User Association based Load Balancing using Q-Learning in 5G Heterogeneous Networks

Gokul G
Department of Electronics Engineering and Centre for Internet of Things
Madras Institute of Technology Campus
Anna University
Chennai, India
gokulg232003@gmail.com

Dhanushree V. S.
Department of Electronics Engineering and Centre for Internet of Things
Madras Institute of Technology Campus
Anna University
Chennai, India
dhanu2003shree@gmail.com

Sahana S
Department of Electronics Engineering and Centre for Internet of Things
Madras Institute of Technology Campus
Anna University
Chennai, India
sahana.saravan13@gmail.com

Parameswaran Ramesh
Department of Electronics Engineering and Centre for Internet of Things
Madras Institute of Technology Campus
Anna University
Chennai, India
parameswaran0789@gmail.com

Bhuvaneswari P. T. V.
Department of Electronics Engineering and Centre for Internet of Things
Madras Institute of Technology Campus
Anna University
Chennai, India
ptvbmit@annauniv.edu

***Abstract*—The employment of Heterogeneous Networks (HetNets), comprising Macro Base Station (MBS) and Small Base Stations (SBSs), emerges as an effective solution that can address various challenges imposed in 5G networks by leveraging the strengths of various technologies to enhance coverage, capacity, Quality of Service (QoS), cost efficiency and adaptability. Efficient methods of associating User Equipment (UE) with the BS in HetNets scenario to maximize overall network performance is referred to as User Association-based Load Balancing (UALB). Through Load Balancing (LB) the number of UEs that require connectivity can be increased using limited available spectrum and BSs. In this research, a three-tier downlink HetNet, consisting of MBS, Pico BSs (PBS) and Femto BSs (FBS), operating under a log-distance channel model in both rural areas in accordance with 3GPP is taken into consideration. The signal model developed in the research, determines the SINR and transmission rates for each UE associated with the BS. UALB is then performed using the Q-Learning approach and the cumulative transmission rate achieved is computed. The performance of the developed UALB, is analysed for different network scenarios by introducing mobility of UEs and varying the number of UEs and BSs in the network. From the python based simulation made, it is observed that the proposed QL based UALB algorithm enhances the average transmission rate of the network from 0.91 to 4.44 Mbps. An extensive survey of the effect of user association load balancing in improving the cumulative transmission rate of HetNet was performed successfully and the variation in performance under different network scenarios is analyzed.**

***Index Terms*—5G, HetNets, User Association based Load Balancing, Q-Learning and Reinforcement Learning**

I. Introduction

The increasing demand for faster data speeds, lower latency, and enhanced connectivity has propelled the development and adoption of 5G technology as the next frontier in wireless communication systems. The expansion of data-intensive applications namely streaming, gaming, and Internet of Things (IoT) devices, along with the exponential growth of Internet-connected devices, presents unprecedented problems for traditional wireless networks in fulfilling the changing needs of users and industry. In response, 5G technology aims to provide faster data rates, negligible levels of latency reduction, and increased capacity to accommodate the growing number of connected devices [1]. This revolutionary step forward not only meets the needs of the present but also lays the groundwork for upcoming breakthroughs and technological developments in a variety of fields.

LTE and LTE-Advanced [2] prioritize wireless communication by offering 500 Mbit/s for uplink and 1 Gbit/s for downlink data transmission, respectively, with reduced latency of less than 10 milliseconds, and 70 MHz (downlink) and 40 MHz (uplink) bandwidth allocations. The architecture has been designed for enhanced data transmission speed at cell edges, reduced power consumption, and optimized spectrum utilization [3]. On the other hand, 5G technology attains a delay reduction of 1 millisecond while increasing data transmission velocities by a factor of ten, from 1 to 10 Gbps. 5G enables connections among thousands of devices in a confined space at velocities of up to 500 kilometers per hour, guarantees one hundred percent coverage, places importance on energy efficiency and sustainable technologies, and features an extended battery life. 5G emerges as an evolution in the area of wireless communication technology, surpassing LTE and LTE-Advanced [4].

The components involved in 5G Architecture are Next

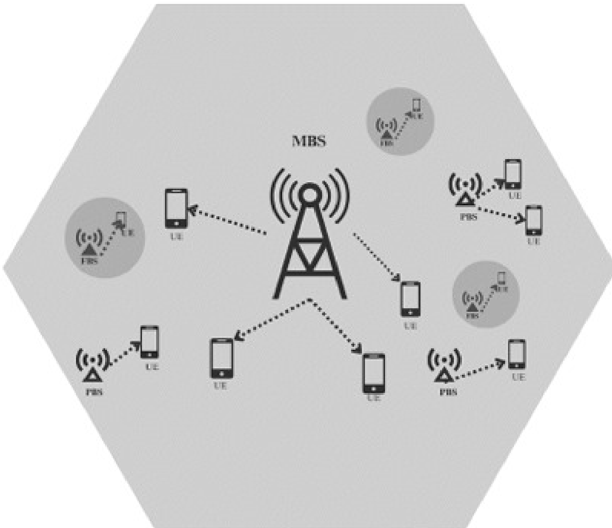

Fig. 1. Architecture of HetNet

Generation - Radio Access Network (NG-RAN), 5G Core Network (5GC), and User Equipment (UEs). The UEs are made up of a Mobile Station and a Universal Subscriber Identity Module (USIM). The UEs wirelessly communicate with Base Stations (BS), or gNBs (Next Generation NodeB) via the NG-RAN. Data speeds, latency, and device density are improved via massive MIMO, beamforming, and dynamic spectrum sharing in the NG-RAN. It supports network slicing, which creates virtualized, isolated networks for specific use cases to optimize service performance [5].

In order to support high mobility users, the 5G cellular architecture needs to be heterogeneous, comprising relays, small cells, macrocells and microcells. Heterogeneous Networks (HetNets) have been considered as the most viable solution to the impending mobile data traffic. An architecture of three-tier 5G HetNet is depicted in Figure 1. It uses macro, micro, pico, and femto cell layers and access technologies. Usually, rooftops, macrocell towers, or other elevated structures are used for the deployment of macrocells. Picocells are frequently placed indoors or in locations that have particular high-density needs. Femtocells are often installed in small areas by end users [6]. They give a localized coverage extension and are linked by the broadband internet [7].

UEs in a 5G wireless network are linked to the backhaul network by BSs spanning a particular cell. Each BS has a maximum number of UEs it can connect to or serve. Due to the rapid growth in the number of UEs, there are instances in which more UEs are present in the region of BS coverage beyond the maximum permissible capacity, resulting in overloading of BS. Under such circumstances, certain UEs will be incapable of establishing a connection with a BS. These UEs can be connected to a neighboring BS that is operating with UEs less than its maximum capacity. This process of distributing the load from overloaded BS to adjacent BSs operating below maximum capacity is called LB [8].

It aims to improve the overall capacity and throughput of a network by efficiently connecting UE to BSs ensuring optimal allocation of spectrum. An efficient load-balancing method should ensure that all the UEs are connected to a BS while also maximizing the overall throughput of the network [9].

The LB mechanism can be executed by using Machine Learning (ML) algorithms. Reinforcement Learning (RL) [10] is a type of ML where agents learn by trial and error to make the correct sequential decisions in an environment. The Agent is given positive and negative numerical rewards when it makes correct and wrong actions respectively. The agent tries to make decisions that maximize the cumulative rewards obtained by it. The three distinguishing features of RL algorithms are that they do not require a dataset as an input as the algorithm learns by experimenting with actions in terms of rewards and penalties. Secondly, the UEs can act as independent agents and can make decisions independently. Therefore, the workload of the MBS will be reduced. Thirdly, it can adapt to a dynamic environment with large state and action spaces as it updates its association strategies based on new experiences [11].

The structure of the paper is outlined as follows. Section 2 offers an overview of the research conducted in this particular field. Section 3 discusses the proposed LB methodology. Section 4 discusses the network parameters, experimental results and their interpretation. Section 5 concludes and presents the scope of future work.

II. Literature Survey

The authors in [12] have analyzed the HetNets architecture utilized within 5G wireless communication systems. They conducted an extensive examination of past studies concerning Resource Allocation (RA) challenges in HetNets, as well as the methods presently employed to address them.

In [13], the authors have proposed leveraging HetNets to enhance the capacity of wireless cellular networks. They advocate for deploying low-power BSs in high demand service regions alongside MBS for broader coverage, aiming to boost network capacity. Furthermore, they delve into the issue of co-channel interference experienced by users served by various BSs.

In [14], the author addresses the coexistence of delay-sensitive and bandwidth-sensitive traffic in the multi-factor LB technique for data center networks. Performance of the algorithm is compared with existing methods in a simulated environment constructed with mininet, and the results demonstrate significant increases in standardized total throughput, average throughput, and connection delay. To further improve the path selection process, the algorithm additionally considers the load level and the cost of bandwidth fragmentation.

In [15], the author developed a joint optimization of interference mitigation and LB which takes into account the coverage rate and coverage probability of a user in a specific tier which is given by the Signal-to-Interference Ratio (SIR) of that tier. It involves optimizing the spectrum partition and user association bias to maximize the average user utility and aims to strike a balance between minimizing the utility percentage of zero-rate users and maximizing the utility percentage with non-zero rate.

In [16], the author evaluated the use of Device-to-Device (D2D) communications for LB in 5G ultra-dense networks. They address the spatial-temporal fluctuation problem of mobile data traffic in cellular networks by offloading congested small cells to underutilized small cells. This approach leverages the underlay property and proximity gain of D2D communication to relay traffic from congested cells to adjacent cells without the need for extra spectrum and analyzes the system sum-rate as the parameter of interest. In 5G ultra-dense small cell networks, the same spectrum band is reused in neighboring small cells. This limits the effectiveness of D2D communication for LB, as the neighboring cells cannot borrow channels from each other and solving this optimization problem efficiently can be challenging and may require iterative algorithms.

In [17], the capacity shortage and cell performance deterioration in LB is addressed using Constriction Factor Particle Swarm Optimization (CFPSO) technique. The proposed CFPSO algorithm enhances the throughput by 44.08% and latency by 4.08% when compared to existing algorithms.

In [18],the author presents a novel decentralized approach for LB in heterogeneous networks. The paper aims to minimize an α-fairness objective function with load and outage constraints This approach is formulated as a minimization of an fairness objective function with load and outage constraints. The proposed distributed learning algorithms namely log-linear learning algorithm and binary log-linear learning algorithm are proven to converge faster. The convergence of the proposed algorithms within a few hundred iterations is compared with other learning algorithms.

The author in [19] addresses the use of two different association methodologies for LB in cellular HetNets. The initial approach seeks to maximise the weighted total of long-term rates in order to balance network loads. The second kind, on the other hand, uses power control to lessen network interference and save energy. They argue that although using the second algorithm results in a slight decrease in data rate when compared to the first, the average energy efficiency of the system has significantly improved.

In previous studies, the authors constructed a number of system models with limited requirements. These works focused on joint optimization algorithms for user association and resource allocation. They considered the fixed deployment of UEs, analyzed system performance based on model performance, and proposed that their model would be useful in situations with certain parameters. However, none of these studies addressed the UARA problem in the context of user mobility. It is important to note that UE locations may vary, making it inefficient to apply the same channel characteristics to all environmental applications. Consequently, the model created by the authors cannot be universally applied in all scenarios with respect to distance.The proposed work considers a signal model for a three-tier HetNet operating under a log-distance channel model in rural areas. It analyzes and plots the performance of load balancing after varying network parameters and introducing mobility to UEs. This work serves as a benchmark for evaluating the performance of various intelligent user association-based load balancing methods.

III. Proposed User Association based Load Balancing (UALB) using QL Algorithm

A. Network Model

A three-tier HetNet in downlink condition consisting of N_m MBSs, N_p Pico BSs (PBSs), and N_f Femto BSs (FBSs) is considered in this research to perform Load Balancing. The Base Stations (BSs) are represented by the set $BS = \{mbs_1, \ldots, mbs_{Nm}, pbs_1, \ldots, pbs_{Np}, fbs_1, \ldots, fbs_{Nf}\}$ and with the indexes $\{1, 2, \ldots, N\}$ where the total number of base station is given by $N = N_m + N_f + N_p$. In the Network, the MBSs and PBSs are network deployed, while the FBSs are user deployed and all the BSs are assumed to be centre-excited.

TABLE I
List of Notations

S. No.	Notation	Description
1	N_m	Number of MBSs
2	N_p	Number of PBSs
3	N_f	Number of FBSs
4	N	Number of BSs
5	L	Number of UEs
6	K	Number of Channels
7	i	BS under Consideration
8	j	UE under Consideration
9	k	Channel under Consideration
10	b_j^i	BS Association Vector
11	c_j^k	Channel Association Vector
12	$PL_{i,j}$	Path Loss
13	$SINR_{i,j}^k$	SINR
14	$p_{i,j}^k$	Transmission Power
15	$h_i^{j,k}$	Channel Gain
16	BW	Bandwidth
17	N_o	Noise Power Density
18	$C_{i,j}$	Capacity
19	$r_{i,j}$	Transmission Rate
20	$\bar{R}$	Average Transmission Rate
21	s	Current State
22	a	Current Action
23	s'	Next State
24	a'	Next Action
25	$Q(s,a)$	Q-Value for given State and Action
26	α	Learning Rate
27	γ	Discount Factor

A total of L number of UEs in the set $UE = \{1, 2, \ldots, L\}$ are considered to be requesting the BSs for connection in the HetNet. A total of K orthogonal channels belonging to the set $Ch = \{1, 2, \ldots, k\}$ is available to the UEs to connect to a BS. Also, one or more UEs can utilize a common channel to connect with the BSs.

Consider $i \in UE$ trying to be associated with a $j \in BS$ through channel $k \in Ch$. The following parameters need to be computed for every BS and UE pair (i, j) to perform user association and check its effectiveness.

In the network, each UE can associate with only one BS and channel at any given instant, and this is governed by the conditions given by:

$$\sum_{i \in BS} b_j^i \leq 1 \tag{1}$$

$$\sum_{k \in Ch} c_j^k \leq 1 \tag{2}$$

Where, b_j^i and c_j^k are the binary BS and channel association vectors respectively. If the j^{th} UE associates with i^{th} BS, then $b_j^i = 1$ or $b_j^i = 0$ otherwise. Similarly, if the j^{th} UE utilizes the k^{th} channel, then $c_j^k = 1$ or $c_j^k = 0$ otherwise.

Path loss is used to account for factors namely signal attenuation, interference, and noise that characterize signal propagation in a real-time environment. This work considers the log-distance path loss, defined in the 3GPP 38.901 standards document [20], which accounts for signal attenuation over distance. The path loss for communication between a BS i and an UE j in rural environment is given by:

$$PL_{i,j}\ (in\, dB) = 34 + 40 \log_{10}(d) \tag{3}$$

where, d is the 2D distance between BS i and UE j.

Signal-to-Interference-plus-Noise Ratio (SINR) represents the ratio of the desired signal power to the combined interference and background noise. In this work, SINR plays a crucial role in determining the suitability of a UE and BS pair and offloading a UE from a BS. The general formula of $SINR$ between BS i and UE j through channel k is given by:

$$SINR_{i,j}^k = \frac{b_j^i h_i^{j,k} c_j^k p_{i,j}^k}{\sum_{l \in BS \setminus \{i\}} b_j^l h_l^{j,k} c_j^k p_{l,j}^k + BW.N_o} \tag{4}$$

Where, $h_i^{j,k}$ is the channel gain derived from $PL_{i,j}$ defined in Equation 3, $p_{i,j}^k$ is the power transmitted by BS i, BW is the channel bandwidth and N_o is the noise power density.

The transmission rate is the speed at which digital data is transferred between UE and BS. It can also be defined as the amount of data transmitted in a given time between a UE and BS. In this work, the transmission rate is an important indicator for representing the efficiency of a communication link because it measures measures the rate of data exchanged over a channel.

According to the Shannon capacity theorem, the maximum transmission rate $r_{i,j}$ for a link between BS i and UE j is less than or equal to the capacity $C_{i,j}$ of the link. The capacity $C_{i,j}$ is given by:

$$C_{i,j}^k = BW \log_2(1 + SINR_{i,j}^k) \tag{5}$$

In this work, the maximum possible transmission rate is taken into consideration. So, the $r_{i,j}$ is given by:

$$r_{i,j}^k\ (in\, bps) = BW \log_2(1 + SINR_{i,j}^k) \tag{6}$$

For the whole network, the average transmission rate $\bar{R}$ is given by:

$$\bar{R}\ (in\, bps) = \frac{\sum_{i \in UE} r_{i,j}^k}{L} \tag{7}$$

Mobility is incorporated into the UE to introduce dynamism to the system. In this study, the Random Walk mobility model is employed to simulate UE movement. This model operates on a stochastic process, where UEs change direction randomly at each time step [21]. The updated coordinates after each step of UE mobility are given by:

$$x' = x + l \cos(\theta) \tag{8}$$

$$y' = y + l \sin(\theta) \tag{9}$$

Where, (x, y) is the initial coordinates, l is the step size and $\theta \in (0, 2\pi)$ is an angle chosen randomly that controls the mobility direction at each step.

B. UALB using Q-Learning Algorithm

Q-learning (QL) is a model-free RL algorithm designed to find the optimal action-selection policy in a given environment. Unlike some other RL methods, the QL algorithm does not require a model of the environment, making it highly adaptable to various situations. Additionally, QL algorithm does not require a dataset for training the ML model as the learning happens on the go. Thus, QL is well-suited for performing UALB in 5G HetNets.

A QL algorithm involves an agent making decisions within an environment and receiving rewards based on those decisions. These rewards can be positive for successful actions and negative for unsuccessful ones. Objective of the agent is to take actions that maximize positive rewards over time. In the context of LB in 5G HetNets, the agent represents a UE i, the environment is the 5G HetNet, and the action involves UE i selecting a BS j and channel k for association. Positive rewards are received when the chosen BS and channel pair establish a successful connection.

The QL algorithm aims to learn the Q-function, which estimates the expected utility or the total accumulated reward of choosing a BS and channel pair in a given state and following the optimal policy thereafter. The Q-function, denoted as $Q_i(s, a_i)$, represents the value of action a_i of a particular UE i picking a BS and channel pair in the current network state s. The goal of the algorithm is to learn the optimal Q-values, $Q^*(s, a)$ which satisfy the Bellman equation:

$$Q^*(s, a) = \mathbb{E}[r_t + \gamma \max_{a'} Q^*(s_{t+1}, a') | s_t = s, a_t = a] \tag{10}$$

where r_t is the reward received after taking action a in state s, and s_{t+1} is the state resulting from this action. The

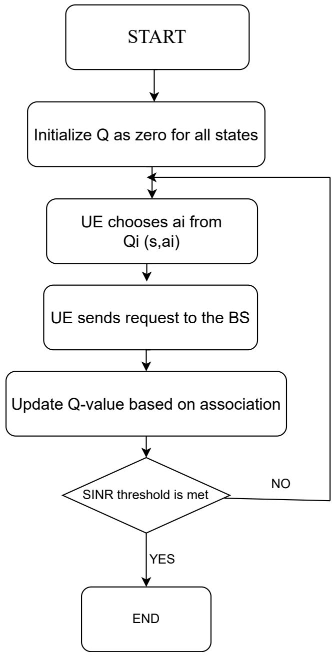

Fig. 2. QL Algorithm

discount factor γ $(0 \leq \gamma \leq 1)$ determines the importance of future rewards, with values closer to 1 making future rewards more significant.

The steps involved in QL algorithm is illustrated in Figure 2. Initially, the Q-function of all the UEs $Q_i(s, a)$ is set to zero. Using the ϵ-greedy approach [22], the algorithm either chooses to explore by gaining more information about the environment (picks a BS and channel pair randomly) or exploit by leveraging the current knowledge of the environment to select actions that are known to be positive (picks BS and channel based on Q-function). Then a connection request is sent by UE and a reward is observed based on the success of the connection. Then, the Q-function is updated based on association success using:

$$Q(s, a) \leftarrow Q(s, a) + \alpha(R(s, a) + \gamma \max_{a'} Q(s', a') - Q(s, a)) \tag{11}$$

where α $(0 \leq \alpha \leq 1)$ is the learning rate that controls how much new information overrides the old information. This equation ensures that the Q-values are progressively refined based on the observed rewards and the estimated value of the subsequent states. The above process is repeated until all the UEs associate with a BS successfully. Q-learning is guaranteed to converge to the optimal Q-functions, $Q(s, a)$, as long as all state-action pairs continue to be updated and the learning rate α decreases appropriately over time.

IV. Results and Discussion

A. Simulation Parameters

The HetNet under consideration consisting of 1 MBs, 2 PBs and 4 FBs along with 30 UEs is simulated using python and the performance of QL for UALB is investigated. The other parameters considered for simulation are given in II.

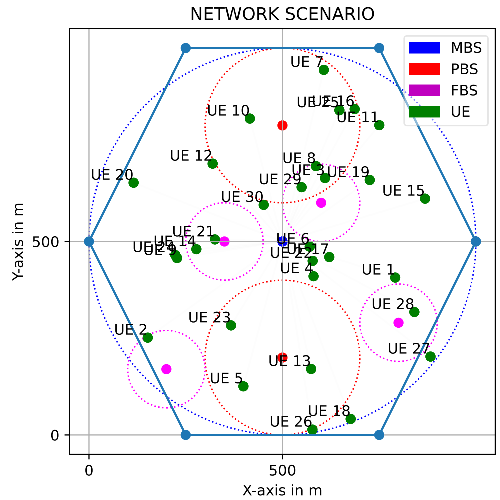

Fig. 3. Network Scenario before performing load balancing

TABLE II
Network Simulation Parameters

Parameter	Value
Channel Bandwidth W	180 KHz
Centre Frequency f	2 GHz
Number of UEs L	10 to 100
Number of Channels K	25
Number of BSs	MBS-1, PBS-2, FBS-4
Channel Model	Rural
Transmit Power of BSs $p^k_{i,j}$	MBS-40, PBS-30, FBS-8 dBm
Coverage of BS	MBS-500, PBS-200, FBS-100 m
Step Size l	10 meters
Noise Power Density N_o	-174 dB

The HetNet under simulation is a hexagonal cellular network of radius 500 meters. The MBS is placed in the centre of network at (500,500) meter. The PBSs are place above 300m above and below the MBS. The FBSs and UEs are deployed randomy inside the area of coverage of the MBS. The overall network under simulation is depicted in Figure 3.

B. Trasmission Rate Analysis

The average transmission rate of the network is a key performance indicator of the proposed algorithm. Initially, before performing LB all the UEs are associated with the MBS. Then, the QL algorithm balances the network load by performing user associations with other PBSs and FBSs. The Firgure 4 depicts the user association after performing LB using QL algorithm. Table III illustrates the comparison of transmission rates before and after implementing LB for 30 UEs using QL algorithm and the same is plotted in Figure 5. A significant increase in transmission rate can be observed after performing LB. According to 3GPP standards, the minimum required rate for audio calls is 64Kbps, and for video calls is 1.5Mbps. The average transmission rate of the network

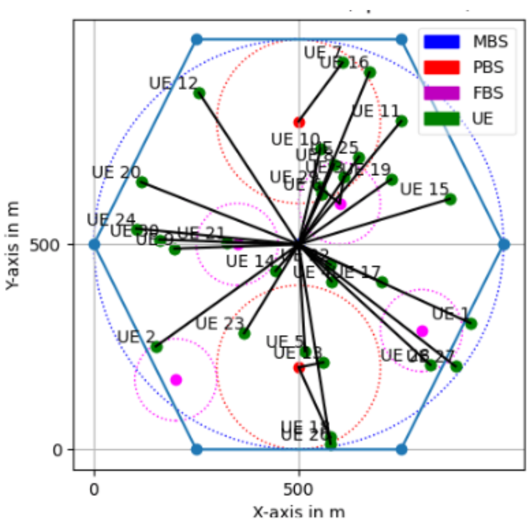

Fig. 4. Network Scenario after performing UALB using QL algorithm

TABLE III
TRANSMISSION RATE OF UEs BEFORE AND AFTER LB

UE	User Association before LB	Rate before LB (Mbps)	User Association after LB	Rate after LB (Mbps)
1	MBS	0.66	MBS	4.40
2	MBS	1.00	FBS4	4.05
3	MBS	0.61	MBS	4.72
4	MBS	1.02	MBS	5.08
5	MBS	0.72	MBS	4.23
6	MBS	0.70	MBS	5.45
7	MBS	1.06	MBS	4.11
8	MBS	1.04	FBS2	4.05
9	MBS	0.83	MBS	4.47
10	MBS	1.00	MBS	4.35
11	MBS	0.95	MBS	4.22
12	MBS	0.94	MBS	4.49
13	MBS	0.67	MBS	4.33
14	MBS	0.88	FBS3	4.22
15	MBS	0.70	MBS	4.23
16	MBS	0.93	PBS1	4.14
17	MBS	1.30	MBS	5.03
18	MBS	1.14	MBS	4.05
19	MBS	1.01	MBS	4.47
20	MBS	0.85	MBS	4.18
21	MBS	1.13	MBS	4.81
22	MBS	1.19	MBS	5.27
23	MBS	0.85	PBS2	4.29
24	MBS	1.15	MBS	4.46
25	MBS	0.83	MBS	4.26
26	MBS	0.98	MBS	4.05
27	MBS	0.57	MBS	4.06
28	MBS	0.73	MBS	4.23
29	MBS	0.91	FBS1	4.33
30	MBS	1.09	MBS	5.17

increased from 0.91 Mbps to 4.44 Mbps after performing LB using the QL algorithm.

C. UE Scalability Analysis

The scalability of the proposed UALB - QL algorithm is verified by varying the number of UE from 10 to 100. The variation in transmission rate for different number of UEs

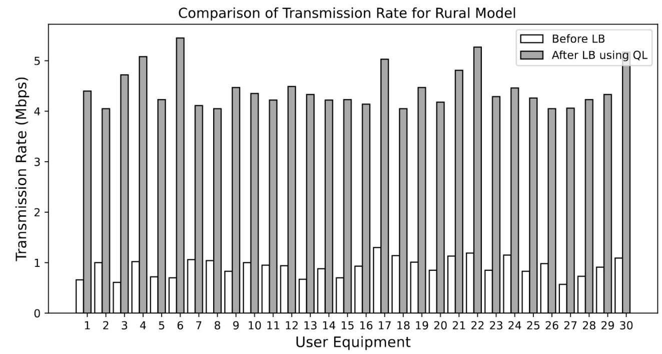

Fig. 5. Comparison of Transmission Rate of UEs before and after LB

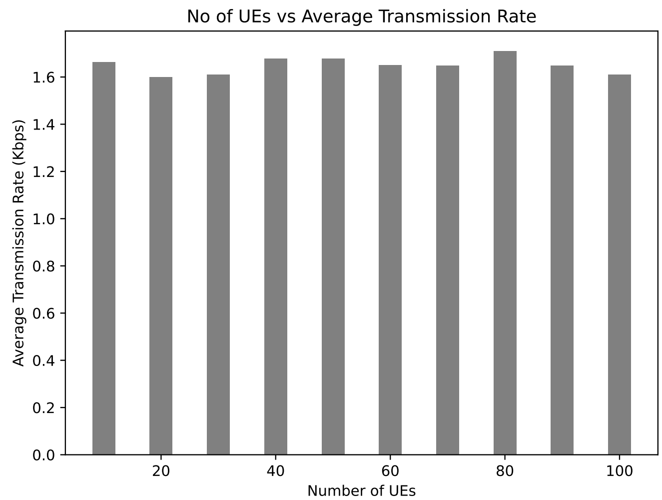

Fig. 6. Average transmission rate for varying number of UEs

in the network is depicted in Figure 6. It suggests that the performance of the developed UALB- QL algorithm remains unaffected by the rising number of UEs. Also, it can be inferred that the QL UALB algorithm is capable of handling over 100 UEs.

D. Computational Complexity Analysis

The time and space complexity of the proposed algorithm is derived in terms of number of UEs L, BSs N, channels K and replay memory size R. The complexity of QL algorithm is derived as $O(LNK)$. It is inferred that the QL algorithm is considered to be moderately computationally expensive. The plot in Figure 7 shows the change in total time taken by UALB - QL algorithm to execute on Windows i5 system with 8GB RAM on scaling the total number of UEs in the network. The gradual increase in the number of UEs results in an exponential increase in simulation time as predicted by the computational analysis.

E. Convergence Analysis

The proposed algorithm is said to converge when all UEs get associated with a BS successfully. This evaluation is essential

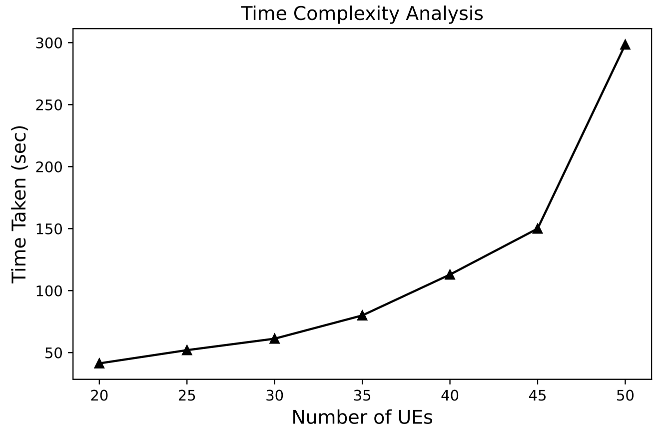

Fig. 7. Time Complexity analysis of the algorithm

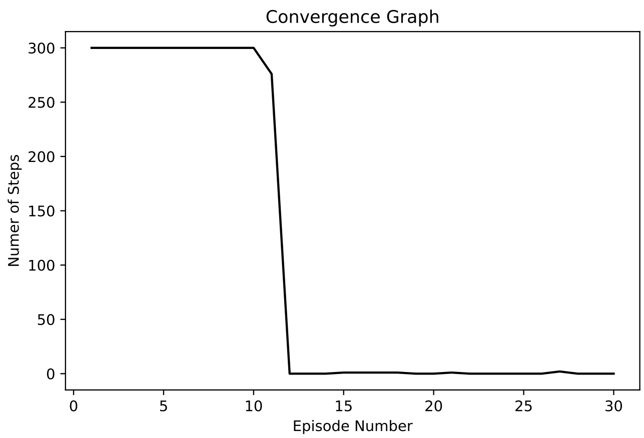

Fig. 8. Convergence Analysis of the algorithm

to understand the stability of the learning algorithms. Covergence can be achieved by fine-tuning the hyperparameters of the QL algorithm. As seen in Figure 8, the QL algorithm starts to converge from the 12th episode. Prior to this, the algorithm operates on a trial-and-error basis to perform associations and accumulate rewards. It then learns from these rewards and starts to converge.

F. Mobility Analysis

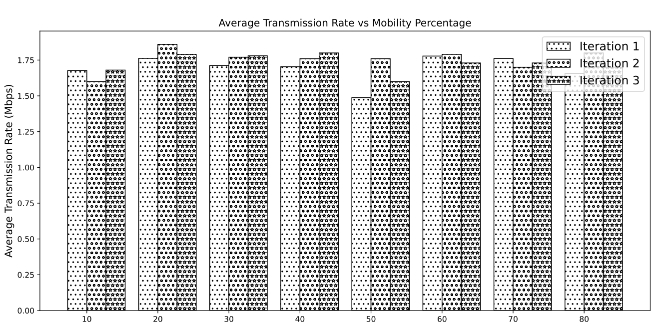

Fig. 9. Transmission rate for varying Mobility in Network

In real-time network environments, users frequently move around and the proposed UALB algorithms need to be robust to this mobility. The Figure 9 shows the transmission rate for the system by varying the percentage of mobile UEs. The transmission rate is not impacted by increasing the number of mobile UEs in the network. From this, it can be inferred that mobility has minimal impact on the average transmission rate. Therefore, the proposed algorithm exhibits robustness in a dynamic environments.

V. Conclusion and Future Work

The design and simulation of a downlink three-tier 5G heterogeneous network, consisting of 30 UEs, 1 MBS, 2 PBSs, and 4 FBSs, has been performed. UALB is achieved using the QL algorithm, quantifying the results with respect to 3GPP in improved transmission rates.The proposed algorithm is stable in dynamic environments with mobile users and a varying number of users.The QL approach demonstrates quicker convergence and reduced computational complexity. The proposed algorithm meets the minimum transmission rate required for eMBB services such as audio and video calls. The existing work does not consider UEs as independent agents while performing LB and the proposed work evaluates the performance of each UE indvidually in the scenario.Additionally, it can allocate a channel and BS pair to UEs, referred to as a resource block. Future research can focus on enhancing transmission rates to support eMBB services such as high-definition video streaming by assigning multiple resource blocks to a single UE and using channel models that simulate real-world scenarios with high UE density, such as stadiums, where effective LB is essential.

References

[1] A. Gohil, H. Modi, and S. K. Patel, "5G technology of mobile communication: A survey," in *2013 international conference on intelligent systems and signal processing (ISSP)*. IEEE, 2013, pp. 288–292.

[2] B. A. Bjerke, "LTE-advanced and the evolution of LTE deployments," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 4–5, 2011.

[3] G. Barb and M. Otesteanu, "4G/5G: A comparative study and overview on what to expect from 5G," in *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2020, pp. 37–40.

[4] J. Isabona, C. C. Ugochukwu, A. L. Imoize, and N. Faruk, "An empirical comparative analysis of 4G LTE network and 5G new radio," in *2022 5th Information Technology for Education and Development (ITED)*. IEEE, 2022, pp. 1–5.

[5] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE communications magazine*, vol. 52, no. 2, pp. 122–130, 2014.

[6] S. M. Abd El-atty and Z. Gharsseldien, "On performance of hetnet with coexisting small cell technology," in *6th*

Joint IFIP Wireless and Mobile Networking Conference (WMNC). IEEE, 2013, pp. 1–8.

[7] B. Rong, X. Qiu, M. Kadoch, S. Sun, and W. Li, *5G heterogeneous networks: Self-organizing and optimization*. Springer, 2016.

[8] A. Sharma, A. Roy, S. Ghosal, R. Chaki, and U. Bhattacharya, "Load balancing in cellular network: A review," in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. IEEE, 2012, pp. 1–5.

[9] T. M. Shami, D. Grace, A. Burr, and J. S. Vardakas, "Load balancing and control with interference mitigation in 5G heterogeneous networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, pp. 1–12, 2019.

[10] A. G. Barto, "Reinforcement learning," in *Neural systems for control*. Elsevier, 1997, pp. 7–30.

[11] M. A. Wiering and M. Van Otterlo, "Reinforcement learning," *Adaptation, learning, and optimization*, vol. 12, no. 3, p. 729, 2012.

[12] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, 2011.

[13] A. Aijaz, H. Aghvami, and M. Amani, "A survey on mobile data offloading: technical and business perspectives," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 104–112, 2013.

[14] S. Yanzhi, L. Yuming, W. Quan, F. Cheng, and J. Liqiong, "Multi-factor load balancing algorithm for 5G power network data center," in *2021 3rd International Conference on Advances in Computer Technology, Information Science and Communication (CTISC)*, 2021, pp. 188–193.

[15] Y. Lin and W. Yu, "Joint spectrum partition and user association in multi-tier heterogeneous networks," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, 2014, pp. 1–6.

[16] H. Zhang, L. Song, and Y. J. Zhang, "Load balancing for 5G ultra-dense networks using device-to-device communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 4039–4050, 2018.

[17] M. K. Hasan, T. C. Chuah, A. A. El-Saleh, M. Shafiq, S. A. Shaikh, S. Islam, and M. Krichen, "Constriction factor particle swarm optimization based load balancing and cell association for 5G heterogeneous networks," *Computer Communications*, vol. 180, pp. 328–337, 2021.

[18] M. S. Ali, P. Coucheney, and M. Coupechoux, "Load balancing in heterogeneous networks based on distributed learning in near-potential games," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 5046–5059, 2016.

[19] T. Zhou, Z. Liu, J. Zhao, C. Li, and L. Yang, "Joint user association and power control for load balancing in downlink heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2582–2593, 2018.

[20] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," Tech. Rep. TR 38.901 V14.3.0 R14, 2017.

[21] B. Jabbari, Y. Zhou, and F. Hillier, "Random walk modeling of mobility in wireless networks," in *VTC'98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No. 98CH36151)*, vol. 1. IEEE, 1998, pp. 639–643.

[22] R. Sutton and A. Barto, *Reinforcement Learning, second edition: An Introduction*. MIT Press, 2018.