Pakistan Journal of Engineering and Technology, PakJET
ISSN (p): 2664-2042, ISSN (e): 2664-2050
Volume: 6, Number: 1, Pages: 92- 97, Year: 2023

Hybrid Learning based Radio Resource Management in 5G Heterogeneous Networks

Muhammad Usman Iqbal[1], Ejaz Ahmad Ansari[1], Saleem Akhtar[1], Muhammad Nadeem Rafiq[1], Muhammad Farooq-i-Azam[1], and Beenish Hassan[2]

[1]Department of Electrical and Computer Engineering, COMSATS University Islamabad - Lahore Campus, Lahore, 54000, Pakistan
[2]Department of Electrical Engineering, Faculty of Engineering, Islamia University Bahawalpur, Bahawalpur, 63100, Pakistan

Corresponding author: Muhammad Usman Iqbal (e-mail: usmaniqbal@cuilahore.edu.pk).

Received: 15/09/2022, Revised: 23/12/2022, Accepted: 06/03/2023

Abstract- **Ultradensification using different types of small cells (SCs) is one of the key enabling solutions to meet the multiple stringent requirements of 5G cellular networks. However, radio resource management (RRM) in ultra-dense heterogeneous networks (HetNets) is not easy due to interferences in multi-tiered architecture and dynamic network conditions. Interferences in 5G HetNets can be efficiently managed only through the techniques which are adaptive and self-organizing to handle dynamic conditions in 5G HetNets. In this article, a machine learning (ML) based self-adaptive resource allocation scheme is proposed based on the combination of independent and cooperative learning and evaluated for ultra-dense 5G HetNets. The proposed scheme aims to improve the QoS of all users associated with different network tiers in ultra-dense HetNets simultaneously. The proposed solution adaptively optimizes the SCs transmit power either through independent learning or cooperative learning based on the varying density of small cells to minimize the interferences and ensure minimum QoS requirements for all users in different network tiers. The proposed scheme not only maintains the minimum required capacities for QoS provision to all users simultaneously but has also shown a significant improvement in the capacities of users in different network tiers in high interference scenarios as compared to the use of a single learning scheme.**

***Index Terms*—5G, Heterogeneous Networks, Q-Learning, Small cells.**

I. INTRODUCTION

Wireless communication evolved from 1G to 5G at an exponential rate in the last three decades. Each of the previous generations from 1G to 4G was a simple enhancement of the previous generations. However, simple enhancements in 4G cannot meet the future demands of users, data rate, and capacity. Therefore, the requirements of 5G are very stringent like nearly zero latency, very high data rate, and capacity to support 100 billion devices [1]–[5]. To meet the requirements of 5G multiple enabling solutions have been proposed in the literature like massive multiple input and multiple outputs (MIMO), millimeter wave (mmW) communication, and ultradensification. Among the proposed enabling solutions for 5G cellular networks, ultradensification is the one which can provide the solution to multiple requirements simultaneously. However, co-tier and cross-tier interference resulting from the multi-tiered architecture of ultra-dense HetNets in the process of ultradensification is a performance-limiting factor. To efficiently utilize ultradensification, co-tier, and cross-tier interferences have to be mitigated simultaneously for QoS provision to all users in the different network tiers [1]–[6].

Researchers have proposed many schemes for efficient ultradensification in 5G cellular networks by considering co-tier and cross-tier interferences. However, most of the schemes either could not handle both types of interferences simultaneously or could not provide minimum QoS to all users in the network as the density of SCs increased in ultra-dense HetNets. From the literature, it can be inferred that non-adaptive schemes for resource management in ultra-dense HetNets are not effective due to continuously changing conditions. Therefore, efficient ultradensification requires an intelligent or cognitive resource allocation scheme that is adaptive to the conditions. In this context, self-organizing networks (SON) [7] combined with ML [8]–[10] are explored in this work to devise an adaptive resource allocation scheme for ultra-dense HetNets to simultaneously mitigate co-tier and cross-tier interferences to provide QoS to all users in the network.

II. RELATED WORK

Recently ML has been integrated into various communication systems to solve the optimization problems which are difficult to solve with other conventional techniques.

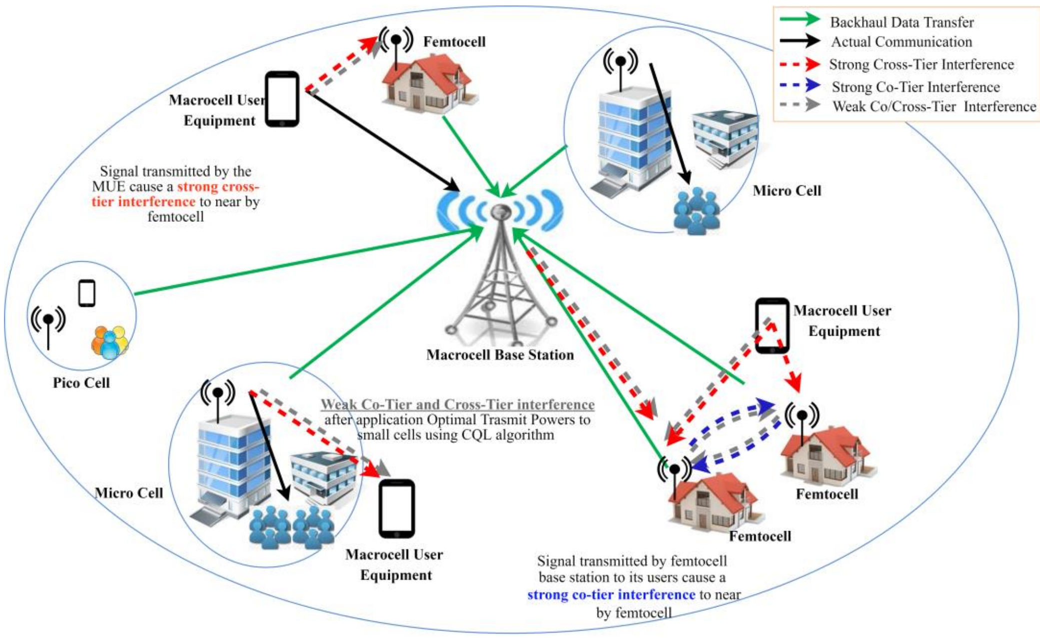

FIGURE 1. System model based on different types of small cells under laid the macrocell.

Artificial intelligence (AI) and ML are being utilized as a source of cognition in communication systems. Reinforcement Learning (RL) which is a subdomain of ML, is recently applied for RRM in ultra-dense HetNets. RL is applied through Q-Learning (QL) which is model-free, robust, and resource-efficient therefore a suitable ML technique for real-time application [8]–[10].

Multiple QL solutions have been proposed since last decade where each QL solution is different from the other based on the modeling of HetNets for implementation of QL and QL reward function (RF) designed to solve the distinct optimization problem. The QL for RRM in HetNets is either implemented through independent learning or cooperative learning [11]–[15]. Although cooperative learning has been proven as a superior learning paradigm when the small cells are deployed in the form of clusters. The improved throughput of cooperative learning is at the cost of increased computational time. However varying cluster size and non-homogeneous distribution of small cells still require investigating an optimal learning paradigm especially when the distribution of small cells is not specified like the strip model in 3GPP TR 36.872 [16].

Therefore, this work aims to handle the limitations of recently proposed solutions to work in a certain defined environment and propose a hybrid learning solution for RRM using QL which can mitigate co-tier and cross-tier interferences simultaneously and provide minimum QoS to all macrocell user equipment (MUEs) and small cell user equipment (SUEs) in a system with any distribution of SC HetNets.

III. SYSTEM MODEL

Ultra-dense HetNets system model for 5G, shown in Fig. 1, is composed of one macrocell (MC) and multiple SCs deployed in co-channel mode. Based on the specifications of 3GPP TR 36.872 [16], all SCs and related SUEs are deployed indoors. Ultradensification, or the deployment of a large number of small cells in a given area, leads to increased interference in wireless communication systems. This is because the small cells operate on the same frequency bands as macrocell, which results in interference between them. Interference can lead to a degradation of the quality of service (QoS) for users in the affected area. To provide the required minimum SINR to the MUEs and SUEs, Γ_M and Γ_K , respectively, we focus on interference mitigation through adaptive power allocation. The adaptive power allocation is considered in the downlink of the ultra-dense SC HetNets using the SON features defined in 3GPP TR 32.500 [17].

The SINR of the signal received by MUE, UE_i^m where $i \in \boldsymbol{I} = \{1,2, \dots I\}$ in the downlink includes cross-tier interference from SCs and thermal noise. The SINR at the UE_i^m, S_m, is [19]-[21]

$$S_m = \frac{P_m \left|h_{\{m,m\}}\right|^2}{\underbrace{\sum_{k \in K} P_k \left|h_{\{k,m\}}\right|^2}_{Cross\ Tier\ Interference} + N_o} \tag{1}$$

where P_m and P_k are the transmitted power by MBS, and k^{th} SC respectively, $h_{m,m}$ and $h_{k,m}$ are the channel gains from the MBS and k^{th} SC to the MUE respectively. N_o represents the variance, σ^2, of the additive white Gaussian noise (AWGN).

Like (1), the SINR at the c^{th} SUE, $UE_{c,k}^s$ where $c \in \boldsymbol{C} = \{1,2, \dots, C\}$ and $k \in \boldsymbol{K} = \{1,2, \dots, K\}$ are number of SCs and related SUEs respectively, S_k, is [19]-[21]

$$S_k = \frac{P_k \left|h_{\{k,k\}}\right|^2}{\underbrace{P_m \left|h_{\{m,k\}}\right|^2}_{Cross\ Tier\ Interference} + \underbrace{\sum_{j \in S\ and\ j \neq k} P_j \left|h_{\{j,k\}}\right|^2}_{Co-Tier\ Interference} + N_o} \tag{2}$$

where $h_{k,k} h_{j,k}$, and $h_{m,k}$ are the channel gains from the k^{th} SC, j^{th} SC, and MBS to the k^{th} SUE respectively, P_j the transmit power of the j^{th} SC, and N_o represents the variance, σ^2, of AWGN. Finally, the normalized capacity at the MUE and SC, C^m and C^k, respectively are given below as in [19]-[21]:

$$C^m = \log_2(1 + S_m) \tag{3}$$
$$C^s = \log_2(1 + S_k) \tag{4}$$

The minimum capacity for providing QoS to MUEs and SUEs, C_M and C_S, respectively, can be calculated using the (3) and (4) by inserting the minimum required SINR of MUEs and SUEs for QoS, i.e. Γ_M and Γ_K.

In the context of the above-presented system model, the adaptive power allocation problem is presented as follows.

$$\max_{P} \quad \mathrm{C^m}, \mathrm{C^s}, \mathrm{C}^s_{sum} \tag{5a}$$
$$Subject\ to \quad p_1 \le p_c^s \le p_{max}, \quad c \in \mathbf{C} \tag{5b}$$
$$\mathrm{C^m} \ge C_M \tag{5c}$$
$$\mathrm{C^s} \ge C_S \tag{5d}$$

where p_1 and p_{max} are the minimum and maximum transmit powers that an SC can select.

IV. PROPOSED METHODOLOGY

In 5G wireless communication networks, resource management, and interference mitigation can be considered as the policy function in a Markov Decision Process (MDP). MDP can be effectively solved through QL. The goal of RRM is to optimize the allocation of wireless resources, such as frequency bands and transmission power, to different users and devices in a network, in order to maximize network performance and user capacity. QL is used to learn a policy for resource allocation, where the policy is a mapping from the current network state to an action (i.e., a resource allocation decision). The goal is to learn a policy that maximizes an RF, which could be a measure of network throughput, energy efficiency, or user satisfaction. To apply QL for RRM in 5G HetNets, the following components of QL are defined by considering the system model in section II and the optimization problem in (5):

A. AGENTS

QL agents are agents that use the QL algorithm to learn the optimal policy for maximizing cumulative rewards. QL agents work by estimating the value of each possible action in a given state. In case of 5G HeNets, small cells act as the agent of QL.

B. STATES

The states of the network can include information such as the number of devices, their communication requirements, their locations, and the available wireless resources. In our case, the state of the agent, i.e. SC, is defined on the basis of the location of the SC with respect to the nearby MUE and MBS. The distances from the MBS and MUE are defined as follows on the basis of the distance rings $\boldsymbol{N_{MBS}}$ and $\boldsymbol{N_{MUE}}$ respectively.

$$\boldsymbol{D_{MBS}} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\ldots, \boldsymbol{N_{MBS}}\} \tag{6}$$
$$\boldsymbol{D_{MUE}} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\ldots, \boldsymbol{N_{MUE}}\} \tag{7}$$

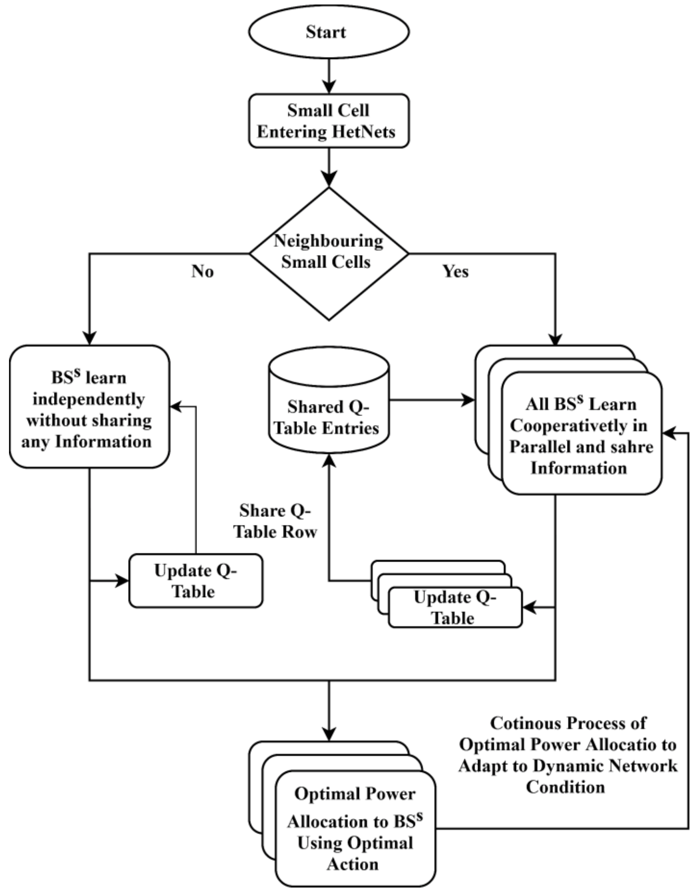

FIGURE 2. Flow chart of hybrid learning algorithm for optimal radio resource allocation in HetNets.

These parameters define the state of $\boldsymbol{i^{th}}$ SC at the time $\boldsymbol{t}$ as

$$\boldsymbol{s_i^t} = (\boldsymbol{D_{MBS}}, \boldsymbol{D_{MUE}}) \tag{8}$$

C. ACTIONS

The action space includes all possible resource allocation decisions that can be made in the current state, such as allocating a frequency band or adjusting transmission power. In case of HetNets, the transmit power for each SC is considered as the action which an agent can take. The transmit power of the SCs can be selected from an equally spaced set of transmit power levels.

$$\boldsymbol{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \ldots, \boldsymbol{a}_{\boldsymbol{N_{power}}}\} \tag{9}$$

between the $\boldsymbol{P_{min}}$ and $\boldsymbol{P_{max}}$.

D. REWARD FUNCTION

The reward signal is a measure of network performance or user satisfaction and is typically a function of the chosen action and the resulting network state.

E. Q-TABLE

In RL, an agent interacts with an environment by taking actions and receiving rewards based on those actions. The goal of the agent is to learn a policy that maximizes the cumulative reward over time. The Q-Table is a table that maps a state-action pair to the expected cumulative reward for taking that action in that state. The Q-Table is updated as the agent gathers more experience and learns the expected rewards for different actions in different states.

During the learning process, the agent explores the environment by taking actions and updating the Q-Table based on the rewards it receives. The agent then uses the Q-Table to select the best action to take in a given state based on the expected cumulative reward for each possible action.

F. PROPOSED HYBRID LEARNING BASED QL ALGORITHM

Here, first, we model the SC HetNets network as the multiagent MDP and then proposed a QL-based algorithm to solve the OP presented in (5a-5d). The QL algorithm used in this research is presented in Fig. 2.

The proposed QL algorithm uses both, independent and cooperative learning paradigms simultaneously. At the start of the algorithm, each SC detects its neighbors through Automatic Neighbor Discovery (AND) operation, which is defined by 3GPP. If an SC does not find any other SCs in its neighborhood, it works in an independent learning paradigm and learns to optimize its RF. In this case, SC does not share learned information with any other SC. On the other hand, if an SC finds other SCs in its neighborhood, then it starts learning in the cooperative paradigm where it continuously shares and accepts information from the neighboring cells. Hybrid learning which is a combination of both independent and cooperative learning in different scenarios may significantly improve the throughput of the network and also reduce the computational time in the case of using cooperative learning only.

V. RESULTS AND DISCUSSION

The proposed hybrid learning-based RRM algorithm for effective interference mitigation and QoS provision to all MUEs and SUEs is evaluated in an ultra-dense scenario, presented in Fig. 3, which is in line with the system model presented in Fig. 1. The simulation scenario is evaluated through the Monte-Carlo simulations where the location of MUEs, SCs and their related SUEs changes in each iteration. The simulation setup is composed of 16 UE_i^m and 20 BS^s where each BS^s supports 02 $UE_{c,k}^s$. The simulation parameters and channel model are according to the

TABLE I
SIMULATION PARAMETERS

Parameter	Quantity
Number of Macrocell Base station	1
Number of Small Cells	20
Number of Macrocell Users	16
Number of Small Cell Users	2 (Each Small Cell)
Coverage Radius of Macrocell	350m
Coverage Radius of Small Cell	10m
Transmit Power of Macrocell	50dBm
Transmit Power of Small Cell	-15dBm to 15dBm
Number of Power Steps, N_p	31
Γ_M	1 b/s/Hz
Γ_K	1b/s/Hz
Q-Learning Rate	0.5
Discount Factor	0.9
Number of Q-Learning Iterations	75000
Operating Frequency	2.0 GHz
Path Loss Model	3GPP TR36.872 [18]

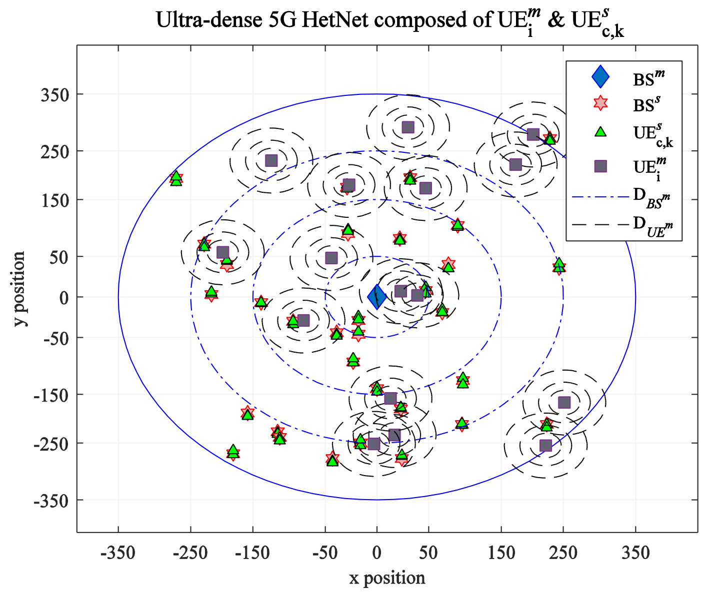

FIGURE 3. Simulation scenario based on multiple MUEs, SCs and SUES.

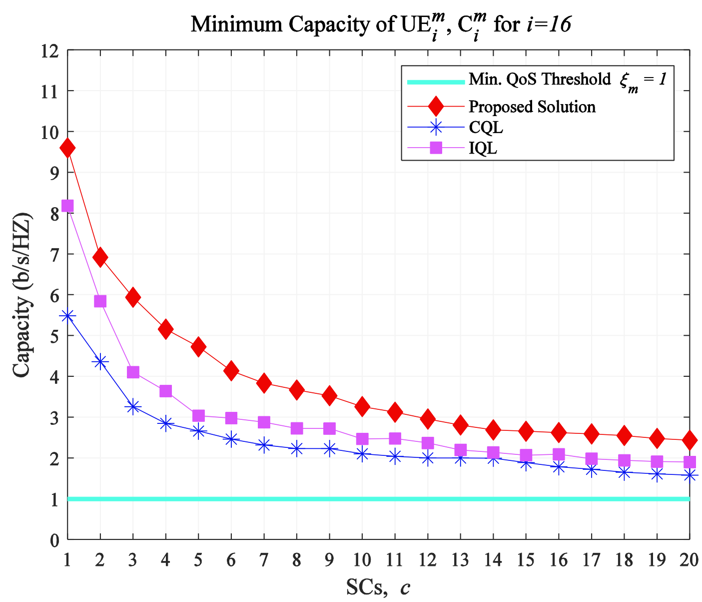

FIGURE 4. Minimum MUE Capacity for 16 MUEs and 20 SCs in the system where each SC has 02 SUES.

3GPP TR36.872 [16] whereas the QL parameters are in line with the other solutions recently proposed in the literature [12], [14], [19]-[24]. Simulation parameters are summarized in Table 1.

The simulation results of various key performance indicators (KPIs) are obtained by adding SCs in the simulation model one by one at the random location in the coverage area of MC whereas the number of UE_i^m remain constant. The locations of UE_i^m and $UE_{c,k}^s$ in the MC coverage area is kept random in each Monte-Carlo iteration. The results obtained through simulation of the proposed solution are analyzed in two ways, firstly, if the proposed solution can meet the minimum QoS requirements for all users in the network, and secondly, the results are compared with other state of the art solutions recently proposed in the literature.

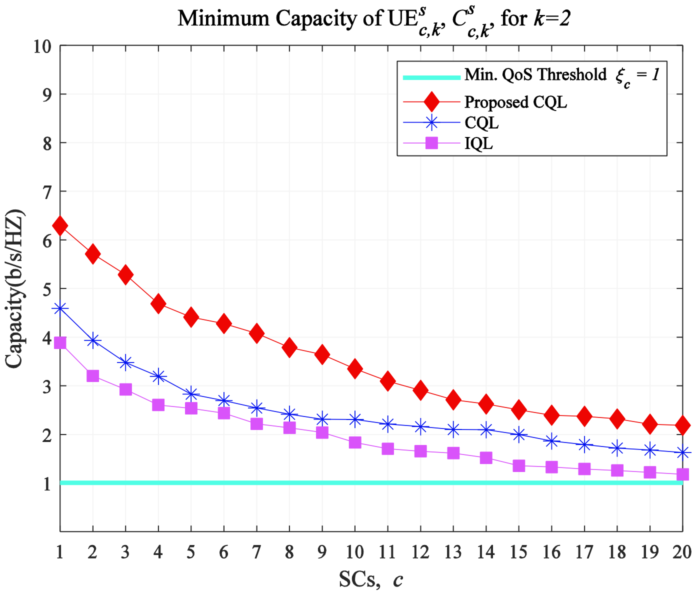

FIGURE 5. Minim SUE Capacity for 16 MUEs and 20 SCs in the system where each SC have 02 SUES.

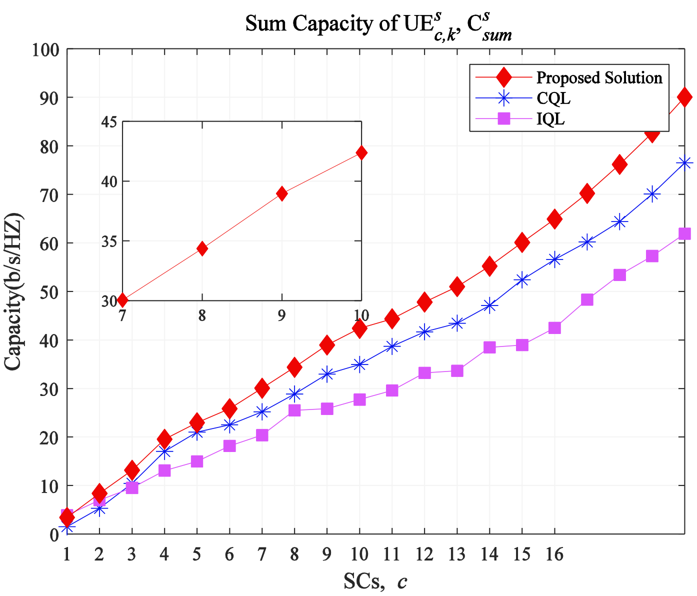

FIGURE 6. SUM SUE Capacity for 16 MUEs and 20 SCs in the system where each SC have 02 SUES.

A. MINIMUM MUE CAPACITY

The capacity of UE_i^m, C^m, is an important KPI in terms of QoS provision in co-channel deployment mode as the UE_i^m is a primary user in the network. The cross tier interference from close-by SCs severely degrades the QoS of UE_i^m. However, the proposed solution successfully handled the cross tier interference and provided minimum C^m significantly higher than the minimum QoS threshold. The simulation results of $C^m$are presented in Fig. 4. It pertains to mention that the minimum C^m is least C^m which any of the 16 UE_i^m in the system obtained. The proposed solution not only provided C^m significantly higher than the minimum QoS threshold but also outperformed the state of the art solutions IQL [19] and CQL [20].

B. MINIMUM SUE CAPACITY

To provide QoS to $UE_{c,k}^s$ the capacity of $UE_{c,k}^s$, C^s, should always be greater than the minimum QoS threshold. However, in case of SCs in ultra-dense HetNets, it is a difficult task due to dynamic network conditions and network density. The performance of the SCs are effected by the co-tier and cross tier interference from near by SCs and MUE respectively. The proposed solution not only successfully provided the minimum required C^s to all 20 $UE_{c,k}^s$ simultaneously but also outperformed other recently proposed solutions, IQL [19] and CQL [20]. The simulation results for C^s are presented in the Fig. 5.

C. SUM CAPACITY OF SUEs

The sum capacity of SUEs, C_{sum}^s is an increasing function of SCs compared to the C^m and C^s. The C_{sum}^s represents how the proposed solution improves the throughput of all SCs together. The simulation results for C_{sum}^s are in line with the C^s. The proposed solution significantly improved the C_{sum}^s as compared to the other recently proposed solutions in the literature. The comparison of C_{sum}^susing the proposed solution and other state of the art solutions is presented in Fig. 6.

VI. CONCLUSION

The hybrid QL based on the combination of IL and CL in different SC density scenarios outperformed both IL and CL when applied as a single learning scheme. The simulation results show that all three KPIs, minimum MUE capacity, minimum SUE capacity, and sum capacity of SUEs are significantly improved in the ultra-dense HetNets as compared to the IL and CL proposed recently in the literature. The proposed solution opts for the use of independent or cooperative learning based on the density of small cells. The results of the proposed hybrid learning further signify the real-time implementation of Q-learning for RRM in HetNets. In the future, the proposed hybrid QL algorithm will be evaluated in a more realistic ultra-dense scenario by considering user load balancing among the different network tiers.

FUNDING STATEMENT

This research is not funded by the university or any other funding agency.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding this research.

REFERENCES

[1] D. P. A. Rupendra Nath Mitra, "5G mobile technology: A survey," ICT Express, vol. 1, pp. 132–137, December 2015.
[2] N. Panwar, S. Sharma, and A. K. Singh, "A survey on 5G: The next generation of mobile communication," Physical Communication, vol. 18, pp. 64 – 84, 2016. Special Issue on Radio Access Network Architectures and Resource Management for 5G.
[3] I. F. Akyildiz, S. Nie, S.-C. Lin, and M. Chandrasekaran, "5G roadmap: 10 key enabling technologies," Computer Networks, vol. 106, pp. 17 – 48, 2016.

[4] A. Gupta and R. K. Jha, “A survey of 5G network: Architecture and emerging technologies,” IEEE Access, vol. 3, pp. 1206–1232, 2015.
[5] S. Manap, K. Dimyati, M. N. Hindia, M. S. Abu Talip, and R. Tafazolli, “Survey of radio resource management in 5G heterogeneous networks,” IEEE Access, vol. 8, pp. 131202–131223, 2020.
[6] K. A. Yau, J. Qadir, C. Wu, M. A. Imran, and M. H. Ling, “Cognition Inspired 5G Cellular Networks: A Review and the Road Ahead,” IEEE Access, vol. 6, pp. 35072–35090, 2018.
[7] 3GPP, “Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN), overall description,” technical report (release8), ts 36.300, 3rd Generation Partnership Project (3GPP), October 2020. Version 16.3.0.
[8] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y. Liang, and D. I. Kim, “Applications of deep reinforcement learning in communications and networking: A survey,” IEEE Communications Surveys Tutorials, vol. 21, pp. 3133–3174, Fourth quarter 2019.
[9] Y. L. Lee and D. Qin, “A survey on applications of deep reinforcement learning in resource management for 5G heterogeneous networks,” in 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1856–1862, 2019.
[10] K.-L. A. Yau, P. Komisarczuk, and P. D. Teal, “Reinforcement learning for context awareness and intelligence in wireless networks: Review, new features and open issues,” Journal of Network and Computer Applications, vol. 35, no. 1, pp. 253 – 267, 2012. Collaborative Computing and Applications.
[11] A. Galindo-Serrano and L. Giupponi, “Distributed Q-Learning for interference control in OFDMA-based femtocell networks,” in 2010 IEEE 71st Vehicular Technology Conference, pp. 1–5, May 2010.
[12] H. Saad, A. Mohamed, and T. ElBatt, “Distributed cooperative Q-Learning for power allocation in cognitive femtocell networks,” in 2012 IEEE Vehicular Technology Conference (VTC Fall), pp. 1–5, Sep. 2012.
[13] H. Saad, A. Mohamed, and T. ElBatt, “A cooperative Q-Learning approach for online power allocation in femtocell networks,” in 2013 IEEE 78th Vehicular Technology Conference (VTC Fall), pp. 1–6, Sep. 2013.
[14] J. R. Tefft and N. J. Kirsch, “Accelerated learning in machine learning based resource allocation methods for heterogenous networks,” in 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), vol. 01, pp. 468–473, 2013.
[15] Bin Wen, Zhibin Gao, Lianfen Huang, Yuliang Tang, and Hongxiang Cai, “A Q-learning-based downlink resource scheduling method for capacity optimization in LTE femtocells,” in 2014 9th International Conference on Computer Science Education, pp. 625–628, Aug 2014.
[16] 3GPP, “Small cell enhancements for E-UTRA and E-UTRAN – physical layer aspects,” Technicl Report (Release12),TR 36.872, 3rd Generation Partnership Project (3GPP), December 2013. Version 12.1.0.
[17] 3GPP, “Telecommunication management; self-organizing networks (SON); concepts and requirements,” Technicl Report (Release 8),TR 32.500, 3rd Generation Partnership Project (3GPP), July 2020. Version 12.1.0.
[18] 3GPP, “Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects,” Technicl Report (Release 9), TR 36.814, 3rd Generation Partnership Project (3GPP), March 2017. Version 9.2.0.
[19] M. U. Iqbal, E. A. Ansari, and S. Akhtar, “Interference mitigation in HetNets to improve the QoS using Q-Learning,” IEEE Access, vol. 9, pp. 32405–32424, 2021.
[20] M. U. Iqbal, E. A. Ansari, S. Akhtar, and A. N. Khan, “Improving the QoS in 5G HetNets through cooperative Q-learning,” IEEE Access, vol. 10, pp. 19654–19676, 2022.
[21] R. Amiri, H. Mehrpouyan, L. Fridman, R. K. Mallik, A. Nallanathan, and D. Matolak, “A machine learning approach for power allocation in HetNets considering QoS,” in 2018 IEEE International Conference on Communications (ICC), pp. 1–7, May 2018.
[22] R. Amiri, M. A. Almasi, J. G. Andrews, and H. Mehrpouyan, “Reinforcement learning for Self-Organization and power control of two-tier heterogeneous networks,” IEEE Transactions on Wireless Communications, vol. 18, pp. 3933–3947, Aug 2019.
[23] Q. Su, B. Li, C. Wang, C. Qin, and W. Wang, “A power allocation scheme based on deep reinforcement learning in HetNets,” in 2020 International Conference on Computing, Networking and Communications (ICNC), pp. 245–250, Feb 2020.
[24] W. AlSobhi and A. H. Aghvami, “QoS-aware resource allocation of two-tier HetNet: A Q-learning approach,” in 2019 26th International Conference on Telecommunications (ICT), pp. 330–334, April 2019.