

RESEARCH

Open Access



Multipoint channel charting-based radio resource management for V2V communications

Hanan Al-Tous^{1*} , Tushara Ponnada¹, Christoph Studer² and Olav Tirkkonen¹

*Correspondence:

hanan.al-tous@aalto.fi

¹Department of Communications and Networking, Aalto University, Espoo, Finland

Full list of author information is available at the end of the article

Abstract

We consider a multipoint channel charting (MPCC) algorithm for radio resource management (RRM) in vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems. A massive MIMO (mMIMO) infrastructure network performs logical localization of vehicles to a MPCC, based on V2I communication signals. Combining logical distances given by channel charting with V2V measurements, the network trains a function to predict the quality of a direct V2V communication link from observed V2I communication signals. In MPCC, the network uses machine learning techniques to learn a logical radio map from V2I channel state information (CSI) samples transmitted from unknown locations. The network extracts CSI features, constructs a dissimilarity matrix between CSI samples, and performs dimensional reduction of the CSI feature space. Here, we use Laplacian Eigenmaps (LE) for dimensional reduction. The resulting MPCC is a two-dimensional map where the spatial distance between a pair of vehicles is closely approximated by the distance in the MPCC. In addition to V2I CSI, the network acquires V2V channel quality information for vehicles in the training set and develops a link quality predictor. MPCC provides a mapping for any vehicle location in the training set. To use MPCC for cognitive RRM of V2I and V2V communications, network management has to find logical MPCC locations for vehicles not in the training set, based on newly acquired V2I CSI measurements. For this, we develop an extension of LE-based MPCC to out-of-sample CSI samples. We evaluate the performance of link quality prediction for V2V communications in a mMIMO millimeter-wave scenario, in terms of the relative error of the predicted outage probability.

Keywords: Massive MIMO, Channel charting, Laplacian Eigenmaps, Out-of-sample extension, V2I, V2V

1 Introduction

Communication technologies are becoming integrated in vehicles for safety applications, such as blind spot warning and forward collision warning, as well as for non-safety-related applications such as toll collection and infotainment [1]. The dedicated short-range communication (DSRC) protocol can be used both for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, and provides a coverage range of about 1 km and achieves data rates in the range of 2–6 Mbps [2]. 4G-LTE connectivity below 6 GHz can be used for V2I, achieving a data rate of up to 100 Mbps [3].

Next generation vehicles are expected to become automated and to contain hundreds of sensor nodes. The increase in the number of sensors will generate a huge amount of data that can be utilized for different applications. It is expected that autonomous cars will comprise 75% of total traffic on the road by the year 2040. There are many benefits of sharing rich sensor data with other vehicles and infrastructure. However, this will require exchanging a large amount of data, from tens to thousands of megabits per second. The state-of-the-art vehicular communication standard DSRC is not sufficient to handle such high data rates in next generation vehicles [4].

The large bandwidth channels at millimeter-wave (mm-Wave) are a promising candidate to realizing high data rates and is of prime interest for 5G and Beyond 5G (B5G) communication [5]. Massive MIMO (mMIMO) is another promising technology for 5G and B5G, with the potential to provide high spectral and power efficiency. In a mMIMO cell, each base station (BS) has a large number of antennas, which can provide a simultaneous use of the resource (e.g., frequency and/or time slots) for multiple user equipments (UEs) in the cell [5–7]. Furthermore, the high spatial resolution exploited by the large-scale antenna arrays used at the mMIMO BSs can be used for many applications, such as UE positioning and environment mapping [8–10].

In [11], test results of mm-Wave for V2V and V2I communications are reported. The results are promising, while it is indicated that much research is still needed to develop the physical (PHY) and medium access control (MAC) layers for mm-Wave systems to provide a reliable basis for V2V and V2I communication. A key challenge in developing mm-Wave systems is the potential for rapid channel dynamics; mm-Wave propagation suffers from high path loss, reduced diversity, and increased effect of blockage by obstacles [12]. mm-Wave BSs have to use beamforming for transmission in order to increase the signal-to-noise ratio, reaching a radius of up to 200 m. Hence, hundreds of BSs will be needed to cover large spaces. Modeling, measuring, and predicting the radio channel characteristics of mm-Wave systems for V2V communications are the currently active research areas [3, 13]. Successful deployment of mm-Wave systems requires new management procedures to handle resource-constrained devices, radio resource management, heterogeneous networking, and computing infrastructures [4, 5, 14, 15]. The level of channel variability in mm-Wave has widespread implications for virtually every aspect of V2V communications.

Motivated by the burgeoning progress of artificial intelligence (AI) and its breakthroughs in a variety of domains, the B5G research community is currently seeking solutions from machine learning (ML) for intelligent control of PHY and 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 Quality of Service (QoS) requirements [16–19].

To efficiently manage B5G networks and to perform cognitive networking tasks, the network states which include the spatial distribution and trajectories of the UEs, neighborhood relationships among the UEs, and handover boundaries among neighboring cells need to be estimated. A novel ML framework called channel charting (CC) based on the massive amounts of channel state information (CSI) available at the base stations is proposed for a single-cell MIMO system in [20]. CC applies unsupervised ML techniques to create a radio map of the cell served by the 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. However, cell edge UEs may not be accurately located in the chart due to their low signal-to-noise ratio (SNR) at the cell edge. In [21], a multipoint CC (MPCC) framework is proposed to support advanced multicell RRM and to accurately map cell edge UEs. First, each BS generates its own dissimilarity matrix between the users it can decode; then, the dissimilarity matrices are fused and used to construct the MPCC. The trustworthiness and continuity measures show that the proposed MPCC is capable to preserve the neighborhood structure between UEs in the network.

MPCC-based approach entails more computational efforts compared to other approach at the BSs to compute the dissimilarity matrix between the UEs seen by the same BS.

In this paper, we consider MPCC in V2I networks, where vehicular UEs communicate with infrastructure BSs. Using only uplink radio channel features, a logical MPCC map is constructed for the network. Furthermore, some of the UEs have the capability of V2V communications. To enable V2V connectivity prediction, radio link quality information of V2V pairs is collected and used to build a link quality prediction (LQP) model utilizing the MPCC distance between V2V pairs.

To use MPCC for online RRM, it is important to generalize the chart, allowing the incorporation of new data to an existing MPCC and/or to estimate the features related to a location in the chart. As the radio channel features of a UE can change rapidly in a small distance, it is important to accurately estimate the MPCC location of data from a UE that was not included in the training data set (out-of-sample UE). In this paper¹, an extension-of-MPCC (EMPCC) to out-of-sample data points is considered. This is a general framework that is needed to implement any online RRM function using CC.

This paper investigates V2V link quality prediction based on an MPCC approach. MPCC-based LQP for V2I/V2V consists of two phases: an offline training and online usage phase. In the training phase, V2I and V2V radio channel features of a large number of UEs are used to construct the MPCC and LQP model, respectively. In the online phase, given the radio features of active vehicles (UEs), the EMPCC algorithm is used to map the UEs to CC locations. Based on the CC distance and LQP model, the possibility of V2V communication for a given pair of vehicles is evaluated. All simulation and modeling are performed in an mm-Wave context, lending credibility for the considered solutions for mm-Wave-based V2I/V2V. It is worth noting here that the proposed MPCC-based LQP for V2I/V2V is not restricted to mm-Wave communications and can be used for other radio frequencies.

In LQP based on MPCC, neither physical location information, downlink channel measurement at the vehicular terminals, nor V2V measurements are needed for predicting V2V connectivity. Advanced power allocation and beam alignment algorithms for V2V communications can be then designed based on LQP and MPCC.

The remainder of this paper is organized as follows. Section 2 presents the system model of V2I and V2V communications. In Section 4, the MPCC and LQP and EMPCC frameworks are presented. Numerical results are presented and discussed in Section 5. Finally, conclusions are drawn in Section 6.

¹Part of the results of this paper were presented in [22].

1.1 Notation

We adopt the following notation: matrices and vectors are set in upper and lower boldface, respectively. $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $|\cdot|$, $\|\cdot\|_p$ denote the transpose, the conjugate, the Hermitian, the absolute value, and the p -norm, respectively. $\text{Tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. Calligraphic letters denote sets, e.g., $\mathcal{G}$, and $|\mathcal{G}|$ denotes the cardinality of $\mathcal{G}$. $\mathbb{R}_+$ is the set of non-negative real numbers, $\mathbb{C}$ is the set of complex numbers, $\mathbb{C}^{N \times M}$ is the space of $N \times M$ matrices and $\mathbb{E}[\cdot]$ denotes expectation, and $\iota = \sqrt{-1}$.

2 System model

The system under consideration is schematically shown in Fig. 1.

Each infrastructure BS $b = 1, \dots, B$ has M antenna elements. In the network, two types of UEs are assumed: V2I UEs and V2V UEs.

Each UE of V2I type has a single antenna element, whereas UE of V2V type has $N + 1$ antenna elements, one is used for V2I communications, and N antennas for V2V communication. In V2I communications, the base station antenna is at an elevated position, 10–25 m above ground. This is not the case in V2V communications; both the transmit (Tx) and receive (Rx) antennas are at the same height relatively close to the ground level, at some 1–2 m above ground, by having antennas close to the ground level, shadowing effects from other vehicles and surrounding buildings are expected to be stronger. To handle this issue, multiple antennas are used at both the Tx and Rx terminals [3]. Note that UEs of type V2I can have more than one antenna; however, it is shown that one element at the UE can be used to construct an accurate MPCC [21].

The V2I channel vector of UE $k = 1, \dots, K$ using a uniform-linear-array (ULA) at BS b for a coherence bandwidth can be modeled as [23]:

$$\mathbf{h}_{b,k} = \sum_{l=1}^{L_k} \beta_{b,k}^{(l)} \mathbf{a}(\phi_{b,k}^{(l)}), \quad (1)$$

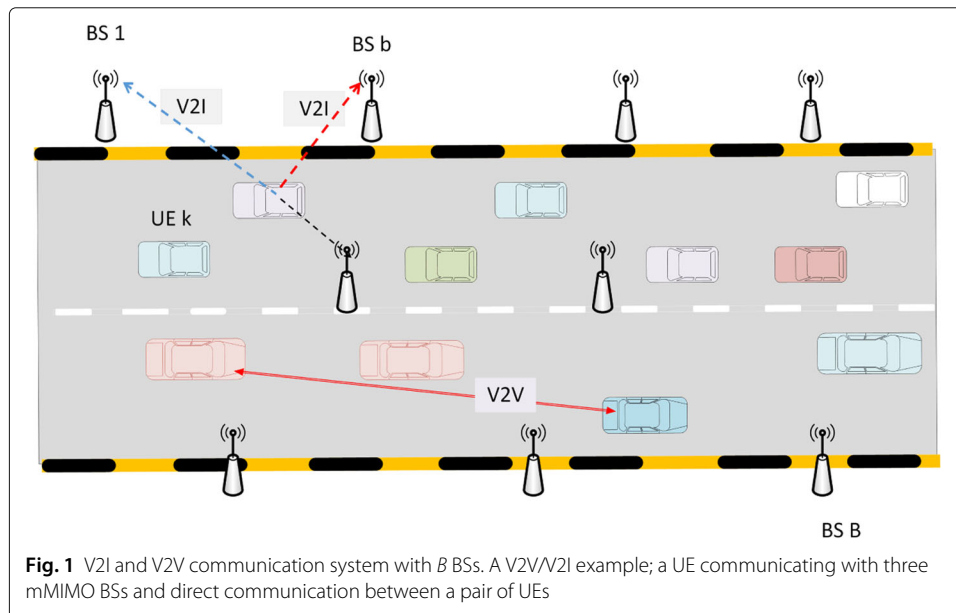


Fig. 1 V2I and V2V communication system with B BSs. A V2V/V2I example; a UE communicating with three mMIMO BSs and direct communication between a pair of UEs

where L_k is the number of multipath components for the wireless channel between UE k and BS b , $\phi_{b,k}^{(l)}$ is the direction of arrival of the l th path, $\beta_{b,k}^{(l)}$ is the complex-valued gain of the l th path, and $\mathbf{a}(\cdot)$ is the BS steering vector. For ULA, the steering vector is:

$$\mathbf{a}(\phi) = [1, e^{i\frac{2\pi}{\lambda}s \sin(\phi)}, \dots, e^{i\frac{2\pi}{\lambda}(M-1)s \sin(\phi)}]^T, \quad (2)$$

where λ is the carrier wavelength, and s is the antenna spacing. The covariance $\mathbf{R}_{b,k} \in \mathbb{C}^{M \times M}$ of the CSI $\mathbf{h}_{b,k}$ used to extract the features at BS b becomes:

$$\mathbf{R}_{b,k} = \mathbb{E}[\mathbf{h}_{b,k}\mathbf{h}_{b,k}^H] = \mathbf{A}_{b,k}\mathbf{S}_{b,k}\mathbf{A}_{b,k}^H, \quad (3)$$

where $\mathbb{E}$ is the expectation operator, $\mathbf{A}_{b,k} = [\mathbf{a}(\phi_{b,k}^{(1)}), \dots, \mathbf{a}(\phi_{b,k}^{(L_k)})]$ is a matrix of array steering vectors, and $\mathbf{S}_{b,k} = \text{diag}(\mathbb{E}[|\beta_{b,k}^{(1)}|^2], \dots, \mathbb{E}[|\beta_{b,k}^{(L_k)}|^2])$ is a diagonal matrix of multipath power components.

For V2V communication between UEs i and j , the channel matrix is denoted as $\mathbf{H}_{i,j} \in \mathbb{C}^{N \times N}$, and the channel covariance matrix at receiver terminal j is:

$$\mathbf{Q}_{i,j} = \mathbb{E}[\mathbf{H}_{i,j}\mathbf{H}_{i,j}^H]. \quad (4)$$

The received signal vector at UE j is:

$$\mathbf{y}_{i,j} = \sqrt{P}\mathbf{H}_{i,j}\mathbf{w}_{i,j}x + \mathbf{n}_j, \quad (5)$$

where x is the transmitted symbol with $\mathbb{E}[|x|^2] = 1$, P is the transmitted power, $\mathbf{n}_j \in \mathbb{C}^{N \times 1}$ is the received white Gaussian noise, and $\mathbf{w}_{i,j}$ is the beamformer weight. Assuming Tx i knows the statistics of the wireless channel, the beamformer weight $\mathbf{w}_{i,j}$ is selected as the Eigenvector $\mathbf{u}$ corresponding to the largest Eigenvalue of the covariance matrix $\mathbf{Q}_{i,j}$. The average V2V SNR at UE j can be computed as [24]:

$$\Gamma_{i,j} = \frac{P}{\sigma_n} \mathbb{E}[\text{Tr}[\mathbf{H}_{i,j}\mathbf{w}_{i,j}\mathbf{w}_{i,j}^H\mathbf{H}_{i,j}^H]] = \frac{P}{\sigma_n} \lambda_{\max}, \quad (6)$$

where Tr is matrix trace operator and σ_n is the received noise power. The latter equality holds for the adopted Eigenbeamformer, and $\lambda_{\max}$ is the maximum Eigenvalue of $\mathbf{Q}_{i,j}$.

3 Channel charting

3.1 Feature extraction and dissimilarity matrix

Large-scale effects of wireless channel are caused by reflection, diffraction, and scattering of the physical environment, whereas small-scale effects are caused by multipath propagation and related destructive/constructive addition of signal components. CC is based on the assumption that statistical properties of MIMO channel vary relatively slowly across space, on a length-scale related to the macroscopic distances between scatterers in the channel, not on the small fading length-scale of wavelengths. In this regard, the CSI covariance matrix can be used to capture large-scale effects of the wireless channel based on the assumption that there is a continuous mapping from the spatial location $\mathbf{p}_k$ of UE k to the covariance CSI $\mathbf{R}_{b,k}$ [20, 21]:

$$\mathcal{H}_b : \mathbb{R}^d \rightarrow \mathbb{C}^{M \times M}; \quad \mathcal{H}_b(\mathbf{p}_k) = \mathbf{R}_{b,k}. \quad (7)$$

Here, d is the spatial dimension which is either 2 or 3.

CC starts by processing the CSI covariance matrix $\mathbf{R}_{b,k}$ into suitable channel features $\mathbf{f}_{b,k}$ that capture large-scale properties of the wireless channel. CC then proceeds by using the set of collected features $\{\mathbf{f}_{b,k}\}_{k=1}^{K_b}$ for the set of UEs $\mathcal{K}_b = \{1, \dots, K_b\}$ seen by BS b to

learn the dissimilarity matrix $\mathbf{D}_b \in \mathbb{R}_+^{K_b \times K_b}$. The pairwise dissimilarity $[\mathbf{D}_b]_{k,m}$ between UEs k and m , for $k, m \in \mathcal{K}_b$ measures the dissimilarity of the radio features between UEs k and m . Different approaches can be used to select the channel features and then computing the dissimilarity matrix (see [20, 21]). In this paper, we select the feature vector $\mathbf{f}_{b,k}$ based on multipath components [21]:

$$\mathbf{f}_{b,k} = [\lambda_{b,k}^{(1)}, \dots, \lambda_{b,k}^{(L_k)}, \phi_{b,k}^{(1)}, \dots, \phi_{b,k}^{(L_k)}], \quad (8)$$

where $\lambda_{b,k}^{(l)} = \mathbb{E}[\beta_{b,k}^{(l)2}]$. The multipath components (power and phase) $\{\lambda_{b,k}^{(l)}\}_{l=1}^{L_k}$ and $\{\phi_{b,k}^{(l)}\}_{l=1}^{L_k}$ of UE k at BS b are estimated from the CSI covariance matrix $\mathbf{R}_{b,k}$ using the multiple signal classification (MUSIC) algorithm [25]. The dissimilarity between two UEs (k, m) is based on identifying multipath components in their feature vectors that are similar. For this, the components of feature vectors are transformed to Cartesian coordinates as [21]:

$$\mathcal{F}\{\mathbf{f}_{b,k}\} = [\mathbf{x}_{b,k}^{(1)}, \dots, \mathbf{x}_{b,k}^{(L_k)}], \quad (9)$$

where $\mathbf{x}_{b,k}^{(l)} = \left[\frac{\cos(\phi_{b,k}^{(l)})}{\sqrt{\lambda_{b,k}^{(l)}}}, \frac{\sin(\phi_{b,k}^{(l)})}{\sqrt{\lambda_{b,k}^{(l)}}} \right]^T$. To cluster multipath components to clusters deemed to be similar, the density-based spatial clustering of applications with noise (DBSCAN) algorithm [26] is used to label the multipath components² $\{\mathcal{F}\{\mathbf{f}_{b,k}\}\}_{k=1}^{K_b}$. This results in a label $\mathcal{L}(\mathbf{x}_{b,k}^{(l)}) \in \{C_1, \dots, C_N\}$ for each multipath component, where C_n is the label of the n th cluster. The dissimilarity coefficient between a pair of UEs (k, m) then is computed taking into consideration multipath components of the UEs that are in the same cluster. The pairwise dissimilarity is computed as:

$$[\mathbf{D}_b]_{k,m} = \begin{cases} \|\mathbf{x}_{b,k}^{(i')} - \mathbf{x}_{b,m}^{(j')}\|_2 & \text{if } \mathcal{L}(\mathbf{x}_{b,k}^{(i')}) = \mathcal{L}(\mathbf{x}_{b,m}^{(j')}) \\ \|\mathbf{x}_{b,k}^{(1)} - \mathbf{x}_{b,m}^{(1)}\|_2 & \text{otherwise,} \end{cases} \quad (10)$$

where $[i', j'] = \arg \max_{i,j} \min(\lambda_{b,k}^{(i)}, \lambda_{b,m}^{(j)})$.

3.2 Multipoint channel charting

MPCC utilizes the different views of the spatially distributed BSs by fusing the BS-specific dissimilarity matrices $\mathbf{D}_b$, $b = 1, \dots, B$ into a global dissimilarity matrix $\mathbf{D}$ [21]. The benefits of having multiple spatially distributed BSs can be utilized by merging the BS-specific dissimilarity matrices $\{\mathbf{D}_b\}_{b=1}^B$ into a global dissimilarity matrix $\mathbf{D}$, where the (k, m) th element $[\mathbf{D}]_{k,m}$ can be computed as:

$$[\mathbf{D}]_{k,m} = \frac{1}{\sum_{b=1}^B \omega_b(k, m)} \sum_{b=1}^B \omega_b(k, m) [\mathbf{D}_b]_{k,m}, \quad (11)$$

where $\omega_b(k, m)$ is a weighting factor computed as $\omega_b(k, m) = \min(\gamma_{b,k}, \gamma_{b,m})^2$ and $\gamma_{b,k}$ is the SNR of the wireless link between UE k and BS b .

3.3 Dimensionality reduction and Laplacian Eigenmaps

CC finds in an unsupervised manner a low dimensional channel chart providing logical locations $\mathbf{Z} = \{\mathbf{z}_k\}_{k=1}^K$ for the sample UEs such that neighboring UEs will be neighboring

²A non-linear transformation can be applied to make clusters of multipath components separable.

points in the channel chart, i.e., CC preserves the local geometry. The relation between the logical and physical locations is approximative:

$$\|\mathbf{z}_k - \mathbf{z}_m\| \approx \alpha \|\mathbf{p}_k - \mathbf{p}_m\|, \text{ for } k, m \in \mathcal{K}, \quad (12)$$

where α is a scaling factor. Note that the UE spatial location $\mathbf{P}$ is not known and BS location is not needed; CC is computed solely based on the dissimilarity matrix. A channel chart is constructed using an unsupervised ML framework that processes the dissimilarity matrix, and manifold learning is used to dimensionally reduce the CSI feature space [20]. For a given dissimilarity matrix, different dimension reduction techniques have been proposed in the literature. The performance of a given technique is problem dependent, as discussed in [27]. The single-cell CC problem has been solved using principle component analysis (PCA), Sammon's mapping (SM), and autoencoder reduction techniques in [20], whereas the MPCC is solved using SM, Laplacian Eigenmaps (LE), and t -Distributed Stochastic Neighbor Embedding (t -SNE) in [21]. Recently, neural networks have been used successfully for dimensionality reduction as in [28, 29].

LE is a computationally efficient non-linear dimensionality reduction technique based on the graph Laplacian. It preserves neighborhood properties and clustering connections [30]. LE constructs a graph from neighborhood information of the dissimilarity matrix. The LE problem is expressed as [30]:

$$\underset{\mathbf{Z}}{\text{minimize}} \text{Tr}(\mathbf{Z}^T \mathbf{L} \mathbf{Z}), \quad (13a)$$

$$\text{subject to } \mathbf{Z}^T \mathbf{S} \mathbf{Z} = \mathbf{I}_{d+1}, \quad (13b)$$

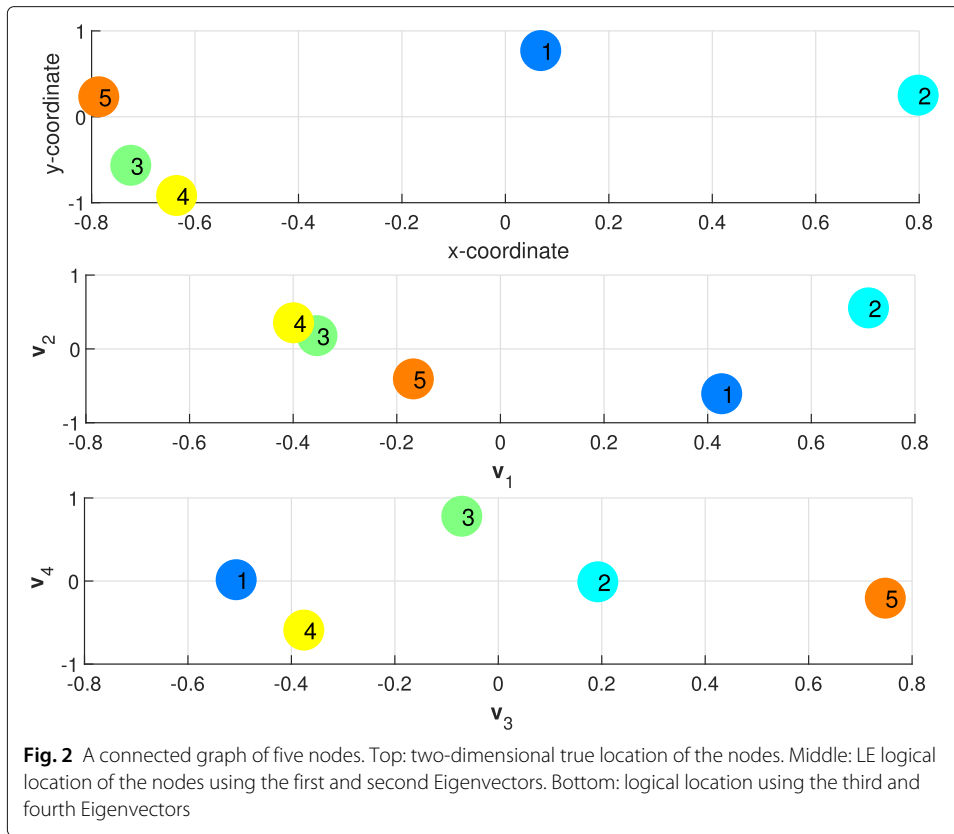
where $\mathbf{Z} = [\mathbf{z}_1^T, \dots, \mathbf{z}_K^T]^T$ represents the optimization variables (CC locations) in a matrix form, $\mathbf{I}_d$ is the identity matrix of order d , $\mathbf{L}$ is the graph Laplacian matrix, and $\mathbf{S}$ is the degree matrix. The graph Laplacian matrix is computed as:

$$\mathbf{L} = \mathbf{S} - \mathbf{W}, \quad (14)$$

where $\mathbf{W}$ is the weight matrix. The degree matrix $\mathbf{S}$ can be constructed using the dissimilarity matrix either by an ϵ -neighborhood, i.e., nodes k and m are connected by an edge if $[\mathbf{D}]_{k,m} \leq \epsilon$, or by N nearest neighbors, i.e., nodes k and m are connected by an edge if m is among the N nearest neighbors (N smallest dissimilarity values of the k th row of $\mathbf{D}$) of k or k is among the N nearest neighbors (N smallest dissimilarity values of the m th row of $\mathbf{D}$) of m . The weight matrix can be constructed using the dissimilarity matrix either by a simple approach, if nodes k and m are connected, $[\mathbf{W}]_{k,m} = 1$, otherwise $[\mathbf{W}]_{k,m} = 0$ or by using the heat kernel with temperature T , if nodes k and m are connected, $[\mathbf{W}]_{k,m} = e^{-\frac{[\mathbf{D}]_{k,m}^2}{T}}$, otherwise $[\mathbf{W}]_{k,m} = 0$. The temperature T can be selected based on the statistics of the dissimilarity matrix.

The Laplacian matrix is a symmetric positive-semidefinite matrix. Every row sum and column sum of $\mathbf{L}$ is zero, consequently $\lambda_0 = 0$ is the smallest Eigenvalue of $\mathbf{L}$, and $\mathbf{v}_0 = [1, \dots, 1]^T$ satisfies $\mathbf{L}\mathbf{v}_0 = \mathbf{0}$. In addition, the elements of an Eigenvector sum to zero, i.e., $\sum_{k=1}^K [\mathbf{v}_i]_k = 0$ for $i = 1, \dots, K-1$.

The solution of (13) can be obtained in closed form as the solution of a generalized Eigenvector problem based on KKT conditions [30]. The CC locations are obtained by finding the $d+1$ Eigenvectors corresponding to $d+1$ smallest Eigenvalues. An example of a connected graph of five nodes is shown in Fig. 2. The dissimilarity matrix is computed



using the true Euclidean distance, three nearest neighbors are used to compute the degree matrix, and the heat kernel temperature is set $T = 1$. The true location of the nodes is shown in the top subfigure. LE is used to find the logical location of the nodes. The second and third Eigenvectors preserve the local neighborhood information as shown in the middle subfigure, whereas the forth and fifth Eigenvectors maximize the difference between the nodes as shown in the bottom subfigure. The neighborhood information is not preserved using Eigenvectors corresponding to the largest Eigenvalues.

Algorithm 1 summarizes how the CC locations can be obtained using LE.

3.4 Out-of-sample extension

Since the MPCC is constructed by processing the data of all UEs from all BSs, it is computationally expensive to repeat the MPCC process if an out-of-sample data item is available, and needs to be inserted into the chart. If the original MPCC is based on a sufficient number of samples, it is expected that the out-of-sample data will not change the MPCC positions of the original samples.

Here, we address out-of-sample extension of MPCC in this sense, aiming to estimate the location of the new sample on the MPCC, to be used for RRM functions, such as V2V LQP. It is worth mentioning that the same feature extraction should be used for the out-of-sample data items as for the original samples, and the data-driven dissimilarity measure found for the original samples should be used to measure dissimilarity of the out-of-sample items to the original samples. For an out-of-sample UE j , the CSI covariance matrix $R_{b,j}$ at BS b is used to find the feature vector $\mathcal{F}\{f_{b,j}\} = [\mathbf{x}_{b,j}^{(1)}, \dots, \mathbf{x}_{b,j}^{(L_j)}]$. The cluster

Algorithm 1 The LE for MPCC

-
- 1: **Given:** the dissimilarity matrix $\mathbf{D}$, ϵ/N , T .
 - 2: **Construct:** the adjacency matrix, two approaches can be considered:
 - The ϵ -neighborhood, nodes k and m are connected by an edge if $[\mathbf{D}]_{k,m} \leq \epsilon$.
 - Nodes k and m are connected by an edge if m is among the N nearest neighbors of k or k is among the N nearest neighbors of m .
 - 3: **Choosing:** the weight matrix $\mathbf{W}$; two approaches can be considered:
 - Using the heat kernel with temperature T , which needs to be chosen based on the dissimilarity statistics; if nodes k and m are connected, $[\mathbf{W}]_{k,m} = e^{-\frac{[\mathbf{D}]_{k,m}}{T}}$, otherwise $[\mathbf{W}]_{k,m} = 0$.
 - Simple approach, if nodes k and m are connected, $[\mathbf{W}]_{k,m} = 1$, otherwise $[\mathbf{W}]_{k,m} = 0$.
 - 4: **Compute:** the Laplacian Matrix $\mathbf{L} = \mathbf{S} - \mathbf{W}$, where $\mathbf{S}$ is the degree matrix (diagonal matrix) with $[\mathbf{S}]_{k,k} = \sum_{i=1}^K [\mathbf{W}]_{k,i}$.
 - 5: **Compute:** the eigenvalues λ_i for $i = 0, \dots, K-1$ and eigenvectors $\mathbf{v}_i$ for $i = 0, \dots, K-1$ for the generalized eigenvector problem: $\mathbf{L}\mathbf{v} = \lambda\mathbf{S}\mathbf{v}$,
 - 6: **Order:** the eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{K-1}$ according to their eigenvalues, with $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{K-1}$.
 - 7: **Return:** the CC position of the k th UE on the MPCC as: $\mathbf{z}(k) = [\mathbf{v}_1(k), \mathbf{v}_2(k)]$ for $d = 2$ and $\mathbf{z}(k) = [\mathbf{v}_1(k), \mathbf{v}_2(k), \mathbf{v}_3(k)]$ for $d = 3$.
-

label for an out-of-sample multipath component is determined based on the cluster label of the nearest multipath component on the original data set, i.e., $\mathcal{L}(\mathbf{x}_{b,j}^{(l)}) = \mathcal{L}(\mathbf{x}_{b,m}^{(l')})$ where $[m, l'] = \underset{k,n}{\operatorname{argmin}} \|\mathbf{x}_{b,k}^{(n)} - \mathbf{x}_{b,j}^{(l)}\|_2$, $k \in \mathcal{K}_b$, and $n = 1, \dots, L_k$. The out-of-sample dissimilarity element $[\mathbf{D}_b]_{j,m}$ at BS b is computed using (10), and then, the global dissimilarity is computed using (11). The relation between MPCC and EMPCC is shown in Fig. 3.

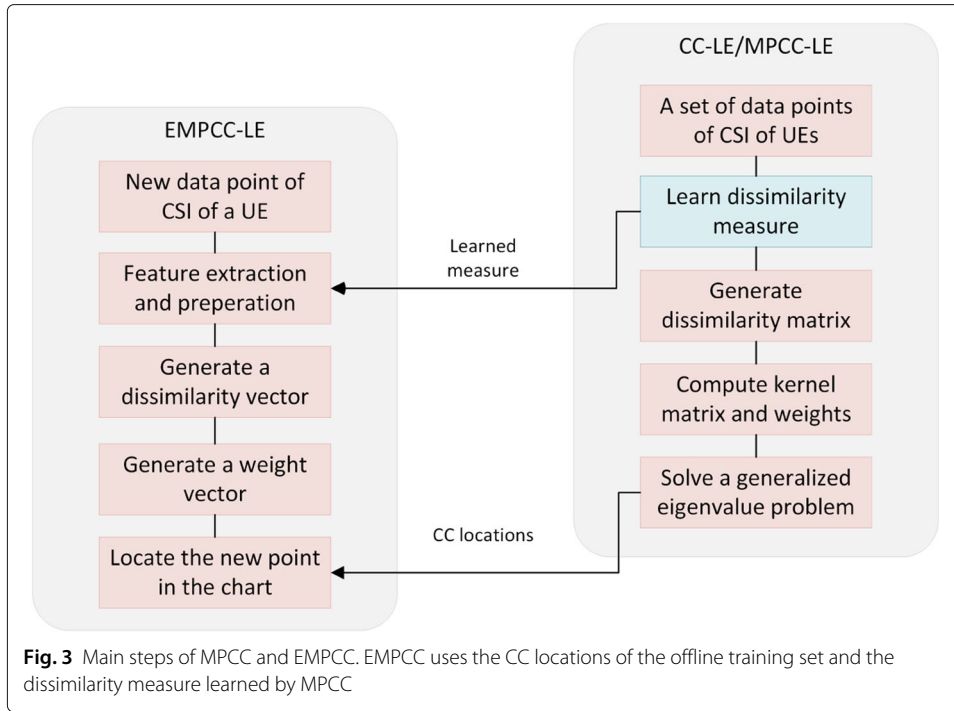
In [31], a generalized framework for out-of-sample extension is proposed for several algorithms, providing that these algorithms learn Eigenfunctions of a data-dependent kernel. The out-of-sample mapping can be formulated as an optimization problem, where the objective is to find a normalized kernel function that minimizes the mean squared error. The normalized kernel vector is used as a weight vector to find the out-of-sample mapping. For LE, the normalized kernel function (weight) is computed as [31]:

$$\hat{W}(k, i) = \frac{1}{K} \frac{W(k, i)}{\sqrt{\mathbb{E}_x[W(k, x)] \mathbb{E}_y[W(i, y)]}}, \quad k, i = 1, \dots, K, \quad (15)$$

where $W(k, i) = [\mathbf{W}]_{k,i}$ and the expectation is taking with respect to the original data set. The EMPCC position of an out-of-sample data $\mathbf{z}(j)$ for $j \notin \{1, \dots, K\}$, $i = 1, \dots, K$, and $d = 2$ can be computed as:

$$\mathbf{z}(j) = \left[\sum_{k=1}^K \hat{W}(j, k) \hat{\mathbf{v}}_1(k), \sum_{k=1}^K \hat{W}(j, k) \hat{\mathbf{v}}_2(k) \right], \quad (16)$$

where the weight $\hat{W}(j, i)$ for $j \notin \{1, \dots, K\}$ is computed based on the dissimilarity of the radio features of UE j with respect to the radio features of all UEs in the original set, and



the Eigenvectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ are computed based on the normalized weighting matrix $\hat{\mathbf{W}}$ of the original data set.

The resulting EMPCC method is summarized in Algorithm 2.

4 MPCC-based V2V link quality prediction

Radio maps can be utilized for RRM functionalities. To construct radio maps, either the physical or the logical location of the UEs in the radio environment and the corresponding CSIs are needed. The physical location can be obtained either by a global navigation satellite system (GNSS) such as GPS or by a triangulation approach. Triangulation can be used for only LOS communications with at least three BSs. The locations of the BSs need to be known, whereas CC has the advantage of being able to be used for both LOS and NLOS communications without the need to know the BS locations. CC can be used with a single BS; however, using more BSs improves the CC accuracy. CC has the advantage

Algorithm 2 The EMPCC for UE $j, j \notin \{1, \dots, K\}$

- 1: **Given** the feature vector $\{\mathcal{F}\{\mathbf{f}_{b,k}\}\}$, cluster label $\{\mathcal{L}(\mathbf{x}_{b,k}^{(l)})\}$, weighting matrix $\hat{\mathbf{W}} \in \mathbb{R}^{K \times K}$ and the eigenvectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$.
 - 2: **Estimate** the multipath components $\{\mathbf{f}_b^{(j)}\}_{b=1}^B$.
 - 3: **Compute** the feature vector $\mathcal{F}\{\mathbf{f}_{b,j}\}$.
 - 4: **Find** the cluster label $\mathcal{L}(\mathbf{x}_{b,j}^{(l)})$ for each multipath components.
 - 5: **Compute** the dissimilarity coefficient $[D_b]_{j,m}$ for $m = 1, \dots, K_b$.
 - 6: **Compute** the dissimilarity fusion vector $\{[D]_{j,m}\}_{m=1}^K$ for out-of-sample UE j .
 - 7: **Compute** the weight vector $[\hat{\mathbf{W}}(j, 1), \dots, \hat{\mathbf{W}}(j, K)]$.
 - 8: **Map** the position $\mathbf{z}(j)$ on the MPCC using (16).
-

of replacing the timely and costly measurement campaign in GNSS fingerprinting-based algorithms by heavily processing ML algorithms (i.e., unsupervised learning plays a key role of mapping radio features to logical locations and preserving neighborhood relations) at the BSs, which has the advantage of being able to be applied for large-scale areas and in an automated manner when the radio environment changes. The back-haul cost of CC is less than the back-haul of GNSS fingerprinting, since the location information is not transmitted. Table 1 compares CC-based radio maps with GNSS-based fingerprinting and triangulation-based fingerprinting in terms of communication scenario, BS location, back-haul load, and computational cost at UEs and BSs.

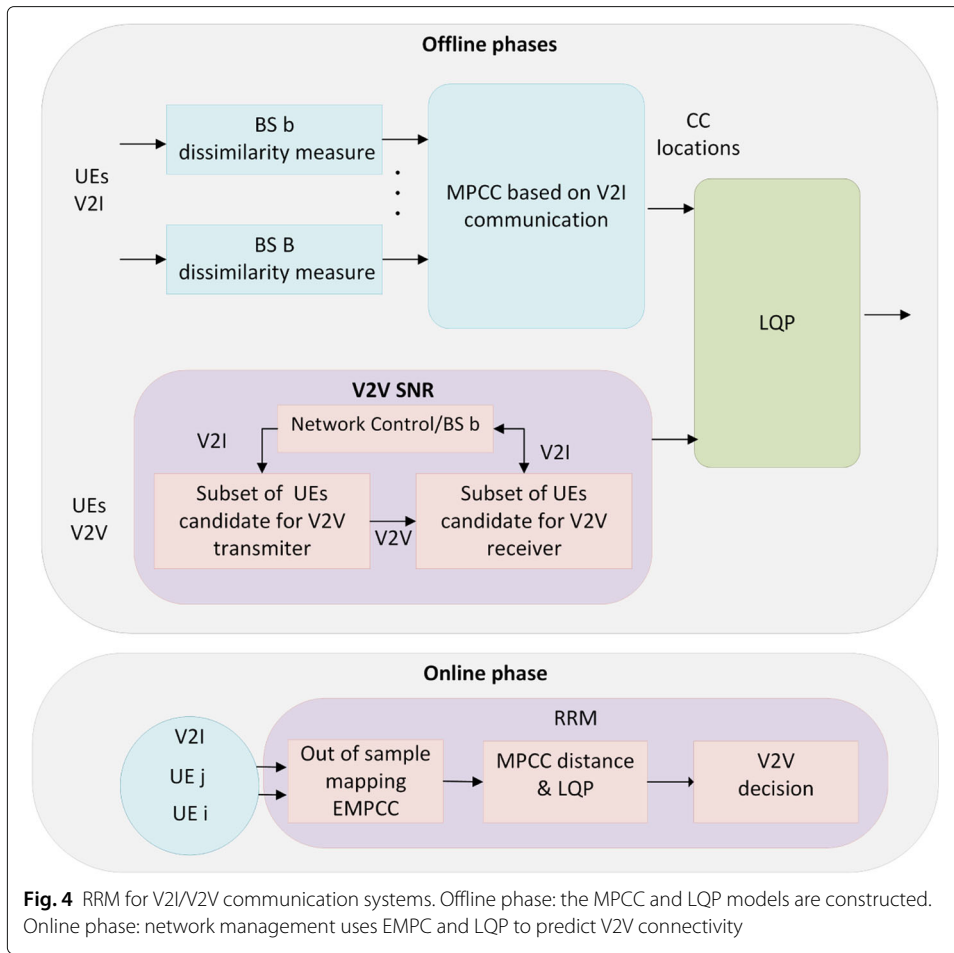
We consider V2I/V2V RRM based on large-scale radio features, i.e., the covariance matrices. A large data set of radio features of V2I is processed to obtain a channel chart of logical locations. In the training phase, the network control unit selects pairs of UEs that have the capability for V2V communications, and asks them to establish connection and measure the link quality. The vehicular terminals then feedback the average SNR of V2V communication to the network. The control unit constructs a LQP model based on the knowledge of CC locations of the vehicular terminals and the received average SNR of V2V pairs. The RRM framework consists of an offline training phase where MPCC and LQP are generated and an online phase where the MPCC and LQP are used to predict connectivity of UEs in the network. In the online phase, out-of-sample extension of MPCC is used to place vehicles to the MPCC, and the LQP model is used to predict V2V connectivity. The block diagram of the considered method to predict V2V connectivity is shown in Fig. 4.

4.1 Link quality prediction model

In wireless communications, the optimal transmission scheme is adaptively selected based on the estimated CSI. Due to the high-mobility nature of V2V, directivity, and blockage of mm-Wave bands, link quality prediction of V2V is a challenging problem. Generally, analytical and theoretical models for LQP are based on simplified bounding assumptions, which cannot be used in practical scenarios. Here, we consider a data-driven probabilistic LQP model, utilizing the MPCC locations and average SNR of a large set of V2V pairs. The LQP of V2V communications is determined by the average SNR at the receiving terminal. The most important characteristic of a V2V channel is whether there is a connection or not. To proceed with predicting connectivity, we assume that there is an SNR threshold for successful reception. Knowing the SNR statistics for V2V

Table 1 Benefits and costs of CC-based RRM

Property	MPCC/CC	GNSS-based fingerprinting	Triangulation-based fingerprinting
Scenario	NLOS/LOS	Signals from three satellites	LOS from three BSs
BSs location	Not needed	Not needed	Needed
Back-haul load	CSIs	CSIs and location coordinates	CSIs
Comp. cost at BSs	High (population-based CSI processing)	Not applicable	Low (point-based CSI processing)
Comp. cost at UEs	Not applicable	GNSS position calculation	Not applicable



communication with a given MPCC distance, one may then predict the probability of the V2V link being in outage with respect to this SNR threshold.

The channel charting distance $d_{ij}^{(C)}$ between UE i and UE j is defined using the Euclidean distance of MPCC locations $\mathbf{z}_i$ and $\mathbf{z}_j$ as:

$$d_{ij}^{(C)} = |\mathbf{z}_i - \mathbf{z}_j|_2. \quad (17)$$

The MPCC distance $d_{ij}^{(C)}$ of the V2V pairs is quantized into a grid with G points, $\mathcal{D} = \{d_0^C, \dots, d_{G-1}^C\}$, such that $d_{ij}^{(C)}$ is assigned to grid point g if $d_{g-1}^C \leq d_{ij}^{(C)} < d_g^C$. The outage probability for CC grid distance d_g^C can then be estimated as:

$$\mathcal{O}(\gamma_{th}|d_g^{(C)}) = \Pr(\Gamma \leq \gamma_{th}|d_g^C), \quad (18)$$

where γ_{th} is an SNR threshold determined for reliable communication at a rate required by the network, and Γ is the average SNR of a V2V communication pair belonging to the sample set with MPCC distance quantized to d_g^C . The outage probability for distances $d_g \in \mathcal{D}$ is empirically computed using the measured SNR of V2V UEs.

5 Simulation results and discussion

A multicell mm-Wave scenario is considered as discussed in [21]. The simulation parameters are shown in Table 2. The UE locations are generated on the streets of a Manhattan

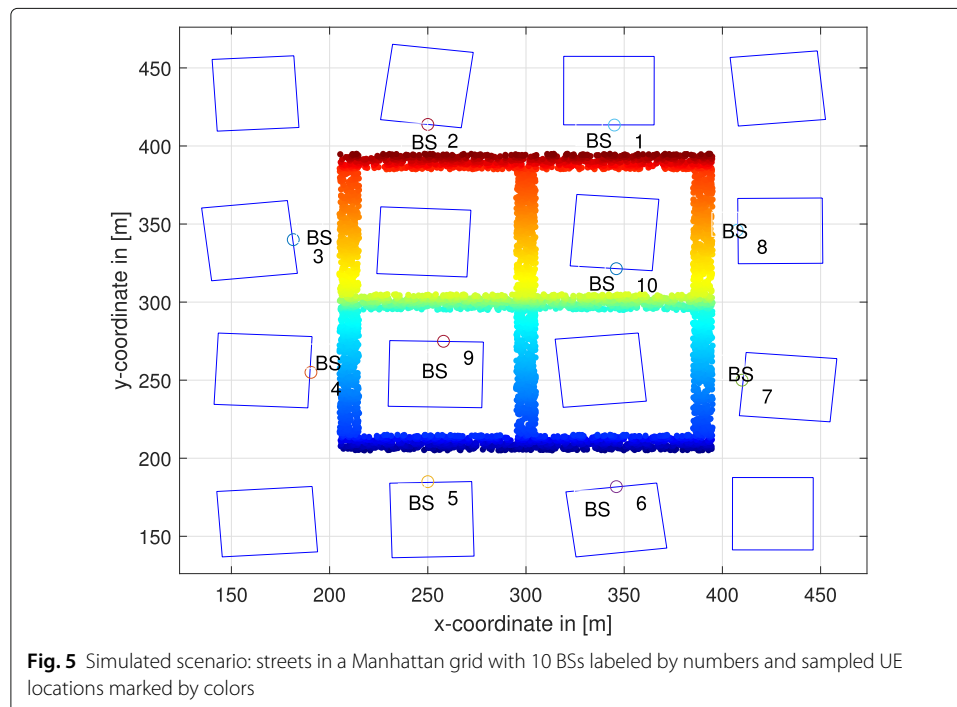
Table 2 Simulation parameters [21]

Parameter	Value	Parameter	Value
Pathloss at 1 m	61.4 dB	Reflection loss	0–15.5 dB
UE tx power	23 dBm	BS noise power	−86 dBm
BS antenna gain	0 or 2	UE antenna gain	1
Bandwidth	200 MHz	OFDM subcarriers	256
BS antenna	64 ULA	V2V antenna	8 ULA
BS array gain	18 dB	V2V array gain	8 dB
Noise figure	6 dB	Noise power	−174 dBm
No. of subrays/cluster	5	Max. no. of bounces	5
Max. no. of multipaths	10	Intra-cluster mean delay	10 ns

grid as shown in Fig. 5. In [21], a ray-tracing mm-Wave cellular channel model was created following the principles of [32, 33]. Here, we use this channel model for V2I and further generalize it to a V2V model. The channel simulator models the path loss experienced by the multipath components using the free-space path loss model with power inversely proportional to the square of the distance. The reflections from obstacles, i.e., the walls, are modeled such that the reflection coefficients are based on Fresnel's equations. The typical value for the wall relative permittivity is between 4 and 6. The channel for each link is then calculated using the ray-traced paths with the path loss, reflection losses, and antenna gain accounted for in the channel. The multipath gain $\beta_{b,k}^{(l)}$ is computed as:

$$\beta_{b,k}^{(l)} = e^{j\psi_l} \sqrt{G_0 \rho d_l^{-2} g_1(\theta_l) g_2(\phi_l) \prod_{i=1}^R |r_l^{(i)}|^2}, \quad (19)$$

where $G_0 = 10^{-6.14}$ is the omnidirectional path gain at a reference distance of 1 m; ρ is the transmit power; ψ_l is the phase modeled as a uniform random variable $\psi_l \sim \mathcal{U}(0, 2\pi)$;

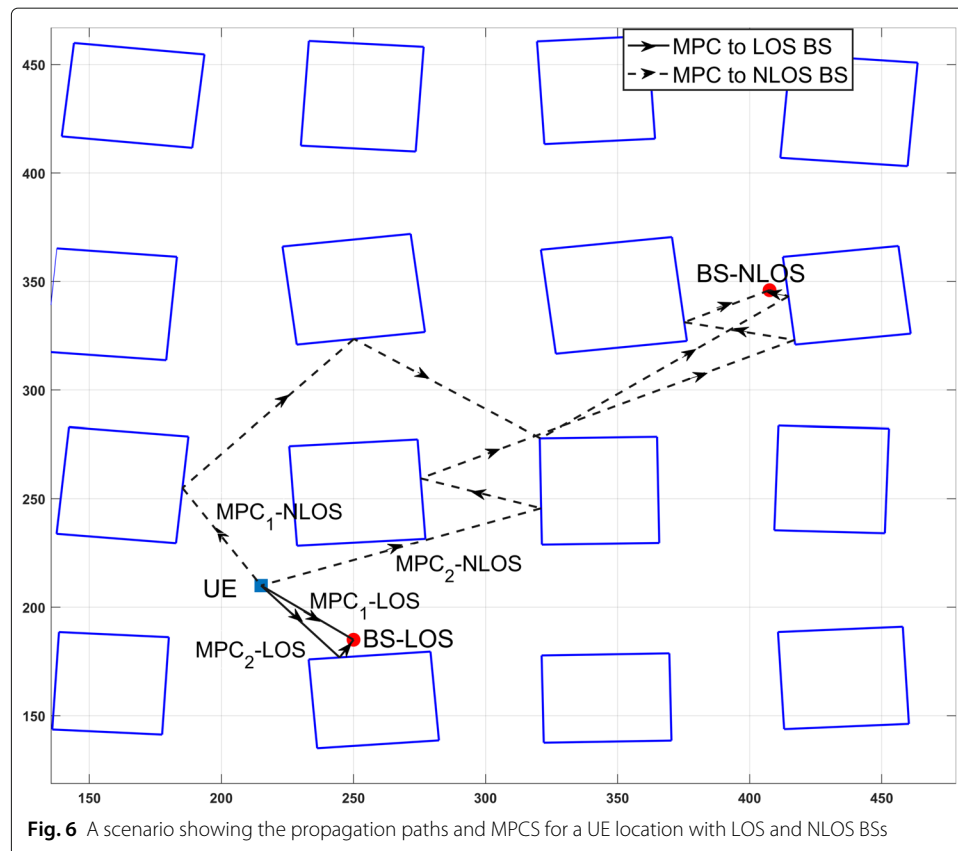


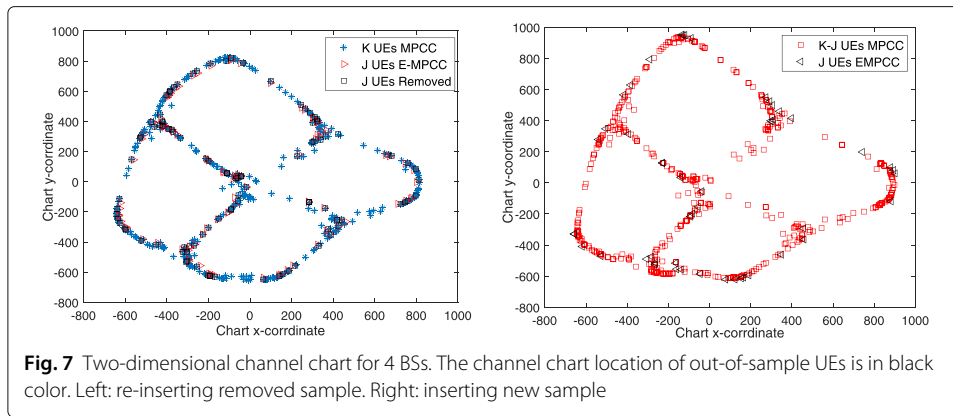
d_l is the propagation distance in meters; $g_1(\theta_l)$ and $g_2(\phi_l)$ are the antenna gain for an angle of departure θ_l at the UE and angle of arrival ϕ_l at the BS, respectively; R is the number of reflections that the l th multipath component undergoes; and $r_l^{(i)}$ is the i th reflection coefficient. For an LOS path, $R = 1$ and $r_l^{(1)} = 1$.

A scenario showing the propagation paths for multipath components using the ray-tracing model is shown in Fig. 6. A UE location has LOS communication with one BS ($BS-LOS$) and a NLOS communication with another BS ($BS-NLOS$). The SNR observed at $BS-LOS$ which is at a distance of 43.01 m is obtained as 38 dB. The SNR at $BS-NLOS$ which is at a distance of 235.7 m is calculated as -36.83 dB.

5.1 Performance of out-of-sample extension algorithm

First, we investigate the performance of EMPCC, which inserts out-of-sample UEs to the chart. There are K UEs, and the number of neighboring UEs used to construct the graph for LE is denoted by N . The number of UEs for which EMPCC is used is denoted by J . Two scenarios are considered to evaluate the performance of EMPCC. In scenario I, the MPCC is generated based on the channel features of K UE locations. Then, J UE locations are removed at random, and EMPCC is used for mapping the J locations to the chart. In scenario II, J UE locations are selected at random and the MPCC is generated based on the channel features of $K - J$ UE locations. EMPCC is used for mapping the J locations to the chart. Both Laplacian Eigenmaps based on a conflict graph and LE based on a weighted graph are used for channel charting.





An example instance for LE-based MPCC/EMPCC for different parameters is shown in Figs. 7 and 8. For Fig. 7, the parameters are $K = 500$, $J = 100$, $N = 25$, and $B = 4$ BSs labeled as $\{1, 3, 5, 7\}$. We select a reference point in the first quadrant for K and $K - J$ MPCC to avoid the possibility of rotation or flipping of the EMPCC compared to MPCC.

In Fig. 8, the parameters are $K = 5000$, $N = 250$, $J = 500$, and $B = 10$. The J out-of-sample locations are accurately mapped by EMPCC.

The performance of MPCC/EMPCC is evaluated using continuity (CT) and trustworthiness (TW) measures as shown in Table 3. For a discussion on these measures, see [34]. CT and TW are computed by considering 50 nearest neighbors. For MPCC, all K UEs are used to generate the chart, whereas for EMPCC, the chart is constructed by $K - J$ UEs and the EMPCC is used to position the remaining J UEs. For weighted Gaussian kernel, $T = 0.05$. The CT and TW and measures of EMPCC are comparable to MPCC, indicating that the out-of-sample extension methodology in EMPCC works.

5.2 Performance of link quality prediction

For V2V link quality prediction, the MPCC is constructed based on V2I communications. For this, we consider a scenario with $K = 5000$ UEs and $B = 10$ BSs, in the Manhattan grid considered above. The LQP model is constructed based on the SNR of V2V pairs with the corresponding Euclidean charting distance computed using the MPCC locations. To construct the V2V channels, 1, 000, 000 random pairs of UEs are selected among

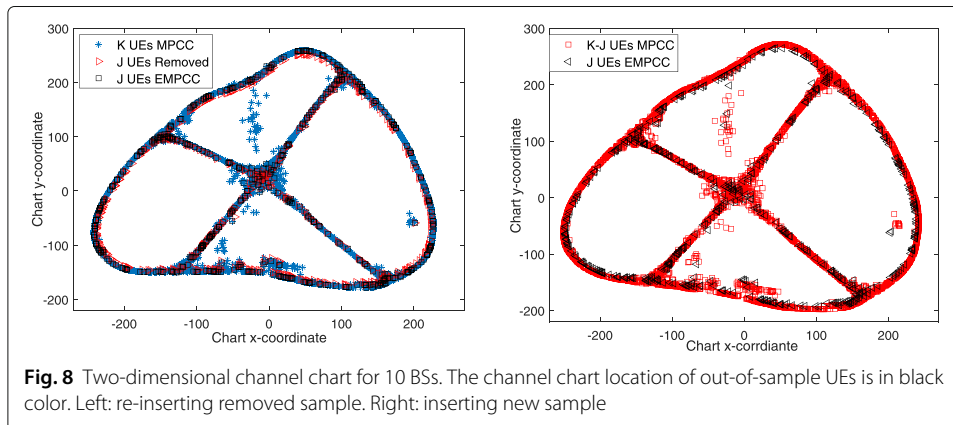


Table 3 Comparison of MPCC and EMPCC in terms of TW and CT measures considering 50 neighbors, weighted LE (w-LE) with $T = 0.05$ is considered

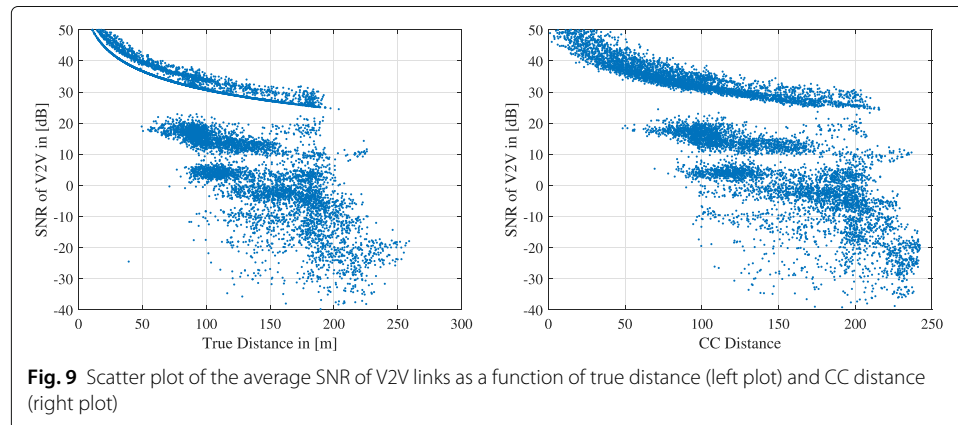
	LE-MPCC		LE-EMPCC		wLE-MPCC		wLE-EMPCC	
	CT	TW	CT	TW	CT	TW	CT	TW
4 BSs	0.9880	0.9883	0.9868	0.9802	0.9777	0.9776	0.9777	0.9776
10 BSs	0.9943	0.9806	0.9788	0.9942	0.9955	0.9954	0.9955	0.9954

the chart locations. The V2V mm-Wave channels are generated by generalizing the ray-tracing channel model of [21] in the same environment where the MPCC is constructed, and the average SNRs for V2V communications are computed as in (6).

Figure 9 shows a scatter plot of the average SNR of the V2V pairs as function of physical and chart distances. As expected, the SNR of a V2V link decreases with increasing physical distance, and the relation of SNR with chart distance also captures this. This figure indicates that MPCC preserves the distance-SNR relation. It can be seen from Fig. 9 that at smaller distances, when charting distance $d^{(C)} < 75$, the probability that an average SNR of a V2V link is below a SNR threshold of $\gamma_{th} = 25$ dB is zero, so for this charting distance, V2V communication is guaranteed to be successful with high data rates, or the transmitted power can be reduced to reduce the interference to other terminals.

Using the collected data of the average SNRs and the physical distances, a benchmark LQP model is constructed. The outage probabilities $\mathcal{O}_P(\gamma_{th}|d_g^{(P)})$ are empirically computed for the true location UEs, for different SNR thresholds γ_{th} using a grid distance $d_g^{(P)}$. Using the collected data of the average SNRs and the corresponding CC distances, a LQP model is constructed. The outage probabilities $\mathcal{O}_{CC}(\gamma_{th}|d_g^{(C)})$ of (18) are empirically computed for the chart UEs, for different SNR thresholds γ_{th} . The trained outage probability model for different physical and CC distances and different thresholds γ_{th} is shown in Fig. 10. The left plot represents the benchmark LQP model that can be used to predict the outage probability of an out-of-sample V2V pair by knowing the true distance. The right plot represents the LQP model that can be used to predict the outage probability of an out-of-sample V2V pair by just knowing the EMPCC (out-of-sample chart) distance between them. The CC LQP relation as a function of the CC distance is similar to the benchmark LQP as a function of true distance.

To estimate the performance of LQP in the online RRM phase, a test set of $J = 1000$ out-of-sample V2I UEs was generated. The large-scale radio features of V2I channels are



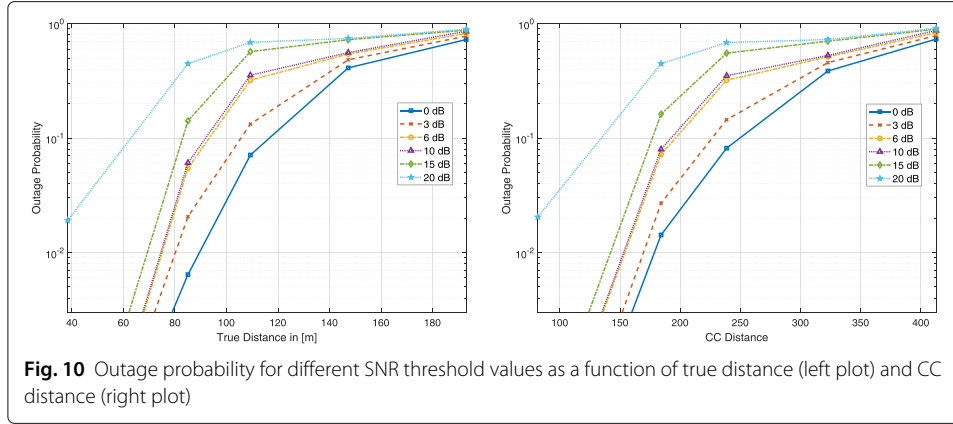


Fig. 10 Outage probability for different SNR threshold values as a function of true distance (left plot) and CC distance (right plot)

used to map these out-of-sample UEs to the existing chart using the EMPCC algorithm. Again, 1,000,000 V2V pairs are constructed at random from these out-of-sample UEs. The V2V mm-Wave channels and the V2V SNRs are generated in the same way as that for the chart UEs.

The true outage probabilities as the function of physical and chart distances are then constructed for this test set. As a result, we get the outage probabilities of out-of-sample UEs as $\mathcal{O}_{OS}(\gamma_{th}|d_g^{(P)})$ and $\mathcal{O}_{OS}(\gamma_{th}|d_g^{(C)})$, respectively. Note that for comparison to the LQP model, the same quantization grid $d_g^{(C)}$ and $d_g^{(P)}$ are used for the physical and chart distances of the test set, as for the original trained UEs, respectively.

The true outage probabilities can be compared to the ones predicted by the trained LQP. The relative mean square error for LQP of the outage probability from the data of chart UEs for a given γ_{th} at chart distance $d_g^{(C)}$ is given by:

$$\delta_{C,g}^2 = \frac{\left(\mathcal{O}_{OS}(\gamma_{th}|d_g^{(C)}) - \mathcal{O}_{CC}(\gamma_{th}|d_g^{(C)})\right)^2}{\left(\mathcal{O}_{CC}(\gamma_{th}|d_g^{(C)})\right)^2}. \quad (20)$$

Similarly, the relative mean square error for the benchmark LQP of the outage probability from the data of UEs for a given γ_{th} at physical distance $d_g^{(P)}$ is denoted as $\delta_{P,g}^2$.

Figure 11 shows the error $\delta_{C,g}^2$ for different SNR thresholds as a function of the CC distance and the error $\delta_{P,g}^2$ as a function of the physical distance. The largest relative

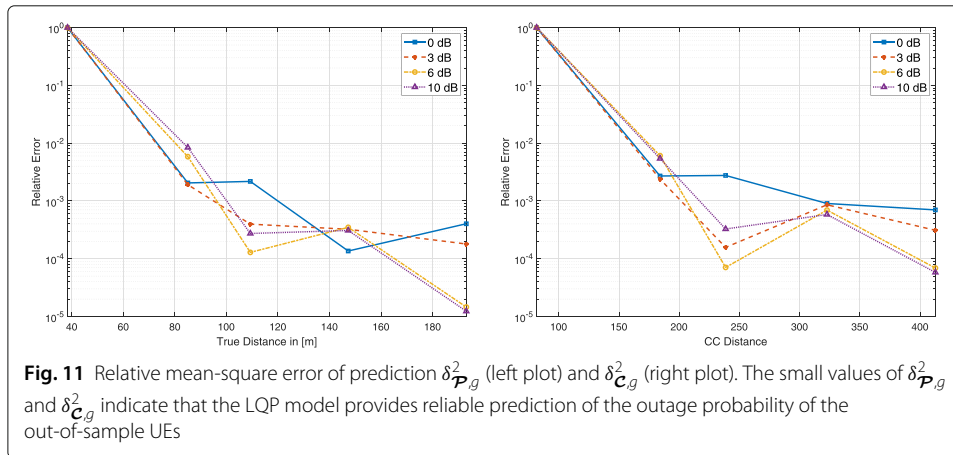


Fig. 11 Relative mean-square error of prediction $\delta_{P,g}^2$ (left plot) and $\delta_{C,g}^2$ (right plot). The small values of $\delta_{P,g}^2$ and $\delta_{C,g}^2$ indicate that the LQP model provides reliable prediction of the outage probability of the out-of-sample UEs

mean square error $\delta_{C,g}^2 = 3.0\%$ is observed for threshold $\gamma_{th} = 6$ dB, for a CC distance larger than 150. This indicates that the trained LQP model provides reliable prediction of the outage probability of the out-of-sample UEs just based on CSI of the V2I links. The relative error $\delta_{P,g}$ based on the true distance is smaller than the relative error $\delta_{C,g}$ based on the chart distance.

6 Conclusion

We have presented the concept of link quality prediction for V2V communications in dynamic environments based on multipoint channel charting. For this, the physical locations of neither the vehicles nor the base stations are required. We have considered a network controlled V2V approach, where vehicles communicate with infrastructure BSs, and the large-scale radio frequency features of the V2I channels have been used to map vehicles to a logical map. A network control unit has been used to manage the selection and collection of enough SNR samples of V2V channels and to construct a LQP model. In order to use the prediction in online RRM, the channel charting principle has to be extended to out-of-sample data CSI features, related to out-of-sample vehicle locations. For this, a MPCC has been constructed first using an original data set of V2I CSIs. The multipath components of the new CSI samples have been estimated at each BS and then processed using the data-driven dissimilarity computation as the original set. The dissimilarity vector of the out-of-sample vehicle has been used to generate the weighting vector for out-of-sample mapping. The resulting EMPCC algorithm has been used to map out-of-sample vehicles to the chart. The trustworthiness and continuity performance measures have been used to evaluate the EMPCC, and we found that out-of-sample extension works in a reliable manner.

The method has wide applicability in cognitive RRM, where predictions of vehicle connectivity parameters would be used. Here, we have used the channel chart to predict V2V connectivity. Based on the Euclidean chart distance, the probability of outage of V2V communication between two out-of-sample vehicles has been predicted. This can be used by the network to identify which vehicles may communicate over direct V2V links. The only input for this prediction is the V2I CSI of the two involved vehicles, as measured by the infrastructure base stations. In simulation modeling of a mm-Wave network, the LQP was found to perform well, with a typical relative mean square error of $< 2\%$.

In future work, the locations of V2V pair, not only the chart distance, are going to be used to improve the LQP model. An advanced LQP model based on deep learning will also be considered to predict the SNR of the link given the CC locations of the V2V pairs. Using channel charting for multihop V2V communication is another RRM problem that can be considered, i.e., selecting the relaying nodes for V2V communication to achieve a desired link quality. Advanced mm-Wave channel models, in which the blockage probability, density, size, and speed of vehicles are taken into consideration, are important components when verifying channel charting-based RRM in such challenging scenarios.

Abbreviations

AI: Artificial intelligence; B5G: Beyond 5G; BS: Base station; CC: Channel charting; CSI: Channel state information; CT: Continuity; DSRC: Dedicated short-range communication; EMPCC: Extension-of-MPCC; LE: Laplacian Eigenmaps; LQP: Link quality prediction; MAC: Medium access control; ML: Machine learning; mMIMO: Massive MIMO; mm-Wave: Millimeter-wave; MPCC: Multipoint channel charting; PHY: Physical; PCA: Principal component analysis; QoS: Quality of

service; RRM: Radio resource management; Rx: Receive; SM: Sammon's mapping; SNR: Signal-to-noise ratio; t-SNE: t-Distributed Stochastic Neighbor Embedding; Tx: Transmit; TW: Trustworthiness; UEs: User equipments; V2I: Vehicle-to-infrastructure; V2V: Vehicle-to-vehicle; w-LE: Weighted-LE

Acknowledgements

Not applicable.

Authors' contributions

OT proposed the idea and revised this paper. HA and TP wrote the manuscript and participated in the simulation. CS gave some suggestions and participated in the paper revision. All authors have contributed to this research work. All authors have read and approved the final manuscript.

Funding

This work was funded in part by the Academy of Finland (grant 319484). The work of C. Studer was supported in part by Xilinx Inc. and by the US NSF under grants ECCS-1408006, CCF-1535897, CCF-1652065, CNS-1717559, and ECCS-1824379.

Availability of data and materials

Data sharing not applicable to this article as no data sets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Communications and Networking, Aalto University, Espoo, Finland. ²School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA.

Received: 2 January 2020 Accepted: 23 April 2020

Published online: 24 June 2020

References

1. V. Va, T. Shimizu, G. Bansal, R. Heath, *Millimeter Wave Vehicular Communications: A Survey*. (Now Publishers Inc., Hanover, 2016)
2. J. Kenney, Dedicated short-range communications (DSRC) standards in the United States. *Proc. IEEE*. **99**(7), 1162–1182 (2011)
3. M. Giordani, A. Zanella, M. Zorzi, in *Proc. of the 6th International Conference on Modern Circuits and Systems Technologies, (MOCAST)*, Millimeter wave communication in vehicular networks: challenges and opportunities, (2017), pp. 1–6. <https://doi.org/10.1109/mocast.2017.7937682>
4. J. Choi, V. Va, N. Gonzalez-Precic, R. Daniels, C. Bhat, R. Heath, Millimeter-Wave vehicular communication to support massive automotive sensing. *IEEE Commun. Mag.* **54**(12), 160–167 (2016)
5. S. Busari, K. Huq, S. Mumtaz, L. Dai, J. Rodriguez, Millimeter-Wave massive MIMO communication for future wireless systems: a survey. *IEEE Commun. Surv. Tuts.* **20**(2), 836–869 (2018)
6. E. Bjornson, E. G. Larsson, T. Marzetta, Massive MIMO: ten myths and one critical question. *IEEE Trans. Commun.* **54**(2), 114–123 (2016)
7. S. Yang, L. Hanzo, Fifty years of MIMO detection: the road to large-scale MIMOs. *IEEE Commun. Surv. Tuts.* **17**(4), 1941–1988 (2015)
8. A. Shahmansoori, G. Garcia, G. Destino, G. Seco-Granados, H. Wymeersch, Position and orientation estimation through Millimeter-Wave MIMO in 5G systems. *IEEE Trans. Wirel. Commun.* **17**(3), 1822–1835 (2018)
9. F. Guidi, A. Guerra, D. Dardari, A. Clemente, R. Errico, in *Proc. of IEEE Globecom Workshops, (GC Wkshps)*, Environment mapping with Millimeter-Wave massive arrays: system design and performance, (2016), pp. 1–6. <https://doi.org/10.1109/glocomw.2016.7848895>
10. N. Garcia, H. Wymeersch, E. Larsson, A. Haimovich, M. Coulon, Direct localization for massive MIMO. *IEEE Trans. Signal Process.* **65**(10), 2475–2487 (2017)
11. A. Kato, K. Sato, M. Fujise, ITS wireless transmission technology. technologies of Millimeter-Wave inter-vehicle communications: Propagation characteristics. *J. Commun. Res. Lab.* **48**(4), 99–110 (2001)
12. M. Giordani, M. Polese, A. Roy, D. Castor, M. Zorzi, A tutorial on beam management for 3GPP NR at mmWave frequencies. *IEEE Commun. Surv. Tuts.* **21**(1), 173–196 (2019)
13. M. Giordani, T. Shimizu, A. Zanella, T. Higuchi, O. Altintas, M. Zorzi, Path loss models for V2V mmWave communication: performance evaluation and open challenges. *CoRR* (2019). <https://doi.org/10.1109/cavs.2019.8887792>
14. M. Giordani, M. Mezzavilla, A. Dhananjay, S. Rangan, M. Zorzi, in *Proc. of the 22th European Wireless Conference, Channel dynamics and SNR tracking in millimeter wave cellular systems*, (2016), pp. 1–8
15. L. Liang, H. Ye, G. Y. Li, Toward intelligent vehicular networks: a machine learning framework. *IEEE Internet Things J.* **6**(1), 124–135 (2019)
16. M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, F. Kojima, Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks. *IEEE Access.* **6**, 32328–32338 (2018)
17. D. Gündüz, P. de Kerret, N. Sidiropoulos, D. Gesbert, C. Murthy, M. Schaar, Machine learning in the air. *CoRR*. **abs/1904.12385** (2019)
18. C. Zhang, P. Patras, H. Haddadi, Deep learning in mobile and wireless networking: a survey. *IEEE Commun. Surv. Tuts.*, 1 (2019). <https://doi.org/10.1109/comst.2019.2904897>
19. Z. Jiang, S. Chen, A. Molisch, R. Vannithamby, S. Zhou, Z. Niu, Exploiting wireless channel state information structures beyond linear correlations: a deep learning approach. *IEEE Commun. Mag.* **57**(3), 28–34 (2019)

20. C. Studer, S. Medjkouh, E. Gonultaş, T. Goldstein, O. Tirkkonen, Channel charting: locating users within the radio environment using channel state information. *IEEE Access*. **6**, 47682–47698 (2018)
21. J. Deng, S. Medjkouh, N. Malm, O. Tirkkonen, C. Studer, in *Proc. of 52nd Asilomar Conference on Signals, Systems, and Computers*, Multipoint channel charting for wireless networks, (2018), pp. 286–290. <https://doi.org/10.1109/acssc.2018.8645281>
22. T. Ponnada, H. Al-Tous, O. Tirkkonen, C. Studer, in *Proc. of 14th EAI International Conference on Cognitive Radio Oriented Wireless Networks, (Crowncom)*, An out-of-sample extension for wireless multipoint channel charting, (2019). https://doi.org/10.1007/978-3-030-25748-4_16
23. M. Akdeniz, Y. Liu, M. Samimi, S. Sun, S. Rangan, T. Rappaport, E. Erkip, Millimeter wave channel modeling and cellular capacity evaluation. *IEEE J. Sel. Areas Commun.* **32**(6), 1164–1179 (2014)
24. A. Goldsmith, *Wireless Communications*. (Cambridge University Press, New York, NY, USA, 2005)
25. R. Schmidt, Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **34**(3), 276–280 (1986)
26. M. Ester, H. Kriegel, J. Sander, X. Xu, in *Proc. of the Second International Conference on Knowledge Discovery and Data Mining, (KDD)*, A density-based algorithm for discovering clusters in large spatial databases with noise, (1996), pp. 226–231
27. L. Maaten, E. Postma, H. Herik, Dimensionality reduction: a comparative review. *J Mach Learn Res.* **10**, 66–71 (2009)
28. P. Huang, O. Castaneda, E. Gonultaş, S. Medjkouh, O. Tirkkonen, T. Goldstein, C. Studer, in *Proc. of the IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications, (SPAWC)*, Improving channel charting with representation-constrained Autoencoders, (2019), pp. 1–5. <https://doi.org/10.1109/spawc.2019.8815478>
29. E. Lei, O. Castaneda, O. Tirkkonen, T. Goldstein, C. Studer, Siamese neural networks for wireless positioning and channel charting. *ArXiv*. **abs/1909.13355** (2019). <https://doi.org/10.1109/allerton.2019.8919897>
30. M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**(6), 1373–1396 (2003)
31. Y. Bengio, J. Paiement, P. Vincent, O. Delalleau, N. L. Roux, M. Ouimet, in *Proc. of the 16th International Conference on Neural Information Processing Systems, (NIPS)*, Out-of-sample extensions for LLE, Isomap, MDS, eigenmaps, and spectral clustering (MIT Press, Cambridge, MA, USA, 2003), pp. 177–184
32. S. Hur, S. Baek, B. Kim, J. Park, A. F. Molisch, K. Haneda, M. Peter, in *Proc. of the 9th European Conference on Antennas and Propagation, (EuCAP)*, 28 GHz channel modeling using 3D ray-tracing in urban environments, (2015), pp. 1–5
33. M. Samimi, T. Rappaport, in *Proc. of IEEE International Conference on Communications, (ICC)*, 3-D statistical channel model for millimeter-wave outdoor mobile broadband communications, (2015), pp. 2430–2436. <https://doi.org/10.1109/icc.2015.7248689>
34. J. Venna, S. Kaski, in *Proc. of the International Conference on Artificial Neural Networks, (ICANN)*, Neighborhood preservation in nonlinear projection methods: an experimental study, (2001), pp. 485–491. https://doi.org/10.1007/3-540-44668-0_68

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)