# FML: Fast Machine Learning for 5G mmWave Vehicular Communications

Arash Asadi*, Sabrina Müller†, Gek Hong (Allyson) Sim*, Anja Klein†, Matthias Hollick*
*Secure Mobile Networking Lab (SEEMOO), Technische Universität Darmstadt
{firstname.lastname}@seemoo.tu-darmstadt.de
†Communication Engineering Lab, Technische Universität Darmstadt
{firstname.lastname}@nt.tu-darmstadt.de

***Abstract*—Millimeter-Wave (mmWave) bands have become the de-facto candidate for 5G vehicle-to-everything (V2X) since future vehicular systems demand Gbps links to acquire the necessary sensory information for (semi)-autonomous driving. Nevertheless, the directionality of mmWave communications and its susceptibility to blockage raise severe questions on the feasibility of mmWave vehicular communications. The dynamic nature of 5G vehicular scenarios, and the complexity of directional mmWave communication calls for higher context-awareness and adaptability. To this aim, we propose the first** *online learning algorithm* **addressing the problem of** *beam selection with environment-awareness* **in mmWave vehicular systems. In particular, we model this problem as a contextual multi-armed bandit problem. Next, we propose a lightweight context-aware online learning algorithm, namely FML, with proven performance bound and guaranteed convergence. FML exploits coarse user location information and aggregates received data to learn from and adapt to its environment. We also perform an extensive evaluation using realistic traffic patterns derived from Google Maps. Our evaluation shows that FML enables mmWave base stations to achieve near-optimal performance on average within** 33 **minutes of deployment by learning from the available context. Moreover,** *FML remains within* $\sim 5\%$ *of the optimal performance* **by swift adaptation to system changes such as blockage and traffic.**

## I. Introduction

Recent studies highlight the necessity of *multi-Gbps links* to enable 5G vehicle-to-everything (V2X) communications [1]–[3]. Such a high data rate link is needed to acquire accurate sensory data (e.g., HD maps, radar feeds), which is crucial for (semi)-autonomous driving. Due to high congestion in sub-6 GHz bands used by 4G LTE-A systems, the 5G community plans to exploit the underutilized mmWave bands (10-300 GHz). This underutilization is due to the impairments of mmWave bands, such as high path loss and penetration loss. Nevertheless, new research demonstrates that: $(i)$ directional transmission and beamforming is the solution to compensate for the high path loss, and $(ii)$ higher deployment density of base stations is the remedy for short communication range in mmWave bands (100-150 m) [4], [5].

These solutions prove the feasibility of mmWave communication. However, they bring about many new challenges in the system design. Firstly, the directional communication requires *accurate beam alignments* between the base station and the vehicle [6], which is unnecessary for the omnidirectional transmission in sub-6 GHz bands. Secondly, mmWave signals are *prone to blockages* (e.g., buildings, foliage) due to high penetration loss (see Fig. 1). Thus, the performance of mmWave systems can be severely hampered by inaccurate *beam selection*. The performance degradation can be mitigated by enabling the base stations to perform beam selection based on their surrounding environment (e.g., to avoid blockages). In today's network, this knowledge is populated via on-site signal measurements (e.g., war-driving tests), which are time-consuming and unscalable for dense 5G deployments. Moreover, this approach cannot account for dynamic traffic patterns and blockages. We believe that the base stations should autonomously *explore, learn from, and adapt to* their environment to make accurate beam selection while maintaining sustainable scalability. To date, there is no proposal fostering such a capability at mmWave base stations [7]. To this aim, a practical approach should allow the base station to characterize its surroundings autonomously by exploiting the available contextual information. In particular, the correlation between this information (e.g., location of the users) and the outcome of a decision (e.g., beam selection) is the key to optimal future decisions. This emphasizes the necessity of autonomous learning more than ever, specifically to cope with the massive densification of 5G networks [8], [9].

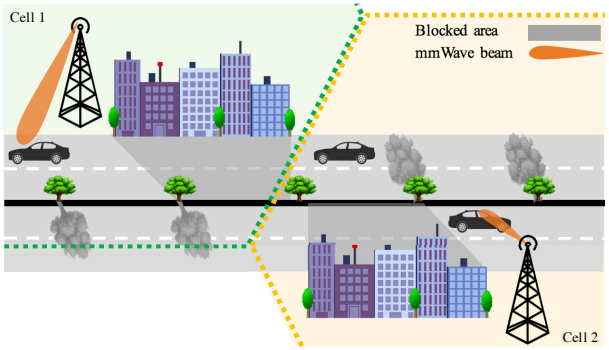


Fig. 1. An example of an mmWave cellular scenario and the impact of different sources of blockage.

**FML algorithm.** In this paper, we propose fast machine learning (FML), which is a low-complexity and a highly scalable online learning algorithm for mmWave base stations. FML is coupled with a practical protocol designed based on the features of the forthcoming 5G cellular network. We model the beam selection as a *contextual multi-armed bandit problem* and propose a *contextual online learning algorithm.* This

algorithm enables the mmWave base stations to autonomously learn from prior decisions and their relations to the available contextual information. In particular, FML explores different beams over time while accounting for contextual information (i.e., vehicles' direction of arrival). The outcome of the exploration is used to adapt to system dynamics such as the appearance of blockages and changes in traffic patterns. FML identifies blockages by evaluating the aggregate received data of each vehicle for each selected beam. It also adapts to traffic patterns by learning the correlation between the direction of arrival and the received data. As a result, FML selects the beams, which maximize the overall network capacity. Consequently, FML provides more coverage to the roads with higher traffic and hence, serves a larger number of vehicles.

FML fights the issues of mmWave vehicular communication on several fronts: $(i)$ it detects permanent blockages (e.g., buildings), and frequently blocked areas due to temporary blockages (e.g., parking spots, bus stations or construction sites frequented by large trucks) using online learning; $(ii)$ it leverages traffic patterns to maximize the system capacity by providing larger coverage (i.e., allocation of more beams) in areas with heavier traffic. This is important because mmWave base stations can transmit simultaneously over a limited number of beams. This limitation depends on the hardware characteristics (e.g., limited number of RF chains), the mmWave channel sparsity, and the beamforming technique [5]; $(iii)$ it infers traffic patterns from the context (i.e., the vehicle's direction of arrival) and selects the best beams. Majority of roads have distinct traffic patterns influenced by the time of the day. For example, the traffic in the main streets tends to go towards the financial center early in the morning and away from it in the evening. While interpreting these patterns are out of the scope of this paper, we design FML to identify and learn from such patterns.

**Our contributions.** The following summarizes the contributions of this paper:

- We model the beam selection at mmWave base stations as a contextual multi-armed bandit problem. Our model is generic, and it is easily adaptable to different contexts.
- We provide the first contextual online learning algorithm for beam selection in mmWave base stations. The algorithm enables the base stations to autonomously learn each beam's data rate, without requiring a training phase.
- We give an analytical upper bound on the regret, i.e., the loss of learning, which proves convergence of FML to the optimal beam selection.
- We demonstrate by means of extensive simulation that – with live and typical traffic patterns obtained from Google Maps at our premises – FML substantially outperforms the benchmark algorithms.

## II. System Model

We consider a heterogeneous cellular system in which mmWave base stations (mmBSs) overlay the coverage area of an LTE eNB (see Fig. 2). This network model is widely expected for forthcoming 5G systems [10]–[13]. The mmBSs

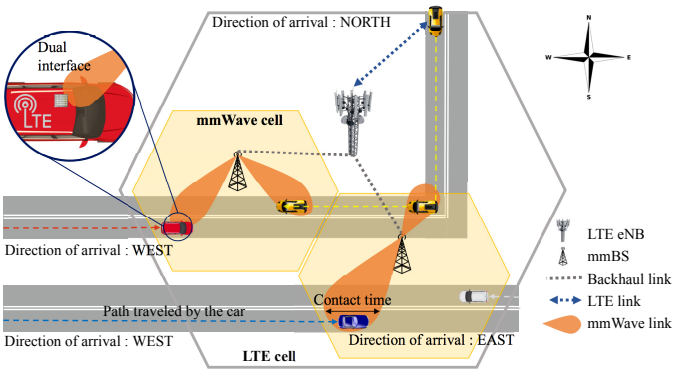


Fig. 2. Illustration of our system model. For clarity, the figure contains only two mmBSs. Each mmBS can transmit over two beams simultaneously. The direction of arrival of the vehicles (shown in dashed line) is derived from the location of the vehicle upon registration to the eNB.

are connected to the eNB via a backhaul link. The vehicles are equipped with $(i)$ an LTE interface to maintain a connection to the eNB, and $(ii)$ an mmWave interface for high-speed data communication. We focus on the downlink in this paper. However, our proposal can be applied in uplink with marginal modifications. We assume neither the eNB nor the mmBSs have any knowledge of their surroundings. We focus on a system with small signaling overhead. On the one hand, the only information available to the mmBS is the direction of arrival of the vehicle (i.e., north, south, east, west), which we define below formally as the vehicle context. On the other hand, the vehicles will only know the location of the mmBS and the selected beam(s).

### A. Choice of Learning Method

We model beam selection in an mmBS as an online learning problem. This is because it allows the mmBS to identify the best beams autonomously over time while accounting for dynamic traffic and environment changes. Specifically, we model the problem as a *multi-armed bandit problem*. Various problems in wireless communications have been treated using multi-armed bandits [14]. In multi-armed bandit problems, a decision maker has to select a subset of actions of unknown rewards with the goal to maximize the reward over time [15]. Such a multi-armed bandit approach is relevant for our problem since an mmBS may only use a limited set of beams simultaneously (as shown in Fig. 2). Therefore, the mmBS needs to identify the best beams by carefully selecting subsets of beams over time. Our approach falls under the category of *contextual multi-armed bandit problems*. These problems additionally include side information that affects the rewards of the actions [16]–[19]. We chose a contextual multi-armed bandit approach since in this way, the mmBS does not simply learn which beams are the best *on average*, but instead it exploits additional information about approaching vehicles to identify which beams are the best under a given traffic situation.

### B. Problem Formulation

The mmBS can use a finite set $\mathcal{B}$ of $B = |\mathcal{B}|$ distinct, orthogonal beams (see Fig. 2). We assume that the mmBS may only select a subset of $m$ beams simultaneously, where $m \in \mathbb{N}$, $m < B$, is a fixed number. This limitation is imposed by the mmWave channel sparsity, beamforming technique, and the hardware characteristics (e.g., number of RF chains) [5]. The goal of the mmBS is to select a subset of $m$ beams that maximizes the amount of data successfully received by the bypassing vehicles in the coverage area. We assume that the mmBS is unaware of its surrounding, i.e., the mmBS does not have prior knowledge about its environment (e.g., street course, blockages). This significantly reduces the complexity of the network implementation as the operator does not need to configure each mmBS based on its surroundings. Hence, the mmBS should learn over time the best subset of beams given its environment. For this purpose, the mmBS should take into account vehicle context, since which are the best beams depends on the context of bypassing vehicles (e.g., their directions of arrival).

We consider a discrete time setting, where the mmBS updates its beam selection in regular time periods. In each period $t = 1, ..., T$, where $T \in \mathbb{N}$ is a finite time horizon, the following events happen:

(i) A set $\mathcal{V}_t = \{v_{t,i}\}_{i=1,...,V_t}$ of $V_t = |\mathcal{V}_t|$ vehicles registers to the mmBS via the LTE eNB. The number $V_t$ of vehicles satisfies $V_t \leq V_{\max}$, where $V_{\max} \in \mathbb{N}$ is the maximum number of supported vehicles in the system. During the registration process, the mmBS receives information about the *context* $x_{t,i}$ of each approaching vehicle $v_{t,i}$. Formally, the context $x_{t,i}$ is an $X$-dimensional vector taken from the bounded context space $\mathcal{X} = [0, 1]^X$.

(ii) The mmBS selects a subset of $m$ beams. We denote the set of selected beams in period $t$ by $\mathcal{S}_t = \{s_{t,j}\}_{j=1,...,m} \subseteq \mathcal{B}$. Then, the vehicles in $\mathcal{V}_t$ are informed about the selected beams via the LTE interface.

(iii) When vehicle $v_{t,i}$ reaches the mmBS's coverage area, the mmBS transmits data to vehicle $v_{t,i}$ and observes the amount of data $r_{s_{t,j}}(x_{t,i}, t)$ vehicle $v_{t,i}$ successfully receives via the selected beams $s_{t,j}, j = 1, ..., m$, until the end of the period $t$.

In general, the amount of data $r_b(x)$ a vehicle with context $x \in \mathcal{X}$ can successfully receive from the mmBS using beam $b \in \mathcal{B}$ during one period is a random variable that depends on the environment of the mmBS (e.g., street conditions and course, blockages, etc.). We call the random variable $r_b(x)$ the *beam performance* (i.e, the aggregate received data by the vehicle) of beam $b$ under context $x$. We assume that this random variable is bounded in $[0, R_{\max}]$, where $R_{\max}$ is the maximum amount of data that can be received by a vehicle. $R_{\max}$ is bounded by the maximum achievable rate of the channel, and it depends on the selected modulation and coding scheme. The contact time as shown in Fig. 2 (i.e., the time within which mmBS can transmit data to the vehicle) is bounded by the coverage area of the beam, which depends on the beam width and speed. By $\mu_b(x)$, we denote the expected value of random variable $r_b(x)$, and we call it the *expected beam performance* of beam $b$ under context $x$. The mmBS aims at selecting a subset of beams which maximizes the expected received data at the vehicles, i.e., maximizes the sum of the expected beam performances. We denote the optimal subset in period $t$, by $\mathcal{B}_t^*(\mathcal{X}_t) = \{b_{t,j}^*(\mathcal{X}_t)\}_{j=1,...,m} \subseteq \mathcal{B}$. The set $\mathcal{B}_t^*(\mathcal{X}_t)$ depends on $\mathcal{X}_t = \{x_{t,i}\}_{i=1,...,V_t}$ and its $m$ beams formally satisfy

$$b_{t,j}^*(\mathcal{X}_t) \in \underset{b \in \mathcal{B} \setminus (\bigcup_{k=1}^{j-1} \{b_{t,k}^*(\mathcal{X}_t)\})}{\operatorname{argmax}} \sum_{i=1}^{V_t} \mu_b(x_{t,i}) \tag{1}$$

for $j = 1, ..., m$. Hence, if the mmBS *knew* the expected beam performances $\mu_b(x)$ for each vehicle context $x \in \mathcal{X}$ and each beam $b \in \mathcal{B}$ *a priori, like an oracle,* it could simply select the optimal subset of beams for each set of approaching vehicles according to (1). Over the sequence $1, ..., T$ of periods, this would yield an expected amount of

$$\sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \mathrm{E}[r_{b_{t,j}^*(\mathcal{X}_t)}(x_{t,i})] = \sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \mu_{b_{t,j}^*(\mathcal{X}_t)}(x_{t,i}) \tag{2}$$

data that can be received in total.

However, the mmBS does not know the environment, and hence it has to learn the expected beam performances $\mu_b(x)$ over time. In order to learn these values, the mmBS has to try out different beams for different vehicle contexts over time. At the same time, it should ensure that those beams that were already proven to be good are used sufficiently often. Hence, the mmBS has to find a trade-off between exploring beams of which it has little knowledge and exploiting beams with high average beam performance. In the next section, we will present a *learning algorithm*, which for each period with approaching vehicles of contexts $\mathcal{X}_t$, selects a subset $\mathcal{S}_t$ of $m$ beams. The selection of the learning algorithm depends on the history of selected beams in previous periods and the corresponding observed beam performances. Given the selections $\mathcal{S}_t$, $t = 1, ..., T$, of the learning algorithm, the expected amount of data received by the vehicles is given by

$$\sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \mathrm{E}[r_{s_{t,j}(\mathcal{X}_t)}(x_{t,i})] = \sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \mathrm{E}[\mu_{s_{t,j}(\mathcal{X}_t)}(x_{t,i})]. \tag{3}$$

The expected difference in the amount of received data achieved by an oracle and by the learning algorithm is called the *regret of learning*. Given (2) and (3), it is defined as

$$\begin{aligned} R(T) &= \mathrm{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \left(r_{b_{t,j}^*(\mathcal{X}_t)}(x_{t,i}) - r_{s_{t,j}(\mathcal{X}_t)}(x_{t,i})\right)\right] \\ &= \sum_{t=1}^{T} \sum_{i=1}^{V_t} \sum_{j=1}^{m} \left(\mu_{b_{t,j}^*(\mathcal{X}_t)}(x_{t,i}) - \mathrm{E}[\mu_{s_{t,j}(\mathcal{X}_t)}(x_{t,i})]\right). \end{aligned} \tag{4}$$

## III. FML ALGORITHM

The above problem formulation corresponds to a contextual multi-armed bandit problem and we propose a contextual online learning algorithm inspired by [19]. Intuitively, the algorithm learns the expected beam performances under different contexts online over time. The algorithm works on the assumption that for similar vehicle contexts, the performance of a particular beam will on average be similar.

The algorithm first uniformly partitions the context space into small sets of similar contexts and learns about the performance of different beams independently in each of these small sets. Then, in each of its discrete periods, the algorithm either enters an exploration phase or an exploitation phase. The phase it enters is decided based on the contexts of approaching vehicles and based on a control function. While in exploration phases, the algorithm selects a random subset of beams, in exploitation phases, the algorithm selects beams that showed the highest performance when selected in previous periods. By observing the amount of data received by vehicles in the system, the algorithm acquires performance estimates of beams; thereby, it learns the performance of the different beams under different vehicle contexts over time.

### A. Detailed Description

In detail, our proposed beam selection algorithm, called *Fast Machine Learning* (FML) works as follows: First, during initialization (lines 2-4), FML uniformly partitions the context space $\mathcal{X} = [0,1]^X$ into $(p_T)^X$ $X$-dimensional hypercubes of size $(\frac{1}{p_T})^X$, where $p_T$ is an input to the algorithm. We call the resulting partition $\mathcal{P}_T$. Moreover, FML initializes the counters $N_{b,h}(t)$ and $\hat{\mu}_{b,h}(t)$ for each beam $b \in \mathcal{B}$ and each hypercube $h \in \mathcal{P}_T$. The counter $N_{b,h}(t)$ represents the total number of vehicles with the context in hypercube $h$ that approached the mmBS whenever beam $b$ was selected in any of the periods $1, ..., t-1$. The counter $\hat{\mu}_{b,h}(t)$ represents the estimated beam performance of beam $b$ for vehicles with the context in hypercube $h$.

In period $t$, FML observes the contexts $\mathcal{X}_t := \{x_{t,i}\}_{i=1,...,V_t}$ of the $V_t$ approaching vehicles and for each context $x_{t,i}$, FML determines to which hypercube this context belongs to (lines 6-7), i.e., it finds $h_{t,i} \in \mathcal{P}_T$ with $x_{t,i} \in h_{t,i}$. Based on the collection $\mathcal{H}_t := \{h_{t,i}\}_{i=1,...,V_t}$ of hypercubes, FML next (line 8) computes the set $\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)$ of under-explored beams via

$$\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t) := \cup_{i=1}^{V_t} \{b \in \mathcal{B} : N_{b,h_{t,i}}(t) \leq K(t)\}, \qquad (5)$$

where $K : \{1, ..., T\} \mapsto \mathbb{R}$ is a deterministic, monotonically increasing control function, which the algorithm gets as input. If there are under-explored beams, FML enters an exploration phase (lines 9-16). In case the number $u(t) := |\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)|$ of under-explored beams is at least $m$, FML randomly selects $m$ of them. In case the number $u(t)$ of under-explored beams is smaller than $m$, FML selects all $u(t)$ beams. In addition, it selects the $(m - u(t))$ beams $\hat{b}_{1,\mathcal{H}_t}(t), ..., \hat{b}_{m-u,\mathcal{H}_t}(t)$ from $\mathcal{B} \setminus \mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)$, which satisfy

$$\hat{b}_{j,\mathcal{H}_t}(t) \in \underset{b \in \mathcal{B} \setminus (\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t) \cup \bigcup_{k=1}^{j-1} \{\hat{b}_{k,\mathcal{H}_t}(t)\})}{\text{argmax}} \sum_{i=1}^{V_t} \hat{\mu}_{b,h_{t,i}}(t) \qquad (6)$$

for $j = 1, ..., m - u(t)$. If there are no under-explored beams, FML enters an exploitation phase (lines 17-19). It selects the $m$ beams $\hat{b}_{1,\mathcal{H}_t}(t), ..., \hat{b}_{m,\mathcal{H}_t}(t)$ from $\mathcal{B}$, which satisfy

$$\hat{b}_{j,\mathcal{H}_t}(t) \in \underset{b \in \mathcal{B} \setminus (\bigcup_{k=1}^{j-1} \{\hat{b}_{k,\mathcal{H}_t}(t)\})}{\text{argmax}} \sum_{i=1}^{V_t} \hat{\mu}_{b,h_{t,i}}(t) \qquad (7)$$

for $j = 1, ..., m$. After beam selection, FML observes the beam performance of each selected beam for each vehicle within this period (line 20). Using these observations, FML updates its internal counters (lines 21-25).

**Algorithm 1** Pseudocode of FML algorithm.

1: Input: $T$, $p_T$, $K(t)$
2: Initialize context partition: Create partition $\mathcal{P}_T$ of context space $[0,1]^X$ into $(p_T)^X$ hypercubes of identical size
3: Initialize counters: For all $b \in \mathcal{B}$ and all $h \in \mathcal{P}_T$, set $N_{b,h} = 0$
4: Initialize estimates: For all $b \in \mathcal{B}$ and all $h \in \mathcal{P}_T$, set $\hat{\mu}_{b,h} = 0$
5: **for each** $t = 1, ..., T$ **do**
6: Observe vehicle contexts $\mathcal{X}_t = \{x_{t,i}\}_{i=1,...,V_t}$
7: Find $\mathcal{H}_t = \{h_{t,i}\}_{i=1,...,V_t}$ such that $x_{t,i} \in h_{t,i} \in \mathcal{P}_T, i = 1, ..., V_t$
8: Compute the set of under-explored beams $\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)$ in (5)
9: **if** $\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t) \neq \emptyset$ **then** ▷ Exploration
10: $u = \text{size}(\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t))$
11: **if** $u \geq m$ **then**
12: Select $s_{t,1}, ..., s_{t,m}$ randomly from $\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)$
13: **else**
14: Select $s_{t,1}, ..., s_{t,u}$ as the $u$ beams from $\mathcal{B}^{\text{ue}}_{\mathcal{H}_t}(t)$
15: Select $s_{t,u+1}, ..., s_{t,m}$ as the $(m-u)$ beams $\hat{b}_{1,\mathcal{H}_t}(t), ..., \hat{b}_{m-u,\mathcal{H}_t}(t)$ from (6)
16: **end if**
17: **else** ▷ Exploitation
18: Select $s_{t,1}, ..., s_{t,m}$ as the $m$ beams $\hat{b}_{1,\mathcal{H}_t}(t), ..., \hat{b}_{m,\mathcal{H}_t}(t)$ from (7)
19: **end if**
20: Observe received data $r_{j,i}$ of each vehicle $v_{t,i}$, $i = 1, ..., V_t$, in each beam $s_{t,j}, j = 1, ..., m$
21: **for** $i = 1, ..., V_t$ **do**
22: **for** $j = 1, ..., m$ **do**
23: $\hat{\mu}_{s_{t,j},h_{t,i}} = \frac{\hat{\mu}_{s_{t,j},h_{t,i}} N_{s_{t,j},h_{t,i}} + r_{j,i}}{N_{s_{t,j},h_{t,i}} + 1}$ and $N_{s_{t,j},h_{t,i}} = N_{s_{t,j},h_{t,i}} + 1$
24: **end for**
25: **end for**
26: **end for**

### B. Regret and Choice of Parameters

The regret of FML in (4) can be bounded from above. The upper bound given below is based on the following assumption, which states that, the expected beam performance of a particular beam is similar in similar contexts:

**Assumption 1.** *There exist $L > 0$ and $\alpha > 0$ such that for all $b \in \mathcal{B}$ and for all $x, y \in \mathcal{X}$, it holds that*

$$|\mu_b(x) - \mu_b(y)| \leq L||x - y||^\alpha,$$

*where $||\cdot||$ denotes the Euclidean norm in $\mathbb{R}^X$.*

The regret of FML can be bounded as follows (see [19]):

**Theorem 1** (Bound for $R(T)$)**.** *Let* $K(t) = t^{\frac{2\alpha}{3\alpha+X}} \log(t)$ *and* $p_T = \lceil T^{\frac{1}{3\alpha+X}} \rceil$*. If FML is executed using these parameters and if Assumption 1 holds true, the leading order of the regret* $R(T)$ *is* $O\left(mV_{\max}R_{\max}BT^{\frac{2\alpha+X}{3\alpha+X}} \log(T)\right)$.

This theorem shows that the regret of FML is sublinear in the time horizon $T$, i.e., $R(T) = O(T^{\gamma})$ with $\gamma < 1$. This implies that $\lim_{T\to\infty} \frac{R(T)}{T} = 0$ holds, which guarantees that the algorithm has an asymptotically optimal performance. Hence, over time, FML converges to the optimal beam selection strategy. Moreover, for finite time horizon $T$, the upper bound on the regret characterizes FML's speed of convergence.

## IV. Numerical Evaluation

Here, we evaluate and benchmark `FML` via numerical simulations. In the following, we first describe the simulation environment and the relevant parameters. Next, we provide the path loss model and other simulation settings, which are chosen according to the 3GPP technical specification in [20]. Then, the benchmark algorithms and results are presented.

### A. Simulation Setup

The simulation scenario (e.g., blockages, roads, traffic patterns) is designed with reference to information obtained from Google Maps in the vicinity of our premises. The mmBS is assumed to have 16 orthogonal beams with variable beam width from $10°$ to $40°$ covering the $360°$ azimuth. The beams are selected according to the recent measurements in [21]. The vehicles enter the system with an arrival rate of $\lambda$ (in vehicles per second) and their speed varies between 20km/h and 70km/h. Each vehicle chooses one of the routes on the map, whose probability is determined by the typical traffic observed within the area in Google Maps (see Fig. 3). We consider two types of blockages: permanent and temporary blockages. Permanent blockages are the buildings that *permanently* block the path between the mmBS and the vehicles on the road. The temporary blockages (e.g., a large vehicle or tram) are modeled by random appearance of objects on either side of the road causing a *temporary* signal blockage. In our implementation, a time period is defined as the time in which the vehicle under observation enters and leaves the cell coverage of the mmBS. Thus, the length of the time periods varies over the simulation depending on the path and speed of the vehicles. Within this time period, the learning algorithms additionally learn from the context and received data of the other vehicles passing through the selected beams. In this way, we ensure that the algorithms have enough samples to learn from. Table I contains the important simulation parameters.

### B. Benchmark Algorithms and Metrics

We provide a thorough performance analysis by comparing `FML` to several other schemes. The following elaborates on each benchmark scheme:

- **`Optimal`.** This algorithm has a priori knowledge about the expected beam performance $\mu_b(x)$ of each beam $b \in \mathcal{B}$ in each context $x \in \mathcal{X}$. In each period, `Optimal` selects the optimal subset $\mathcal{B}_t^*(\mathcal{X}_t)$ of $m$ beams as in (1) and hence gives an upper bound to the other algorithms.
- **`UCB`.** This is a variant of the classical learning algorithm `UCB` [15], which we adapted to our use-case. It learns from previously observed beam performance, but without taking into account context information. In each period, `UCB` selects $m$ beams with the highest estimated upper confidence bounds on their expected beam performance.
- **`MaxRate`.** This algorithm first explores each beam once. Then, it sticks to the beam with the highest received data.
- **`Random`.** This algorithm selects $m$ random beams in each period.

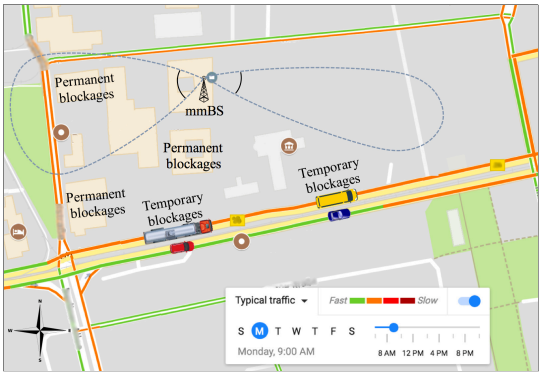


Fig. 3. The map of the simulation environment. The tram and bus temporarily block the red vehicle and black vehicle, respectively.

**Performance metrics.** The performance metrics used in the evaluation are aggregate and cumulative received (rx) data, the number of served vehicles, and average learning time (i.e., the time required for `FML` to reach a certain percentage of the `Optimal`'s performance). The aggregate received data is defined as the data received (in bits) by all the vehicles in the system in time period $t$. The cumulative received data is defined as the data received by all the vehicles in the system up to time period $t$.

### C. Numerical Evaluation

Here, we first evaluate a generic scenario. Next, we analyze the impact of several parameters, i.e., the arrival rate of the vehicles, the number of selected beams, the frequency of blockages, and the underlying traffic patterns. Unless otherwise stated, we consider the case where (i) the percentage of permanent and temporary blockages each corresponds to $20\%$

TABLE I
Channel parameters as specified in [20].

| Parameter | Value |
|---|---|
| Carrier frequency | 28 GHz |
| System Bandwidth | 1 GHz |
| Transmit power | 30 dBm |
| Noise figure | 4 dB@mmBS, 7 dB@Vehicle |
| Vehicle's beam width | $30°$ |
| Thermal noise | $-174$ dBm/Hz |
| Path loss model (dB) | $32.4 + 17.3 \log_{10} d(\text{m}) + 20 \log_{10}(f_c(\text{GHz})) + \xi$<br>$\xi \sim \mathcal{N}(0, \sigma), \sigma = 1.1\text{dB}$ |

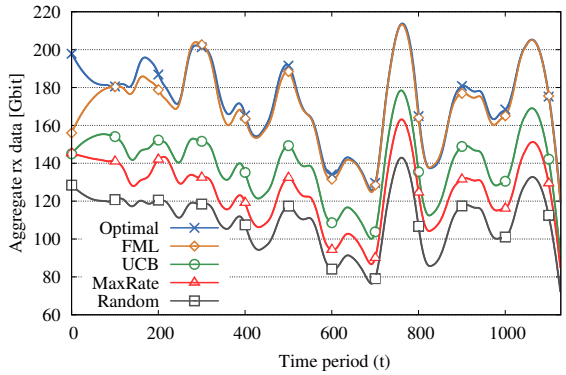


Fig. 4. Aggregate received data for arrival rate with $\lambda = 0.4$ and $m = 4$.

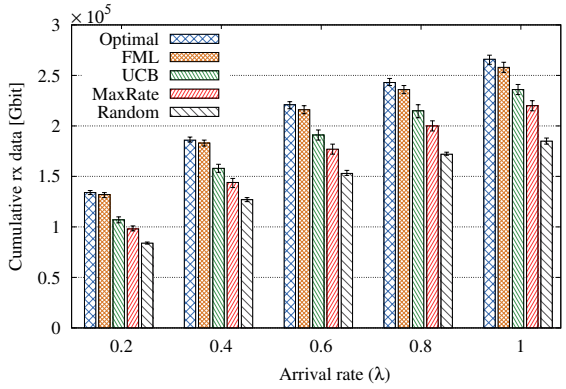


Fig. 5. Impact of arrival rate $\lambda$ on cumulative received data for $m = 4$ after.

of all paths and (ii) traffic patterns (i.e., the vehicle arrival rate and the route probabilities) change based on the typical patterns provided by Google Maps. Since Google Maps only provides the typical daily traffic patterns for 17 hours of the day (from 06:00 to 22:00), we run algorithms over a 17-hour simulation. Each simulation is repeated 20 times for which we show 95% confidence intervals in the figures.

*1) **Average received data**:* In the following, we analyze the aggregate received data achieved by the algorithms over a time horizon of 17 hours. Fig. 4 shows the aggregate received data per time period for an arrival rate of $\lambda = 0.4$ in the case of $m = 4$ selected beams per period. The fluctuations in the graph result from the number of vehicles in the system. Specifically, the aggregate received data increases with the number of vehicles. The impact of vehicle arrival rate and traffic patterns are evaluated in the following sections in detail. As expected, `Optimal` gives an upper bound to the other algorithms due to a priori knowledge of the expected beam performance. Our proposed algorithm `FML` clearly outperforms the other algorithms `UCB`, `MaxRate`, and `Random`. We observe that `FML`'s performance quickly approaches that of `Optimal` within the first 100 periods while the other algorithms perform at least $\sim$ 20% worse than `FML`. This behavior is even more pronounced starting from the $256^{th}$ period when `FML` remains very close to `Optimal`'s aggregate received data. `FML` experiences small divergence (below 3%) from `Optimal` at a few points within the simulation. These small variations are due to the occurrence of new events, which are not learned from or the re-exploration of past events. Average performance within 17 hours of simulation indicates that the average aggregate received data achieved by `FML` is 21.99%, 36.08% and 54.76% higher than that achieved by `UCB`, `MaxRate` and `Random`, respectively. Moreover, on average, `FML` performs only 1.73% below that achieved by `Optimal`.

*2) **Impact of arrival rate**:* Next, we investigate the impact of the arrival rate on the cumulative received data achieved by the different algorithms in the case of $m = 4$ selected beams per period for different arrival rates $\lambda \in \{0.2, 0.4, 0.6, 0.8, 1\}$. From Fig. 5, we can observe that the cumulative received data grows as the number of vehicles in the system increases. Over the whole range of $\lambda$, the cumulative received data achieved by `FML` lies between 9.36% and 23.06% higher than that achieved by the next-best algorithm `UCB` and only up to 3.06% lower than that achieved by the `Optimal`.

Fig. 6 shows the time required for `FML` to achieve $\{80\%, 85\%, 90\%\}$ of the `Optimal`'s performance for different $\lambda$, respectively. It is observed from our evaluation that `FML` achieves $\{80\%, 85\%, 90\%\}$ of `Optimal`'s performance within $\{13, 25, 56\}$ minutes for all arrival rates, respectively. The large confidence interval is due to the randomness of certain parameters both in the learning algorithm (exploration decisions) and in the evaluation scenario (e.g., location of temporary blockages, selected routes, and speed). As a result, `FML` may approach near-optimal performance below seven minutes if all random effects are in favor, and up to 75 minutes otherwise. Note that even manual configuration and war-driving tests require much more than 75 minutes. Furthermore, war-driving tests may only capture effects of permanent, but not from temporary blockages. The subplot in Fig. 6 shows the average received (rx) data with $\lambda = 0.4$. Average rx data is the average data over all the vehicles in the system up to this time period. This figure illustrates `FML`'s quick learning and adaptation capability. In particular, `FML` achieves 90% of the performance `Optimal` within 30 time periods. This result tallies with the performance figure shown for $\lambda = 0.4$ in Fig. 6 and the trend observed in Fig. 4. This shows how quickly `FML` converges to near-optimal beam selection. Moreover, the general trend shows that the time to converge to near-optimal results reduces when the vehicle density in the system increases. This is due to the fact that with higher vehicle density, `FML` has more examples to learn from simultaneously.

*3) **Impact of the number of selected beams**:* Here, we analyze the impact of the number of selected beams $m$ per period on the cumulative received data. Fig. 7 shows the cumulative received data achieved with an arrival rate of $\lambda = 0.4$ for different $m \in \{1, 2, 4, 8\}$. As the number of simultaneously selected beams increases, the cumulative received data increases as well. This increase is due to the enhanced coverage area. However, as mentioned earlier, the higher the number of beams, the higher is the hardware complexity and energy consumption at the mmBS [5]. For different values of $m$, the cumulative received data achieved by `FML` is between 10.45% and 18.98% higher than that achieved

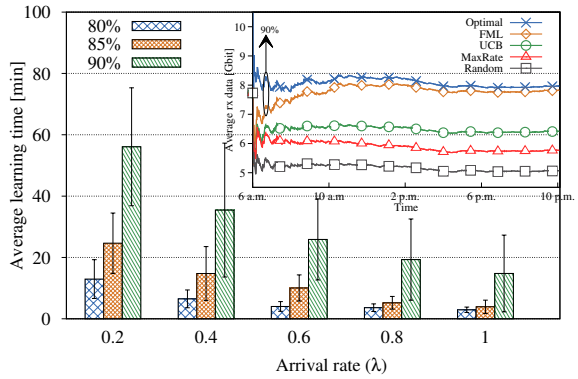


Fig. 6. Impact of arrival rate $\lambda$ on required time for FML to achieve a certain percentage of Optimal's performance for $m = 4$.

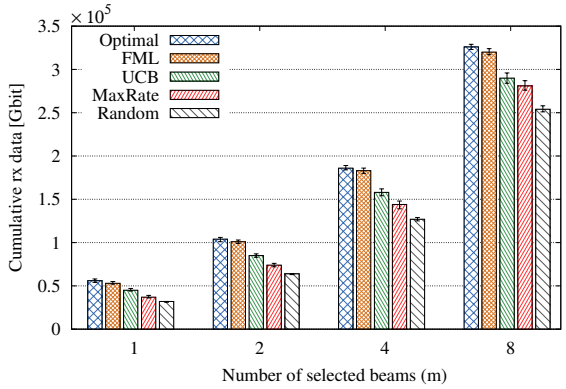


Fig. 7. Impact of number $m$ of selected beams per period on cumulative received data for arrival rate $\lambda = 0.4$ after 17 hours.

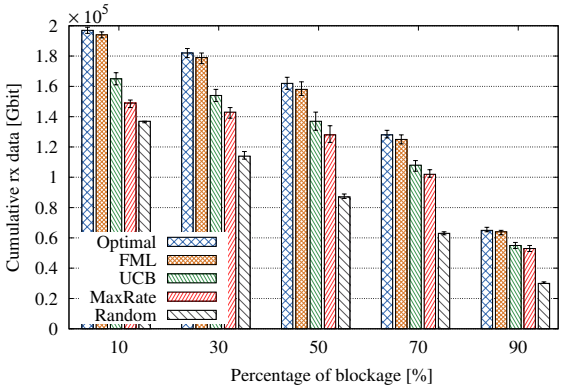


Fig. 8. Impact of blockages on cumulative received data for arrival rate $\lambda = 0.4$ after 17 hours.

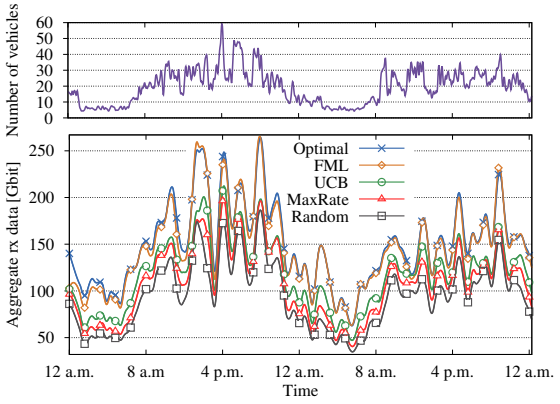


Fig. 9. Data for 48 hours of live daily traffic pattern.

by the next-best algorithm UCB and only up to 4.71% lower than that achieved by the Optimal.

*4) **Impact of blockages**:* Here, we investigate the impact of blockages on the cumulative received data. Fig. 8 shows the cumulative received data with an arrival rate of $\lambda = 0.4$ in case of $m = 4$ selected beams per period for $\{10\%, 30\%, 50\%, 70\%, 90\%\}$ of permanent blockages in the system. Clearly, as the percentage of permanent blockages in the system increases, the cumulative received data decreases. For any percentage of permanent blockages, FML outperforms all non-optimal algorithms. The cumulative received data achieved by FML lies between 15.55% and 17.42% higher than that achieved by the next-best algorithm UCB. Moreover, FML's achieved results deviate from that of Optimal merely by at most 2.61%.

*5) **Live daily traffic pattern**:* The prior evaluation was based on the typical traffic pattern as in Fig. 3. Due to the averaging effect, Google's typical traffic does not capture the quick changes in traffic patterns which are visible in the live traffic report. To this aim, we recorded the observed live traffic reports of Google for a period of 48 hours in 30-minute intervals. We fed this data to the simulator to evaluate the performance of FML in live traffic conditions. The top plot of Fig. 9 shows the number of vehicles in the system within a 48-hour-period. Clearly, the arrival rate has characteristic peaks during the course of a day (especially between 6 a.m. and 10 p.m.), which lead to an increase in vehicle density. For better readability, the graph in Fig. 9 is smoothed by 5% and 1% for the figures at the top and bottom, respectively, with a local regression using weighted linear least squares and a $2^{nd}$-degree polynomial model. The bottom plot of Fig. 9 shows the aggregate received data for $m = 4$ selected beams per period. We can see that FML achieves near-optimal aggregate received data within at most two hours. This tallies with our observation of $256^{th}$ time periods in Fig. 4. We also observe that FML can capture the effect of traffic fluctuations and leverage it to make better decisions. Since the other algorithms do not adapt to the change in traffic, they perform worse than FML. Averaging over 48 hours of simulation, FML performs 24.96%, 39.61%, and 60.51% better than UCB, MaxRate, and Random, respectively. Further, the performance of FML only lies within 2.47% below that of Optimal.

## V. Discussion

Our evaluation results confirm the superiority and adaptability of FML compared to the benchmark algorithms. In this section, we discuss the impact of some of our system model assumptions (e.g., beam orthogonality) and the capabilities of FML, which were not detailed previously.

**Beam orthogonality.** We assumed that the available beams at the mmBS are all orthogonal. Nevertheless, FML's per-

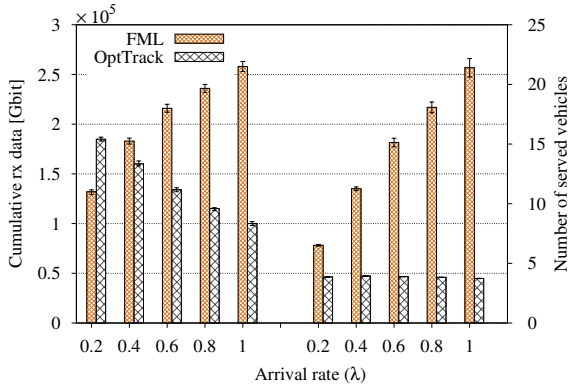


Fig. 10. Impact of number $m$ of selected beams per period on cumulative received data for arrival rate $\lambda = 0.4$ after 17 hours.

formance is not restricted to this limitation. In fact, non-orthogonality can be formulated as an additional constraint in our model, where overlapping beams cannot be used simultaneously. In this case, the set of actions (i.e., different *beam combinations*) naturally increases. However, not all possible combinations are allowed due to the constraint that overlapping beams may not be used simultaneously. Thus, the number of actions remains feasible.

**Interference.** In this paper, we assume that there is no interference among orthogonal beams. However, *theoretically* orthogonal beams may interfere with each other due to reflections from surrounding objects (e.g., walls, buildings). Such scenarios are very rare in high frequency ranges of mmWave bands such as 60 GHz due to low reflection coefficient but more plausible in 28 GHz and 38 GHz. Although this factor is not considered in our simulation, its impact will be interpreted as a blockage by FML, thus refraining from communicating over those beams.

**Co-located vehicles.** There could be multiple vehicles in the coverage area of a beam at the same time. Given the focus of this paper is on unicast communication, only one of these vehicles is allowed to communicate with the mmBS. In our simulation, we select one of the vehicles at random for communication. Nevertheless, it should be noted that this phenomenon has a high potential for multicast scenarios.

**Number of selected beams.** We assumed that the number $m$ of simultaneously selectable beams at the mmBS is limited due to current hardware limitation [5]. Assuming fully digital beamforming materializes in the future, this limitation will present itself as overlapping beam patterns, which cannot be selected simultaneously. FML can also be used in such scenarios to select the most suitable beam pattern according to the available context (e.g., vehicle's location).

**Location reporting.** FML infers traffic patterns from a very coarse location information (i.e., direction of arrival). This information is chosen to emphasize the potential of FML to infer traffic patterns based on coarse geo-locations. If the network can afford the extra overhead for high-resolution location reporting, FML's performance would only improve.

## VI. Related Work

Beam selection issues have been addressed before in conventional vehicular networks operating at sub-6 GHz frequencies to achieve maximum rate using multi-lobe beam patterns [22], [23]. Unlike our proposal, these works rely on accurate GPS location reporting in order to perform beam switching. The complexity of this method grows exponentially with variable vehicular speed and channel conditions. Moreover, as mentioned, the signal propagation characteristics in sub-6 GHz frequencies significantly differs from mmWave bands.

This is the first paper to propose an adaptable learning algorithm for mmWave vehicular scenarios in which blockages, and traffic are taken into account. Specifically, our algorithm does not require accurate localization information, and its performance is independent of the aforementioned variabilities. In what follows, we provide a short overview of the ongoing efforts in mmWave vehicular research. For more details, we encourage the readers to refer to a recent survey on the topic in [7]. The body of the works on mmWave V2X can be categorized in channel characterization, PHY design, and MAC design.

In [24], Kato et al. provide propagation characteristics of mmWave communication in inter-vehicular scenarios. On the other hand, the authors of [25] derive closed-form approximation for coherence time and beam width with consideration of directional communication. The feasibility of mmWave communication is analyzed in [4] via an extensive measurement campaign. The results indicate the low number of scattering cluster in mmWave bands, which imply that the number of supported data streams are significantly less than the antenna array size. The challenges of enabling mmWave V2X communication is elaborated in [26]. In addition, the authors describe possible solutions to these challenges from PHY and MAC perspective. The works in [27]–[30] focus on mmWave beam adaptation in vehicular scenarios. These works exploit DSRC to estimate the location of the vehicles. This estimation allows mmBS to track the moving vehicles and to adapt the mmWave beam accordingly. While feasible, this technique requires a complex transceiver chain and accurate localization information. Further, modification to the MAC protocol to allow the exchange of this information between the interfaces is crucial.

While prior work agrees that blockage is the Achilles heel of mmWave communication [4]–[7], [26], none of the above provide a method to detect blockages in vehicular scenarios. Furthermore, traffic-awareness is another aspect which has not been addressed to the best of our knowledge. Finally, this work provides the first online learning algorithm for vehicular mmWave communications.

## VII. Conclusions

In this paper, we address the problem of beam selection at mmWave base stations where the outcome of the selection is highly dependent on the traffic and the blockages in the network. To this aim, we propose FML, an online

learning algorithm based on contextual multi-armed bandits that operates on minimal contextual network information (i.e., a vehicle's direction of arrival). In addition, we analyze the implementation feasibility of FML in the cellular network by proposing a protocol within the definition of 3GPP standard. The advantage of FML is twofold: $(i)$ it enables mmWave base stations to autonomously learn from the context to understand their surrounding environment and $(ii)$ it provides a scalable solution to increase the deployment density of mmWave base stations with minimal setup overhead for the operators. Our evaluation results show that FML requires on average only 33 minutes to achieve near-optimal performance. Noteworthily, without the overhead of tracking of OptTrack, FML achieves $61.37\%$ and $82.55\%$ gain in terms of the cumulative received data and number of served vehicles, respectively. The results demonstrate the capability of online bandit learning and emphasize on the relevance of context-awareness in 5G scenarios.

This is the first attempt to incorporate learning algorithms in such dynamic vehicular scenarios. However, our analytical modeling of the system can be extended to observe and learn from richer context (e.g., coordinates and type of the vehicle). Also extending FML to non-orthogonal code-books and evaluating the combinatorial overhead would be an interesting future research direction. Given the low complexity of contextual bandit learning, this approach can be applied in other areas of mmWave cellular networking such as initial access and user handover.

## VIII. Acknowledgements

This work has been co-funded by the LOEWE initiative (Hessen, Germany) within the NICER project and the German Research Foundation (DFG) in the Collaborative Research Center (CRC) 1053 - MAKI.

## References

[1] K. Sakaguchi, T. Haustein, S. Barbarossa, E. C. Strinati, A. Clemente, G. Destino, A. Pärssinen, I. Kim, H. Chung, J. Kim, W. Keusgen, R. J. Weiler, K. Takinami, E. Ceci, A. Sadri, L. Xain, A. Maltsev, G. K. Tran, H. Ogawa, K. Mahler, and R. W. Heath Jr, "Where, When, and How mmWave is Used in 5G and Beyond," *arXiv.org*, Apr 2017.

[2] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-Wave Vehicular Communication to Support Massive Automotive Sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, Dec 2016.

[3] A. Loch, A. Asadi, G. H. Sim, J. Widmer, and M. Hollick, "mm-Wave on Wheels: Practical 60 GHz Vehicular Communication without Beam Training," in *COMSNETS*, 2017, pp. 1–8.

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[5] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave Cellular Wireless Networks: Potentials and Challenges," in *Proceedings of the IEEE*, Feb 2014.

[6] A. Alkhateeb, Y.-H. Nam, M. S. Rahman, J. Zhang, and R. W. Heath, "Initial Beam Association in Millimeter Wave Cellular Systems: Analysis and Design Insights," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2807–2821, Mar 2017.

[7] V. Va, T. Shimizu, G. Bansal, and R. W. Heath Jr, "Millimeter Wave Vehicular Communications: A Survey," *Foundations and Trends in Networking*, vol. 10, no. 1, pp. 1–113, Jun 2016.

[8] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Key Elements to Enable Millimeter Wave Communications for 5G Wireless Systems," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 136–143, Dec 2014.

[9] M. Series, "Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced," *Report ITU*, no. 2135-1, Dec 2009.

[10] H. Shokri-Ghadikolaei, F. Boccardi, E. Erkip, C. Fischione, G. Fodor, M. Kountouris, P. Popovski, and M. Zorzi, "The Impact of Beamforming and Coordination on Spectrum Pooling in MmWave Cellular Networks," in *ASILOMAR*, Nov 2016.

[11] M. Giordani, M. Mezzavilla, and S. Rangan, "Multi-Connectivity in 5G mmwave cellular networks," in *IFIP Med-Hoc-Net*, Jun 2016.

[12] G. H. Sim, A. Asadi, A. Loch, M. Hollick, and J. Widmer, "Opp-relay: Managing Directionality and Mobility Issues of Millimeter-wave via D2D Communication," in *COMSNETS*, Jan 2017, pp. 144–151.

[13] H. Shokri-Ghadikolaei, F. Boccardi, C. Fischione, G. Fodor, and M. Zorzi, "Spectrum Sharing in mmWave Cellular Networks via Cell Association, Coordination, and Beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2902–2917, Nov 2016.

[14] S. Maghsudi and E. Hossain, "Multi-armed bandits with application to 5G small cells," *IEEE Wireless Communications*, vol. 23, no. 3, pp. 64–73, Jun 2016.

[15] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time Analysis of the Multiarmed Bandit Problem," *Journal of Machine Learning*, vol. 47, no. 2-3, pp. 235–256, May 2002.

[16] T. Lu, D. Pal, and M. Pal, "Contextual Multi-Armed Bandits," in *AISTATS*, May 2010.

[17] A. Slivkins, "Contextual Bandits with Similarity Information," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 2533–2568, Jan 2014.

[18] C. Tekin, S. Zhang, and M. van der Schaar, "Distributed Online Learning in Social Recommender Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 638–652, Aug 2014.

[19] S. Müller, O. Atan, M. van der Schaar, and A. Klein, "Context-Aware Proactive Content Caching with Service Differentiation in Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1024–1036, Feb 2017.

[20] 3GPP, "Technical Specification Group Radio Access Network: Study on channel model for frequencies from 0.5 to 100 GHz ," Tech. Rep., Mar 2017.

[21] T. Nitsche, G. Bielsa, I. Tejado, A. Loch, and J. Widmer, "Boon and Bane of 60 GHz Networks: Practical Insights into Beamforming, Interference, and Frame Level Operation," in *ACM CoNEXT*, May 2015.

[22] V. Navda, A. P. Subramanian, K. Dhanasekaran, A. Timm-Giel, and S. Das, "Mobisteer: Using steerable beam directional antenna for vehicular network access," in *ACM MobiSys*, Jun 2007.

[23] K. Ramachandran, R. Kokku, K. Sundaresan, M. Gruteser, and S. Rangarajan, "R2D2: Regulating Beam Shape and Rate As Directionality Meets Diversity," in *ACM MobiSys*, Jun 2009.

[24] A. Kato, K. Sato, and M. Fujise, "ITS Wireless Transmission Technology. Technologies of Millimeter-wave Inter-vehicle Communications: Propagation Characteristics," *Journal of the Communications Research Laboratory*, vol. 48, no. 1, pp. 99–110, Nov 2001.

[25] V. Va and R. W. Heath, "Basic Relationship Between Channel Coherence Time and Beamwidth in Vehicular Channels," in *IEEE VTC*, Jan 2015.

[26] M. Giordani, A. Zanella, and M. Zorzi, "Millimeter Wave Communication in Vehicular Networks: Challenges and Opportunities," in *IEEE MOCAST*, May 2017.

[27] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Beam Design for Beam Switching Based Millimeter Wave Vehicle-to-infrastructure Communications," in *IEEE ICC*, Jul 2016.

[28] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-Wave Vehicular Communication to Support Massive Automotive Sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, Dec 2016.

[29] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. R. Nix, "MmWave System for Future ITS: A MAC-layer Approach for V2X Beam Steering," *CoRR*, vol. abs/1705.08684, May 2017. [Online]. Available: http://arxiv.org/abs/1705.08684

[30] ——, "Beam Alignment for Millimetre Wave Links with Motion Prediction of Autonomous Vehicles," *CoRR*, vol. abs/1702.04264, Feb 2017. [Online]. Available: http://arxiv.org/abs/1702.04264