

Enhancing Energy Efficiency for Cellular-Assisted Vehicular Networks by Online Learning-Based mmWave Beam Selection

Jinsong Gui (✉ jsgui06@163.com)

Central South University <https://orcid.org/0000-0001-9399-9092>

Yao Liu

Central South University

Research

Keywords: energy efficiency, vehicular networks, fast machine learning, millimeter wave, beam selection

Posted Date: August 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-774513/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at EURASIP Journal on Wireless Communications and Networking on January 3rd, 2022. See the published version at <https://doi.org/10.1186/s13638-021-02080-5>.

Enhancing Energy Efficiency for Cellular-assisted Vehicular Networks by Online Learning-based mmWave Beam Selection

Jinsong Gui, Yao Liu

School of Computer Science and Engineering, Central South University, Changsha, 410083, China

Correspondence should be addressed to J.S. Gui, jsgui06@163.com

The second author: liuyao816@csu.edu.cn

Abstract—Millimeter-Wave (mmWave) technology is deemed as a feasible approach for future vehicular communications. However, mmWave signals are characterized by high path loss and penetration loss, which can be alleviated by directional communication. Directional transmission performance depends on beam alignment between transmitter and receiver, which is not easy to achieve in highly dynamic vehicular communications. The existing works addressed beam alignment problem by designing online learning-based mmWave beam selection schemes, which can be well adapted to high dynamic vehicular scenarios. However, this type of works does not take energy efficiency into account. Therefore, we propose an Energy efficiency-based FML (EFML) scheme to compensate for this shortfall, where the power consumption can be reduced as far as possible under the premise of meeting the basic data rate requirements of vehicle users and the users requesting the same content in close proximity can be organized into the same receiving group to share the same mmWave beam. The simulation results show that the EFML scheme improves both the network energy efficiency and the amount data of cellular-assisted vehicular networks at the cost of more beam performance update overhead. However, there is no difference in the cost of updating beam performance after adequate online learning.

Keywords—energy efficiency, vehicular networks, fast machine learning, millimeter wave, beam selection

I. INTRODUCTION

Connected vehicles and Cooperative Intelligent Transport System (C-ITS) systems will depend on Vehicle to Everything (V2X) communications to improve traffic safety, driving efficiency, and infotainment experience. The Long Term Evolution (LTE) network and the Fifth-Generation (5G) cellular network have been widely recognized as the main infrastructure to adapt to the characteristics of connected vehicles, such as high mobility, high time sensitivity, and high transmission reliability. With the rapid growth of connected vehicles and vehicle users' continuous pursuit of experience

quality, higher requirements are put forward for network capacity. Although the sub-6 GHz frequency band adopted in the LTE network cannot meet the increasing capacity demand, it has the longer communication distance. On the contrary, the millimeter Wave (mmWave) frequency band adopted in the 5G system can support the higher transmission capacity, but it has the shorter communication distance. Therefore, the 5G system and the LTE network are widely recognized as the main driving forces for fully supporting all the V2X applications.

However, the high mobility of connected vehicles will lead to the complexity of directional mmWave communications. The quality of directional mmWave communications depends on whether and to what extent the beams at both ends of the communication are aligned [1]. The high mobility of connected vehicles will force alignment operations to be performed frequently, resulting in high overhead. For the usually used beam alignment mechanisms (e.g., beamforming training, beam tracking, and beam selection), it will be a promising option to always select a reasonable beam pair for a mmWave link in time. Due to the many communication parameters involved in beam selection, pure mathematical modeling is very complex. So some researchers are trying to use machine learning to figure out how to solve this problem, where beam selection methods based on online learning [2,3] are used without prior training and very suitable for dynamic vehicular networks. In the existing online learning-based mmWave beam selection methods, based on the assumption that the performance of a specific beam is similar under similar contexts, by self-exploration, learning and adapting to the dynamic communication environment, each mmWave Small Base Station (mmSBS) can select right beams, where it independently learns from its previous decisions and the relationship to available vehicular contexts, and becomes conscious of the performance of each mmWave beam.

In the Fast Machine Learning (FML) scheme in [2], the assumption that there are no overlapping mmWave beams limits the optional beam width and azimuth range. In the Improved Fast Machine Learning (IFML) scheme in [3], the beams are allowed to overlap and the virtual beam concept is

proposed. Unlike the number of Radio Frequency (RF) chains, which is limited by form-factor and manufacturing cost [4], a virtual beam is only a region that limits energy propagation and its shape is usually a cone (or sector), where the number of virtual beams can be infinite at each mmSBS. When an RF chain is assigned to a virtual beam and also a specific channel is assigned to it, this virtual beam becomes the actual available physical beam and its use is not limited by beam overlap. When the overlapping beams share the same channel, the corresponding transmission powers should be adjusted to control the mutual interference. Not all the cases of mutual interference can be well controlled, so the specific solution for power adjustment is not covered in [3]. In addition, in the existing related works, the amount of data received by vehicles is considered as the performance measure of the beam, which is not suitable for the development demand of green communication due to the lack of power consumption concerns. Also, the existing related works do not consider aggregating users in close proximity who request the same data content (e.g., the latest traffic congestion information, real-time high-definition electronic maps, current events and news) to serve as a multicast group, resulting in the possibility of consuming resources repeatedly to send the same content.

Therefore, to address the above problems, we propose an Energy efficiency-based FML (EFML) scheme and list the major contributions as follows.

1) Different from the existing online learning-based mmWave beam selection schemes aiming to maximize the overall aggregated received data, our scheme aims at enhancing energy efficiency of cellular-assisted vehicular networks.

2) In our scheme, the transmission power of each mmSBS is allowed to be adjusted as long as the energy efficiency can be improved. Therefore, the power consumption can be reduced as far as possible under the premise of meeting the basic data rate requirements of vehicle users.

3) In order to further reduce energy consumption and save communication resources, the users requesting the same data content in close proximity are organized into the same receiving group to share the same mmWave beam and reduce the occupation of RF chains.

4) The simulation results show that, compared with the existing online learning-based mmWave beam selection schemes, the EFML scheme substantially improves the energy efficiency and the amount data of cellular-assisted vehicular networks at the cost of more system overhead. However, after a certain time, there is no difference in the cost of updating the beam performance of the system.

In the rest of the paper, the related works are presented in Section II. The system model and the detail of EFML algorithm are addressed in Sections III and IV, respectively. Simulations results are discussed in Section V. Finally, we conclude this paper in Section VI.

II. RELATED WORK

There have been many solutions to the problem of beam selection in traditional networks (e.g., the works in [5,6]). However, they need complex transceiver links and accurate location information and thus undoubtedly cause high overhead and delay. Unlike the above works in the sub-6 GHz bands, there are many works on beam selection in mmWave networks. The authors in [7] proposed a mmWave beam selection method based on deep learning which utilizes the channel characteristics of the Sub-6 GHz band to solve the mmWave beam selection problem, while the authors in [8] presented a beam alignment algorithm based on machine learning for the beam management problems in the mmWave massive Multi-input Multi-output (MIMO) networks. The author in [9] proposed an iterative order minimum optimization training scheme based on the simulated beam selection of machine learning. The above schemes require a large number of prior data samples and beams training processes.

In [10], the authors studied the problem of multiple RF chains of mmWave transceivers in mobile mmWave communication systems and developed a codebook-based beam tracking strategy, which shows that the performance of the beam tracking strategy can be improved by optimizing the transmit power of the training beam. The authors in [11] gave an overview of current beam management approaches based on 5G standardization, including some of the major challenges and future trends for mmWave communications in current 5G New Radio (NR) standards. By analyzing the average search delay of two different mmWave network models, the authors in [12] found that the average number of searches is related to the number of search sectors. The authors in [13] designed a low-cost beamforming module-assisted hybrid architecture and proposed a fast beam training method. The authors in [14] studied the sensitivity of the beam stability selected by the base station. By observing different operating frequencies, dynamic channel characteristics, and different user mobility, they found that the perceived time-of-stay of the beam will be affected by beam management parameterization.

The authors in [15] leveraged the advantage of the mmWave characteristics in ultra-dense networks and proposed a method for joint optimization and resource allocation between base stations and users. Specifically, they aimed at maximizing user throughput in the system while also considering fairness. To reduce the overhead and complexity of the wireless backhaul and access process, the authors in [16] proposed a hybrid beamforming multi-stage design scheme based on channel feedback. To improve the efficiency of user-provided networks through resource allocation of links, the authors in [17] proposed a joint incentive and resource allocation algorithm, which considered the restriction of network resources, incentive system and user fairness.

To reduce the overhead of establishing a mmWave link in vehicular-to-everything networks, the authors in [18] proposed a beam training method based on the assistance of out-of-band information. The authors in [19] proposed a beamforming scheme based on deep learning for high mobility mmWave systems. To further reduce the beamforming overhead of the mmWave system, the authors in [20] proposed an intelligent prediction beam alignment algorithm from the Multiple Access Control (MAC) layer of the mmWave vehicle system. The authors in [21] proposed a machine learning approach based on situational awareness to predict mmWave beams. Specifically, this approach learns beam information from some past observations including the position of the vehicles and the optimal beam. The authors in [22] proposed a neural network-based algorithm for beam alignment in vehicular networks. However, this scheme needs to learn more information about the channel state and can only select the best beam direction for a single user. Considering the propagation characteristics of mmWave in 5G vehicular networks, the authors in [23] proposed a simulated anneal-based beam management model to improve the effective communication of the system. The challenges of mmWave communication for the vehicular networks are also investigated in [24,25].

Moreover, Multi-Armed Bandit (MAB) is a classic and general online learning method and has been used to solve various problems in wireless communication networks [26]. The author in [27] developed an equivalent structured MAB model to solve the beam alignment problem in the mmWave system. However, this method requires an exhaustive search for beam alignment between transceivers, which will cause great system overhead due to the large search space. The authors in [2] proposed FML to address the context-awareness beam selection issue. Specifically, they model the problem of beam selection as a contextual MAB problem and proved the convergence of FML. However, they only consider one-dimensional contextual information and only one vehicle can be served within the beam range.

The authors in [28] model the problem of beam selection as a contextual combinatorial MAB problem with delayed feedback and Quality of Service (QoS) constraints and proposed an online learning algorithm that achieves a good balance between satisfying the performance guarantee of the system and maximizing the network capacity. However, since this prediction mechanism requires the view information of the source mmWave base station, it will cause greater system overhead. In addition, the fast mobility of the vehicle scenarios is also a big challenge to this prediction method. The authors in [29] developed an online learning algorithm for beam selection by using the MAB framework that requires learning rough beam orientation in the pre-defined codebook.

The authors in [3] proposed an online learning-based algorithm for mmWave beam selection to improve the network capacity of the vehicular communication systems. In addition,

the algorithm selects a more appropriate beam direction and beam width for the mmWave base station by setting and learning more dimensional context information. However, the above researches are all only considered maximizing the system throughput or achieving the maximum network rate, but they did not consider the power adjustment to reduce the energy consumption. On the whole, the work in [3] is the most relevant to our work. However, due to the reasons mentioned above, there is still room for improvement in IFML, and the power allocation problem is not described in detail in the text. In addition, consideration of user multicast groups and power requirements for green and energy-saving communication is also the motivation for this article.

III. SYSTEM MODEL

A. Network Architecture

An integrated mmWave/sub-6 GHz cellular network is considered in this paper, in which some mmSBSs are overlapped in the coverage area of an LTE eNB. As shown in Fig. 1, by a wired or wireless backhaul link, a mmSBS can communicate with its associated LTE eNB. Each vehicle is equipped with the two kinds of radio interfaces, where an LTE interface is used to keep a connection to the LTE eNB, and a mmWave interface is adopted for high-speed data transmission. From a theoretical point of view, an infinite number of virtual beams can be programmed per mmSBS, and the beam width of each beam can be set between 0 and 360°, where beams are allowed to overlap and the number of RF chains is much less than the number of virtual beams due to the manufacturing cost and the limitation of form-factor.

All the vehicles will be grouped according to how close they are to each other and whether the same data content is requested. In this paper, each mmSBS only provides service for each vehicle group. Even if there is only one vehicle in a vehicular network, it must form a vehicle group. Each vehicle group has a unique identifier and the other parameters associated with this group include the number of vehicles, the identifiers of vehicles, the requested data content, the central coordinates of the distribution of vehicles within the group, and the identifier of the vehicle farthest from the target mmSBS within the group.

The maximum number of RF chains at a mmSBS determines the maximum number of vehicle groups that this mmSBS can simultaneously serve. If the number of vehicle groups in the coverage of a mmSBS and the number of the virtual beams of it exceed the number of RF chains of this mmSBS respectively, the mmSBS should select the best subset of beam-power pairs in order to provide the best system performance. To reach that target, we formulate each mmSBS's beam selection as a MAB problem, where each mmSBS can identify the subset of best beams with the matching transmission power values over time. According to the

description in [30], a decision maker of a MAB problem has to choose a subset of actions of unknown expected rewards to maximize the reward over time, but those which have already generated high rewards should also be exploited, where how to deal with the exploration vs. exploitation dilemma is a challenging problem.

B. Problem Statement

Like the work in [3], the number of virtual beams at a mmSBS is not limited in this paper. Also, unlike the work in [3], the transmission power of each mmSBS is allowed to be adjusted in this paper. Therefore, besides preserving the coverage division in [3], we must also focus on how to find an appropriate transmission power for each mmSBS from the set of available transmission powers.

Firstly, for the purpose of reducing the search time of the online learning process, each mmSBS's coverage area is divided into L non-overlapping sectors (e.g., $L=4$ in Fig. 1), where there are no more than M^l virtual beams for the l -th sector ($l \in \{1, \dots, L\}$) and these virtual beams are allowed to overlap.

In the l -th sector, each mmSBS uses a set \mathcal{M}^l and a set \mathfrak{N}^l , which includes $M^l = |\mathcal{M}^l|$ virtual beams and $N^l = |\mathfrak{N}^l|$ transmission power levels. Therefore, there are $M^l \times N^l$ beam-power pairs for the l -th sector.

For each sector, the mmSBS can choose a subset of no more than n beam-power pairs to serve no more than n ($n < M^l \times N^l, \forall l \in \{1, \dots, L\}$) vehicle groups simultaneously, in which the number of the served vehicle groups is limited by the maximum number of RF chains at the mmSBS and thus the maximum number of them is limited to n .

From the perspective of all the sectors, the n beam-power pairs in the same sector may not be the best n beam-power pairs. Thus, first of all, the mmSBS should choose no more than n best beam-power pairs from every sector separately to provide them to no more than n vehicle groups in every sector respectively. Then, it chooses no more than n best beam-power pairs from all the chosen beam-power pairs to provide them to no more than n vehicle groups in the whole coverage area of the mmSBS.

The LTE eNB is capable of providing the vehicle group context information to the mmSBS. With the help of the LTE eNB, a vehicle group will know the location of the mmSBS and the chosen beam-power pair for it. Fig. 2 shows the whole process of information interaction between vehicles, the LTE eNB, and mmSBSs. When a vehicle wants to communicate with a mmSBS via a mmWave link, it firstly sends a registration request message (refer to "1: registration request" in Fig. 2) to the LTE eNB with which it keeps a continuous connectivity via its LTE interface. This registration request message contains the description of the vehicle's velocity, location and request data content.

The LTE eNB may receive a large number of registration requests from the vehicles in the service areas it covers. So it will periodically analyze and handle the received registration requests based on certain policies, where the interval between successive processing operations can be adjusted according to the delay requirements of registration requests. If the interval is longer, the number of registration requests processed at a single time period may be larger, but the response per vehicle will be slower, and vice versa.

Once the time to process the registration requests arrives, the LTE eNB firstly analyzes the received registration requests to build the vehicle groups one by one according to the vehicles' velocity, location and request data content, and then sends a potential mmSBS a mmWave service request message (refer to "2: service request" in Fig. 2). This message contains each vehicle group's identifier, each vehicle's cellular system identifier in a vehicle group, the identifier of the vehicle farthest from the target mmSBS within a vehicle group, and the expected direction of arrival at the mmSBS.

By using the EFML scheme, the mmSBS will respond to the LTE eNB's mmWave service request (refer to "3: service response" in Fig. 2) with the chosen beam-power pairs. Upon receipt of service response from the mmSBS, the LTE eNB will send each vehicle in each vehicle group a registration response message about the mmSBS (refer to "4: registration response" in Fig. 2). This message contains the mmSBS's location and the chosen beam-power pairs.

Once each vehicle in the vehicle group reaches the covered area, it sends the mmSBS an associating request to start a mmSBS associating process, and then it receives an associating response from the mmSBS (refer to "5: association" in Fig. 2). Then, each vehicle obtains the Channel State Information (CSI) by analyzing the associating response message from mmSBS and feedbacks the CSI to the mmSBS. Based on the CSI feedback from all the vehicles in a group, the mmSBS can know whether the beam-power pair it chooses can meet the data rate requirement of the vehicle with the worst channel quality in the group.

After the associating operations, the mmSBS starts the data transfer process (refer to "6: communication" in Fig. 2), and then it will get acknowledgments of the transferred data frames if the data transfer process is successful, where any other feedback is not required. If any vehicle in the vehicle group cannot detect the mmSBS within the chosen beam-power pair, it will send the feedback to the LTE eNB (refer to "7: service feedback" in Fig. 2). Finally, in order to help mmSBS make better decisions in the future, the LTE eNB will send the feedback to the mmSBS (refer to "8: service feedback" in Fig.2).

The selection results of beam-power pairs should be adjusted in time to serve the most suitable set of vehicle groups, so each mmSBS uses a discrete time setting, where system time

is divided into time slices with equal length and denoted as t ($t \in \{1, \dots, T\}$).

When each time slice t passes, all the selection results of beam-power pairs will be updated. If each time slice is relatively shorter, the selection results of beam-power pairs are updated more timely, but it generates the higher system overhead. Thus, how to get a reasonable tradeoff will be very critical, and one option is to determine the specific value through experience. The detailed process of selection and update for beam-power pairs is described below.

1) At the first time slot of each time slice t , a set $\mathcal{G}_t^l = \{g_{t,i}^l | i \in \{1, \dots, G_t^l\}\}$ of vehicle groups will be registered in the l -th ($l \in \{1, \dots, L\}$) sector of the mmSBS via the LTE eNB, in which G_t^l is the number of vehicle groups and meets $G_t^l = |\mathcal{G}_t^l| \geq n$. The parameter n is determined by the maximum number of RF chains that mmSBS can support, so it also represents the maximum number of vehicle groups that can simultaneously obtain downlink transmission services in the entire coverage of mmSBS.

As mentioned above, the mmSBS obtains the information about the group context $o_{t,i}^l$ of each incoming vehicle group $g_{t,i}^l$. The group context $o_{t,i}^l$ may be described by D context dimensions, which is regarded as an D -dimensional vector $O_{G_t^l} = \{o_{t,i}^l = \langle o_{t,i}^{l,1}, \dots, o_{t,i}^{l,D} \rangle | i \in \{1, \dots, G_t^l\}\}$ and acquired by the mmSBS after the first time slot of each time slice t . In this paper, we only consider vehicle group distance (which is determined by the vehicle farthest from the target mmSBS within the group) and direction of arrival as the context for a vehicle group, so the context vector is two-dimensional (i.e., $D=2$).

2) The mmSBS chooses a subset of no more than n best beam-power pairs from the l -th sector, in which the set of chosen beam-power pairs in each time slice t is indicated as $\mathcal{BP}_t^l = \{\langle b_{t,j}^l, p_{t,j}^l \rangle | j \in \{1, \dots, n\}, n < M^l \cdot N^l\}$. After this, it reselects no more than n beam-power pairs from $\bigcup_{i=1}^{G_t^l} \mathcal{BP}_t^l$ to serve no more than n vehicle groups within the mmSBS's coverage area. Finally, no more than n vehicle groups in $\bigcup_{i=1}^{G_t^l} \mathcal{G}_t^l$ are selected to accept service, and each vehicle of each selected vehicle group is informed about the chosen beam-power pair by the associated LTE eNB by adopting their LTE interfaces.

3) When any vehicle of each chosen vehicle group (e.g., $g_{t,i}^l$) reaches its expecting coverage of mmSBS, it receives communication data from this mmSBS and feeds this situation back to it. The mmSBS only observes the amount of data successfully received by the vehicle with the worst channel quality within each group, and then regarded it as the amount of data $r_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)$ that the vehicle group $g_{t,i}^l$ successfully receives via the chosen beam-power pair $\langle b_{t,j}^l, p_{t,j}^l \rangle$, until the time slice t is over or the vehicle with the worst channel quality in the group is not covered by its beam.

The amount of data $r_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)$ is usually limited to r_{max} , in which r_{max} is the maximum amount of data that can be received by the vehicle with the worst channel quality in the group. The contact time and the Shannon theorem can be employed to estimate r_{max} . The contact time is considered to be the time during that mmSBS can send data to the vehicle with the worst channel quality in the group, which is bounded by the coverage area of the chosen beam-power pair and relies on vehicle speed, beam direction, beam width and transmission power size.

In this paper, the performance of the chosen beam-power pair $\langle b_{t,j}^l, p_{t,j}^l \rangle \in \mathcal{BP}_t^l$ for the vehicle group with the context $o_{t,i}^l \in O_{G_t^l}$ during the time slice t is estimated by

$$e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l) = \frac{r_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)}{p_{t,j}^l \cdot t} \quad (1)$$

This performance measure in (1) is also approximated as the energy efficiency of the vehicle group with the context $o_{t,i}^l \in O_{G_t^l}$ when this vehicle group gets the beam-power pair $\langle b_{t,j}^l, p_{t,j}^l \rangle$ during the time slice t .

We consider $e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)$ as a random variable, and denote its expected value as $\tilde{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)$, which is also seen as the expected performance of the beam-power pair $\langle b_{t,j}^l, p_{t,j}^l \rangle$ under the group context $o_{t,i}^l$.

The goal of the mmSBS's choosing a subset of the beam-power pairs is to maximize the expected energy efficiency at a subset of vehicle groups. In other words, its goal is to maximizing the average expected beam-power pair performance. We denote the optimal subset of beam-power pairs in the time slice t of the l -th sector of the mmSBS as $\mathcal{BP}_t^{l^*}(O_{G_t^l}) = \{\langle b_{t,j}^{l^*}, p_{t,j}^{l^*} \rangle \in (O_{G_t^l}) | j \in \{1, \dots, n\}, n < M^l \cdot N^l\}$, which depends on $O_{G_t^l} = \{o_{t,i}^l = \langle o_{t,i}^{l,1}, \dots, o_{t,i}^{l,D} \rangle | i \in \{1, \dots, G_t^l\}\}$ and its n beam-power pairs formally satisfy the following formula.

$$\begin{aligned} & \langle b_{t,j}^{l^*}, p_{t,j}^{l^*} \rangle \\ \in & \underset{b_{t,j}^l \in \mathcal{M}^l, p_{t,j}^l \in \mathcal{N}^l \setminus (\bigcup_{k=1}^{j-1} \{\langle b_{t,k}^{l^*}, p_{t,k}^{l^*} \rangle\})}{\arg \max} \tilde{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l) \quad (2) \\ & o_{t,i}^l \in O_{G_t^l} \setminus (\bigcup_{k=1}^{j-1} \{o_{t,k}^{l^*}\}) \end{aligned}$$

At the beginning of system initialization, if the mmSBS already knows the expected beam-power pair performance $\tilde{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)$ for each vehicle group context $o_{t,i}^l \in O_{G_t^l}$ and each beam-power pair $\langle b_{t,j}^l, p_{t,j}^l \rangle \in \mathcal{M}^l \times \mathcal{N}^l$, it could be easy to choose the optimal subset of beam-power pairs for each set of the reaching vehicle groups in the l -th sector by (2). The average energy efficiency can be obtained through the total amount of data expected to be received for all time slices.

$$\begin{aligned} & \sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n E \left[\frac{e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)}{T \cdot n} \right] = \\ & \sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n \frac{\tilde{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^{l^*}(o_{t,i}^l)}{T \cdot n} \quad (3) \end{aligned}$$

Usually, the mmSBS has no information about the communication environment, so it must learn the expected beam-power pair performance $\bar{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l)$ over time for each vehicle group $g_{t,i}^l$ with the group context $o_{t,i}^l$.

That is, in order to learn these performances, the mmSBS must explore different beam-power pairs for different group contexts over time.

Also, it should exploit the beam-power pairs proved to have good performance. Thus, the mmSBS must make a trade-off between exploring the beam-power pairs with the unknown performance and exploiting those with the known high performance.

Next, we will elaborate the EFML scheme, in which the best n beam-power pairs are selected from $\cup_{l=1}^L \mathcal{M}^l \times \mathcal{N}^l$ based on the incoming vehicle groups with the contexts $\cup_{l=1}^L \mathcal{O}_{G_t^l}$ under each time slice. The EFML's choice depends on the historical record of the chosen beam-power pairs in previous time slices and the corresponding observed beam-power performance values. For any set of vehicle groups in any group context in l -th sector, the expected average energy efficiency of amount of received data of this sector is estimated by

$$\sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n E \left[\frac{e_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} \right] = \sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n E \left[\frac{\bar{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} \right] \quad (4)$$

We regard the regret of learning as the expected difference in the average energy efficiency achieved by a vehicle group and by the learning algorithm. According to (3) and (4), it can be estimated as follows

$$R(T) = E \left[\sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n \left(\frac{e_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} - \frac{\bar{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} \right) \right] \\ = \sum_{t=1}^T \sum_{i=1}^{G_t^l} \sum_{j=1}^n \left(\frac{\bar{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} - E \left[\frac{e_{\langle b^l, p^l \rangle}(o_{t,i}^l)}{T \cdot n} \right] \right) \quad (5)$$

IV. THE ENERGY EFFICIENCY-BASED FML

Each mmSBS executes the EFML scheme independently. To begin with, the context space of the vehicle group in each sector of the mmSBS will be evenly divided into context subspaces of the same size. Then, the EFML learns the performance in each subspace of different beam-power pairs separately. Furthermore, the EFML executes either an exploration action or an exploitation action in each time slice, and it depends on the control function of the system and the contexts of reaching vehicle groups. If the EFML executes an exploration action, the scheme randomly chooses a subset of beam-power pairs. Furthermore, in any exploitation process, it will choose the beam-power pairs that performed best in the previous time slices. Finally, by observing the average energy

efficiency achieved by the vehicle groups in its coverage area, the EFML scheme obtains performance estimating values of the chosen beam-power pairs. Thus, the algorithm learns the performance of each beam-power pair under each vehicle group context over time.

The pseudo-code description of the EFML scheme is listed in Algorithms 1~3. In the lines 1~5 of Algorithm 1, the EFML evenly divides the vehicle group context space $\cup_{l=1}^L \mathcal{O}_{G_t^l}$ into $L \cdot O_T$ D -dimensional subspaces with the equal size, in which L and O_T are the input constants. Then, the EFML scheme initializes each counter $N_{\langle b, p \rangle, s}^l$ for each beam-power pair $\langle b, p \rangle \in \cup_{l=1}^L \mathcal{M}^l \times \mathcal{N}^l$ and each subspace $s \in \cup_{l=1}^L \mathcal{O}_{G_t^l}$. The function of these counters is to record how many vehicle groups of a certain group context have entered the coverage area of the mmSBS in previous time slices, where the mmSBS had chosen a certain beam-power pair.

The counter $N_{\langle b, p \rangle, s}^l(t)$ represents the total number of vehicle groups with the group context in subspace s that entered the coverage area of the mmSBS whenever the beam-power pair $\langle b, p \rangle$ had been chosen in any of the time slices 1, ..., $t-1$ of any of the sectors 1, ..., L . Moreover, Algorithm 1 also initializes each estimator $\hat{e}_{\langle b, p \rangle, s}^l$ for each beam-power pair $\langle b, p \rangle \in \cup_{l=1}^L \mathcal{M}^l \times \mathcal{N}^l$ and each subspace $s \in \cup_{l=1}^L \mathcal{O}_{G_t^l}$, which represents the estimated performance of beam-power pair $\langle b, p \rangle$ for the vehicle groups with the group contexts in subspace s .

In each time slice t , the EFML scheme observes the vehicle group contexts $\cup_{l=1}^L \mathcal{O}_{G_t^l}$ of the $\sum_{l=1}^L G_t^l$ incoming vehicle groups. In the lines 3-4 in Algorithm 2, for each vehicle group context $o_{t,i}^l$, the EFML scheme decides to which subspace the vehicle group context belongs, where it finds out $s_{t,i}^l \in \cup_{l=1}^L \mathcal{O}_{G_t^l}$ with $o_{t,i}^l \in s_{t,i}^l$. Based on the set $\cup_{l=1}^L \mathcal{GH}_t^l := \cup_{l=1}^L \{s_{t,i}^l | i \in \{1, \dots, G_t^l\}\}$ of subspaces, the EFML scheme solves the set $\mathcal{BP}_{\mathcal{GH}_t^l}^{ue}(t)$ of under-explored beam-power pairs (see line 5 in Algorithm 2) by the following formula.

$$\mathcal{BP}_{\mathcal{GH}_t^l}^{ue}(t) := \cup_{l=1}^L \{ \langle b, p \rangle \in \mathcal{M}^l \times \mathcal{N}^l : N_{\langle b, p \rangle, s_{t,i}^l}^l(t) \leq K(t) \} \quad (6)$$

In (6), $K: \{1, \dots, T\} \mapsto \mathbb{R}$ is a deterministic monotone increasing control function, which is adopted to decide whether to execute an exploration process or run an exploitation process. The control function $K(t)$ should be adequately chosen to ensure that the EFML scheme obtains an expected good performance with respect to its regret. The Theorem 1 in [2] provides an appropriate choice for the control function and it is also described in [3]. For the convenience of readers, we repeat it as follows.

Theorem 1 (Bound for $R(T)$): For the l -th sector, let $K(t) = t^{\frac{2\alpha}{3\alpha+D}} \log(t)$ and $O_T = \lceil T^{\frac{1}{3\alpha+D}} \rceil$. If the EFML scheme is executed by adopting these parameters and if Assumption 1 is

true, the leading order of the regret $R(T)$ is $O\left(nG_t^l \tau_{\max} M^l N^l T^{\frac{2\alpha+D}{3\alpha+D}} \log(T)\right)$.

There is a detailed proof of Theorem 1 in [2], and the parameters are slightly different but basically the same. For the reader's convenience, we also repeat the description of *Assumption 1* here.

Assumption 1: there is $\alpha > 0$ and $\beta > 0$ so that for all $\langle b, p \rangle \in \mathcal{M}^l \times \mathcal{N}^l$ and for all $x, y \in O_{G_t^l}$ in the l -th sector, it is true that $|\tilde{e}_{\langle b, p \rangle}(x) - \tilde{e}_{\langle b, p \rangle}(y)| \leq \beta \|x - y\|^\alpha$, where $\|\cdot\|$ represents the Euclidean norm in \mathbb{R}^D .

For *Assumption 1*, although the parameters are somewhat different from those described in [2,3], they are essentially the same.

The lines 6-14 in Algorithm 2 show that the EFML scheme will execute the exploration process if there are under-explored beam-power pairs. If the number $u^l(t) := \mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ of under-explored beam-power pairs is at least n , the EFML scheme randomly chooses n beam-power pairs from them. If the number $u^l(t)$ of under-explored beam-power pairs is less than n , the EFML scheme chooses all the $u^l(t)$ under-explored beam-power pairs. Furthermore, it chooses the $(n - u^l(t))$ beam-power pairs $\langle \hat{b}_{1, G\mathcal{H}_t^l}^l(t), \hat{p}_{1, G\mathcal{H}_t^l}^l(t) \rangle, \dots, \langle \hat{b}_{n-u^l, G\mathcal{H}_t^l}^l(t), \hat{p}_{n-u^l, G\mathcal{H}_t^l}^l(t) \rangle$ from $\mathcal{M}^l \times \mathcal{N}^l \setminus \mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$, which satisfy the following formula.

$$\begin{aligned} & \langle \hat{b}_{j, G\mathcal{H}_t^l}^l(t), \hat{p}_{j, G\mathcal{H}_t^l}^l(t) \rangle \\ &= \langle b^l, p^l \rangle \in \mathcal{M}^l \times \mathcal{N}^l \setminus (\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t) \cup \bigcup_{k=1}^{j-1} \langle \hat{b}_{k, G\mathcal{H}_t^l}^l(t), \hat{p}_{k, G\mathcal{H}_t^l}^l(t) \rangle) \\ & \quad o^l \in O_{G_t^l} \setminus (\bigcup_{k=1}^{j-1} \{\delta_{k, G\mathcal{H}_t^l}^l(t)\}) \end{aligned} \quad \arg \max_{\langle b^l, p^l \rangle} \tilde{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l) \quad (7)$$

In (7), $j=1, \dots, (n - u^l(t))$. As shown in lines 15-17 in Algorithm 2, the EFML scheme will conduct an exploitation action when there are no under-explored beam-power pairs, and it will choose the n beam-power pairs $\langle \hat{b}_{1, G\mathcal{H}_t^l}^l(t), \hat{p}_{1, G\mathcal{H}_t^l}^l(t) \rangle, \dots, \langle \hat{b}_{n, G\mathcal{H}_t^l}^l(t), \hat{p}_{n, G\mathcal{H}_t^l}^l(t) \rangle$ from $\mathcal{M}^l \times \mathcal{N}^l$, which satisfy the following formula.

$$\begin{aligned} & \langle \hat{b}_{j, G\mathcal{H}_t^l}^l(t), \hat{p}_{j, G\mathcal{H}_t^l}^l(t) \rangle = \\ & \langle b^l, p^l \rangle \in \mathcal{M}^l \times \mathcal{N}^l \setminus (\bigcup_{k=1}^{j-1} \langle \hat{b}_{k, G\mathcal{H}_t^l}^l(t), \hat{p}_{k, G\mathcal{H}_t^l}^l(t) \rangle) \\ & \quad o^l \in s^l \in O_{G_t^l} \setminus (\bigcup_{k=1}^{j-1} \{\delta_{k, G\mathcal{H}_t^l}^l(t)\}) \end{aligned} \quad \arg \max_{\langle b^l, p^l \rangle} \tilde{e}_{\langle b^l, p^l \rangle}(o_{t,i}^l) \quad (8)$$

In (8), $j=1, \dots, n$. After choosing the n beam-power pairs from each sector respectively, the EFML scheme will reselect no more than n beam-power pairs from all the chosen beam-power pairs of all the sectors, as described in the lines 8-18 in Algorithm 1. After this, the EFML scheme observes the received data of each vehicle group $g_{t,i}^l$ with the context $o_{t,i}^l \in s_{t,i}^l \in O_{G_t^l}$ in each beam-power $\langle b_{t,i}^l, p_{t,i}^l \rangle$, and estimates the energy efficiency of each vehicle group $g_{t,i}^l$ according to the formula (1) (see lines 2-3 in Algorithm 3).

Based on these observations and estimations, the EFML scheme updates its internal counters (see lines 4-13 in

Algorithm 3), where the weight coefficient ζ denotes the contribution of the recently observed beam-power pair performance to the current updated beam-power pair performance, and usually depends on the empirical values with respect to the communication performance of the system.

Algorithm 1: Pseudocode of the EFML scheme

Input: $T, L, O_\tau, K(t)$

Output: the chosen beam-power pairs and the specified vehicle groups using the beam-power pairs

```

1: For each  $l=1, \dots, L$  do
2:   Initialize context partition: Create partition  $O_{G_t^l}$  into  $O_\tau$  subspaces with same size.
3:   Initialize counters:
   for all  $\langle b, p \rangle \in \mathcal{M}^l \times \mathcal{N}^l$  and  $s \in O_{G_t^l}$ , set  $N_{\langle b, p \rangle, s}^l = 0$ 
4:   Initialize estimators:
   for all  $\langle b, p \rangle \in \mathcal{M}^l \times \mathcal{N}^l$  and  $s \in O_{G_t^l}$ , set  $\hat{e}_{\langle b, p \rangle, s}^l = 0$ 
5: End for
6: For each  $t=1, \dots, T$  do
7:   Invoke Algorithm 2 to get  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$  and  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$ 
8:   If  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t) \neq \emptyset$  then ► Exploration
9:      $u = \text{size}(\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t))$ 
10:    If  $u \geq n$  then
11:      Select the  $n$  beam-power pairs randomly from  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ 
12:    Else
13:      Treat all the members of  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$  as the  $u$  beam-power pairs
14:      Select the  $(n - u)$  best beam-power pairs from  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$ 
15:    End if
16:  Else ► Exploitation
17:    Select the  $n$  best beam-power pairs from  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$ 
18:  End if
19:  Invoke Algorithm 3 to update estimators and counters
20: End for

```

Algorithm 2: Getting $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ and $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$

Input: *null*

Output: $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ and $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$

```

1:  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t) = \mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t) = \emptyset$ ;
2: For each  $l=1, \dots, L$  do
3:   Observe vehicle group contexts  $O_{G_t^l} = \{o_{t,i}^l | i \in \{1, \dots, G_t^l\}\}$ 
4:   Find  $G\mathcal{H}_t^l = \{s_{t,i}^l | i \in \{1, \dots, G_t^l\}\}$  such that  $o_{t,i}^l \in s_{t,i}^l \in O_{G_t^l}$ 
5:   Solve the set of under-explored beam-power pairs  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$  in (6)
6:   If  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t) \neq \emptyset$  then ► Exploration
7:      $u^l = \text{size}(\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t))$ 
8:     If  $u^l \geq n$  then
9:       Select  $\{\langle b_{t,i}^l, p_{t,i}^l \rangle | i \in \{1, \dots, n\}\}$  randomly from  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ 
10:      Add  $\{\langle b_{t,i}^l, p_{t,i}^l \rangle | i \in \{1, \dots, n\}\}$  to  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ 
11:     Else
12:       Select  $\{\langle b_{t,i}^l, p_{t,i}^l \rangle | i \in \{1, \dots, u^l\}\}$  as the  $u^l$  beam-power pairs from  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$  and then add them to  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ue}(t)$ 
13:       Select  $\langle b_{t,u^l+1}^l, p_{t,u^l+1}^l \rangle, \dots, \langle b_{t,n}^l, p_{t,n}^l \rangle$  as the  $(n - u^l)$  beam-power pairs  $\langle \hat{b}_{1, G\mathcal{H}_t^l}^l(t), \hat{p}_{1, G\mathcal{H}_t^l}^l(t) \rangle, \dots, \langle \hat{b}_{n-u^l, G\mathcal{H}_t^l}^l(t), \hat{p}_{n-u^l, G\mathcal{H}_t^l}^l(t) \rangle$  from (7) and then add them to  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$ 
14:     End if
15:   Else ► Exploitation
16:     Select  $\langle b_{t,1}^l, p_{t,1}^l \rangle, \dots, \langle b_{t,n}^l, p_{t,n}^l \rangle$  as the  $n$  beam-power pairs  $\langle \hat{b}_{1, G\mathcal{H}_t^l}^l(t), \hat{p}_{1, G\mathcal{H}_t^l}^l(t) \rangle, \dots, \langle \hat{b}_{n, G\mathcal{H}_t^l}^l(t), \hat{p}_{n, G\mathcal{H}_t^l}^l(t) \rangle$  from (8) and then add them to  $\mathcal{B}P_{G\mathcal{H}_t^l}^{ve}(t)$ 
17:   End if
18: End for

```

Algorithm 3: Updating estimators and counters

Input: the chosen beam-power pairs

Output: the updated estimators and counters

```

1: The random variables  $r_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l)$  and  $e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l)$  are initialized to a
   negative number
2: Observe the received data of each vehicle group  $g_{t,i}^l$  with the context
    $o_{t,i}^l \in S_{t,i}^l \in O_{g_t^l}$  in each beam-power  $\langle b_{t,j}^l, p_{t,j}^l \rangle$ , and assign it to
    $r_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l)$ 
3: Estimate the energy efficiency of each vehicle group  $g_{t,i}^l$  according to the
   formula (1), and assign it to  $e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l)$ 
4: For each  $l=1, \dots, L$  do
5:   For each  $i=1, \dots, G_t^l$  do
6:     For each  $j=1, \dots, n$  do
7:       If  $e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l) \geq 0$  then
8:          $\hat{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l) = (1 - \zeta) \cdot \hat{e}_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l) + \zeta \cdot e_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l(o_{t,i}^l)$ 
9:          $N_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l = N_{\langle b_{t,j}^l, p_{t,j}^l \rangle}^l + 1$ 
10:       End if
11:     End for
12:   End for
13: End for

```

V. PERFORMANCE EVALUATION

A. Simulation Settings

Fig. 3 shows the simulation scenario adopted in this paper, where the mmSBS's coverage is partitioned into the four sectors (i.e., $L=4$) with the equal size and there are differences in terms of road length and road distribution between the different sectors. Each sector is equally divided from two dimensions, where the mmSBS is firstly taken as the central point for equiangular division and then it is used as the starting point for equal length division. Based on the above spatial division principle, besides paying attention to vehicles' requesting data content, the LTE eNB should also make the registered vehicles located in the same angle range be the same group, where the furthest group member from the mmSBS is used to determine how far the group is from the mmSBS. For simplicity without loss of generality, the distance from the mmSBS is classified into three levels: Near, Moderate, and Far. The group distance from the mmSBS is regarded as 'Near' if it is less than R_N . Otherwise, it is regarded as 'Moderate' if it is less than R_M . Except for these two cases, everything else is considered to be 'Far'.

Like in [3], the seven types of beams are adopted in the EFML scheme based on the different beam widths (i.e. from 30° to 90° with the step size of 10°) for each sector, in which the number of beams per type is one. Thus, M^l is equal to 7. In addition, we set N^l is equal to 10, which means that the discrete transmission power values are set from 0.1 to 1 Watts with the step size of 0.1. In the simulations, a time slice is set to a fixed length of time by the mmSBS. For each vehicle group, its two-dimensional context vector involves the arrival direction dimension and the vehicle group distance dimension (i.e., $D=2$). The arrival direction of a vehicle group is defined as the angle between the positive X -axis in the plane coordinate system with the mmSBS as the origin and the line connecting the center point of the vehicle group with a mmSBS.

We denote the number of two-dimensional subspaces in each sector of the mmSBS by O_T and set $O_T = 18$. Furthermore, the parameter α and the length of a time slice t are set to $\alpha = 0.34$ and $t = 3$ s, respectively. Thus, according to Theorem 1, the time horizon T is approximately 6000, and the value of the control function $K(t)$ is approximately 2.03. We implement our simulation environment through the event-based network simulator OMNeT++, in which we also use the vehicular network simulation framework of Veins and the road traffic simulator of SUMO. In addition, we consider the vehicle's speed between 5m/s and 10m/s.

The mmWave channel propagation model in our simulation as the following formulation.

$$p_i^r = p_i^t G_i^t G_i^r G_i^c \quad (9)$$

where p_i^r is the received power of the vehicle i while p_i^t is the transmission power of the mmSBS, G_i^t and G_i^r are the directional transmission gain and directional reception gain between the vehicle i and the mmSBS, respectively, and G_i^c is the channel gain between the vehicle i and the mmSBS. The estimation of the gain parameters in the above channel model can be found in [31,32]. But for the convenience of readers, they are briefly stated as follows.

When a mmSBS selects a certain beam-power pair to the vehicle i , the transmission gain and the reception gain of this mmWave channel can be estimated by

$$G_i^t = \begin{cases} \frac{2\pi - (2\pi - \theta_i^t)\xi}{\theta_i^t}, & \text{if } |\varphi_i^t| \leq \frac{\theta_i^t}{2} \\ \xi, & \text{otherwise} \end{cases} \quad (10)$$

and

$$G_i^r = \begin{cases} \frac{2\pi - (2\pi - \theta_i^r)\xi}{\theta_i^r}, & \text{if } |\varphi_i^r| \leq \frac{\theta_i^r}{2} \\ \xi, & \text{otherwise} \end{cases} \quad (11)$$

In (10) and (11), θ_i^t denotes the transmitting beam width of the mmSBS, while θ_i^r is the receiving beam width of the vehicle i . In addition, θ_i^t and θ_i^r is the gain of the main lobe, while ξ represents the gain in the side lobe and $0 < \xi \ll 1$. φ_i^t denotes the angle between the line connecting the mmSBS with the vehicle i and the center line of the transmitting beam of the mmSBS, and φ_i^r is the angle between the line connecting the vehicle i with the mmSBS and the center line of the receiving beam of the vehicle i . According to [33], the channel gain G_i^c can be given by

$$G_i^c = |\chi_i^c \delta(\tau - \tau_i)|^2 \quad (12)$$

where $\delta(\cdot)$ denotes the Dirac delta function, and χ_i^c is the amplitude of the path from the vehicle i to the mmSBS, while τ_i is the propagation delay of the path from the vehicle i to the mmSBS. And τ_i can be estimated by the following expression.

$$\tau_i = \frac{d_i}{c} \quad (13)$$

In (13), c and d_i are the speed of light and the distance of the path from the vehicle i to the mmSBS, respectively. Wireless signal transmission methods include Line-of-Sight (LOS) transmission and Non-Line-of-Sight (NLOS) transmission. When there are buildings and plants between the

transmitter and the receiver, the NLOS path will have some problems such as high pathloss, reflection and penetration loss. Here, we consider only one reflection of a given path. According to [33], we can get the estimation of the amplitude of LOS and NLOS path as follows.

$$\chi_i^c = \begin{cases} \frac{\lambda}{4\pi d_i}, & \text{in LOS path} \\ \frac{\lambda}{4\pi d_i} \partial, & \text{in NLOS path} \end{cases} \quad (14)$$

where λ denotes the wavelength of the mmWave in this simulation, which can be estimated by $\lambda = c/f_c$ and f_c is the carrier frequency. ∂ is the reflection coefficient of the path between the vehicle and the mmSBS.

B. Simulation Schemes and Performance Metrics

The EFML scheme is most similar to the works in [2,3]. However, the work in [2] only considers the one-dimensional context vector. Although the work in [3] considers the two-dimensional context vector, it uses the identifier of road and the direction of arrival as the context instead of the arrival direction and the vehicle group distance. Thus, in order to compare the difference between considering the energy efficiency as optimization goal and taking the amount of received data as optimization goal under the same context, we design a comparison scheme, which is called NFML for convenience. The NFML scheme keeps the common core ideas in [2,3] (i.e., taking the amount of received data as optimization goal, adopting the consistent transmission power, no using group mode), but the other will be the same as our EFML scheme.

Also, we adopt a User Experience Quality Assurance (UEQA) for vehicle group as another comparison of our EFML scheme. Based on the approximate estimation formula $f(\gamma_i) = 1 - e^{-0.5\gamma_i}$ in [34], the bit transmission success rate from the mmSBS to a vehicle can be easily estimated if the signal-to-noise ratio (SNR) γ_i at this vehicle is known. If we know that the receiving bit error rate (BER) threshold BER_i^{th} of each vehicle, we can estimate the corresponding SNR threshold value γ_i^{th} by letting BER_i^{th} be equal to $e^{-0.5\gamma_i^{th}}$, which can be expressed by

$$\gamma_i^{th} = -2 \ln BER_i^{th} \quad (15)$$

In order to ensure that the BER level of each vehicle's receiving data from the mmSBS is not higher than BER_i^{th} , the transmission power of the mmSBS should not be lower than the transmission power threshold p_i^{th} , which is estimated as follows.

$$p_i^{th} = \frac{\gamma_i^{th} W N_0}{G_i^t G_i^r G_i^c} \quad (16)$$

In (16), W and N_0 represent the bandwidth of mmWave band and the background noise power spectrum density, respectively. In other words, when the mmSBS adopts p_i^{th} to send data to the vehicle i , the bit transmission success rate from the mmSBS to the vehicle i can be expressed by

$$f(\gamma_i^{th}) = 1 - e^{-0.5\gamma_i^{th}} \quad (17)$$

Thus, combined with Shannon theorem, the energy efficiency of the UEQA from the mmSBS to the vehicle i can be expressed by

$$e(\gamma_i^{th}) = \frac{W \log_2(1 + \gamma_i^{th})}{p_i^{th}} \quad (18)$$

The energy efficiency of a vehicle group is determined by the vehicle that has the minimum energy efficiency of bit transmission in all the members of a vehicle group.

The performance metrics adopted in the simulation experiments are the energy efficiency, the online learning cost, the cumulative received data, and the aggregate received data. The energy efficiency of the EFML and NFML is defined in the formula (1) while that of the UEQA is defined in the formula (18). The definition of online learning cost is the number of explorations rounds that each of the three schemes to achieve a certain percentage of the performance of the optimal solution, and all exploration operations in each discrete time slice are regarded as one round of exploration. The cumulative received data for all the three schemes is defined as the amount of data received by all the vehicles during the time horizon T , while the aggregate received data for all the three schemes is defined as the amount of data (in Gbits) received by all the vehicles during a time slice. Unless otherwise stated below, the values of the rest simulation parameters are listed in Table 1.

Table 1: Simulation parameters

Symbol	Description	Value
f_c	Carrier frequency	60 GHz
W	System bandwidth	1 GHz
p_{max}^t	The mmSBS's maximum transmission power	1 W
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 at a mmSBS	6
t	The length of a time period	3 s
T	The number of time periods	6000
ζ	The weighted coefficient	0.125
BER_i^{th}	The receiving bit error rate threshold	10^{-10}
R_F	Radius of simulation area of a mmSBS	100 m

C. Analysis of Simulation Results

We evaluate the performance metrics of the EFML scheme compared with the benchmark schemes such as the NFML scheme and the UEQA scheme. In Fig.4~Fig.6, we investigate the impact of the number of vehicles in the simulation area on the performance metrics, in which no more than 6 selected beam-power pairs are employed simultaneously in each time slice and the number of vehicles ranges from 35 to 95 with the step size of 15.

As shown in Fig. 4, we can see that the cumulative received data increases the number of vehicles in the simulation area. The reason is that the fewer the number of vehicles, the less contextual information in the system, which ultimately leads to a poorer learning effect. Relatively speaking, it is less

likely that the system can accurately select the beam-power pair that maximizes the received data. However, the increase in the number of vehicles means that more contextual information can be provided, which is conducive to getting a better learning effect. In this case, the system is more likely to precisely select the beam-power pair that maximizes the received data.

It can also be observed in Fig. 4 that the EFML outperforms the NFML and the UEQA in the cumulative received data. There are two reasons for this occurrence: for one thing, compared with the NFML, the EFML can provide service for each vehicle group with the same request (i.e., multicast), while the NFML focuses on unicast communication (i.e., only one of the vehicle group can be serviced). For another thing, the UEQA only considers the power that meets the worst SNR within the vehicle group, while the EFML allows the power to be adjusted while satisfied the BER threshold of the system. That is, in addition to finding a more appropriate beam orientation and beam width for vehicle groups, it can also provide a more suitable power to reduce the power consumption.

Fig. 5 shows the online learning costs of EFML, NFML and UEQA under different vehicle numbers with the same simulation configurations as Fig. 4. We can observe from Fig. 5 that the number of exploration costs of online learning decreases with the number of vehicles in the simulation area. The main reason is that the greater the number of vehicles that enter the system in a scheduling period, the greater the number of corresponding context subspaces, so the performance of each beam-power pair in more subspaces can be detected. If it is found that there is no historical performance data or the recorded historical data is not sufficient, the exploration schedule should start as soon as possible. Therefore, more beams can be scheduled for detection in a scheduling time slice, which is beneficial to speed up the detection process of the performance of each beam-power pair in each context subspace, and thus effectively reduce the number of exploration rounds.

In Fig. 5, we can also see that the cost of the online learning of the EFML is higher than that of the NFML and the UEQA. This is because neither the NFML nor the UEQA considers the dimension of power adjustment and thus they have the smaller learning context space than the EFML considering the adjustment of transmission powers. That is, the EFML needs to explore a larger learning space, which aims to select the more appropriate transmission power for each mmSBS at higher online learning costs.

In Fig. 6, we can see that as the number of vehicles in the simulation area increases, the energy efficiency of the network also increases. The main reason is that under the condition that the number of concurrent beams is limited, the more vehicles means that the probability of selecting the vehicles with relatively high data rate and relatively low power consumption is greater.

Moreover, Fig. 6 shows that the energy efficiency achieved by the EFML is better than that achieved by the NFML and the UEQA. This phenomenon can be explained in the two aspects. Firstly, the NFML adopts a fixed transmission power while the UEQA only considers meeting the worst SNR in a certain vehicle group. That is, the adjustability of the transmission power is not considered in these two schemes. However, in addition to meeting the BER threshold for vehicles within the group, the EFML can adjust a more appropriate transmission power for each mmSBS and thus it improves the energy efficiency of the system. Secondly, in the EFML, the vehicles with the same content request and in close proximity are constantly grouped together to share the same mmWave beam-power pair and thus they save power consumption and improve the energy efficiency of the system.

In Fig.7~Fig.9, we investigate the impact of the number of selected beam-power pairs per time slice on the performance metrics, in which the number of vehicles in the simulation area is set to 65 and the number of selected beam-power pairs per time slice varies from 2 to 6 with the step size of 1. In Fig. 7, it can be seen that as the number of beam-power pairs that can be used concurrently in each time slice increases, the cumulative received data of all schemes also increases. This is because the more number of selected beam-power pairs that can be used concurrently means that the more vehicles can be served at the same time. We can also find from Fig.7 that cumulative received data achieved by the EFML is higher than that achieved by the NFML and the UEQA. The explanation of the difference of cumulative received data between different schemes is similar to that of the result in Fig. 4.

From Fig. 8, we observe that the number of exploration rounds decreases with the number of selected beam-power pairs per time slice. The more number of beam-power pairs that can be used concurrently in each time slice means that the more number of beam-power pairs with unknown performance or uncertain performance information that may be explored in the same time slice is more. Therefore, when the number of context spaces of the system is fixed, the number of exploration rounds will decrease as the number of selected beam-power pairs per time slice increases. It can also be seen that the cost of online learning of the EFML is higher than that of other schemes, and the explanation of the difference among different schemes is similar to that of the results in Fig. 5.

Fig. 9 shows that the energy efficiency slightly decreases with the number of selected beam-power pairs per time slice. As mentioned earlier, in any exploitation process, the EFML will select the beam-power pairs that have been proved to have the best performance in the previous time slices. In this case, the average energy efficiency will be higher if a smaller number of optimal beam-power pairs are selected. We can also see from Fig. 9 that the energy efficiency achieved by the EFML is higher than that achieved by the NFML and the UEQA, and the

explanation for this difference is similar to that of the results in Fig. 6.

In Fig.10~Fig.12, we investigate the effect of different thermal noise power density on the performance metrics, in which the number of selected beam-power pairs per time slice is set to 6, the number of vehicles in the simulation area is set to 65, and the thermal noise power density ranges from -170 dBm/Hz to -150 dBm/Hz in steps of 5 dBm/Hz. It is obvious from Fig. 10 that cumulative received data decreases with thermal noise power density. This is due to the fact that the SNR of any receiver will be affected by thermal noise density. It is easy to know that the data rate will decrease with the SNR by the Shannon theorem.

Furthermore, we can see from Fig. 10 that the cumulative received data of UEQA is almost unaffected when thermal noise power density is relatively small. This is because that the UEQA adjusts the transmission power to maintain a certain SNR to meet the BER threshold of the system according to the formula (16), and thus the cumulative received data remains almost unchanged. However, when thermal noise power density is too large, the transmitter may not meet the BER threshold of the system even if the transmission power is adjusted to the maximum value. At this time, the cumulative received data will decrease with thermal noise power density. It can also be seen from Fig. 10 that the cumulative received data of the EFML is more than that of the NFML and the UEQA, where the explanation of the reason for the difference is similar to the explanation of the result of Fig. 4.

It can be seen from Fig. 11 that thermal noise density has almost no effect on the cost of online learning. This is because that the size of context space required by the online learning algorithm will not change with the change of thermal noise density. Moreover, the explanation of the difference in the number of the online learning costs among different schemes is similar to the explanation of the results in Fig. 5.

From Fig. 12, we can observe that the network energy efficiency will decline significantly as thermal noise density increases. This decrease is due to the two reasons. On the one hand, we can know that the cumulative received data decreases with thermal noise power density based on the results in Fig. 10. On the other hand, the NFML always maintains a fixed transmitting power, while the EFML and the UEQA have to increase the transmitting power of each mmSBS to ensure the SNR of each vehicle satisfies the BER threshold of the system as thermal noise density increases and thereby decrease the network energy efficiency. We can also see from Fig. 12 that the difference in the network energy efficiency among different schemes, and the explanation of this difference is similar to the result in Fig. 6.

In Fig.13~Fig.15, we analyze the performance metrics achieved by the schemes over the time horizon with 6000 time slices, in which the number of selected beam-power pairs per time slice is 6 and the number of vehicles in the simulation area

is set to 65. Fig. 13 shows that the aggregate received data achieved by different schemes over a time horizon of 6000 time slices. Specifically, it also shows that the EFML can achieve better performance than the NFML and the UEQA after the 2200th time slice. The fluctuations on the graph are caused by the number of the vehicle groups, the contact time and the speed of each vehicle.

We can see from Fig. 13 that the aggregate received data per time slice achieved by the EFML began to show an upward trend after the 1300th time slice and higher than the NFML and the UEQA after the 2200th time slice. This is because that the context space of the EFML is larger than that of the NFML and the UEQA. Due to the insufficient online learning before the 1300th time slice, most of the beam-power pairs allocated by the EFML to the vehicles are selected randomly. Moreover, the NFML and the UEQA may have entered the exploitation phase while the EFML is still in the exploration phase. So the cumulative received data of the EFML may not be as good as the other two schemes before the 1300th time slice. However, after a period of sufficient online learning, the EFML can choose a set of beam-power pairs that are more reasonable than the other schemes, including beam directions, beam widths and transmission powers.

By using the same simulation parameters as in Fig. 13, Fig. 14 shows the number of exploration operations per time slice over the time horizon of 6000 time slices. We can observe from Fig. 14 that the number of exploration operations in each time slice will decrease as the number of time slices increases. This is because the number of under-explored beam-power pairs decreases in the system over time. It can also be seen from Fig. 14 that the EFML requires more time slices to explore beam performance than the NFML and the UEQA. That is, it will take the longer time for the EFML to enter the exploitation phase. The reason is that the context subspace of the EFML is larger than that of the NFML and the UEQA.

As can be seen from Fig. 15, after a certain number of time slices, the energy efficiency per time slice achieved by the EFML is higher than that of the UEQA. The main reason is as mentioned above. Since the EFML has more context space than the UEQA, it needs more time slices to explore beam performance. As long as the learning is sufficient, the EFML can adjust and select a better power for the mmSBS to reduce the power consumption and improve the energy efficiency of the system. Also, we see that the NFML is always less energy efficient than the other schemes over a time horizon of 6000 time slices. This is because it always selects the maximum transmission power for each mmSBS without reasonable power adjustment.

VI. CONCLUSIONS

In this paper, we proposed the EFML scheme to improve network energy efficiency in cellular-assisted vehicular networks based on the MAB theory. By reducing power

consumption as far as possible under the premise of meeting the basic data rate requirements of vehicle users, the EFML scheme avoids unnecessary power consumption. By grouping the users requesting the same data content in close proximity into the same receiving group, the EFML scheme save mmWave beams and reduce the occupation of RF chains. The simulation results show that, compared with the existing online learning-based mmWave beam selection schemes, the EFML scheme not only improves the energy efficiency but also the amount data in cellular-assisted vehicular networks at the cost of more system overhead. However, there is no difference in terms of beam performance update cost after a certain number of time slices between the EFML scheme and the comparison schemes.

Methods/Experimental

The simulation scenario is shown in Fig. 3, where the coverage radius of each mmSBS is 100 meters and each mmSBS coverage is partitioned into the four sectors (i.e., $L=4$). Each sector is equally divided from two dimensions, where the mmSBS is firstly taken as the central point for equiangular division and then it is used as the starting point for equal length division. We implement our simulation environment through the event-based network simulator OMNeT++, in which we also use the vehicular network simulation framework of Veins and the road traffic simulator of SUMO. In addition, we consider the vehicle's speed between 5m/s and 10m/s. In order to compare the difference between considering the energy efficiency as optimization goal and taking the amount of received data as optimization goal under the same context, we design a comparison scheme, which is called NFML for convenience. Also, we adopt the UEQA scheme as another comparison of our EFML scheme.

Availability of data and materials

Not applicable

Competing interests

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

Funding

This work was supported in part by the National Natural Science Foundation of China (61873352).

Authors' contributions

JSG presented the scheme and designed the experiments. YL did the experiments, analyzed the data, and explained the simulation results. JSG drafted this paper, and all authors read and approved it.

Acknowledgements

The authors would like to acknowledge the anonymous reviewers for their thoughtful comments.

Authors' information

Jinsong Gui received the PhD from Central South University, China, in 2008. He is currently a Professor in School of Computer Science and Engineering, Central South University, China. Yao Liu is currently a Master Student working in School of Computer Science and Engineering, Central South University, China.

Abbreviations

C-ITS: Cooperative Intelligent Transport System; V2X: Vehicle to Everything; LTE: Long Term Evolution; 5G: Fifth-Generation; mmWave: millimeter Wave; mmSBS: mmWave Small Base Station; FML: Fast Machine Learning; IFML: Improved Fast Machine Learning; RF: Radio Frequency; EFML: Energy efficiency-based FML; MIMO: Multi-input Multi-output; NR: New Radio; MAC: Multiple Access Control; MAB: Multi-Armed Bandit; QoS: Quality of service; CSI: Channel State Information; LOS: Line-of-Sight; NLOS: Non-Line-of-Sight; UEQA: User Experience Quality Assurance.

REFERENCES

- [1] A. Alkhateeb, Y.H. Nam, M.S. Rahman, C. Zhang, R. Heath, Initial beam association in millimeter wave cellular systems: analysis and design insights. *IEEE Trans. Wireless Commun.* 16(5), 2807–2821(2017)
- [2] G.H. Sim, S. Klos, A. Asadi, A. Klein, M. Hollick, An online context-aware machine learning algorithm for 5G mmWave vehicular communications. *IEEE/ACM Transactions on Networking* 26(6), 2487–2500 (2018)
- [3] J.S. Gui, Y. Liu, X.H. Deng, B. Liu, Network capacity optimization for cellular-assisted vehicular systems by online learning-based mmWave beam selection. *Wireless Communications and Mobile Computing* 2021, Article ID 8876186, 26 pages (2021)
- [4] S. Han, C.-L. I, Z. Xu, C. Rowell, Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Commun. Mag.* 53(1), 186–194 (2015)
- [5] P. Kela, M. Costa, J. Turkka, M. Koivisto, J. Werner, A. Hakkarainen, M. Valkama, R. Jantti, K. Leppanen, in Proceedings of IEEE 84th Vehicular Technology Conference (VTC-Fall), Montreal, QC, Canada. Location Based beamforming in 5G ultra-dense networks (2016), pp. 1-7
- [6] P. Kela, M. Costa, K. Leppänen, R. Jäntti, in Proceedings of IEEE Conference on Standards for Communications and Networking (CSCN), Helsinki, Finland. Location-aware beamformed downlink control channel for ultra-dense networks (2017), pp. 7-11
- [7] M.S. Sim, Y.G. Lim, S.H. Park, L.L. Dai, 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* 8(1), 51634-51646 (2020)
- [8] W.Y. Ma, C.H. Qi, G.Y. Li, Machine learning for beam alignment in millimeter wave massive MIMO. *IEEE Wireless Communications Letters* 9(6), 875-878 (2020)
- [9] Y. Yang, Y. He, D. He, Z. Gao, Y. Luo, in Proceedings of IEEE Globecom Workshops (GC Wkshps), Waikoloa, HI, US. Machine learning based analog beam selection for 5G mmWave small cell networks (2019), pp. 1-5
- [10] D.Y. Zhang, A. Li, M. Shirvanimoghaddam, P. Cheng, Y.H. Li, B. Vucetic, Codebook-based training beam sequence design for millimeter-wave tracking systems. *IEEE Transactions on Wireless Communications* 18(11), 5333-5349 (2019)
- [11] Y.N.R. Li, B. Gao, X.D. Zhang, K.B. Huang, Beam management in millimeter-wave communications for 5G and beyond. *IEEE Access* 8(1), 13282-13293 (2020)
- [12] J.C. Fan, L.Y. Han, X.M. Luo, J. Huang, Delay analysis and optimization of beam scanning-based user discovery in millimeter wave systems. *IEEE Access* 8(1), 25075-25083 (2020)

- [13] J. Yang, S. Jin, C.K. Wen, X. Yang, M. Matthaiou, Fast beam training architecture for hybrid mmWave transceivers. *IEEE Transactions on Vehicular Technology* 69(3), 2700-2715, (2020)
- [14] F. Fernandes, C. Rom, J. Harrebek, G. Berardinelli, in *Proceedings of IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, Helsinki, Finland. Beam management in mmWave 5G NR: an intra-cell mobility study (2021), pp. 1-7
- [15] H. T. Nguyen, H. Murakami, K. Nguyen, K. Ishizu, W.J. Hwang, Joint user association and power allocation for millimeter-wave ultra-dense networks. *Mobile Networks and Applications* 25(1), 274–284 (2020)
- [16] G. Kwon, H. Park, Joint user association and beamforming design for millimeter wave UDN with wireless backhaul. *IEEE Journal on Selected Areas in Communications* 37(12), 2653–2668 (2019)
- [17] Y. Liu, A. Tang, X. Wang, Joint incentive and resource allocation design for user provided network under 5G integrated access and backhaul networks. *IEEE Transactions on Network Science and Engineering* 7(2), 673–685 (2020)
- [18] A. Ali, N. González-Prelcic, R.W. Heath, Millimeter wave beam selection using out-of-band spatial information. *IEEE Transactions on Wireless Communications* 17(2), 1038-1052 (2018)
- [19] A. Ahmed, A. Sam, V. Paul, L. Ying, Q. Qu, T. Djordje, Deep learning coordinated beamforming for highly-mobile millimeter wave systems. *IEEE Access* 6(1), 37328-37348 (2018)
- [20] I. Mavromatis, A. Tassi, R.J. Piechocki, A. Nix, in *Proceedings of IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Toronto, ON, Canada. MmWave system for future ITS: a MAC-layer approach for V2X beam steering (2017), pp. 1-6
- [21] Y.Y. Wang, M. Narasimha, R. Heath, in *Proceedings of IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Greece. MmWave beam prediction with situational awareness: a machine learning approach (2018), pp. 1-5
- [22] S.B. Chen, Z.Y. Jiang, S.Z. Zhou, Z.S. Niu, in *Proceedings of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, USA. Time sequence channel inference for beam alignment in vehicular networks (2018), pp. 1199-1203
- [23] R. Benelmir, S. Bitam, A. Mellouk, in *Proceedings of IEEE 22nd International Conference on High Performance Switching and Routing (HPSR)*, Paris, France. Simulated annealing-based beam management for 5G vehicular networks (2021), pp. 1-4
- [24] K. Z. Ghafoor, L. Kong, S. Zeadally, A.S. Sadiq, G. Epiphaniou, M. Hammoudeh, A.K. Bashir, S. Mumtaz, Millimeter wave communication for internet of vehicles: status, challenges, and perspectives. *IEEE Internet of Things Journal* 7(9), 8525-8546 (2020)
- [25] S. Gyawali, S. Xu, Y. Qian, R. Q. Hu, Challenges and solutions for cellular based V2X communications. *IEEE Communications Surveys & Tutorials* 23(1), 222-255 (2021)
- [26] S. Maghsudi, E. Hossain, Multi-armed bandits with application to 5G small cells. *IEEE Wireless Commun.* 23(3), 64–73 (2016)
- [27] M. Hashemi, A. Sabharwal, C. E. Koksal, N. B. Shroff, in *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*, Honolulu, HI, USA. Efficient beam alignment in millimeter wave systems using contextual bandits (2018), pp. 1–9
- [28] X.T. Li, R.T. Zhou, Y. Zhang, L. Jiao, Z.P. Li, Smart vehicular communication via 5G mmWaves. *Computer Networks* 172, 107173 (2020)
- [29] V. Va, T. Shimizu, G. Bansal, R. Heath, Online learning for position-aided millimeter wave beam training. *IEEE Access* 7(1), 30507-30526 (2019)
- [30] P. Auer, N. Cesa-Bianchi, P. Fischer, Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.* 47(2), 235–256 (2002)
- [31] O. Semiari, W. Saad, M. Bennis, Z. Dawy, Inter-operator resource management for millimeter wave multi-hop backhaul networks. *IEEE Trans. Wireless Commun.* 16(8), 5258–5272 (2017)
- [32] Q. Xue, X. Fang, C. X. Wang, Beamspace SU-MIMO for future millimeter wave wireless communications. *IEEE Journal on Selected Areas in Communications* 35(7), 1564–1575 (2017)
- [33] P. Liu, J. Blumenstein, N. S. Perovic, M. di Renzo, A. Springer, Performance of generalized spatial modulation MIMO over measured 60GHz indoor channels. *IEEE Transactions on Communications* 66(1), 133–148 (2018)
- [34] H.L. Ren, M.Q.H. Meng, Modeling of joint topology control and power scheduling for wireless heterogeneous sensor networks. *IEEE Transactions on Automation Science and Engineering* 6(4), 610-625 (2009)

Fig. 1. A mmWave/sub-6 GHz vehicular communication scenario with grouping mode

Fig. 2. Integration feasibility of EFML in a mmWave/sub-6 GHz cellular system

Fig. 3. The example of our simulation scenario

Fig. 4. Impact of vehicle density on cumulative received data

Fig. 5. Impact of vehicle density on online learning costs

Fig. 6. Impact of vehicle density on energy efficiency

Fig. 7. Impact of the number of selected beam-power pairs per time slice on cumulative received data

Fig. 8. Impact of the number of selected beam-power pairs per time slice on online learning costs

Fig. 9. Impact of the number of selected beam-power pairs per time slice on energy efficiency

Fig. 10. Impact of thermal noise power density on cumulative received data

Fig. 11. Impact of thermal noise power density on online learning costs

Fig. 12. Impact of thermal noise power density on energy efficiency

Fig. 13. Aggregate received data over the time horizon

Fig. 14. The number of exploration operations per time slice over the time horizon

Fig. 15. The energy efficiency over the time horizon

Figures

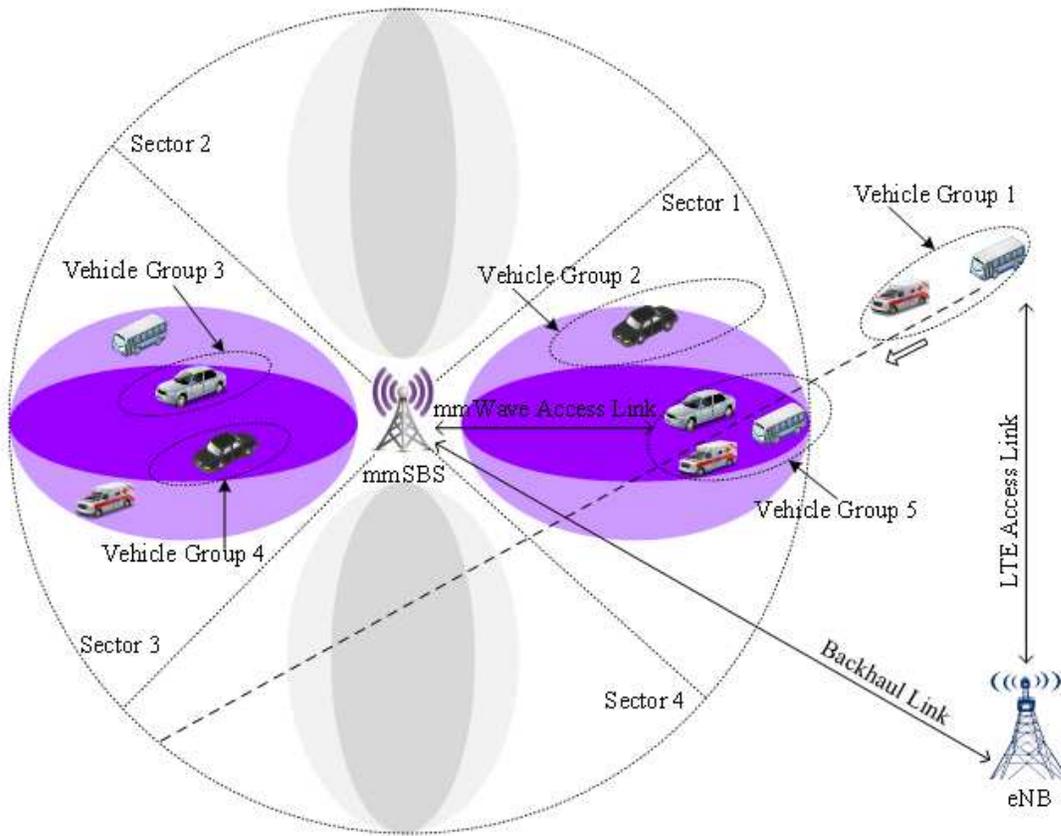


Figure 1

A mmWave/sub-6 GHz vehicular communication scenario with grouping mode

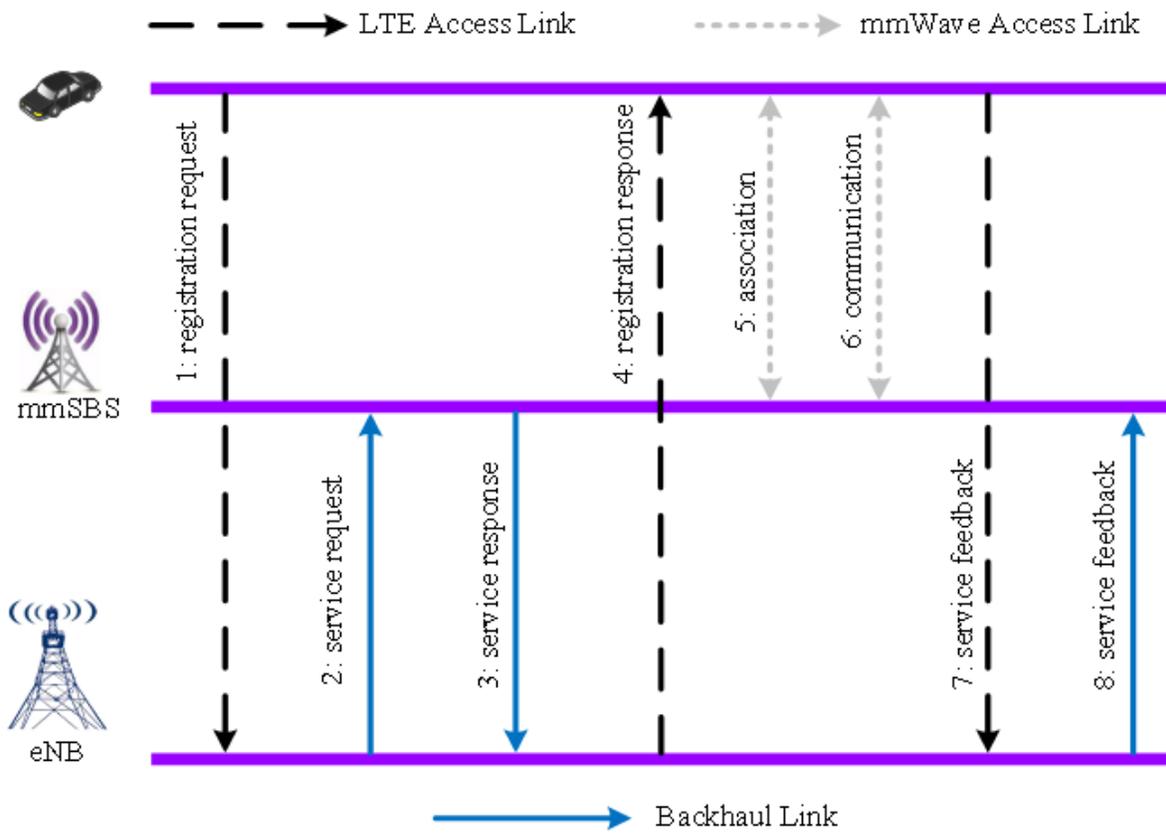


Figure 2

Integration feasibility of EFML in a mmWave/sub-6 GHz cellular system

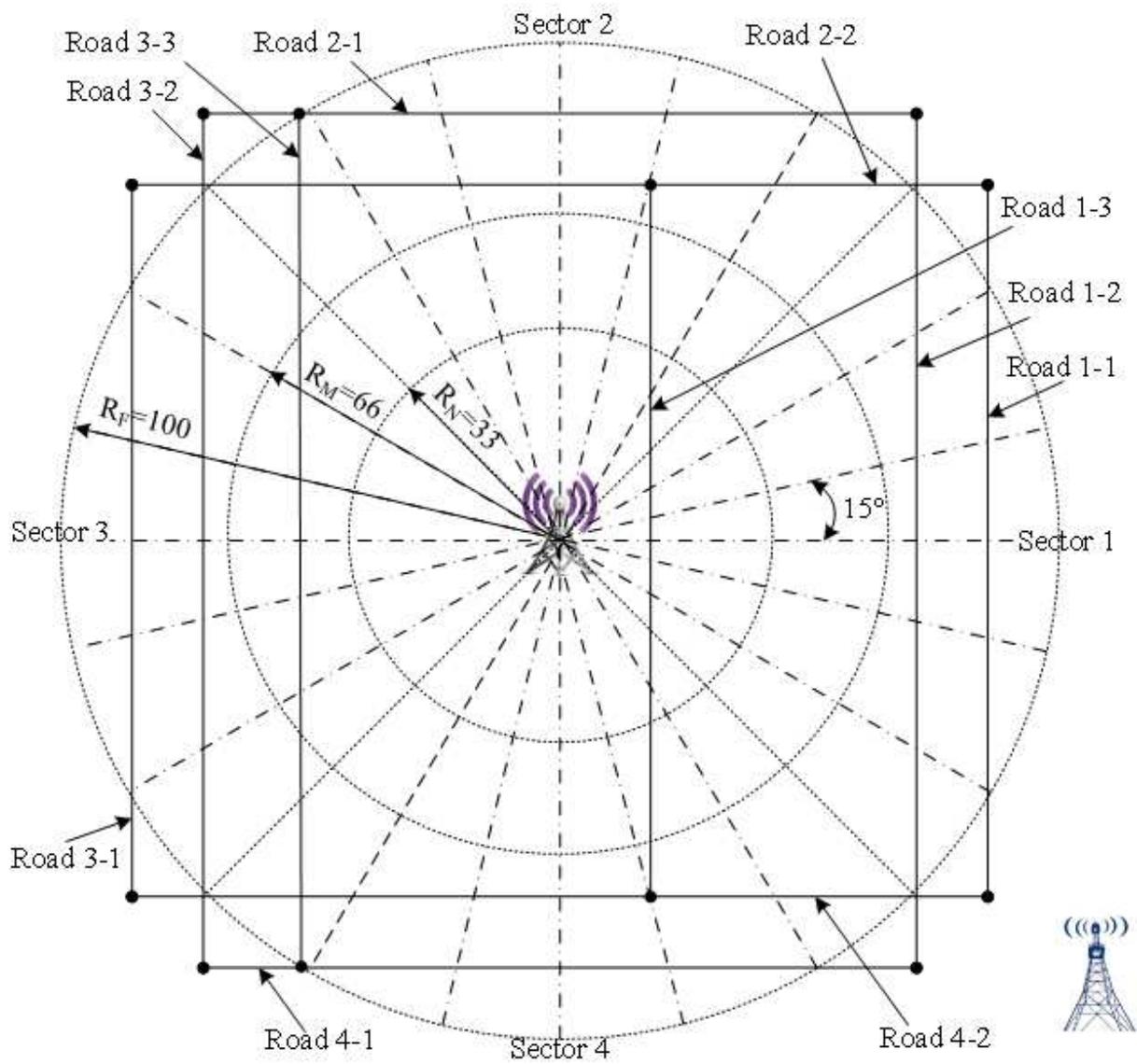


Figure 3

The example of our simulation scenario

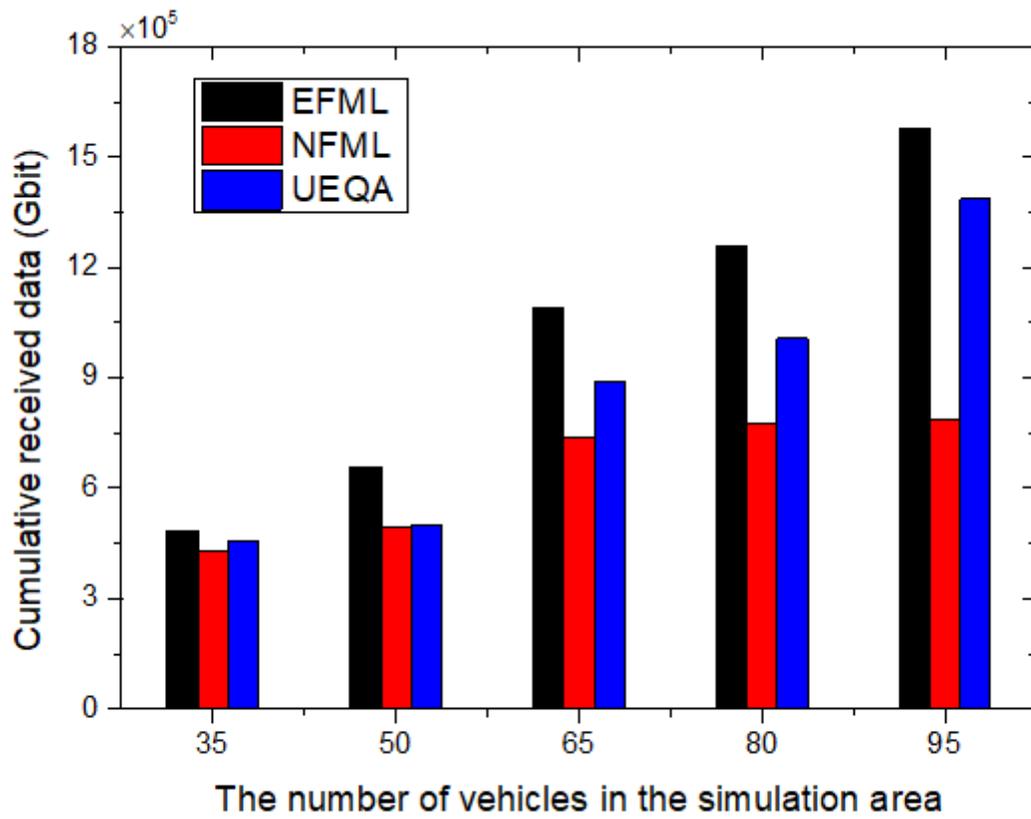


Figure 4

Impact of vehicle density on cumulative received data

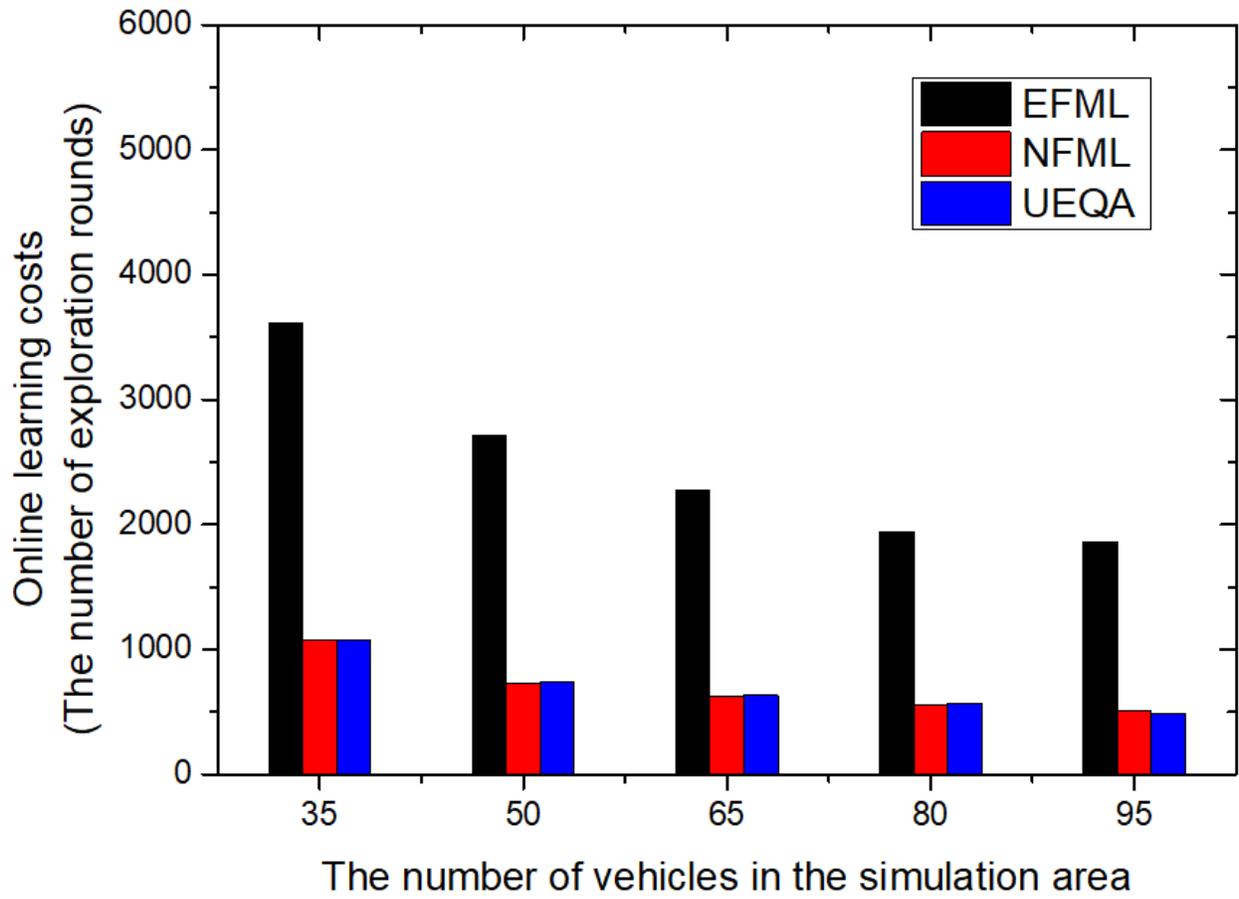


Figure 5

Impact of vehicle density on online learning costs

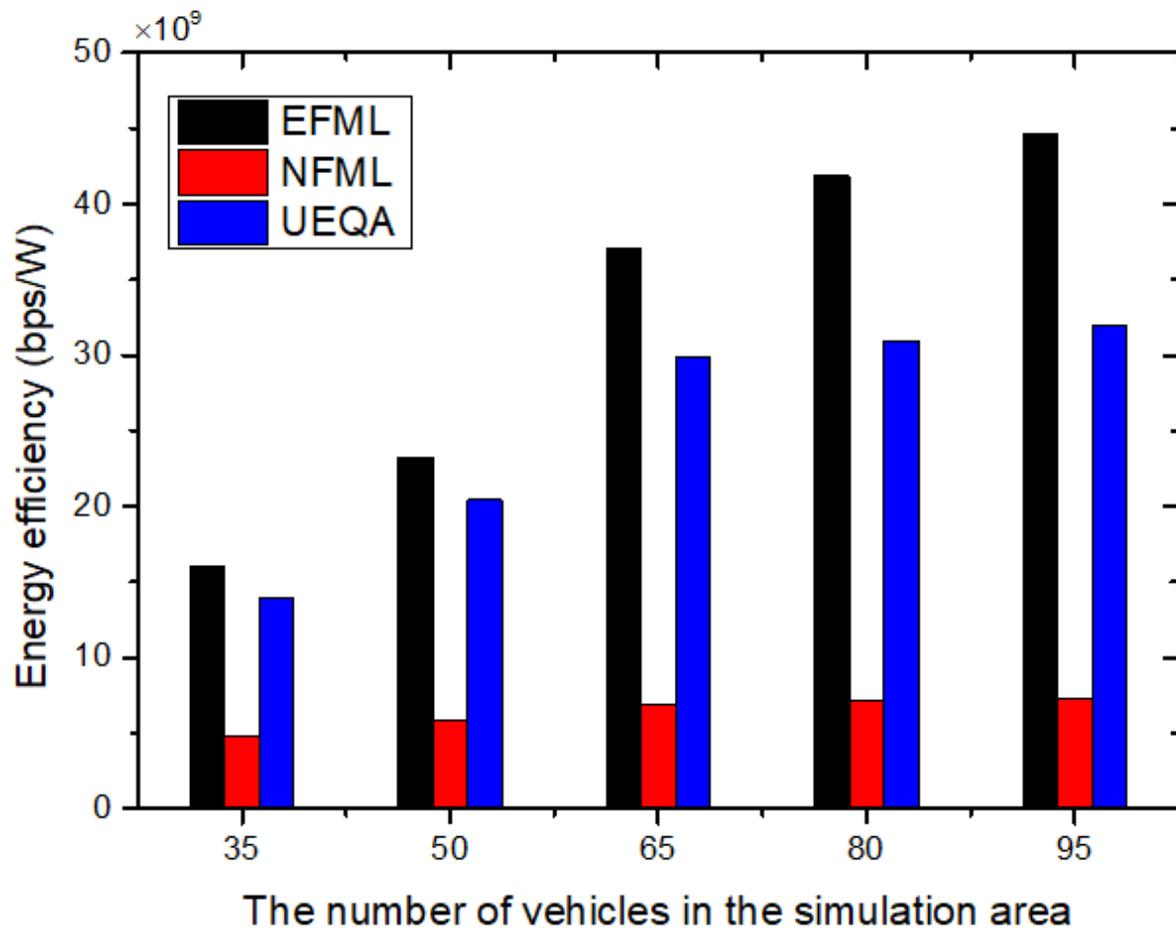


Figure 6

Impact of vehicle density on energy efficiency

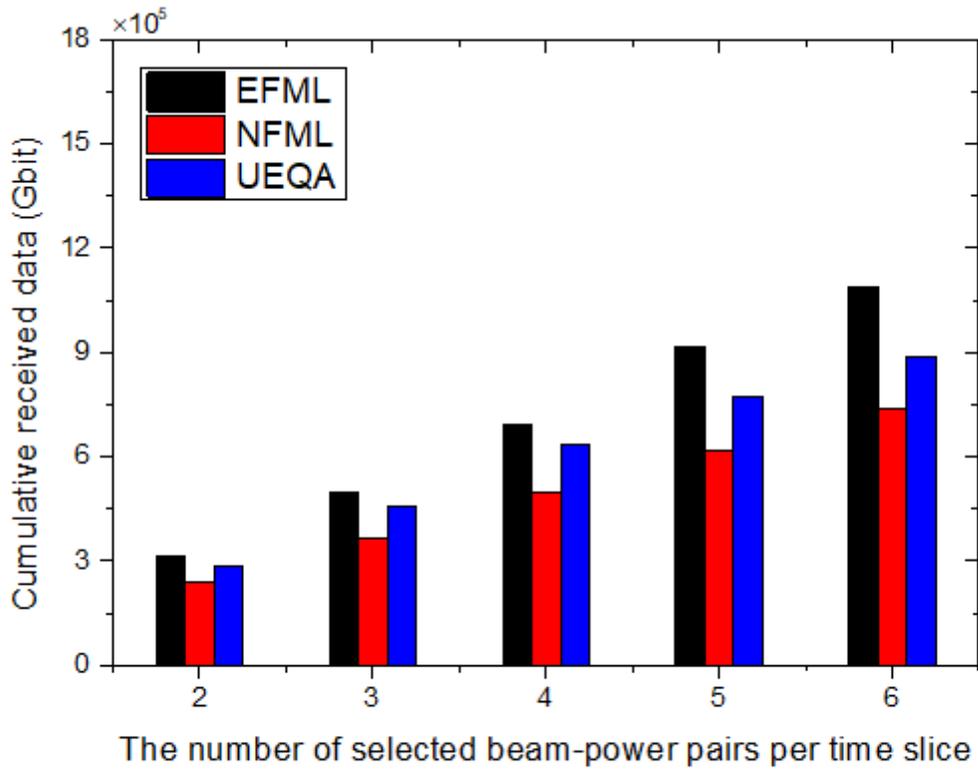


Figure 7

Impact of the number of selected beam-power pairs per time slice on cumulative received data

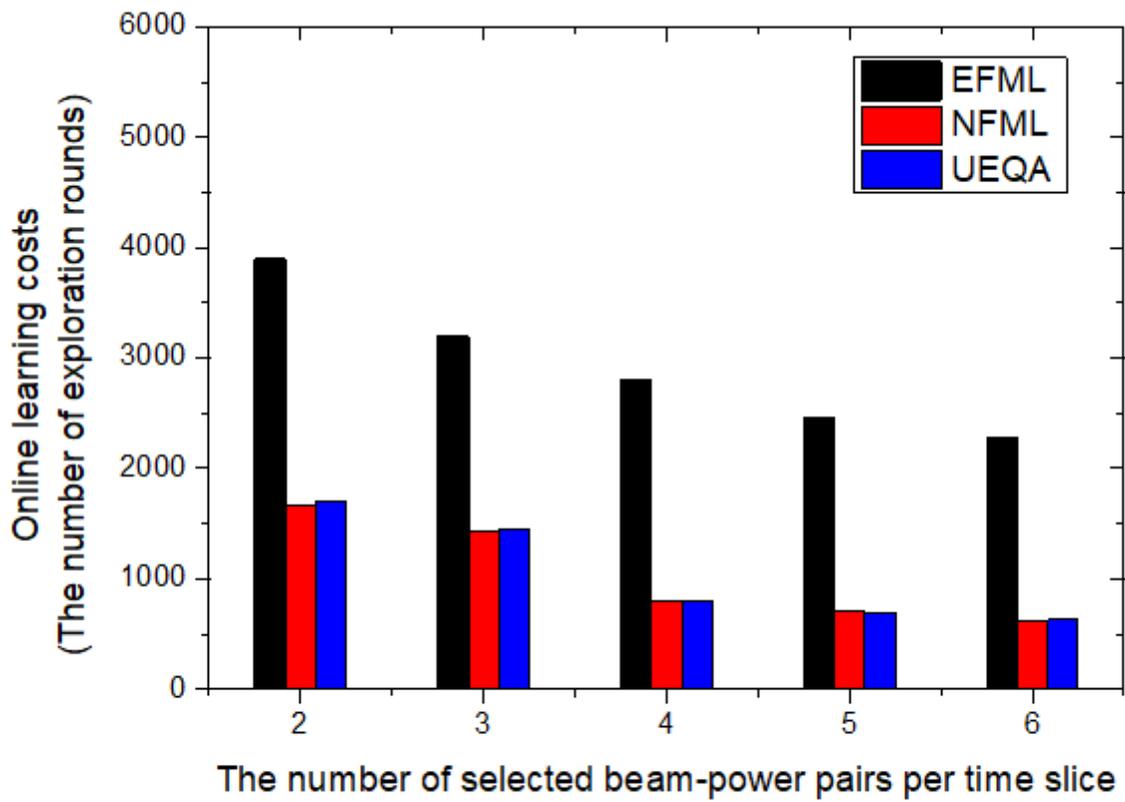


Figure 8

Impact of the number of selected beam-power pairs per time slice on online learning costs

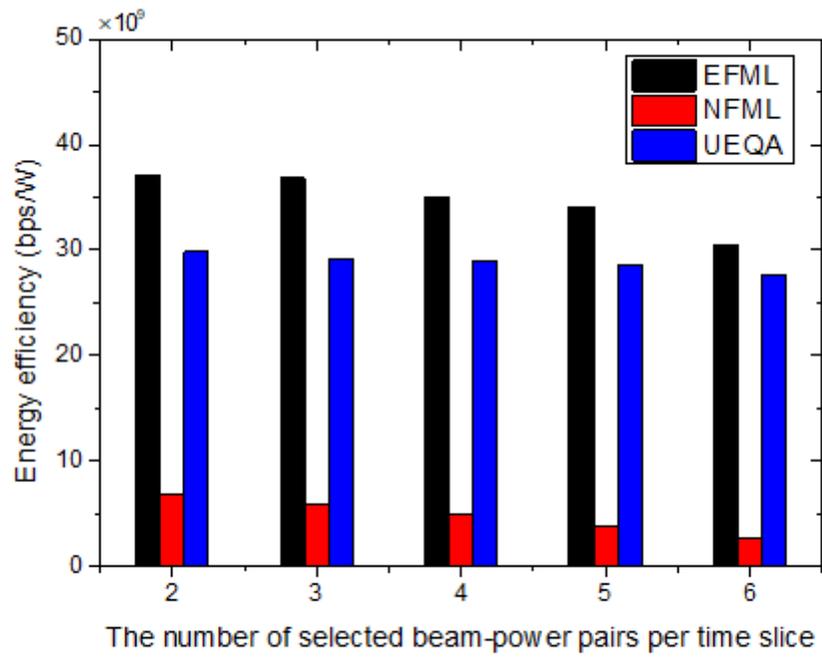


Figure 9

Impact of the number of selected beam-power pairs per time slice on energy efficiency

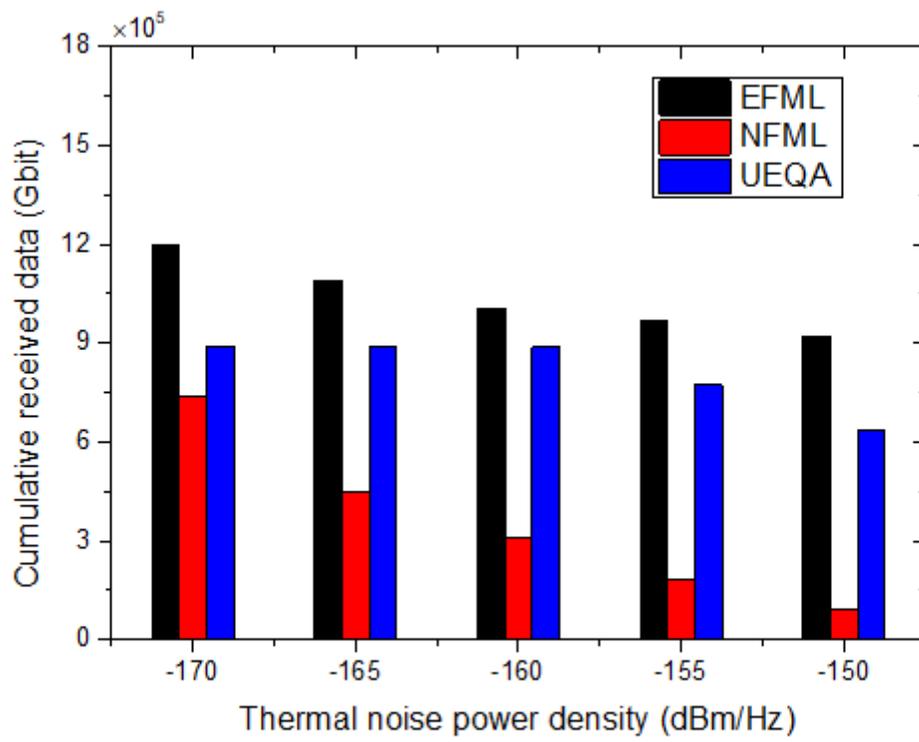


Figure 10

Impact of thermal noise power density on cumulative received data

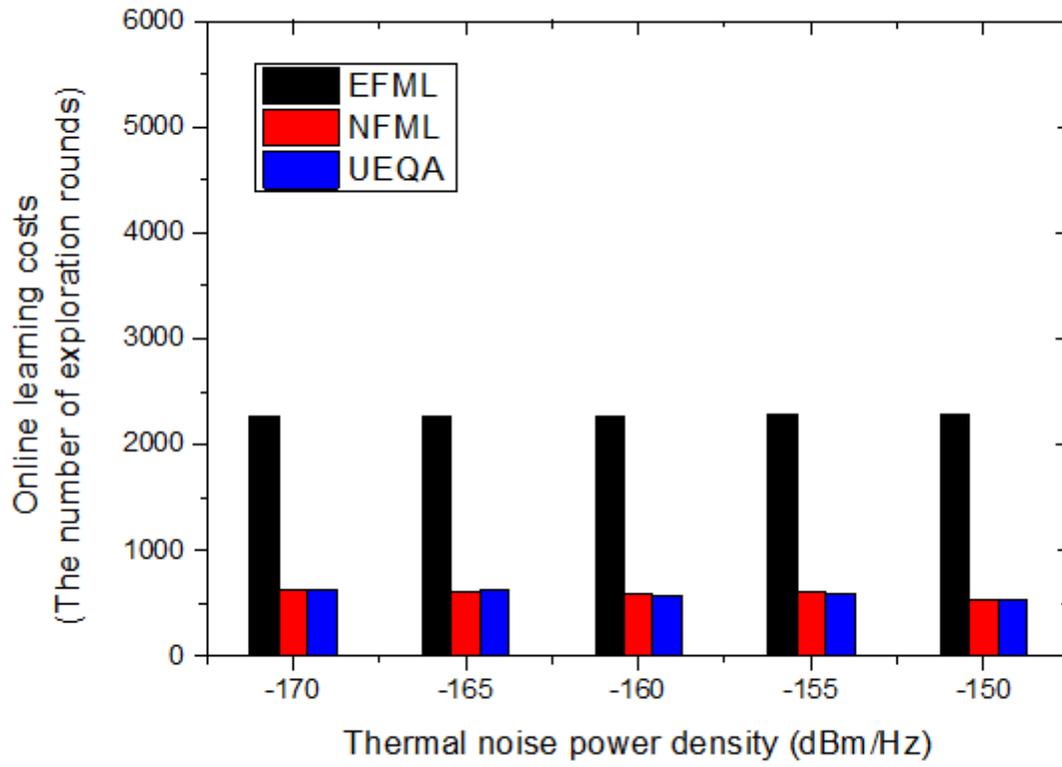


Figure 11

Impact of thermal noise power density on online learning costs

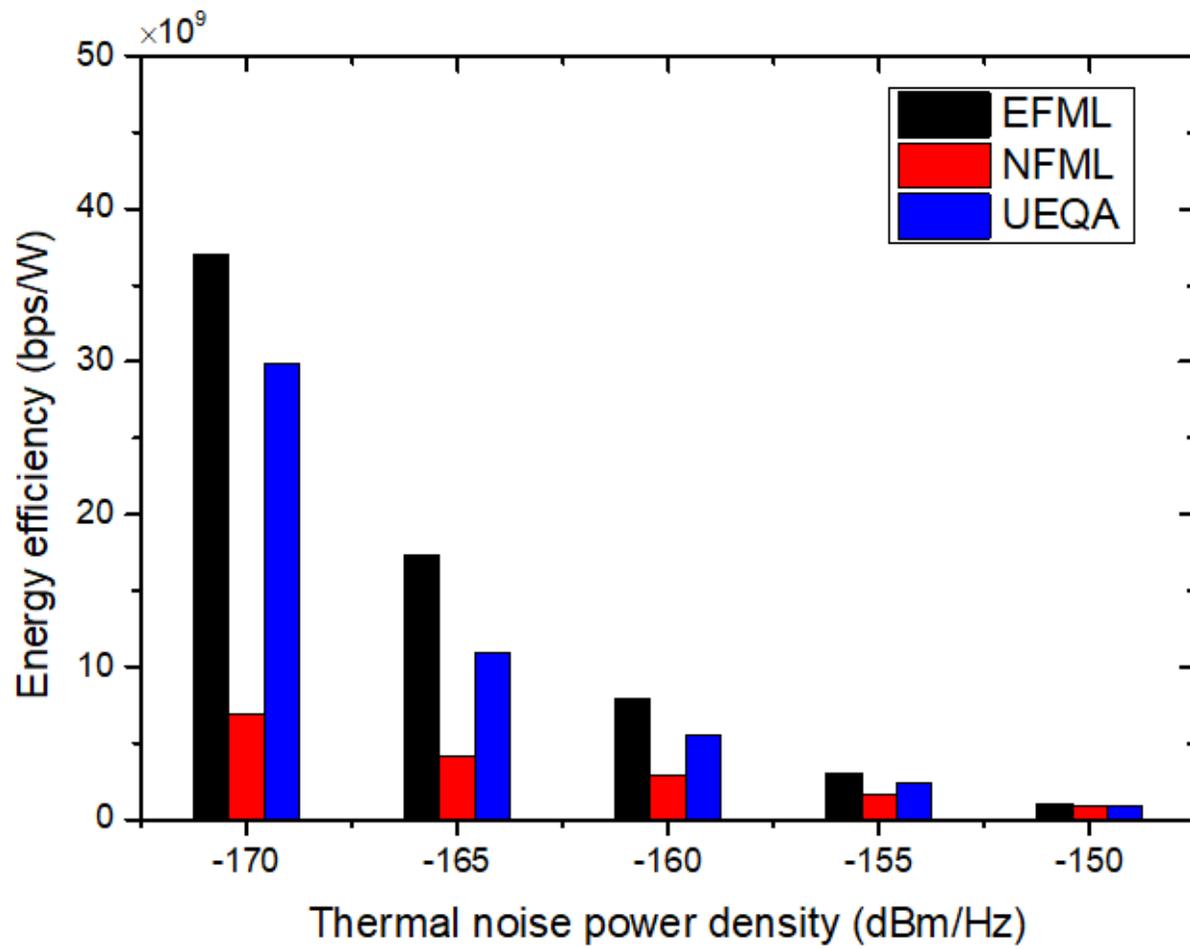


Figure 12

Impact of thermal noise power density on energy efficiency

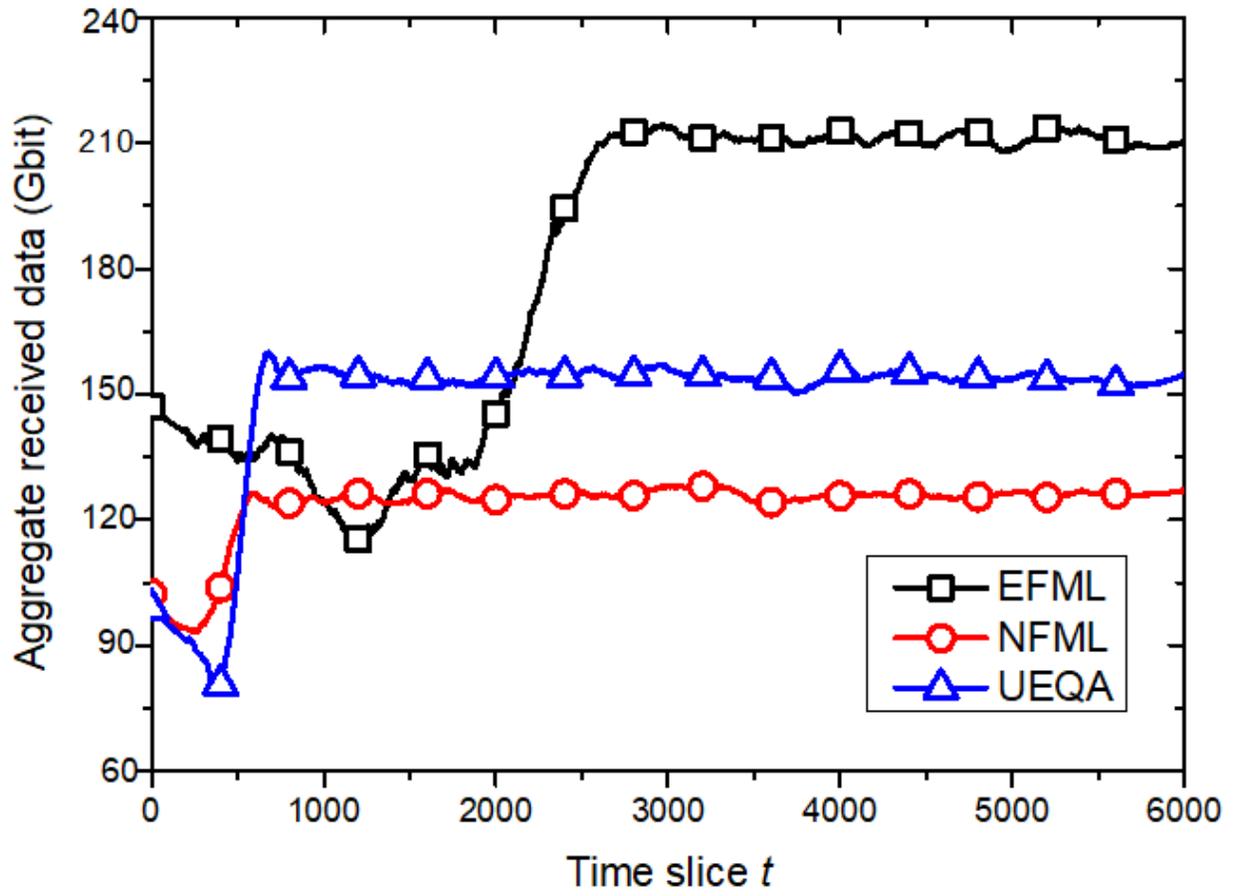


Figure 13

Aggregate received data over the time horizon

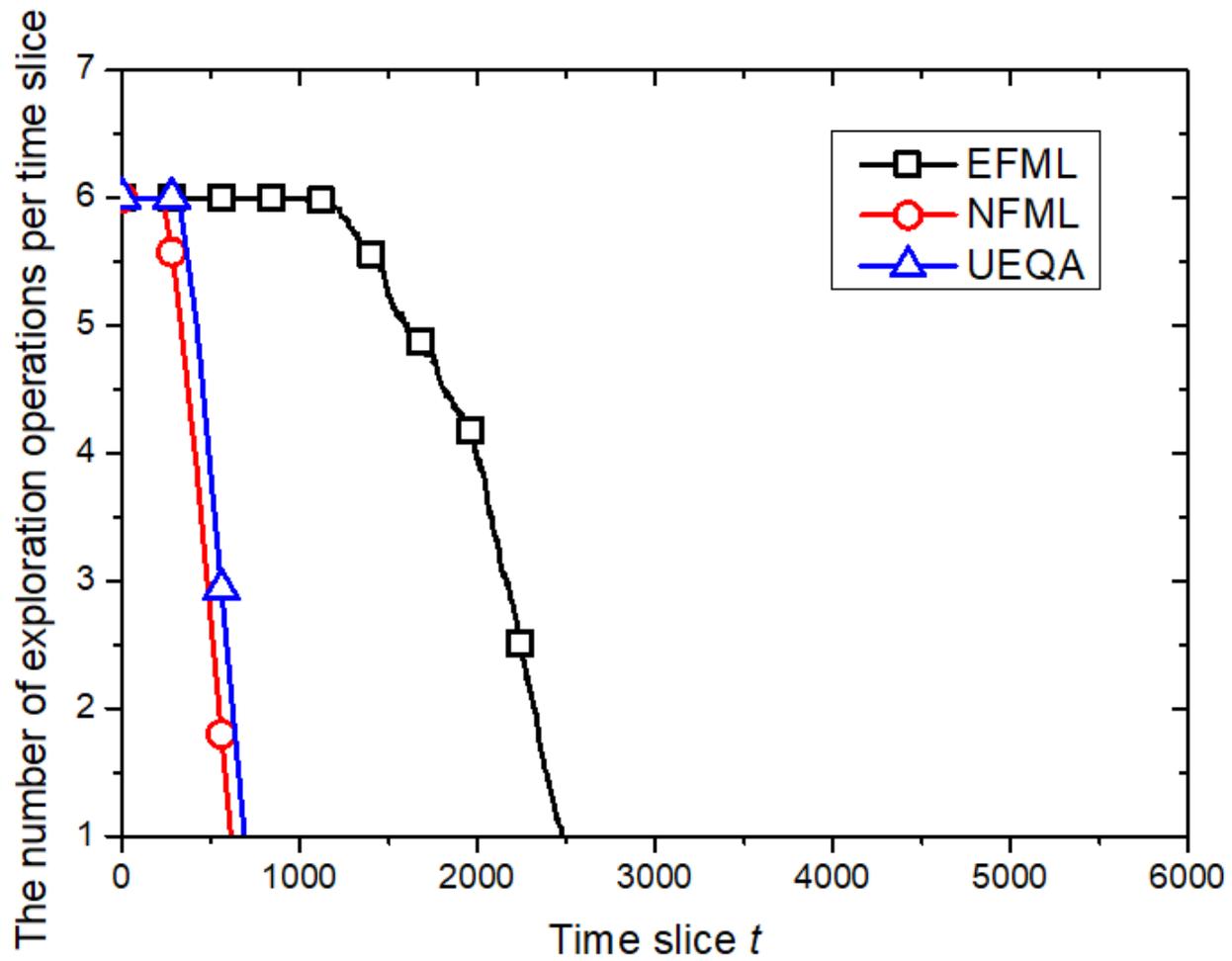


Figure 14

The number of exploration operations per time slice over the time horizon

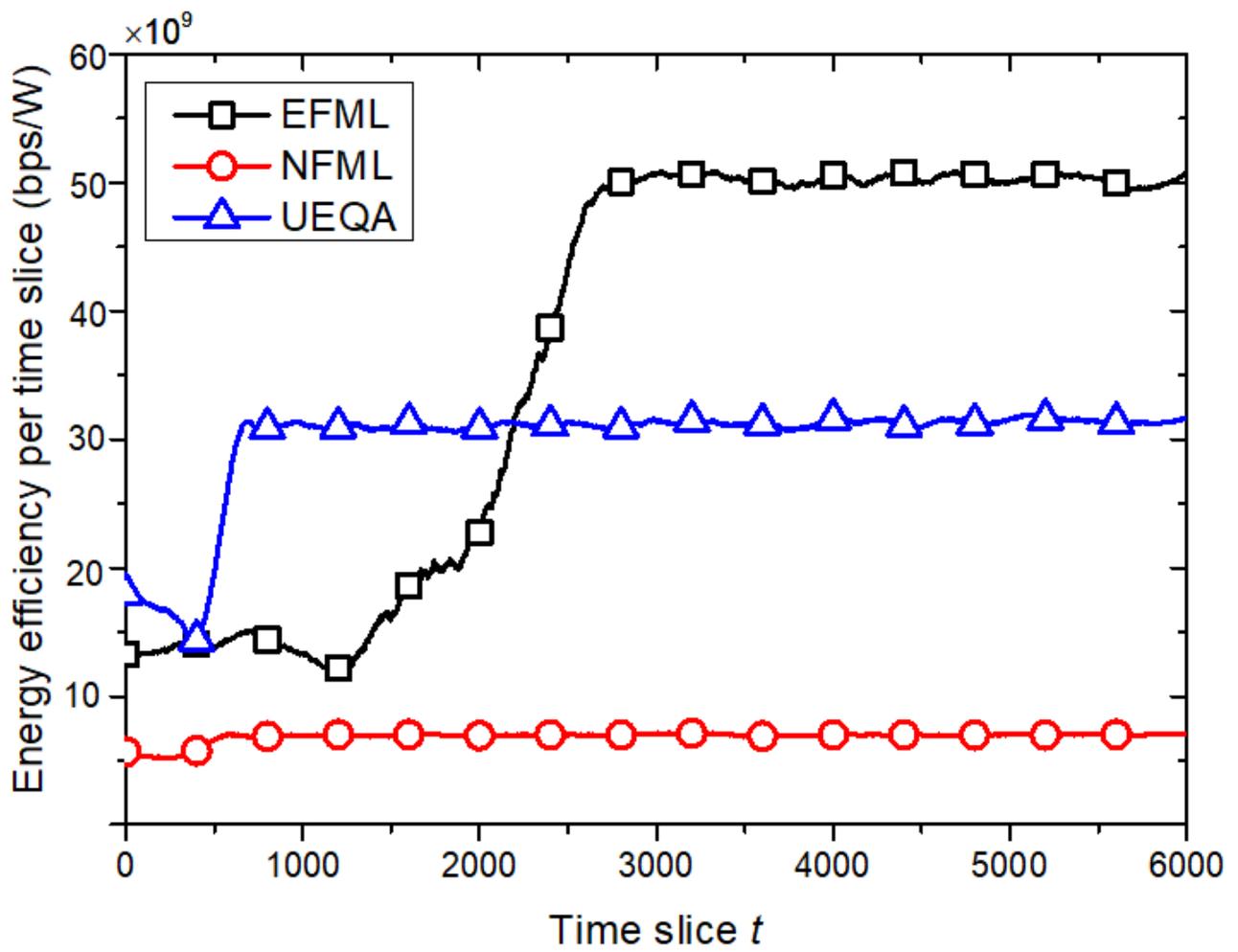


Figure 15

The energy efficiency over the time horizon