






Intelligent Handover Algorithm for Vehicle-to-Network Communications With Double-Deep Q-Learning

Kang Tan , *Student Member, IEEE*, Duncan Bremner, *Senior Member, IEEE*, Julien Le Kernec , *Senior Member, IEEE*, Yusuf Sambo , *Member, IEEE*, Lei Zhang , *Senior Member, IEEE*, and Muhammad Ali Imran , *Senior Member, IEEE*

Abstract—For vehicle-to-network communications, handover (HO) management enables vehicles to maintain the connection with the network while transiting through coverage areas of different base stations (BSs). However, the high mobility of vehicles means shorter connection periods with each BS that leads to frequent HO, hence raises the necessity for optimal HO decision making for high quality infotainment services. Machine learning is capable of capturing underlying pattern via data driven methods to find optimal solutions to complex problems, and much learning-based HO optimization research has been conducted focusing on specific network setups. However, attention still needs to be paid to the actual deployment aspect and standardized datasets or simulation environments for evaluation. This paper proposes a deep reinforcement learning-based HO algorithm using the input parameters that are configurable in the existing measurement report of cellular networks. The performance of the proposed algorithm is evaluated using the well-known network simulator ns-3 with its official LTE module. A realistic network setup in the city center of Glasgow (U.K.) is configured with vehicle trajectories generated by the routes mobility model using Google Maps Directions API. Evaluation results reveal that the proposed algorithm significantly outperforms the A3 RSRP baseline with an average of 25.72% packet loss reduction per HO, suggesting significant improvement in quality of service of phone call and video streaming, etc. The proposed algorithm also has a small implementation cost compared to some state-of-the-art and should be deployed by a software update to a local BS controller.

Index Terms—Deep learning, handover optimization, reinforcement learning, vehicular networks.

I. INTRODUCTION

THE future intelligent transportation systems (ITS) [1] are becoming a key component in society to improve everyday life, addressing the overarching goal to improve on-road safety and traffic congestion while providing various utility-based on-board services [2]. Vehicular networks are a critical enabler

for information sharing in the ITS arena, and ITS applications introduce a high level of expectation for vehicular communications and networking such as throughput, reliability and latency. This brings new challenges to traditional wireless networks' key performance indicators, and have been receiving considerable attention from the research community.

The objective of vehicular networks is to ensure road safety, increase traffic efficiency, and provide a new level of on-board entertainment. To achieve these goals, vehicles need to connect with other communication entities for information exchange, referred to as vehicle-to-everything (V2X) [3]. Various types of communication are defined in V2X networks based on which entity a vehicle connects to, including vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-network (V2N) and vehicle-to-pedestrian (V2P) [4]. Depending on the application, vehicular networks have strict and differentiated quality of service (QoS) requirements. The technological factors of QoS such as reliability, scalability, and network congestion, can be measured by throughput, packet loss, errors, and latency at a network level [5]. In addition, different radio access technologies (RATs) also exist for vehicles to select based on specific scenarios, further increasing the complexity of V2X. Fig. 1 gives an example V2X scenario showing the different vehicular communication types utilising various RATs.

According to the network blueprint [6], V2N communications will form a major aspect to enable future vehicular networks. V2N enables on-board infotainment services via frequent Internet access, which necessitates high data throughput with large bandwidth between vehicles and base stations (BSs), just as traditional cellular network users' equipment (UEs) do. In cellular networks, the mobility of UEs is handled by a handover (HO) mechanism, which reassigns ongoing communications sessions from one BS to the next, enabling UEs to move between coverage areas of different BSs without breaking the session.

5 G is the latest generation of cellular network currently under deployment and contains some new system-level features. One such feature is the inclusion of higher frequency bands such as the millimetre wave (mmWave) bands to higher bandwidth and higher data rate [7]. These frequency bands are characterised by high directionality, relatively short transmission range, and are sensitive to blockage, hence require dense deployment to ensure adequate coverage, bringing new challenges to traditional

Manuscript received November 23, 2021; revised March 15, 2022; accepted April 15, 2022. Date of publication April 25, 2022; date of current version July 18, 2022. This work was supported by the U.K. EPSRC under Grant EP/S02476X/1. The review of this article was coordinated by Dr. Tao Dusit Niyato. (*Corresponding author: Kang Tan.*)

The authors are with the James Watt School of Engineering, University of Glasgow, G12 8QQ Glasgow, U.K. (e-mail: k.tan.3@research.gla.ac.uk; duncan.bremner@btinternet.com; Julien.LeKernec@glasgow.ac.uk; yusuf.sambo@glasgow.ac.uk; lei.zhang@glasgow.ac.uk; muhammad.imran@glasgow.ac.uk).

Digital Object Identifier 10.1109/TVT.2022.3169804

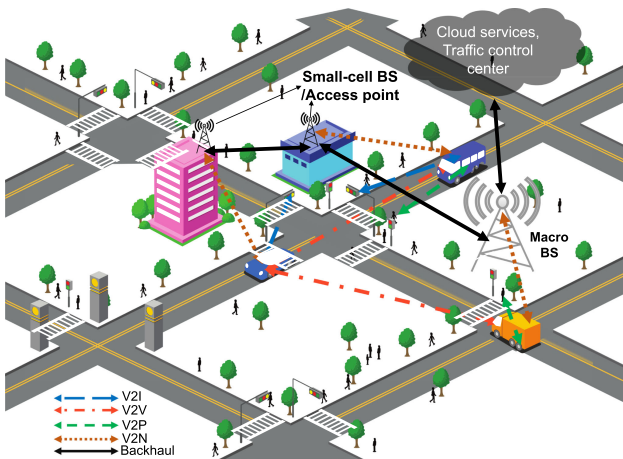


Fig. 1. A heterogeneous V2X network with different communication types and multiple available RATs for each type of communication.

fixed parameter-based HO algorithms. Radio access network slicing [8] is another technology for 5 G and beyond that virtualizes the physical equipment resources into logically independent network slices with specific configurations. It can satisfy QoS requirement for different applications with set physical infrastructures, while also addressing the dual challenges of high network complexity and spectrum resource efficiency in current wireless networks, and is an important enabler for future V2X applications [9]. However, network slicing also introduces the requirement for HO among network slices, as UEs may face the necessity of inter-slice HOs between BSs.

The substantially increasing numbers of BSs and connections with more HO entities to be considered will lead to highly complicated HO management. This rising complexity and the high mobility of on-road vehicles brings challenges to traditional solutions with fixed parameters for HO decision making. Machine learning (ML) that extracts underlying patterns from collected data can cope with the uncertainty and dynamics in V2X networks to provide more intelligent and flexible solutions. In particular, reinforcement learning (RL) has the ability to learn to make optimal decisions by interacting with an environment via trial and error. Additionally, deep learning techniques utilize artificial neural networks (ANNs) capable of further exploiting data patterns and further improving task-specific performances [10]. ANN enabled deep reinforcement learning (DRL) has made significant advances in applications with high complexity and variability such as continuous control and wireless resource allocation [11], [12] and is a promising technique to develop effective solutions to the HO challenges and has gained research momentum [13].

To date, considerable research has been done on the topic of ML-based HO optimization utilizing different input parameters, with various network architecture designs and system setups [14], [15]. However, most of the existing research focuses on specific use cases and system architecture designs, while real-life deployment scenarios for the current network is seldom investigated [16]. In addition, ML-based HO solutions' performance evaluation using a standardized dataset and/or simulation environment has had relatively little research.

Consequently, this paper chose to focus on a DRL-based HO algorithm to tackle the HO optimization problem in cellular V2N networks based on a current cellular network architecture, aiming to explore how DRL may improve the system's HO decision-making performance, and evaluate the performance using the well-established full-stack network simulator ns-3. The simulation developed a realistic scenario setup for the Glasgow city center, U.K., using BS locations and vehicle trajectories that emulate reality. The official ns-3 cellular network module was used in accordance with the 3GPP standards [17]. The algorithm only utilized the reference signal received power (RSRP) parameter available in the current measurement reports for HO as the input parameter, and the dataset was gathered directly from the corresponding network layer of the cellular protocol stack via simulation for training. After offline training, the algorithm was then deployed in the mobility management entity for online performance analysis directly using the ns-3 simulator. Performance comparisons showed a 11.56-second HO time delay reduction per HO and 25.73% packet loss reduction per HO. The contributions of this paper can be summarized as follows:

- 1) A HO algorithm is developed using DRL utilizing the standard input parameters list as available in cellular network configurations, hence can be deployed via a software upgrade with small system level modifications.
- 2) The proposed algorithm is validated on the discrete-event network simulator ns-3 with realistic scenario setups, in contrast to high-level proof-of-concept simulations.
- 3) Result evaluation demonstrates a significant HO delay reduction of 11.56 seconds per HO and 25.73% packet loss reduction per HO compared to the A3 RSRP HO baseline, offering improved performance for 5 G networks.

The rest of this paper is organized as follows: A literature review in Section II; the problem formulation and proposed algorithm in Section IV, followed by the simulation setup and performance evaluation results in Section V. Section VI presents the implementation cost discussion of the proposed algorithm and comparison with the state-of-the-art, culminating the conclusion and future research insights in Section VII.

II. RELATED WORK

In recent years, ML-based solutions have been widely explored in various wireless communication research, including resource management, mobility, and HO management for different systems [15]. ML techniques can utilize the rich dataset generated by wireless systems and extract hidden patterns in the dataset that are usually difficult to derive using analytical optimization techniques [18]. To improve the performance during HO (triggering and decision making), a variety of research has been conducted [14]. ML-based solutions for HO optimization can be classified to three main types: ML-based HO parameter optimization, direct ML-based HO decision making, and ML assisted HO optimization.

To optimize HO parameters, a Q-learning based algorithm was proposed by [19]; by setting the reward function to consider the number of HOs, HO delays and throughput system-wise, the proposed algorithm optimized the value for time-to-trigger

and hysteresis. Similarly, the work of Goyal and Kaushal [20] also utilized Q-learning, combined with an Analytic Hierarchy Process Technique for Order of Preference by Similarity to Ideal Solution, to optimize two HO parameters; hysteresis and time-to-trigger. This scheme utilized information including RSRP, reference signal received quality (RSRQ), signal-to-interference-and-noise ratio (SINR), the UE's location and direction of movement, and the load on each BS to rank neighboring BSs for the Q-learning algorithm to make effective HO decisions.

For ML-based HO decision making, A K-means clustering algorithm was developed in [21] to cluster UEs based on the mobility pattern, followed by an asynchronous multi-agent DRL algorithm for optimal HO decision making. A unified HO algorithm for LTE-A systems was developed in [22] based on discrete stochastic dynamic programming. This algorithm considered both UE measurements (RSRP and RSRQ), and overall resource utilization to produce load-balanced HO decisions. Mollel *et al.* used simulation generated SINR maps and deep Q-learning to calculate adaptive HO decisions in a mmWave vehicular network [23]. Their work used event A2 to trigger HOs as it could indicate a blockage in mmWave networks while also accelerating ANN training by skipping states that were not points of interest. Furthermore, a joint HO and power allocation scheme was developed for heterogeneous networks utilizing multi-agent DRL [24]. Using a reward design based on system throughput and introducing a penalty for HO, the algorithm optimized BS and power level selections for each UE. Tackling the additional HO requirement accompanying network slicing, Sun *et al.* [25] explored a distributed Q-learning method for HO decision making in a network slicing setup where a UE needed to decide whether a HO was required for BSs, network slices, or both. This also proposed a data-sharing mechanism to improve the local training results of each UE and thus the overall network performance. In their later work, Liu *et al.* [26] explored a federated machine learning (FL) [27] setup to further improve the learning architecture using a local aggregator that kept and updated a global model shared by all UEs in the local area.

In the area of ML assisted HO management, prediction-based algorithms that accurately predict metrics for HO or the future location of a vehicle were also being used in HO optimization. An FL training setup for future signal-to-noise (SNR) prediction utilizing both the macro BS and local UEs was proposed [28]. The SNR prediction was added to a conventional HO algorithm to proactively trigger HOs in a mmWave vehicular network. By matching the vehicle's predicted future location with known BS locations, the algorithm proposed in [29] could proactively trigger optimal HOs while reducing the complexity of HO decision making. A recurrent neural network (RNN) based auto-encoder and a multi-layer perception neural network were developed in [30] for HO optimization based on the quality of experience after HO by combining various information gathered throughout the LTE protocol stack for a regression task to assist in optimal HO decision making and provided generalization ability across different scenario setups.

In addition, some HO research combined the previously discussed ML HO optimization techniques to form multi-tier

learning-based HO solutions. For example, a long short-term memory-based RNN was trained in [31] to predict future received signal strength indication (RSSI) values to trigger HO predictively. After triggering, a Hidden Markov Model (HMM) [32] was used to optimize the HO decision making.

According to the literature review, extensive research has been conducted into ML applications in HO management from different optimization aspects, with some researchers focusing on specific network setups such as mmWave networks and network slicing. Most literature considered a scenario of generic cellular network with slower moving UEs whilst vehicular UE scenarios with much higher moving speed and strict QoS requirements were less common. However, it is noteworthy that relatively little research evaluating the performance of an ML-based solution with conventional methods using the same set of information has been considered. Although it is essential to evaluate and compare performance of different ML-based solutions using standardized dataset and/or test environments [33], surprisingly little research considered this aspects of different HO algorithms with only [22], [30] and [31] implementing their proposed algorithms on a full-stack simulator (ns-2 and ns-3). Consequently, our work implements a DRL-based HO algorithm utilizing the same input parameters as existing cellular HO deployments, and builds a realistic simulation environment of a cellular vehicular network using the full-stack network ns-3 simulator.

III. SYSTEM MODEL AND HO PROBLEM FORMATION

This paper considers a cellular V2N network architecture that consists of vehicular UEs and BSs for HO management optimizations. The scenario can be reflected by simplifying Fig. 1 to only on-road vehicles and BSs with the V2N communication type.

For cellular networks, a UE reports a set of measurements of the BSs that it is able to perceive: RSRP and RSRQ. Presently, these two are the most important metrics in cellular networks to infer signal strength and quality from a BS to UE for HO decision making. According to the 3GPP standard [34], RSRP is the linear average over the power contributions of the resource elements of BS-specific reference signals within the measurement frequency bandwidth, while the RSRQ also includes channel interference and thermal noise. The relationship between RSRP and RSRQ is shown as (1).

$$RSRQ = N_{rb} \times \frac{RSRP}{RSSI} \quad (1)$$

where N_{rb} is the number of resource blocks of the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) carrier RSSI measurement bandwidth.

There are also several important events for HO decision-making based on UEs' measurement reports; specifically events A2 and A3 that are used for intra-RAT HO initialization [35]. When a serving BS becomes worse than the predefined threshold in terms of RSRP or RSRQ it satisfies the conditions to enter event A2. For event A3, the entering condition is when the RSRP or RSRQ from a neighbour BS becomes better than that of the serving BS by a predefined threshold. The opposite of such conditions indicate the leaving condition of both events.

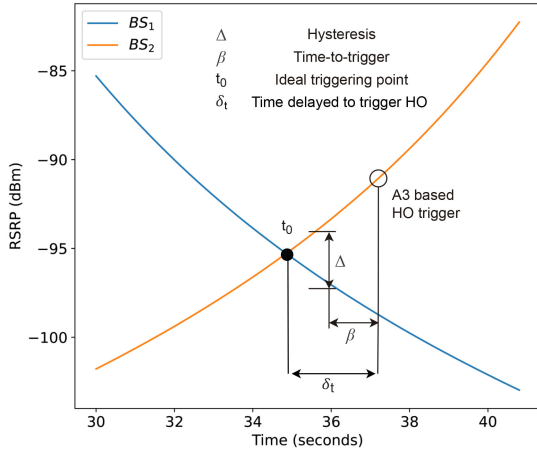


Fig. 2. Event A3-based HO trigger, in this case, t_0 specifies the optimal HO triggering instant, Δ and β represent the two parameters of the 3 RSRP HO algorithm, hysteresis and time-to-trigger, while δ_t indicates the time delay for this HO with respect to t_0 .

Two standardized HO algorithms for cellular networks are the event A2A4 and event A3 based HO algorithms for HO trigger and decision making [36]. Fig. 2 demonstrates the concept of A3 based HO. As also shown in the figure, t_0 represents the ideal HO triggering instant in the demonstrative case, while Δ and β represent the two HO parameters, hysteresis and time-to-trigger. A3 RSRP HO algorithm utilizes these two parameters to reduce unnecessary HOs and overcome the ping-pong effect [37]. Finally, δ_t represents the delayed time to trigger this HO, with respect to t_0 . During δ_t , the UE experiences suboptimal signal strength while waiting for the confirmation of HO trigger.

A. The HO Delay Cost

In cellular networks including LTE and the current 5 G implementation, hard HO is applied such that the connection between the UE and its serving BS is severed before the new connection is established [38]. The optimal BS selection during hard HO becomes critical as no useful data will be transmitted for the UE during the HO process. Following the HO trigger shown in Fig. 2, the delay to complete the hard HO process for the UE to switch connection from its serving BS to the target BS, is known as the HO delay time t_d . The accumulation of t_d will lead to a degradation effect on the average throughput of the UE. The cumulative t_d and total number of HO in a given trajectory is known as HO delay cost D_{HO} , and is the function of the number of HOs and HO delay time, according to [39]:

$$D_{HO} = N_{HO} \times t_d \quad (2)$$

For a given time period T of a moving UE, the normalized HO delay cost β_{HO} can then be derived:

$$\beta_{HO} = \min\left(\frac{D_{HO}}{T}, 1\right) \quad (3)$$

β_{HO} is expressed as a factor between 0 and 1, indicating the percentage of total time consumed on HO operations such as radio link switching between BSs, and as β_{HO} tends towards 1, thus indicating that the UE has spent almost the whole period T on HOs, and hence the average throughput derived from

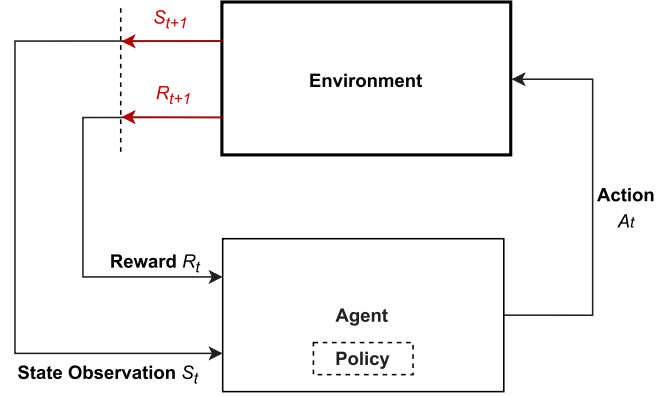


Fig. 3. The generic reinforcement learning framework, Fig. 3.1, [40].

the Shannon capacity formula when considering β_{HO} can be expressed as:

$$Throughput = BW \times \log_2(1 + SINR) \times (1 - \beta_{HO}) \quad (4)$$

where BW is the bandwidth and $SINR$ is the average signal-to-noise ratio. Therefore, maximizing *throughput* indicates maximizing BW and $SINR$ while minimizing β_{HO} . If constant BW and t_d are considered, $SINR$ and N_{HO} then play a vital role to increase the average throughput for a defined time period T . Therefore, given T and a know trajectory in a hard HO setup, the objective is to maximize average throughput by optimizing the HO decision making to increase the average $SINR$ while reducing N_{HO} for a constant BW and t_d . Based on this objective, it is acceptable to have more HOs to further increase $SINR$ as long as the resultant *Throughput* can also increase following (4).

IV. PROPOSED DRL SOLUTION

As discussed in Section III, this research focuses on developing a DRL solution to optimize the decision for BS selection after a HO is triggered, to maximize throughput along a given UE's trajectory. In this section, a DRL based intelligent HO decision making algorithm is proposed for cellular V2N networks. The background to RL and DRL is introduced in Section IV-A, followed by a detailed algorithm design in Section IV-B for the formulated problem.

A. The RL Framework and DRL

The concept of RL is shown in Fig. 3, where a decision-making agent takes actions based on its observation of the current environmental state, receiving the observation of the following state, and a numeric reward generated by the environment. After many rounds of training through trial and error, the agent develops an optimal policy to adopt in taking action within the given environment that maximizes the long-term accumulated reward. For interested readers, detailed information of RL can be found in a well-known book on the field [40].

Traditional RL algorithms are successful in the tabular environment, but such approaches become inefficient when representing a highly complex environment with very large state space or high-dimensional state inputs such as an outdoor

radio environment requiring optimal HO decision making. As a result, approximate solution methods were proposed such as linear value function approximation or policy gradient methods, etc. [41].

With the development of deep learning techniques using ANN, DRL algorithms utilizing the strong approximation abilities of an ANN were developed as a promising candidate for approximation-based RL methods as first proposed by Mnih *et al.* [42], which is known as Deep Q-network (DQN). DQN combines Q-learning, an RL algorithm that does not require a model of the environment [43], and ANN to achieve RL training in complex environments where traditional RL encounters difficulty. A technique named Experience Replay [44] is also introduced to speed up the learning process.

According to the Bellman equation [45], the essential mechanism for the RL training, the calculation of the Q value in DQN is shown in (5). In this equation, s, a represent the current state and current action taken respectively, s', a' stand for the resultant state and action taken in that state. Besides, r and $\gamma \in [0, 1]$ represent the received reward and the discount factor respectively, while θ is the ANN approximating the Q table. For DQN, the prediction Q value $Q(s, a; \theta)$ and updated Q value $Q(s, a)$ are calculated via the same ANN θ . While updating the current $Q(s, a; \theta)$ will change the value of future states (as parameters of θ are updated), it leads to potential instability during training that may result in non-convergence.

$$Q(s, a; \theta) = r + \gamma Q(s', \arg \max_{a'} Q(s', a'; \theta); \theta) \quad (5)$$

To overcome this problem, a new DRL algorithm Double deep Q-learning (DDQN) has been developed [46]. In DDQN, another ANN θ' is introduced as the target network, which calculates predicted value $Q(s, a)$, alongside the training network θ , which calculates current value $Q(s, a)$, changing (5) to (6). Throughout the training process of DDQN, only θ will be updated with each training iteration, while θ' only synchronizes periodically with θ by copying all parameters from θ to keep information updated. This design can greatly stabilize the DRL training and improves the chances of convergence.

$$Q(s, a; \theta) = r + \gamma Q(s', \arg \max_{a'} Q(s', a'; \theta); \theta') \quad (6)$$

B. Proposed DDQN-Based HO Algorithm

Because of DDQN's advantage over the original DQN, this research aims to further develop a DDQN-based HO algorithm and deploy it in a cellular network architecture and compare with the results of the presently implemented A3 RSRP HO algorithm. A centralized agent is hence designed that utilizes the DDQN HO algorithm, following the same HO decision making setup as the existing cellular network. The HO process derived from the current X2-based HO [47] can be found in Fig. 4.

In addition, the state observation and reward design must be based on the E-UTRAN measurement report entities from the UE, specifically, the mapped RSRP index values (integer values between 0 and 97) for the serving BS and neighbour BSs and corresponding BS IDs [34]. According to 3GPP, raw RSRP measurement values will first go through layer-3 filtering before being reported to the serving BS by a UE. The layer-3 filtering is shown in (7), where F_n and F_{n-1} are the current and old filtered

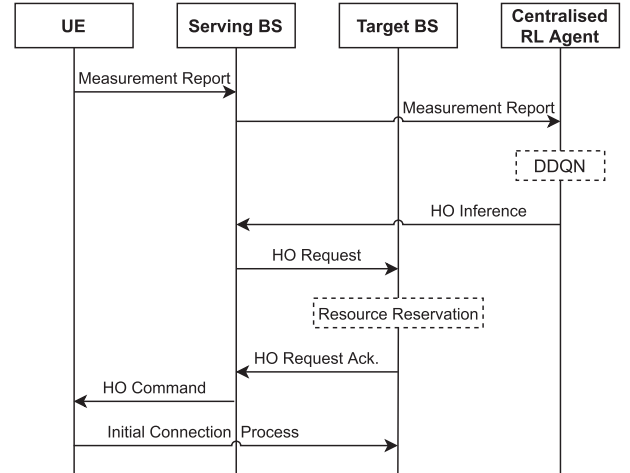


Fig. 4. The HO process with a centralized DDQN agent.

RSRP values to report, M_n is the latest received measurement result from the physical layer, and $a = \frac{1}{2^{\frac{k}{4}}}$ where k is the filter coefficient for corresponding measurement quantity received by the quantity configuration parameter.

$$F_n = (1 - a) \cdot F_{n-1} + a \cdot M_n \quad (7)$$

1) *State Space*: Mobility-based BS selection strategies have been intensively studied for vehicular networks, utilizing a UE's location and speed [14]. Measuring the exact location of a UE is impractical, while RSRP information can be utilized to estimate a UE's location [48], [49]. This provides a strong mapping between a geographical location in a defined area to a set of RSRP values from the BSs within the area, while more BSs available in the area may further improve the location estimation's precision. Therefore, this study considers the combination of the RSRP values measured by a UE from all surrounding BSs to represent the HO location-of-interest instead of the precise location of the UE (i.e. geo-coordinates of UE's location), and vehicle UEs are assumed to be of the same height to reinforce such representation.

Following the E-UTRAN configuration, converted RSRP indexes will be reported by a UE to its serving BS for HO inference [50] however, in contrast to the E-UTRAN configuration deployed in the current network, our study requires that all RSRP indexes of listed BSs within an area to be reported to form a state observation vector.

For a given local area containing n BSs and for a UE at position p , the RSRP measurements of all BSs \overline{RSRP}_p is given as:

$$\overline{RSRP}_p = \{rsrp_p^1, rsrp_p^2, \dots, rsrp_p^n\} \quad (8)$$

Hence, the state observation vector s_p is the combination of \overline{RSRP}_p and the serving BS ID $\{RSRP_p; BS_{serving}\}$.

However, instead of using the converted decimal value to represent the serving BS ID (as it may be confused with a RSRP index value), this information is designed to be represented via one-hot encoding [51]. For example, if the serving BS of a UE has a local ID of 2, with totally 5 BSs in the local area, then the serving BS ID after one-hot encoding becomes the vector $\overline{BS}_{serving} = \{0, 1, 0, 0, 0\}$. Therefore, s_p can then be formally

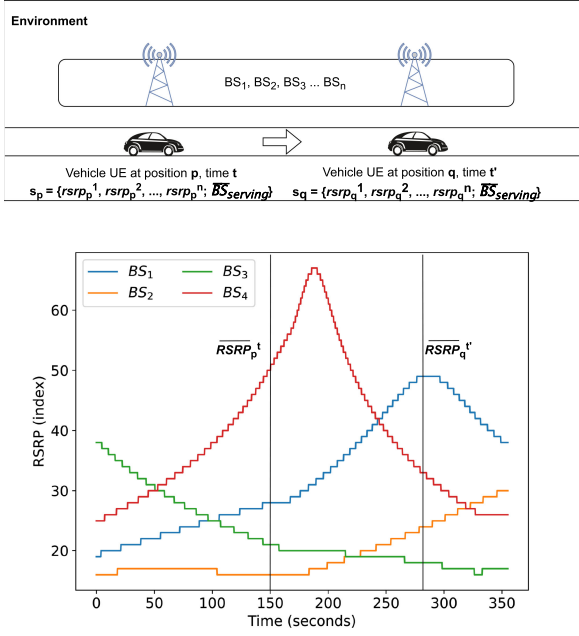


Fig. 5. A simplified environmental representation of n BSs and a vehicular UE in the upper graph, with the lower graph giving a demonstrative RSRP record of $n = 4$ BSs for a UE trajectory.

defined as:

$$s_p = \{rsrp_p^1, rsrp_p^2, \dots, rsrp_p^n; \overline{BS}_{serving}\} \quad (9)$$

and the complete state space S is then defined as the collection of all possible states. To observe the environment fully, the algorithm assumes that a state observation is periodically reported by a UE for both training and performance evaluation.

Fig. 5 demonstrates the environment-state relationship following the state design. At the top of the figure, a conceptual geographic environment is represented by n BSs and a moving vehicle UE between two locations, p and q , at time t and t' , to show the formulation of s_p . The lower part of the diagram shows the \overline{RSRP}_p formulation in a graphical example of an RSRP record assuming $n = 4$.

2) *Action Space*: An action a to take in each state can be defined as the BS to connect to the next state for a UE, i.e. all listed BSs in the local area including the serving BS (if the decision is not to HO). Therefore, the action space can be defined as a vector consisting of local BS IDs, $A = \{BS_0, BS_1, \dots, BS_n\}$. Note: a HO will only be executed if the action taken specifies a neighboring BS to connect to, while an action of serving BS ID indicates no HO required for the UE and to remain on the current the serving BS.

3) *Reward Design*: In line with the design of the state observation space, the reward design should also utilize the information from the E-UTRAN report only. A straightforward design uses the RSRP of the target BS specified by the action taken after state transition. In this work, the proposed reward design is to normalize this value with the highest reported value to emphasize the RSRP difference between the current BS choice and the local maximum RSRP, following the premise that higher signal strength correlates with a higher SINR and thus larger throughput.

A constant HO ‘‘punishment’’ is introduced in the reward design to enable the agent to consider the negative impact of performing HOs. Because the simplified approach of maximizing cumulated RSRP-based reward may lead to unwanted HOs resulting in a ‘‘ping-pong’’ effect due to noisy measurement reports causing pauses in data transmission as occur in the current hard HO implementation of LTE and 5 G cellular systems.

As shown in (10), $r(s_p, a_p; s_{p+1})$ is the reward gained after taking action a_p in state s_p and observing the next state s_{p+1} . $\max(\overline{RSRP}_{p+1})$ is the largest RSRP value from the measurement report in state s_{p+1} . In addition, $rsrp_{p+1}^a$ is the RSRP value of the target cell decided by the action taken, and C_{HO} is the introduced punishment on HO, a positive number with its specific value configuration depending on the environment.

$$r(s_p, a_p; s_{p+1}) = \begin{cases} \max(\overline{RSRP}_{p+1}) - rsrp_{p+1}^a - C_{HO}, & \text{if HO is triggered} \\ \max(\overline{RSRP}_{p+1}) - rsrp_{p+1}^a, & \text{otherwise} \end{cases} \quad (10)$$

4) *Experience Replay*: During Q-learning (hence DQN and DDQN) training, an experience consisting of the current state, action taken, reward, and the resultant state observations (s, a, r, s') is used once to update the value function parameters and then discarded. This is inefficient and can also cause instability as only the latest experience sample is being considered during parameter updates. Therefore, experience replay aims to improve the efficiency and reduce the potential instability by re-utilizing all experience samples. This is achieved by setting up a replay buffer \mathbb{B} that stores all experience tuples (s, a, r, s') until the maximum capacity of \mathbb{B} is reached, and then the oldest experience can be deleted. To update parameters, experiences in \mathbb{B} are uniformly sampled so that both current and previous experiences are considered during the algorithm’s training process. Using experience replay can significantly improve the performance of the algorithm [42].

5) *Algorithm Design*: The proposed DDQN algorithm implementation contains two phases: exploration (training) and exploitation (execution) phases. During the exploration phase, the algorithm is trained offline; the dataset is first collected, pre-processed, and then used to train the DDQN, without directly interacting with the environment. The dataset is collected along different UE trajectories and the ϵ -greedy strategy [52] is used to explore various actions in every state to update the ANN approximating the optimal HO value function. Algorithm 1 summarizes the DDQN algorithm’s training process during the exploration phase.

In the exploitation phase, the training process completes and the ANN update is terminated. The trained ANN is then used to emulate the optimal HO policy to take HO actions, with the ϵ -greedy strategy also disabled (i.e. $\epsilon = 0$). In a direct comparison with the A3 RSRP baseline, event A3 is also used in this phase to trigger the HO inference. Fig. 6 demonstrates the algorithm in the exploitation phase. It is noteworthy that the new data generated during the exploitation phase can also be stored and processed to update the DDQN algorithm to learn the underlying patterns

Algorithm 1: The Proposed DDQN Algorithm.**Initialization:**

θ - Training Q network; θ' - Target Q network;
 \mathbb{B} - The reply buffer (empty);
 N_r - Replay buffer capacity;
 N_b - Training mini-batch size;
 N_f - Step size to update target network;
 γ - The discount factor;
 ϵ - The probably to take a random action.

For $episode \leftarrow 1, 2, \dots, N_{episode}$:

Set the initial state s_1 ;

For $i \leftarrow 1, 2, \dots, end\ of\ trajectory$:

Observe s_i ;

$a_i \leftarrow \begin{cases} \text{a random action, with probability } \epsilon \\ \text{argmax}_a Q(s, a; \theta), \text{ with probability } 1 - \epsilon \end{cases}$

Execute a_i

Observe s'_i and r_i

Store (s_i, a_i, r_i, s'_i) in \mathbb{B}

If N_r is reached:

Delete the oldest experience sample

If \mathbb{B} has more than N_b samples:

Sample a mini-batch of data from \mathbb{B} ;

Construct target value tuple:

$$y_i \leftarrow \begin{cases} r_i, & \text{if } s' \text{ is the terminal state} \\ r_i + \gamma Q(s', \text{argmax}_{a'} Q(s', a'; \theta); \theta'), & \text{else} \end{cases}$$

Do gradient descent with loss $\|y_i - Q(s, a; \theta)\|^2$

If $\text{mod}(\text{step}, N_f) = 0$:

$\theta' \leftarrow \theta$

(Replace the parameters of θ' by those of θ);

Update parameters of θ

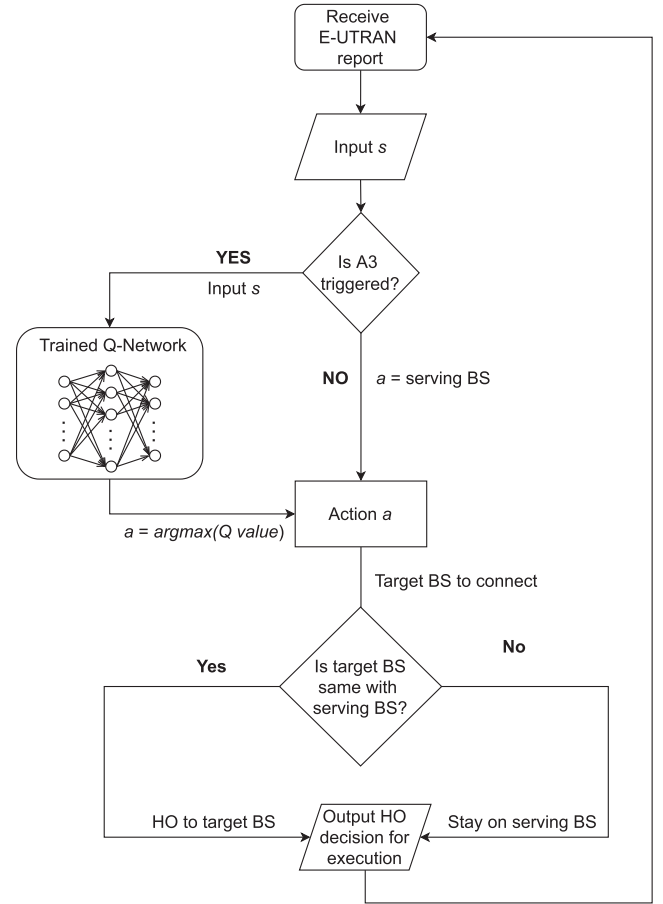


Fig. 6. A flowchart of the proposed DDQN HO algorithm in execution phase.

in the new datasets. However, the design and implementation of such future updates to the algorithm is beyond the scope of this work.

V. SIMULATION SETUP AND PERFORMANCE EVALUATION

A. Scenario Design and Simulation Setup

Due to the complexity of real-world UE handover data collection, the dataset for this research was generated using the ns-3 network simulator [53]. The ns-3 simulator is an open-source, discrete-event full-stack simulator, that allows tracing internal events with flexible configurations and supports multiple communication technologies. The ns-3 official standard-compliant LTE module LENA [54] was chosen to configure an LTE cellular network scenario to investigate a cellular V2N network. This approach was adopted as the 5G and LTE network HO mechanisms are very similar and the 5G-LENA [55] (the 5G version of the LENA module) is still under development at present.

The need for the simulation data is to train and evaluate the DDQN HO algorithm in an environment that is a close analogue to a real-world network. The simulation scenario was a 2×2 km local area in the city center of Glasgow, U.K.. For the implementation of a realistic mobility simulation for vehicles,

the routes mobility model [56] was selected, which utilized the Google Maps' directions API and the way-point mobility model provided by ns-3. By specifying the target area and a trajectory's start and end points, the routes mobility model enabled real-world trajectory generation to be directly used by ns-3. For BSs, the location references were taken via the Cell Mapper website [57], which recorded the approximate real-world BS deployment locations based on measured data from participants. For the scenario in this paper, eight BSs from the U.K. mobile operator Vodafone were chosen. Fig. 7 demonstrates the scenario setups, including the environment setup, BS locations, and the network architecture on the map of the selected area. Each BS is connected to the core network (i.e. the Mobility Management Entity and Serving Gateway) via the S1 interface, and to other BSs via the X2 interface. A vehicle trajectory is also included in Fig. 7, whose current serving BS is the red BS, and will HO to the green BS following its route. The red and green circles are simplified indications that the signal strength from the corresponding BSs are the same, while the intersection of the circles represents a HO location-of-interests between the two BSs.

After establishing the scenario, the detailed network configuration is listed in Table I. An isotropic antenna model is used at the current stage, demonstrating the same HO strategy while simplifying the scenario, and hence the state space to be considered by the DDQN HO algorithm. Other network configurations,

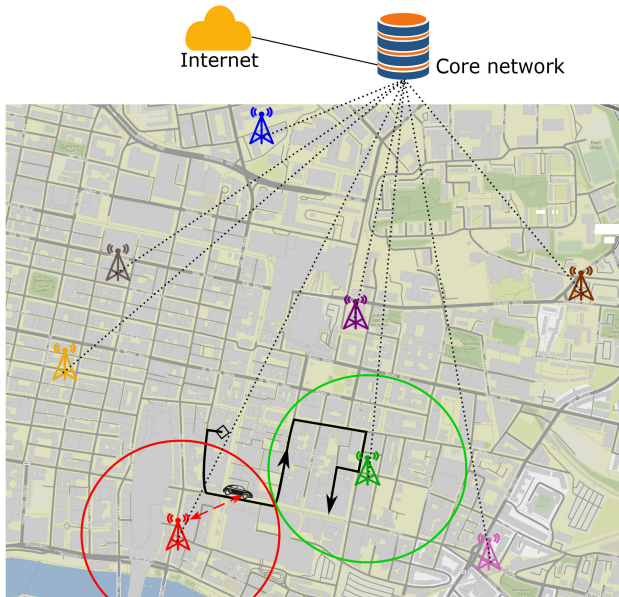


Fig. 7. Glasgow city center scenario setup with 8 BSs marked in different colors, and 1 demonstrative trajectory of an on-road vehicle. The network setup and a potential HO along the trajectory is also demonstrated.

TABLE I
SIMULATION CONFIGURATION

Setup properties	Parameters and values
LTE Network	8 BS sites, UK operator Vodafone [57]
Vehicle mobility model	Waypoint mobility model generated by Routes mobility model [56]
Antenna model	Isotropic antenna model
Pathloss model	3GPP Uma Macro, shadowing disabled
Scheduling algorithm	Proportional Fair
BS transmission power	40 dBm
Carrier frequency	2115 MHz, downlink only
Noise Figure	UE: 5 dB BS: 9 dB
HO algorithm	A3 RSRP (baseline): hysteresis - 3 dB, time-to-trigger - 320 ms; DDQN (proposed): HO cost - 3 (equivalent to dB)
Application setup	UDP, downlink only packet interval: 20 ms (50 packets / sec) individual packet size: 1024 bits

such as the carrier frequency, are set according to the 3GPP standards. However, these setups did not consider the effect of fading except for pathloss. While it is recognized that multi-path fading and Doppler spread can introduce significant variations during propagation, particularly in a vehicular networks of high UE mobility, it is important to introduce such effects to the simulation to emulate the real-world situation. Therefore, trace-based fading was generated via the MATLAB script provided in the LENA module [58], using the the fading model derived from that of [59], and loaded into ns-3. The “Vehicular” mode with nodes’ moving speed of 60 km/h was chosen to generate the fading traces and hence introduced small scale fading within the simulation scenario. This 60 km/h speed was chosen as a worse case scenario for Doppler spread and thus fading, while being in line with the maximum moving speed in the area according to the area’s speed limit to generate more realistic RSRP values.

TABLE II
ANN CONFIGURATION

Hyperparameters	Values
Hidden layers, Neuron size	3, $256 \times 128 \times 64$
Hidden layer activation function	GELUs [60]
Output layer activation function	Linear
Target network update frequency	Every 100 episodes
Optimizer	AdamW [61]
Learning rate	0.001
Learning rate decay ratio	0.98
Replay buffer size	100,000
Mini batch size	1,024

To train and evaluate the performance of the proposed DDQN HO algorithm, 18 trajectories covering the major routes across the selected area of Glasgow city center were generated via the routes mobility model [56] for the ns-3 simulation, to form a training dataset for the exploration phase of the DDQN algorithm. The maximum simulation time is set to 800 seconds (13.3 minutes) so that all vehicles can have sufficient time to complete their trajectories. For performance evaluation, overall the same trajectories were used for the exploitation phase of the algorithm, while slight modifications were applied to the detailed way-points representing the trajectories to also test the algorithm’s generalization ability. Performance evaluation was implemented in an online manner using the ns-3 simulator directly and the connection between a UE and BS pair was terminated shortly after the UE reached its end point to avoid redundant data collection.

For the DDQN setup, a fully connected feed-forward ANN with 3 hidden layers was used, and the detailed hyperparameter configurations are summarized in Table II. The Gaussian Error Linear Units (GELUs) was used as the activation function of hidden layers [60], while the optimization algorithm was set to the “Adam with decoupled weight decay” (AdamW) optimizer [61], for their better performances in general in ANN training compared with their predecessors (other linear unit activation functions and the original Adam optimizer). The initial learning rate was set to 0.001, and an exponential learning rate decay was configured with the decay ratio set to 0.98 to stabilize the ANN training convergence. The replay buffer was set to have a maximum capacity of 100,000 observation samples, and a mini batch sampling size of 1024 was configured for ANN training. After completing the exploration phase, the trained ANN was then turned into exploitation phase and deployed to directly interact with the ns-3 simulator using the ns3-ai module [62] for direct online performance evaluation.

1) *Data Collection and the Evaluation Metrics:* To train and evaluate the DDQN HO algorithm, RSRP data of all BSs need to be collected, following the design in Section IV-B. The data is collected directly from the ns-3 LENA module’s radio resource control (RRC) layer of the vehicle UEs where E-UTRAN measurements are performed and reported [63]. An RSRP index record of all BSs is generated for each trajectory to form a dataset to train the algorithm.

To evaluate the performance of the proposed HO algorithm, several metrics are chosen to compare performances between the proposed algorithm and A3 baseline. The time delay δ_t as expressed in Fig. 2 indicates when a HO decision is made with

respect to the optimal time instant. Therefore, the gain in time delay \mathbb{G}_{δ_t} becomes a clear metric to measure how much faster the DDQN HO algorithm makes the HO decision towards the optimal HO instant.

According to (4), it is essential to maximize the $SINR$ while minimizing β_{HO} in order to maximize throughput. Given a constant t_d in (2), the objective of maximizing throughput becomes maximizing the $SINR$ while minimizing the N_{HO} . Therefore, the $SINR$ traces of serving BSs are collected for signal quality comparison between the proposed DDQN HO algorithm and the A3 RSRP baseline during the HO periods-of-interests, while the number of HOs N_{HO} is also collected for all trajectories. The $SINR$ traces are collected by the ns-3 LENA module and presented as raw linear-scale values [64], and are converted to decibels for performance comparisons. This $SINR$ metric is defined as the normalized $SINR$ value (gain in $SINR$) \mathbb{G}_{SINR} using the DDQN HO algorithm with respect to A3 RSRP baseline.

In addition, the Packet Data Convergence Protocol (PDCP) packet loss is also used as the metric because the PDCP layer in the LTE protocol stack is responsible for the transfer of data on the control/user planes [65] hence is a clear metric to evaluate throughput performance. A smaller packet loss during a HO period-of-interests suggests higher throughput during that period. The PDCP packet loss is calculated by the ns-3 simulator with the statistics recorded by the LENA module [54] while the calculation of packet loss follows ns-3's data plane error model utilizing the link-to-system technique and block error rate mapping [66].

The analysis of \mathbb{G}_{δ_t} and \mathbb{G}_{SINR} for all 18 trajectories cannot be presented clearly in a graphical manner in the simulation scenario due to a number of HOs occurring. Therefore, for the remainder of this section, the results from 1 demonstrative trajectory selected from the full simulation are presented in Section V-B, to present a detailed graphical demonstration of metric analysis for a single HO instance. Then, the statistical results for the whole simulation scenario (i.e. all 18 trajectories including the demonstrative one) is then presented in Section V-C with some edge-case discussions. Excluding the reference results introduced in the remainder of this section, all presented results are collected via the same round of training and evaluation.

B. Result Analysis of One HO Case

Following the scenario and simulation setup, this subsection presents an exemplar performance analysis case of a single trajectory within the 18 trajectories in the simulation scenario. The visualization for this trajectory's geographical information is shown in Fig. 7, and the performance of the proposed DDQN HO algorithm was evaluated against the A3 RSRP baseline with all other network settings kept constant. Different from the learning-based algorithms, the A3 RSRP HO algorithm triggers a HO based on event A3, when a neighbor BS's RSRP becomes larger than that of the serving BS by a predefined offset value. The two parameters of this HO algorithm, hysteresis and time-to-trigger, are used to avoid the "ping-pong" effect, and Fig. 2 demonstrates this HO decision-making. To permit easy comparison in this performance evaluation, the A3 RSRP

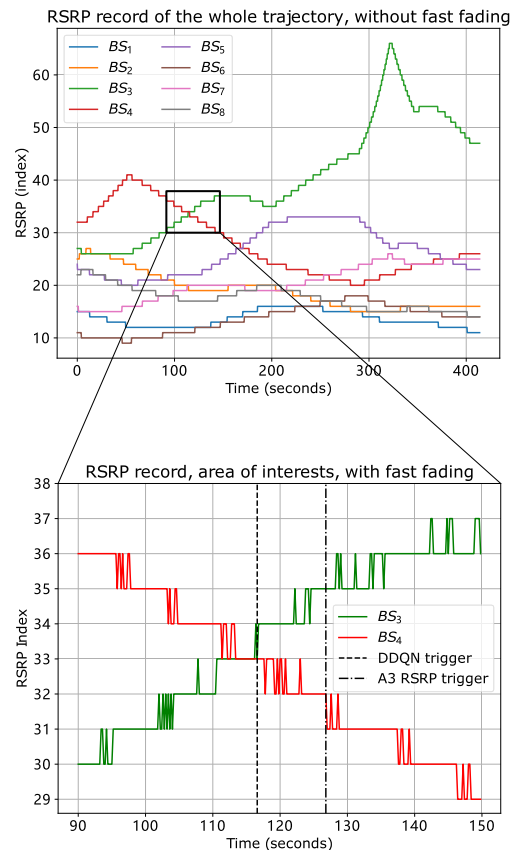


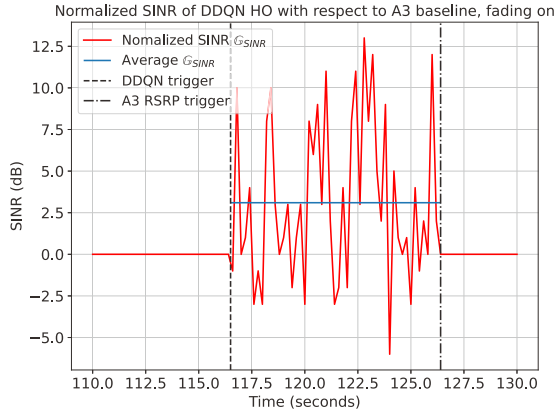
Fig. 8. Case study: The corresponding RSRP record for the selected trajectory. The top half of the figure includes the RSRP record for all 8 BSs (without fading for a clear visual presentation). The bottom half of the figure is the zoomed-in period of interests when a HO was triggered, with the triggering instants of DDQN and A3 baseline plotted as vertical lines (fading enabled as it was in the simulation).

baseline used the same parameter configuration as current cellular networks as stated in Table I.

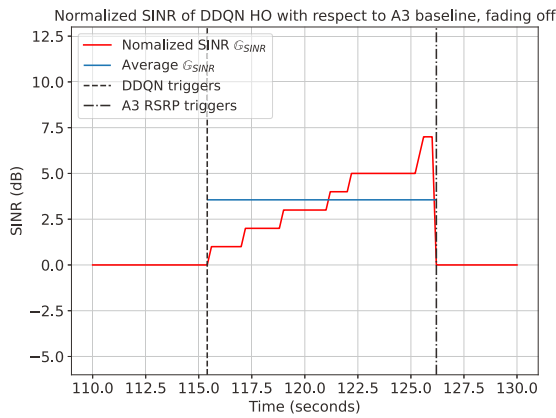
The RSRP record for all BSs throughout this trajectory is shown in the top half of Fig. 8. For improved visual clarity, this part of Fig. 8 is presented without fading. Accordingly, 1 HO should be triggered for this trajectory, that also corresponds with the simulation results. This HO period-of-interests is highlighted by a black rectangle in the top half of Fig. 8, and the zoomed-in of this area is presented in the bottom half of the figure. The bottom half of Fig. 8 is presented with fading enabled to reflect the actual simulation.

It is clear that the optimal triggering instant of this HO is around 116 seconds after the simulation starts (0 s) based on Fig. 8. This is when the RSRP of BS_3 becomes greater than that of BS_4 , the initial serving BS, and remains so until the end of the simulation. Due to the 3 dB hysteresis and time-to-trigger, the A3 RSRP baseline has to wait until both thresholds are satisfied to trigger the HO at 126.72 seconds, while the DDQN-based HO algorithm triggers the HO as soon as the optimal triggering point is reached (116.68 seconds). As a result, the DDQN-based HO improves the triggering instant by having a $\mathbb{G}_{\delta_t} = 10.04$ seconds compared to A3 RSRP baseline.

After presenting \mathbb{G}_{δ_t} , Fig. 9 includes the normalized serving BS $SINR$ (\mathbb{G}_{SINR}) using the proposed DDQN HO with respect



(a)



(b)

Fig. 9. The SINR gain \mathbb{G}_{SINR} during the HO period-of-interests of the DDQN HO with respect to the A3 RSRP baseline for the trajectory. Both results were produced by the same DDQN trained using data with fading. (a) Original result with fading enabled (b) Reference result with fading disabled.

to using the A3 baseline to show the SINR gain. When not in the HO period, the normalized SINR = 0 as the experience SINR using both HO algorithms are the same, while the differences are clearly evident during the HO period-of-interests. As the simulation setup includes fading that introduces fluctuations in Fig. 9 a, a reference curve of normalized SINR is also included in Fig. 9 b using the same simulation configuration but disabling fading, to give a clearer visual presentation on \mathbb{G}_{SINR} . Note that the same DDQN trained using the dataset generated in the scenario with fading is deployed to produce results in this reference, and in both sub-figures, the average \mathbb{G}_{SINR} is calculated and plotted as well. Also, by disabling fading, the RSRP record becomes smooth, leading to some state changes and hence different HO triggering instances when using the two HO algorithms. However, it is clear that the DDQN's HO start at 115.8 seconds is still an optimal trigger referring to Fig. 8.

The results shown in Fig. 9 demonstrate an obvious HO performance gain with respect to \mathbb{G}_{SINR} during the HO period-of-interests, with a maximum SINR gain of over 12.5 dB compared with the A3 RSRP baseline. According to the reference curve without fading, an average \mathbb{G}_{SINR} of 3.51 dB is obtained by the DDQN HO algorithm during the HO period-of-interests.

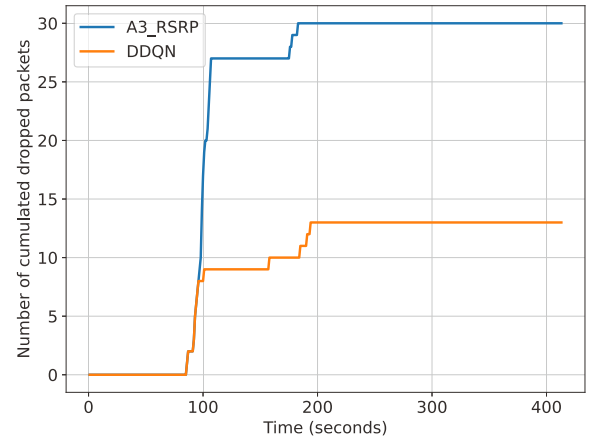


Fig. 10. Cumulated packet loss comparison between the proposed DDQN HO algorithm and A3 RSRP baseline for the demonstrative trajectory.

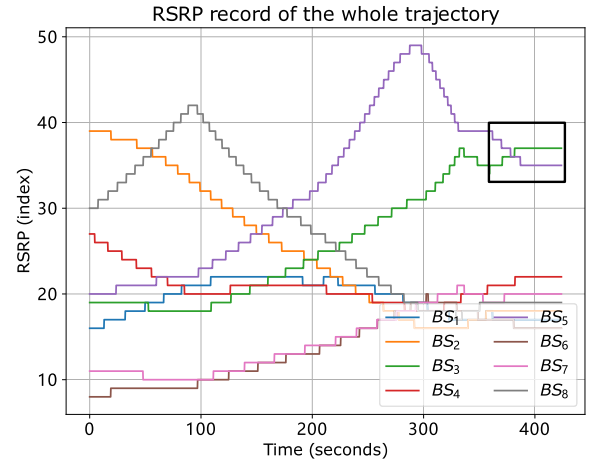


Fig. 11. The edge case that leads to DDQN's 1 more HO trigger, marked in the black rectangle. Fading is disabled for a clear visual presentation.

The improvement of HO performance is also reflected in packet loss, as shown in Fig. 10, where the PDCCP packet loss is counted for this HO for A3 RSRP baseline and the DDQN algorithm respectively. For the HO period-of-interests between 116 and 127 seconds shown in Fig. 9, there are overall 625 packets sent. When using the A3 RSRP baseline, the UE experiences a packet loss of 30 packets due to waiting for HO to trigger and the connection drop during HO, while that number for DDQN is 13 packets, indicating a 56.7% better performance with respect to packet loss and thus throughput gain.

C. Results Analysis for All Trajectories

In total, there are 46 HOs using the A3 RSRP baseline, and 47 HOs using the proposed DDQN HO algorithms for all 18 trajectories throughout the simulation over the scenario. The edge case for the additional HO happened when using DDQN is that the RSRP difference between the best neighboring BS and the serving BS, although stable, could not satisfy the A3 RSRP baseline's 3 dB hysteresis, as shown in the black rectangle area in Fig. 11. In contrast, the DDQN algorithm learned the RSRP features along the whole trajectory and performed an additional HO to improve optimal signal strength, demonstrating

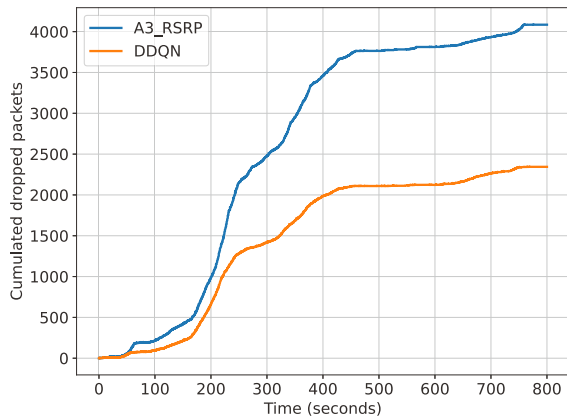


Fig. 12. Cumulated packet loss comparison for all trajectories between the A3 RSRP baseline and the proposed DDQN HO algorithm.

the superiority of having the knowledge of the environment. Excepting this one edge case, both HO algorithms have the same number of HOs, indicating that DDQN manages the ping-pong effect as effectively as A3 RSRP baseline does. However, instead of using predefined parameters (hysteresis and time-to-trigger) as A3 RSRP HO, the proposed DDQN algorithm learns with the HO cost punishment of unnecessary HOs, while being able to avoid staying connected to a BS with suboptimal signal strength until the static conditions are satisfied as the A3 baseline.

Throughout the whole evaluation, using the DDQN-based HO algorithm results in 1741 fewer accumulated lost packets over the 47 HOs compared to the A3 RSRP baseline, showing an improvement in cumulated packet loss by 42.62%, as shown in Fig. 12. On average, the corresponding numbers are 35.37 fewer packet loss and 25.72% packet loss reduction for each HO. As the HO mechanism of 5 G is very similar to that of the current LTE networks, the evaluation results suggest a potential for significant performance gain in a dense 5 G mmWave network that will have many more HOs due to the short range of mmWave beams. The 25.72% fewer packet losses per HO in the proposed simulation scenario will result in more significant packet loss reduction in a dense 5 G network.

If the TCP protocol is used instead of UDP, the 25.72% packet loss reduction means that many packet re-transmissions are saved, leading to a further improvement of 51.44% more packets transferred compared with the A3 baseline.

As demonstrated in Section V-B, using DDQN-based HO algorithm makes the HO triggering instant closer to an optimal triggering instant for a HO, compared to the A3 RSRP baseline. For all HOs happened in the simulation scenario, the DDQN-based HO algorithm may improve a HO trigger instant from 4.52 to 25.64 seconds towards the optimal HO instant. The average time improvement to HO triggering instants in the scenario is 11.56 seconds per HO, suggesting a significant performance gain. Both metrics are summarized in Table III. Note that the results of these metrics also depend on the UE's moving speed. For a given trajectory, the slower a UE moves along it, the longer it takes to satisfy the A3 RSRP baseline's hysteresis and time-to-trigger, resulting in a larger δ_t for the baseline and thus higher \mathbb{G}_{δ_t} and \mathbb{G}_{SINR} .

TABLE III
STATISTICAL RESULTS FOR METRICS \mathbb{G}_{δ_t} AND \mathbb{G}_{SINR} FOR ALL 18 TRAJECTORIES

Metric / Value	Average	Maximum	Minimum
Time delay gain: \mathbb{G}_{δ_t} (seconds)	11.56	25.64	4.52
SINR gain: \mathbb{G}_{SINR} (dB)	3.03	3.41	1.84

VI. DISCUSSION

After presenting the performance evaluation results, this section discusses the proposed algorithm's implementation aspects, followed by a qualitative comparison of the proposed algorithm with some state-of-the-art research.

A. Discussion on Deployment Aspects

As this research looks at the deployment aspect of an intelligent HO algorithm, it is essential to consider the implementation cost alongside the performance gain. For the proposed DDQN algorithm, the fundamental requirement is that RSRP measurement of all listed BSs in a local area to be reported for HO decision making. Because the related configuration options are available in the current E-UTRAN setup [63], implementing the algorithm locally requires only a software patch.

Compared to the current HO solution (A3 RSRP), the proposed algorithm requires proportionally larger data input to operate. In the proposed scenario of 8 BSs in this paper, the A3 RSRP baseline usually gets a report of two entries, the RSRP and IDs of the serving BS and the best neighboring BS. In contrast, the DDQN requires the reports for all 8 listed BSs to operate, resulting in quad times the data input. Linear complexity yields a moderate cost with respect to computation and storage. Furthermore, the E-UTRAN measurement reporting uses the dedicated control channel [67], and the increase in data transmission will result in a larger communication overhead in the control channel. With the deployment of 5 G, ultra-dense network deployment is also scheduled and in progress. The future large number of BSs also adds to this potential issue of linear complexity and communication overhead for the proposed DDQN, the effect of which demands further research and quantitative analysis. Similar issues in resource allocation has previously attracted attention and research to reduce input state space of a learning-based algorithm to reduce communication overhead [68]; which may also provide a valuable exploration for learning-based HO algorithms.

As the proposed DDQN algorithm considers the scenario of vehicular networks and the vehicular UE mobility, a major feature of vehicular networks compared to generic cellular networks may also impact the algorithm's performance and needs discussions. Fading in wireless channels can be broadly classified as large-scale and small-scale fading, and Doppler spread from the small-scale fading has a positive correlation to the UEs' moving speed [69]. Consequently the fading setup in Section V-A, and the results in Figs. 8 and 9 indicate that the same DDQN is able to produce the optimal HO decision in both the theoretical worst fading situation and in the ideal situation where only pathloss is considered. The results in such extreme cases suggests that the proposed DDQN holds the resilience against the speed dependent Doppler spread within the

TABLE IV
COMPARISON BETWEEN THE PROPOSED DDQN HO ALGORITHM AND SOME STATE-OF-THE-ART

Algorithm	A3 RSRP [70]	Proposed DDQN HO	Q-learning-based HO parameter optimization [19]	Multiuser DRL [21]	FL proactive HO trigger [28]	Two-tier proactive HO optimization [31]
Methodology	Static control scheme	RL	RL	Unsupervised learning, RL	Supervised learning	Supervised learning, RL
Key idea	Fixed parameter-based HO trigger to strongest BS	HO decision optimization	Parameter optimization for A3-based HO	Mobility-based UE clustering, cluster-level optimal HO policies	SNR prediction for HO trigger	RSSI prediction for HO trigger, RL-based HO decision optimization
Learning technique setups	/	Deep Q-learning	Traditional Q-learning	K-means clustering, A3C policy gradient	Feedforward ANN-based prediction, Federated Learning	RNN (prediction), HMM (HO decision)
Input parameters	RSRP	RSRP	Throughput, Packet delay, HO frequency	UE mobility, RSRQ, Throughput, HO frequency	SNR	RSSI
Execution agent	Core network	Core network	Core network	Core network	Distributed	UE (prediction), Core network (HO decision)
Impact on network architecture	No	Low	Moderate	High	Very high	High
Communication overhead	Very low	Moderate	Low	Moderate	High	Moderate

considered UE speed range however, it may be necessary to train the algorithm using data generated at the speed limit to provide the DDQN knowledge of the worst fading case. Additionally, as the proposed DDQN is based on the model-free Q-learning technique that does not require a model of the environment, and trained using data generated via known trajectories, the state transition variation caused by a UE's moving at different speeds along a given trajectory will not influence the overall decision making.

For the real-world deployment aspect, the HO algorithm is designed to work on a defined local area, leading to the requirement of dedicated local area specifications and individual ANNs to be trained and stored locally for HO decision making that are heavily dependant on mobile operators' physical network deployments. Finally, as for all learning-based solutions, the DDQN-based HO algorithm requires a significant amount of data to form an effective training dataset. Therefore, the related data collection and processing for individual local area that influence the scalability of the algorithm, remains an important aspect for the algorithm's implementation.

In summary, the proposed DDQN-based HO algorithm, after learning from collected data of the environment, can improve the performance of handover significantly compared to the baseline. Importantly, it only requires a small change to the existing network architecture setup, therefore the implementation cost should be small requiring only a software patch. However, other important deployment aspects need to be considered, including state space reduction, actual network deployment, and data collection and processing.

B. Comparison With the State-of-the-Art

After discussing the deployment aspects, comparison of the proposed DDQN HO algorithm with the A3 RSRP baseline, and some of the state-of-the-art research is presented. The contents of this comparison include the methodologies and key design concepts, followed by the input parameters required to operate the algorithms, and what level within the network the algorithms execute. Impact on the cellular network architecture (LTE and 5 G) is also compared on a qualitative level based on the proposed

algorithm and system architectures in the original literature. Finally, the communication overhead using the selected algorithms is compared by analyzing the type and amount of information required to be transferred on wireless channels for operation. The full comparison is presented in Table IV.

Note that with the exception of the A3 RSRP baseline, all state-of-the-art works are ML-based solutions, and are selected based on their ML application types for HO optimization as discussed in Section II. Algorithm designs that considered multiple optimization objectives such as [22] (joint optimization of HO and radio resource management), have not been selected in this comparison in order to focus on HO optimization.

According to Table IV, ML-based HO parameter optimization [19] and HO decision making [21] exploit information from various input parameters for optimization, and require some network adjustment for the HO algorithms to operate. In contrast, predictive HO triggering considers one type of input parameter for accurate predictions, while the prediction is usually performed at UE level. However, using additional training setups (i.e. the FL setup in [28]) will require major change to the network architecture for deployment, while also demanding that the ANN model is transmitted through the wireless channel among training participants. Similarly, the two-tier design in [31] also requires the RNN model to be transmitted via wireless channels and may lead to a large communication overhead, while its HMM-based centralized HO decision require less modification to the network architecture compared to [28]. In comparison, the proposed DDQN HO design, although aimed at HO decision making, requires only 1 type of input parameter while maintaining a low impact on the existing network architectures (utilizing existing deployment design and input parameters). Therefore, it is a lightweight upgrade taking real-world deployment aspects into consideration while delivering near-optimal decisions.

VII. CONCLUSION

Vehicles have higher moving speeds, which leads to reduced connection time between a vehicle UE and a BS and more HOs as a result, especially for 5 G enabled dense V2N networks currently under deployment, where small-cell BSs with reduced

cover range are deployed. ML-based HO optimization research has utilized various input parameters and enabling technologies, while relatively less attention was paid to implementation in unified test environments as well as real-world deployment aspects. In this paper, a double deep Q-network-based HO algorithm is proposed and evaluated using the ns-3 full-stack network simulator with the LENA module and a realistic simulation setup. The results analysis from 47 HOs throughout the simulation show a 44% reduction in accumulated packet loss compared to the A3 RSRP HO algorithm baseline, and an average 11.56-second improvement of the optimal HO triggering instant in the simulation scenario. The proposed algorithm also aimed to utilize the existing cellular network configuration with only minor additional information requirements; reporting all neighboring BS's RSRP instead of only those satisfying predefined conditions (as the A3 RSRP baseline). This makes the algorithm implementable via a software patch, while some small modifications in the network configuration greatly reduced the implementation cost.

Future research extending this work will implement and evaluate the proposed algorithm using ns-3's latest standardised cellular module, 5G-LENA [55], which has become publically available but is still under active development requiring updating with the 5G HO interfaces. Other research optimizing learning-based algorithms, such as state space reduction, is also relevant to this research to reduce foreseeable communication overheads in the control plane. In addition, extending the HO algorithm to also combine inter-slice HO in network slice is also viable to fully exploit the benefit of new enabling technologies.

REFERENCES

- [1] L. Figueiredo, I. Jesus, J. A. T. Machado, J. R. Ferreira, and J. L. Martins de Carvalho, "Towards the development of intelligent transportation systems," in *Proc. IEEE Intell. Transp. Syst.*, 2001, pp. 1206–1211. [Online]. Available: <https://ieeexplore.ieee.org/document/948835/>
- [2] J. Misener, "Smart transportation," Aug. 25, 2020. [Online]. Available: <https://www.qualcomm.com/media/documents/files/smart-transportation-presentation.pdf>
- [3] P. K. Singh, S. K. Nandi, and S. Nandi, "A tutorial survey on vehicular communication state of the art, and future research directions," *Veh. Commun.*, vol. 18, Aug. 2019, Art. no. 100164. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2214209618300901>
- [4] 3GPP, "Study on LTE-Based V2X services," 3rd Generation Partnership Project (3GPP), Technical report (TR) 36.885, version 14.0.0, Jun. 2016. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2934>
- [5] G. Miao, J. Zander, K. W. Sung, and S. Ben Slimane, *Fundamentals of Mobile Data Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [6] 5GAA, "5G V2X: The automotive use-case for 5G," 2017. [Online]. Available: https://www.3gpp.org/ftp/information/presentations/Presentations_2017/A4Conf010_Dino%20Flore_5GAA_v1.pdf
- [7] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges," *Wireless Netw.*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015. [Online]. Available: <http://link.springer.com/10.1007/s11276-015-0942-z>
- [8] L. Zhang, A. Farhang, G. Feng, and O. Onireti, *Radio Access Network Slicing and Virtualization for 5G Vertical Industries*. Hoboken, NJ, USA: Wiley, 2020.
- [9] L. Zhang, A. Farhang, G. Feng, and O. Onireti, *Radio Access Network Slicing and Virtualization for 5G Vertical Industries*. Hoboken, NJ, USA: Wiley, 2020, pp. 167–208.
- [10] K. Tan, D. Bremner, J. L. Kerneç, and M. Imran, "Federated machine learning in vehicular networks: A summary of recent applications," in *Proc. Int. Conf. U.K.-China Emerg. Technol.*, 2020, pp. 1–4.
- [11] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*.
- [12] L. Liang, H. Ye, G. Yu, and G. Y. Li, "Deep-learning-based wireless resource allocation with application to vehicular networks," *Proc. IEEE*, vol. 108, no. 2, pp. 341–356, Feb. 2020.
- [13] M. S. Mollel *et al.*, "A survey of machine learning applications to handover management in 5G and beyond," *IEEE Access*, vol. 9, pp. 45770–45802, 2021.
- [14] N. Aljeri and A. Boukerche, "Mobility management in 5G-enabled vehicular networks: Models, protocols, and classification," *ACM Comput. Surv.*, vol. 53, no. 5, Sep. 2021, Art. no. 92.
- [15] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surv. Tut.*, vol. 21, no. 4, pp. 3072–3108, Oct.–Dec. 2019.
- [16] V. Yajnanarayana, H. Rydén, and L. Hévízi, "5G handover using reinforcement learning," in *Proc. IEEE 3rd 5G World Forum*, 2020, pp. 349–354.
- [17] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, Jul. 2011.
- [18] H. Tabassum, M. Salehi, and E. Hossain, "Mobility-aware analysis of 5G and B5G cellular networks: A tutorial," *CoRR*, 2018. [Online]. Available: <http://arxiv.org/abs/1805.02719>
- [19] A. Abdelmohsen, M. Abdelwahab, M. Adel, M. S. Darweesh, and H. Mostafa, "LTE handover parameters optimization using Q-learning technique," in *Proc. IEEE 61st Int. Midwest Symp. Circuits Syst.*, 2018, pp. 194–197.
- [20] T. Goyal and S. Kaushal, "Handover optimization scheme for LTE-Advance networks based on AHP-TOPSIS and Q-learning," *Comput. Commun.*, vol. 133, pp. 67–76, Jan. 2019. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0140366417312549>
- [21] Z. Wang, L. Li, Y. Xu, H. Tian, and S. Cui, "Handover control in wireless systems via asynchronous multiuser deep reinforcement learning," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4296–4307, Dec. 2018.
- [22] S. H. Srikantamurthy and A. Baumgartner, "A novel unified handover algorithm for LTE-A," in *Proc. 17th Int. Conf. Netw. Serv. Manage.*, 2021, pp. 407–411.
- [23] M. S. Mollel *et al.*, "Intelligent handover decision scheme using double deep reinforcement learning," *Phys. Commun.*, vol. 42, Oct. 2020, Art. no. 101133. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S187449072030210X>
- [24] D. Guo, L. Tang, X. Zhang, and Y.-C. Liang, "Joint optimization of handover control and power allocation based on multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13124–13138, Nov. 2020.
- [25] Y. Sun *et al.*, "Efficient handover mechanism for radio access network slicing by exploiting distributed learning," *IEEE Trans. Netw. Service Manag.*, vol. 17, no. 4, pp. 2620–2633, Dec. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9223714/>
- [26] Y.-J. Liu, G. Feng, Y. Sun, S. Qin, and Y.-C. Liang, "Device association for RAN slicing based on hybrid federated deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15731–15745, Dec. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9237167/>
- [27] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning, concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, Feb. 2019.
- [28] K. Qi, T. Liu, and C. Yang, "Federated learning based proactive handover in millimeter-wave vehicular networks," in *Proc. 15th IEEE Int. Conf. Signal Process.*, 2020, pp. 401–406.
- [29] N. Aljeri and A. Boukerche, "An efficient movement-based handover prediction scheme for hierarchical mobile IPv6 in VANETs," in *Proc. 15th ACM Int. Symp. Perform. Eval. Wireless Ad Hoc Sensor Ubiquitous Netw.*, 2018, pp. 47–54.
- [30] 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 *Proc. IEEE 31st Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, 2020, pp. 1–6.
- [31] N. Aljeri and A. Boukerche, "A two-tier machine learning-based handover management scheme for intelligent vehicular networks," *Ad Hoc Netw.*, vol. 94, Nov. 2019, Art. no. 101930. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S157087051930160X>
- [32] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *IEEE ASSP Mag.*, vol. 3, no. 1, pp. 4–16, Jan. 1986. [Online]. Available: <https://doi.org/10.1109/MASSP.1986.1165342>
- [33] K. Tan, D. Bremner, J. Le Kerneç, L. Zhang, and M. Imran, "Machine learning in vehicular networking: An overview," *Digit. Commun. Netw.*

2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352864821000870>
- [34] 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management, 3rd Generation Partnership Project (3GPP) Technical specification (TS) 36.133, version 17.2.0, Jul. 2021. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2420>
- [35] 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification, 3rd Generation Partnership Project (3GPP) Technical specification (TS) 36.331, version 16.5.0, Jul. 2021. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440>
- [36] A. Orsino, G. Araniti, A. Molinaro, and A. Iera, "Effective RAT selection approach for 5G dense wireless networks," in *Proc. IEEE 81st Veh. Technol. Conf.*, 2015, pp. 1–5.
- [37] K. Ghanem, H. Alradwan, A. Motermawy, and A. Ahmad, "Reducing ping-pong handover effects in intra EUTRA networks," in *Proc. 8th Int. Symp. Commun. Syst. Netw. Digit. Signal Process.*, 2012, pp. 1–5.
- [38] J. Sultan, M. S. Mohsen, N. S. G. Al-Thobhani, and W. A. Jabbar, "Performance of hard handover in 5G heterogeneous networks," in *Proc. 1st Int. Conf. Emerg. Smart Technol. Appl.*, 2021, pp. 1–7.
- [39] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri, and M.-S. Alouini, "Handover management in dense cellular networks: A stochastic geometry approach," in *Proc. IEEE Int. Conf. Commun.*, 2016, pp. 1–7.
- [40] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.
- [41] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018, pp. 195–341.
- [42] V. Mnih *et al.*, "Playing Atari with deep reinforcement learning," CoRR, 2013. [Online]. Available: <http://arxiv.org/abs/1312.5602>
- [43] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018, pp. 131–133.
- [44] L.-J. Lin, "Self-improving reactive agents based on reinforcement learning, planning and teaching," *Mach. Learn.*, vol. 8, no. 3, pp. 293–321, 1992.
- [45] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018, pp. 62–64.
- [46] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. AAAI Conf. Artif. Intell.*, 2016. [Online]. Available: <https://ojs.aaai.org/index.php/AAAI/article/view/10295>
- [47] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, Jul. 2011, pp. 49–53.
- [48] G. Çelik, H. Çelebi, and G. Tuna, "A novel RSRP-based E-CID positioning for LTE networks," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf.*, 2017, pp. 1689–1692.
- [49] E. Rastorgueva-Foi, M. Costa, M. Koivisto, K. Leppänen, and M. Valkama, "User positioning in mmW 5G networks using beam-RSRP measurements and Kalman filtering," in *Proc. 21st Int. Conf. Inf. Fusion*, 2018, pp. 1–7.
- [50] H. Hendrawan, A. R. Zain, and S. Lestari, "Performance evaluation of A2-A4-RSRQ and A3-RSRP handover algorithms in LTE network," *Jurnal Elektronika dan Telekomunikasi*, vol. 19, no. 2, pp. 64–74, Dec. 2019. [Online]. Available: <https://www.jurnalet.com/jet/article/view/272>
- [51] D. Harris and S. Harris, *Digital Design and Computer Architecture*. San Mateo, CA, USA: Morgan Kaufmann, 2010.
- [52] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [53] G. F. Riley and T. R. Henderson, "The ns-3 network simulator," in *Modeling and Tools for Network Simulation*. Berlin, Germany: Springer, 2010, ch. 2, pp. 15–34. [Online]. Available: http://link.springer.com/10.1007/978-3-642-12331-3_2
- [54] N. Baldo, M. Miozzo, M. Requena-Esteso, and J. Nin-Guerrero, "An open source product-oriented LTE network simulator based on ns-3," in *Proc. 14th ACM Int. Conf. Model. Anal. Simul. Wireless Mobile Syst.*, 2011, pp. 293–298.
- [55] N. Patriciello, S. Lagen, B. Bojovic, and L. Giupponi, "An E2E simulator for 5G NR networks," *Simul. Modelling Pract. Theory*, vol. 96, Nov. 2019, Art. no. 101933. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1569190X19300589>
- [56] T. Cerqueira and M. Albano, "RoutesMobilityModel: Easy realistic mobility simulation using external information services," in *Proc. Workshop ns-3*, 2015, pp. 40–46.
- [57] Cell Mapper, "Cellular tower maps," Accessed: Apr. 2021. [Online]. Available: <https://www.cellmapper.net/Index>
- [58] NS-3, "Design documentation of the LTE module, fading model." Accessed: 2021. [Online]. Available: <https://www.nsnam.org/docs/models/html/lte-design.html#fading-model>
- [59] G. Piro, N. Baldo, and M. Miozzo, "An LTE module for the ns-3 network simulator," in *Proc. 4th Int. ICST Conf. Simul. Tools Techn.*, 2011, pp. 415–422.
- [60] D. Hendrycks and K. Gimpel, "Bridging nonlinearities and stochastic regularizers with Gaussian error linear units," CoRR, 2016. [Online]. Available: <http://arxiv.org/abs/1606.08415>
- [61] I. Loshchilov and F. Hutter, "Fixing weight decay regularization in adam," CoRR, 2017. [Online]. Available: <http://arxiv.org/abs/1711.05101>
- [62] H. Yin *et al.*, "NS3-AI: Fostering artificial intelligence algorithms for networking research," in *Proc. Workshop ns-3*, 2020, pp. 57–64.
- [63] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, Jul. 2011, pp. 63–77.
- [64] NS-3, "User documentation of the LTE module, simulation outputs." Accessed: 2021. [Online]. Available: <https://www.nsnam.org/docs/models/html/lte-user.html#simulation-output>
- [65] ns-3, "Design documentation of the LTE module, PDCP." Accessed: 2021. [Online]. Available: <https://www.nsnam.org/docs/models/html/lte-design.html#pdcp>
- [66] ns-3, "Design documentation of the LTE module, data PHY error model." Accessed: 2021. [Online]. Available: <https://www.nsnam.org/docs/models/html/lte-design.html#data-phy-error-model>
- [67] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, Jul. 2011, pp. 513–518.
- [68] L. Wang, H. Ye, L. Liang, and G. Y. Li, "Learn to compress CSI and allocate resources in vehicular networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3640–3653, Jun. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9026965/>
- [69] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [70] K. Dimou *et al.*, "Handover within 3GPP LTE: Design principles and performance," in *Proc. IEEE 70th Veh. Technol. Conf. Fall*, 2009, pp. 1–5.



Kang Tan (Student Member, IEEE) received the B.Eng. degree in electronics and electrical engineering from the Glasgow College, University of Electronic Science and Technology of China, Chengdu, China, in 2017, and the M.Sc. degree in artificial intelligence from the University of Edinburgh, Edinburgh, U.K. in 2018. He is currently working toward the Ph.D. degree with the Communications, Sensing and Imaging Research Group, James Watt School of Engineering, University of Glasgow, Glasgow, U.K. His research interests include machine

learning applications and vehicular networks.



Duncan Bremner (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electrical and electronic engineering from The University of Edinburgh, Edinburgh, U.K., and the Ph.D. and M.B.A. degrees from the University of Glasgow, Glasgow, U.K. He has more than 30 years in the semiconductor industry and has held operational and strategic executive roles in product development and technology planning within leading organisations, such as National Semiconductor and The Intel Corporation. Originally trained and recognised as an accomplished analog design engineer, he has held both product management and customer facing roles before establishing Intel's high speed development centre for global communication products. He was a Member of CTO office with Intel, for communications products to developing product roadmaps before being invited to define and lead a \$15M program to develop ultra-wideband 10 GHz wireless technology for ITI Techmedia.

He was active in promoting change and improvements to Engineering education and is a Senior Lecturer, teaching several electronics courses to students in Glasgow and Chengdu in collaboration with the University of Electronics Science and Technology China, Chengdu, China. As a Senior Member of staff, he has also been the Program Director of two of the programmes in the TNE partnership. He is cited as inventor on 17 patents.



Julien Le Kerne (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees in electronic engineering from the Cork Institute of Technology, Cork, Ireland, in 2004 and 2006, respectively, and the Ph.D. degree in electronic engineering from University Pierre and Marie Curie, Paris, France, in 2011. He is currently a Senior Lecturer with the School of Engineering, University of Glasgow, Glasgow, U.K. He is also a Senior Lecturer with the University of Electronic Science and Technology of China, Chengdu, China, and an Adjunct Associate Professor

with the ETIS Laboratory, University of Cergy-Pontoise, Cergy, France. His research interests include radar system design, software-defined radio/radar, signal processing, and health applications.



Yusuf Sambo (Member, IEEE) received the M.Sc. degree (with distinction) in mobile and satellite communications, and the Ph.D. degree in electronic engineering from the Institute for Communication Systems formally known as CCSR, University of Surrey, Guildford, U.K, in 2011 and 2016, respectively. Between 2016 and 2017, he was a Lecturer in telecommunications engineering with Baze University, Abuja, Nigeria. He was a Postdoctoral Research Associate with the Communications, Sensing and Imaging Research Group, University of Glas-

gow, Glasgow, U.K. Since 2019, he has been a Lecturer with the James Watt School of Engineering, University of Glasgow, and 5G Testbed Lead managing the Scotland 5G Centre testbed with the University of Glasgow. His research interests include self-organising networks, cell-free massive MIMO, EMF exposure, and green communications.



Lei Zhang (Senior Member, IEEE) received the Ph.D. degree from the University of Sheffield, Sheffield, U.K. He is currently a Lecturer with the University of Glasgow, Glasgow, U.K. He also holds a Visiting Position in 5G IC with the University of Surrey, Guildford, U.K. His research interests include communications and array signal processing, including radio access network slicing (RAN slicing), wireless blockchain networks, new air interface design, Internet of Things, V2X, multiantenna signal processing, and massive MIMO systems. He is holding 16

U.S., U.K., EU, and China granted patents on wireless communications. He is an Associate Editor for the IEEE ACCESS.



Muhammad Ali Imran (Senior Member, IEEE) is currently a Professor of wireless communication systems. He heads the Communications, Sensing and Imaging Research Group, University of Glasgow, Glasgow, U.K. He is an Affiliate Professor with The University of Oklahoma, Norman, OK, USA, and a Visiting Professor with 5G Innovation Centre, University of Surrey, Guildford, U.K. He has more than 20 years of combined academic and industry experience with several leading roles in multimillion pounds funded projects. He has authored or coauthored more

than 400 journal and conference publications, was the Editor of three books and author of more than 20 book chapters, has filed 15 patents and has successfully supervised more than 40 Postgraduate students at doctoral level. His research interests include self organised networks, wireless networked control systems, and wireless sensor systems. He was a consultant to international projects and local companies in the area of self-organised networks. He is a Fellow of IET and a Senior Fellow of HEA.