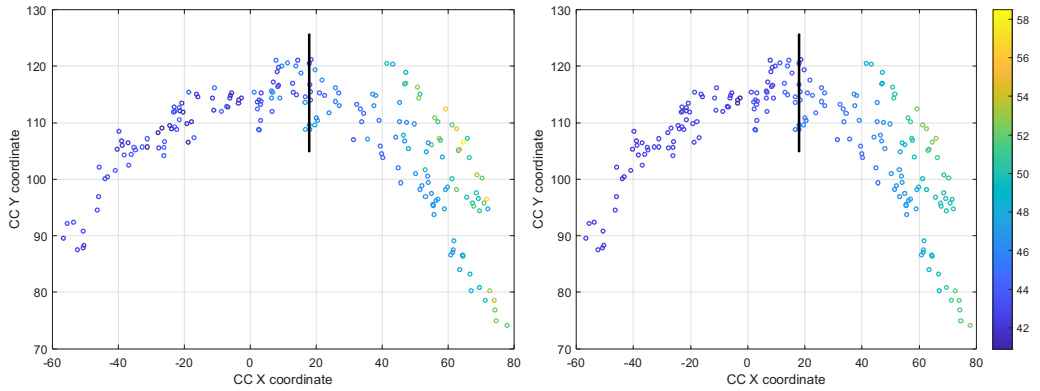

Fig. 6: Target BS SNR as function of CC location (arbitrary length scale). Left: Ground truth SNR. Right: Predicted SNR. Approximative cell boundary drawn.

The Root Mean Squared Error (RMSE) of the predictors are measured. The standard deviation of the RMSEs of the predictors generated for different (s,t) pairs is also measured. Both of these are measured in dBs. In Table II, a comparison of the best NN to GPR and SVM is shown. NN outperforms the other regression methods in terms of RMSE.

In the full network simulation, the data set for learning the SNR prediction function was rather limited. For an ordered pair (s,t) on average some 550 UE locations are sampled. Already with this limited training set, NN outperforms the other predictors. To further clarify the merits of the considered predictors, we construct a larger data set for a pair of cells. We drop a large number of UEs in the street where BSs 1 and 2 are located, and select at random a set of 7000 UE positions that belong to $\mathcal{G}_1^{(2)}$. Channel charting is performed based on the CSI of these UEs, as measured at BS 1, and the CC is annotated with the SNRs of BSs 1 and 2.

TABLE II: Performance comparison for the whole network; a data set of 5000 users in the 10 cells.

Algorithm	NN	GPR	SVM
RMSE	1.41	1.53	1.68
std (RMSE)	0.26	0.16	0.20

The effect of the number of initializations on NN performance is investigated in Figure 7. The average RMSE performance (with standard deviation error bars) for a NN with $[20,20]$ hidden layers is plotted as a function of the number of initializations n. Performance improves considerably when increasing the number of initializations from $n=1$ to $n=50$. Increasing the number of initializations to $n>150$, only marginal performance gain can be achieved as compared to $n=150$, which is used to create the results below.

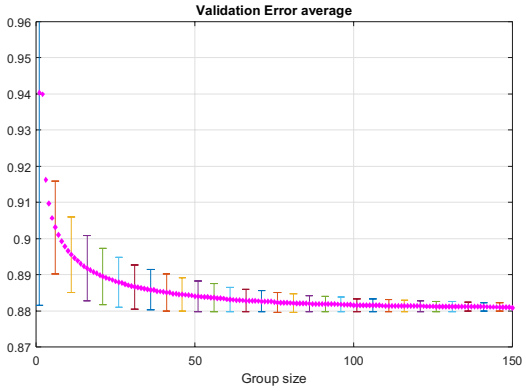

Fig. 7: Validation error vs. number of initializations for NN.

TABLE III: Performance of different NN structures & algorithms; a data set of 7000 users in $\mathcal{G}_1^{(2)}$.

Hidden layer structure / Algorithm	RMSE
10	0.95
20	0.93
30	0.91
[10 10]	0.89
[20 20]	0.87
[30 30]	0.88
[10 10 10]	0.86
[20 20 20]	0.88
[30 30 30]	0.88
[10 10 10 10]	0.88
[20 20 20 20]	0.88
GPR	1.01
SVM	1.55

We examine different NN structures in this setting. Results for neural networks with different number of neurons and hidden layers are shown in Table III. The results show the benefit of a deeper network for the prediction task at hand. Adding more layers to the NN we can better capture the non-linearity of the predicted SNR. Based on RMSE, a NN with three hidden layers and 10 neurons in each hidden layer outperforms other structures. Growing the network beyond this seems counterproductive. When choosing the best of $n = 150$ initializations, the standard deviation between different initializations is ~ 0.03 for all of these NN structures. GPR also benefits from the larger data set, but not as much as NN.

A sample of BS $s = 1$ predicting the SNR of $t = 2$ is depicted in Figures 5 and Figure 6. Figure 5 shows the ground truth and predicted BS 2 SNR values in dB, plotted against the ground truth locations. Figure 6, in contrast, shows the same ground truth and predicted SNR values, plotted against the CC locations $\mathbf{z} \in \mathcal{C}_s^{(t)}$, which are used as input for predictor training. In addition, an approximate cell boundary is drawn, to sketch the place where BS 2 becomes better than BS 1.

For HO, between cells s and t, the crucial variable to control is the ratio of the SNRs γ_s and γ_t, or their dB-domain difference $\gamma_s - \gamma_t$. Figure 8 plots the predicted SNR difference against the ground truth for $s = 1, t = 2$, with prediction based on GPR. Similarly, Figure 9 shows the predicted SNR difference based on a NN with structure [10 10 10], against the ground truth. In addition, the 95% confidence interval for the SNR difference prediction using GPR is plotted in Figure 10.

Interestingly, the edge of NN against GPR seems to arise in the domain where γ_1 is comparable or larger than γ_2, while for γ_2 much larger than γ_1, the algorithms perform similarly. This is encouraging for the prospect of using an NN-based SNR predictor for CCNCH, as the ideal handover location would be $\gamma_2 = \gamma_1$, and this border would be approached from a direction where $\gamma_1 > \gamma_2$.

VI. Conclusion

We have considered an algorithm for learning the SNR of a user in a neighboring cell from the signal received in a

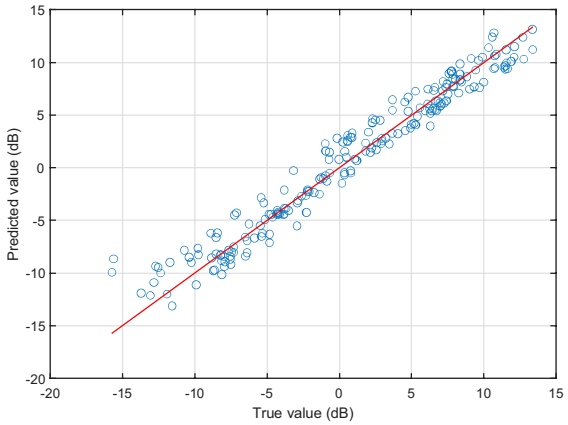

Fig. 8: Predicted $\gamma_{1,k} - \gamma_{2,k}$ value vs ground truth for GPR predictor.

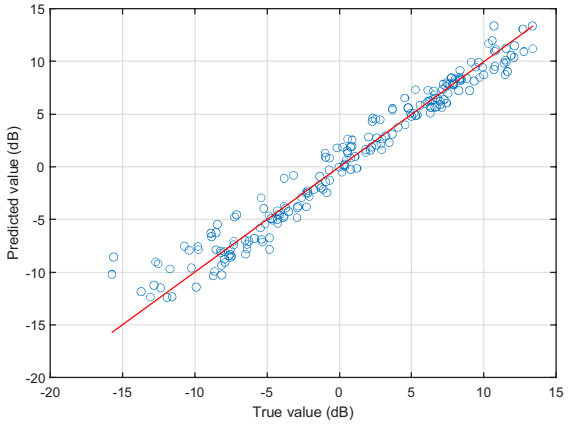

Fig. 9: Predicted $\gamma_{1,k} - \gamma_{2,k}$ value vs ground truth for NN predictor.

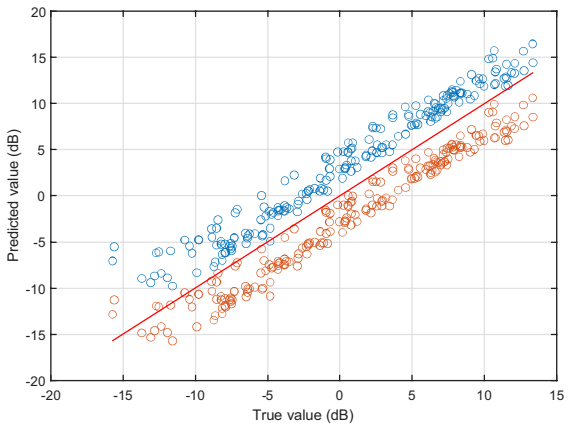

Fig. 10: Upper and lower bounds of the 95% confidence intervals for predicted value of $\gamma_{1,k} - \gamma_{2,k}$ vs. ground truth using GPR predictor.

serving cell. The learning is based on a channel chart, which is a dimensional reduction of multiuser radio channel state information measured at the serving BS. A handover algorithm can be designed based on the predicted SNR of the target BS. We apply a network centric handover model, based on uplink CSI measurements at the serving cell. In online operation, when SNRs of users are predicted for handover management purposes, neither downlink channel measurements at the UE terminal nor measurements performed at neighboring base stations, nor physical location information is needed to predict the SNR of a user at a neighboring BS. Three different regression learners have been considered for SNR prediction; Gaussian Process Regression, Support Vector Machines, and Neural Networks. Performance of each learner is evaluated based on RMSE. Simulation results show that NN outperforms other methods. In future work, the NN structure will to be refined in order to reduce the prediction error. Refined cost functions which target the prediction capability towards the region most important for handovers, will be considered, and beam handover, pertinent for 5G networks, will be investigated.

ACKNOWLEDGMENT

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

REFERENCES

[1] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd quarter 2018.

[2] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, 4th quarter 2015.

[3] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 1st quarter 2019.

[4] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kurner, "The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 10–27, 1st quarter 2019.

[5] M. Koivisto, A. Hakkarainen, M. Costa, P. Kela, K. Leppanen, and M. Valkama, "High-efficiency device positioning and location-aware communications in dense 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 188–195, Aug. 2017.

[6] N. Malm, L. Zhou, E. Menta, K. Ruttik, R. Jäntti, O. Tirkkonen, M. Costa, and K. Leppänen, "User localization enabled Ultra-dense network testbed," in *Proc. IEEE 5G World Forum, (5GWF)*, Jul. 2018, pp. 405–409.

[7] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2019.

[8] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE Access*, vol. 6, pp. 32 328–32 338, 2018.

[9] C. Studer, S. Medjkouh, E. Gönültaş, 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.

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

[11] T. Ponnada, H. Al-Tous, O. Tirkkonen, and C. Studer, "An out-of-sample extension for wireless multipoint channel charting," in *Proc. EAI International Conference on Cognitive Radio Oriented Wireless Networks, (Crowncom)*, Jun. 2019.

[12] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE J. Selected Areas in Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[13] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[14] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Internat. Conf. on Knowledge Discovery and Data Mining*. AAAI Press, 1996, pp. 226–231.

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

[16] G. P. Pollini, "Trends in handover design," *IEEE Commun. Mag.*, vol. 34, no. 3, pp. 82–90, Mar. 1996.

[17] M. Tayyab, X. Gelabert, and R. Jäntti, "A survey on handover management: From LTE to NR," *IEEE Access*, vol. 7, pp. 118 907–118 930, 2019.

[18] D. Xenakis, N. Passas, L. Merakos, and C. Verikoukis, "Handover decision for small cells: Algorithms, lessons learned and simulation study," *Computer Networks*, vol. 100, pp. 64 –74, 2016.

[19] R. Arshad, H. Elsawy, S. Sorour, T. Y. Al-Naffouri, and M. Alouini, "Handover management in 5g and beyond: A topology aware skipping approach," *IEEE Access*, vol. 4, pp. 9073–9081, 2016.

[20] A. Alhammadi, M. Roslee, M. Y. Alias, I. Shayea, and S. Alraih, "Dynamic handover control parameters for LTE-A/5G mobile communications," in *Proc. of Advances in Wireless and Optical Communications, (RTUWO)*, Nov. 2018, pp. 39–44.

[21] R. Parada and M. Zorzi, "Context-aware handover in mmWave 5G using UE's direction of pass," in *Proc. European Wireless Conference*, May 2018, pp. 1–6.

[22] F. B. Mismar and B. L. Evans, "Partially blind handovers for mmWave new radio aided by Sub-6 GHz LTE signaling," in *Proc. IEEE Internat. Conference on Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–5.

[23] T. Mahmoodi and S. Seetharaman, "On using a SDN-based control plane in 5G mobile networks," in *Proc.Wireless World Res. Forum, Meeeting*, vol. 32, 2014, pp. 1–6.

[24] N. Sung and Y. Choi, "Fast intra-beam switching scheme using common contention channels in millimeter-wave based cellular systems," in *Proc. Internat. Conf. on Advanced Commun. Technology, (ICACT)*, Jan. 2016, pp. 760–765.

[25] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial ue discovery in Millimeter-Wave MIMO systems," *IEEE J. Selected Topics in Sign. Proc.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[26] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. Cambridge, MA: MIT press, 2006.

[27] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.

[28] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2019.