# Ultra-density Aware Learning-based Handover Management in High-mobility 5G Vehicular Networks

Mubashir Murshed, Glaucio H. S. Carvalho, and Robson E. De Grande
Department of Computer Science
Brock University, Canada
Email: {mmurshed,gdecarvalho,rdegrande}@brocku.ca

***Abstract*—Ensuring connection stability is crucial for both vehicular safety and user experience. With the increasing amount of data sharing among connected vehicles, there is a need for more bandwidth, stability, and reliability. While 5G technology can offer these benefits with its small cellular range and densification, it also presents a challenge in frequent handovers (HOs). This issue can result in unnecessary HO, HO failures, and ping-pong effects, negatively impacting service delivery and compromising safety data sharing. To this end, we present High-mobility and Ultra-density Aware Handover decision-making (HMUD-H) approach using the SARSA Reinforcement Learning algorithm for connection management, which efficiently makes HO decisions to ensure stable connectivity. The HMUD-H algorithm is adaptable and can handle dynamic, highly mobile, and ultra-dense vehicular networks. Realistic simulated analyses have demonstrated that our algorithm significantly reduces the number of HOs, average cumulative HO time, HO failures, and ping-pong effects, thus improving overall connection stability.**

***Index Terms*—5G, SARSA, Handover, High-mobility, Ultra-density of Networks**

## I. Introduction

Connection stability plays a significant role in ensuring vehicular safety and infotainment services [1]. Vehicles generate vast data with the enriched On-Board-Units (OBU), and exchange information with other vehicles, traffic infrastructure, and smart devices. A stable connection is crucial for safety-related data, as even a brief connection drop can lead to disastrous consequences. In addition, with the rising expectations for user experience, a reliable, high bandwidth, and low latency connection is necessary for vehicular networks to ensure a seamless experience [2].

In urban and highway scenarios, high mobility and random movement are common, adding to the complexity of ensuring a stable connection [2]. The growing number of vehicles on the road leads to an ultra-dense network, posing a significant challenge in maintaining a stable connection in terms of network load balancing. To address this challenge, the 5G standard offers a potential solution for vehicular networks. Designed to deliver high data rates, ultra-low latency, small cell densification, improved energy efficiency, and robustness, the 5G standard shows great potential as a data communication technology for vehicular networks [3].

The handover (HO) operation is ultimately needed and executed for vehicles to keep connected and support ongoing services. However, HO can affect connection stability, especially for low coverage ranges of the 5G network where frequent HO may be necessary to remain connected and support ongoing services [2]. It is essential to make intelligent HO decisions to ensure a stable connection and enhance user experience. However, frequent HO and HO times can impact service delivery; HO failures, unnecessary HO, and ping-pong effects can also affect connection stability and degrade user experience [2].

Numerous studies have addressed the issue of connection stability. Various techniques based on different methods and technologies, such as software-defined networks (SDN), heterogeneous networks (HetNet), edge computing, hybrid networks, and virtual cell (VC), have been proposed to reduce HOs and improve connection stability [2], [4]–[6]. Despite these efforts, optimizing stable connections in 5G networks remains challenging, especially for high-mobility vehicles and ultra-dense networks in extensive scenarios. To tackle the challenge of maintaining stable connections in 5G vehicular networks by reducing HO, HO time, HO failure, and ping-pong effects in HO, we present an adaptive learning technique utilizing the SARSA Reinforcement Learning algorithm. Our approach involves continuous communication between vehicles and towers while traversing various road segments, enabling real-time parameter adjustment of the adaptive learning model. Our algorithm can make efficient HO decisions by maintaining connection stability in both urban and highway scenarios.

The structure of this paper is as follows. Section II provides an overview of previous works on connection stability. Section III presents our proposed handover technique for achieving and maintaining stable 5G connections. Section IV defines the performance analysis, presenting and discussing the results. Finally, Section V concludes the paper and provides future research directions.

## II. Related Works

Stable connection for high mobility vehicles is always a concerning research area. Many strategies with different approaches have attempted to support HOs for ensuring stable connection [4]. Previous works showed that connection stability improved through their methods by managing communication cells and achieving a reduction of HOs and HO times. However, high mobility and ultra-dense networks are

This work is partially funded by NSERC Discovery grants.

not considered simultaneously in all handover management approaches for 5G networks, and there is not much work can be found that deals with adaptive learning strategies and considers real-world scenarios [2].

Some approaches have considered the diversity of network technologies and communication media in their strategies to ensure network stability. One approach based on HetNet is proposed to optimize overall service delivery while maintaining reliability and minimizing unnecessary handovers [7]. This approach considers factors, such as connection signal, power, bandwidth, location, and velocity, within a Markov Decision Process (MDP) model. The model explores stochasticity by considering the user's current and future service experience. This is achieved by computing a weighted sum of throughput and handover cost and using an action elimination method to optimize beam width and network actions.

An SDN-based approach is another strategy to ensure continuous service and seamless coverage with minimal handovers. One such approach involves using a Multiaccess Edge Computing (MEC) enabled strategy to manage service-aware handovers to reduce the number of required handovers [8]. Another approach uses SDN in a multi-level view by dividing the network's support into edges and core to optimize handover times [9].

Two models are proposed for improving download time and handover management, which are based on multi-layer many-to-one LSTM architecture and a multi-layer LSTM AutoEncoder (AE) combined with a MultiLayer Perceptron (MLP) neural network [10]. These models leverage heterogeneous data to make more informed handover decisions, taking into account the expected Quality of Experience after the handover rather than relying solely on signal strength before the handover.

Studies have investigated the use of virtual cell (VC) management as a means of maintaining stable connections in 5G networks. The use of probabilities to manage VC formation and updates in a user-centric manner is described in [11]. This approach enables decision-making in a decentralized manner and reduces network complexity. Probabilities are employed to represent network status to define the size of the virtual cell by analyzing the network betweenness, base station degree, and distances. Similarly, in 5G V2X networks, the VC paradigm has been utilized to handle handovers, utilizing probabilities and comparable parameters to select cells [5]. In another approach, an anchor is selected using a static cluster of small cells with localized mobility, assisted by the VC scheme for achieving stable connections [12]. Furthermore, a dynamic approach is employed to update VC by tracking the mobility of vehicles [13], where it models resource management in vehicle-to-everything communication as max-min optimization.

Machine learning techniques have also been applied to handover management in intelligent vehicular networks [2]. A handover management approach proposed to predict signal strength and decide on handovers using recurrent neural network model to [14]. A stochastic Markov model is then used to determine the new access point based on the predicted signal strength. This approach was presented in a study focusing on improving handover management in intelligent vehicular networks.

Our study finds that many previous works on handover management have faced challenges such as slow decision-making, high computational requirements, and network complexity [2]. Furthermore, a few previous approaches incorporate adaptive learning strategies, and most of them relied on limited connectivity parameters to estimate general communication conditions and did not provide a guarantee of stable connections.

## III. High-mobility and Ultra-density Aware Handover decision-making

Handover is essential for ensuring stable connections, and the decision-making process must be executed efficiently to enhance the user experience. We propose an intelligent handover decision-making approach using SARSA Reinforcement Learning (SRL), named High-mobility and Ultra-density Aware Handover decision-making (HMUD-H), to ensure stable connections in 5G vehicular networks. We assume that vehicles run in a scenario where multiple 5G towers (gNodeB) are placed. Vehicles continue exchanging beacon messages with every gNodeB while they remain within the communication range. The beacon message contains measurement reports, including connectivity parameters for handover decision-making. We also assume the beacon message with the serving gNodeB contains a measurement report with connectivity parameters and SRL information. HMUD-H considers several connectivity parameters, such as Received Signal Strength Indicator (RSSI), Reference Signal Received Quality (RSRQ), and Signal to Interference & Noise Ratio (SINR), distance, density, and vehicle speed.

We assume a backbone controller runs in the system to perform all necessary HO decision-making operations. We refer to the controller as gNodeB for the sake of simplification. Each gNodeB runs the SRL algorithm and keeps a Q-table, updated over time and shared among the network. A vehicle registers for service as soon as it is within the range of a gNodeB. The gNodeB then updates the storage area for the Q-table with measurement reports of the cell for the vehicle. Vehicles periodically send beacon messages to the connected serving gNodeB for necessary computation and operation of connection management. The serving gNodeB performs SRL operation, updates Q-table for each vehicle, and forwards it to the next selected gNodeB at the time of HO.

Let us consider a finite number of vehicles denoted as $V = \{v_1, v_2, ..., v_i\}$ are traversing in an intelligent transportation scenario where a number of 5G towers (gNodeBs) are deployed, represented as $\mathcal{TT} = \{tt_1, tt_2, ..., tt_j\}$. Vehicle $v_i$ is within the communication range of gNodeBs of $\mathcal{TT}$ at time $k$; such gNodeBs are represented as $\mathcal{T} = \{\tau_1, \tau_2, ..., \tau_j\}$, where $\tau_j \in \mathcal{T}$, and $\mathcal{T} \subset \mathcal{TT}$.

### A. Connectivity Parameters

Handover decision depends on the Connectivity Parameters (CPs). We employ CPs with several signal parameters, including RSSI, RSRQ, and SINR, as well as mobility factors, such as distance and speed. We consider multiple connectivity parameters because of the parameters' relevant impact on network connectivity differently [15] [16].

*1) Signal Measurement:* We consider $\chi^i_{rssi_{\tau_j}}$, $\chi^i_{rsrq_{\tau_j}}$, and $\chi^i_{sinr_{\tau_j}}$ signal levels to represent the quality of RSSI, RSRQ, and SINR reading estimates, respectively. Many signal measurements in CPs allow dealing with the diversity and dynamics of environments, which have a direct impact on the stability of network connections [15]. The exchanged beacon messages of a vehicle $v_i$ and gNodeB $\tau_j$ have signal values in the measurement report $\mathcal{M}$. Beacon messages carry the generated RSSI, RSRQ, and SINR, as well as their parameters; an evenly spaced time interval $k_1, k_2, ..., k_n$ defines the frequency of the messages. The beacon message exchanged with the serving gNodeB contains information of $\mathcal{M}$ of every $\tau_j \in \mathcal{T}$ for the computation of SRL. To minimize computational cost, this exchanged beacon message takes place in vehicle-gNodeB pairs with at least minimum signal readings.

We convert $\chi^i_{rssi_{\tau_j}}$, $\chi^i_{rsrq_{\tau_j}}$, and $\chi^i_{sinr_{\tau_j}}$ to the form where a higher value means better connectivity. We have converted in this form so that our approach takes decisions in a fixed direction for SRL computation; a higher value means a better result, and a lower value is the opposite. Among all factors related to networks and communication, $\chi^i_{rssi_{\tau_j}}$, $\chi^i_{rsrq_{\tau_j}}$, and $\chi^i_{sinr_{\tau_j}}$ present high significance to the connectivity status, being, as signal measurement CPs. They present additional information when existing decision factors are insufficient for connectivity management decisions and the overall functioning of a connection [17].

*2) Mobility:* We consider distance and speed as the mobility parameters for HMUD-H. For calculating distance, each $v_i$ keeps $(x, y)$ positions of its traversing path in the storage area of its OBU. We presume that gNodeB positions are known across the entire system. A vehicle $v_i$ uses the Euclidean distance according to Equation (1) to determine its distance to $\tau_j \in \mathcal{T}$ at time $k$.

$$d(v_i, \tau_j) = \sqrt{\left(x_{v_i} - x_{\tau_j}\right)^2 + \left(y_{v_i} - y_{\tau_j}\right)^2} \tag{1}$$

The exchanged beacon messages contains $d(v_i, \tau_j)$ in the $\mathcal{M}$ of every gNodeB $\tau_j \in \mathcal{T}$. This information is shared with the current $\tau^\xi_j$ for further operation of SRL.

Speed is another factor considered as mobility CP. The speed of each $v_i$ is measured and stored in the OBU. The speed data is added to the $\mathcal{M}$ and sent with the beacon message to the serving gNodeB $\tau^\xi_j$ for further processing of SRL.

*3) Density:* The number of vehicles linked to a gNodeB is referred to as its density when a high density compromises network stability. In our proposed approach, beacon messages are used to exchange $\mathcal{M}$ with the density of $\tau_j \in \mathcal{T}$. It is assumed that all gNodeBs in $\mathcal{TT}$ shared data related to density among themselves.

For each vehicle and corresponding gNodeB, an estimate of the density ratio $R_{\tau_j}$ is calculated using Equation (2), where $l\tau_j$ denotes the load of $\tau_j$ (number of associated vehicles), and $\sum_{l_{tt_j} \in \mathcal{TT}}(l_{tt_j})$ represents the sum of the number of associated vehicles with gNodeBs in $\mathcal{TT}$ in the given scenario. The ratio $R_{\tau_j}$ is a measure of relative density compared to a single gNodeB. A lower value of $R_{\tau_j}$ indicates a higher density of the gNodeB, reducing the likelihood of selecting $\tau_j$ as the serving gNodeB.

$$R_{\tau_j} = 1 - \frac{l_{\tau_j}}{\sum_{l_{tt_j} \in \mathcal{TT}}(l_{tt_j})} \tag{2}$$

Equation (3) describes a dynamically updated value of density $\chi^i_{density_k}$ which is calculated by summing average density $\overline{R_{\tau_j}}$ and hysteresis value is determined by dividing the density ratio $R_{\tau_j}$ by the hysteresis value from the previous time period $\chi^i_{density_{k-1}}$.

$$\chi^i_{density_k} = \overline{R_{\tau_j}} + \frac{R_{\tau_j}}{\chi^i_{density_{k-1}}} \tag{3}$$

### B. SRL Model

Based on the recently collected connectivity parameters, our proposed approach HMUD-H determines which gNodeB $\tau^\xi_j$ among the towers $\tau_j \in \mathcal{T}$ is most suited to serve the vehicle $v_i$ based on the recently gathered connection metrics. We integrate adaptive learning in the dynamic vehicular environment using the SRL algorithm. By employing SARSA, which is an online temporal difference learning technique, agents can effectively learn the value of policies and associated actions during state-action transitions [18]. SRL's unique advantage lies in its ability to function without a predefined dataset or prior training or testing, making it an ideal fit for the complex and rapidly evolving vehicular network. With the high dynamic behaviour of vehicles, relying on pre-defined static models for decision-making can be challenging to achieve efficiently.

The SRL approach involves taking an action $a_k$ in state $s_k$ according to policy $\pi$, receiving a reward of $r_{k+1}$ and transitioning to the next state $s_{k+1}$ [18]. Subsequently, the action $a_{k+1}$ is taken in the next state $s_{k+1}$ based on policy $\pi$. This is done in an online fashion using the tuple $(s_k, a_k, r_{k+1}, s_{k+1}, a_{k+1})$. The Q-values are updated using $\alpha$ and $\gamma$ – learning rate and discount factor, respectively – under the state-action transitions [19].

We define the components of our proposed HMUD-H model as vehicular connectivity requirements to cope with high mobility and ultra-density to guarantee stability. HMUD-H performs its operation each time $k$.

*1) State:* We define states as gNodeB serving a vehicle. The current serving gNodeB $\tau^\xi_j$ is the current state for $v_i$ at time $k$. The state $s_{k+1}$ can be any in-range gNodeB of $\tau_j \in \mathcal{T}$.

*2) Action:* In state $s_k$, there are multiple actions $a_k$ available, which can be denoted as $a_{k,1}, a_{k,2}, ..., a_{k,m}$, with the same number as the gNodeBs that are in range. The action $a_{k,m}$ may lead to migration from the current gNodeB to available gNodeBs $\tau_j \in \mathcal{T}$ at time $k$. The available actions in state $s_k$ can be expressed as $a_{k,m}\tau^\xi_j \rightarrow \tau_j \in \mathcal{T}$. Upon performing the action $a_{k,m}$, the state transitions to $s_{k+1}$, where the serving tower $\tau^\xi_j$ is changed to the gNodeB of $s_{k+1}$.

*3) Reward:* By transitioning from state $s_k$ to $s_{k+1}$ through the action $a_k$, the corresponding reward $r_{k+1}$ is obtained. The scalar combination $\mathcal{CPR}_{k+1}$ is calculated using $\chi^i_{rssi_{\tau_j}}$, $\chi^i_{density_{\tau_j}}$ and $d(v_i, \tau_j)$ of an action $a_k$ as the reward $r_{k+1}$ using Equation (4).

$$\begin{aligned} \mathcal{CPR}_{k+1} = (\chi^i_{rssi_{\tau_j}} * \hat{w}_{rssi}) + (\chi^i_{density_j} * \hat{w}_{denscity}) \\ + ((d^x - d(v_i, \tau_j)) * \hat{w}_d) \end{aligned} \tag{4}$$

Here, $\hat{w}_{rssi}$, $\hat{w}_{dencity}$ and $\hat{w}_d$ are the weights of RSSI, density, and distance, respectively, and $d^x$ conditions $d(v_i, \tau_j)$ where the higher value of the distance, better is the connection. This relation simplifies the reward calculation - a higher reward provides better performance with a higher Q-value.

*4) Policy:* Our approach utilizes a well-known $\epsilon$-greedy policy in reinforcement learning to balance the exploration-exploitation trade-off [18] [19]. At the time $k$, we calculate $\epsilon_k$ as a scalar combination of the average CPs $\sum_b^e CP_j/e$ of all gNodeBs within range over a specific time interval. Initially, an action is chosen randomly at time $k$, and the exploration probability is determined by $\epsilon_k$. The action with the highest reward is selected as the iterations continue and follow the exploitation probability $(1 - \epsilon_k)$.

*5) Q-value:* RSRQ and SINR are regarded as Q-values, while their scalar combination is considered the overall Q-value. Q-value is calculated using Equation (5), where $\hat{w}_{rsrq}$ and $\hat{w}_{sinr}$ are the weights of RSRQ and SINR, respectively.

$$Q_{val} = (\chi^i_{rsrq_{\tau_j}} * \hat{w}_{rsrq}) + (\chi^i_{sinr_{\tau_j}} * \hat{w}_{sinr}) \tag{5}$$

The signal is better when RSRQ and SINR have higher scalar values. Hence, a gNodeB is better for connection stability if its Q-value is higher. Moreover, we employ the scalar combination of CPs as the reward, maintaining the variety of connectivity-related variables in HMUD-H. It is noted that all the parameters that have impacts on connection stability are added in our proposed approach, which makes it more adaptive to the real-world dynamic scenario.

### C. HMUD-H Functioning

Upon the registration of $v_i$ in a network, $v_i$ is connected to the first $\tau_j$ that comes to its range. The gNodeB $\tau_j$ starts updating the Q-table and executing the Q-value of SRL for each $v_i$. At the initial stage, a default Q-value is selected, and state, action, and reward are chosen according to the first registered gNodeB of a vehicle. In subsequent iterations, the Q-value of HMUD-H is updated by taking an action in a state, observing its corresponding reward, and utilizing the policy function. This process is performed on gNodeBs.

In HMUD-H, the serving gNodeB $\tau_j^\xi$ calculates Q-value for each $v_i$. $\tau_j^\xi$ maintains Q-table for each $v_i$ separately. The Q-value for each $v_i$ is generated for every gNodeB $\tau_j \in \mathcal{T}$. $v_i$ exchanges beacon messages with every gNodeB that is in the range that contains $\mathcal{M}$, and updates its storage of $\mathcal{M}$ at each time $k$. The $\mathcal{M}$ is exchanged with the current $\tau_j^\xi$ at each time $k$ for performing SRL. When HO takes place, it exchanges the beacon message with $\mathcal{M}$ that contains the latest Q-value of the Q-table to the new $\tau_j^\xi$. The calculation of the Q-value update is defined by Equation (6).

$$Q_{val}(s_k, a_k) \leftarrow Q_{val}(s_k, a_k)^\xi + \alpha[\mathcal{CPR}_{k+1} + \gamma Q^{mbr}_{val_{\tau_j}} - Q^{srv}_{val_{\tau_j}}] \tag{6}$$

Suppose that a handover to gNodeB $\tau_j^\xi$ occurs at time $k$. The Q-value of the selected serving gNodeB is updated to $Q_{val}(s_k, a_k)^\xi$ during this period. The Q-value of gNodeB $\tau_j \in \mathcal{T}$ is defined as $Q^i_{val}(s_{k+1}, a_{k+1})$, which is equivalent to the next state's Q-value. Consequently, $Q^i_{val}(s_{k+1}, a_{k+1}) = Q^{mbr}_{val_{\tau_j}}$ defines the Q-value of in-range gNodeBs $\tau_j \in \mathcal{T}$.

**ALGORITHM 1:** High-mobility and Ultra-density Aware Handover decision-making (HMUD-H)

**Data** : $\chi^i_{rssi_{\tau_j}}$, $\chi^i_{rsrq_{\tau_j}}$, $\chi^i_{sinr_{\tau_j}}$, $d(v_i, \tau_j)$, $\chi^i_{density_j}$, $\chi^i_{speed}$

**Result:** $\tau_j^\xi$

$InitializeCondition$;
**while** ($\tau_j \in \mathcal{T} \neq \emptyset$) **do**
  $a_k \leftarrow \tau_j \in \mathcal{T}$;
  $\mathcal{CPR}_{k+1} \leftarrow sca(\chi^i_{rssi_{\tau_j}}, \chi^i_{density_j}, d(v_i, \tau_j))$;
  $\epsilon_k \leftarrow calPolicy()$;
  $\alpha$, $\gamma \leftarrow setBasedOnSpeed$;
  $Q_{val} \leftarrow sca(\chi^i_{rsrq_{\tau_j}}, \chi^i_{sinr_{\tau_j}})$;
  $Q_{val}(s_k, a_k) \leftarrow Q_{val}(s_k, a_k)^\xi + \alpha[\mathcal{CPR}_{k+1} + \gamma Q^{mbr}_{val_{\tau_j}} - Q^{srv}_{val_{\tau_j}}]$;
  **while** ($\tau_j^\xi$ $remainedSame$) **do**
    $Q_{val}(s_k, a_k)^\tau \leftarrow \overline{Q_{val}}$ $of$ $cur$ $\tau_j^\xi$;
  **if** ($Q_{val}(s_k, a_k)_{\tau_j} > Q_{val}(s_k, a_k)^\tau$) **then**
    $\tau_j^\xi \leftarrow \tau_j$ – execute handover;
  **else**
    remain connected to same $\tau_j^\xi$;

Equation (6) contains $Q^{mbr}_{val_{\tau_j}}$, the scalar form of RSRQ and SINR of a gNodeB, and $Q^{srv}_{val_{\tau_j}}$, the scalar form of RSRQ and SINR of the serving tower that is currently connected. It updates these terms each period $k$. It is worth noting that the update frequency of $Q^{srv}_{val_{\tau_j}}$ and $Q^{mbr}_{val_{\tau_j}}$ in each time $k$ while $Q_{val}(s_k, a_k)^\xi$ refers to the value during last HO.

Each vehicle updates the Q-value respective to a state transition $s_k$ to $s_{k+1}$ of every available action $a_k$. A table is shared with gNodeB $\tau_j^\xi$ to store the updated Q-values. The Q-value of current $\tau_j^\xi$ is averaged $Q_{val}(s_k, a_k)^\tau$ and stored in the Q-table until HO is performed to a new $\tau_j^\xi$. When a Q-value $Q_{val}(s_k, a_k)_{\tau_j}$ for actions $a_k$ at time $k$ within in range $\tau_j \in \mathcal{T}$ is greater than averaged Q-value $Q_{val}(s_k, a_k)^\tau$, HMUD-H performs HO operation to the new $\tau_j^\xi$.

We have categorized vehicle speeds into three groups, namely faster (121 km/h to 180 km/h), fast (61 km/h to 120 km/h), and medium (0 km/h to 60 km/h), and we use the vehicle speed as a tuning factor for the SRL learning rate ($\alpha$) and discount factor ($\gamma$). Learning must be applied at a higher frequency for higher-speed vehicles to accommodate the fast-paced changes [19]. Specifically, for faster vehicles, we set slightly lower values of $\alpha$ and $\gamma$, while for fast- and medium-speed vehicles, we choose relatively higher values for these parameters. The exact values of $\alpha$ and $\gamma$ for each speed category are shown in Table I, and this approach is applied during the time interval $k' - k$. HMUD-H functioning is summarized in Algorithm 1.

## IV. Performance Analysis and Results

A series of simulations have been conducted using Veins, OMNet++, SUMO, and Simu5G to analyze the proposed HMUD-H approach for connection stability in vehicular networks. Our simulation environment employs the Cologne, Germany map, which presents both urban and highway scenarios. This map is useful in simulating ultra-dense networks and high-mobility and dynamic random situations that vehicles may encounter.

### A. Parameters

We assess the performance of our proposed approach using various parameter settings, as summarized in Table I. The performance of handover (HO) in terms of connectivity is influenced by factors, such as vehicle density and speed. We randomly place ten 5G gNodeBs over the map, strategically inside and outside each other's range to ensure different overlapping communication ranges.

### B. Performance Metrics

The effectiveness of our approach has been evaluated using a variety of performance indicators on vehicles with various densities and levels of mobility. The **number of HO** estimates the overall number of HO. The amount of time spent on HO is shown by the **average cumulative HO time**. They are determined by dividing the total number of HOs by the cumulative HO time. The **number of HO failure** is used as another metric by calculating the number of HO failed after attempting. We calculate **number of ping-pong effects during HO**. This is calculated by determining the number of the switching serving tower of HO operation between the previous tower as the immediate last serving tower and a new tower. In other words, HO performed to the immediate last serving tower in the next time step. The lower the value the better the user experience. To assess the frequency of HOs across different speeds, we compute the **percentage of HO**. This is calculated by dividing the total number of HO by the total count of vehicles that traveled within specific speed ranges. A lower HO percentage indicates better performance, as it measures the variability of HO incidents across various speed ranges. Moreover, we calculate the **percentage of HO failures** by dividing the total number of failures by the total number of handovers that occurred within each speed range, a smaller proportion of HO failures indicates an improved performance.

### C. Results

We evaluate our proposed approach across various vehicle densities ranging from 100...3100. The results are the averages obtained from $30+$ runs with different seeds, with $95\%$ confidence intervals.

The performance of HMUD-H is compared with a dynamic RSSI-based approach named Naive Signal RSSI (NSIG-R). It is a simple threshold-based naive HO decision-making approach to maintain connectivity. In each time $k$, the threshold of RSSI is adjusted by dividing the current RSSI by an average RSSI of a time interval. When a vehicle observes an RSSI value lower than the threshold – HO occurs. For comparison, we also consider a previous work named a two-tier Machine Learning-based scheme (TTML) [14]. TTML observes the stochasticity of the environment for HO decision-making.

The total number of handovers (HO) increases with the density of vehicles, as shown in Figure 1a. In this regard, the number of HO for HMUD-H ranges between 500...15000, while TTML and NSIG-R have ranges of 2000...60000 and 2500...85000, respectively. The graph depicts an almost linear increase in the number of HOs as the density of vehicles increases. However, the slope is lower for HMUD-H compared to TTML and NSIG-R, indicating that HMUD-H performs better than the other two methods in reducing the number of HO. The number of HOs directly impacts the total HO time, as shown in Figure 1b. The average cumulative time for HO in HMUD-H ranges from 10...380, while TTML and NSIG-R have average ranges of 20...600 and 23...560, respectively. TTML requires longer time due to the need for additional time for LSTM training and testing purposes during a certain period. Among three of them, HMUD-H performs better.

Table I: Simulation Parameters.

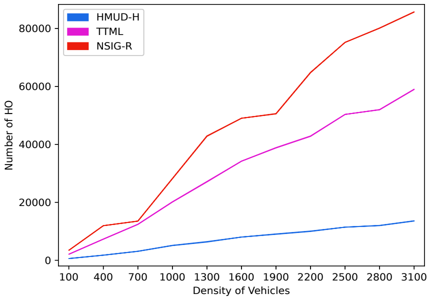
| Parameter | Value Range |
|---|---|
| Simulation Area | $5 * 5$ km$^2$ |
| Density of Vehicle | $100 - 3100$ |
| Speed of Vehicle | 0km/h - 180km/h |
| Num of gNodeB | 10 |
| Distribution of gNodeB | Random |
| Comm. Range of gNodeB | 1000m |
| PHY Model | $5G$ |
| Transmission Power (gNodeB) | 46dbm |
| Transmission Power (Vehicle) | 26dbm |
| Medium, Fast, Faster ($\alpha$) | 0.8, 0.5, 0.3 |
| Medium, Fast, Faster ($\gamma$) | 0.8, 0.5, 0.1 |

Figure 1c depicts the total number of HO failures for different densities, where NSIG-R exhibits a linear increase with fluctuations. However, HMUD-H and TTML maintain a low, almost constant value of HO failures. This demonstrates that both HMUD-H and TTML can maintain stable connections without experiencing extensive HO failures. Specifically, the number of HO failures for HMUD-H and TTML ranges from 15...460, while NSIG-R ranges from 85...1370.

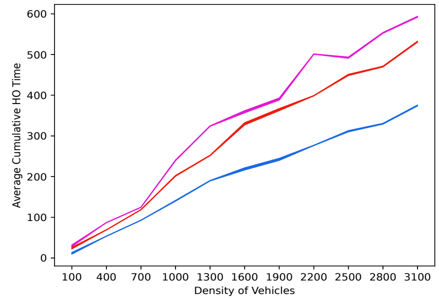
Figure 1d illustrated the total ping-pong effects on HO for different densities. It remains almost flat with a low value for HMUD-H 20...400. It increases linearly for TTML. NSIG-R has a larger value with some fluctuations. The number of ping-pong effects on HO for TTML and NSIG-R is 100...3200 while 200...3900.

We analyzed the impact of speed on the HO decision-making process by examining different speed ranges in 20km/h intervals, from $0-20$km/h to $160-180$km/h in a scenario with 1000 vehicles. The results are presented in Figure 1e, showing the HO percentage for each speed range. Our proposed approach, HMUD-H, exhibits a slightly increasing pattern with respect to speed, which is expected since high-speed vehicles require more HO to maintain stable connections. In contrast, TTML displays fluctuations in the percentage of HO, while NSIG-R shows a sharply linear increasing pattern. The difference in HO percentage between HMUD-H and the other methods is significant.

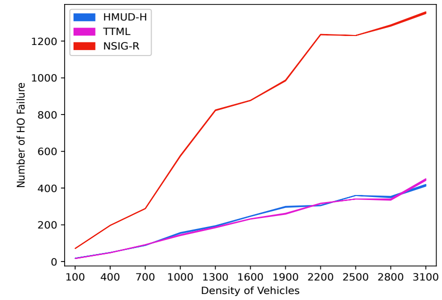
The percentage of the HO failure for every speed range is shown in Figure 1f. HMUD-H remains below 0.3, which means it maintains stability with proper HO decision-making. TTML maintains stability with a range of 0.28...0.43. The previous work with LSTM and MDP leans well toward HO decision-making. NSIG-R has a sharp increasing pattern of starting at a comparatively high value of 0.38...0.85. NSIG-R depends on RSSI for HO-decision making. High-speed vehicles frequently face varied signal values from different gNodeBs. For this reason, it is unstable with speed ranges with a higher value. It is noted that HMUD-H performs better than TTML and NSIG-R.



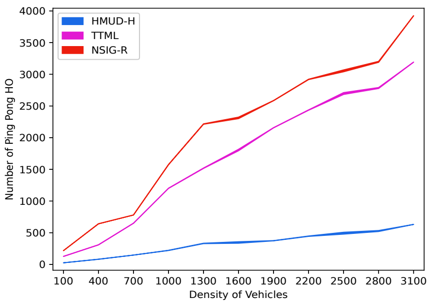

(a) # of HOs vs vehicle density



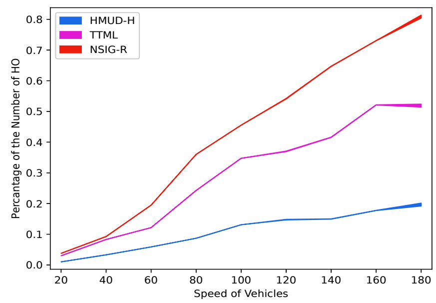

(b) Avg cum. HO time vs vehicle density



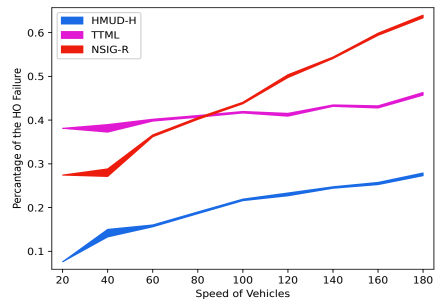

(c) # of HO failures vs vehicle density




(d) # of ping-pong HOs vs vehicle density




(e) Percentage of HOs vs speed




(f) Percentage of HO failures vs speed

Figure 1: Simulation Results of Proposed HMUD-H.

## V. Conclusion

The stability problem of vehicular networks in 5G networks has been addressed in this paper, particularly in ultra-dense networks and high mobility scenarios. To mitigate issues related to HO, such as HO number, HO time, HO failure, and ping-pong effects, an adaptive approach called HMUD-H using SRL has been proposed. The study results demonstrate that HMUD-H provides stable network connectivity while significantly reducing HO overhead, surpassing the performance of TTML and NSIG-R methods. For future work, various SRL learning rates and discount factors will be explored across many scenarios and setups where different topologies, 5G gNodeB placements, and coverage ranges will be varied. In addition, long- and short-term time dependencies will be explored in other learning techniques.

## References


[1] F. Cunha, G. Maia, H. S. Ramos, B. Perreira, C. Celes, A. Campolina, P. Rettore, D. Guidoni, F. Sumika, L. Villas, R. Mini, and A. Loureiro, *Vehicular Networks to Intelligent Transportation Systems*. Springer, 2018, pp. 297–315.

[2] N. Aljeri and A. Boukerche, "Mobility management in 5g-enabled vehicular networks: Models, protocols, and classification," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–35, 2020.

[3] E. A. Oyekanlu, A. C. Smith, W. P. Thomas, G. Mulroy, D. Hitesh, M. Ramsey, D. J. Kuhn, J. D. Mcghinnis, S. C. Buonavita, N. A. Looper, M. Ng, A. Ng'oma, W. Liu, P. G. Mcbride, M. G. Shultz, C. Cerasi, and D. Sun, "A review of recent advances in automated guided vehicle technologies: Integration challenges and research areas for 5g-based smart manufacturing applications," *IEEE Access*, vol. 8, pp. 202 312–202 353, 2020.

[4] L. Tuyisenge, M. Ayaida, S. Tohme, and L.-E. Afilal, "Handover mechanisms in internet of vehicles (iov): Survey, trends, challenges, and issues," in *Global Advancements in Connected and Intelligent Mobility: Emerging Research and Opportunities*. IGI Global, 2020, pp. 1–64.

[5] C. R. Storck, E. d. O. Efrem, G. d. O. Guilherme, R. A. Mini, and F. Duarte-Figueiredo, "Fivh: A solution of inter-v-cell handover decision for connected vehicles in ultra-dense 5g networks," *Elsevier Vehicular Communications*, vol. 28, p. 100307, 2021.

[6] T. Sahin, M. Klugel, C. Zhou, and W. Kellerer, "Virtual cells for 5g v2x communications," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 22–28, 2018.

[7] S. Zang, W. Bao, P. L. Yeoh, B. Vucetic, and Y. Li, "Managing vertical handovers in millimeter wave heterogeneous networks," *IEEE Trans. on Communications*, vol. 67, no. 2, pp. 1629–1644, 2018.

[8] S. D. A. Shah, M. A. Gregory, S. Li, R. Fontes, and L. Hou, "Sdn-based service mobility management in mec-enabled 5g and beyond vehicular networks," *IEEE Internet of Things Journal*, 2022.

[9] A. Gharsallah, F. Zarai, and M. Neji, "Exploring software-defined network (sdn) for seamless handovers in future vehicular networks: Mobility management approach," *Int. Journal of Smart Security Technologies (IJSST)*, vol. 9, no. 1, pp. 1–16, 2022.

[10] Z. Ali, M. Miozzo, L. Giupponi, P. Dini, S. Denic, and S. Vassaki, "Recurrent neural networks for handover management in next-generation self-organized networks," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2020, pp. 1–6.

[11] C. R. Storck, G. d. O. Guilherme, and F. Duarte-Figueiredo, "A vehicle-centric probabilistic approach to virtual cell management in ultra-dense 5g networks," in *Proc. of the IEEE ISCC*, 2020, pp. 1–7.

[12] N. Meng, H. Zhang, and B. Lin, "User-centric mobility management based on virtual cell in ultra-dense networks," in *Proc. of the IEEE/CIC ICCC*, 2016, pp. 1–6.

[13] H. Xiao, X. Zhang, A. T. Chronopoulos, Z. Zhang, H. Liu, and S. Ouyang, "Resource management for multi-user-centric v2x communication in dynamic virtual-cell-based ultra-dense networks," *IEEE Trans. on Communications*, vol. 68, no. 10, pp. 6346–6358, 2020.

[14] N. Aljeri and A. Boukerche, "A two-tier machine learning-based handover management scheme for intelligent vehicular networks," *Ad Hoc Networks*, vol. 94, p. 101930, 2019.

[15] M. Cicioğlu, "Performance analysis of handover management in 5g small cells," *Computer Standards & Interfaces*, vol. 75, p. 103502, 2021.

[16] V. Sharma, F. Song, I. You, and H.-C. Chao, "Efficient management and fast handovers in software defined wireless networks using uavs," *IEEE Network*, vol. 31, no. 6, pp. 78–85, 2017.

[17] G. Kaur, R. K. Goyal, and R. Mehta, "An efficient handover mechanism for 5g networks using hybridization of lstm and svm," *Springer Multimedia Tools and Applications*, pp. 1–29, 2022.

[18] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[19] S. Zhang, T. Wu, M. Pan, C. Zhang, and Y. Yu, "A-sarsa: A predictive container auto-scaling algorithm based on reinforcement learning," in *Proc. of the IEEE Int. Conf. on Web Services*, 2020, pp. 489–497.