

# A Deep Reinforcement Learning Framework to Combat Dynamic Blockage in mmWave V2X Networks

Sheng Chen, Kien Vu, Sheng Zhou, Zhisheng Niu, Mehdi Bennis, and Matti Latva-aho

**Abstract**—Millimeter Wave (mmWave) systems are considered as one of the key technologies in future wireless systems due to the abundant spectrum resources in mmWave band. With the aim of achieving the capacity requirements in vehicular networks, large antenna arrays can be deployed at both the road side units (RSUs) side and the vehicles side. However, dynamic blockage caused by mobile obstacles in mmWave bands may hinder the system reliability. In this work, we study the temporal effects of dynamic blockage in vehicular networks and propose a deep reinforcement learning framework to overcome dynamic blockage. By dynamically adjusting blockage detection parameters and making intelligent handover decisions according to the observed states, system reliability can be significantly improved. Simulation results based on ray-tracing channel data show that the proposed scheme reduces the violation probability by 28.9% over conventional schemes.

## I. INTRODUCTION

Recently, vehicular networks are gaining worldwide attention due to their significant potential in autonomous driving and cooperative vehicles infrastructure systems [1]. In vehicular networks, intelligent vehicles communicate with road side units (RSUs) and other vehicles for message delivery, task offloading, sensor sharing and so forth. To support the diverse requirements of these applications, vehicular networks are supposed to provide ultra-reliable and low latency communications (URLLC) for safety issues, as well as large capacity to deliver massive sensor data. To this end, the evolution of wireless communication technologies in vehicular networks is in urgent need.

Extensive works have been devoted to realizing URLLC in vehicular networks [2]–[5]. While sub-6 GHz bands are widely considered in these works, millimeter wave (mmWave) bands have also been included as potential bands for vehicular networks in 3GPP standards [6]. In mmWave systems, beamforming is applied on the transceivers to compensate

This work is sponsored in part by the Nature Science Foundation of China (No. 61871254, No. 91638204, No. 61861136003), National Key R&D Program of China 2018YFB0105005 and 2018YFB1800800, the Academy of Finland 6Genesis Flagship project under grant 318927 and Intel Collaborative Research Institute for Intelligent and Automated Connected Vehicles. The Nokia Foundation, the UniOGS travel grant and the NVIDIA Corporation are also acknowledged.

S. Chen, S. Zhou and Z. Niu are with Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. Emails: {chen-s16@mails., sheng.zhou@, niuzhs@}tsinghua.edu.cn.

K. Vu was with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland, and now is with Nokia Networks, Finland. Email: kien.vu@nokia.com.

M. Bennis and M. Latva-aho are with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland. Emails: {mehdi.bennis, matti.latva-aho}@oulu.fi.

for the high path loss, which requires the accurate alignment of the beam directions between the transceivers. High mobility in vehicular networks leads to more frequent beam alignment and larger training overhead. Meanwhile, mmWave signals can be easily blocked by obstacles, which hurts the system reliability. In vehicular networks, dynamic blockage caused by pedestrians and vehicles occurs more frequently and is harder to be predicted compared with static blockers like buildings. The work in [7] analyzes the blockage and coverage probability in vehicular networks. The work in [8] studies the temporal effects of mobile blockers in mmWave cellular systems, and formulates the transmission process as an on-off process. Authors in [9] exploit stochastic geometry theory to analyze the impact of mobile blockers in mmWave cellular systems. However, existing work mainly focuses on the scenario of static transceivers or the impact of human blockage. Further study on the dynamic blockage in vehicular networks is still an open problem.

To deal with dynamic blockage, the work in [10] considers using dual-band systems, wherein the sub-6 GHz band can provide services when the mmWave link is blocked. However, supporting two separate bands increases the system complexity and hardware cost. Authors in [11] propose a deep learning based method to predict the blockage and proactively handover the vehicles to nearby RSUs. The method requires data collection and model update in an online manner. The work in [12] proposes to involve multi-connectivity to improve the performance of mmWave cellular networks at the cost of deploying extra resources. Essentially, these works find backup transmission links, which can be non-line-of-sight (NLoS) paths, lower frequency bands and nearby relay nodes. However, choosing appropriate backup links is still an unsolved problem considering the different switching cost and the temporal effects of dynamic blockage.

In this work, we apply deep reinforcement learning (DRL) to jointly optimize beam training and data transmission to improve the reliability of a vehicle-to-infrastructure (V2I) communication system. To cope with blockage, we leverage NLoS paths or nearby RSUs to ensure system reliability, in which the switching decisions are followed by the proposed DRL algorithm. The main contributions are listed as follows:

- We analyze the temporal effects of dynamic blockage in V2I networks and provide analytical expressions of the duration of blockage and non-blockage intervals. This finding reveals the relationship between channel states and physical states and inspires the usage of context information for decision making.
- We propose a deep reinforcement learning framework to deal with dynamic blockage in mmWave V2X networks.

The proposed online framework can detect dynamic blockage during the transmission phase and make appropriate switching decisions to improve system reliability.

- We evaluate the performance of the proposed framework on ray-tracing based channel data, in which simulation results show a 28.9% reduction of violation probability over baseline algorithms.

## II. SYSTEM MODEL

Considering a V2I communication system, a target vehicle is served by one RSU out of the total  $N_{\text{rsu}}$  RSUs. The vehicle needs to download contents through the V2I link continuously, and the goal is to guarantee the reliability of the downlink transmission process. The RSUs are equipped with a uniform linear array (ULA) and have  $M_t$  antennas, while the vehicle is equipped with  $M_r$  antennas forming a ULA.

### A. Channel Model

Considering a geometry stochastic channel model [13], the downlink channel between the RSU and the vehicle can be expressed as:

$$\mathbf{H} = \sqrt{M_t M_r} \sum_{l=1}^L \alpha_l \beta_l \mathbf{a}_r(\theta_l) \mathbf{a}_t^T(\psi_l). \quad (1)$$

Here  $\alpha_l$  denotes the complex impulse response and  $\beta_l$  is a parameter representing whether the propagation path is blocked by other vehicles. We assume that if the path is blocked,  $\beta_l = 0$ , otherwise  $\beta_l = 1$ .  $\theta_l$  and  $\psi_l$  denote the angle of arrival (AoA) and angle of departure (AoD) of the  $l$ -th path, respectively.  $\mathbf{a}$  is the steering vector and for a ULA,

$$[\mathbf{a}(\theta)]_i = \frac{1}{\sqrt{M}} e^{-j \frac{2\pi}{\mu} i d \sin \theta}, (i = 0, 1, \dots, M-1). \quad (2)$$

where  $d$  is the antenna spacing and  $\mu$  is the wavelength.

### B. Beamforming Model

For simplicity, we adopt analog beamforming for the transceivers, i.e., the beamforming weights are given as

$$[\mathbf{w}]_i = \frac{1}{\sqrt{M}} e^{j\phi_i}, i = 0, 1, \dots, M-1 \quad (3)$$

where  $\phi_i$  denotes the phase of the  $i$ -th antenna. The beamforming codewords are from a discrete Fourier transform (DFT) matrix, i.e.  $\phi_i = i\theta_w$ , where  $\theta_w \in \{(-1 + \frac{2n-1}{M})\pi, n = 1, 2, \dots, M\}$ . The beamforming gain can then be expressed as:

$$G = \|\mathbf{w}_R^T \mathbf{H} \mathbf{w}_T\|^2, \quad (4)$$

where  $\mathbf{w}_R$  and  $\mathbf{w}_T$  denote the beamforming weights of the receiver and the transmitter, respectively.

We adopt a two-stage beam training (BT) protocol, including a sector level sweeping (SLS) phase and a beam refinement (BR) phase<sup>1</sup>. In the sector-level training, vehicles can receive the pilot signals from all RSUs and measure the corresponding SNRs. These measurements are reported

<sup>1</sup>In this work, we assume that the resource allocated for each user is divided by bandwidth instead of time slots, therefore the problem is simplified to single user case. The multiple user case is left for future work.

to the active RSU for reducing the search space in the next stage as well as making handover decisions. In the dedicated BR phase, narrower beams are trained between the vehicle and the active RSU. The length of each BR phase can be adjusted to optimize the system performance. As for the beam training vectors, a naive method is adopted, by which the neighborhoods of the current beam direction are first trained, and the remaining beam directions are uniformly sampled based on the training overhead. After the BT phase, the active RSU chooses the beam directions for data transmission (DT) and determines two lists recording backup beam pairs and backup RSUs. Specifically, the first list records potential beam pairs corresponding to the NLoS paths, while the second list records RSUs with good channel conditions. The service rate can be expressed as:

$$r(t) = \mathbf{1}(t \in \mathcal{T}_{\text{DT}}) W \log(1 + \frac{P G_t}{W N_0}). \quad (5)$$

Hereinafter,  $\mathcal{T}_{\text{DT}}$  denotes the set of all time slots in the DT phase.  $W$ ,  $P$ ,  $G_t$  and  $N_0$  denote the bandwidth, the transmission power, beamforming gain at time  $t$  and noise power, respectively. If a dynamic blockage happens in the DT phase, the transceivers will try backup beam pairs first and then try the backup RSUs.

### C. Queueing Model

Consider a stochastic queuing network that operates in a slotted time  $t \in \{0, 1, 2, \dots\}$ . We assume that there are  $N_{\text{rsu}}$  queues in the network, and the queuing vector at an active RSU follows the following evolution:

$$Q_{U(t)}(t+1) = \max\{Q_{U(t)}(t) - r_{U(t)}(t), 0\} + \lambda(t), \quad (6)$$

where  $U(t)$  indicates the active RSU in the slot  $t$  and  $\lambda(t)$  is the data arrival of the target vehicle during slot  $t$ , which is i.i.d. over time with a mean value  $\bar{\lambda}$  [14].

During the handover process, the buffered data in the source RSU will be transferred to the target RSU. In other words, if the vehicle is switched to RSU  $U'$  from RSU  $U$  starting from time slot  $t$  to time slot  $t + t_{\text{HO}}$ , the queue length of the target RSU can be expressed as:

$$Q_{U'}(t + t_{\text{HO}}) = Q_U(t) + \sum_{\tau=0}^{t_{\text{HO}}-1} \lambda(t + \tau). \quad (7)$$

## III. DYNAMIC BLOCKAGE IN VEHICULAR NETWORKS

In this section, we study the temporal effects of dynamic blockage in mmWave-enabled vehicular networks to gain some insights on blockage statistics.

As shown in Fig. 1, a RSU is located at  $[0, y_{\text{rsu}}]$  with a height of  $H_{\text{rsu}}$ . Assume that there are totally  $N_{\text{la}}$  lanes and the width of each lane is  $W_{\text{la}}$ . Target vehicle moves on the center of the  $m$ -th lane, with a size of  $[L_t, W_t, H_t]$  and a uniform speed of  $v_t$ . Assuming that a vehicle on the  $n$ -th lane with a size of  $[L_b, W_b, H_b]$  and a uniform speed of  $v_b$  is blocking the line-of-sight (LoS) path of the target vehicle. Using the similar methods in [7], blockage conditions can be obtained.

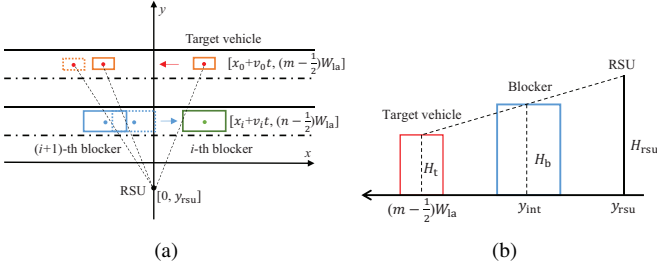


Fig. 1. (a) Top view of the scenario. (b) Side view of the scenario.

**Lemma 1.** A vehicle on the  $n$ -th lane that blocks the LoS path of target vehicle should satisfy the following two constraints:

$$\begin{cases} n < m, \\ \frac{H_{rsu} - H_b}{H_{rsu} - H_t} \leq \frac{(2n-1)W_{la} + W_b - 2y_{rsu}}{(2m-1)W_{la} - 2y_{rsu}}. \end{cases} \quad (8)$$

*Proof.* The first constraint results from the fact that the LoS can only be blocked by the vehicles on the closer lanes to the RSU. As for the latter, according to geometry relationship in Fig. 1(b),

$$\frac{H_{rsu} - H_b}{H_{rsu} - H_t} = \frac{y_{int} - y_{rsu}}{(m - 1/2)W_{la} - y_{rsu}}. \quad (9)$$

Herein  $y_{int}$  denotes the coordinate of the intersections in the  $y$  axis, and  $y_{int} \leq (n - 1/2)W_{la} + 1/2W_b$ . Substitute  $y_{int}$  into (9) and obtain the second constraint.  $\square$

**Lemma 2.** The duration for a single blocker can be approximated as:

$$t_b = \frac{L_b}{\|v_b - kv_t\|}, \quad (10)$$

where  $k = \max \left\{ \frac{H_{rsu} - H_b}{H_{rsu} - H_t}, \frac{(2n-1)W_{la} - W_b}{(2m-1)W_{la}} \right\}$ . The interval between two successive blockers on the same lane can be approximated as:

$$t_{nb} = \frac{L_{dis} - L_b}{\|v_b - kv_t\|}, \quad (11)$$

where  $L_{dis}$  is the distance between two blockers.

*Proof.* The proof is given in Appendix A.  $\square$

According to (10) and (11), the blockage duration is related to the sizes, speeds and the lanes of the vehicles, therefore this context information is useful for blockage detection and making corresponding decisions. Combining the blockage states of all lanes, the blockage state of the target vehicle can be obtained. We can then formulate the transmission process of the target vehicle as a renewal process with on phase and off phase. During the on phase when the LoS path exists, the channel condition is good and a high SNR can be achieved. During the off phase, the LoS path is blocked, and the corresponding SNR is relatively low.

#### IV. DEEP REINFORCEMENT LEARNING FRAMEWORK TO COMBAT DYNAMIC BLOCKAGE

In this section, a deep reinforcement learning framework is proposed to tackle dynamic blockage in vehicular networks.

##### A. Problem Formulation

According to *Little's law*, the average latency in  $t$ -th frame is proportional to the average queue length, which is expressed as  $\frac{Q_U(t)}{\lambda}$ . However, this average metric cannot guarantee low latency and high reliability, which are stringently required in 5G networks and beyond [15]. Moreover, latency and reliability are two coupled factors, and thus, in this work we focus on the reliability of vehicular networks by taking queue statistics into account, as a proxy for latency constraints. In particular, we impose the probabilistic queuing constraint to measure the reliability by maximizing the probability that the queue length is not greater than a pre-defined threshold. The studied problem can be formulated as follows:

$$\max_{\mathcal{T}_{DT}, \mathbf{w}_R, \mathbf{w}_T, U(t)} \frac{1}{T} \sum_{t=1}^T \mathbb{1}(Q_U(t) \leq Q^{th}) \quad (12)$$

where  $Q^{th}$  denotes the maximum queue length.

##### B. Reinforcement Learning Framework

The active RSU makes decisions on allocating beam training resources and switching to backup links, which affects the achievable rate and queue length. Since blockage can happen at any time slot in a frame, the optimal decision should be made at each time slot. However, per-slot decision is hard to implement, therefore in this work the decisions are made at the start of each frame.

During the  $k$ -th frame, the state can be expressed as  $s_k = (Q_k, r_{SNR,k}, (l_k, v_k)) \in \mathcal{S} = \mathcal{Q} \times \mathcal{R}^{L_o} \times \mathcal{C}$ , including the queue length  $Q_k$ , SNR feedback from previous time slots  $r_{SNR,k}$ , location  $l_k$  and speed  $v_k$ .  $\mathcal{Q}$ ,  $\mathcal{R}$  and  $\mathcal{C}$  denote the states sets of queue length, SNR feedback and context information, respectively, while  $L_o$  is the length of SNR feedback records. Provided the system states, the RSU decides the beam training overhead and blockage detection parameters, which are described as  $a_k = (n_{bt,k}, (T_{d,k}, SNR_{th,k}, \eta_k)) \in \mathcal{A} = \mathcal{N}_{bt} \times \mathcal{P}_{bd}$ . Here,  $n_{bt}$  denotes the overhead in the BR phase, and the corresponding action set  $\mathcal{N}_{bt}$  contains the integers smaller than the beam codebook size.  $\mathcal{P}_{bd}$  is the set of blockage detection parameters, including the detection time length  $T_d$ , the SNR threshold  $SNR_{th}$  and frequency threshold  $\eta$ . A blockage is assumed to be detected when the following formula holds:

$$\frac{1}{T_d} \sum_{t=1}^{T_d} \mathbb{1}(SNR(t) < SNR_{th}) \geq \eta. \quad (13)$$

For instance, given  $T_d = 1$  and  $\eta = 1$ , a blockage is assumed to happen if current SNR is smaller than  $SNR_{th}$ , and the transceivers will try backup beam pairs. If however (13) is still satisfied, representing the possible blockage of the NLoS path, the vehicle will be switched to the backup RSU. By adjusting these parameters according to the system states, the switching behaviours can be controlled, which may improve the system performance.

The cumulative reward is the sum of the per time slot reward in the whole frame and can be expressed as

$$R_k = \sum_{t=\tau_{k-1}}^{\tau_k-1} (\mathbb{1}(Q(t) \leq Q_{th}) - \gamma Q(t)), \quad (14)$$

where  $\tau_k$  denotes the start of the  $k$ -th frame and  $\gamma$  is a trade-off factor. The first part indicates the reliability of the current slot, while the second part reflects the latency performance according to *Little's law*.

### C. Deep Reinforcement Learning Implementation

Faced with the large state space and continuous actions, we adopt a model-free RL algorithm, namely Deep Deterministic Policy Gradient (DDPG) [16], as shown in Alg. 1. Two networks are included in the method, namely critic network  $Q(s, a|\theta^Q)$  and actor network  $\mu(s|\theta^\mu)$ , where  $s$  and  $a$  denote the states and actions respectively. The actor network can generate deterministic actions based on the input states, while the critic network is a neural network (NN) function approximator.

In this work, each episode is defined as the sojourn of a target vehicle in the area of interest. At the start of the  $k$ -th frame, the actor network chooses the action  $a_k = \mu(s_k|\theta^\mu) + \mathcal{N}_k$  based on the current state  $s_k$  and exploration noise  $\mathcal{N}_k$ . After executing the action and collecting the reward, transitions including state  $s_k$ , action  $a_k$ , reward  $r_k$  and the corresponding next state  $s_{k+1}$  will be stored in an experience buffer. A minibatch of the buffered record transitions will then be used for training. The critic network is updated to minimize the loss:

$$L_c = \frac{1}{N} \sum_i (r_i + \gamma_d Q(s_{i+1}, \mu(s_{i+1})) - Q(s_i, a_i))^2, \quad (15)$$

while the actor network is updated based on the sampled policy gradient:

$$\nabla_{\theta^\mu} J = \frac{1}{N} \sum_i \nabla_a Q(s, a)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s)|_{s=s_i}. \quad (16)$$

Here  $N$  is the batch size and  $\gamma_d$  is the discounting factor.

---

#### Algorithm 1: DDPG for Dynamic Blockage

---

Initialize actor network  $\mu$  and critic network  $Q$  with parameters  $\theta^\mu$  and  $\theta^Q$  randomly;  
**for** episodes = 1, 2,  $\dots$ ,  $N_{\text{epi}}$  **do**  
  Initialize the random exploration noise  $\mathcal{N}$ ;  
  Receive the initial observation  $s_0$ ;  
  **for**  $k = 0, 1, \dots, K - 1$  **do**  
    Select action  $a_k$  based on the output of actor network and the exploration noise;  
    Set beam training and data transmission parameters according to  $a_k$ ;  
    Observe the reward  $r_k$  and new state  $s_{k+1}$  after data transmission in the  $k$ -th frame;  
    Store transition  $(s_k, a_k, r_k, s_{k+1})$  in the experience buffer;  
    Sample a minibatch from the experience buffer;  
    Update critic network based on (15);  
    Update actor network based on (16);

---

## V. SIMULATION RESULTS

We evaluate the performance of the proposed algorithm on ray-tracing based channel data. In the simulation settings, a

TABLE I  
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
$[N_{\text{la}}, W_{\text{la}}]$	$6 \times 3.5$ m	$W$	500 MHz
$H_{\text{rsu}}$	5 m	$N_0$	-173 dBm/Hz
$[L_{\text{t}}, W_{\text{t}}, H_{\text{t}}]$	$5 \times 2 \times 1.6$ m	$P$	23 dBm
$[L_{\text{b}}, W_{\text{b}}, H_{\text{b}}]$	$13 \times 2.6 \times 3$ m	$M_{\text{t}}$	32
$v_{\text{t}}$	$[50, 60]$ km/h	$M_{\text{r}}$	1
$v_{\text{b}}$	$[50, 60]$ km/h	$\lambda$	100 Mbps

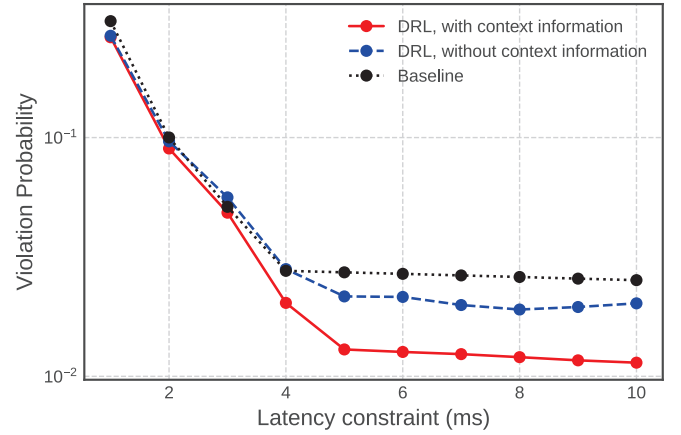


Fig. 2. Violation probability versus latency constraint.

vehicle is going straight on an urban street with surrounding buildings under the service region of 4 RSUs. 100 records are generated for training, leaving another 50 records for test. Simulation parameters are listed in Table I. The channel parameters of each ray are generated from ray-tracing software Wireless inSite [17], while the blockage state of each ray is simulated based on (10) and (11). The carrier frequency is set to be 28 GHz. Traffic density is characterized by the distance between two blockers on the same lane, which follows an exponential distribution with a minimal distance of 2 m. Handover time  $t_{\text{HO}}$  is uniformly distributed between 15 ms and 25 ms, and the queue length threshold  $Q_{\text{th}}$  is set to be  $\lambda T_{\text{th}}$ , where  $T_{\text{th}}$  denotes the latency constraint. The scaling factor  $\gamma$  in the reward function (14) is  $1/Q_{\text{th}}$ . The context information includes the estimated location and speed of the target vehicle, with an estimation error of 1 m and 1 m/s, respectively.

We implement the proposed reinforcement learning framework on Tensorflow [18]. Both the actor and critic network have two dense layers with 128 and 64 hidden nodes each layer. The learning rates for the actor network and critic network are  $10^{-4}$  and  $10^{-3}$ , respectively. The batch size is 256 and the discounting factor is 0.95.

Fig. 2 shows the trade-off between reliability and latency. We compare our scheme with a baseline algorithm wherein beam training overhead and blockage detection parameters are fixed. It can be observed that the proposed DRL-based method outperforms the baseline scheme, owing to dynamic beam training overhead and flexible handover decisions. Besides, the performance can be further improved leveraging context information.

Fig. 3 shows the violation probability under different settings of traffic density, represented by the average distance

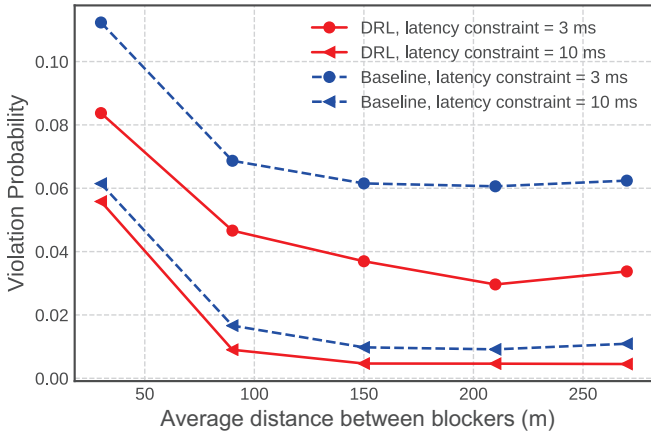


Fig. 3. Violation probability versus average distance between two blockers.

between two blockers. It can be observed that the violation probability increases with the traffic density, while the performance gain of the proposed scheme decreases with the latency constraint. For example, the proposed scheme can reduce violation probability by 28.9% when the average distance between blockers is 30 m and latency constraint is 3 ms.

## VI. CONCLUSION

This work studied the problem of dynamic blockage in mmWave-enabled V2I networks. By leveraging the DRL framework, a novel handover mechanism was proposed, which taking channel states, physical states, and context information into account. Based on ray tracing channel data, numerical results show significant improvements in system reliability.

There are many future directions to be explored. The extension to the multi-user case is challenging, since multiple access schemes should be included. Another interesting topic is using intelligent vehicles as relay nodes, calling for the joint optimization of V2I and V2V transmission.

## APPENDIX A

As shown in Fig. 1(a), the RSU is located at  $[0, y_{rsu}]$ . Assumes that all vehicles move on a straight line at a uniform speed. At time  $t$ , the coordinate of target vehicle can be expressed as  $[x_0 + v_0 t, (m-1/2)W_{la}]$ , and the coordinate of the  $i$ -th blocker on the  $n$ -th lane is  $[x_i + v_i t, (n-1/2)W_{la}]$ . Assuming that the blockage interval of the  $i$ -th blocker is from  $t_i$  to  $\tilde{t}_i$ , then the following equations hold:

$$\frac{x_i + v_i t_i \pm L_b/2}{x_0 + v_0 t_i} = k, \quad (17)$$

$$\frac{x_i + v_i \tilde{t}_i \mp L_b/2}{x_0 + v_0 \tilde{t}_i} = \tilde{k}, \quad (18)$$

where  $k = \frac{y_{int} - y_{rsu}}{(m-1/2)W_{la} - y_{rsu}}$ . According to the geometric relationship in Fig. 1(b),  $k$  can also be expressed as  $\frac{H_{rsu} - H_b}{H_{rsu} - H_i}$ , unless  $y_{int} < (m-1/2)W_{la} - 1/2W_b$ . Therefore,  $k = \max \left\{ \frac{H_{rsu} - H_b}{H_{rsu} - H_i}, \frac{(2n-1)W_{la} - W_b}{(2m-1)W_{la}} \right\}$ .

$\tilde{k}$  has the similar expression and can be approximated as  $k$ , i.e.  $\tilde{k} \approx k$ . Combining (17) and (18), we have  $\tilde{t}_i - t_i =$

$\frac{L_b}{\|v_i - kv_0\|}$ , which is the duration of a single blocker. Replacing  $i$  with  $i+1$  in (17), we obtain  $\frac{x_{i+1} - x_i + v_{i+1}t_{i+1} - v_i t_i}{v_0(\tilde{t}_i - t_i)} = k$ . Assuming that  $v_{i+1} = v_i$ , we obtain  $t_{i+1} - t_i = \frac{x_{i+1} - x_i}{\|v_i - kv_0\|}$ . Subtracting the blockage time, the duration between two successive blockers is obtained.

## REFERENCES

- [1] M. Gerla and L. Kleinrock, "Vehicular networks and the future of the mobile internet," *Computer Networks*, vol. 55, no. 2, pp. 457–469, 2011.
- [2] M. K. Abdel-Aziz, S. Samarakoon, C. Liu, M. Bennis, and W. Saad, "Optimized Age of Information Tail for Ultra-Reliable Low-Latency Communications in Vehicular Networks," *IEEE Transactions on Communications*, vol. 68, 2020. Accepted.
- [3] M. K. Abdel-Aziz, S. Samarakoon, M. Bennis, and W. Saad, "Ultra-Reliable and Low-Latency Vehicular Communication: An Active Learning Approach," *IEEE Communications Letters*, pp. 1–1, Nov. 2019.
- [4] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Federated Learning for Ultra-Reliable Low-Latency V2V Communications," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Dec. 2018.
- [5] M. Bennis, M. Debbah, and V. Poor, "Ultra-reliable and low-latency wireless communication: Tail, Risk and Scale," in *Proceedings of the IEEE*, vol. 106, pp. 1834–1853, Oct. 2018.
- [6] *3GPP TR 37.885: Study on evaluation methodology of new Vehicle-to-Everything (V2X) use cases for LTE and NR*, Jun. 2019.
- [7] Y. Wang, K. Venugopal, A. F. Molisch, and R. W. Heath, "Blockage and Coverage Analysis with MmWave Cross Street BSs Near Urban Intersections," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017.
- [8] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy, "On the Temporal Effects of Mobile Blockers in Urban Millimeter-Wave Cellular Scenarios," *IEEE Transactions on Vehicular Technology*, vol. 66, pp. 10124–10138, Nov. 2017.
- [9] I. K. Jain, R. Kumar, and S. S. Panwar, "The Impact of Mobile Blockers on Millimeter Wave Cellular Systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, pp. 854–868, Apr. 2019.
- [10] M. Hashemi, C. E. Koksal, and N. B. Shroff, "Out-of-Band Millimeter Wave Beamforming and Communications to Achieve Low Latency and High Energy Efficiency in 5G Systems," *IEEE Transactions on Communications*, vol. 66, pp. 875–888, Feb. 2018.
- [11] A. Alkhateeb, I. Beltagy, and S. Alex, "Machine learning for reliable mmwave systems: Blockage prediction and proactive handoff," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1055–1059, Nov. 2018.
- [12] V. Petrov, D. Solomitckii, A. Samuylov, M. A. Lema, M. Gapeyenko, D. Moltchanov, S. Andreev, V. Naumov, K. Samouylov, M. Dohler, and Y. Koucheryavy, "Dynamic Multi-Connectivity Performance in Ultra-Dense Urban mmWave Deployments," *IEEE Journal on Selected Areas in Communications*, vol. 35, pp. 2038–2055, Sep. 2017.
- [13] A. Molisch, "A Generic Model for MIMO Wireless Propagation Channels in Macro- and Microcells," *IEEE Trans. Signal Process.*, vol. 52, pp. 61–71, Jan. 2004.
- [14] T. K. Vu, M. Bennis, M. Debbah, and M. Latva-Aho, "Joint Path Selection and Rate Allocation Framework for 5G Self-Backhauled mm-wave Networks," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 2431–2445, Apr. 2019.
- [15] T. K. Vu, C. Liu, M. Bennis, M. Debbah, M. Latva-aho, and C. S. Hong, "Ultra-reliable and low latency communication in mmwave-enabled massive mimo networks," *IEEE Communications Letters*, vol. 21, pp. 2041–2044, Sep. 2017.
- [16] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [17] Remcom, "Wireless InSite." <http://www.remcom.com/wireless-insite>.
- [18] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., "Tensorflow: a system for large-scale machine learning," in *OSDI*, vol. 16, pp. 265–283, Nov. 2016.