

## Research Article

# Network Capacity Optimization for Cellular-Assisted Vehicular Systems by Online Learning-Based mmWave Beam Selection

Jinsong Gui<sup>1</sup>, Yao Liu<sup>1</sup>, Xiaoheng Deng<sup>1</sup> and Bin Liu<sup>2</sup>

<sup>1</sup>School of Computer Science and Engineering, Central South University, Changsha 410083, China

<sup>2</sup>Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

Correspondence should be addressed to Jinsong Gui; [jsgui06@163.com](mailto:jsgui06@163.com)

Received 2 September 2020; Revised 7 February 2021; Accepted 10 March 2021; Published 20 March 2021

Academic Editor: Nabil Benamar

Copyright © 2021 Jinsong Gui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Directional communication is helpful to improve the performance of millimeter Wave (mmWave) links. However, the dynamic nature of vehicular scenarios raises the complexity of directional mmWave vehicular communications. Also, a mmWave link is susceptible to blockages. Therefore, a mmWave vehicular communication system requires high environmental adaptability and context-awareness. Due to inadequate context information and insufficient beam settings in the existing related algorithm, it is difficult to pick out the set of beams with more reasonable widths and directions, which hinders the further promotion of network capacity in vehicular networks. Therefore, we propose an improved fast machine learning (IFML) algorithm to overcome this shortcoming. In order to improve network capacity while suppressing the additional beam search overhead, a partitioned search method is designed in the IFML. Also, in order to be robust to occasional fluctuations and timely adapt to significant changes in communication environments, the IFML adopts a flexible beam performance update approach based on adjustable weight coefficient. The simulation results show that the IFML significantly outperforms the existing related algorithm in terms of aggregate received data after a certain number of online learning time periods.

## 1. Introduction

As the Internet of Things (IoT) is applied to the automobile industries, the vehicles equipped with communication modules are networked to their surroundings and further connected to the Internet, which can improve user experience and promote safe driving. Such vehicular communicating modes as Vehicle to Network (V2N), Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), and Vehicle to Pedestrian (V2P) are collectively referred to as Vehicle to Everything (V2X). In Intelligent Transport System (ITS), the V2X applications include autonomous driving services, traffic management services and road safety, and infotainment V2X services.

To ensure Quality of Service (QoS) for these V2X applications, it is essential to provide adequate radio spectrum resources to increase network capacity, which can support the rapid exchange of information between infrastructure, vehicles, and pedestrians. Since radio spectrum resources are limited, the Long-Term Evolution V2X (LTE-V2X) sys-

tems and Dedicated Short-Range Communication (DSRC) systems mainly focus on the life-critical V2X applications (e.g., autonomous driving services and road safety management). Due to the controllable QoS and wide radio coverage, the LTE-V2X system outperforms the DSRC system in terms of providing connectivity for the V2X applications.

With the rapid growth of interconnected vehicles and the continuous pursuit of experience quality by vehicle users, higher requirements are put forward for network capacity. However, sub-6 GHz frequency bands used by the LTE system cannot meet the increasing capacity demand, so the Fifth-Generation (5G) communication system adopts the underutilized millimeter Wave (mmWave) frequency bands to make up for the shortfall. This underutilization is mainly caused by the penetration loss and high path loss of mmWave frequency bands. Fortunately, the existing researches [1, 2] indicate that (1) the high path loss is compensated by directional transmission and beamforming and (2) the short communication range in mmWave bands can be formed by higher deployment density of base stations.

The sub-6 GHz frequency bands have the wider coverage but the smaller capacity, while the mmWave frequency bands have the larger capacity but the narrower coverage. Therefore, the 5G systems and the LTE systems are widely considered to be the main driving forces for full support for all the V2X applications. However, the dynamic nature of vehicular scenarios raises the complexity of cellular-based directional mmWave vehicular communications. This is mainly because the directional communication needs an accurate beam alignment between a vehicle and a base station [3]. This alignment needs to be constantly adjusted due to vehicle mobility, which is not necessary for sub-6 GHz frequency bands since the omnidirectional transmission mode is adopted. Also, a mmWave link is susceptible to blockages (e.g., foliage, trucks, and buildings). Therefore, inaccurate beam selection raises high penetration loss and thus leads to the performance degradation of mmWave systems, since it does not guarantee effective avoidance of obstacles.

The related means and mechanisms like beamforming training, beam tracking, and beam selection are usually used to ensure beam alignment. With the increase of the number of beams and users, the search space of choosing a reasonable beam pair for communication link is increasing, which greatly increases the delay of establishing mmWave communication link. Because beam selection involves many communication parameters, pure mathematical modeling is very complicated. Although beam selection methods based on machine learning are widely accepted, most require training before they can be used. However, online learning-based beam selection methods can be used without prior training, so it is well suited to dynamic vehicular communication systems. In this respect, an excellent example is the work in the literature [4].

The authors in [4] argue that, through self-exploration, learning, and adaptation to communication environments, mmWave base stations can select beam accurately and maintain continuous scalability. Through modeling the beam selection as a contextual multi-armed bandit (MAB) problem, they propose a contextual online learning algorithm based on contextual information (i.e., a fast machine learning (FML) algorithm) to solve it. By this algorithm, any mmWave base station can autonomously learn from its previous decisions and the relationship to available contextual information, and thus, it knows the performance of every beam. Since this algorithm learns the expected beam performance in different contexts over time, it does not require a training phase. Here, the assumption is that a particular beam's performance is basically similar under similar vehicle contexts.

In [4], the beams are not overlapping, which limits the optional beam width and azimuth range. Although the different beams can serve the vehicles in their respective coverage at the same time, the same beam can only serve one of the vehicles in its coverage at any time since only unicast is considered. That is, only one vehicle is lucky to be served if there are more vehicles in the same fine place. For brevity, we call such a case the location conflict.

Therefore, in order to reduce the possibility that the vehicles will be excluded from the scope of simultaneous services due to the location conflict, we must get rid of the limitations

of beam width and azimuth range, which is helpful to increase the amount of data received by the vehicles in the system during each time period.

Currently, only a limited number of vehicles can be selected to receive data from a base station at a time. This is because a base station only has a limited number of radio frequency (RF) chains (including power amplifiers, low noise amplifiers, up and down converters, and RF switches), which limits a base station's ability to concurrently transmit data. Unlike a RF chain, in this paper, a beam is only viewed as a region that limits energy propagation and its shape is usually a cone (or sector), which we call a virtual beam for convenience.

Therefore, the number of virtual beams can be unlimited at a base station. If the overlapped virtual beams are allowed at a base station, they can be used concurrently by assigning the different frequency bands and RF chains to them, where the number of the different frequency bands is not more than that of RF chains. Due to the abundant spectrum resources in mmWave frequency bands, it is not difficult to provide these frequency bands. Under such conditions, the selected beams for concurrent usage have the more reasonable beam width and beam direction and, thus, serve for the more appropriate set of vehicles at the same time.

On the other hand, if we do not want to adopt the additional frequency bands, overlapping beams may be used simultaneously on the same frequency band, where the corresponding transmission powers have to be coordinated to control the interference between each other. However, it is not always the case that mutual interference is well controlled. If this case happens, overlapping beams cannot be used simultaneously in the same frequency band, and thus, its effect is reduced to that of any nonoverlapping beam scheme.

As the number of vehicles that want to be connected increases, it is time-consuming to select the right beam for each vehicle from the large number of virtual beams at a base station, especially when the transmission power of the RF chain allocated to a virtual beam also needs to be selected. When a virtual beam is assigned an RF chain and the RF chain's transmission power is specified, it is called a physical beam. In addition, the extension of the context dimension is beneficial to the construction of richer context decision space. The authors in [4] claim to consider only the small amount of vehicle arrival direction information, though they believe that richer contextual information will further improve system performance.

Finally, in [4], with the increasing times that a beam is used, the latest observed beam performance value contributes less and less to the current updated beam performance than the previous updated values. This update method is robust to occasional fluctuations in beam performance. However, when the latest observed performance of a beam reflects a significant change in its communication environment, the real-time performance of this update method is very poor. Therefore, to address the above challenges, we propose an Improved FML (IFML) algorithm and the main contributions are as follows.

- (1) By using the IFML, we can select the set of beams with the more reasonable beam widths and beam

directions. These selected beams can serve for the more appropriate set of vehicles at the same time during a time period, which is helpful to increase the amount of data received by vehicles

- (2) In the IFML, the richer contexts are exploited besides allowing beam overlap, which is helpful to pick out the desired beams. Also, the coverage of a base station is divided into multiple subcoverages to search for the desired beams from them, respectively. The used beams in a time period will come from the beams selected from these subcoverages. This partitioned search method reduces the beam search overhead
- (3) The beam performance update approach in the IFML is more flexible, which can be robust to occasional fluctuations and timely adapt to significant changes in communication environments by adjusting the weight coefficient in terms of the contribution of the latest observed value to the current updated beam performance value
- (4) Compared with the existing related algorithm, we demonstrate by means of the simulation results that the IFML substantially improves the amount of data received by the vehicles in the system at a slightly increased overhead. However, there is no difference in terms of beam performance update cost after a certain number of time periods

The rest of the paper is organized as follows. Section 2 gives an overview of related works. In Section 3, we describe the system model, including network architecture and problem formulation. In Section 4, we propose an improved algorithm, including theoretical analysis of regret. The algorithm performance analysis is provided in Section 5, while the simulation results are presented in Section 6. Section 7 concludes the paper.

## 2. Related Work

*2.1. V2X Communications and Architectures.* The authors of [5] proposed a context-awareness heterogeneous vehicle network architecture, which is compatible with the current European Telecommunications Standards Institute (ETSI) and International Organization for Standardization (ISO) standardized ITS station reference architectures. The architecture is particularly suitable for this kind of applications that connected vehicles require the different radio communication technologies to satisfy the requirements in diverse vehicular scenarios, since it has the ability for dynamic selection and configuration of communication profiles based on the context conditions and the application requirements.

The authors of [6] described the advantage of cellular infrastructure used for V2X applications alongside the architectures for cellular-assisted V2X systems and specifically discussed the security requirements of V2X in cellular networks. Besides the security challenges and requirements of V2X, the authors of [7] surveyed the key features of V2X and focused on the standardization techniques used for

vehicular communication technologies. The V2X entities include vehicles equipped with various radio interfaces (e.g., DSRC, LTE, and mmWave), V2X application servers, LTE base stations (evolved node B, eNodeB), and mmWave small base stations. They are connected by various short radio communication technologies (e.g., IEEE 802.11p [8], IEEE 1609.x [9], ETSI ITS-G5 [10], and 5G New Radio [11]) and long-range radio communication technologies (e.g., LTE-V2X [6]).

Due to only supporting short radio communication and adopting carrier sense multiple access with collision avoidance mechanism, the DSRC technology has small coverage and suffers from high collision probabilities under high vehicle density scenarios. Moreover, it has a limited bandwidth in the DSRC frequency spectrum and thus hardly meets the high traffic demand for V2V applications [9, 12]. The 3rd Generation Partnership Project (3GPP) release 14 [13] specifies that the LTE should both support V2I and V2N communications for latency tolerant V2X services and extend the current proximity service (ProSe) (e.g., Device-to-Device (D2D) communication) to implement direct V2V communications for delay-sensitive vehicular applications.

The authors of [14] envisaged the potential benefits of using LTE infrastructure for V2X services and the superiority of LTE-V2X technologies over DSRC-based approaches, including wide coverage capacity, high reliability, low latency, high data rate, and controlled QoS. The 3GPP release 16 will specify the advanced solutions for V2X communications in the 5G New Radio access technology [15], which can support ultrahigh reliability, ultrahigh data rates, and ultralow latency. Also, the 3GPP radio access network working group agreed that the 5G-based V2X system supplements LTE-based V2X system to provide advanced V2X services instead of replacing the services offered by LTE-based V2X system [16].

*2.2. mmWave Beam Management.* Beam selection problems have been researched in traditional vehicular networks which operate at sub-6 GHz frequency bands to achieve maximum throughput. As mentioned above, the propagation characteristics of radio signals in sub-6 GHz frequency bands are essentially different from those of mmWave frequency bands.

There are also studies on beam selection in mmWave networks. The authors of [17] addressed a hierarchical beam search problem, while the authors of [18] exploited the sparse multipath structure of mmWave channels to optimally select beam directions. The authors of [19] proposed a cell discovery method, where a base station periodically transmits synchronization signals to scan its entire angular space in time-varying random directions. Based on adaptive subspace sampling and hierarchical beam codebooks, the authors of [20] designed a beam alignment technique. The authors of [21] addressed the beam alignment issue in mmWave networks, where the transmitting and receiving antenna arrays require to frequently find the optimal beam pair. Therefore, when the number of transmitting and receiving beams is very large, the search space for beam matching is also large.

The authors of [22] proposed a new beamforming scheme to minimize the transmission power and weighted

symbol error rates, while the authors of [23] described a hybrid beamforming structure for small cell base station and formulated the joint allocation of the transmission power and beam into a mixed integer nonlinear programming problem. In [24], the authors proposed a codebook-based beam tracking scheme for mobile mmWave networks, which shows that the average tracking error probability can be lowered by improving the power allocation algorithm over the training beams.

The authors of [25] analyzed the average search delay of mmWave systems, where the average search times for discovering user equipment is taken as the number of search sectors in the beamspace. The authors of [26] proposed a deep learning-based beam selection, which exploits sub-6 GHz channel information to select a mmWave beam. The authors of [27] provided the overview on beam-forming architecture and deployment scenarios in the 5G New Radio (NR) standard. To address power-efficient beam design, the authors of [28] proposed a feasible point search method and then developed a hybrid analog-digital mapping algorithm. To lay the foundation for beam tracking, the authors of [29] used machine learning to track user mobility, in which a deep neural network is learned and then used to predict the direction of user movement.

To accelerate the beam training process, the authors of [30] proposed a low-cost hybrid architecture assisted by a digital beamforming module and designed a fast beam training method by utilizing the proposed architecture and using the sparsity of mmWave channels. To improve the latency performance of mmWave networks for industrial automation, the authors of [31] proposed an adaptive beam selection strategy to select the best set of beams among multiple users to reduce the overall latency for all users. The authors of [32] studied the beam alignment problem of multiuser mmWave massive Multi-input Multi-output (MIMO) system and proposed an alignment algorithm with partial beams by using machine learning. The authors of [33] proposed an efficient hierarchical beamforming training mechanism to establish directional links in dense mmWave cellular networks.

Given the severe interbeam interferences caused by mobility of users, the authors of [34] proposed a beam resource allocation and mobility-aware subband scheme for the mmWave subband-beam massive MIMO systems. The authors of [35] focused on beam training problem of wireless local area networks in the 60-GHz band and designed an advanced beamforming training protocol for multiple access points and multiple users. The authors of [36] applied per-beam synchronization for a massive MIMO system to alleviate the delay dispersion effects and channel Doppler due to mobility and blockage.

*2.3. Solutions of mmWave Vehicular Networks.* The above works are not specific to mmWave vehicular networks, and thus, the corresponding solutions cannot be directly applied to the field. For the characteristics of mmWave vehicular networks, the readers are encouraged to refer to the overview in [37]. The authors of [38] analyzed the mmWave propagation characteristics in V2V communication scenarios and derived the relationship between beam width and channel coherence

time in vehicular networks. The authors of [39–42] focused on the adaptability of mmWave beams in vehicular scenarios, and they believed that mmWave base stations can track vehicles by exploiting DSRC and adjusting the beam accordingly.

The authors of [43] utilized Kernel-based machine learning algorithms and channel state information of sub-6 GHz bands to predict vehicles' positions and then used them to preactivate target mmWave remote radio units. Also, they used K-nearest neighbor machine learning algorithms and historical handover data to predict handover decisions.

Since achieving perfect beam-alignment is challenging in vehicular scenarios, the authors of [44] proposed the method of fingerprint-based beam-alignment, where the fingerprint of a given location is viewed as a set of beam pairs. Based on the proposed multifingerprint database for a given location, the base station intelligently adapts the fingerprints with the aid of learning, which is capable of maintaining the target performance in mmWave vehicular environments.

The authors of [45] aimed to maximize the overall network throughput for highly-dynamic vehicular communications. They proposed a reinforcement learning approach to handle the beam selection problem in a high-mobility vehicular environment. The authors of [46] addressed optimal beam selection in mmWave-based V2X communications by designing the multitype-2 fuzzy inference system.

The authors of [4] addressed the beam selection issue at mmWave base stations in vehicular scenarios and proposed the FML algorithm. The FML is essentially an online learning algorithm, and it does not require either accurate localization information or prior statistical knowledge of dynamic vehicular scenarios. If the work in [21] is used to extend the FML to achieve longer connectivity, the additional learning cost is required for beam alignment, which increases the complexity of the FML and, thus, is unsuitable for vehicular scenarios.

In [47], the authors used a multiarmed bandit framework to develop online learning algorithms for beam pair selection and refinement. They used a type of sensor data (e.g., position) to identify the good mmWave beam directions, which is one approach to reduce the overhead since such information is widely available in vehicular applications. However, the work in [47] only deals with how to choose the more promising beam directions and refine them but does not focus on the optimal selection of beam width, where beams are spaced by the 3 dB beam widths in a fixed manner.

Although the work in [21] investigated an online stochastic optimization problem and proposed an equivalent structured multiarmed bandit model to exploit contextual information, it is difficult to adapt to vehicular scenarios since it requires accurate localization information. The work in [4] is most related to our work. However, for the reasons mentioned in the introduction, the FML still has room for improvement, which is the motivation for the work in this paper.

### 3. Network Model

*3.1. Network Architecture.* As shown in Figure 1, we consider an integrated mmWave/sub-6 GHz cellular system, where mmWave small base stations (mmSBSs) are deployed in the coverage area of an LTE eNode B (eNB). Each mmSBS is

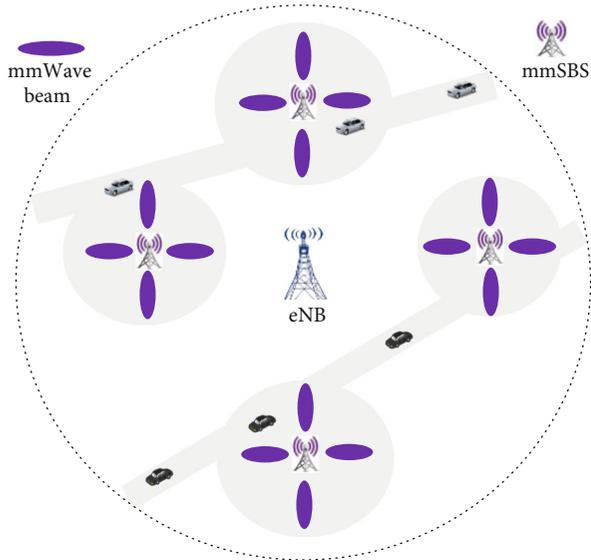


FIGURE 1: A mmWave vehicular communication scenario.

connected to the eNB via a backhaul link in a wired mode (e.g., an optical fiber) or a wireless mode (e.g., a mmWave link). The vehicles are equipped with two types of radio frequency interfaces (i.e., an LTE interface, which is used to connect to the eNB, and a mmWave interface, which is used for ultrahigh-speed data communication).

We assume that the eNB has the function that can assist each vehicle in its coverage to obtain the information of the mmSBS with which the vehicle should be associated. Since the LTE interface has the wider coverage, each vehicle can easily request help from the eNB via the LTE interface. If the eNB judges that there is not any appropriate mmSBS for the vehicle to associate with, it will reject the request, which will occur when the eNB judges that the vehicle is about to move out of its coverage. Even if the eNB agrees to handle a request that it should not have accepted, it can also transfer the request to an infrastructure component (e.g., a V2X server). With the help of the Home Subscriber Server (HSS) and Mobility Management Entity (MME) in the LTE infrastructure, this component will specify another suitable eNB to assist the vehicle.

In theory, each mmSBS can plan an infinite number of virtual beams, and the beam width of each beam can be set to the range from 0 to 360°. Therefore, the overlap between virtual beams should be allowed. In fact, if all the virtual beams of a mmSBS are used simultaneously, its potential transmission capacity can be multiplied. To this end, each virtual beam must actually have an exclusive RF chain. Moreover, if the virtual beams overlap each other, they must be assigned to different frequency bands. Since there are abundant bandwidth resources in mmWave frequency bands, it is not difficult to allocate different channels for overlapping virtual beams. If each mmSBS coverage area only shares the same frequency band in an underlay mode, the RF chains' transmission powers on overlapping beams must be reasonably selected to control mutual interference.

Due to the limitation of form-factor and manufacturing cost, it is undesirable that the number of RF chains is propor-

tional to the number of virtual beams when the number of virtual beams is set to a very large value. The literatures [2, 48] discuss the technical details on this limitation. Therefore, the number of vehicles that a mmSBS can simultaneously serve is limited by the maximum number of RF chains which a mmSBS can be equipped with. When the number of the virtual beams of a mmSBS and the number of the vehicles in its coverage exceed the number of its RF chains, respectively, it should select the best subset of virtual beams to provide the best downlink sum-data.

In order to achieve this goal, we model a virtual beam selection problem in a mmSBS as a MAB problem, which allows this mmSBS to identify the best virtual beams by carefully selecting subsets of virtual beams over time. Also, each selected virtual beam must have its reasonable RF transmission power to control mutual interference if the same frequency band is shared in an underlay mode.

As described in [49], a decision-maker must select a subset of actions with unknown expected earnings for the purpose of maximizing the earnings over time in MAB problems. The challenge in MAB problems is to address the exploration and exploitation dilemma, since all the actions should be fully explored to learn their earnings, but those which have already generated high earnings should also be exploited. Based on the contextual online learning idea, the authors in [4] address the abovementioned challenge. However, in our concerns, the online learning algorithm will have a larger search space since the number of possible virtual beams is larger, especially when the corresponding powers also need to be reasonably selected. Moreover, as the number of vehicles increases, the contextual space search overhead also increases accordingly. Therefore, we propose the IFML algorithm that tackles this problem.

**3.2. Problem Formulation.** We do not limit the number of virtual beams which a mmSBS can possess, since a larger virtual beamspace is more likely to find a more suitable beam for a vehicle. However, in order to keep this advantage while reducing the search time of online learning algorithm, we evenly divide the coverage area of a mmSBS into  $L$  sectors (e.g.,  $L = 4$  in Figure 2, that is, Sector 1, Sector 2, Sector 3, and Sector 4), where at most  $M^l$  virtual beams are set in the  $l$ th sector and  $l \in \{1, \dots, L\}$ . The virtual beams among different sectors do not overlap, while the virtual beams within the same sector are allowed to overlap.

The mmSBS can employ a finite set  $\mathcal{M}^l$  of  $M^l = |\mathcal{M}^l|$  virtual beams in the  $l$ th sector. For any sector, the mmSBS may simultaneously select a subset of at most  $n$  virtual beams to serve at most  $n$  ( $n < M^l, \forall l \in \{1, \dots, L\}$ ) vehicles, where  $n$  corresponds to the maximum number of RF chains at the mmSBS. However, considering all the sectors, the  $n$  virtual beams in the same sector are not necessarily the best  $n$  virtual beams. Therefore, the mmSBS should firstly select at most  $n$  best virtual beams from each sector separately to serve at most  $n$  vehicles in each sector, respectively, and then it reselects at most  $n$  best virtual beams from all the selected virtual beams to serve at most  $n$  vehicles in the entire coverage area of the mmSBS.

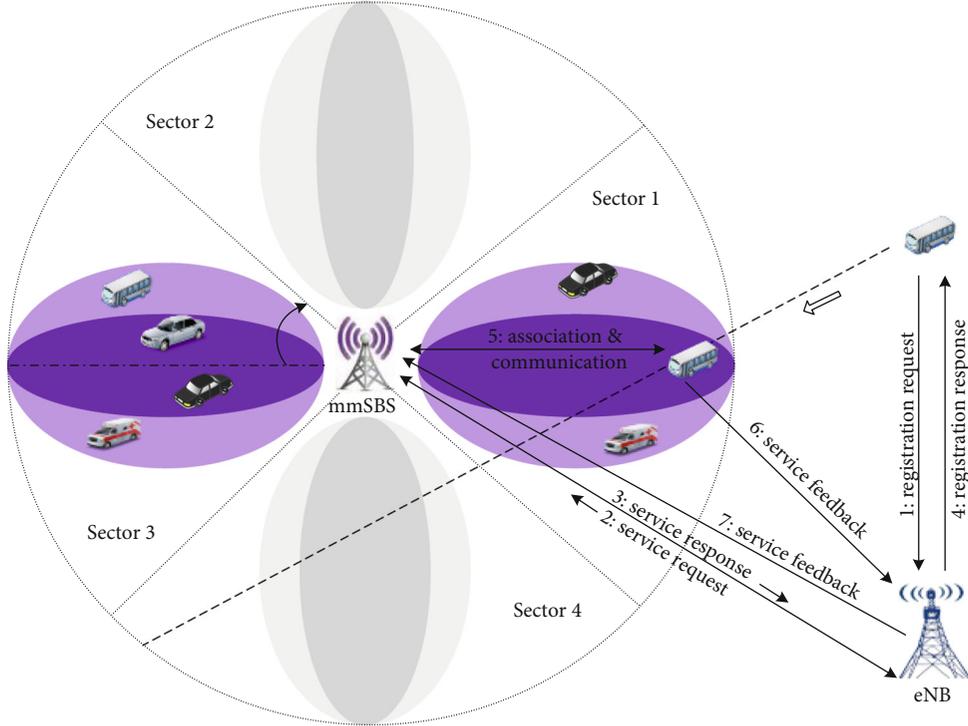


FIGURE 2: Integration feasibility of online learning method.

When a virtual beam is selected to limit the energy propagation area of the transmitted signal, the RF chain and frequency band are assigned to it. At this point, the selected virtual beam becomes the physical beam. For the sake of simplicity, we will not distinguish between a virtual beam and a physical beam in the following text.

We assume that the eNB can provide the necessary information to the mmSBS, which will be regarded as the vehicle context below. With the aid of the eNB, a vehicle will know the location of the mmSBS and the selected beam for it. Figure 2 shows how the information is communicated within the network. Since a vehicle keeps a continuous connectivity with the eNB via its LTE interface, it can send a registration request message (see “1: registration request” in Figure 2) to this eNB when it requires a mmWave connectivity between itself and a mmSBS, which contains the vehicle’s location and velocity.

Upon receipt of the registration request, the eNB sends a mmWave service request message (see “2: service request” in Figure 2) to a potential mmSBS after it makes a certain decision and judgment based on the vehicle’s location and velocity. This message contains the vehicle’s identifier in the cellular system, the identifier of the road on which the vehicle is travelling, and its expected direction of arrival at the mmSBS. By running the IFML algorithm, the mmSBS responds to the mmWave service request (see “3: service response” in Figure 2) with the information with respect to the selected beam. Next, the eNB forwards the related information about the mmSBS (i.e., the mmSBS’s location and its selecting beam) to the vehicle (see “4: registration response” in Figure 2).

Once reaching the coverage area, the vehicle starts a mmSBS associating process by sending an associating request, and then, it is replied by an associating response from the mmSBS, in which the vehicle measures the channel state information (CSI) from the associating response message and sends the CSI feedback for modulation and coding assignment. Next, the mmSBS enters the data transmission phase (see “5: association & communication” in Figure 2). When the data transmission phase is successful, the mmSBS will get acknowledgments of the transmitted data frames and thus any other feedback is not required. If a vehicle cannot detect the mmSBS within a selected beam, it will send the feedback to the eNB (see “6: service feedback” in Figure 2). This feedback will be forwarded to the mmSBS as a reference for future decisions (see “7: service feedback” in Figure 2).

Because the beam selection results need to be adjusted in time to serve the most appropriate set of vehicle users, we assume that the mmSBS adopts a discrete time setting. In other words, system time is discretized into equal time periods, which is denoted as  $t (t \in \{1, \dots, T\})$ . After each time period  $t$ , all the beam selection results will be updated. If a time period is shorter, the beam selection results are updated more timely, but it leads to more system overhead. Therefore, a reasonable tradeoff is a wise choice, and the specific value needs to be selected empirically. The detailed process of beam selection and update is described below.

- (1) At the first time slot of each time period  $t$ , a set  $\mathcal{U}_t^l = \{v_{t,i}^l \mid i \in \{1, \dots, V_t^l\}\}$  of vehicles will be registered to the  $l$ th ( $l \in \{1, \dots, L\}$ ) sector of the mmSBS via

the eNB, where  $V_t^l$  is the number of vehicles and satisfies  $V_t^l = |\mathcal{U}_t^l| > n$ .  $n$  is the maximum number of supported RF chains in the mmSBS, and it corresponds to the maximum number of vehicles that can simultaneously get downlink transmission service within the coverage area of the mmSBS. As the foregoing, during the registration, the mmSBS gets the information about the context  $c_{t,i}^l$  of each incoming vehicle  $v_{t,i}^l$ , which is formally regarded as a  $D$ -dimensional vector. Since the information about a vehicle can be described by using  $D$  context dimensions, it is encoded as a bit string with the fixed length in each dimension. After the first time slot of each time period  $t$ , the new context space is acquired by the mmSBS, which is represented as  $\mathcal{C}_{V_t^l} = \{c_{t,i}^l = \langle c_{t,i}^{l,1}, \dots, c_{t,i}^{l,D} \rangle | i \in \{1, \dots, V_t^l\}\}$ . In this paper, the context vector is two-dimensional (i.e.,  $D=2$ ), since we only consider the identifier of road and direction of arrival as the context for a vehicle

- (2) The mmSBS selects a subset of at most  $n$  best beams from the  $l$ th sector, where the set of selected beams in each time period  $t$  is denoted as  $\mathcal{B}_t^l = \{b_{t,j}^l | j \in \{1, \dots, n\}, n < M^l\}$ . Then, it reselects at most  $n$  beams from  $\bigcup_{i=1}^L \mathcal{B}_t^i$  to serve at most  $n$  vehicles in the entire coverage area of the mmSBS. Finally, at most  $n$  vehicles in  $\bigcup_{i=1}^L \mathcal{U}_t^i$  are informed about the selected beams by the associated eNB through using their LTE interfaces
- (3) When each selected vehicle (e.g.,  $v_{t,i}^l$ ) enters a mmSBS's coverage area, it receives data from this mmSBS and feeds back this situation to it. Therefore, the mmSBS will observe the amount of data  $r_{b_{t,j}^l}(c_{t,i}^l)$  that vehicle  $v_{t,i}^l$  successfully receives via the selected beams  $b_{t,j}^l$ , until the time period  $t$  is over or the vehicle is out of the reach of its beam

In general, the amount of data  $r_{b_{t,j}^l}(c_{t,i}^l)$  that a vehicle with context  $c_{t,i}^l \in \mathcal{C}_{V_t^l}$  can successfully receive from the mmSBS by using beam  $b_{t,j}^l \in \mathcal{B}_t^l$  during the time period  $t$  is a random variable that relies on the communication environment of the mmSBS (e.g., road conditions, vehicle routes, and blockages). The random variable  $r_{b_{t,j}^l}(c_{t,i}^l)$  is also regarded as the beam performance (i.e., the aggregate data received by the vehicle) of the beam  $b_{t,j}^l$  under the context  $c_{t,i}^l$ , which is usually bounded in  $[0, r_{\max}]$ , where  $r_{\max}$  is the maximum amount of data that can be received by a vehicle. The Shannon theorem and the contact time can be used to estimate  $r_{\max}$ . The Shannon theorem determines the maximum achievable data rate of the channel, which also relies on the selected modulation and coding scheme. The contact time is defined as the time during which mmSBS can transmit data to the vehicle, which is bounded by the coverage area of the beam and depends on beam width, beam direction, and vehicle speed.

The expected value of the random variable  $r_{b_{t,j}^l}(c_{t,i}^l)$  can be denoted as  $\mu_{b_{t,j}^l}(c_{t,i}^l)$ , which is also seen as the expected beam performance of the beam  $b_{t,j}^l$  under the context  $c_{t,i}^l$ . The purpose of the mmSBS's selecting a subset of beams is to maximize the expected received data at a subset of vehicles. That is, it aims at maximizing the sum of the expected beam performances. The optimal subset of beams in the time period  $t$  of the  $l$ th sector is denoted as  $\mathcal{B}_t^{l*}(\mathcal{C}_{V_t^l}) = \{b_{t,j}^{l*}(\mathcal{C}_{V_t^l}) | j \in \{1, \dots, n\}, n < M^l\}$ , which relies on  $\mathcal{C}_{V_t^l} = \{c_{t,i}^l = \langle c_{t,i}^{l,1}, \dots, c_{t,i}^{l,D} \rangle | i \in \{1, \dots, V_t^l\}\}$  and its  $n$  beams formally meet the following relational expression.

$$b_{t,j}^{l*} \in = \underset{\substack{b_{t,j}^l \in \mathcal{M}^l / \left( \bigcup_{k=1}^{j-1} \{b_{t,k}^l\} \right) \\ c_{t,i}^l \in \mathcal{C}_{V_t^l} / \left( \bigcup_{k=1}^{i-1} \{c_{t,k}^{l*}\} \right)}}{\operatorname{argmax}} \mu_{b_{t,j}^l}(c_{t,i}^l). \quad (1)$$

If the mmSBS knows the expected beam performance  $\mu_{b_{t,j}^l}(c_{t,i}^l)$  for each vehicle context  $c_{t,i}^l \in \mathcal{C}_{V_t^l}$  and each beam  $b_{t,j}^l \in \mathcal{M}^l$  in advance, it can be easy to select the optimal subset of beams for the set of incoming vehicles in the  $l$ th sector according to (1). For all the time periods, this will generate the amount of data expected to be received in total.

$$\sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n E \left[ r_{b_{t,j}^{l*}}(c_{t,i}^{l*}) \right] = \sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n \mu_{b_{t,j}^{l*}}(c_{t,i}^{l*}). \quad (2)$$

Unfortunately, the mmSBS does not understand the communication environment, so it must learn the expected beam performance  $\mu_{b_{t,j}^l}(c_{t,i}^l)$  over time for each vehicle  $v_{t,i}^l$  with the context  $c_{t,i}^l$ . The mmSBS must attempt different beams for different vehicle contexts over time to learn these values. Meanwhile, it should also ensure that those beams that have been proved to have good performance are fully utilized. Therefore, the mmSBS must seek out a trade-off between exploring the beams with the unknown performance and exploiting the beams with the proven high average beam performance.

In the following section, we will describe the IFML algorithm, where for each time period with incoming vehicles of contexts  $\bigcup_{i=1}^L \mathcal{C}_{V_t^i}$ , the best  $n$  beams are selected from  $\bigcup_{i=1}^L \mathcal{M}^i$ . The IFML's selection relies on the historical record of selected beams in previous time periods and the corresponding observed beam performance values. Given any set of the vehicles arriving in any contexts in the  $l$ th sector, the expected amount of received data is estimated as follows:

$$\sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n E \left[ r_{b_{t,j}^l}(c_{t,i}^l) \right] = \sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n E \left[ \mu_{b_{t,j}^l}(c_{t,i}^l) \right]. \quad (3)$$

The expected difference in the amount of data received by vehicles that are also observed by the mmSBS from the vehicles' feedback and that learned by the learning algorithm is

called the regret of learning. Based on (2) and (3), the regret of learning can be estimated by the following formula:

$$\begin{aligned} R(T) &= E \left[ \sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n \left( r_{b_{t,i}^*} \left( c_{t,i}^{l*} \right) - r_{b_{t,i}^l} \left( c_{t,i}^l \right) \right) \right] \\ &= \sum_{t=1}^T \sum_{i=1}^{V_t^l} \sum_{j=1}^n \left( \mu_{b_{t,i}^*} \left( c_{t,i}^{l*} \right) - E \left[ \mu_{b_{t,i}^l} \left( c_{t,i}^l \right) \right] \right). \end{aligned} \quad (4)$$

#### 4. The Improved Fml Algorithm

The proposed beam selection algorithm (i.e., the IFML algorithm) is run in each mmSBS. It first evenly partitions the context space into small sets of similar contexts in each of the sectors of a mmSBS and, then, learns about the individual performance of each beam in each of these small sets. Next, in each time period, the IFML goes through either an exploration process or an exploitation phase, which relies on the contexts of arrival vehicles and the preselected control function. For an exploration process, the IFML selects a random subset of beams. However, in an exploitation phase, it selects the beams that performed best during the previous time periods. Finally, by observing the amount of received data from the vehicles in the coverage area of the mmSBS, the IFML gets performance estimating values of selected beams. Therefore, it learns the performance of each beam under each vehicle context over time.

The pseudocode of the IFML algorithm is described in Algorithms 1–3. In the lines 1~5 of Algorithm 1, the IFML evenly partitions the context space  $\bigcup_{l=1}^L \mathbb{C}_{V_t^l}$  into  $L \cdot \mathbb{C}_T \cdot D$ -dimensional subspaces with the same size, where  $L$  and  $\mathbb{C}_T$  are the inputting values to the algorithm. Then, the IFML initializes each counter  $N_{b,s}^l(t)$  for each beam  $b \in \bigcup_{l=1}^L \mathcal{M}^l$  and each subspace  $s \in \bigcup_{l=1}^L \mathbb{C}_{V_t^l}$ . The purpose of setting these counters is to describe how many vehicles of a certain context have reached the mmSBS in previous time periods, in which the mmSBS had selected a certain beam. Moreover, the counter  $N_{b,s}^l(t)$  formally denotes the total number of vehicles with the context in subspace  $s$  that reached the mmSBS whenever beam  $b$  had been selected in any of the time periods  $1, \dots, t-1$  of any of the sectors  $1, \dots, L$ . The algorithm also initializes each estimator  $\hat{\mu}_{b,s}^l(t)$  for each beam  $b \in \bigcup_{l=1}^L \mathcal{M}^l$  and each subspace  $s \in \bigcup_{l=1}^L \mathbb{C}_{V_t^l}$ , which denotes the estimated performance of beam  $b$  for vehicles with the context in subspace  $s$ .

In a time period  $t$ , the IFML observes the contexts  $\bigcup_{l=1}^L \mathbb{C}_{V_t^l}$  of the  $\sum_{l=1}^L V_t^l$  incoming vehicles. For each context  $c_{t,i}^l$ , the IFML decides to which subspace this context belongs (lines 3-4 in Algorithm 2). That is, it seeks  $s_{t,i}^l \in \bigcup_{l=1}^L \mathbb{C}_{V_t^l}$  with  $c_{t,i}^l \in s_{t,i}^l$ . According to the set  $\bigcup_{l=1}^L \mathcal{H}_t^l := \bigcup_{l=1}^L \{s_{t,i}^l \mid i \in \{1, \dots, V_t^l\}\}$  of subspaces, the IFML computes the set  $\mathcal{B}_{\mathcal{H}_t^l}^{ue}(t)$  of underexplored beams (line 5 in Algorithm 2) by using the following formula:

$$\mathcal{B}_{\mathcal{H}_t^l}^{ue}(t) := \bigcup_{i=1}^{V_t^l} \left\{ b \in \mathcal{M}^l : N_{b,s_{t,i}^l}^l(t) \leq K(t) \right\}. \quad (5)$$

In (5),  $K : \{1, \dots, T\} \mapsto \mathbb{R}$  is a deterministic, monotonically increasing control function, which is used to determine whether to go through an exploration process or enter an exploitation phase. The control function  $K(t)$  needs to be selected adequately to guarantee that the IFML achieves a desired fine performance in terms of its regret. The Theorem 1 in [4] provides a suitable selection for the control function, which is also repeated below for the convenience of readers.

Theorem 1 (bound for  $R(T)$ ): for the  $l$ th sector, let  $K(t) = K(t) = t^{2\alpha/3\alpha+D} \log(t)$  and  $\mathbb{C}_T = \lceil T^{1/3\alpha+D} \rceil$ . If the IFML is run by using these parameters and if Assumption 1 holds true, the leading order of the regret  $R(T)$  is  $O(nV_t^l r_{\max} M^l T^{2\alpha+D/3\alpha+D} \log(T))$ .

The detailed proof of Theorem 1 can be found in [4] (although the parameters are somewhat different, they are essentially the same), where Assumption 1 is repeated below for the convenience of readers.

Assumption 1: there exist  $\alpha > 0$  and  $\beta > 0$  such that for all  $b \in \mathcal{M}^l$  and for all  $x, y \in \mathbb{C}_{V_t^l}$  in the  $l$ th sector, it holds that  $|\mu_b(x) - \mu_b(y)| \leq \beta \|x - y\|^\alpha$ , where  $\|\cdot\|$  denotes the Euclidean norm in  $\mathbb{R}^D$ .

When there exist underexplored beams, the IFML goes through an exploration phase (lines 6-14 in Algorithm 2). If the number  $u^l(t) := |\mathcal{B}_{\mathcal{H}_t^l}^{ue}(t)|$  of underexplored beams is at least  $n$ , the IFML randomly selects  $n$  beams from them. If the number  $u^l(t)$  of underexplored beams is smaller than  $n$ , the IFML selects all  $u^l(t)$  underexplored beams. Moreover, it selects the  $(n - u^l(t))$  beams  $\hat{b}_{1,\mathcal{H}_t^l}^l(t), \dots, \hat{b}_{n-u^l,\mathcal{H}_t^l}^l(t)$  from  $\mathcal{M}^l \setminus \mathcal{B}_{\mathcal{H}_t^l}^{ue}(t)$ , which meet the following formula:

$$\begin{aligned} \hat{b}_{j,\mathcal{H}_t^l}^l(t) &= \operatorname{argmax}_{b^l \in \mathcal{M}^l / \left( \mathcal{B}_{j,\mathcal{H}_t^l}^{ue}(t) \cup \bigcup_{k=1}^{j-1} \left\{ \hat{b}_{k,\mathcal{H}_t^l}^l(t) \right\} \right)} \mu_{b^l} \left( c_{t,i}^l \right) \\ &\quad c^l \in \mathbb{C}_{V_t^l} / \left( \bigcup_{k=1}^{j-1} \left\{ \hat{c}_{k,\mathcal{H}_t^l}^{l*}(t) \right\} \right) \end{aligned} \quad (6)$$

In (6),  $j = 1, \dots, (n - u^l(t))$ . If there are no underexplored beams, the IFML adopts an exploitation action (lines 15-17 in Algorithm 2), it selects the  $n$  beams  $\hat{b}_{1,\mathcal{H}_t^l}^l(t), \dots, \hat{b}_{n,\mathcal{H}_t^l}^l(t)$  from  $\mathcal{M}^l$ , which meet the following formula:

$$\begin{aligned} \hat{b}_{j,\mathcal{H}_t^l}^l(t) &= \operatorname{argmax}_{b^l \in \mathcal{M}^l / \left( \bigcup_{k=1}^{j-1} \left\{ \hat{b}_{k,\mathcal{H}_t^l}^l(t) \right\} \right)} \mu_{b^l} \left( c_{t,i}^l \right) \\ &\quad c^l \in \mathbb{C}_{V_t^l} / \left( \bigcup_{k=1}^{j-1} \left\{ \hat{c}_{k,\mathcal{H}_t^l}^{l*}(t) \right\} \right) \end{aligned} \quad (7)$$

In (7),  $j = 1, \dots, n$ . After selecting the  $n$  beams from each sector, respectively, the IFML will reselect at most  $n$  beams from all the selected beams of all the sectors (lines 8-18 in Algorithm 1). After beam reselection, the IFML observes

```

Input:  $T, L, \mathbb{C}_T, K(t)$ 
Output: the selected beams and the specified vehicles using the beams
1: For each  $l=1, \dots, L$  do
2:   Initialize context partition: Create partition  $\mathbb{C}_{V_i^l}$  into  $\mathbb{C}_T$  subspaces with same size.
3:   Initialize counters: for all  $b \in \mathcal{M}^l$  and  $s \in \mathbb{C}_{V_i^l}$ , set  $N_{b,s}^l = 0$ 
4:   Initialize estimators: for all  $b \in \mathcal{M}^l$  and  $s \in \mathbb{C}_{V_i^l}$ , set  $\hat{\mu}_{b,s}^l = 0$ 
5:   End for
6:   For each  $t=1, \dots, T$  do
7:     Invoke Algorithm 2 to get  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$  and  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
8:     If  $\mathcal{B}_{\mathcal{H}}^{ue}(t) \neq \emptyset$  then           ■ Exploration
9:        $u = \text{size}(\mathcal{B}_{\mathcal{H}}^{ue}(t))$ 
10:      If  $u \geq n$  then
11:        Select the  $n$  beams randomly from  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$ 
12:      Else
13:        Treat all the members of  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$  as the  $u$  beams
14:        Select the  $(n - u)$  best beams from  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
15:      End if
16:    Else           ■ Exploitation
17:      Select the  $n$  beams with the best performance from  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
18:    End if
19:    Invoke Algorithm 3 to update estimators and counters
20:  End for

```

ALGORITHM 1: Pseudocode of the IFML Algorithm.

```

Input: null
Output:  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$  and  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
1:  $\mathcal{B}_{\mathcal{H}}^{ue}(t) = \mathcal{B}_{\mathcal{H}}^{ve}(t) = \emptyset$ ;
2: For each  $l=1, \dots, L$  do
3:   Observe vehicle contexts  $\mathbb{C}_{V_i^l} = \{c_{i,i}^l \mid i \in \{1, \dots, V_i^l\}\}$ 
4:   Find  $\mathcal{H}_i^l = \{s_{i,i}^l \mid i \in \{1, \dots, V_i^l\}\}$  such that  $c_{i,i}^l \in s_{i,i}^l \in \mathbb{C}_{V_i^l}$ 
5:   Compute the set of under-explored beams  $\mathcal{B}_{\mathcal{H}_i^l}^{ue}(t)$  in (5)
6:   If  $\mathcal{B}_{\mathcal{H}_i^l}^{ue}(t) \neq \emptyset$  then           ■ Exploration
7:      $u^l = \text{size}(\mathcal{B}_{\mathcal{H}_i^l}^{ue}(t))$ 
8:     If  $u^l \geq n$  then
9:       Select  $\{b_{i,i}^l \mid i \in \{1, \dots, n\}\}$  randomly from  $\mathcal{B}_{\mathcal{H}_i^l}^{ue}(t)$ 
10:      Add  $\{b_{i,i}^l \mid i \in \{1, \dots, n\}\}$  to  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$ 
11:     Else
12:       Select  $\{b_{i,i}^l \mid i \in \{1, \dots, u^l\}\}$  as the  $u^l$  beams from  $\mathcal{B}_{\mathcal{H}_i^l}^{ue}(t)$  and then add them to  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$ 
13:       Select  $b_{i,u^l+1}^l, \dots, b_{i,n}^l$  as the  $(n - u^l)$  beams  $\hat{b}_{1,\mathcal{H}_i^l}^l(t), \dots, \hat{b}_{n-u^l,\mathcal{H}_i^l}^l(t)$  from (6) and then add them to  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
14:     End if
15:   Else           ■ Exploitation
16:     Select  $b_{i,1}^l, \dots, b_{i,n}^l$  as the  $n$  beams  $\hat{b}_{1,\mathcal{H}_i^l}^l(t), \dots, \hat{b}_{n,\mathcal{H}_i^l}^l(t)$  from (7) and then add them to  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ 
17:   End if
18: End for

```

ALGORITHM 2: Algorithm for getting  $\mathcal{B}_{\mathcal{H}}^{ue}(t)$  and  $\mathcal{B}_{\mathcal{H}}^{ve}(t)$ .

the beam performance of each selected beam for each vehicle within this time period (line 1 in Algorithm 3). According to these observations, the IFML updates its internal counters (lines 2-11 in Algorithm 3), where the weight coefficient  $\zeta$

represents the contribution of the recently observed beam performance to the updated beam performance and is usually determined by the empirical values in terms of the communication performance of the system.

```

Input: the selected beams
Output: the updated estimators and counters
1: The random variable  $r_{b_{t,i}^l}(c_{t,i}^l)$  is initialized to a negative number
2: Observe the received data of each vehicle  $v_{t,i}^l$  with the context  $c_{t,i}^l \in \mathcal{C}_{V_{t,i}^l}$ , in each beam  $b_{t,i}^l$ , and assign it to  $r_{b_{t,i}^l}(c_{t,i}^l)$ 
3: For each  $l=1, \dots, L$  do
4:   For each  $i=1, \dots, V_t^l$  do
5:     For each  $j=1, \dots, n$  do
6:       If  $r_{b_{t,i}^l}(c_{t,i}^l) \geq 0$  then
7:          $\hat{\mu}_{b_{t,i}^l, s_{t,i}^l}^l = (1 - \zeta) \cdot \hat{\mu}_{b_{t,i}^l, s_{t,i}^l}^l + \zeta \cdot r_{b_{t,i}^l}(c_{t,i}^l)$ 
8:          $N_{b_{t,i}^l, s_{t,i}^l}^l = N_{b_{t,i}^l, s_{t,i}^l}^l + 1$ 
9:       End if
10:     End for
11:   End for
12: End for

```

ALGORITHM 3: Algorithm for updating estimators and counters.

## 5. Algorithm Performance Analysis

For a mmSBS with coverage area partition, its online learning overhead is  $O(\bigcup_{l=1}^L (\mathcal{C}_T \cdot M^l))$ . However, For a mmSBS without coverage area partition, the corresponding overhead is  $O(L \cdot \mathcal{C}_T \cdot (\bigcup_{l=1}^L M^l))$ . Obviously, the overhead of the former is much smaller than that of the latter, especially when the number of the divided coverage areas (i.e., the value of  $L$ ) is larger.

When the number of beams at a mmSBS is larger and there are more context subspaces in its coverage area, it will take a longer time to learn the beam performance under each context subspace. In the scheme proposed in this paper, the corresponding learning overhead is  $O(\bigcup_{l=1}^L (\mathcal{C}_T \cdot M^l))$ , since coverage area partition is employed. The overhead savings are due to the neglect of a large number of unreasonable beam-context matching operations. For example, a beam obviously cannot be assigned to any vehicle in any context subspace outside its coverage, since such beam-context matching operation makes no sense.

When the number of beams at a mmSBS is not large and there are not many context subspaces in its coverage area, the coverage partition for a mmSBS is not necessary since it may not reduce overhead significantly. In fact, only if the number of beams is not limited, it is possible to set more types of beam widths and beam directions, and thus, it is more possible to select the beam with a more reasonable width and direction for each vehicle.

## 6. Performance Evaluation

**6.1. Simulation Scenario and Parameter Settings.** The IFML is evaluated via simulation experiments. The IFML is divided into two categories (i.e., the IFML-I and the IFML-II), depending on whether the overlapping beams use the different frequency bands or the same one. In the IFML-I, the overlapping beams use the different frequency bands, while they use the same one in the IFML-II.

In the following text, we first describe the simulation scenario and, then, introduce the mmWave channel propagation model that is used in our simulation experiments. Next, we describe the benchmark algorithms and performance metrics. Finally, we give the relevant simulation parameter settings according to the 3GPP technical specification in [50].

Our simulation scenario is shown in Figure 3, where the mmSBS's coverage is divided into the four sectors (i.e.,  $L = 4$ ) with the same size. For each sector, the two roads pass through it, where the location and size of the blockages on each road are fixed and account for about 10 percent of the road length. These blockages block the mmWave channels between the mmSBS and the vehicles on the roads.

There are seven types of beams according to the various beam width values (i.e., from  $30^\circ$  to  $90^\circ$  with the step size of  $10^\circ$ ) for each sector, where the number of beams per type is one. Therefore, the number of the mmSBS's beams in the  $l$ th sector is set to  $M^l = 7$  in the IFML-I, while  $M^l$  is  $7 |P|$  in the IFML-II. The set  $P$  is a set of discrete values with respect to transmission power of a RF chain, while  $|P|$  is the number of the members in the set  $P$ . For the IFML-II, the discrete power values are set from 0.1 to 1 Watts with the step size of 0.1, and thus,  $|P|$  is equal to 10.

In our implementation, a time period is defined by the mmSBS as a fixed length of time. During this time period, the IFML-I and the IFML-II receive registration information and, then, learn from the context and received data of the other vehicles passing through the selected beams. We select the identifier of the road and the direction of arrival as context. Therefore, the context vector is a two-dimensional vector (i.e.,  $D = 2$ ). The arrival direction of a vehicle is defined as the angle between the line connecting this vehicle with a mmSBS and the positive  $x$ -axis in the plane coordinate system with this mmSBS as the origin.

The length of a time period  $t$ , the parameter  $\alpha$ , and the number of two-dimensional subspaces in each sector of the mmSBS are set to  $t = 3$  s,  $\alpha = 0.33$ , and  $\mathcal{C}_T = 10$ , respectively. Therefore, based on Theorem 1, the time horizon

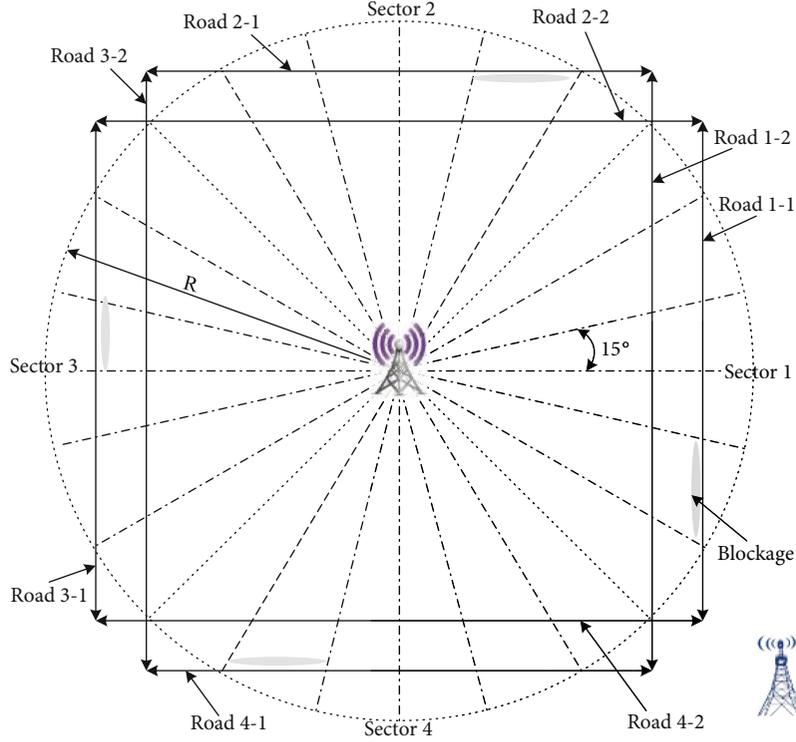


FIGURE 3: Illustration of our simulation scenario.

$T$  is about 1000, while the value of control function  $K(t)$  is about 2.02

When the  $n$  beams with the best performance are selected, the contexts in which these beams match are also determined at the same time. If the number of vehicles in such contexts is greater than the number of the selected beams, a strategy for vehicle selection needs to be determined in advance. For example, depending on the order in which vehicles enter a context space, a First-Come-First-Served (FCFS) policy can be used, or a First-Come-Last-Served (FCLS) policy can be adopted. In addition, a Random Selection (RASE) policy can also be employed.

When the vehicles under the specific beam coverage in the  $l$ th sector are counted as the members of  $V_t^l$  in the first time slot of the  $t$ th time period, every counted vehicle has entered this coverage area before the beginning of this time slot but does not leave before the end of this time slot. So the earlier they enter, the earlier they may leave.

The above simulation scenario is implemented on the OMNeT platform (i.e., omnetpp-5.4.1) with the application package (i.e., veins-4.7.1) in terms of vehicle communication simulation framework, where the traffic pattern is generated by the third-party software: simulation of urban mobility (i.e., sumo-0.30.0). The speed of vehicles is set to the range from 5 to 10 m/s.

The following mmWave channel propagation model is adopted in our simulation experiments:

$$p_i^r = p_i^t G_i^t G_i^r G_i^c. \quad (8)$$

In (8),  $p_i^t$  is the transmission power at a directional beam

from the mmSBS to the vehicle  $i$ , while  $p_i^r$  is the received power at the vehicle  $i$  side when the mmSBS transmits at the power  $p_i^t$ ;  $G_i^t$ , and  $G_i^r$  are the directional transmitting gain and directional receiving gain, respectively;  $G_i^c$  is the channel gain between the mmSBS and the vehicle  $i$ . When the beam between the mmSBS and the vehicle  $i$  is aligned,  $G_i^t$  and  $G_i^r$  can be estimated by the following formula [51]:

$$\begin{cases} G_i^t = \frac{2\pi - (2\pi - \theta_i^t)\partial}{\theta_i^t}, & (9a) \\ G_i^r = \frac{2\pi - (2\pi - \theta_i^r)\partial}{\theta_i^r}, & (9b) \end{cases} \quad (9)$$

In (9),  $\theta_i^t$  is the beam width of the transmitter, while  $\theta_i^r$  is the beam width of the receiver;  $\partial$  is the gain of the side lobe and  $0 < \partial \ll 1$ , while  $\theta_i^t$  and  $\theta_i^r$  are the main lobe in radian. The channel gain  $G_i^c$  can be estimated by the following formula [52]:

$$G_i^c = |\chi_i^c \delta(\tau - \tau_i)|^2. \quad (10)$$

In (10),  $\delta(\cdot)$  is the Dirac delta function;  $\tau_i$  and  $\chi_i^c$  are the propagation delay and the amplitude of the path from the mmSBS to the vehicle  $i$ , respectively.  $\tau_i$  is estimated by the following formula:

$$\tau_i = \frac{d_i}{c}. \quad (11)$$

In (11),  $d_i$  is the distance of the path from the mmSBS to the vehicle  $i$ , and  $c$  is the speed of light. When there is a line-of-sight (LOS) path between the mmSBS and the vehicle  $i$ , the amplitude is estimated by the following formula [52]:

$$\chi_i^{c,\text{los}} = \frac{\lambda}{4\pi d_i}. \quad (12)$$

In (12),  $\lambda$  is the wavelength, and  $\lambda = c/f_c$ ,  $f_c$  is the carrier frequency. When there is not any LOS path between the mmSBS and the vehicle  $i$ , the amplitude is related to both path loss and reflection coefficients. Due to very high reflection loss in mmWave band [53], only one reflection of a given path is considered, which is estimated by the following formula [52]:

$$\chi_i^{c,\text{nlos}} = \frac{\lambda\Gamma}{4\pi d_i}. \quad (13)$$

In (13),  $\Gamma$  is the reflection coefficient of the mmWave reflection path. Therefore, based on (12) and (13),  $\chi_i^c$  is estimated by

$$\chi_i^c = \begin{cases} \chi_i^{c,\text{los}}, & \text{In unblocked area} & (14a) \\ \chi_i^{c,\text{nlos}}, & \text{In blocked area} & (14b) \end{cases}. \quad (14)$$

In the IFML-II, when the overlapping beams use the same frequency band, there is the same frequency interference between them. Let the beam of the vehicle  $i$  and the beam of the vehicle  $j$  be overlapping in the same mmSBS coverage. When the mmSBS transmits data to vehicle  $j$ , the interference power received at vehicle  $i$  should be

$$p_{j \rightarrow i}^r = p_j^t G_{j \rightarrow i}^t G_{j \rightarrow i}^r G_i^c. \quad (15)$$

In (15),  $G_{j \rightarrow i}^t$  and  $G_{j \rightarrow i}^r$  are the directional transmitting gain and directional receiving gain of the mmWave link between the mmSBS and vehicle  $i$  when the mmSBS transmits data to vehicle  $j$ . The directional transmit-receive gain of each mmWave link can be derived by the following formula:

$$G_{j \rightarrow i}^t G_{j \rightarrow i}^r = \begin{cases} \frac{2\pi - (2\pi - \theta_j^t)\partial}{\theta_j^t} \cdot \frac{2\pi - (2\pi - \theta_i^r)\partial}{\theta_i^r} & \text{Condition A} \\ \frac{2\pi - (2\pi - \theta_j^t)\partial}{\theta_j^t} \cdot \partial & \text{Condition B} \\ \partial \cdot \frac{2\pi - (2\pi - \theta_i^r)\partial}{\theta_i^r} & \text{Condition C} \\ \partial \cdot \partial & \text{Condition D} \end{cases}. \quad (16)$$

Let  $\phi_{j \rightarrow i}^t$  and  $\phi_{j \rightarrow i}^r$  be the beam offset angle from the mmSBS's (the mmSBS transmits to vehicle  $j$ ) transmitting beam direction to the position of vehicle  $i$  and that from vehi-

cle  $i$ 's (vehicle  $i$  receives from the mmSBS) receiving beam direction to the position of the mmSBS, respectively; Condition A is both  $|\phi_{j \rightarrow i}^t| \leq \theta_j^t/2$  and  $|\phi_{j \rightarrow i}^r| \leq \theta_i^r/2$ ; Condition B is both  $|\phi_{j \rightarrow i}^t| \leq \theta_j^t/2$  and  $|\phi_{j \rightarrow i}^r| > \theta_i^r/2$ ; Condition C is both  $|\phi_{j \rightarrow i}^t| > \theta_j^t/2$  and  $|\phi_{j \rightarrow i}^r| \leq \theta_i^r/2$ ; Condition D is both  $|\phi_{j \rightarrow i}^t| > \theta_j^t/2$  and  $|\phi_{j \rightarrow i}^r| > \theta_i^r/2$ .

The work in [4] is most similar to ours. However, in the implementation in [4], the context vector is one-dimensional since the only direction of arrival is considered as the context. Also, a time period is defined as the time during which the observed vehicle enters and leaves the cell coverage area. Since this type of time period is influenced by a variety of factors (e.g., speed of vehicle, size of cell coverage area, and road orientation), it has great uncertainty and thus is difficult to determine its value in practical applications.

Furthermore, the direction of arrival in [4] only takes some simple and abstract values (i.e., north, south, east, and west) due to the consideration for wide adaptation. Therefore, in order to make a fairer comparison, we design a variant of the algorithm (i.e., FML) in [4] and call it VFML, which basically retains the core idea of the FML algorithm except for the fixed time period (i.e.,  $t = 3$  s) and the adoption of vehicle arrival direction defined in our scheme.

Based on the same bound for  $R(T)$  as that in the IFML, the time horizon  $T$  in the VFML should be about 1000. In addition, based on our simulation scenario in Figure 3, the simulation area for the VFML is divided into 24 one-dimensional subspaces according to the arrival direction of a vehicle. Therefore, the other parameters in the VFML should be set to  $\alpha = 0.39$  and  $K(t) = 2.35$ . Since the maximum number of supported RF chains in a mmSBS is set to six and the beams do not overlap in the VFML, the six orthogonal beams with variable beam width from  $30^\circ$  to  $60^\circ$  cover the  $360^\circ$  azimuth.

In addition, we also consider an Optimal Scheme (OPSC) as a comparison of our scheme, which has a priori knowledge about the expected beam performance  $\mu_{b_{i,j}^l}(c_{t,i}^l)$  of each beam  $b_{i,j}^l \in \mathcal{B}_t^l$  in each context  $c_{t,i}^l \in \mathcal{C}_{V_i^l}$ . In each period, the OPSC selects the optimal subset  $\mathcal{B}_t^{l*}(\mathcal{C}_{V_i^l})$  of  $n$  beams as in (1). Therefore, the results achieved by the OPSC is the expected performance upper bound of the system.

The performance metrics used in the evaluation are the aggregate received data, the cumulative received data, and the online learning cost. We define the aggregate received data as the received data (in Gbits) from all the vehicles during a time period  $t$ , while we define the cumulative received data as the received data from all the vehicles during the time horizon  $T$ . The online learning cost is defined as the number of exploration rounds required for the IFML-I, the IFML-II, and the VFML to reach a certain percentage of the OPSC's performance, where all the exploration operations during a time period are classified as one round. The values of the remaining simulation parameters are listed in Table 1, unless otherwise stated in the following text.

**6.2. Simulation Results and Analysis.** Firstly, we evaluate the performance metrics of the IFML-I and the IFML-II

TABLE 1: The simulation parameters.

Symbol	Description	Value
$f_c$	Carrier frequency	60 GHz
$W$	System bandwidth	1 GHz
$P_{\max}^t$	The maximum transmission power of a mmSBS	30 dBm
$N_0$	Noise power density at a vehicle	-174 dBm/Hz
$BW^r$	Each vehicle's receiving beam width	90°
$BW^t$	A mmSBS's transmitting beam width	30°~90°
$n$	The maximum number of supported RF chains in a mmSBS	6
$\Gamma$	The reflection coefficient of mmWave reflection path	0.3
$t$	The length of a time period	3 s
$T$	The number of time periods	1000
$\zeta$	The weighted coefficient	0.125
$R$	Radius of simulation area of a mmSBS	100 m

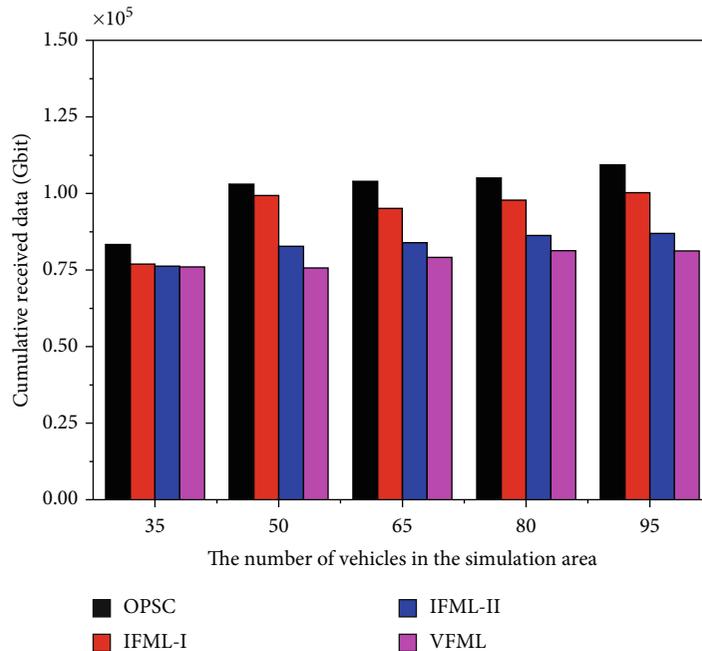


FIGURE 4: Impact of vehicle density on cumulative received data.

compared to the benchmark algorithms such as the VFML and the OPSC. The corresponding results are shown in Figures 4–11. When there are multiple vehicles in the coverage area of a selected beam at the same time, the VFML selects one of the vehicles at random for communication. Therefore, in this set of simulations, the IFML-I, the IFML-II, and the OPSC also adopt the RASE strategy. Next, we evaluate the performance metrics of the IFML-I in terms of the FCFS strategy, the FCLS strategy, and the RASE strategy. The corresponding results are shown in Figures 12–19. Finally, we compare the impact of learning information space size on the performance metrics of the IFML-I and the VFML, respectively. The corresponding results are shown in Figures 20–23.

In Figure 4, we investigate the impact of the number of vehicles in the simulation area on the cumulative received data achieved by the different algorithms, where at most 6 selected beams are used simultaneously in each time period and the number of vehicles varies from 35 to 95 with the step size of 15. We can see from Figure 4 that as the number of vehicles in the simulation area increases, the cumulative received data also increase. This is because the smaller number of vehicles means the more insufficient context information and thus results in a poorer learning effect. Therefore, the probability of accurately selecting the beams that can make vehicles maximize the cumulative received data is relatively small. With the increase of the number of vehicles, more context information is provided, which is conducive

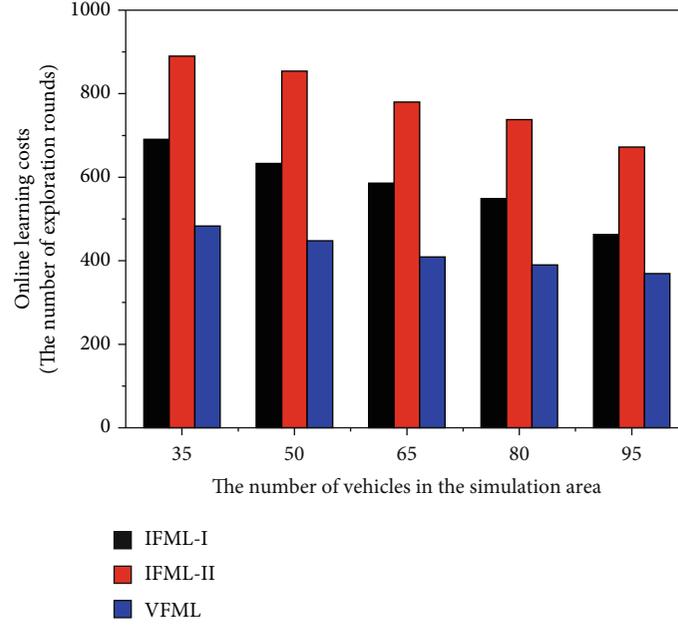


FIGURE 5: Impact of vehicle density on online learning costs.

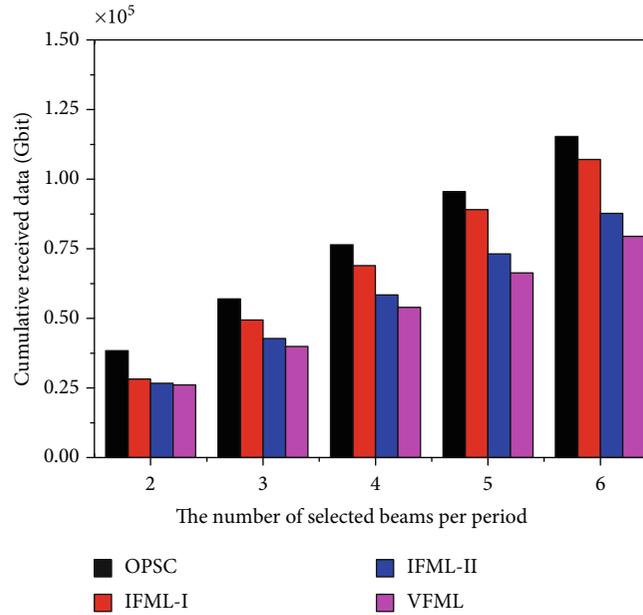


FIGURE 6: Impact of the number of selected beams per period on cumulative received data.

to improving the learning effect. Therefore, the probability of accurately selecting the beams that can make vehicles maximize the cumulative received data is relatively large.

From Figure 4, we also see that the cumulative received data hardly increase when the number of vehicles in the simulation area increases to a certain amount. This is because the number of selected beams per time period is fixed. At the same time, when the context information is large enough, the specified number of served vehicles can be selected reasonably through online learning. Figure 4 shows that the IFML-I is closer to the optimal one (i.e., the OPSC) than the IFML-II and the VFML, where the IFML-II outperforms

the VFML. The main reason lies in two aspects. On the one hand, the IFML-I and the IFML-II can find the more reasonable beam width and beam direction for each served vehicle when compared with the VFML. On the other hand, the two IFMLs' two-dimensional context can bring more accurate information than the VFML's one-dimensional context, which can make a more reasonable decision for each vehicle. Compared with the IFML-II, the IFML-I pays more frequency band resources for better performance.

Based on the same simulation configuration as that in Figure 4, the online learning costs required for the IFML-I, the IFML-II, and the VFML under the different number of

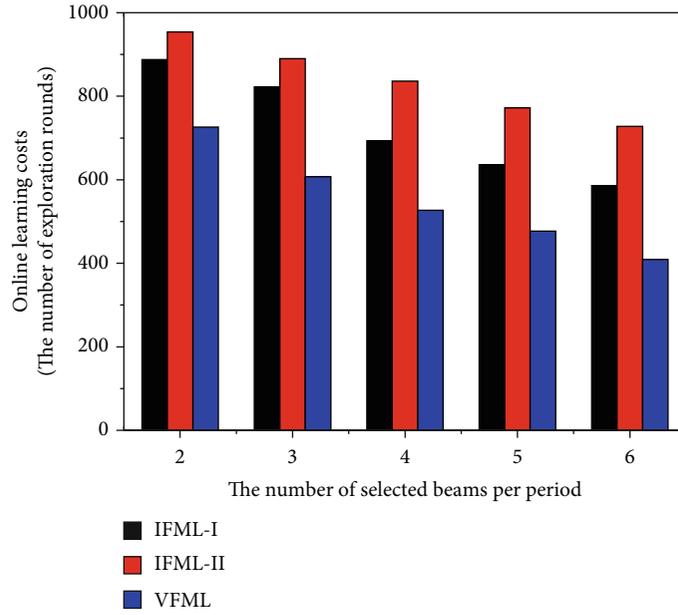


FIGURE 7: Impact of the number of selected beams per period on online learning costs.

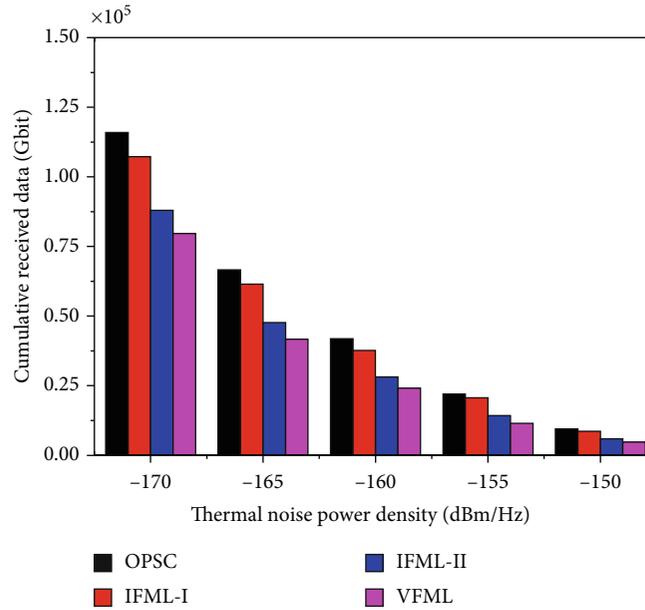


FIGURE 8: Impact of thermal noise power density on cumulative received data.

vehicles are shown in Figure 5. We can observe from Figure 5 that the changes of the online learning costs result from the number of vehicles in the system. As the number of vehicles increases, the costs increase as well. This is because the more vehicles that enter the system in a time period will give each beam the opportunity to probe the more context subspaces. Therefore, the performances of each beam on the more subspaces can be detected in a time period. If there is no performance history or the recorded history is insufficient, the corresponding beams are scheduled for exploration as soon as possible in a time period, which is helpful to speed up the detection process of the performance of each beam

in each context subspace, and thus reduce the number of exploration rounds.

Moreover, we can see from Figure 5 that the two IFMLs' online learning costs are larger than those of the VFML. This is because the two IFMLs have more beams and context subspaces than those of the VFML. By setting the more number of beams and using the more context information, each selected vehicle can be assigned the beam with a more reasonable width and direction. However, there are larger combinations of beams and contexts that need to be explored. As the number of vehicles increases, the difference in the number of exploration rounds between the two

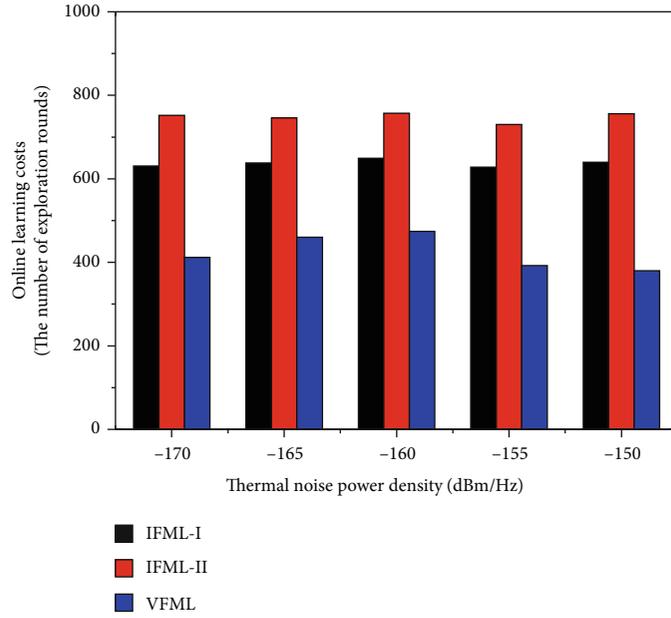


FIGURE 9: Impact of thermal noise power density on online learning costs.

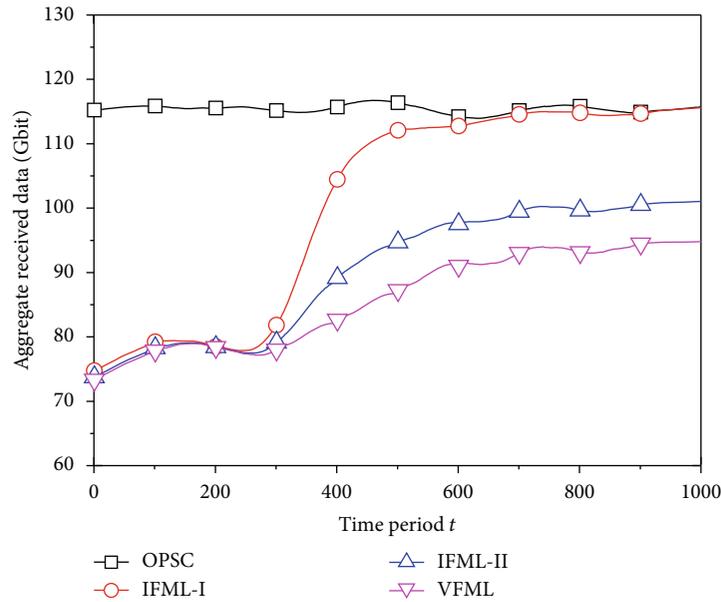


FIGURE 10: Aggregate received data over the time horizon.

schemes is narrowed. This indicates that the operation mode (that the two IFMLs first perform a partitioned search for a mmSBS coverage area and then select the exploration beams from the partitioned search results) has higher exploration efficiency with the increase of the number of vehicles. In the IFML-II, in order to control cochannel interference, the beam transmission power must be adjusted, so it needs to be discretized into various values, which makes the beam exploration space be larger and the online learning cost be higher.

In Figure 6, we analyze the impact of the number of selected beams per time period on the cumulative received data, where the number of vehicles in the simulation area is

set to 65 and the number of selected beams varies from 2 to 6 with the step size of 1. It can be seen from Figure 6 that as the number of simultaneously selected beams increases, the cumulative received data in the system increase as well. The reason for this increase is that the increasing number of beams used in a concurrent manner leads to the enhanced coverage area and thus the more vehicles can be served at the same time. However, as mentioned earlier, the larger the number of concurrent beams is, the higher the hardware complexity and energy consumption at a mmSBS are. It can also be seen that the two IFMLs achieve a larger cumulative received data than the VFML, where the explanation for the

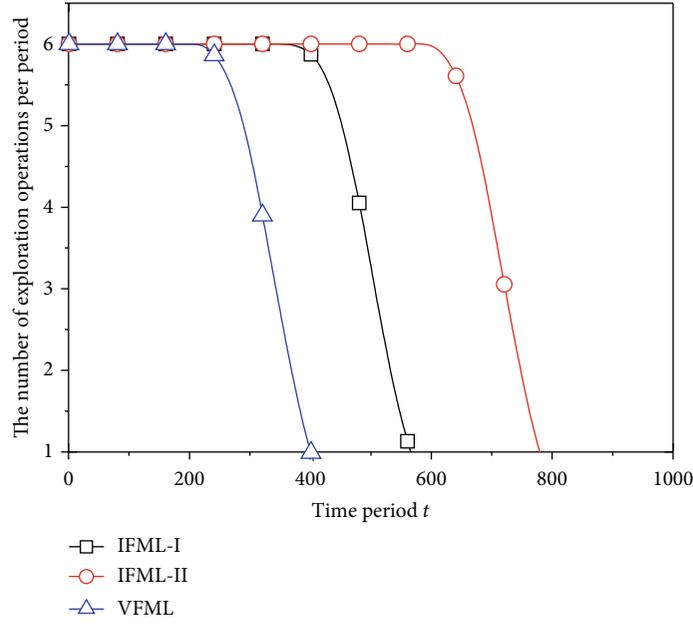


FIGURE 11: The number of exploration operations per period over the time horizon.

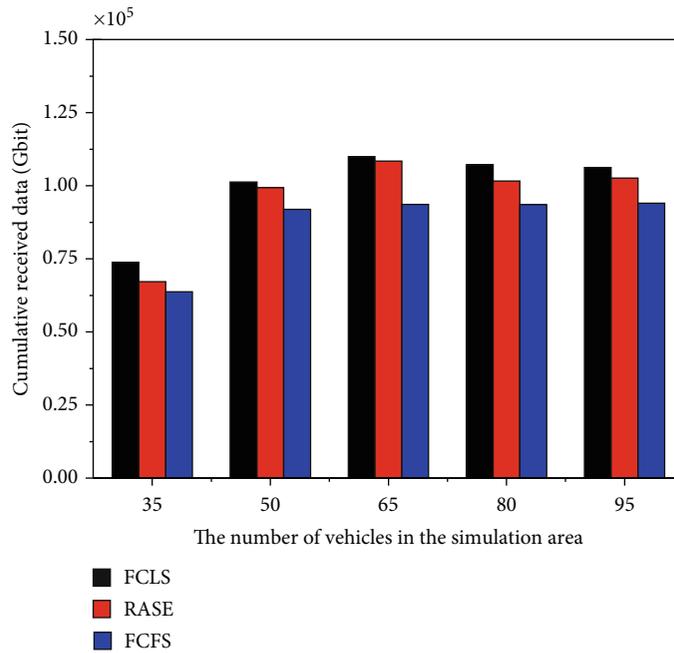


FIGURE 12: Impact of vehicle density on cumulative received data under the three vehicle scheduling strategies.

difference in terms of cumulative received data between the different schemes is similar to the explanation of the results in Figure 4.

Based on the same simulation configuration as that in Figure 6, the online learning costs required for the two IFMLs and the VFML under the different number of selected beams per period are shown in Figure 7. From Figure 7, as the number of selected concurrent beams increases, the number of exploration rounds decreases. The larger number of selected

concurrent beams per time period means that the larger number of beams with unknown or uncertain performance can be explored in each time period. When the combined space size of beams and context subspaces remains unchanged, the number of required exploration rounds will decrease. In Figure 7, we see that the online learning costs of the two IFMLs are larger than that of the VFML, and the explanation of this difference is similar to the explanation of the results in Figure 5.

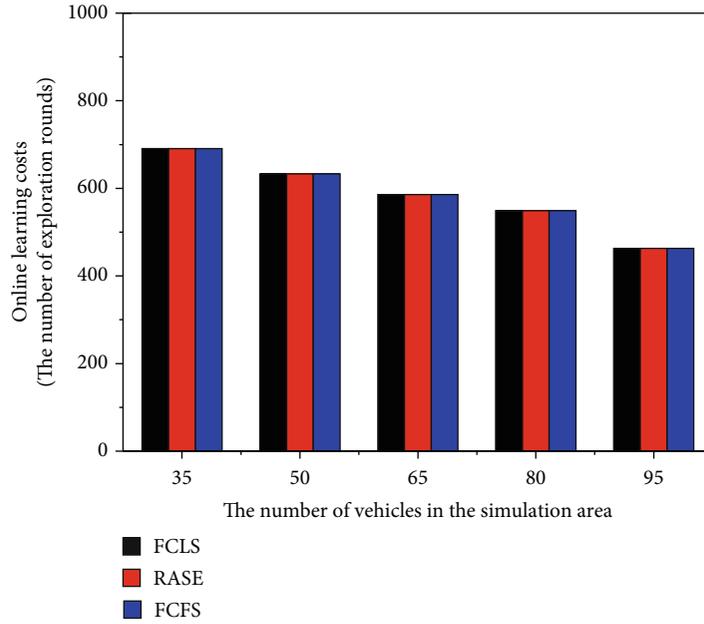


FIGURE 13: Impact of vehicle density on online learning costs under the three vehicle scheduling strategies.

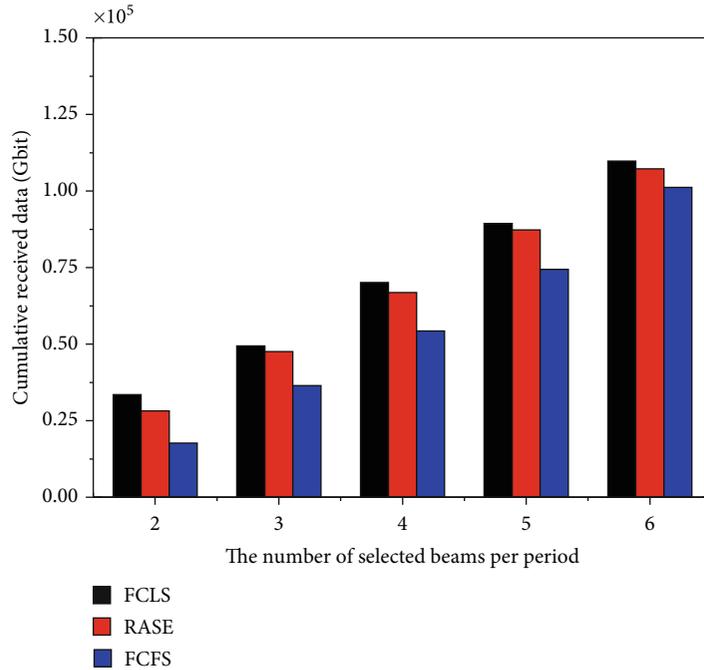


FIGURE 14: Impact of the number of selected beams per period on cumulative received data under the three vehicle scheduling strategies.

In Figure 8, we analyze the impact of thermal noise power density on the cumulative received data, where the number of vehicles in the simulation area is set to 65, the number of selected beams is set to 6, and the thermal noise power density varies from -170 dBm/Hz to -150 dBm/Hz with the step size of 5 dBm/Hz. Obviously, as the thermal noise power density in the system increases, the cumulative received data decrease. This is because the increase of thermal noise power density will reduce the signal-to-noise ratio. When the transmission power of a mmSBS is constant, the data rate will

decrease according to the Shannon theorem. Also, we can see from Figure 8 that the cumulative received data of the two IFMLs are higher than that of the VFML, where the explanation for this difference is similar to the explanation of the results in Figure 4.

Based on the same simulation configuration as that in Figure 8, the online learning costs required for the two IFMLs and the VFML under the different thermal noise power density are shown in Figure 9. It is observed from Figure 9 that the impact of the thermal noise power density

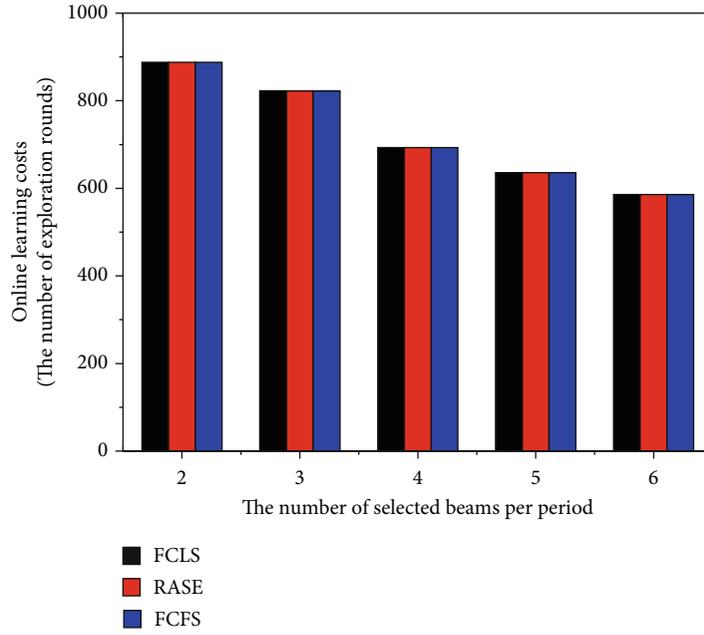


FIGURE 15: Impact of the number of selected beams per period on online learning costs under the three vehicle scheduling strategies.

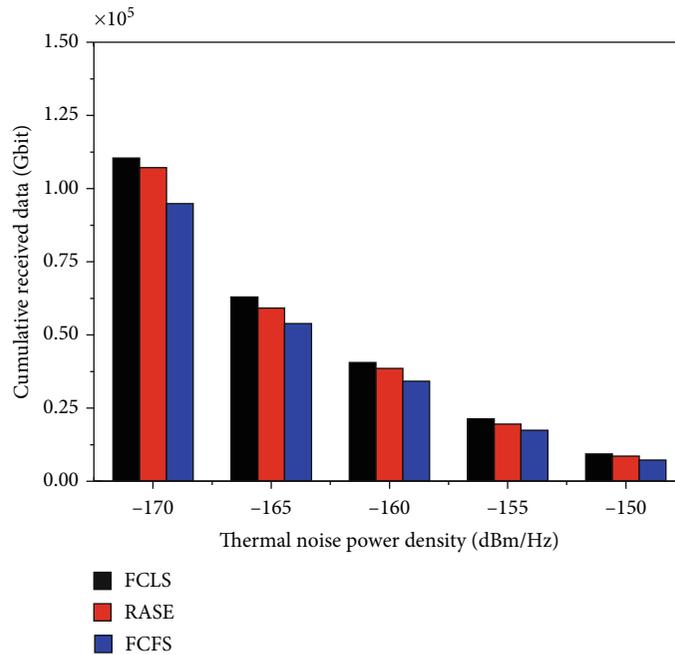


FIGURE 16: Impact of thermal noise power density on cumulative received data under the three vehicle scheduling strategies.

on the online learning costs of the algorithms is not obvious in general. The reason is that the variation of thermal noise power density does not change the size of the information space of online learning.

However, at higher thermal noise density, the number of the VFML’s beam exploration rounds decreases slightly. As shown in Figure 8, the thermal noise power density has a great impact on the cumulative received data. Therefore, the possible reason is that, due to the combination of the

high thermal noise environment and the small number of optional beams in the VFML, the performance of each beam under each context subspace is not easily affected by the changes of other factors in the environment, and thus, a beam performance is considered stable and does not require additional exploration.

In Figure 10, we analyze the aggregate received data achieved by the algorithms over the time horizon with 1000 time periods, where the number of vehicles in the simulation

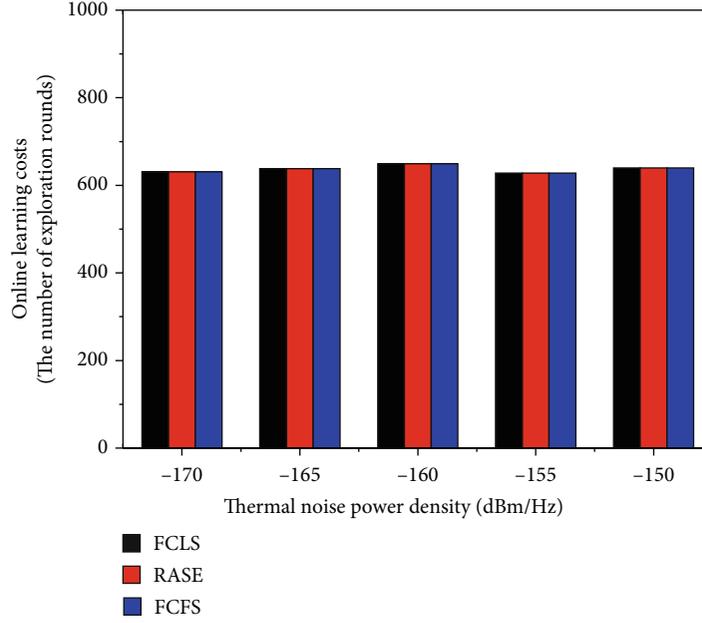


FIGURE 17: Impact of thermal noise power density on online learning costs under the three vehicle scheduling strategies.

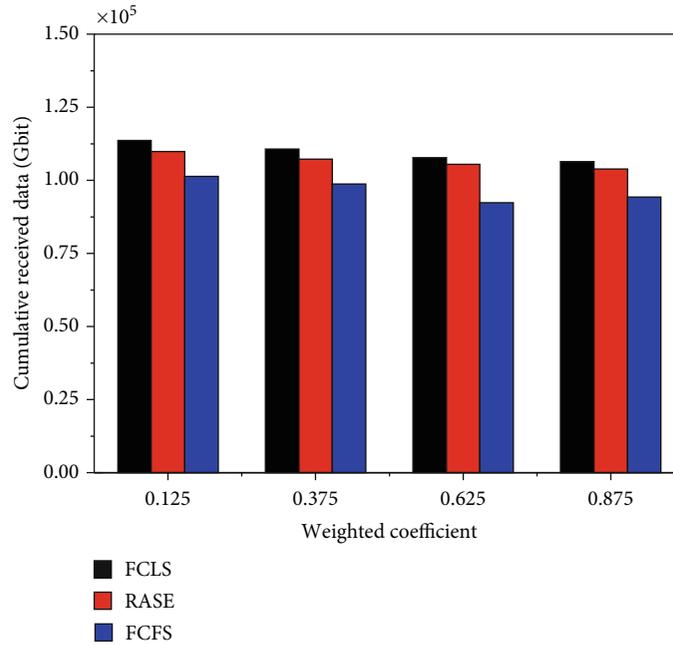


FIGURE 18: Impact of weight coefficient  $\zeta$  on cumulative received data under the three vehicle scheduling strategies.

area is set to 65. Figure 10 illustrates that the IFML-I can achieve performance closer to the OPSC than the IFML-II and the VFML. The fluctuations on the graph result from the number of vehicles in the system, the speed of each vehicle, and the contact time. Clearly, the OPSC gives an upper bound to the other algorithms due to a priori knowledge of the expected beam performance. So, the OPSC can select the optimal subset of beams according to the relational expression (1).

The two IFMLs obviously outperform the VFML in terms of aggregate received data after 300 time periods, and the IFML-I is close to the OPSC's aggregate received data after 600 time periods. The main reason is that after sufficient online learning, the two IFMLs can select a set of beams with more reasonable widths and directions than the VFML, which can provide greater transmission capacity for the vehicles in their coverage. However, before 300 time periods, due to insufficient online learning, most of the beams assigned to

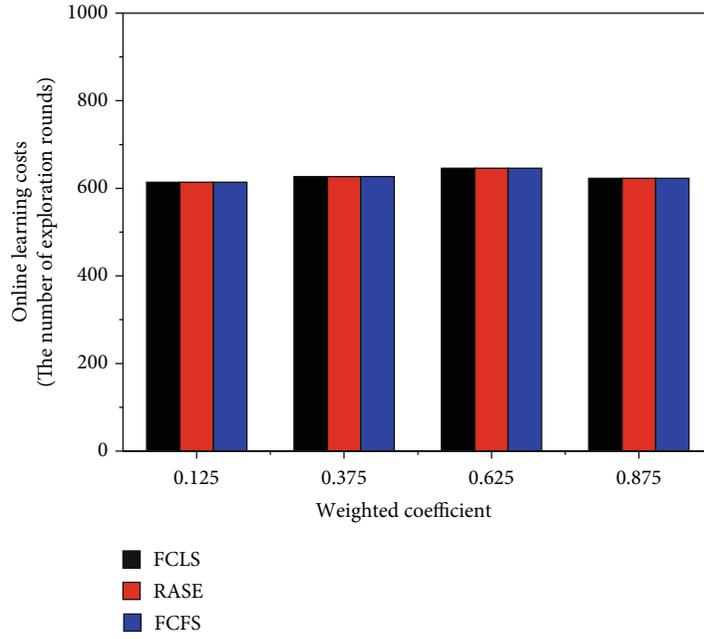


FIGURE 19: Impact of weight coefficient  $\zeta$  on online learning costs under the three vehicle scheduling strategies.

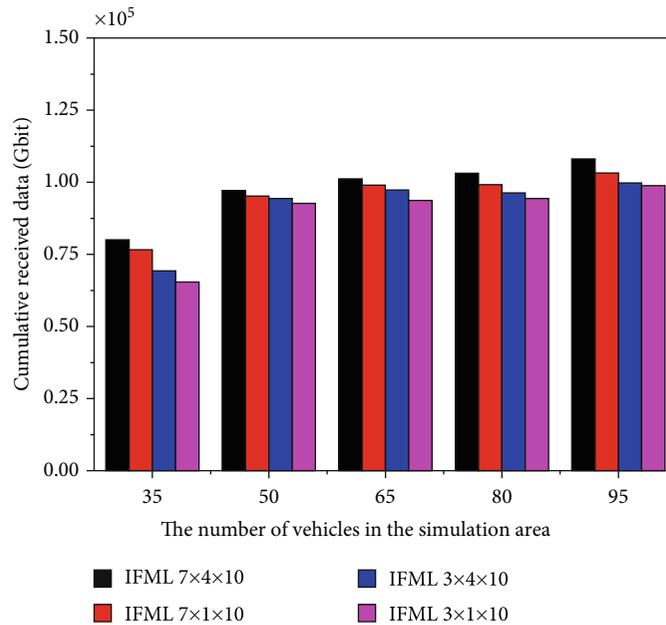


FIGURE 20: Impact of learning information space size on the IFML-I's cumulative received data.

the vehicles are selected at random, and thus, there is no significant difference between the two IFMLs' aggregate received data and the VFML's.

Based on the same simulation configuration as that in Figure 10, the number of exploration operations per time period required for the algorithms over the time horizon with 1000 time periods are shown in Figure 11. We can see from Figure 11 that the two IFMLs needs more time periods for beam performance exploration than the VFML. This indicates that the two IFMLs takes longer time to sufficient online

learning effect than the VFML. The main reason is that there is a difference of the learning information space resulting from the combination of beams and context subspaces between the two IFMLs and the VFML.

In the two IFMLs, the larger learning information space leads to higher aggregate received data but increases the online learning costs. Fortunately, after a certain amount of time periods, there is no difference in the online learning costs between the three schemes, where they still have a similar beam performance update process after each beam is used.

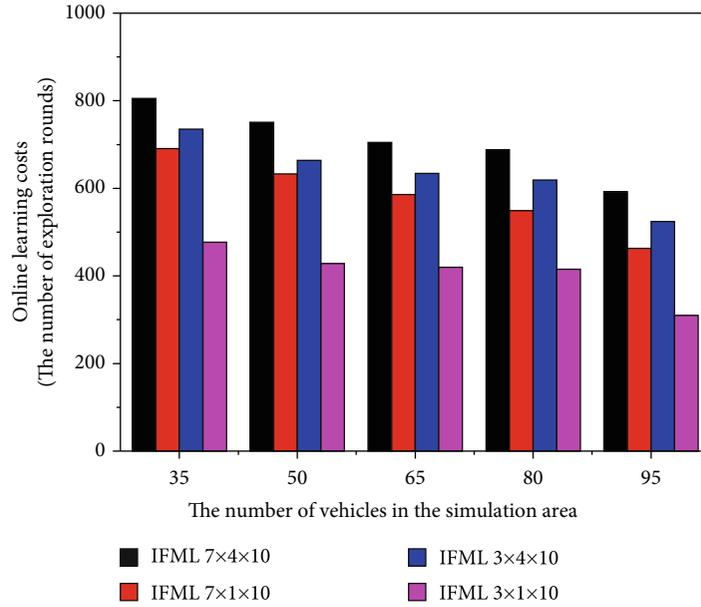


FIGURE 21: Impact of learning information space size on the IFML-I’s online learning costs.

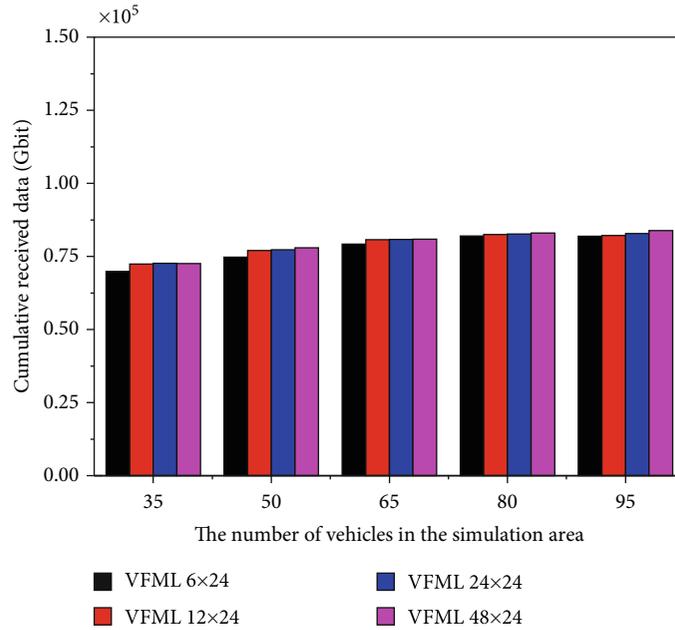


FIGURE 22: Impact of learning information space size on the VFML’s cumulative received data.

Based on the same simulation configuration as that in Figure 4, under the three vehicle scheduling strategies adopted by the IFML-I, Figure 12 shows that the impact of the number of vehicles in the system on the cumulative received data, while Figure 13 shows that the impact of the number of vehicles in the system on online learning costs.

As explained in Figure 4, the cumulative received data under all the vehicle scheduling strategies in Figure 12 varies with the number of vehicles in the simulation area. However, there is a difference in terms of cumulative received data between the three vehicle scheduling strategies. The main

for this difference is that the different strategies result in the different contact time, during which a vehicle stays in a beam coverage area. Among the three strategies, the FCLS leads to the longest average contact time, while the FCFS leads to the shortest one. So the cumulative data received under the FCLS is the largest, while that under the FCFS is the least.

The variation trend of the online learning costs with the number of vehicles in the simulation area under all the vehicle scheduling strategies in Figure 13 is similar to that in Figure 5. Therefore, the explanation in Figure 5 applies here as well. There is no difference in the online learning

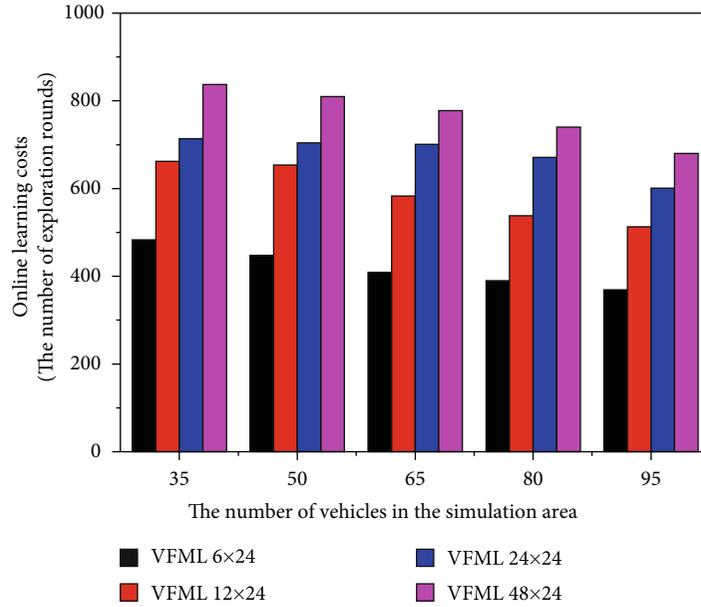


FIGURE 23: Impact of learning information space size on the VFML's online learning costs.

costs among the three strategies under the different number of vehicles. The main reason is that the different vehicle scheduling strategies do not affect the size of the learning information space.

Based on the same simulation configuration as that in Figure 6, under the three vehicle scheduling strategies adopted by the IFML-I, Figure 14 shows that the impact of the number of selected beams per period on cumulative received data, while Figure 15 shows that the impact of the number of selected beams per period on online learning costs.

The variation trend of cumulative received data with the number of selected concurrent beams under all the vehicle scheduling strategies in Figure 14 is similar to that in Figure 6. Therefore, the explanation in Figure 6 applies here as well. Also, in the case of different number of beams that can be used simultaneously, the main reason for the difference in cumulative received data under the three strategies is similar to that in Figure 12.

As explained in Figure 7, the online learning costs under all the vehicle scheduling strategies in Figure 15 vary with the number of beams that could be used simultaneously. However, there is no difference in the online learning costs under the three vehicle scheduling strategies. The main reason is that the different vehicle scheduling strategies do not affect the size of the information space for online learning.

Based on the same simulation configuration as that in Figure 8, under the three vehicle scheduling strategies adopted by the IFML-I, Figure 16 shows that the impact of thermal noise power density on cumulative received data, while Figure 17 shows that the impact of thermal noise power density on online learning costs.

The variation trend of cumulative received data with the thermal noise power density under all the vehicle scheduling strategies in Figure 16 is similar to that in Figure 8. Therefore, the explanation in Figure 8 applies here as well. Also, in the case of different thermal noise power density, the main reason

for the difference in cumulative received data under the three vehicle scheduling strategies is similar to that in Figure 12.

The variation trend of the online learning costs with the thermal noise power density under all the vehicle scheduling strategies in Figure 17 is similar to that in Figure 9. However, there is no difference in the online learning costs under the three vehicle scheduling strategies. The main reason is that the different vehicle scheduling strategies do not affect the size of the information space for online learning.

For the three vehicle scheduling strategies adopted by the IFML-I, we investigate the impact of the weight coefficient  $\zeta$  on the cumulative received data and the online learning costs, where the number of vehicles in the simulation area is set to 65 and  $\zeta$  varies from 0.125 to 0.875 with the step size of 0.125. Figure 18 shows that the impact of  $\zeta$  on cumulative received data, while Figure 19 shows that the impact of  $\zeta$  on online learning costs.

Combined with Figures 18 and 19, we can see that when the IFML-I adopts the weight coefficient value of 0.125, the cumulative received data is relatively large and the corresponding online learning costs are relatively small, where the ratio of performance to cost is relatively high.

For different communication environments, the weight coefficient value of achieving a relatively high ratio of performance to cost may be different. In the IFML-I, it is easy to adjust the weight coefficient value to adapt to different communication environments, while in the VFML, this flexibility is not available because of an incremental absolute average evaluation approach used in beam performance update.

Based on the same simulation configuration as that in Figure 4, we compare the impact of learning information space size on the performance metrics of the IFML-I, where "IFML  $7 \times 4 \times 10$ ", "IFML  $7 \times 1 \times 10$ ", "IFML  $3 \times 4 \times 10$ ", and "IFML  $3 \times 1 \times 10$ " represent the different learning information spaces used by the IFML-I, respectively. "IFML  $7 \times 4 \times 10$ " means that there are 7 types of beam widths (i.e.,

from  $30^\circ$  to  $90^\circ$  with the step size of  $10^\circ$ ) for each sector, the number of each type of beam widths is 4, and each sector is divided into 10 contextual subspaces. By analogy, the meanings of “IFML  $7 \times 1 \times 10$ ”, “IFML  $3 \times 4 \times 10$ ”, and “IFML  $3 \times 1 \times 10$ ” are also easy to understand, so the detailed description is omitted.

The cumulative received data of the IFML-I with different learning information space sizes under the different number of vehicles are shown in Figure 20, while the online learning costs required for the IFML-I with different learning information space sizes under the different number of vehicles are shown in Figure 21.

Combined with Figures 20 and 21, we can see that the IFML-I with a larger learning information space usually performs better than those with a smaller learning information space in terms of cumulative received data except for “IFML  $7 \times 1 \times 10$ ”. However, in terms of online learning costs, the IFML-I with a larger learning information space performs worse than those with a smaller learning information space.

Although “IFML  $7 \times 1 \times 10$ ” has less learning information space than “IFML  $3 \times 4 \times 10$ ”, it performs better than “IFML  $3 \times 4 \times 10$ ” in terms of both cumulative received data and online learning costs. This indicates that setting more types of beam widths is helpful to pick out more reasonable beam widths and thus improve performance. In other words, only increasing the number of beams with the same width has a limited improvement on performance. Also, we see that the cumulative received data of “IFML  $7 \times 1 \times 10$ ” is slightly worse than that of “IFML  $7 \times 4 \times 10$ ”, but the online learning costs of the former are much less than that of the latter, especially in the case of a large number of vehicles. This indicates that “IFML  $7 \times 1 \times 10$ ” has the relatively high ratio of performance to cost.

Based on the same simulation configuration as that in Figure 4, we compare the impact of learning information space size on the performance metrics of the VFML, where “VFML  $6 \times 24$ ”, “VFML  $12 \times 24$ ”, “VFML  $24 \times 24$ ”, and “VFML  $48 \times 24$ ” represent the different learning information spaces used by the VFML, respectively. “VFML  $6 \times 24$ ” means that there are 6 types of beam widths (i.e., from  $60^\circ$  to  $360^\circ$  with the step size of  $60^\circ$ ) and the entire base station coverage area is divided into 24 contextual subspaces. By analogy, the meanings of “VFML  $12 \times 24$ ”, “VFML  $24 \times 24$ ”, and “VFML  $48 \times 24$ ” are also easy to understand, so the detailed description is omitted.

The cumulative received data of the VFML with different learning information space sizes under the different number of vehicles are shown in Figure 22, while the online learning costs required for the VFML with different learning information space sizes under the different number of vehicles are shown in Figure 23.

Combined with Figures 22 and 23, we can see that the VFML with a smaller larger learning information space performs worse than those with a larger learning information space in terms of cumulative received data, while the opposite is true in terms of online learning costs.

Also, we see from Figures 22 and 23 that the cumulative received data gains from expanding the learning information space are extremely limited for the VFML. This indicates that

without changing the design structure or idea of the VFML algorithm, it is difficult to improve the cumulative received data by expanding the learning information space, which further illustrates the necessity of the work in this paper.

## 7. Conclusion

In this paper, we proposed the two improved fast machine learning algorithms (i.e., the IFML-I and the IFML-II) to improve network capacity in mmWave vehicular networks based on contextual multi-armed bandits. To cope with unknown performance values of mmWave beams assigned to the dynamically arriving vehicles, the two IFMLs periodically observe the amount of data received by the connected vehicles in each time period and update performance values of the mmWave beams under the context subspaces corresponding to these vehicles by adopting the proposed beam performance update approach based on adjustable weight coefficient. In this way, the two IFMLs learn context-specific mmWave beam performances online. Moreover, the two IFMLs exploit the richer contexts and more beam types so that the set of beams with the more reasonable beam width and beam direction is picked out. Also, the additional beam search overhead can be suppressed by the partitioned search method. Compared with the benchmark algorithm, the simulation results showed that the two IFMLs substantially improve the amount of data received by the vehicles in the system during each time period at a slightly increased online learning overhead. However, there is no difference in terms of beam performance update cost after a certain number of time periods between the two IFMLs and the benchmark algorithm.

## Data Availability

The simulation data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61873352, No. 61803387).

## References

- [1] T. S. Rappaport, R. Mayzus, H. Zhao et al., “Millimeter wave mobile communications for 5G cellular: it will work!,” *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [2] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: potentials and challenges,” *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [3] Y. H. Alkhateeb, M. S. Nam, C. Rahman, R. 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, 2017.

- [4] G. H. Sim, S. Klos, A. Asadi, A. Klein, and M. Hollick, "An online context-aware machine learning algorithm for 5G mmWave vehicular communications," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2487–2500, 2018.
- [5] M. Sepulcre and J. Gozalvez, "Context-aware heterogeneous V2X communications for connected vehicles," *Computer Networks*, vol. 136, pp. 13–21, 2018.
- [6] M. Muhammad and G. A. Safdar, "Survey on existing authentication issues for cellular-assisted V2X communication," *Vehicular Communications*, vol. 12, pp. 50–65, 2018.
- [7] M. Ghosal and M. Conti, "Security issues and challenges in V2X: a survey," *Computer Networks*, vol. 169, article 107093, 2020.
- [8] IEEE Computer Society LAN/MAN Standards Committee, "IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 2007, IEEE Std. 802.11-2012.
- [9] G. Karagiannis, O. Altintas, E. Ekici et al., "Vehicular networking: a survey and tutorial on requirements, architectures, challenges, standards and solutions," *IEEE communications surveys & tutorials*, vol. 13, no. 4, pp. 584–616, 2011.
- [10] Intelligent Transport Systems (ITS), "European Profile Standard for the Physical and Medium Access Control Layer of Intelligent Transport Systems Operating in the 5 GHz Frequency Band, ETSI ES 202 663 V1.1.0, 2010," 2017, [http://www.etsi.org/deliver/etsi\\_es/202600\\_202699/202663/01.01.00\\_60/es\\_202663v010100p.pdf](http://www.etsi.org/deliver/etsi_es/202600_202699/202663/01.01.00_60/es_202663v010100p.pdf).
- [11] R. Molina-Masegosa and J. Gozalvez, "LTE-V for sidelink 5G V2X vehicular communications: a new 5G technology for short-range vehicle-to-everything communications," *IEEE Vehicular Technology Magazine*, vol. 12, no. 4, pp. 30–39, 2017.
- [12] A. Bazzi, B. M. Masini, A. Zanella, and I. Thibault, "Beaconing from connected vehicles: IEEE 802.11p vs. LTE-V2V," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, Valencia, Spain, 2016.
- [13] D. Hoymann, M. Astely, G. Stattin et al., "LTE release 14 outlook," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 44–49, 2016.
- [14] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for vehicular networking: a survey," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 148–157, 2013.
- [15] S. Chen, J. Hu, Y. Shi et al., "Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 70–76, 2017.
- [16] "3GPP TR 38.913, v.14.0.0, Study on scenarios and requirements for next generation access technologies," 2016.
- [17] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, "Initial beamforming for mmWave communications," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, pp. 1926–1930, Pacific Grove, CA, USA, 2014.
- [18] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2670–2683, 2015.
- [19] C. N. Barati, S. A. Hosseini, S. Rangan et al., "Directional cell discovery in millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6664–6678, 2015.
- [20] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [21] M. Hashemi, A. Sabharwal, C. E. Koks, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pp. 2393–2401, Honolulu, HI, USA, 2018.
- [22] K. C. Ho and S. H. Tsai, "A novel multiuser beamforming system with reduced complexity and beam optimizations," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4544–4557, 2019.
- [23] W. B. Zhang, Y. M. Wei, S. C. Wu, W. X. Meng, and W. Xiang, "Joint beam and resource allocation in 5G mmWave small cell systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10272–10277, 2019.
- [24] D. Y. Zhang, A. Li, M. Shirvanimoghaddam, P. Cheng, Y. H. Li, and B. Vucetic, "Codebook-based training beam sequence design for millimeter-wave tracking systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5333–5349, 2019.
- [25] J. C. Fan, L. Y. Han, X. M. Luo, and J. Huang, "Delay analysis and optimization of beam scanning-based user discovery in millimeter wave systems," *IEEE Access*, vol. 8, pp. 25075–25083, 2020.
- [26] M. S. Sim, Y. G. Lim, S. H. Park, L. L. Dai, and C. B. Chae, "Deep learning-based mmWave beam selection for 5G NR/6G with sub-6 GHz channel information: algorithms and prototype validation," *IEEE Access*, vol. 8, pp. 51634–51646, 2020.
- [27] Y. N. R. Li, B. Gao, X. D. Zhang, and K. B. Huang, "Beam management in millimeter-wave communications for 5G and beyond," *IEEE Access*, vol. 8, pp. 13282–13293, 2020.
- [28] J. J. Zhang, Y. M. Huang, J. H. Wang, R. Schober, and L. X. Yang, "Power-efficient beam designs for millimeter wave communication systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1265–1279, 2020.
- [29] X. L. Liu, J. D. Yu, H. R. Qi et al., "Learning to predict the mobility of users in mobile mmWave networks," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 124–131, 2020.
- [30] J. Yang, S. Jin, C. K. Wen, X. Yang, and M. Matthaiou, "Fast beam training architecture for hybrid mmWave transceivers," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2700–2715, 2020.
- [31] J. Mazgula, J. Sapis, U. S. Hashmi, and H. Viswanathan, "Ultra reliable low latency communications in MmWave for factory floor automation," *Journal of the Indian Institute of Science*, vol. 100, no. 2, pp. 303–314, 2020.
- [32] W. Y. Ma, C. H. Qi, and G. Y. Li, "Machine learning for beam alignment in millimeter wave massive MIMO," *IEEE Wireless Communications Letters*, vol. 9, no. 6, pp. 875–878, 2020.
- [33] J. S. Gui and J. L. Liu, "An efficient radio access resource management scheme based on priority strategy in dense mmWave cellular networks," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8891660, 19 pages, 2020.
- [34] L. H. Shen and K. T. Feng, "Mobility-aware subband and beam resource allocation schemes for millimeter wave wireless

- networks,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11893–11908, 2020.
- [35] L. H. Shen, K. T. Feng, and L. Hanzo, “Coordinated multiple access point multiuser beamforming training protocol for millimeter wave WLANs,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13875–13889, 2020.
- [36] L. You, X. Chen, X. H. Song et al., “Network massive MIMO transmission over millimeter-wave and terahertz bands: mobility enhancement and blockage mitigation,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 12, pp. 2946–2960, 2020.
- [37] 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, 2015.
- [38] V. Va and R. W. Heath, “Basic relationship between channel coherence time and beamwidth in vehicular channels,” in *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, pp. 1–5, Boston, MA, USA, 2015.
- [39] J. Choi, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, R. W. Heath, and V. Va, “Millimeter-wave vehicular communication to support massive automotive sensing,” *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, 2016.
- [40] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, “Beam design for beam switching based millimeter wave vehicle-to-infrastructure communications,” in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, 2016.
- [41] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. Nix, “mmWave system for future ITS: a MAC-layer approach for V2X beam steering,” in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–6, Toronto, ON, Canada, 2017.
- [42] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. Nix, “Beam alignment for millimetre wave links with motion prediction of autonomous vehicles,” in *Antennas, Propagation & RF Technology for Transport and Autonomous Platforms 2017*, pp. 1–8, Birmingham, 2017.
- [43] L. Yan, H. C. Ding, L. Zhang et al., “Machine learning-based handovers for sub-6 GHz and mmWave integrated vehicular networks,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4873–4885, 2019.
- [44] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, “Deep learning aided fingerprint-based beam alignment for mmWave vehicular communication,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10858–10871, 2019.
- [45] A. S. Nasim, S. Ibrahim, and S. Kim, “Learning-based beamforming for multi-user vehicular communications: a combinatorial multi-armed bandit approach,” *IEEE Access*, vol. 8, pp. 219891–219902, 2020.
- [46] F. Rasheed and F. Hu, “Intelligent super-fast vehicle-to-everything 5G communications with predictive switching between mmWave and THz links,” *Vehicular Communications*, vol. 27, article 100303, 2021.
- [47] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, “Online learning for position-aided millimeter wave beam training,” *IEEE Access*, vol. 7, pp. 30507–30526, 2019.
- [48] S. Han, C. L. I, Z. Xu, and C. Rowell, “Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G,” *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186–194, 2015.
- [49] P. Auer, N. Cesa-Bianchi, and P. Fischer, “Finite-time analysis of the multi-armed bandit problem,” *Machine Learning*, vol. 47, no. 2/3, article 391590, pp. 235–256, 2002.
- [50] Technical Specification Group Radio Access Network, “Study on channel model for frequencies from 0.5 to 100 GHz,” 2017, document 138 901 V14.0.0, 3GPP.
- [51] Q. Xue, X. Fang, and C. X. Wang, “Beamspace SU-MIMO for future millimeter wave wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1564–1575, 2017.
- [52] P. Liu, J. Blumenstein, N. S. Perovic, M. di Renzo, and A. Springer, “Performance of generalized spatial modulation MIMO over measured 60GHz indoor channels,” *IEEE Transactions on Communications*, vol. 66, no. 1, pp. 133–148, 2018.
- [53] M. Xiao, Y. Huang, L. Dai et al., “Millimeter wave communications for future mobile networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.