Received: December 1, 2023. Revised: January 20, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Distributed Reinforcement Learning–based Seamless Multi-Connectivity Solution for Heterogeneous 5GNR and Wi-Fi Networks

Haider D. Albonda[1]*

Department of Control and System Engineering, University of Technology, Baghdad, Iraq
* Corresponding author's Email: *haider.d.albonda@uotechnology.edu.iq*

Abstract: The integration of 5G New Radio (5G-NR) and Wi-Fi networks in heterogeneous environments holds significant promise for delivering critical services with relatively high reliability and low latency. In this paper, we tackle the critical challenge of establishing robust multi-connectivity within heterogeneous systems integrating 5G NR and Wi-Fi systems. In this context, our focus is on optimizing a new, distributed multi-Connectivity solution, with a main emphasis on identifying the ideal base stations (BSs) and Wi-Fi access points (APs) with their respective spectrum associations. This solution is essential for ensuring low-latency and reliable communication required, particularly for mission-critical services. The simulation results demonstrate the significant advantages of the proposed algorithm compared to reference schemes. It leads to network performance improvements by enhancing system utility function, reducing service access time, and increasing packet success rates, through high-quality connectivity with relatively low interference probability. In particular, for mission-critical services, the proposed algorithm leads to a 16 % and 41% increase in effective network throughput when the offered load is 190 Mbps, which outperforms the benchmark approaches. Furthermore, it improves the utility function by approximately 0.12, and 0.19 compared to the reference approaches. This reflects a substantial improvements in success delivery rates for mission-critical services by 10 % and 18 %, respectively, when compared to the reference approaches, particularly evident at an offered load of approximately 190Mbps.

Keywords: 5G-NR, Cellular networks, Reinforcement learning, Wi-Fi networks, Licensed and unlicensed spectrum.

1. Introduction

The increasing demand for mission-critical communication services such as public safety, transportation, and healthcare that require consistent and efficient communications during emergencies and mission-critical situations has been experienced over the past few years [1, 2]. The mission-critical services require persistent connectivity, low latency, and high reliability to guarantee a safe environment, and the smooth operation of critical systems. While the cellular networks available for these services are expensive and restricted, Wi-Fi systems are becoming increasingly congested and susceptible to interference.

The problem addressed in this study is lies in the limitations closely associated to systems that rely on a single radio access technology (RAT) and a single specific spectrum band used for data transmission. As a result, this approach is unable to manage communications among users effectively and appropriately, exposing the reliability of services to be relatively low, and ultimately leading to network congestion, service outages, and greatly increased delay [3, 4]. In response to addressing this issue, the heterogeneous networks (Het-Nets) turns out to be a practical strategic solution, which are made possible by multi-connectivity and incorporate a variety of access technologies, such as wifi and cellular systems. The coexistence of Wi-Fi with mobile cellular networks that enable multi-connectivity has the potential to reduce the inefficiencies associated to single-spectrum radio access technology and meet the growing needs for high data rates and low latency. Potential multi-connectivity, which uses multiple links over mobile cellular systems and Wi-Fi networks, opens up new opportunities and capabilities for reliable communications [5]. The capability of Wi-Fi systems have proven to significantly impact the data exchange process by providing enhanced radio beams for secure communications, reinforcing their position as an effective option for a wide range of applications. The cost effectiveness and flexibility offered by Wi-Fi systems improve their ability to survive in a dynamic scalable communications environment.

The coexistence of multi-connectivity-enabled cellular and Wi-Fi communications systems is receiving increasing attention, driven by the demand for more efficient, adaptive and robust communications solutions. Multi-connectivity provides dynamic load balancing and traffic optimization. Since traffic is distributed across multiple connections, congested networks are avoided, communications interference is mitigated, and network outages are reduced [6, 7]. This means that mission-critical services are able to access the bandwidth and resources they need even in high demand scenarios. Multi-connectivity also supports seamless transition between networks or technologies when switching between the base station and access point. Furthermore, seamless switching between different technologies is supported when moving between base stations or access points. However, implementing multi-connectivity in mission-critical communications is crucial and can pose challenges. These challenges include optimal allocation of spectrum bands, coordination and management of network interference, maintaining network compatibility and interoperability, and addressing spectrum utilization and regulatory constraints.

To tackle these challenges and ensure the reliability of mission-critical services across multiple networks, different solutions have been proposed including spectrum allocation [8, 9], capacity management strategies [10, 11], and interference control strategies [12, 13].

These thoughtful solutions aim to facilitate optimize resource utilization; minimize interference, and appropriate spectrum allocation among users. In addition, strict restrictions on the use of unlicensed frequencies in safety-critical communications have been suspended, ensuring the reliability of these services, while also taking considering the proposal and development of various regulatory frameworks. For instance, the European telecommunications guidelines institute (ETSI) and the US federal communications commission (FCC) released guidelines for the use of unlicensed spectrum in industrial applications and public safety applications, respectively [14, 15]. Nevertheless, despite these initiatives, the ongoing challenge lies in the process of choosing the most appropriate communication technology to meet the unique requirements of various services.

The motivation behind the call for proposing multi-communication strategy driven by the need to effectively address these challenges. The proposed Multi-connectivity is rooted in the pursuit of reliability and enhanced the user requirement especially in terms of throughput and latency. Through the utilization multiple communication links and diverse technologies, multi-connectivity ensures reliable service performance, provides flexibility to address scalable demands, and enhances system resilience against the interference and the effects of network outages even in challenging situations. Indeed, multi-connectivity is emerging as the preferred choice for mission-critical applications where the imperatives of high-quality data transmission, congestion prevention through load balancing, reliability, and reduced latency are crucial.

The remaining sections of this paper are arranged as follows: in section 2, a comprehensive review of existing works in the relevant field is presented. Section 3 provides a system model based on different considerations. A proposed solution for Multi-Connectivity within heterogeneous 5GNR and Wi-Fi Networks is presented in section 4. In section 5, the performance evaluation of the proposed solution is conducted. Finally, based on the insights and contributions of the research, section 6 draws conclusions.

2. Related works

In heterogeneous environments involving 5G-NR and Wi-Fi, extensive research has been conducted to achieve seamless integration between WiFi and cellular networks. The researchers have particularly focus on critical parameters such as throughput, interference, and latency, which significantly impact network performance. Considering a connectivity solutions for 5G NR heterogeneous cellular networks, the authors in [16] proposed an improved sleep control solution for new heterogeneous radio (NR) 5G cellular networks (HetNets) consisting of different base stations, including macro and small utilize licensed spectrum for transmission. The proposed solution is applied to enhance wireless coverage and network capacity, and energy efficiency. Similarly, a seamless communication solution is proposed in [17, 18] to enhance the quality of service (QoS) of heterogeneous wireless networks (HWNs) involving various applications using licensed spectrum. However, the solutions proposed in [16-18] face scalability issues particularly challenges when scaling up to a larger network or a more complex environment, especially since it relies on the use of a single radio spectrum band, which may be insufficient, causing service outages as user requests increase. Furthermore, an approach based on utility-based handover decisions involves selecting the network with a lower load is proposed in [19] to provide better performance for vehicular network, particularly in terms of throughput and handover rate. The utility is designed based on network traffic load and SINR using licensed spectrum. An aspect that has not been widely included in their proposed solution is its susceptibility to challenges in dynamic environments, due to the assumption of network stability that may not hold. This could pose potential challenges to facility capacity management and affect the adaptability and effectiveness of the considered approach. In addition, a RAT selection scheme is proposed in [20], which involves efficient multiple communication (MC) configuration to enhance reliability. By incorporating distributed reinforcement learning (RL), each device has the capability to learn the policy for configuring an effective MC and selecting RATs appropriately. However, the work in [20] does not thoroughly explore the potential benefits and challenges associated with RAT selection and dynamic spectrum access. Incorporating dynamic spectrum sharing of different spectrum bands could significantly enhance resource utilization and contribute to the overall efficiency.

On the other hand, in [21], an optimization model for LTE-LAA/WiFi coexistence in both 5GHz and 5GHz NR-U (6GHz) bands is proposed and then analyzed. Although the proposed optimization model is effective within its scope, it exhibits limitations because it relies on a simplified indoor environment. This simplification introduces unrealistic scenarios, overlooking crucial factors including varying user densities, interference conditions, and mobility patterns. Furthermore, [22] explores recent improvements to LTE-LAA coexistence with WiFi systems, and compares the performance for different access priority classes. Nevertheless, the study does not thoroughly explore the influence of coexistence on quality of service (QoS) metrics i.e., the impact of coexistence on parameters such as latency, packet loss, and other QoS factors that would enhance a comprehensive understanding of LTE-LAA coexistence with WiFi systems. The impact of a unique phenomenon related to Physical Cell Id (PCI) on integrated system including LTE, LAA, and Wi-Fi is demonstrated in [23] using machine learning algorithms. Although [23] contributes significantly to enhance LTE-LAA coexistence with WiFi systems, the study does not explicitly examine the impact of varying user densities on the observed PCI scenarios. Particularly in scenarios marked by high traffic density, fluctuations in the dynamics of coexistence and PCI-related effects may vary, which may potentially affect the robustness of identified solutions.

Following the conventional LBT introduced by 3GPP, the coexistence performance between 5G and Wi-Fi operators is demonstrated in [24] to provide optimal performance in terms of latency and throughput. However, the work may not extensively address the dynamic nature of resource sharing among different RATs, including varying levels of spectrum contention and adaptability RATs to dynamic spectrum conditions. This investigation is essential, since it affects the optimization of resources and effectively enhances the overall performance of the system. The authors in [25] proposed two solutions for a dynamic sharing of multiple radio resources to mobile network operators (MNOs) involving cellular technologies and Wi-Fi systems. Although it is noted that the solutions are scalable to 5G NR-U, the specific challenges in implementing 5G NR-U scenarios are not explicitly addressed in the study. Clearly addressing the unique characteristics of 5G NR-U, including factors such as delay and packet drop, would greatly contribute to enhancing the relevance and completeness of the proposed solutions. Considering system fairness constraints, the overall throughput of integrated LTE-LAA and Wi-Fi networks is maximized in [26] based on reinforcement learning strategies applied to intelligently adjust the size of the contention window for both LTE-LAA and Wi-Fi users. Similarly, the authors in [27, 28] proposed a reinforcement learning (RL) based subchannel selection strategy that allows users to make informed decisions about the appropriate subchannels, BSs, and access points (APs) while considering physical layer characteristics and media access control (MAC) channel protocols. Notably, the study in [26-28] lacks a sustained exploration of these algorithms in dynamic and complex network environments. This includes scenarios with multiple RATs involving a number of cells, where users can seamlessly switch to work within the best cells and take advantage of different spectrum bands. In addition, considering user mobility, changing interference conditions, and combining heterogeneous deployment settings may significantly impact system performance.

A dynamic transmission opportunity period (TxOP) is introduced in [29] to provide better fairness and higher total aggregate throughput for LTE-

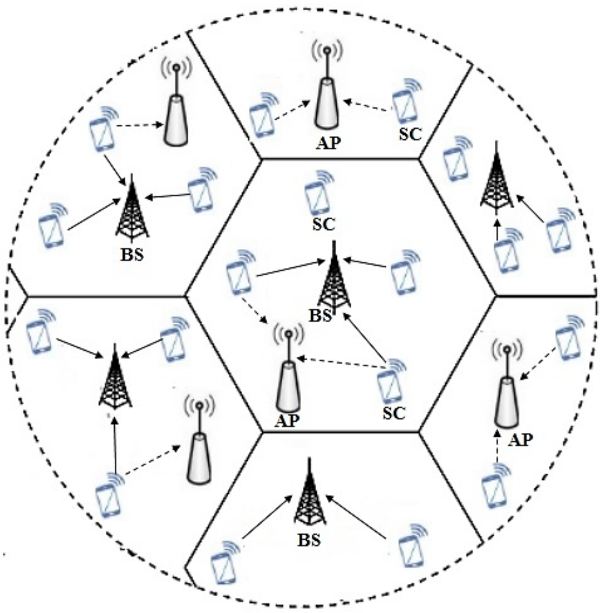

Figure. 1 System model enabled multi-connectivity capabilities

LAA/Wi-Fi coexistence networks. The coexistence of Wi-Fi/ LAA networks operating in dense deployment scenarios is investigated in [30, 31] when the authors presented a solution to maximize throughput and avoid session interruption. Additionally, in order to analyze the collision probability, system throughput, and the access probability of LTE and Wi-Fi networks, two coexistence strategies based on the Markov model are investigated in [32]. However, the study in [30-32], mainly focuses on throughput LTE license assisted access (LAA), and overlooks other critical metrics such as resource utilization, delay, and packet loss that are crucial to evaluate the overall QoS for vehicular scenarios.

Different from the aforementioned works, this study addresses the intricate dynamics of implementing a multi-connectivity strategy for mission-critical communications within vehicular scenarios. In particular, a heterogeneous scenario involving 5GNR and Wi-Fi networks with multiple cells facilitating multi-connectivity across various technologies is considered, efficiently utilizing various spectrum bands involving licensed and unlicensed spectrum resources.

In this context, the key contributions of this paper in addressing the identified gap contributions can be outlined as:

1- We propose a new strategy to address the challenge of determining the most appropriate connectivity option for mission-critical communication in a multi-connectivity scenario involving 5GNR cells and Wi-Fi Networks. This strategy combines Q-learning and softmax decision-making methods to improve the decision-making process. In this scenario, users have the ability to be connected to a high-quality BSs/APs and utilize both licensed and unlicensed spectrum.
2- The proposed multi-connection mechanism is optimized to achieve high utility function efficiently and ensure reliable communication with high probability of success rate. This optimization reduces delays while effectively meeting the unique needs of safety-critical services under varying traffic loads.
3- We conduct an evaluation of the effectiveness of the proposed mechanism across a range of system parameters through extensive simulations. This evaluation mainly focuses on utility function, network throughput, and packet success rate.

3. System model

The scenario considered in this paper is demonstrated in Fig. 1. It assumes a heterogeneous radio access network (RAN) made up of multiple radio access technologies (Multi-RAT) includes N individual cells per RAT, represented by the index n = {1, 2, ..., N}. Similar to [7,33], these RATs incorporates of 5G NR cells and the existence of potential Wi-Fi access points distributed within the network coverage area. The scenario involves the provision of a mission-critical service support specific set of vehicular users, denoted by an index m that ranges from 1 to M given access to these RATs. Each user m generates video session traffic following a Poisson process with an arrival rate density, λ_s. This paper assumes communication in the uplink direction, meaning the transmission is from users to BSs or APs. However, it can be easily extended and applied to include considerations of downlink direction. In the considered scenario, 5G NR cells are expected to be localized at higher carrier frequencies such as the 5 GHz band considered by 3GPP [34]. The frequency band of bandwidth B is divided into a set of orthogonal channels, denoted as $f = \{1, 2, \ldots F\}$. At a given time t, a specific set of sub channels is allocated to cell n. On the other hand, the access points operate in the unlicensed 5.8GHz band.

Additionally, our model operates under the assumption that mission-critical users have multi-connectivity capabilities. This enables them to simultaneously establish connections across various RATs and multiple cells per RAT. Additionally, mission-critical users have the ability to utilize the available licensed and unlicensed channels, sharing them with Wi-Fi networks. The achievable bit rate for each user m in cell n, indexed by $D_{n,m}^{(t)}$, can be determined as follows:

$$D_{n,m}^{t} = \sum_{s} \frac{L_{m,s(n)} \cdot \delta_s}{T_s} \quad (1)$$

Where $L_{m,s}$ is the number of session generated by mission-critical user m in cell n, and δ_s is the size of each session s generated by user m in cell n. T_s is the duration of each session s generated by user m. The total traffic demand for cell n, given by $D_{n}^{(t)}$ can then be determined as the aggregate of the traffic demands of all users accessing cell n as follows:

$$D_{n,m}^{(t)} = \sum_{m} \sum_{f} \rho_{m,f} \cdot D_{n,m}^{t} \quad (2)$$

$\rho_{m,f} = 1$ if the channel f is assigned to the mission-critical user m, and zero otherwise. In order to meet traffic requirements and ensure satisfactory service quality for mission-critical users, it is essential to optimize a utility function that incorporates key parameters such as delay and achieved data rate for each cell. Therefore, the utility function for mission-critical users in cell n, denoted by U_n, is considered and defined as follows:

$$\mathrm{U}_{n} = \sum_{m} \left[\alpha . Min\left(\mathrm{C}_{n,m}, D_{n,m}^{t}\right) . \left(\beta . \sum_{L} \frac{D_{m,L}(n)}{P_{m,L}}\right)\right] \quad (3)$$

Where $C_{n,m}$ is the capacity of user m provided by the total capacity of user m when connected to BS n denoted by $C_{n,m}^{(L)}$ and the capacity of user m when connected to wifi access points j, denoted by $C_{j,m}^{(U)}$ i.e., $C_{n,m} = C_{n,m}^{(L)} + C_{j,m}^{(U)}$. Then, $C_{n,m}^{(L)}$ and $C_{j,m}^{(U)}$ are given by the Shannon- theorem as follows:

$$C_{n,m}^{(U)} = \sum_{u} \rho_{m,u} . B_{n}^{(U)} \left(1 + \frac{P_{m,u}^{(U)} \cdot G_{m,j}^{(U)}}{B_{n}^{(U)} . N_o}\right) \quad (4)$$

$$C_{n,m}^{(L)} = \sum_{f} \rho_{m,f} . B_{n}^{(L)} . \left[1 + \frac{P_{m,f}^{(L)} \cdot G_{m,n}^{(L)}}{\sum_{\substack{\overline{m}=1 \\ \overline{m} \neq m}} \rho_{\overline{m},f} P_{\overline{m},f}^{(L)} \cdot G_{\overline{m},n}^{(L)} + B_{n}^{(L)} . N_o}\right] \quad (5)$$

where $P_{m,f}^{(L)}$ and $P_{m,u}^{(U)}$ are the transmission power of mission critical user $m \in$ M when connected to BS n on licensed subchannel f, and when connected to AP j on unlicensed subchannel u, respectively. $G_{m,n}^{(L)}$ is the channel gain between mission critical user m and BS n, and $G_{m,j}^{(U)}$ is the channel gain between mission critical user m and Wifi access point j. N_o is the noise power spectrum density. $G_{\overline{m},n}^{(L)}$ is the interference link gain from the transmitter $\overline{m} \neq m$ to BS n. $B_{n}^{(L)}$ and $B_{n}^{(U)}$ are licensed bandwidth and unlicensed bandwidth that can be allocated to users in cell n, respectively. $\rho_{m,u}$ are channel assignment indicator i.e., $\rho_{m,u}= 1$ if the unlicensed channel u is assigned to the user m otherwise its value is zero. The notations utilized in this paper are organized and indexed in Table1.

4. The proposed multi-connectivity strategy

The proposed distributed multi-connectivity solution is designed to address the selection of the optimal co

-nnectivity option by identifying the best BSs and APs along with their associate licensed and unlicensed sub-bands for uplink transmission.

To achieve these objectives, we adopt a distributed approach based on a reinforcement learning (RL), where each user m independently selects the most appropriate BSs/APs and spectrum based on their communication requirements. In particular, our RL-based solution enables an agent, seamlessly integrated within the user equipment, to actively interact with the network environment as depicted in Fig. 2. This environment consists of multiple BSs and APs with the capability to access licensed/unlicensed spectrum resources. At each time step t, the agent of each user m, evaluates the current state, s_t. Subsequently, it initiates a decision to take an action, a_t. This process involves the selection of optimal BSs and APs along with their associate licensed and unlicensed sub-bands for uplink transmission based on a softmax policy [35], which allows for probabilistic selection of actions based on the Q-values. Following the selected action a_t, the environment state, $s_t \in$ S transits to a new state, s_{t+1} and the agent receives a reward, $R(s_t, a_t)$, which is determined by the respective link capabilities and response time constraints. In our system, the state observed by each link, which characterizes the environment between transmitters and receivers, consists of several components. These include the user location, the instantaneous received signal to noise ratio when user m connected to BS and AP, and the set of available Licensed and unlicensed sub channels for UL transmissions of user m denoted respectively as $r_f \in f$ and $W_u \in U$. Hence, the state can be expressed as s_t = [UE location, , r_f, W_u, $\lambda_{n,u}^{(T)}$, $\lambda_{n,f}^{(T)}$)] $\in$ S.

Table1. List of notations

Definition	Notation
δ_s	The size of each session s generated by user m.
T_s	The duration of each session s
$L_{m,s}$	Number of sessions generated by user m
$D_{n,m}^{(t)}$	The achievable bit rate for each user m in cell n.
$D_{n}^{(t)}$	The average number of required RBs by the IoT service
$\rho_{m,f}$, $\rho_{m,u}$	Channel assignment indicators
$C_{n,m}^{(L)}$	the total capacity of user m when connected to BS n.
$C_{j,m}^{(U)}$	the capacity of user m when connected to wifi access points j
$P_{m,f}^{(L)}$	The transmission power of user m on licensed subchannel f
$P_{m,u}^{(U)}$	The transmission power of user m on unlicensed subchannel u
$G_{m,n}^{(U)}$	Channel gain between mission-critical user m and BS n
$G_{m,j}^{(U)}$	Channel gain between mission-critical user m and AP j
$B_{n}^{(L)}$	Licensed bandwidth of cell n
$B_{n}^{(U)}$	Unlicensed bandwidth of cell n
r_f	The set of available licensed sub channels.
W_u	The set of available unlicensed sub channels.
α , β	Weight parameters
θ	Learning rate
λ_s	Session arrival rate

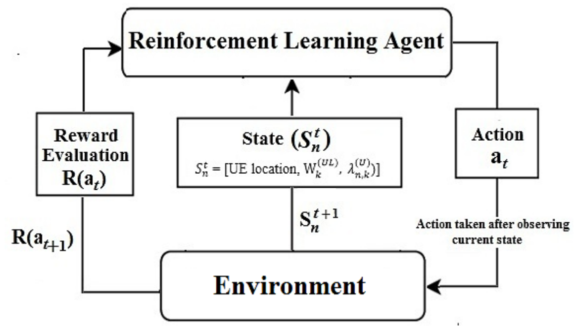

Figure. 2 Reinforcement learning model

To facilitate learning and decision-making of RL following Algorithm 1, each UE m maintains a record of its experiences when connecting to each of the BSs and APs through the available sub channels r_f and W_u. Whenever a user m utilizes a specific BS n using sub channel r_f and wifi access point j using W_u, the corresponding Q(s_t, a_t) value is updated for a given state-action pair (s_t, a_t), with a null discount rate (line 19). This update process allows the UE m to continuously learn and improve its decision-making regarding the choice of BSs/APs and associated with their subchannels r_f and W_u, based on the historical experiences stored in the $Q(s_t, a_t)$ values. The computation of $Q(s_t, a_t)$, involves updating the value based on the rewards received and the learning rate. Specifically, the new value of $Q(s_t, a_t)$, is calculated as follows:

$$Q(s_t.a_t) \leftarrow (1-\theta).Q(s_t.a_t) + \theta.R(s_t.a_t) \quad (6)$$

In this eq. (6), a= {(Select BS n , Select AP j) | , $n \in$ Base Stations, $j \in$ Access Points}, $\theta \in (0, 1)$ represents the learning rate. The rewards, denoted by $R(a_t)$, is designed to reflect the degree of achievement in meeting the optimization goal and constraints. Specifically, our objective is to achieve high data rate and minimize delay. The reward is determined based on the achieved capacity, $C_{n,m}$ and actual required bitrate, $D_{n,m}$ obtained since the last BS and AP selection.

BSs and APs that contribute to less delay and greater achieved capacity receive higher rewards, leading to larger values for $Q(s_t, a_t)$.Taking into account the above considerations, the reward function when establishing a connection between user m to BS n is determined as follows:

$$R(s,a) = \begin{bmatrix} 1-\left|\frac{max[(C_{n,m}(a)-D_{n,m}),0]}{D_{n,m}}\right| & C_{n,m} \geq D_{n,m} \\ \frac{max[(C_{n,m}(a)-D_{n,m}),0]}{D_{n,m}} & Otherwise \end{bmatrix} \quad (7)$$

The reward function, as given in Eq. (7), captures the performance of the system by assigning a value between 0 and 1. It exhibits an exponential increase when the achieved bit rate, $C_{n,m}$ is below the required bit rate, $D_{n,m}$. Conversely, if the service requirement is not met, the reward decreases. It provides a feedback to the agent, influencing its decision-making process throughout the system operation (lines 8-13). Once the Q-values are computed, the agent can use them to choose the action with the highest Q-value, indicating the optimal choice in that particular state (lines 14-18). This selection process, based on maximizing the Q-values, allows the agent to make more effective and informed decisions. Mathematically, this can be expressed as:

$$a_t = arg\max_{a \in A} Q(s_t.a_t) \quad (8)$$

The probability of selecting a particular action is determined by its Q-value relative to other available actions.

$$P_r\left(\left(BS_n, AP_j \middle| s_t\right)\right) = \frac{e^{Q\left(s_t, BS_n, AP_j\right)/\tau}}{\sum_{j=1}^{J} e^{Q\left(s_t, AP_j\right)/\tau} + \sum_{n=1}^{N} e^{Q(s_t, BS_n)/\tau}} \quad (9)$$

Specifically, the probability of selecting action a_t, $P(BS_n, AP_j)$ is defined as,Where τ is the temperature parameter that plays a crucial role in balancing exploration and exploitation. Actions associated with higher Q values, including the selection of the corresponding APs or BSs and spectrum, are assigned higher probabilities. This prioritizes the exploitation of proven successful actions, as per Eq. (8). To identify the sets of available unlicensed channels, users actively measure the received power and perform channel sensing. This process follows to the Category 4 listen-before-talk (LBT) approach specified in [36].

5. Results and discussion

In the given scenario, we consider a simulation model with an urban environment characterized by 5G NR cells and Wi-Fi access points randomly distributed across their coverage areas. The essential simulation parameters are detailed in Table 2. In this configured environment, we simulate mission-critical traffic generated by users that are enabled by the multi-connectivity feature, and actively maintain Video sessions. The generation of these sessions occurs randomly, following a Poisson process at a rate of λs. It is worth noting that these UEs have the ability to connect to a maximum of two RATs, and within each RAT, they can establish a connection to a single cell. To facilitate our comparative analysis, we compare our proposed multi-connectivity strategy based on RL (MCS-RL) with two reference approaches. The first reference approach, denoted as Dynamic-Based access strategy (D-BAS) inspired by [18], enforces exclusive connectivity to multiple BSs of different RATs, with dynamic sharing of resources using only licensed spectrum. The second approach, referred to as Random access strategy (RAS) and inspired by [20], enables users to autonomously select their radio access technology and resources based on signal strength.

Fig. 3 shows the utility function versus offered load which clearly demonstrates the superior performance of the proposed MCS-RL approach compared to RAS and D-BAS. In Fig. 3, we can see that the proposed MCS-RL strategy achieves significantly higher utility improvement, with approximately 129 at the offered load of 225 Mbps. On the other hand, the D-BAS approach achieves a utility of about 119 Mbps when the offered load is 225 Mbps, which is significantly lower than that of the proposed strategy.

The RAS approach performs less than the proposed RL-BCS strategy and D-BAS approach, achieving a utility of about 108 Mbps when the offered load is 225 Mbps. Compared with the proposed strategy, the different utilities achieved by the RAS and D-BAS approaches indicate that it is less effective in providing utility improvements than the proposed strategy, as detailed in Table 3. The

Algorithm: Multi-connectivity based RL strategy

1. ***Data:*** $r_f^{(T)}$, $W_u^{(T)}$, $\lambda_{n,u}^{(T)}$, $\lambda_{n,f}^{(T)}$, γ, ε.
2. ***Initialization:***
3. Choose an initial state, s_0.
4. Initialize BSs /APs allocation P_r(s , a)
5. Initialize approximated Q-value $Q(s_t, a_t)$;
6. ***Iteration***
7. ***While*** learning period is active do
8. ***for*** each selection epoch *t* do
9. Evaluate the current state, $s = s_t$.
10. Compute the softmax probabilities for . selecting each action $a_t \in A$ based on . Eq.(9)
11. Generate a random number between 0 and 1.
12. Select $a_t = \max_{a \in A} Q(s_t, a_t)$ according to P_r(s , a) in Eq. (9) ;
13. Connect to the selected BS *n* or AP *j* : a_t.
14. ***If*** $C_{n,m} > D_{n,m}$
15. $R(s,a) = 1 - \left|\frac{max[(C_{n,m}(a) - D_{n,m}),0]}{D_{n,m}}\right|$
16. ***Else***
17. $R(s,a) = \left|\frac{max[(C_{n,m}(a) - D_{n,m}),0]}{D_{n,m}}\right|$
18. ***End***
19. ***Update*** $Q(s_t, a_t)$ based on equation (6)
20. ***Update State:*** $s_t \leftarrow s_{t+1}$
21. ***Update Iteration Index:*** $t \leftarrow t + 1$
22. ***End***
23. ***End***

Output: BSs *n* and APs *j* associated for UE *m*

Table 2. Network configuration parameters

Parameter	Value
Number of Cells	5
Cell radius	600m
Licensed Channel bandwidth	50 MHz
RBs per cell	100 RBs
UE antenna gain	10 dB
UE noise figure	9 dB
height of the BS	10m
UE height	1.5 m
Subcarrier separation	60kHz
Spectral Efficiency (*Seff*)	5.6 b/s/Hz.
Path loss model for licensed Network	$128.1+37.6\log_{10}(d[km])$
Path loss model for Unlicensed Network	$148.1+\alpha x$ $10\log_{10}(R)$, $\alpha = 4$
Number of UEs	80
λa , session arrival rate	1 session/s
w_{th}	30 ms
Session size, δ_s	5 Mbit/s
Transmit power of CU	23 dBm
Number of WiFi APs	6
Unlicensed Channel bandwidth	400MHz
α	0.4
β	0.6
Learning rate, θ	0.1
Temperature parameter τ	0.1

Table 3. System utility of the proposed MCS-RL and reference approaches

Offered Load	RAS	D-BAS	MCS-RL
System Utility			
21.645	20.23	19.87	20.928
77.1134	52.26	58.06	68.91
120	75.83	83.78	98.78
143.2075	84.03	94.14	107.62
190	98.86	112.12	122.67
225	108.45	119.71	129.42
300	117.61	129.92	140.25

notable improvement in the utility function realized by our proposed MCS-RL solution can be attributed to its ability to adeptly accommodate large user requests even under heavy network loads. Achieving this in the shortest time is made possible by the proposed strategy, which consistently selects the optimal radio access technology and spectrum band. This strategic selection not only allows for more efficient radio spectrum utilization but also ensures a reliable communication between users, contributing to enhanced user experience and overall network performance significantly.

In turn, Fig. 4 makes use of cumulative distribution functions (CDFs) of the utility function to give significant insight into the system's performance for transmissions of mission critical services. These outcomes unequivocally demonstrate the superiority of our proposed MCS-RL strategy compared to the RAS and D-BAS approaches. In particular, the proposed MCS-RL strategy stands out by consistently maintaining a higher utility function. The analysis shows that the proposed strategy achieves a remarkable utility of up to 143 Mb/s with a probability of 95%.

The D-BAS comes with a utility of only 100 Mb/s, whereas the RAS reference system manages to obtain a lower value of 80 Mb/s.

In Fig. 5, the throughput measured in megabits per second (Mbits/sec) is presented for the proposed MCS-RL strategy against the RAS and D-BAS approaches.

The results unequivocally demonstrate that the MCS-RL strategy performs better than the RAS and D-BAS approaches. In particular, we can observe that the proposed MCS-RL strategy performs superbly, reaching a maximum throughput of around 255 MB/s when the offered load is 318 MB/s.

In contrast, the D-BAS approach achieves a maximum throughput of 203 Mb/s when the load is 318 MB/s. This constitutes a significant benefit for the proposed strategy. Furthermore, when the offered load is 318 MB/s, the system achieves a maximum throughput of 180 Mb/s when the RAS approach is considered. The proposed MCS-RL strategy outperforms both RAS and D-BAS approaches in this scenario, showing a noteworthy relative increase of 21% and 27%, respectively. Our proposed MCS-RL technique has consistently demonstrated superior throughput performance compared to the RAS and D-BAS reference methods. This observation indicates improved data transfer efficiency and its ability to accommodate higher data loads seamlessly. The significant improvement in throughput not only exceeds current standards, but also positions our proposed solution as a promising innovation, enhancing overall network performance and data transfer rates. Table 4 provides a comprehensive performance comparison of the proposed MCS-RL strategy with the RAS and D- BAS reference approaches, specifically with respect to throughput under various offered loads. The improvement results in Fig. 5 can be attributed to the ability of the proposed MCS-RL strategy to establish

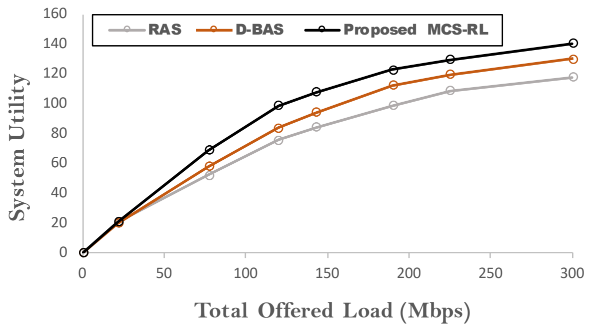

Figure. 3 The system utility versus the offered load

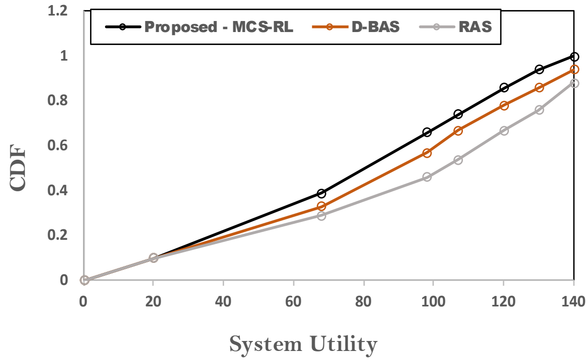

Figure. 4 The CDF of the utility system

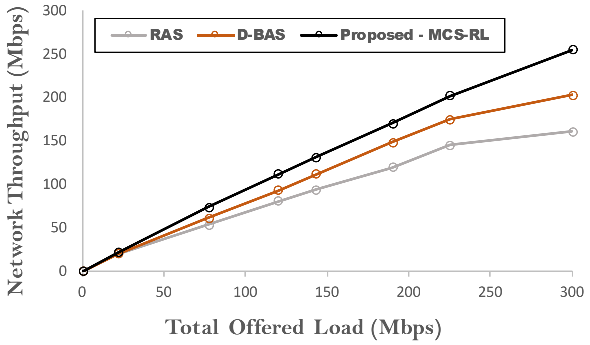

Figure 5. The network throughput versus the offered load

highly reliable communications with high signal-to-noise ratio (SINR), which is a key factor in achieving larger capacity and thus higher data rates. This feature solidifies it as an effective solution for enhancing network performance in data transmissions.

On the other hand, Fig. 6 illustrates the use of CDF to perform a comprehensive throughput analysis. These findings clearly highlight the superior performance of our proposed MCS-RL strategy when placed alongside the RAS and D-BAS schemes. In particular, our proposed approach consistently outperforms by maintaining higher throughput. The figure demonstrates that, with a probability of 98%, our proposed scheme achieves a remarkable

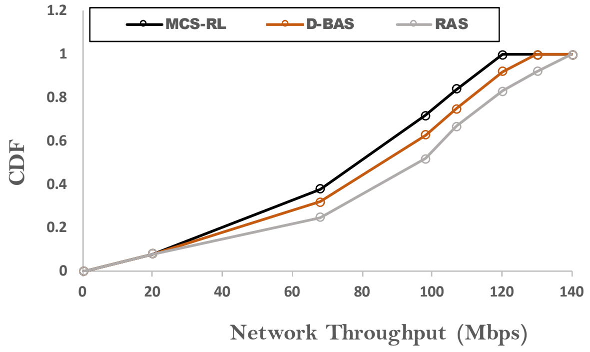

Figure. 6 The CDF of the network throughput

throughput of up to 120 Mb/s. Instead, D-BAS approach only manages a throughput of 100 MB/s, while RAS approach only manages a lower throughput of 80 MB/s.

In Fig. 7, the performance of the proposed MCS-RL solution is compared with the RAS and D-BAS reference approaches, with particular emphasis on the packet success rate. This rate depicts the percentage of packets that are successfully delivered while considering their delivery in time before reaching the threshold, w_{th}. In particular the proposed strategy effectively maintains a greater packet success rate than the RAS and D-BAS approaches i.e., the proposed strategy yields a remarkable packet success rate of about 98% when the offered load reaches 120 Mbps. This indicates that a large number of packets with high probability are almost delivered successfully and satisfied the requisite requirements. On the other hand, for the D-BAS reference approach, the packet success rate experiences a slight drop, reaching around 89%, under the offered load of 120 Mbps. The RAS reference approach showed a lower packet success rate, around 81% when the offered load reaches 120 Mbps. The proposed solution consistently outperforms reference schemes, ensuring a higher rate of successful packet delivery, which is crucial for reliable communication. The significant increases in packet success rate observed in our proposed MCS-RL, which are attributed to improved network congestion management and efficient error handling by our proposed solution as a result of exploiting different radio access technologies and spectrums which frees up more available resources and also increases overall communication reliability. This validation further solidifies the effectiveness of our proposed MCS-RL in different scenarios with different loads and conditions, as detailed in Table 5.

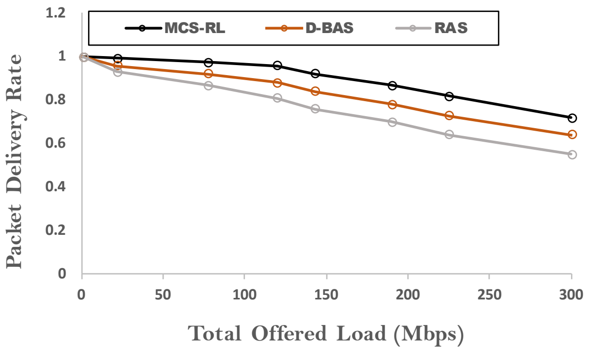

Figure. 7 The packet delivery rate versus the offered load

Table 4. System throughput of the proposed MCS-RL and reference approaches

Offered Load	RAS	D-BAS	MCS-RL
Network Throughput (Mbps)			
21.645	20.4	20.6	21.8
77.1134	53.60	61.85	74.10
120	81.25	93.75	112.25
143.2075	94.25	111.75	131.19
190	119.64	148.50	170.37
225	145.3	174.76	202.23
300	160.75	203.34	255.12

Table 5. Packet delivery rate of the proposed MCS-RL and reference approaches

Offered Load	RAS	D-BAS	MCS-RL
Packet Delivery Rate			
21.64	0.93	0.96	0.99
77.11	0.87	0.92	0.97
120	0.81	0.88	0.96
143.20	0.76	0.84	0.92
190	0.7	0.78	0.87
225	0.64	0.73	0.82
300	0.55	0.64	0.72

6. Conclusion

In our study, we proposed a novel multi-connection strategy to optimize the selection of the most efficient connectivity option. This strategy involves identifying and selecting the optimal BSs and APs along with their associate licensed and unlicensed sub-bands for uplink transmission. To achieve this, the proposed strategy incorporates RL and softmax decision-making methods, which enhance the overall decision-making process. This strategy is compared with a reference scheme that imposes exclusive connectivity to multiple base stations from different RATs, with dynamic sharing of resources using only licensed spectrum. In addition, the strategy has been compared against a reference scheme enables users to randomly select their radio access technology and resources based on signal strength. Our simulation outcomes validate the effectiveness and ability of the proposed multi-connection strategy to in efficiently allocating BSs/APs along with their associated sub-bands. This contributes to enhanced network performance in terms of in terms of utility function, network throughput, and packet success rate. Significantly, our proposed MCS-RL strategy outperforms reference schemes, exhibiting gains of up to 0.12, and 0.19 in achieved utility compared to the RAS and D-BAS schemes, respectively. The results also show substantial bit rate improvements, up to 16 % and 41% in our proposed MCS-RL approach compared to RAS and D-BAS, respectively. Moreover, the proposed algorithms effectively mitigate outage probability associated with increased traffic demands, achieving success delivery rates of approximately 0.10 and 0.18 compared to RAS and D-BAS, respectively.

Conflicts of Interest

The author declares and confirm that there is no conflict of interest to declare.

Author contributions

Haider Albonda comprehensively conducted the research, analyzed the data, contributed to writing the manuscript. Additionally, he actively covered all sections including the introduction, survey of the relevant literature, the system model, results, and discussions.

References

[1] J. Li, K. Nagalapur, and E. Stare, "5G New Radio for Public Safety Mission Critical Communications", *IEEE Communications Standards Magazine*, Vol. 3, pp. 1-8, 2021.

[2] 3GPP TS 22.282 V17.0.0, "Mission Critical Data (MCData); Stage 1", April. 2022.

[3] 4G Americas, "Developing and Integrating a High Performance HET-NET, Bellevue", *WA*, USA, Oct. 2012.

[4] A. F. Isnawati and M. A. Afandi, "Performance Analysis of Game Theoretical Approach for Power Control System in Heterogeneous Network", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 4, 2022, doi: 10.22266/ijies2022.0630.33.

[5] M. Yan, G. Feng, J. Zhou, and S. Qin, "Smart Multi-RAT Access Based on Multiagent Reinforcement Learning", *IEEE Transactions on Vehicular Technology*, Vol. 67, No. 5, pp. 4539-4551, May 2018.

[6] M. Gerasimenko "Adaptive Resource Management Strategy in Practical Multi-Radio Heterogeneous Networks", *IEEE Access*, February 2017.

[7] V. F. Monteiro, "Distributed RRM for 5G Multi-RAT Multiconnectivity Networks", *IEEE Systems Journal*, Vol. 13, No. 1, March 2019.

[8] A. Shamso, M. Khaledi, and F. Afghah, "A Solutionfor Dynamic Spectrum Management in Mission-Critical UAV Networks", *16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Boston, MA, USA, 2019.

[9] W. Ning, Y. Wang, and M. Liu "Mission-Critical Resource Allocation With Puncturing in Industrial Wireless Networks Under Mixed Services", *IEEE Access*, Vol. 9 , No. 1, 2021.

[10] Q. Cui, Y. Gu, and W. Ni, "Effective Capacity of Licensed-Assisted Access in Unlicensed Spectrum for 5G", *IEEE Journal on Selected Areas in Communications*, Vol. 35, No. 8, pp. 1754-1767, 2017.

[11] H. M. Ali, J. Liu, and W. Ejaz, "Planning capacity for 5G and beyond wireless networks by discrete fireworks algorithm with ensemble of local search methods", *EURASIP Journal on Wireless Communications and Networking*, pp. 1-24, 2020.

[12] T. Shiohara and T. Murase, "QoS Control for Mission critical Communication on Vehicles with IEEE802.11 Wireless LAN under Dynamic Interference", In: *Proc. of 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) New York*, NY, USA, 2020.

[13] H. D. Albonda, "Radio Control Optimization for Enhanced QoS-based Spectrum Selection in Heterogeneous 5G Environments with eMBB and uRLLC", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 6, 2023, doi: 10.22266/ijies2023.1231.45.

[14] Federal Communications Commission, "Second Report and Order: Expanding Flexible Use of the 3.7 to 4.2 GHz Band", *Federal Communications Commission*, 2018.

[15] ETSI TS 103 625, *Emergency Communications (EMTEL); Transporting Handset Location to PSAPs for Emergency Communications - Advanced Mobile Location*, V1.2.1, 2022.

[16] A. Dataesatu1, K. Sanada, and H. Hatano "System Performance Enhancement with Energy Efficiency Based Sleep Control for 5G Heterogeneous Cellular Networks", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 2, 2022, doi: 10.22266/ijies2022.0430.21.

[17] R. Kaur and S. Mittal "Enhanced Handoff Decision Making for Application-Aware Environment by Using Blended Approach", *International Journal of Intelligent Engineering and Systems*, Vol. 14, No. 1, 2021, doi: 10.22266/ijies2021.0228.40.

[18] R. Pirmagomedov, D. Moltchanov, and A. Samuylov, "Characterizing throughput and convergence time in dynamic multi-connectivity 5G deployments", *Computer Communications Journal*, Vol. 187, 2022.

[19] S. Alam, S. Sulistyo, and W. Mustika, "Utility-Based Horizontal Handover Decision Method for Vehicle-to-Vehicle Communication in VANET", *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 2, 2020, doi: 10.22266/ijies2020.0430.01.

[20] I. Cinemre, and T. Mahmoodi "Learning-Based Multi Attribute Network Selection in Heterogeneous Wireless Access", *Wireless Personal Communications*, Vol. 125, 2022.

[21] K. Naidu and V. Sathya, "Impressive Resources distribution in LAA/NR-U with Wi-Fi Concurrent Operation", In: *Proc. of IEEE IAS Global Conference on Emerging Technologies (GlobConET), London, United Kingdom*, pp. 1-6, 2023.

[22] V. Sheyhus, A. Furtado, and R. Oliveira "LTE LAA Coexistence with Wi-Fi: Performance Evaluation for Priority Classes", In: *Proc. of IEEE Conference on Standards for Communications and Networking (CSCN)*, Thessaloniki, Greece, pp. 1-6, 2022.

[23] S. M. Kala, V. Sathya, and K. Dahiya "Identification and Analysis of a Unique Cell Selection Phenomenon in Public Unlicensed Cellular Networks Through Machine Learning", *IEEE Access*, Vol. 10, No. 3, 2022.

[24] M. Haghshenas and M. Magarini "NR-U and Wi-Fi Coexistence Enhancement Exploiting Multiple Bandwidth Parts Assignment", In: *Proc. of IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, pp. 1-6, 2022.

[25] D. C Ventureira, F. G. Castaño, and F. J. Castiñeira "Coordinated Allocation of Radio Resources to WiFi and Cellular Technologies

in Shared Unlicensed Frequencies”, *IEEE Access*, Vol. 9, 2021.

[26] M. Han, S. Khairy, and L. X. Cai, “Reinforcement Learning for Efficient and Fair Coexistence Between LTE-LAA and Wi-Fi”, *IEEE Transactions on Vehicular Technology*, Vol. 69, Issue. 8, 2020.

[27] S. Mosleh, Y. Ma, and J. D. Rezac “Dynamic Spectrum Access with Reinforcement Learning for Unlicensed Access in 5G and Beyond”, In: *Proc. of IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Antwerp, Belgium, pp. 1-6, 2020.

[28] C. Wang, Y. Liu, H. Gao, and T. Y. Chiang “Reinforcement-Learning Based Radio Resources Allocation in Licensed Assisted Access”, In: *Proc. of IEEE International Conference on Smart Cloud (SmartCloud)*,Washington, DC, USA, pp. 1-6, 2020.

[29] M. Alhulay and M. L. Benitez “LTE/Wi-Fi Coexistence in Unlicensed Bands Based on Dynamic Transmission Opportunity”, In: *Proc. of IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Seoul, Korea (South), pp. 1-6, April 2020.

[30] S. Muhammad, H. H. Refai, and M. O. Al Kalaa “5G NR-U: Homogeneous Coexistence Analysis”, In: *Proc. of IEEE Global Communications Conference*, Taipei, Taiwan, pp. 1-6, 2020.

[31] R. Saleem, S. A. Alvi, and S. Durrani “Performance-Fairness Trade-off for Wi-Fi and LTE-LAA Coexistence”, *IEEE Access*, Vol. 9, April 2021.

[32] Z. Qin, A. Li, and H. Wang “Modeling and Performance Analysis of LTE Coexisting with Wi-Fi”, In: *Proc. of IEEE 8th International Conference on Information, Communication and Networks (ICICN)*, Xi'an, China, pp. 1-6, 2020.

[33] H. Albonda, K. Majdi, and B. Abbas, “Development Capacity Sharing Strategy for Vehicular Networks with Integrated Use of Licensed and Unlicensed Spectrum”, *EUREKA, Physics and Engineering Journal*, No. 3, pp. 149-158, 2022.

[34] 3GPP TR 38.101-1, 5G; NR, *User Equipment (UE) radio transmission and reception*, Part 1: Range 1 Standalone, Version 18.3.0. Sep. 2023.

[35] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Cambridge, MA, USA: MIT Press, 1998.

[36] 3GPP TS 36.213, *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*, V17.0.0, 2023.