

# Blockage-aware Multi-armed Bandits for Proactive HO in mmWave Vehicular Networks

1<sup>st</sup> Weiqi Chi and 2<sup>nd</sup> Manabu Tsukada\*

*The Graduate School of Information and Science Technology*

*The University of Tokyo, Tokyo, Japan*

email address: weiqichi@g.ecc.u-tokyo.jp, mtsukada@g.ecc.u-tokyo.ac.jp

**Abstract**—Connected Autonomous Vehicles (CAVs) require ultra-reliable, low-latency communication for safety-critical applications, making millimeter-wave (mmWave) technology essential for achieving high data rates. However, mmWave signals suffer from severe blockage sensitivity, particularly in dynamic vehicular environments where buildings, trees, and other vehicles can cause sudden communication failures. Traditional reactive handover (HO) schemes prove inadequate for real-time CAV applications due to inherent delays and service interruptions. This paper proposes a novel blockage-aware proactive handover mechanism based on Contextual Multi-Armed Bandit (CMAB) theory for mmWave vehicular networks. The proposed approach leverages Cooperative Awareness Messages (CAM) to collect comprehensive vehicle dynamics and environmental information, enabling predictive blockage analysis through trajectory forecasting. Unlike existing reactive schemes that respond to signal degradation after blockages occur, our method anticipates future blockage events by considering vehicle kinematics and the proximity of large blocking objects. The system models each base station as an individual arm in the CMAB framework, utilizing Upper Confidence Bound (UCB) algorithm to balance exploration and exploitation. Simulation results demonstrate improved communication reliability and reduced handover frequency compared to conventional reactive approaches, making it suitable for safety-critical autonomous driving applications.

**Index Terms**—Vehicle to Everything communication, mmWave, Multi-armed bandits, Blockage prediction, Handover management, Online Learning

## I. INTRODUCTION

Autonomous Driving (AD) is a crucial component of Intelligent Transportation Systems (ITS), which has gained intensive research attention due to its potential to address fundamental road challenges and provide comprehensive solutions for critical issues, including traffic safety enhancement, optimal trajectory planning, fuel consumption reduction, and overall transportation efficiency improvement. Connected Autonomous Vehicles (CAVs) utilize Vehicle-to-Everything (V2X) communications to facilitate seamless wireless information exchange, enabling coordinated autonomous functions across the entire transportation ecosystem. The European Telecommunications Standards Institute (ETSI) has standardized the Cooperative Awareness Message (CAM) [1], which serves as a fundamental protocol enabling CAVs to systematically share comprehensive status information about both themselves and their surrounding environment.

The CAM protocol contains detailed attributes of the originating Intelligent Transportation System Station (ITS-S),

encompassing vehicle dynamics, positioning data, and environmental awareness parameters. This standardized message framework allows CAVs to maintain continuous situational awareness by broadcasting and receiving critical information, including vehicle speed, heading, acceleration, dimensions, etc. The cooperative nature of this communication paradigm enables vehicles to make informed decisions based not only on their local sensor data but also on the collective intelligence gathered from neighboring vehicles and infrastructure elements.

The diverse and computationally intensive tasks that CAVs are expected to perform impose exceptionally high demands on both throughput and latency performance of the underlying communication infrastructure. Fifth-generation (5G) millimeter-wave (mmWave) vehicular networks have emerged as a promising solution, designed to provide ultra-high data transmission rates, enhanced spectral efficiency, and ultra-reliability communication capabilities [2]. The mmWave spectrum offers unprecedented bandwidth availability, which is essential for supporting the massive data flows generated by CAVs equipped with multiple high-resolution sensors.

However, the practical deployment of mmWave technology in vehicular environments faces significant technical challenges. Due to inherent characteristics of high-frequency electromagnetic propagation, mmWave signals suffer from substantial penetration loss and exhibit extreme sensitivity to physical blockages, such as buildings, trees, and other vehicles. This results in significantly reduced coverage areas and compromised connection stability [3]. Furthermore, as mmWave communications are highly directional in nature, maintaining Line-of-Sight (LoS) connectivity between the transmitter (Tx) and receiver (Rx) becomes absolutely critical to ensure satisfactory communication quality and reliable data transmission [4]. Unfortunately, guaranteeing consistent LoS connectivity presents challenges in dynamic, high-mobility vehicular scenarios.

To illustrate the blockage impact, Fig. 1 shows the CDF of both LOS and NLOS communication rates. The simulation was conducted using a realistic urban scenario based on the actual map of Hachioji, Japan, with vehicle movements simulated for 500 seconds using Simulation of Urban Mobility (SUMO) [5] and 12 uniformly distributed mmWave Base Stations (BSs). The CDF comparison reveals the significant impact of blockage on vehicular communication performance,

where the blue curve (LOS) achieves notably higher communication rates than the red curve (NLOS). The substantial separation between these curves demonstrates how blockage from surrounding buildings and other vehicles severely degrades communication reliability in realistic urban vehicular networks.

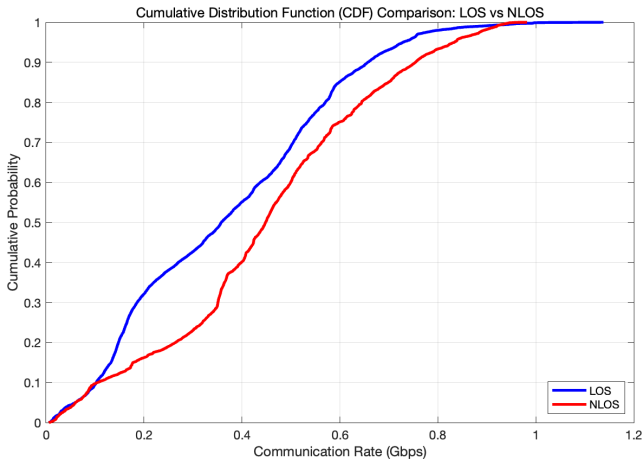


Fig. 1. Blockage impact analysis in mmWave vehicular system

## II. RELATED WORK

The application of Machine Learning (ML) in vehicular communications is particularly compelling due to the inherently dynamic and unpredictable nature of mobile environments, where traditional rule-based algorithms often fail to adapt to rapidly changing conditions.

In [6], authors leverage wireless metrics like signal strength and channel status as contextual information to make real-time HO decisions using reinforcement learning approaches. This work tackles the problem of improving HO success rates, reducing latency, and minimizing HO failures by incorporating real-time channel conditions into the decision-making process. The proposed framework demonstrates improved performance in maintaining connectivity stability compared to traditional threshold-based HO mechanisms, particularly in scenarios with rapid channel variations.

Reference [7] focuses on proactive HO strategies that predict the next base station choice through machine learning techniques, with reward functions designed to maximize predicted signal-to-noise ratio (SNR) stability while minimizing unnecessary HOs and service interruptions. The approach utilizes historical channel quality indicators and mobility patterns to anticipate optimal HO timing. However, the HO decision-making process requires extensive training effort and relies heavily on observed SNR histories, which may not adequately capture the complex dynamics of mmWave propagation characteristics in vehicular environments.

Authors in [8] developed a Deep Double Q-Network (DDQN) agent specifically designed for 5G mmWave Dual-Active Protocol Stack (DAPS) HOs, learning optimal timing to trigger HOs and avoid sudden radio link failures under volatile mmWave channel conditions. Their approach addresses the unique challenges of mmWave communications, including

beam misalignment and rapid signal fluctuations. However, this method represents a reactive HO scheme that primarily responds to prior signal degradation patterns rather than proactively anticipating future network conditions. Importantly, it does not explicitly predict future obstacles or environmental changes, such as motion tracking of vehicles or prediction of moving blockages, which are critical factors in dynamic vehicular network scenarios where proactive decision-making could significantly improve communication reliability.

Among the various ML paradigms, the Contextual Multi-Armed Bandit (CMAB) theory has gained considerable attention as a sophisticated subclass of Reinforcement Learning (RL) problems that specifically addresses sequential decision-making challenges under conditions of uncertainty and incomplete information. CMAB frameworks are particularly well-suited for HO management applications because they can effectively balance the exploration of new HO strategies with the exploitation of proven successful approaches, while continuously adapting to evolving network conditions. In [9], the authors employed MAB algorithms to address joint user association and handover problems in mmWave-enabled heterogeneous networks.

However, despite the technical sophistication and demonstrated effectiveness of this approach, it fundamentally remains a reactive HO scheme that operates based on prior signal degradation patterns and historical network performance data. The future obstacle prediction represents a significant gap in current HO management approaches. The ability to predict the future positions and trajectories of potential blocking objects would enable much more sophisticated proactive HO strategies that could anticipate communication challenges before they occur, rather than simply reacting to degraded signal conditions after blockages have already impacted communication quality.

## III. RESEARCH MOTIVATION AND CONTRIBUTION

To address the challenge of stabilizing mmWave connectivity under dynamic blockage conditions, conventional reactive handover schemes that rely on passive signal quality monitoring prove inadequate for real-time CAV applications due to inherent delays during the time-to-trigger period. Moreover, the high-mobility nature of vehicular networks exacerbates handover frequency, leading to service disconnections and communication gaps that create potential safety hazards. Therefore, developing proactive handover mechanisms becomes essential for maintaining continuous, reliable connectivity in CAV deployments.

In this work, we focus on the future blockage-aware proactive HO issue. In particular, our original contribution can be concluded as:

- This work includes an original attempt to design mmWave HO for V2X communication service under the high-mobility scenario that aims to avoid sudden outage of the communication signal due to vehicle blockages.
- The proposed proactive ML-assisted HO scheme is based on Contextual Multi-Armed Bandit (CMAB) that not only

considers the vehicular kinematics, but also involves the predictive blockage impact concerning the presence of big vehicles.

- The comparison of the overall performance among the reactive HO schemes, the proactive schemes that are purely based on current observation, and the proposed method that considers the future blockage rate will be made.

#### IV. BLOCKAGE-AWARE MAB FOR PROACTIVE HO

##### A. System Overview

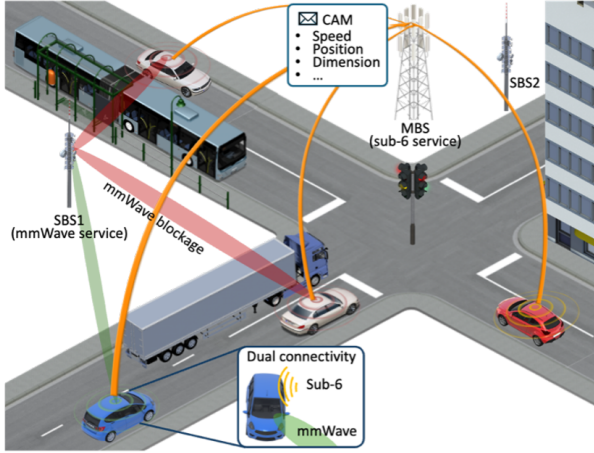


Fig. 2. Research Scenario

The research scenario is shown in Fig. 2. We examine a heterogeneous mmWave vehicular network with regions where a sub-6 GHz macro-Base Station (MBS) overlays multiple mmWave sub-BSSs (SBSs). The MBS functions as both an access point and a central processor for learning tasks, while SBSs enable advanced V2X applications. We assume varying traffic densities and large vehicles. Entering vehicles update their status via Cooperative Awareness Message (CAM) through V2I communication, providing input for mmWave blockage prediction and mitigation strategies.

This scenario is modeled as a CMAB problem with each SBS as an individual arm. The MBS, serving as a centralized learner, executes the following process (Fig. 3):

- **Context collection stage:** MBS collects vehicle CAMs in lower frequency bands, extracts context, and predicts future blockage rates based on trajectory forecasting. The context space is divided into uniform hypercubes containing classified context. Simultaneously, serving SBSs reports vehicle QoS.
- **HO detection stage:** Based on blockage status and current QoS, MBS triggers HO processes for cases including long future blockage duration or imminent significant QoS drops.
- **Execution and learning stage:** Before HO execution, MBS chooses to either explore (random arm selection) or exploit (select highest-reward arm). This trade-off enables system exploration while maintaining performance. As context inputs and reward tables update over time, MBS

learns cross-scenario similarities for improved decision-making.

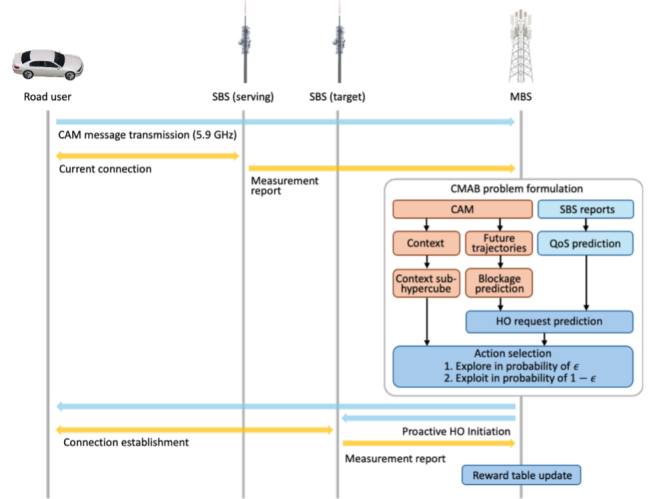


Fig. 3. Blockage-aware Proactive HO System Architecture

##### B. System Model

In the MBS-governed mmWave vehicular network architecture with multiple  $N_b$  Small Base Stations (SBSs) strategically deployed throughout the coverage area, let  $\mathcal{B} = \{1, \dots, j, \dots, N_b\}$  represent the comprehensive set of available SBSs within the geographical region. The MBS serves as the central coordination entity that collects detailed vehicle dynamics information through standardized CAM via Vehicle-to-Infrastructure (V2I) communication links at each discrete time frame  $t = 1, \dots, T$  within a finite time horizon  $T \in \mathbb{N}$ .

The dynamic nature of vehicular networks necessitates flexible vehicle set definitions, where  $\mathcal{U}(t) = \{1, \dots, i, \dots, |\mathcal{U}(t)|\}$  represents the time-varying vehicle population as vehicles enter and exit the network coverage area. Each vehicle  $v_{t,i}$  is characterized by a multi-dimensional context vector  $x_{t,i}$  drawn from the  $X$ -dimensional context space  $\mathcal{X} = [0, 1]^X$ , where normalization to the unit hypercube ensures consistent scaling across different contextual attributes. These attributes encompass instantaneous velocity profiles, acceleration patterns, directional heading information, current lane position, and critically, the proximity of large vehicles that may cause potential communication blockages.

Let  $\gamma_{i,j}(t)$  denote the actual Signal-to-Interference-plus-Noise Ratio (SINR) measurement that vehicle  $v_{t,i}$  experiences when communicating with SBS  $j \in \mathcal{B}$  during time slot  $t$ . This SINR measurement serves as the fundamental metric for assessing communication quality. The binary decision variable  $y_{i,j}(t)$  represents the SBS selection process: if SBS  $j$  is selected as the serving base station for vehicle  $i$  in time slot  $t$ , then  $y_{i,j}(t) = 1$ , otherwise  $y_{i,j}(t) = 0$ . The network architecture enforces that each vehicle maintains communication with only one SBS at any given time slot.

Under these conditions, the instantaneous transmission rate achievable by vehicle  $i$  when communicating with SBS  $j$  follows the standard Shannon capacity formula:

$$R_{i,j}(t) = W \log_2(1 + \gamma_{i,j}(t)) \quad (1)$$

where  $W$  represents the available communication bandwidth.

Let  $\beta_i(t)$  indicate the index of the currently serving BS for vehicle  $i$  at time  $t$ . The total system capacity, serving as the primary performance metric, is calculated by aggregating transmission rates of all active vehicle-SBS pairs:

$$r(t) = \sum_{j \in \mathcal{B}} \sum_{i \in \mathcal{U}(t)} y_{i,j}(t) R_{i,j}(t) \quad (2)$$

The central optimization problem is formulated as maximizing Equation 2, representing the fundamental goal of achieving optimal network-wide communication performance through intelligent handover management and resource allocation strategies.

### C. Contextual Multi-Armed Bandit

To address these challenges, we propose a future blockage-aware CMAB framework for proactive HO management in mmWave vehicular networks. In this approach, each BS represents an "arm" in the bandit formulation, where the context input is derived from real-time CAM containing comprehensive vehicle dynamics and environmental information. The framework leverages the Upper Confidence Bound (UCB) algorithm [10], which addresses the exploration-exploitation trade-off by maintaining confidence intervals for each arm's expected reward and selecting the arm with the highest upper confidence bound at each time step.

The UCB algorithm operates by computing an index for each base station  $j$  as:

$$\hat{\mu}_{j,N_j(t)} + \sqrt{\frac{\alpha \log t}{2N_j(t)}} \quad (3)$$

where  $\hat{\mu}_{j,N_j(t)}$  represents the empirical mean reward from base station  $j$  after  $N_j(t)$  selections, and the confidence radius:

$$\sqrt{\frac{\alpha \log t}{2N_j(t)}} \quad (4)$$

ensures systematic exploration of potentially optimal actions. The reward function corresponds to the transmission rate  $R_{i,j}(t)$  calculated in Equation 2, providing direct feedback on communication quality for each base station selection.

By analyzing vehicle trajectories, velocities, and the proximity of large vehicles that may cause communication obstructions, the system can anticipate potential blockage events before they occur [11]. This predictive capability enables context-dependent arm selection that accounts for future blockage scenarios, allowing the algorithm to proactively select base stations that are less likely to experience communication disruptions. The contextual information enables differentiated treatment of similar vehicles in different environmental conditions, leading to more robust handover decisions that maintain communication continuity in dynamic vehicular environments.

## V. CONCLUSION AND FUTURE WORK

This work has presented a comprehensive CMAB-based handover scheme specifically designed for mitigating dynamic blockage challenges in mmWave V2X communication scenarios. The proposed approach represents a significant advancement in proactive handover management by incorporating predictive blockage analysis and contextual decision-making capabilities that extend beyond traditional reactive schemes. The integration of contextual multi-armed bandit theory with vehicular communication systems provides a robust framework for handling the inherent uncertainty and dynamic nature of mobile environments. By leveraging comprehensive vehicle context information, including mobility patterns and blockage indicators, our approach enables more informed handover decisions that can improve overall system performance while reducing unnecessary handover frequency.

Future research directions will focus on exploiting predictive vehicle information and advanced mobility forecasting techniques to enable truly proactive handover management strategies. The incorporation of machine learning models capable of predicting future vehicle trajectories, environmental changes, and potential blockage scenarios will enhance the system's ability to make optimal handover decisions well in advance of communication quality degradation. The practical validation of the proposed approach will be conducted using comprehensive real-world vehicular datasets collected from urban environments in Japan, providing realistic traffic patterns, vehicle density distributions, and mobility characteristics necessary for thorough performance evaluation under diverse operational conditions, including peak traffic hours, complex urban geometries, and varying weather conditions that affect mmWave propagation characteristics.

## VI. ACKNOWLEDGEMENT

This work was supported by JST ASPIRE Grant Number JPMJAP2325, Japan.

The authors acknowledge the use of AI-assisted tools for language editing and writing improvement during the preparation of this manuscript.

## REFERENCES

- [1] ETSI, "EN 302 637-2 V1.4.1 Intelligent Transport Systems (ITS); Communications Architecture; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service," 2019.
- [2] 3GPP, "3rd generation partnership project; study on enhancement of 3gpp support for 5g v2x services (release 16)," 3GPP, Tech. Rep. TR 22.886 V16.2.0 (2018-09), 2018.
- [3] C. G. Ruiz et al., "Analysis of blocking in mmwave cellular systems: Application to relay positioning," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1329–1342, 2021.
- [4] A. K. Gupta et al., "Macrodiversity in cellular networks with random blockages," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 996–1010, 2018.
- [5] P. A. Lopez et al., "Microscopic Traffic Simulation using SUMO," in *Proc. 21st IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2018.
- [6] M. Christopoulou et al., "Artificial Intelligence and Machine Learning as key enablers for V2X communications: A comprehensive survey", *Vehicular Communications*, Volume 39, 2023, 100569, ISSN 2214-2096.

- [7] K. Qi et al., "Federated Learning Based Proactive HO in Millimeter-wave Vehicular Networks," 2020 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 2020, pp. 401-406.
- [8] C. Lee et al., "Intelligent Dual Active Protocol Stack HO Based on Double DQN Deep Reinforcement Learning for 5G mmWave Networks," in IEEE Transactions on Vehicular Technology, vol. 71, no. 7, pp. 7572-7584, July 2022.
- [9] A. Alizadeh and M. Vu, "Reinforcement Learning for User Association and Handover in mmWave-Enabled Networks," IEEE Transactions on Wireless Communications, vol. 21, no. 11, pp. 9712-9728, Nov. 2022.
- [10] A. Slivkins, "Introduction to Multi-Armed Bandits," arXiv preprint arXiv:1904.07272, 2019.
- [11] W. Chi, J. Nakazato, T. Murakami and M. Tsukada, "V2I Blockage Modeling and Performance Evaluation for Connected Autonomous Vehicle," in Proc. IEEE 99th Vehicular Technology Conference (VTC2024-Spring), Singapore, 2024, pp. 1-5.