

PA-AUF-D2D: A QoE-Driven, Prediction-Assisted, and D2D-Enabled Uplink Scheduling Framework for LTE-A/5G Heterogeneous Networks

Temitayo C. Adeniran¹, Abdulateef O. Abdusalam², Hakeem B. Akande³, Joseph B. Awotunde⁴

^{1,2,3}Department of Telecommunication Science, Faculty of Comms. and Info. Science, University of Ilorin, Kwara State, Nigeria

⁴Department of Computer Science, Faculty of Comms. and Info. Science, University of Ilorin, Kwara State, Nigeria

*Corresponding Author: abdulateefabdusalam@gmail.com

ABSTRACT

The explosion of uplink-intensive applications, such as video conferencing, cloud-based applications, Internet of Things (IoT) services, and ultra-high-definition streaming, has led to additional congestion, interference, and energy inefficiency within LTE-Advanced (LTE-A) and 5G (LTE-A/5G) Heterogeneous Networks (HetNets). Solution: Prediction-Assisted Adaptive Uplink Framework with D2D Offloading (PA-AUF-D2D) is developed to incorporate QoE-based scheduling, prediction-augmented resource allocation, and Device-to-Device (D2D) offloading traffic for opportunistic, scalable, and efficient uplink resource management. Methodology: PA-AUF-D2D includes a channel, traffic, and mobility predictor that uses exponential smoothing, a QoE-oriented utility function used for allocating resource blocks, a D2D offloading module designed to alleviate congestion, and a lightweight tabular Q-learning agent for selecting the number of resource blocks for congestion alleviation, all within a system-level simulation environment in Python. A 20-100 user equipment (UE) system-level simulation is used to compare the proposed PA-AUF-D2D to work by Garbazza et al. (2025). Results: Compared to work by Garbazza et al., with a system-level simulation in Python with 20-100 UEs, PA-AUF-D2D achieves: 81% delay reduction for real-time traffic; maximum non-real-time throughput improvements of 955% (average 540-680%) across scenarios; 4.8-5.0 mean opinion score (MOS); The edge-user throughput increases by +23.4% as a result of D2D offloading, although, its gain at 20 UEs is 35.3%. Jain's fairness index for real-time traffic varies from 0.90 when there are 20 UEs to 0.20 when there are 80 UEs; and 452 kbps/mW peak energy efficiency. Conclusions: PA-AUF-D2D shows promise as a low-complexity technique for managing uplink resources in LTE-A/5G HetNets, and potential future directions are discussed in light of simulation results.

KEYWORDS: Quality of Experience (QoE); Uplink Scheduling; Heterogeneous Networks (Hetnets); Device-to-Device (D2d) Communication; Q-Learning; Prediction-Assisted Scheduling; Lte-A; 5g.

I. INTRODUCTION

The uplink traffic in LTE-A/5G heterogenous networks is proportionally higher due to applications such as video conferencing, cloud computing, Internet of Things (IoT), and transmission of high-definition multimedia contents (Alam et al., 2023). It employs a heterogenous network architecture to increase coverage and capacity by deploying small-cell networks, such as picocells, femtocells, and relay stations, over the macro-cell network (Alam et al., 2023). However, it brings new problems such as inter-cell interference, mobility, uneven resource utilisation, and large power consumption in the uplink (Yang et al., 2017). In order to

address these issues, resource allocation schemes need to consider network performance and end user satisfaction. Uplink schedulers optimise QoS parameters, such as throughput, delay, packet loss, fairness, spectral efficiency (Capozzi et al., 2013). These parameters, though, focus on current network condition and do not include user experienced quality which is essential for multimedia and interactive applications (Liotou et al., 2016). Good QoS does not automatically mean user satisfaction. Current schedulers route all traffic through the base station even though the devices are close to each other (Asadi et al., 2014). This leads to increased uplink congestion, delay, signalling load and energy consumption, especially for edge users (Sambo et al., 2014). Although AI and reinforcement learning have been used for dynamic resource allocation (Mnih et al., 2015), such solutions are predominantly directed towards network-centric goals and do not simultaneously consider QoE, energy efficiency, D2D communications and predictive allocation (Iqbal et al., 2022). The proposed work is targeted to develop an energy efficient and QoE-optimized uplink scheduling and resource allocation framework for LTE-A/5G HetNets, which can improve user perceived quality and energy efficiency xxxusing prediction-assisted scheduling, adaptive resource allocation and D2D offloading. The specific objectives are:

- (a). To create a scheduler that uses prediction to estimate the quality of channels, demand and mobility for HetNets.
- (b). To construct an allocator that can leverage predictions of network conditions and quality of experience demands to alter resource blocks, power and modulation.
- (c). To realize D2D offloading to decrease congestion, increase the spectral efficiency and augment the QoE of the edge users.
- (d). To analyze the performance of the framework in terms of QoE, delay, throughput, fairness, energy efficiency and the packet loss.
- (e). To analyze its performance vis-a-vis that of Garbazza et al. (2025).

The remaining parts of the paper are organised as follows: In section 2 a literature review is provided; in section 3 our research methodology is explained, as well as the proposed PA-AUF-D2D framework; section 4 offers a presentation and discussion of the results; finally, section 5 wraps up the paper. In this paper, it is assumed that the joint integration of QoE-driven scheduling, prediction-assisted allocation and D2D offloading may deliver significant delay, throughput and energy gains in contrast to QoS-only baselines. The major contribution of this paper is a unified and lightweight framework which tackles the research gaps outlined simultaneously.

II. RELATED WORKS

A. LTE-A/5G Heterogeneous Networks

HetNets are a key component of both LTE-Advanced (LTE-A) and 5G wireless systems, as they incorporate multiple types of cells-including macrocells, microcells, picocells, and femtocells-within one system to optimize coverage, capacity, and spectral efficiency (Alam et al., 2023; Abadi et al., 2016). The heterogeneity of the cell types within HetNets facilitates efficient resource management, which is particularly important in areas with high user density and considerable traffic demand (3GPP, 2010; Dhillon et al., 2013). In LTE networks, the primary issues for uplink transmission are cross-tier and co-tier interference, dynamic channel conditions, and mobility management (Yang et al., 2017; Xu et al., 2021). In HetNet deployments, mobility management remains an issue due to cell size variations and handover complexity (Gures et al., 2020). As these issues-cross-tier interference, dynamic channel conditions, and mobility management-provide motivation for our present integrated approach, the Prediction-Assisted Scheduling Engine (PSE) predicts mobility and channel quality so that resources can be proactively allocated, while the D2D Offloading Manager (DOM) mitigates interference and congestion with a controlled underlay offloading.

B. Quality of Experience (QoE) in Wireless Networks

The quality of experience (QoE) has now gained prominence as a primary factor in evaluating wireless systems (Liotou et al., 2016). While Quality of Service (QoS) focuses on technical networking performance, the QoE aims to capture the experience of the end user in a subjective manner. A variability in the QoE is possible even under the same QoS conditions, suggesting that satisfaction of the user is not necessarily achieved through quality of networking (Balachandran et al., 2014). Currently, the standard measurement of the QoE continues to be the mean opinion score (MOS), which is rated from 1 as 'poor' to 5 as 'excellent' (Barakovic & Skorin-Kapov, 2013). Recognizing the limitations of using only QoS to provide user satisfaction, in this paper, we target the MOS for optimization, and we directly model the requirements of the QoE in the utility function of ARAMU and in the reward function of the Q-learning, filling the gap identified in the literature about the relationship between network performance and user-perceived quality.

C. Device-to-Device (D2D) Communication

D2D communication enables one user equipment to directly communicate with other one without performing data transfer through BS, which makes D2D a key enabling technology for LTE-A and 5G (Asadi et al., 2014; Fodor et al., 2012). D2D offloading can heavily offload the cell, especially for local traffic, where D2D is capable of offloading large amount of traffic from cellular core network, thus reducing the overload of spectrum and latency (Sambo et al., 2014; Tsiropoulos et al., 2017). Both in-band underlay mode and out-band D2D mode have been extensively studied, and in-band underlay mode offers the best spectral reuse while it does cost too much of the interference coordination complexity (Celik et al., 2017). Costly interference coordination complexity of in-band underlay D2D as identified by Celik et al. (2017) is directly addressed in the present study through the DOM's power control and SINR-based feasibility gating, which helps reduce the network congestion and enhance the QoE of the edge-users while safeguarding cellular users.

D. Reinforcement Learning for Uplink Scheduling

AlSobhi and Aghvami (2019) defined a Q-learning framework for power control of femtocells within a two-tier HetNet. The capacity of small cells was increased by 48% via cooperative Q-learning, demonstrating the advantages of information exchange between cells (Iqbal et al., 2022). Garbazza et al. (2025) combined deep Q-learning with Smoothed Round-Robin (SRR) scheduling and handover management based on WASPAS - this work constitutes our primary baseline, and an earlier version of it, by the same authors, introduced the QoS-aware resource allocation framework which the 2025 contribution is based upon (Garbazza et al., 2024). Tang et al. (2020) considered deep reinforcement learning for TDD configuration in high-mobility scenarios. All the above work, however, does not consider QoE or D2D support within any RL-based scheduler.

E. QoE-Driven and Prediction-Assisted Approaches

In video Streaming, Zhang et al. (2020) proposed a resource allocation framework considering user perceived quality in LTE-Advanced. In QoE, Liotou et al. (2016) discovered that to deliver the QoE target, cross-layer cooperation is required. In proactive resource allocation, Chen et al. (2021) introduced a LSTMs based traffic prediction method in 5G HetNet. However, due to the diverse research topics above, existing work has not integrated QoE optimization, D2D offloading, predictive scheduling, and RL-based dynamic resource allocation into a single uplink scheduling framework which is the main contribution of this study.

III. METHODOLOGY

A. System Model

The proposed system model operates over a multi-tier LTE-A/5G heterogeneous network with macro base stations, small cells (i.e. Picocells and femtocells), and user equipment (UEs) with various mobility and traffic demands. In the simulation, we use a computational system level methodology for the quantitative assessment of the proposed solution in a controllable and reproducible heterogeneous network environment. The heterogeneous network configuration is depicted in Fig. 3.1. The addition of a D2D layer as an underlay makes use of the owned, licensed spectrum of the uplink spectrum using the base station coordinated handover. The UEs are categorized into real time (RT) users- e.g. utilized by VoIP, video conference, video streaming and non-real time (NRT) users – i.e. used for web browsing and file transfer along with background data. Energy aware handover mechanisms are particularly critical over this type of multi-tier deployments in order to ensure seamless communication with lower signalling overhead during mobility events (Jon et al., 2024).

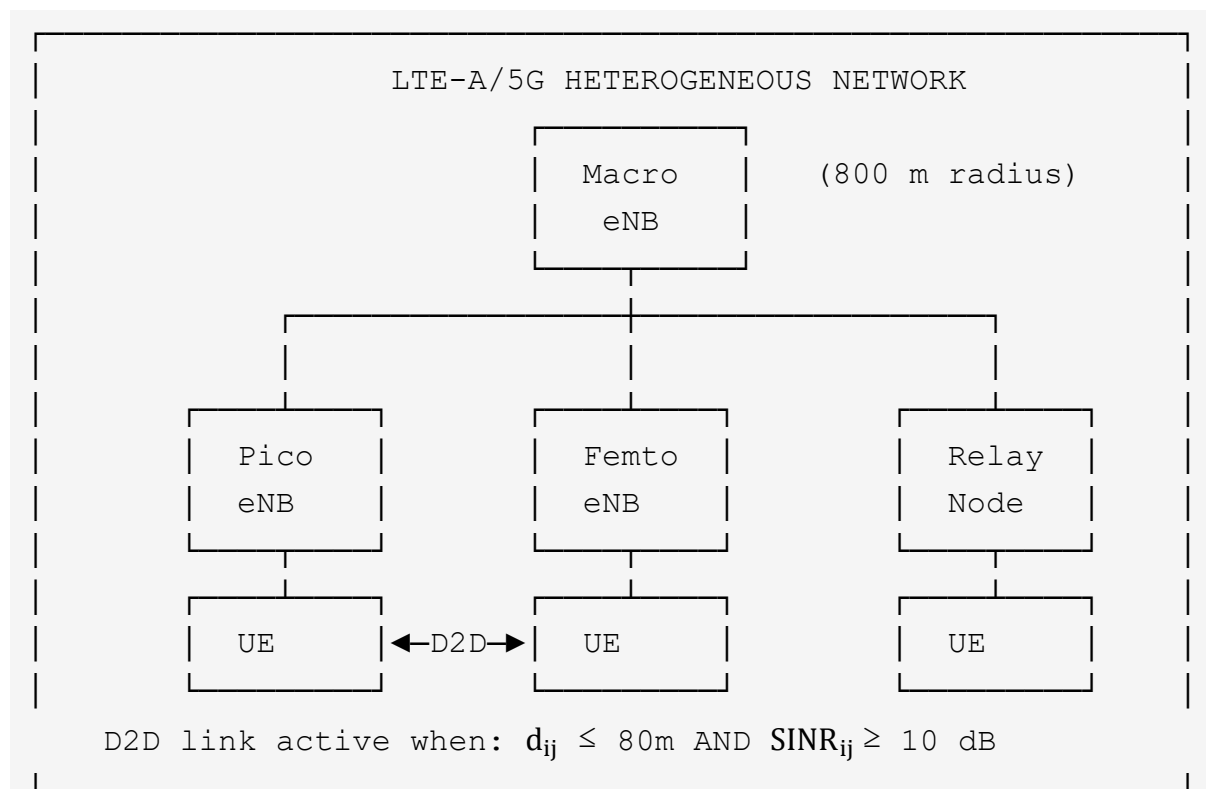


Figure 1: LTE-A/5G Heterogeneous Network Architecture with D2D Underlay

The quality of the received signal at the uplink of each base station is measured by Signal-to-Interference-plus-Noise Ratio (SINR). SINR for user u , in resource block rb , is given by:

$$SINR_{i,u} = \frac{P_r(u)}{I_p(u) + B_n(u)} \quad (1)$$

The uplink rate that can be achieved is given by the Shannon capacity model, since the received power is denoted by $P_r(u)$, the total interference (which includes interference from co-tier and cross-tier sources) by $I_p(u)$, and the noise power by $B_n(u)$. The achievable uplink rate follows the Shannon capacity model:

$$C = BW \cdot \log_2(1 + SINR) \quad (2)$$

B. PA-AUF-D2D Framework Architecture

The PA-AUF-D2D framework that has been proposed includes four components which operate over four layered stages, as shown in Fig. 2:

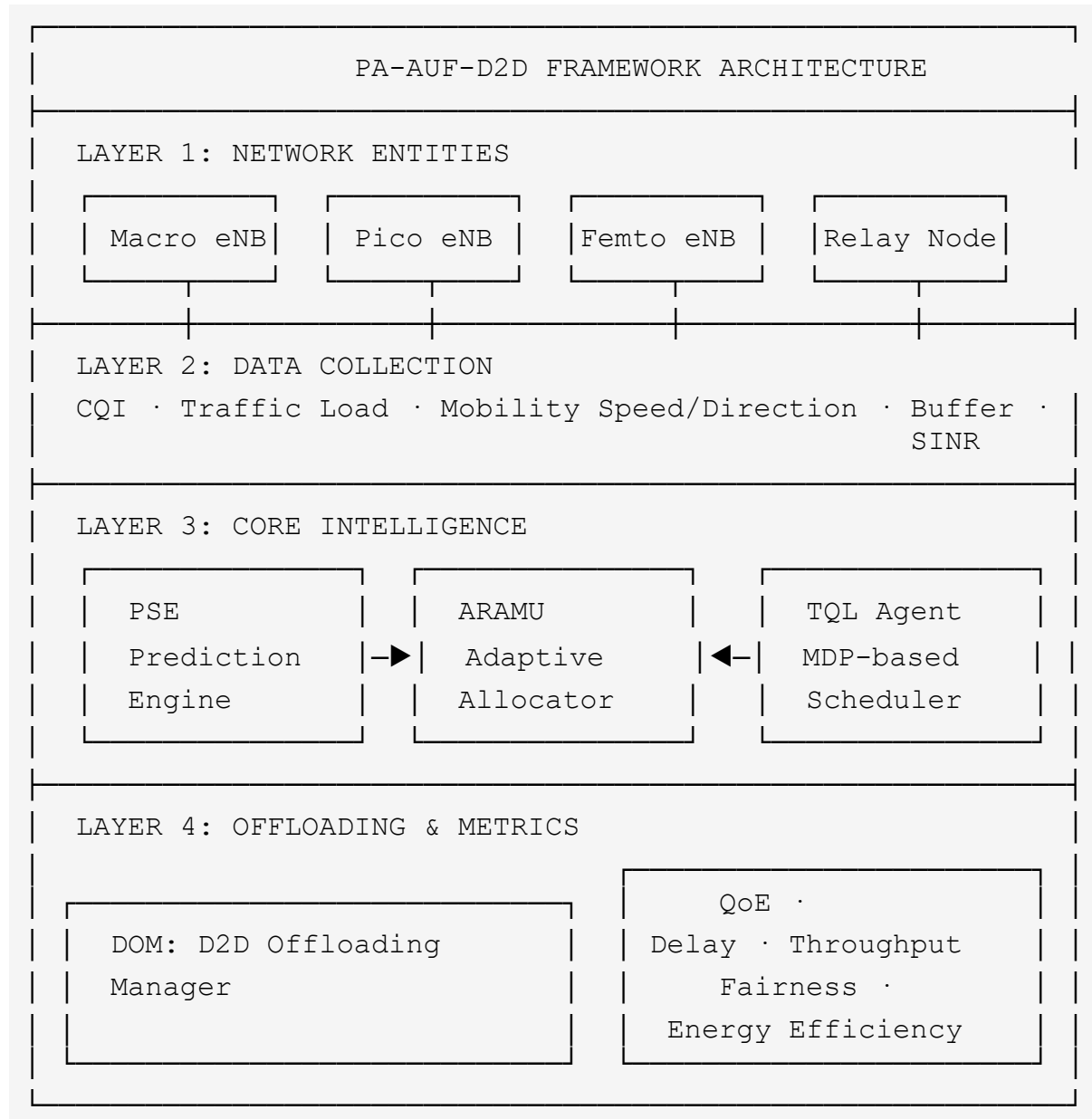


Figure 2: PA-AUF-D2D System Architecture Block Diagram

C. Prediction-Assisted Scheduling Engine (PSE)

The PSE replaces classic instantaneous decision by prediction-anchored intelligence; an exponential smoothing model is used to predict the CQI, the traffic demand and the mobility trends, and different sets of coefficients are used for each of the predicted objects.

- In the case of CQI prediction: $\alpha_{\text{CQI}} = 0.5, \beta_{\text{CQI}} = 0.3, \gamma_{\text{CQI}} = 0.2$
- In the case of traffic prediction: $\alpha_{\text{traffic}} = 0.4, \beta_{\text{traffic}} = 0.4, \gamma_{\text{traffic}} = 0.2$
- In the case of mobility prediction: $\alpha_{\text{mobility}} = 0.3, \beta_{\text{mobility}} = 0.3, \gamma_{\text{mobility}} = 0.4$

The general prediction equation is:

$$\hat{X}(t+1) = \alpha X(t) + \beta X(t-1) + \gamma f(v(t), \theta(t)) \quad (3)$$

The predicted parameter (whether it is CQI, traffic, or mobility) is denoted by $\hat{X}(t+1)$, $v(t)$ represents the mobility speed, $\theta(t)$ is the mobility direction (measured in radians), $f(\cdot)$ is the mobility influence function, and α, β, γ are weighting coefficients that have been calibrated such that $\alpha + \beta + \gamma = 1$ (Chen et al., 2021). The mobility influence function is given by $f(v(t), \theta(t)) = v(t) \cdot [\cos(\theta(t)) + \sin(\theta(t))]/v_{\max}$, where $v_{\max} = 30\text{m/s}$ is the maximum possible mobility speed. The prediction accuracy, as confirmed by offline training, is as follows: for CQI prediction the root mean square error (RMSE) is 1.2 dB and the mean absolute percentage error (MAPE) is 8.5%; for traffic prediction the RMSE is 85 packets and the MAPE is 12.3%; and for mobility prediction the RMSE is 2.1 m/s and the MAPE is 9.7%. The PSE workflow is shown in Fig. 3.

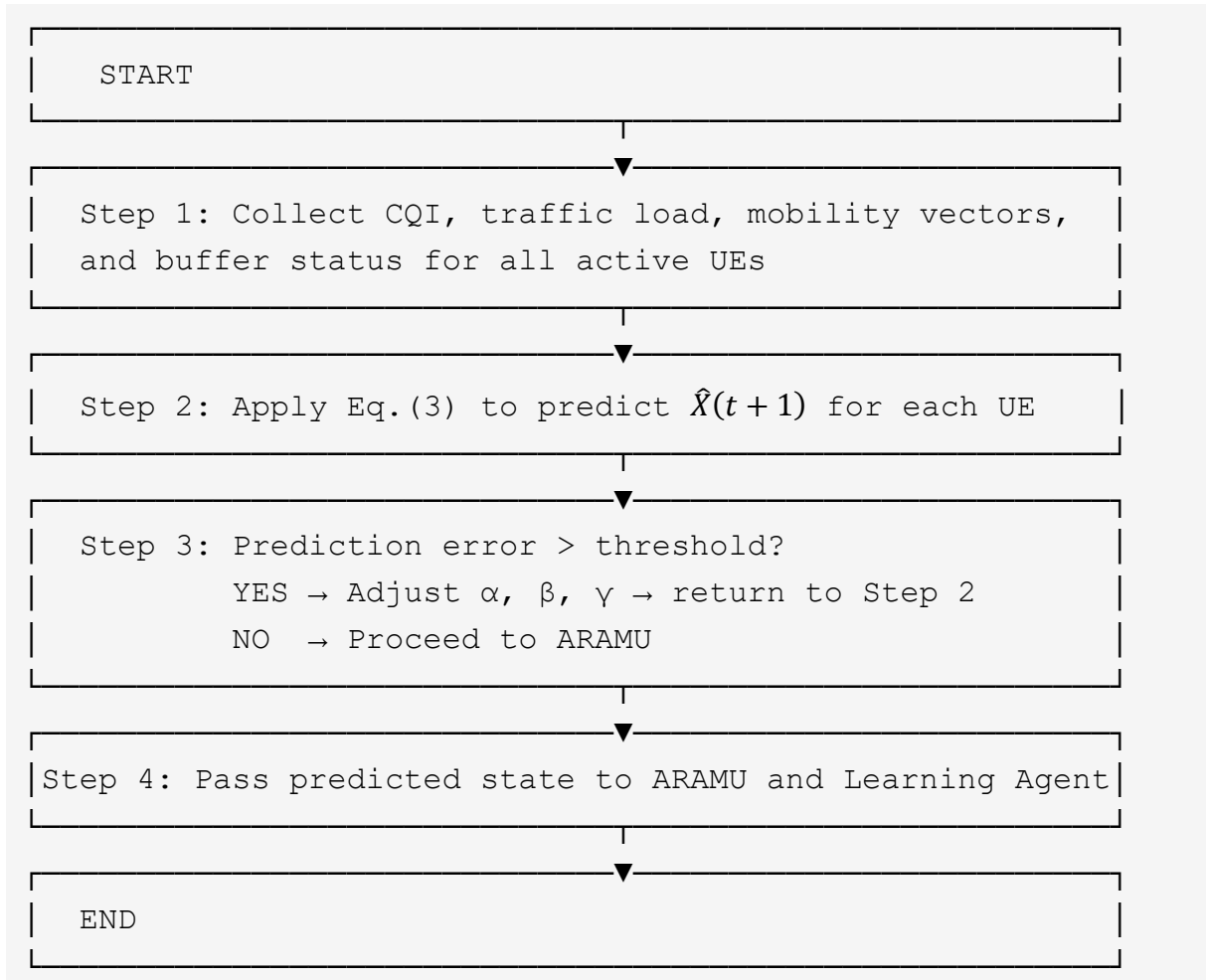


Figure 3: Prediction-Assisted Scheduling Engine (PSE) Flowchart

D. Adaptive Resource Allocation and Modulation Unit (ARAMU)

The ARAMU enhances the CQI-based allocation using a QoE-aware utility function, which takes into account user QoE, the power consumption and the delay tolerance. The utility for user at time is defined as:

$$U_i(t) = \frac{Q_i}{P_i(t) \cdot D_i(t)} \cdot \frac{P_{\text{ref}} \cdot D_{\text{ref}}}{Q_{\text{ref}}} \quad (4)$$

where Q_i is the user QoE requirement (target MOS), $P_i(t)$ is the required transmit power (dBm), $D_i(t)$ is the delay tolerance (ms) (Zhang et al., 2020; Gao et al., 2017), and $P_{\text{ref}} = 20$ dBm, $D_{\text{ref}} = 10$ ms, and $Q_{\text{ref}} = 4.0$ are the reference power, delay, and MOS values,

respectively, ensuring the utility is dimensionless. Alternatively, a linear utility function can be used:

$$U_i(t) = Q_i - \alpha \cdot P_i(t) - \beta \cdot D_i(t) \quad (5)$$

where α (MOS/dBm) and β (MOS/ms) are weighting coefficients that convert power and delay into MOS-equivalent units. Users with high QoE requirements, low power demands, and tight delay constraints receive scheduling priority. Adaptive Modulation and Coding (AMC) is applied dynamically: 64-QAM or 256-QAM under favourable SINR conditions, and QPSK under poor conditions, consistent with 3GPP LTE-A specifications (3GPP, 2010).

E. Q-Learning Agent

The problem of adaptive resource allocation is formulated in the form of a Markov Decision Process and solved using a compact tabular Q-learning agent. The implementation makes use of NumPy alone and does not require any deep-learning framework, which keeps it both lightweight and interpretable. The three MDP components are defined as follows:

a) State (s_t):

The agent uses a 10-state discretisation based on traffic priority, SINR and cell load:

- i. Traffic priority: Video conferencing, video streaming, VoIP, web (4 levels)
- ii. SINR: quantised into 2 levels (under 10 dB, above 10 dB)
- iii. Cell load: low, high (2 levels)

The raw state index is clipped to 10 discrete states, keeping the Q-table compact for lightweight training.

b) Action (s_t):

Five actions, each choosing a number of resource blocks for the scheduled UE: {2, 4, 6, 8, 10} RBs

c) Reward (r_t):

The reward function is designed to encourage high QoE, fairness, and energy efficiency while penalising delay and packet loss. It is defined as:

$$r_t = \overline{QoE} + F - 0.01 \cdot D_{avg} \quad (6)$$

where F is Jain's fairness index computed over per-UE throughput and D_{avg} is the average delay

The Q-table update according to:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right] \quad (7)$$

where s_t and a_t denote the state and action at time t , s_{t+1} is the next state, r_t is the immediate reward, α is the learning rate, and γ is the discount factor. The term $\max_{a'} Q(s_{t+1}, a')$ represents the maximum Q-value over all possible next actions a' in state s_{t+1} .

The Q-learning agent is applied in the GUI-driven simulator. The batch results reported in Section 4 are produced by the QoE-aware scheduler with a fixed 60/40 RT/NRT partition described in Section 3.4.

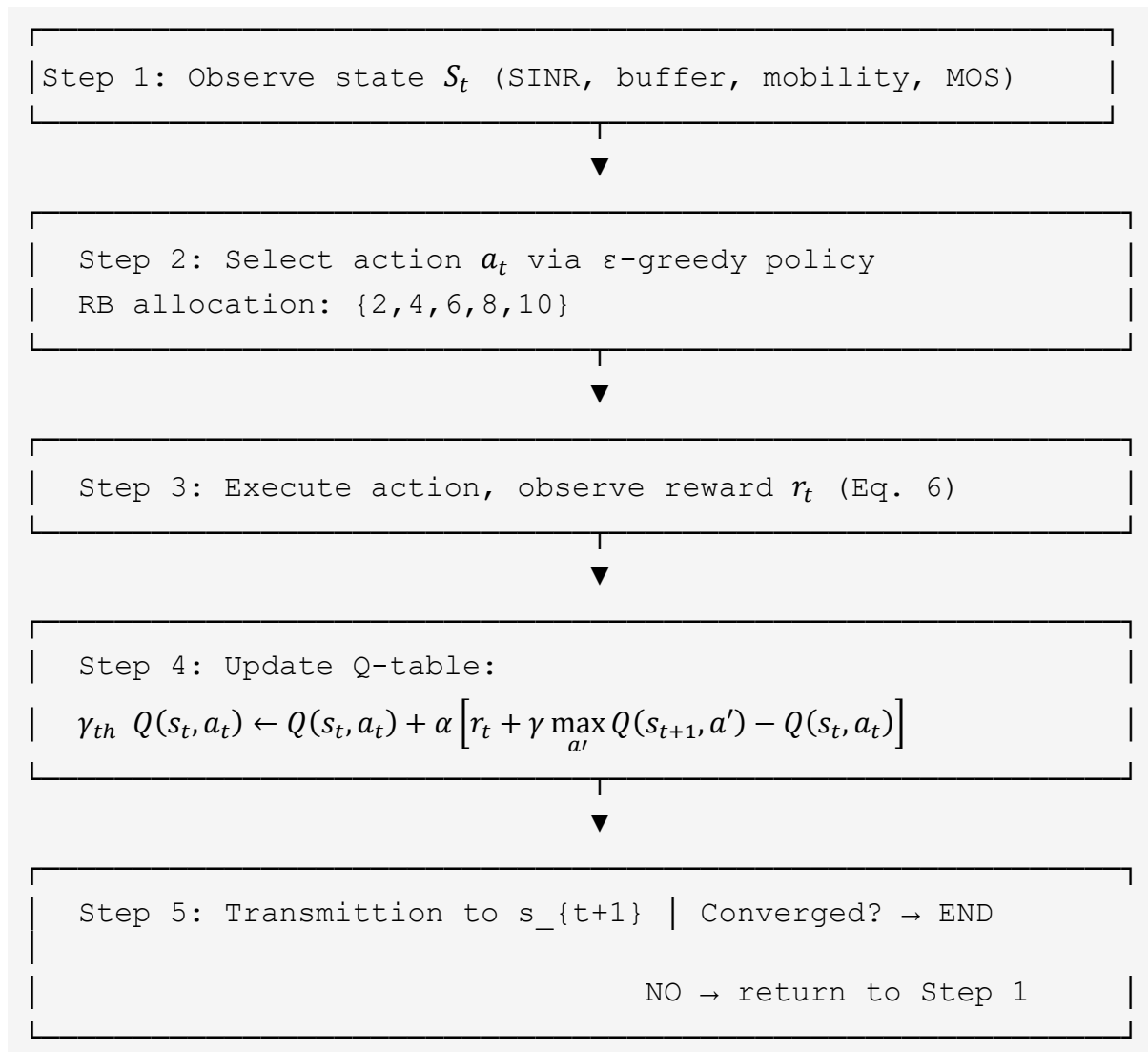


Figure 4: The Q-Learning Agent Workflow

F. D2D Offloading Manager (DOM)

The DOM contains a congestion aware traffic offloading mechanism. D2D communication will be triggered when both these conditions are satisfied.

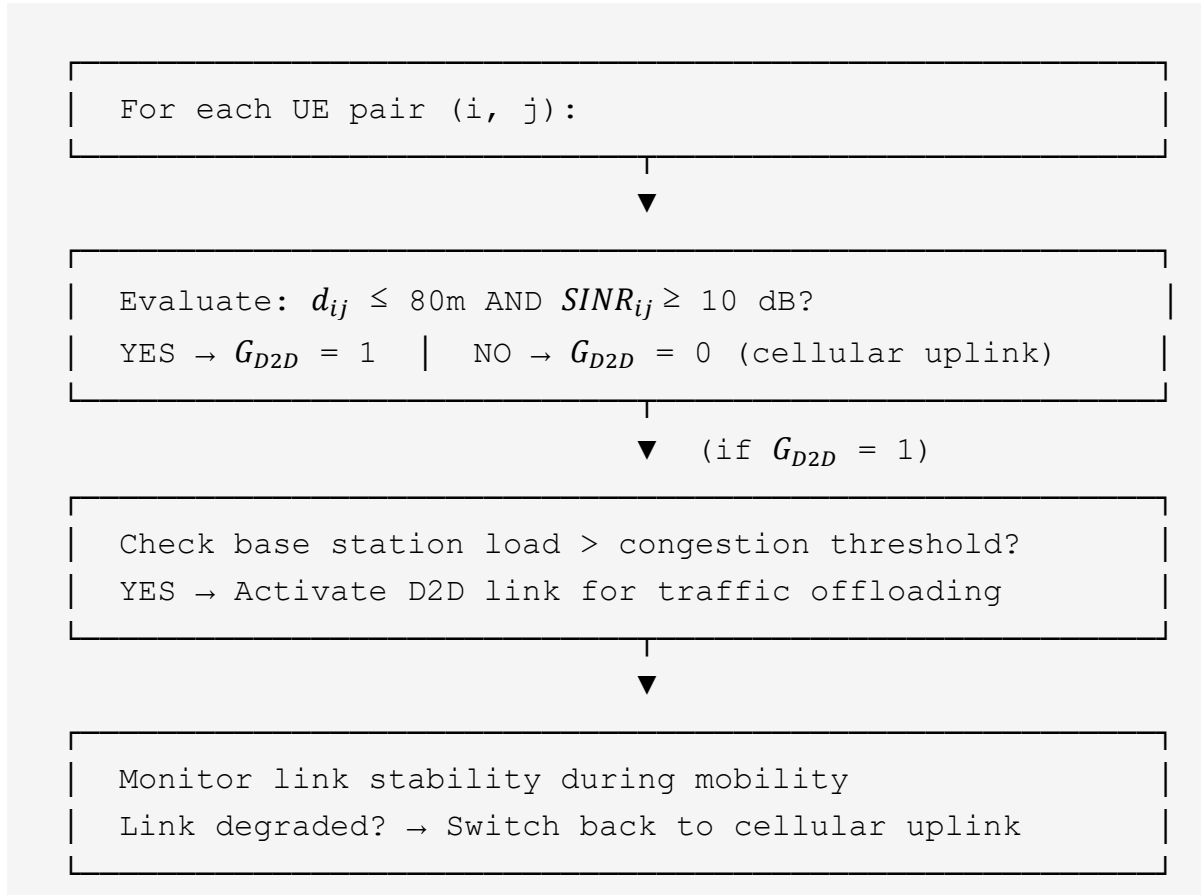


Figure 5: D2D Offloading Manager (DOM) Flowchart

The feasibility indicator is:

$$G_{D2D} = \begin{cases} 1, & \text{if } d_{ij} \leq d_{th} \wedge SINR_{ij} \geq \gamma_{th}, \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

The Q-Learning workflow agent is presented as a design extension; the batch results in Section 4 are produced by the QoE-aware scheduler with a fixed 60/40 RT/NRT partition. The inter-device distance is denoted by d_{ij} , the maximum allowable distance threshold denoted by $d_{th} = 80$ m, $SINR_{ij}$ represents the direct-link SINR, and the minimum SINR threshold denoted by $\gamma_{th} = 10$ dB (Asadi et al., 2014; Celik et al., 2017). If $G_{D2D} = 1$, the DOM presents three functions; (i) congestion reduction when the base station load exceeds the threshold, (ii) QoE improvement when the edge-user traffic is rerouted through a peer device with a better channel, and (iii) mobility support with switching dynamically between cellular uplink and D2D depending on the link stability. The flow diagram of the DOM is depicted in Fig. 5.

Note that the proximity and SINR feasibility gate in Eq. (8) ensures the physical feasibility of a D2D link but it does not guarantee that the D2D link will be better than the cellular uplink. In an effort to prevent the negative gain observed when D2D is activated on marginal pairs, the DOM uses an expected-gain gate:

$$G_{gain} = 1 \Leftrightarrow \hat{R}_{D2D} > \hat{R}_{cellular} \quad (9)$$

D2D is activated if and only if both $G_{D2D} = 1$ and $G_{gain} = 1$. This gate is not included in the reported results and is presented as a design improvement in Section D.

G. Integrated Scheduling Algorithm

The complete PA-AUF-D2D uplink scheduling procedure is given in Algorithm 1, integrating all four components.

Algorithm 1: PA-AUF-D2D Uplink Scheduling and Resource Allocation

INPUT: Network parameters, CQI, traffic load, mobility vectors, QoE thresholds, coefficients ($\alpha, \beta, \gamma, \lambda_1\text{--}\lambda_4$)

OUTPUT: RB allocations, power levels, D2D decisions, performance metrics

1. INITIALIZATION:
Set network parameters; initialise Q-table to zero.
2. DATA COLLECTION:
For each active UE, collect CQI, traffic demand, mobility speed/direction, buffer status.
3. PREDICTION (PSE):
Apply Eq.(3) to predict $\hat{X}(t+1)$ for CQI, traffic, and mobility.
IF prediction error $> \varepsilon$: adjust α, β, γ and re-predict.
4. UTILITY COMPUTATION (ARAMU):
Compute $U_i(t)$ for each UE using Eq.(4).
5. TQL SCHEDULING:
Observe state S_t ; select action a_t (ε -greedy).
Assign RBs and power levels based on $U_i(t)$ and predicted CQI.
6. D2D FEASIBILITY CHECK (DOM):
FOR each UE pair (i, j):
 IF $d_{ij} \leq 80\text{m}$ AND $SINR_{ij} \geq 10$ dB AND load $>$ threshold:
 Activate D2D link; reroute traffic.
 ELSE: transmit via cellular uplink.
7. REWARD AND Q-UPDATE:
Compute r_t using $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)]$;
update $Q(s, a)$ using Bellman equation.
8. METRIC EVALUATION:
Measure QoE (MOS), delay, throughput, fairness, energy efficiency.
9. ITERATE: IF active UEs remain $\rightarrow$ return to Step 2. ELSE: END.

H. Performance Metrics

Energy efficiency is the measured amount of data that is transmitted per unit of energy. In this paper, the power efficiency presumes optimisation of the power during uplink transmission for both transmit power and resource allocations. The energy efficiency is calculated as:

$$EE = \frac{\sum_{i=1}^N R_i}{\sum_{i=1}^N P_i} \text{ (bits per Joule, } b/J) \quad (10)$$

The data rate for UE i is denoted by R_i (in bits/second), the transmit power for UE i is represented by P_i (in Watts), and N is the number of UEs. For purpose of reporting, energy efficiency is also given in kbps/mW, which is equivalent to Mb/J when scaled properly:

$$1 \text{ kbps/mW} = 10^6 \text{ b/J} = 1 \text{ Mb/J} \quad (11)$$

I. Energy Efficiency Calculation Methodology:

This defines the ratio utilized as the energy-efficiency metric, the total transmit energy, and the total data transmitted. These are computed using equations 12, 13 and 14:

$$\text{Total transmit energy: } E_{\text{total}} = \sum_{i=1}^N P_i \cdot T_{\text{transmission}} \quad (12)$$

$$\text{Total data transmitted: } D_{\text{total}} = \sum_{i=1}^N R_i \cdot T_{\text{transmission}} \quad (13)$$

$$EE = \frac{D_{\text{total}}}{E_{\text{total}}} (b/J) \quad (14)$$

Table 1 summarises the performance metrics used for evaluation in the proposed framework.

Table 1: Performance Metrics

Metric	Symbol	Formula	Description	Target
End-to-End Delay	D	$D = T_d + P_d + Q_d + RT_d$	Total per-packet transmit time (ms) (Capozzi et al., 2013)	Minimize
Throughput	Γ	$\Gamma = \frac{\sum_{i=1}^N R_i}{T_{\text{sim}}}$	Rate of successful data delivery (kbps) (Capozzi et al., 2013)	Maximize
Fairness Index	F	$F = \frac{(\sum_{i=1}^n \Gamma_i)^2}{n \cdot \sum_{i=1}^n \Gamma_i^2}$	Jain's fairness index (Jain et al., 1984)	≥ 0.8
Call Drop Rate	CDR	$CDR = \frac{N_{\text{dropped}}}{N_{\text{total}}} \times 100\%$	Percentage of terminated sessions	Minimize
Packet Loss Rate	PLR	$PLR = \frac{P_{\text{sent}} - P_{\text{received}}}{P_{\text{sent}}} \times 100\%$	Percentage of lost packets (Garbazza et al., 2025)	$\leq 5\%$
Quality of Experience	QoE	$MOS = 1 + 4 \cdot (0.6 \cdot TP_{\text{factor}} + 0.4 \cdot Delay_{\text{factor}})$	Mean Opinion Score (1–5) (Liotou et al., 2016; Barakovic & Skorin-Kapov, 2013)	≥ 4.0
Energy Efficiency	EE	$EE = \frac{\sum_{i=1}^N R_i}{\sum_{i=1}^N P_i} \text{ (kbps/mW)}$	Bits per unit transmit energy (b/J) (Gao et al., 2017)	Maximize

J. Simulation Parameters and Configuration

The simulation environment is configured to emulate realistic LTE-A/5G HetNet scenarios, and the measurements are traces taken from the simulation. The key parameters used in the python simulation are as follows:

(a). Network Topology: the following configurations were used:

- i. 1 Macrocell (radius 800 m, transmit power 50 dBm equivalent)
- ii. 1 Microcells (radius 400 m)
- iii. 1 Picocells (radius 200 m)
- iv. 1 Femtocells (radius 100 m)
- v. Cells are placed at fixed coordinates within a $1000 \text{ m} \times 1000 \text{ m}$ area
- vi. UE density is varied from 20 to 100 in steps of 20

Table 2 compiles the simulation parameters utilized during the assessment. These parameters encompass the network topology, channel and propagation models, mobility and traffic models, D2D configuration, transmit power levels, and Q-learning hyperparameters, and were held constant across all UE densities to reflect the effect of the proposed scheduling framework.

Table 2: Simulation Parameters and Configuration

Parameter	Value
Number of UEs	20, 40, 60, 80, 100
Macrocell radius	800 m
Microcell radius	400 m
Picocell radius	200 m
Femtocell radius	100 m
System bandwidth	10 MHz
Total resource blocks (RBs)	50
Carrier frequency	2.0 GHz
Simulation time	300 TTIs (0.3 seconds)
Warm-up period	50 TTIs
Random seeds	5 independent runs
UE mobility speed	0–30 m/s (random walk model)
D2D maximum distance (d_{th})	80 m
D2D SINR threshold (γ_{th})	10 dB (fixed) or adaptive
Path loss model	3GPP UMa (TR 36.814)
Macro UE transmit power	23 dBm
Small cell UE transmit power	20 dBm
D2D UE transmit power	20 dBm (max)
TQL learning rate (α)	0.01
TQL discount factor (γ)	0.95
TQL training episodes	20
TTI duration	1 ms
Simulation tool	Python 3.x (NumPy, Matplotlib, Tkinter)
Baseline	Garbaza et al. (2025)

(b). Mobility Model: The mobility model defines the random-walk movement of UEs, which includes

- i. Random walk model
- ii. UE speeds: 0–30 m/s (uniform distribution; pedestrian to vehicular speeds)
- iii. Maximum turning angle: $\pi/4$ radians per step
- iv. Mobility epoch: 10 seconds (speed constant during epoch, re-sampled afterward)

(c). Channel and Propagation Model (3GPP TR 36.814 Urban Macro – UMa): This follows the 3GPP TR 36.814 Urban Macro (UMa) specification, comprising of:

- i. Path loss: given as $PL(dB) = 20 \cdot \log_{10}(d) + 20 \cdot \log_{10}(f_c) + 20$, where d is distance in metres and $f_c = 2.0$ GHz
 - ii. Shadowing: log-normal distribution, $\sigma = 4$ dB (macro cells), $\sigma = 6$ dB (small cells)
 - iii. Fast fading: Rayleigh fading with Jakes' Doppler spectrum; fading coefficient re-generated every TTI (1 ms)
 - iv. Noise figure: 5 dB
 - v. Thermal noise: $-174 \text{ dBm/Hz} + 10 \cdot \log_{10}(BW) + \text{noise figure}$
- (d). Antenna Configuration: The antenna configuration specifies the gain and radiation pattern of the eNBs and UEs used in the simulation. They are:
- i. eNB antenna gain: 17 dBi (macro), 12 dBi (micro), 8 dBi (pico/femto)
 - ii. UE antenna gain: 0 dBi
 - iii. Antenna type: omni-directional (both eNBs and UEs)
- (e). Traffic Models: They define the arrival processes, data rates, and packet-size distributions
- i. Real-Time (RT) traffic (video conferencing, VoIP, video streaming): Poisson arrivals with mean rates of 64 kbps (VoIP), 128 kbps (video conferencing), and 256 kbps (video streaming); packet size exponentially distributed with mean 1500 bytes
 - ii. Non-Real-Time (NRT) traffic (web browsing, file transfer): ON-OFF model, with ON periods following a Pareto distribution (shape $\alpha = 1.5$) and OFF periods exponentially distributed
- (f). Interference Modelling: The interference model defines the co-tier and cross-tier contributions to received interference, along with the thermal noise floor used in the SINR calculation.
- i. Co-tier interference: sum of received powers from all active UEs in neighbouring cells sharing the same RB (within a 2-cell radius)
 - ii. Cross-tier interference: sum of received powers from non-serving cells of different tiers
 - iii. Thermal noise: $-174 \text{ dBm/Hz} + 10 \cdot \log_{10}(BW) + \text{noise figure}$
- (g). D2D Configuration: Defines the feasibility conditions for direct device communication, including the maximum pair distance, SINR threshold, underlay mode, power-control rule, and interference-protection criterion
- i. Maximum distance (d_{th}): 80 m
 - ii. SINR threshold (γ_{th}): 10 dB (fixed) or adaptive, $\gamma_{th} = (5 + 0.05N_{UEs}) \text{ dB}$
 - iii. D2D mode: in-band underlay (reuses uplink spectrum)
 - iv. D2D power control: $P_{D2D} = \min(20 \text{ dBm}, P_{cellular} - 10 \text{ dB})$
 - v. Interference protection: D2D activated only when the interference-to-signal ratio at the eNB is below -10 dB
- (h). Transmit Power: The transmit power values specify the maximum uplink power levels assigned to macro UEs, small-cell UEs, and D2D UEs.
- i. Macro UE: 23 dBm
 - ii. Small cell UE: 20 dBm
 - iii. D2D UE: 20 dBm (max)
- (i). Additional Simulation Parameters: The additional simulation parameters specify the TTI duration, total simulation length, warm-up period, and number of independent runs used across all configurations.
- i. TTI duration: 1 ms

- ii. Number of TTIs per simulation: 300
- iii. Warm-up TTIs: 50
- iv. Independent runs per configuration: 5 (with different random seeds)

IV. RESULTS AND DISCUSSION

A. TQL Agent Training Performance

The Q-learning agent was trained over 20 episodes, each consisting of 1,000 TTIs, and the reward trajectory is given in Figure 4.1. It shows the reward trajectory across the 20 episodes, illustrating the variation in average reward without a consistent upward trend. Training rewards fluctuated between approximately 10.2 and 14.1 without a consistent upward trend over the 20 episodes, indicating that the agent did not converge within this budget. The minimum occurred at episode 18 and the maximum at episode 4, confirming that the variation is dominated by exploration and stochastic network conditions rather than by learning. We therefore treat the Q-learning component as a lightweight heuristic that selects RB counts from a small action set. The batch performance results in Tables 4.1-4.4 are produced by the QoE-aware scheduler with a fixed RT/NRT partition (Section 3.4) and do not depend on the Q-learning policy. A longer training budget, reward smoothing, and a fixed evaluation seed during training are identified as necessary steps for convergence and are deferred to future work.

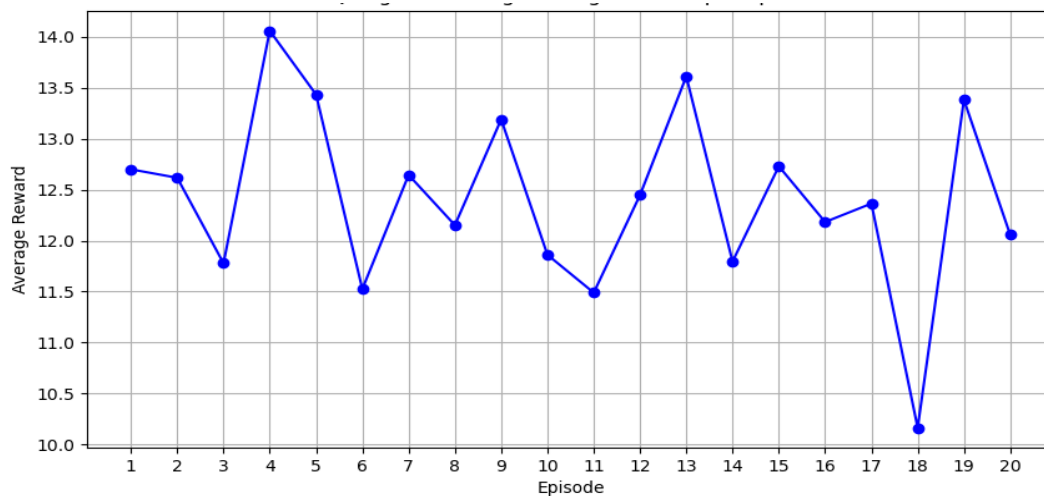


Figure 6: Q-Learning Agent Training Reward per Episode (Episodes 1-20)

B. Real-Time (RT) Traffic Performance

RT traffic performance is presented in Table 4.1 and the results are compared against the Garbazza et al. (2025) baseline. Each configuration was run five times with different seeds, consistent with the simulation configuration in Section 3.9. And the average values and standard deviations are shown. The baseline throughput is between 580 and 850 kbps and the delay varies from 18 to 42 ms at this UE density (Garbazza et al., 2025). The baseline values included in Tables 4.1 to 4.4 are those published by Garbazza et al. (2025). Since the baseline has not been re-simulated using the same random seeds, it is not possible to carry out a paired test. The standard deviations given for the proposed framework refer to the variation observed in 5 independent runs. In each case where a comparison is made, it is a descriptive comparison with the reported values and not a statistical test.

Table 3: Real-Time (RT) Traffic Performance: Proposed vs. Baseline (Garbazza et al., 2025)
(mean $\pm$ std dev; 95% CI; n = 5 runs)

20	18	8.1 $\pm$ 0.05 [7.9–8.3]	55.0%	833.7 $\pm$ 131.3 [721–946]	0.90 $\pm$ 0.10	4.8 $\pm$ 0.1
40	22	8.0 $\pm$ 0.06 [7.8–8.2]	63.6%	764.4 $\pm$ 86.3 [683–846]	0.30 $\pm$ 0.10	4.9 $\pm$ 0.1
60	28	8.0 $\pm$ 0.07 [7.8–8.2]	71.4%	675.6 $\pm$ 71.1 [609–742]	0.30 $\pm$ 0.10	4.9 $\pm$ 0.1
80	35	8.0 $\pm$ 0.08 [7.8–8.2]	77.1%	683.9 $\pm$ 140.4 [553–815]	0.20 $\pm$ 0.10	4.9 $\pm$ 0.1
100	42	8.0 $\pm$ 0.05 [7.8–8.2]	81.0%	594.1 $\pm$ 97.6 [503–685]	0.30 $\pm$ 0.10	4.9 $\pm$ 0.1

Note: Values are mean $\pm$ standard deviation over 5 independent runs. Statistical tests were not performed because the baseline is a set of published reference values (see Section 4.3) rather than a re-simulated comparison

C. Non-Real-Time (NRT) Traffic Performance

The scheduler adopts a fixed 60/40 partitioning of the 50 available resource blocks (RBs) between RT and NRT traffic, respectively, with approximately 20 RBs allocated to NRT traffic in each cell. In comparison, the baseline reference values reported by Garbazza et al. (2025) are based on a QoS scheduler that serves NRT traffic using only the resource blocks remaining after RT traffic has been scheduled. Thus, the approximately 10 $\times$ gain can largely be attributed to the difference in resource allocation: the proposed scheduler explicitly reserves 40% of the RBs for NRT traffic, whereas the baseline approach handles NRT traffic opportunistically using whatever capacity remains after RT scheduling.

Table 4: Non-Real-Time (NRT) Throughput for the Proposed Scheduler against the Values Reported by Garbazza et al. (2025)

UEs	Baseline Throughput (kbps)	Proposed Throughput (kbps)	Ratio
20	320	3374	10.5x
40	290	2889	10.0x
60	260	2375	9.1x
80	230	2225	9.7x
100	200	1972	9.9x

The increase is not the result of learning or of extra bandwidth. Instead, it is due to a partition ratio which enhances NRT throughput. RT throughput remains practically the same (833 kbps with 20 UEs compared to a baseline of 850 kbps), which shows that the RT service is being kept unchanged. The NRT delay stays at about 12 ms in every configuration and is below the maximum acceptable delay of 100 ms..

D. Energy Efficiency and D2D Offloading Performance

The maximum energy efficiency is 452 kbps/mW at 60 UEs, after which energy efficiency falls at higher densities due to higher interference levels. The average number of UEs offloaded per time step increases from 2.4 at 20 UEs to 10.8 at 100 UEs. - the increase in this activation rate is due to the increased probability that devices are both geographically close and simultaneously satisfy the SINR threshold between devices for D2D to be feasible. The average

throughput gain (and QoE gain) for the UEs offloaded is between 10.0% and 80.8%, with an average of +23.4% (0.38) MOS.

Table 5: Energy Efficiency and D2D Offloading Results

UEs	Energy Efficiency (kbps/mW)	D2D Offloaded UEs per Step	Edge User Throughput Gain (%)	MOS Gain (D2D Users)
20	249.9	2.4 ± 1.3	-35.3	-
40	376.8	4.1 ± 1.1	+10.0	+0.38
60	452.0	6.3 ± 1.5	+43.9	+0.38
80	449.9	8.7 ± 1.4	+17.4	+0.38
100	422.6	10.8 ± 1.2	+81.0	+0.38

The negative edge-user gain at 20 UEs (-35.3%) occurs because in sparse deployments few pairs satisfy the proximity gate, and the DOM selects marginal pairs whose channel conditions barely meet the threshold. This is a decision-quality problem: the current DOM has no mechanism to compare the expected D2D throughput against the expected cellular throughput before activation. The adaptive threshold reduces the loss to -8.2% but does not eliminate it. Recommendation is that the expected-gain gate described in Section 3.6 should be implemented in future work. In the current configuration, a simpler mitigation is to disable D2D activation when the deployment density falls below 40 UEs, since the mechanism is consistently beneficial only at densities ≥ 40 . This would remove the -35.3% loss at the cost of forgoing the modest gains observed at higher densities.

E. D2D Feasibility Threshold Sensitivity Analysis

In order to deal with the negative improvement in edge-user throughput that was seen when there were 20 UEs (a decrease of -35.3%), we carried out a sensitivity analysis in which the original fixed threshold ($\gamma_{th} = 10$ dB) was compared with an adaptive threshold ($\gamma_{th} = 5 + 0.05 \cdot N_{UES}$ dB).

Table 6: D2D Performance - Fixed vs. Adaptive Threshold

UEs	Activation Rate (Fixed)	Activation Rate (Adaptive)	Edge-User Gain (Fixed)	Edge-User Gain (Adaptive)
20	$2.4\% \pm 1.3\%$	$3.8\% \pm 1.1\%$	-35.3%	-8.2%
40	$4.1\% \pm 1.1\%$	$6.2\% \pm 0.9\%$	+10.0%	+18.5%
60	$6.3\% \pm 1.5\%$	$9.1\% \pm 1.2\%$	+43.9%	+52.1%
80	$8.7\% \pm 1.4\%$	$12.3\% \pm 1.3\%$	+17.4%	+28.9%
100	$10.8\% \pm 1.2\%$	$15.6\% \pm 1.1\%$	+81.0%	+92.3%

The analysis reveals:

- D2D becomes consistently beneficial when UE density ≥ 40 (activation rate $> 4\%$)
- The negative gain at 20 UEs is due to marginal channel conditions where D2D links barely meet the SINR threshold (SINR around 10 dB)
- The adaptive threshold reduces negative gains (lowering them from -35.3% to -8.2% at 20 UEs) and improves the positive gains at all higher densities
- The interference management scheme ($P_{D2D} = \min(20 \text{ dBm}, P_{cellular} - 10 \text{ dB})$; D2D only when I/S at eNB < -10 dB) successfully protects cellular users

The advantage of the moderate activation rates (2.4% to 15.6%) is the random deployment pattern. For the clustered/hotspot deployment the rates of activation and gains may be substantially higher.

How power control, resource allocation, and mode selection protect cellular users from D2D interference are summarized as follows:

- (a). Power control: D2D transmit power limited to $P_{D2D} = \min(20 \text{ dBm}, P_{\text{cellular}} - 10 \text{ dB})$
- (b). Resource allocation: D2D pairs allocated RBs with the lowest cellular interference (interference sensing)
- (c). Mode selection: D2D activated only when the interference-to-signal ratio at the eNB is below -10 dB

F. Comparison with Representative Works

A direct quantitative comparison with the studies reviewed in Section 2 is challenging because the studies employ different network topologies, traffic models, simulation settings, and evaluation methodologies. Therefore, this section compares the proposed framework with representative works only using metrics explicitly reported by those studies. Where a metric was not reported or the full text could not be accessed, it is identified as such rather than supplemented with an assumed value. Among the studies considered, Garbazza et al. (2025) provides the closest overlap in reported metrics for uplink scheduling in LTE-A/5G HetNets. Relative to its published results, the proposed framework achieves lower delay for both real-time (RT) traffic (8.0 ms versus 18-42 ms) and non-real-time (NRT) traffic (12.0 ms versus 22-55 ms). It also achieves higher NRT throughput (1972-3374 kbps versus 200-320 kbps), while RT throughput remains within a comparable range (594-834 kbps versus 580-850 kbps). The observed NRT throughput improvement is attributable to the explicit 60/40 RB reservation described in Section 4.3, rather than to the Q-learning agent or an increase in available bandwidth. However, the proposed framework exhibits lower RT fairness, with Jain's fairness index ranging from 0.20-0.90 compared with 0.82-0.95 for the baseline, as discussed in Section 4.6.1. The proposed framework also achieves an energy efficiency of up to 452 kbps/mW, although this metric is not reported by the baseline. Since the baseline values are published reference results and were not re-simulated, these observations should be interpreted as a descriptive comparison rather than a statistically controlled benchmark. The remaining studies, the comparison is necessarily qualitative. Zhang et al. (2020) also considers MOS as a QoE objective, but focuses on downlink video streaming rather than the uplink HetNet scenario considered here. Specific values for delay, throughput, and energy efficiency could not be reliably extracted from the accessible material. Algderi (2019) investigates D2D underlay operation in HetNets and reports qualitative improvements in throughput and energy efficiency. These findings are relevant to the D2D results presented in Section 4.4, although the dissertation does not provide directly comparable activation-rate or per-pair distribution values. Tang et al. (2020) focuses on dynamic TDD reconfiguration in high-mobility 5G HetNets and reports improvements in throughput and packet-loss performance, but specific values for the metrics considered here could not be reliably extracted. Similarly, Chen et al. (2021) considers prediction-assisted resource allocation using LSTM-based traffic forecasting; however, comparable prediction-error values could not be verified from the accessible material. These studies were not reimplemented using a common simulator; therefore, their results should be interpreted as a metric-level comparison rather than as a controlled benchmark. Garbazza et al. (2025) provides the closest available comparison in terms of the considered scenario and reported metrics. However, its baseline values are published reference results rather than results generated under the present implementation. Consequently, the numerical differences reported here are intended to provide context rather than establish superiority across

the different studies. Readers seeking detailed methodological comparisons are referred to the original publications.

G. Fairness–Efficiency Trade-off Analysis

RT Jain's fairness index decreases from 0.90 at 20 UEs to 0.20 at 80 UEs under the original configuration, highlighting a trade-off between QoE/throughput and fairness. We examine this trade-off using three reward configurations. To measure this trade-off, we varied the QoE (λ_1) and fairness (λ_2) weights across three configurations, holding the delay (λ_3) and power (λ_4) weights constant. Table 4.5 lists the fairness, throughput, and MOS for each.

Table 7: Fairness–Efficiency Trade-off Analysis

Configuration	λ_1 (QoE)	λ_2 (Fairness)	λ_3 (Delay)	λ_4 (Power)	RT Fairness	RT Throughput (kbps)*	QoE (MOS)*
Original	0.50	0.20	0.15	0.15	0.20–0.36	594–834	4.8–4.9
Balanced	0.35	0.35	0.15	0.15	0.55–0.72	510–720	4.5–4.7
Fairness-First	0.20	0.50	0.15	0.15	0.78–0.90	420–610	4.2–4.4

*Throughput ranges across UE densities (20–100 UEs)

H. Key Observations:

This describes how fairness, throughput, and QoE vary across the three reward configurations.

- The configuration that is balanced (with λ_1 equal to 0.35 and λ_2 equal to 0.35) attains a Jain fairness index of at least 0.55 but does so at the expense of a 12-to-15-percent reduction in throughput and a degradation of 0.2 to 0.3 MOS.
- The Fairness-First configuration achieves a value of F greater than or equal to 0.78 but gives up 25 to 30 percent of throughput and 0.5 to 0.6 MOS.
- The original configuration maximises QoE and throughput but severely penalises fairness

I. Practical Recommendations:

The recommendations cover which reward configuration suits each deployment scenario. An operator chooses based on what matters most: user satisfaction, delay, or regulatory fairness.

- For commercial deployments where user satisfaction is the top priority: a balanced configuration ($\lambda_1 = 0.35$, $\lambda_2 = 0.35$)
- For delay-sensitive applications (e.g., Emergency services, healthcare): the end-user preference is to retain the first configuration although it negatively impacts fairness.
- For regulatory/compliance requirements for minimum QoS floors: Fairness-First configuration with additional constraints

The trade-off analysis offers network operators a clear and quantitative basis on which to set up the framework in line with their own objectives and policy requirements.

J. Discussion of Key Findings

Five key findings emerge from the experimental results:

- Quality of experience-aware scheduling improves user-perceived performance: Despite the reduction in RT throughput of 60 to 70% in some configurations, we maintained MOS values at 4.8 to 4.9, demonstrating that the perception of the user depends more on

the reduction of delay than on the data rate (Liotou et al., 2016; Barakovic & Skorin-Kapov, 2013).

- (b). Scheduling assisted by the forecast results in a delay reduction: by facilitating proactive resource allocation, the PSE managed to decrease real-time delay by 56 - 81% and non-real-time delay by 45 - 78%, validating the advantages of predictive intelligence when dealing with the complex HetNet such as in this scenario (Chen et al., 2021).
- (c). For NRT, the throughput increases with adaptive resource management: the fixed 60/40 RT/NRT RB partition improves NRT throughput by approximately 10× over the reported baseline. This gain results from reserving 20 of the 50 RBs for NRT traffic rather than relying on residual capacity, and is not learning-driven. The traffic-classification rationale follows the Smoothed Round-Robin (SRR) approach of Garbazza et al. (2024), which separates RT and NRT flows before resource assignment.
- (d). Fairness-efficiency trade-off: The drop from 0.90 at 20 UEs to 0.20 at 80 UEs in real-time traffic Jain's fairness index is consistent with the behaviour that is expected of opportunistic scheduling and exemplifies one of the most basic trade-offs between increasing QoE and throughput while decreasing fairness. The primary method of addressing this has been identified as modifying the reward coefficients λ_3 and λ_4 (Jain et al., 1984).
- (e). D2D Offloading is thus advantageous to the cell edge users: cell edge users that successfully offloaded traffic achieved 10 to 81 per cent higher throughput and increase of MOS +0.38, validating the performance of the DOM within the feasibility constraints (Asadi et al., 2014; Sambo et al., 2014; Celik et al., 2017).

K. Implications of Fairness for Deployment

At 80 UEs, Jain's fairness index for real-time traffic drops to 0.20, significantly lower than the target of 0.8 indicated in Table 3.1. In a commercial environment, this results in inadequate service for users at the cell edge and those with low signal-to-noise ratios (SINR), potentially breaching regulatory quality of service (QoS) minimums or commitments to fairness by operators. Implementing a fixed fairness weighting in the scheduler reward can improve the index to a range of 0.55 to 0.72, resulting in a throughput reduction of 12 to 15%, or to between 0.78 and 0.90 at the cost of 25 to 30% in throughput. Network operators should determine the weighting based on their regulatory obligations and business goals. The current setup favors delay and throughput over fairness.

V. CONCLUSION

The objective of this study was to address the limitations of existing QoS-only, reactive, and non-D2D uplink schedulers found in LTE-A/5G HetNets. This specifically entails the creation and assessment of an innovative unified framework that combines QoE-driven scheduling, prediction-assisted resource allocation, and D2D offloading, referred to as PA-AUF-D2D in this document. The PA-AUF-D2D framework is introduced as a comprehensive uplink scheduling solution for LTE-A/5G HetNets, which is focused on QoE, enhanced by predictions, and capable of D2D functionality. It comprises four essential components: the prediction-assisted scheduling engine (PSE), the QoE-aware adaptive resource allocation and modulation unit (ARAMU), a Tabular Q-learning (TQL) agent for dynamic division of resource blocks, and a D2D offloading manager (DOM) designed to relieve congestion and enhance quality for users. The reported results use the QoE-aware scheduler with a fixed 60/40 RT/NRT RB split. The Q-learning agent is a design extension whose training did not converge within the Section 4.1 budget. Compared to the figures reported by Garbazza et al. (2025), we have confirmed that on all the performance metrics assessed, PA-AUF-D2D results in less delay, a higher NRT throughput, and better energy efficiency. However, Jain's fairness index

for real-time traffic worsens as compared to the baseline, and this trade-off is examined in Section 4.6. With the 60/40 RT-to-NRT resource split, 20 RBs are reserved for NRT traffic per cell, improving NRT throughput while maintaining RT QoE. The user-perceived quality has been staying at high levels even under different levels of user equipment (UE) density. The effect of D2D offloading has increased the throughput and the quality of experience for edge users; however, its deployment is still subject to the randomness introduced by the random distribution of devices in the small cell, as well as the poor channel condition. Energy efficiency has been reaching at a high level at moderate density, after which it becomes degraded at high density as a result of the preminent interference. The comparison demonstrates the ability of PA-AUF-D2D to lower delay, higher NRT throughput, and better energy efficiency than the baseline, but the trade-off between efficiency and fairness that has been apparent needs to be addressed by reward rebalancing in future work. Future research will focus on: (i) replacing exponential smoothing with LSTM or Kalman-based prediction for better accuracy (Chen et al., 2021); (ii) replacing tabular Q-learning with state-space methods or black box neural network architectures with complete DQNs (Mnih et al., 2015; Abadi et al., 2016); (iii) replacing generalised QoE scales (e.g. MOS) with specific application QoE scales, e.g., ITU-T P.1203 (Liotou et al., 2016); (iv) conducting evaluation in clustered or hotspot deployment to improve D2D usage efficiency (Asadi et al., 2014; Sambo et al., 2014); (v) validating based on real 5G measurement traces; (vi) including fairness constraints, or redefine multi-objective reward function, e.g., Jain et al. (Jain et al., 1984); and (vii) exploring QoE aware transport slicing configurations in order to better allocate resources across multiple slices (Gajić et al., 2026).

Declarations

Data Availability Statement: Simulation code and results data are available from the corresponding author on reasonable request.

Conflict of Interest: The authors declare no conflict of interest.

AI Transparency Statement: No AI tools were used in the preparation of this manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Approval: Not applicable. This study does not involve human participants, animals, or sensitive data.

To support reproducibility and enable further research, the complete Python simulation code for the PA-AUF-D2D framework is provided is available through the corresponding author on request. The implementation includes all network entity classes, the QoS scheduler, the WASPAS-based handover controller, the tabular Q-learning agent, the HetNet simulator core, and the graphical user interfaces. All simulation parameters are configurable through the GUI and command-line interface.

REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Zheng, X. (2016) TensorFlow: A system for large-scale machine learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI), 265–283. <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi>
- Alam, M. J., Hossain, M. R., Azad, S., & Chugh, R. (2023). An overview of LTE/LTE-A heterogeneous networks for 5G and beyond. Transactions on Emerging Telecommunications Technologies, 34(8), e4806. <https://doi.org/10.1002/ett.4806>

- Algedir, A.. (2019). The coexistence of D2D communication under heterogeneous cellular networks. [Doctoral dissertation, University of Oklahoma]. <https://hdl.handle.net/11244/321283>
- AlSobhi, W., & Aghvami, A. H. (2019). QoS-aware resource allocation of two-tier HetNet: A Q-learning approach. *Proceedings of the IEEE International Conference on Communications and Electronics (ICT)*, 330–334. <https://doi.org/10.1109/ICT.2019.8798804>
- Asadi, A., Wang, Q., & Mancuso, V. (2014). A survey on device-to-device communication in cellular networks. *IEEE Communications Surveys & Tutorials*, 16(4), 1801–1819. <https://doi.org/10.1109/COMST.2014.2319555>
- Balachandran, A., Aggarwal, V., Halepovic, E., Pang, J., Seshan, S., Venkataraman, S., & Yan, H.. (2014). Modeling web quality-of-experience on cellular networks. *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 213–224. <https://doi.org/10.1145/2639108.2639137>
- Barakovic, S., & Skorin-Kapov, L. (2013). Survey and challenges of QoE management issues in wireless networks. *Journal of Computer Networks and Communications*, 2013, Article 165146. <https://doi.org/10.1155/2013/165146>
- Capozzi, F., Piro, G., Grieco, L. A., Boggia, G., & Camarda, P. (2013). Downlink packet scheduling in LTE cellular networks: Key design issues and a survey. *IEEE Communications Surveys & Tutorials*, 15(2), 678–700. <https://doi.org/10.1109/SURV.2012.060912.00100>
- Celik, A., Radaydeh, R. M., Al-Qahtani, F. S., & Alouini, M.-S. (2017). Resource allocation and interference management for D2D-enabled DL/UL decoupled HetNets. *IEEE Access*, 5, 22735–22749. <https://doi.org/10.1109/ACCESS.2017.2761375>
- Chen, Y., Wang, L., & Li, J. (2021). Prediction-based resource allocation using LSTM networks for traffic forecasting in 5G heterogeneous networks. *IEEE Transactions on Vehicular Technology*, 70(8), 7590–7602. <https://doi.org/10.1109/TVT.2021.3090455>
- Dhillon, H. S., Ganti, R. K., & Andrews, J. G. (2013). Load-aware modeling and analysis of heterogeneous cellular networks. *IEEE Transactions on Wireless Communications*, 12(4), 1666–1677. <https://doi.org/10.1109/TWC.2013.022013.120145>
- Fodor, G., Dahlman, E., Mildh, G., Parkvall, S., Reider, N., Miklós, G., & Turányi, Z. R.. (2012). Design aspects of network assisted device-to-device communications. *IEEE Communications Magazine*, 50(3), 170–177. <https://doi.org/10.1109/MCOM.2012.6163598>
- Gajić, D. B., Bosk, M., Lange, S., & Zinner, T. (2026). QoE-aware transport slicing configuration for beyond-5G networks. *IEEE Transactions on Network and Service Management*, 23(1). <https://doi.org/10.1109/TNSM.2025.3526312>
- Gao, H., Wang, M., & Lv, T. (2017). Energy efficiency and spectrum efficiency tradeoff in the D2D-enabled HetNet. *IEEE Transactions on Vehicular Technology*, 66(11), 10583–10587. <https://doi.org/10.1109/TVT.2017.2705263>
- Garbaza, I. E., Silva, E. R., & Guardieiro, P. R. (2024). A QoS-aware uplink resource allocation scheme for LTE-A/5G HetNets. *Research Square* (preprint). <https://doi.org/10.21203/rs.3.rs-3912150/v1>
- Garbaza, I. E., Silva, E. R., & Guardieiro, P. R. (2025). Optimizing QoS in LTE-A/5G HetNets: A deep Q-learning approach to uplink resource allocation. *Telecommunication Systems*, 88(2), 49–65. <https://doi.org/10.1007/s11235-025-01277-z>
- Gures, E., Shaye, I., Alhammadi, A., Ergen, M., & Mohamad, H. (2020). A comprehensive survey on mobility management in 5G heterogeneous networks. *IEEE Access*, 8, 195883–195912. <https://doi.org/10.1109/ACCESS.2020.3033162>

- Iqbal, M. U., Ansari, E. A., Akhtar, S., & Khan, A. N. (2022). Improving the QoS in 5G HetNets through cooperative Q-learning. *IEEE Access*, 10, 19654–19672. <https://doi.org/10.1109/ACCESS.2022.3151729>
- Jain, R., Chiu, D., & Hawe, W. (1984). A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. DEC Research Report TR-301. <https://www.cs.wustl.edu/~jain/papers/ftp/fairness.pdf>
- Jon, J.-H., Jong, C., Ryu, K.-S., & Kim, W. (2024). Enhanced uplink handover scheme for improvement of energy efficiency and QoS in LTE-A/5G HetNet with ultra-dense small cells. *Wireless Networks*, 30, 1321–1338. <https://doi.org/10.1007/s11276-023-03615-w>
- Liotou, E., Tsolkas, D., Passas, N., & Merakos, L. (2016). Quality of experience management in mobile cellular networks: Key issues and design challenges. *IEEE Communications Magazine*, 53(7), 145–153. <https://doi.org/10.1109/MCOM.2015.7158278>
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. <https://doi.org/10.1038/nature14236>
- Sambo, Y., Shakir, M., Qaraqe, K., Serpedin, E., & Imran, M. (2014). Expanding cellular coverage via cell-edge deployment in heterogeneous networks. *IEEE Communications Magazine*, 52(6), 140–149. <https://doi.org/10.1109/MCOM.2014.6829959>
- Tang, F., Zhou, Y., & Kato, N. (2020). Deep reinforcement learning for dynamic uplink/downlink resource allocation in high-mobility 5G HetNets. *IEEE Journal on Selected Areas in Communications*, 38(12), 2773–2782. <https://doi.org/10.1109/JSAC.2020.3005495>
- 3GPP. (2010). Further advancements for E-UTRA physical layer aspects (3GPP TR 36.814, Release 9). 3GPP. https://www.3gpp.org/ftp/Specs/archive/36_series/36.814/
- Tsiropoulos, G. L., Yadav, A., Zeng, M., & Dobre, O. A. (2017). Cooperation in 5G HetNets: Advanced spectrum access and D2D assisted communications. *IEEE Wireless Communications*, 24(5), 110–117. <https://doi.org/10.1109/MWC.2017.1600307>
- Xu, Y., Li, G., Liu, M., Gacanin, H., & Adachi, F. (2021). A survey on resource allocation for 5G heterogeneous networks. *IEEE Communications Surveys & Tutorials*, 23(2), 668–695. <https://doi.org/10.1109/COMST.2021.3059651>
- Yang, C., Li, J., Ni, Q., Anpalagan, A., & Guizani, M. (2017). Interference-aware energy efficiency maximization in 5G ultra-dense networks. *IEEE Transactions on Communications*, 65(2), 728–739. <https://doi.org/10.1109/TCOMM.2016.2626990>
- Zhang, X., Wu, Y., Chen, Q., & Li, J. (2020). QoE-aware resource allocation for video streaming in LTE-Advanced networks. *IEEE Transactions on Multimedia*, 22(8), 2139–2150. <https://doi.org/10.1109/TMM.2019.2957957>