Learning-Aided Demand-Driven Elastic Architecture for 6G & Beyond

Shahrukh Khan Kasi*, Umair Sajid Hashmi†, Sabit Ekin‡, Ali Imran*§
*AI4Networks Research Center, Electrical & Computer Engineering, University of Oklahoma, USA
†Electrical Engineering & Computer Science, National University of Sciences & Technology, Pakistan
‡Electrical & Computer Engineering, Texas A&M University, USA
§James Watt School of Engineering, University of Glasgow, UK
*{shahrukhkhankasi, ali.imran}@ou.edu, †umair.hashmi@seecs.edu.pk, ‡sabitekin@tamu.edu

2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring) | 979-8-3503-1114-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/VTC2023-Spring57618.2023.10200531

***Abstract*—With highly heterogeneous application requirements, 6G and beyond cellular networks are expected to be demand-driven, elastic, user-centric, and capable of supporting multiple services. A redesign of the one-size-fits-all cellular architecture is needed to support heterogeneous application needs. This paper addresses this need by proposing an intelligent, demand-driven, elastic user-centric cloud radio access network (UCRAN) architecture capable of providing services to a diverse set of use cases ranging from augmented/virtual reality to high-speed rails to industrial robots to E-health applications, and more. The proposed framework leverages deep reinforcement learning to adjust the size of a user-centered virtual cell based on each application's heterogeneous throughput and latency requirements. Finally, numerical results are presented to validate the convergence and network adaptability of the proposed approach against the brute-force method.**

***Index Terms*—User-centric, elasticity, demand-driven, deep reinforcement learning, spectral and energy efficiency.**

I. Introduction

A key feature of 6G and beyond networks will be ultra-dense networks offering seamless coverage, very high throughput, and ultra-low latency. Network operators are exploring ultra-dense networks to meet the ever-growing demand for throughput and latency envisioned for 6G and beyond users. While researchers in both academia and industry agree that network densification will enhance the coverage and capacity of current cellular networks, it has its own complications [1]. By densifying the network, the average distance between users and the interferring base stations reduces. This causes a shift in pathloss exponent leading to a scenario where increase in the interference from neighboring base stations overshadows benefits of the decreased average distance from serving base stations. Further, 6G communications are envisioned to cater to a wide range of user services with assorted throughput and latency requirements [2]. In order to meet this requirement, there is a need for an elastic architecture that can tailor to the needs of each service, as opposed to traditional one-size-fits-all architecture. This, along with the interference-limited nature of dense networks, has prompted a shift to a user-centric network paradigm from traditional networks.

UCRAN's ability to abate inter-cell interference and reduce deployment/operational costs makes it the ideal architecture for supporting user-centric services in dense cellular networks [3]. A typical UCRAN consists of a tier of low-density large coverage control base station (CBS) underlaid by a tier of high-density intermediate coverage switchable data base stations (DBS). UCRAN introduces a new degree of freedom that is elastic in nature, referred to as Service Zone or S-zone in this paper. S-zone is defined as the size of the user-centric virtual cell centered around scheduled user equipment(s) (UEs). In each transmission time interval (TTI), CBS activates the best DBS constituting a S-zone centered on the scheduled UE while ensuring no overlap among S-zones. With this concept, the macro-diversity gain is easily achieved through the activation of the best DBS for a scheduled UE.

With user-centric services being considered as an essential feature of future cellular communications, 6G in particular, an elastic and demand-driven UCRAN is needed in which UEs with various throughput and latency requirements are assigned different S-zones. In this study, we present such an elastic and demand-driven UCRAN model, detailed in Section III. We formulate a multi-objective optimization problem to maximize important KPIs such as area spectral efficiency, network energy efficiency, user service rate, and throughput satisfaction. The S-zone size serves as a control parameter to form a Pareto-optimal trade-off among these KPIs. The core research objective of this work is to develop a solution that can dynamically solve this multi-objective optimization problem in UCRAN to achieve a Pareto-optimal solution in real-time based on changes in the varying application demands and user mobility. This paper studies the deep reinforcement learning approach owing to its ability to adapt to dynamic environments to determine the optimal S-zone size for each QoS category intelligently so that network KPIs are maximized. The contributions of this paper are summarized as follows.

- An architecture for demand-driven elastic user-centric communication is proposed with the aim of providing on-demand services to a diverse set of user applications.
- Considering the heterogeneous user requirements in future cellular communications, a multi-objective problem is formulated to optimize KPIs as a function of S-zone size for respective QoS categories.
- Given the non-stationarity of user application demands, we propose a deep reinforcement learning framework to accurately learn the mapping of environment state and

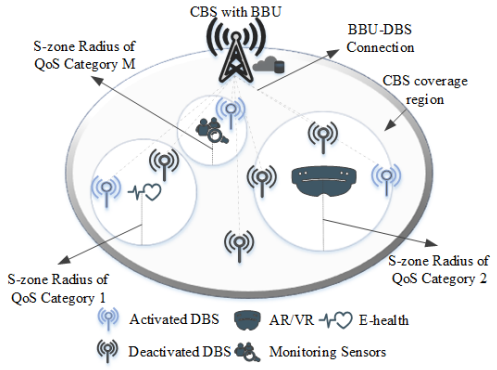


Fig. 1: Dynamic S-zone UCRAN architecture with M different S-zone region of radius R_c for scheduled UE's.

action instilling intelligence in the demand-driven elastic user-centric architecture.

The remainder of the paper is organized as follows. We discuss the related work in Section II. The system model is discussed in Section III. A multi-objective optimization problem is formulated in Section IV. The details of the proposed approach and the results of the numerical analysis are presented in Section V and Section VI, respectively. Finally, the paper is concluded in Section VII.

II. Related Work

In recent works, the impact of S-zone size in UCRAN is investigated using analytical models for both sub 6 Gigahertz and millimeter frequency bands [3]–[5]. For instance, Hashmi et. al. [3] using a statistical framework showed that there exists an optimal user-centric virtual cell size at which both the area spectral efficiency and energy efficiency can be maximized in UCRAN. The authors also noted that this user-centric virtual cell size depends on both DBS and user density variations, thus requiring adaptation with variations in these parameters. In an analytical study, we studied the interaction between spectral and energy efficiencies in a coordinated multipoint-enabled UCRAN architecture as the size of the UCRAN's virtual cell and the density of its DBSs is changed [5]. Humadi et. al. [4] have proposed a user-centric model for combining base stations for millimeter-wave networks and used stochastic geometry to determine the coverage probability and optimal area spectral efficiency performance. They propose a framework for optimizing the clustering parameter, leading to an increase in area spectral efficiency. There is insightful information available in the existing literature on UCRAN, but these methods use an analytical modeling approach that does not study the impact of control parameters (S-zone) with spatiotemporal changes in wireless networks, such as mobility and dynamic user application requirements. Owing to the cellular network's complexity, we propose a learning framework based on deep reinforcement learning to accurately learn the complex mapping of environment state and action in a demand-driven elastic user-centric architecture.

III. System Model

A. UCRAN Architecture

Fig. 1 provides a graphical illustration of a UCRAN network with virtual user-centric cell boundaries for UEs belonging to different QoS categories. These categories are classified according to the UEs' latency and throughput requirements as illustrated in Fig. 1. The DBSs are connected to the pool of base band units (BBUs) via flexible front haul (an optical fiber network) [6], [7]. Most of the signal processing at baseband level is delegated to the BBUs.

A critical design parameter in UCRAN is the size of the S-zone, which is determined by the radius of the circular disk around the UE. In the proposed model, the DBSs falling within the S-zone of a UE are only allowed to associate with that UE in a given TTI. Increasing the S-zone size ensures (i) larger distances between a UE and interfering DBSs resulting in high link-level SINR (hence, link-level high throughput and spectral efficiency); (ii) yields high macro diversity gain through selection among the larger number of DBSs in the S-zone and (iii) offers high energy efficiency as large S-zones keep more DBSs deactivated as compared to small S-zones. However, larger S-zones also yield low user scheduling ratio and low spectrum reuse resulting in negative impact on the system-level capacity. By selectively turning on DBSs in S-zones instead of keeping them always on helps reduce the burden on the network for mobility management, as activating a DBS with the maximum channel gain in each TTI within the S-zone region can provide service to UE regardless of their mobility. Given these insights, the S-zone size serves as a controlling parameter that yields an ideal tradeoff between area spectral efficiency, energy efficiency, and other system-level KPIs.

B. UE Scheduling Algorithm

In this work, we propose a scheduling mechanism to meet the heterogeneous latency requirement of UEs in UCRAN. Latency requirements of UEs are drawn from a uniform distribution and rounded off to specified bins of latency requirements corresponding to the QoS categories. Each UE x is marked with $p_x^{latency} \sim U(a,b)$ by the BBU where a and b are measured in milliseconds (ms) and are determined by the minimum and maximum latency of the considered QoS categories. The lower the value of mark $p_x^{latency} \sim U(a,b)$, the higher will be the scheduling priority. The BBU based on these scheduling priorities schedules a UE x if and if only the scheduling priority of UE x is highest in the neighborhood which is characterized by the S-zone size for a specific QoS category. This means that within a circle of radius centered at UE x, no other UE has a higher priority than UE x. Once the UE is scheduled, a single DBS providing the highest channel gain within the S-zone of the respective UE is activated by the BBU to serve the UE.

C. Network & Channel Model

A downlink of a two-tier ultra-dense network is considered consisting of a CBS and DBSs operating on sub 6 GHz frequencies. The DBSs and UEs are randomly distributed following two independent and homogeneous Poisson point

processes Π_{DBS} and Π_{UE} with intensities λ_{DBS} and λ_{UE} respectively. The location of each UE acts as a centering point for the user-centric virtual cell (S-zone) which bounds the UE to be associated with DBS only within the S-zone region. This work defines the S-zone as a disk of radius R_c, where $c \in C$ is a QoS category present in the network model. The communication channel between an arbitrary user $x \in \Pi_{UE}$ and activated DBS $i \in \Pi^{'}_{DBS}$ is modeled to experience both large-scale and small-scale fading given by hl^{-PLE}, where h is an exponential random distribution with unit mean, l_{xi} represents the propagation distance between x and i, $PLE < 2$ is the pathloss exponent, and $\Pi^{'}_{DBS}$ is the Poisson point process of activated DBSs. UE and DBS are equipped with a single antenna and the transmission power of DBS is assumed to be equal. The scheduled user SINR (Γ_x) is given as:

$$\Gamma_x = \frac{h_{xi} l_{xi}^{-PLE}}{\sum\limits_{j \in \Pi^{'}_{DBS}} h_{xj} l_{xj}^{-PLE} + n_o}, \tag{1}$$

where $i \neq j$ and n_o denotes the additive white Gaussian noise.

IV. Problem Formulation

This section characterizes the KPIs, followed by the formulation of a multi-objective optimization problem.

A. Characterizing Key Performance Indicators

This work measures system performance in terms of area spectral efficiency, network energy efficiency, user service rate, and throughput satisfaction as the desired set of KPIs. We selected these KPIs to reflect that the objective is to meet throughput and latency requirements while maximizing area spectral efficiency and network energy efficiency.

1) Area Spectral Efficiency

The area spectral efficiency refers to the amount of information that can be transmitted from a DBS per unit bandwidth channel per unit area to a UE, which can be defined as follows for each QoS category c:

$$\mathsf{A}_c = \frac{\sum\limits_{x \in N_c} \log_2(1 + \Gamma_x)}{\text{Å}}, \tag{2}$$

where N_c is the set of UEs belonging to QoS category c, and Å is the target area considered in the simulations model.

2) Energy Efficiency

According to [3], [8], [9], the network-wide energy efficiency is defined as the ratio of area spectral efficiency and total power consumed for all scheduled UE's. The power consumption model in this paper is inspired by project Earth [10], in that it represents the power consumption of CBS and DBSs as a linear combination of fixed power and load-dependent power consumption components. The total power consumption can be mathematically calculated as follows:

$$P = \lambda_{DBS} P_f + \lambda'_{DBS} \Delta_{DBS} P_{DBS} + \lambda'_{UE} (\Delta_{UE} P_{UE} + P_{disc}), \tag{3}$$

where λ_{DBS} is the density of all deployed DBSs, λ'_{DBS} is the density of activated DBSs, λ'_{UE} is the density of scheduled UEs, P_f is the fixed DBS power consumption required for DBS to operate in listening mode, P_{DBS} is the DBS transmission power, Δ_{DBS} is the radio frequency component power at DBS, P_{UE} is the UE transmission power, Δ_{UE} is the radio frequency component power at UE, P_{disc} is the power required at UE for discovery of the DBS with the highest channel gain. The typical values of these variables are summarized in [3]. The energy efficiency therefore can be given as:

$$\mathsf{E} = \frac{\text{Å} \times \sum\limits_{c \in C} \mathsf{A}_c}{P}. \tag{4}$$

3) UE Service Rate

The UEs' heterogeneous latency requirements necessitate scheduling more UEs within each TTI while meeting UE quality of experience requirements. The mean UE service rate for any QoS category c can be calculated as:

$$\mathsf{U}_c = \frac{\lambda_{UE_c}^{service}}{\lambda_{UE_c}}, \tag{5}$$

where λ_{UE_c} and $\lambda_{UE_c}^{service}$ are the densities of all UEs and whose minimum throughput requirement is met, respectively.

4) Throughput Satisfaction

There can be a wide variety of throughput requirements for UEs belonging to different QoS categories. Operators must satisfy the minimum throughput requirements of each QoS category as part of their objective. Moreover, network operators must ensure that they are utilizing their resources efficiently by avoiding scenarios in which excess throughput is allocated to a few UEs (or categories of UEs) while other UEs' minimum requirements are not met. For this reason, this work uses the difference between required and obtained throughput, a metric we define as throughput satisfaction, to measure system performance. Throughput satisfaction for a specific QoS category c is given as:

$$\mathsf{T}_c = \left(\prod_{x \in N_c} \left| tp_x^{\star} - tp_x^{\Diamond} \right| \right)^{\frac{1}{|N_c|}}, \tag{6}$$

where $tp_x^{\star}$ and $tp_x^{\Diamond}$ are the obtained and required throughput for an arbitrary UE x respectively.

5) Multi-objective Optimization Problem Formulation

Hitherto, the above definition of KPIs demonstrate the need for optimizing S-zone size of QoS categories to maximize area spectral efficiency, energy efficiency, UE service rate and throughput satisfaction individually. The challenge from a network operator's perspective is that all these KPIs should be optimized simultaneously, leading to a Pareto-optimal tradeoff between them. To account for this tradeoff, this study defines the multi-objective optimization problem as follows:

$$\max_{R_c} \quad \frac{\left(\sum\limits_{c \in C} \mathsf{A}'_c \right)^{\alpha} \left(\sum\limits_{c \in C} \mathsf{U}'_c \right)^{\beta} \left(\mathsf{E}' \right)^{1-\alpha-\beta}}{\sum\limits_{c \in C} \mathsf{T}'_c} \tag{7}$$

$$\text{s.t.} \quad R_{min} \leq R_c \leq R_{max},$$

where $0 \leq \alpha, \beta \leq 1$, $\alpha + \beta \leq 1$, A'_c is area spectral efficiency normalized between $[0, 1]$, E' is energy efficiency normalized between $[0, 1]$, U'_c is UE service rate normalized between

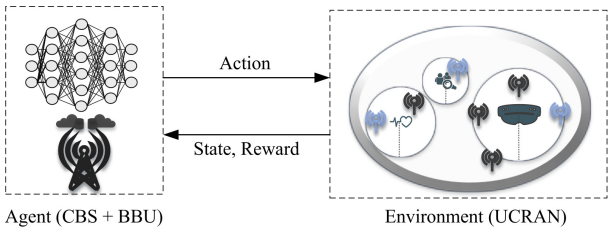


Fig. 2: Block diagram of the proposed framework.

$[0, 1]$, T'_c is throughout satisfaction normalized between $[1, 2]$, R_{min} and R_{max} are the minimum and maximum allowable size for S-zone of QoS categories. The rationale behind the proposed objective function formulation is to optimize holistic system-level performance by combining network operators' four most important and common KPIs of interest. However, these KPIs have different scales/units. This issue makes combining the multiple KPIs in a single objective function far from a straightforward problem. In this work, we address this problem by normalizing each KPI value with its minimum and maximum value. These minimum and maximum KPI values are determined through pseudo brute force method. The pseudo brute force method sweeps the solution space (with a pre-defined step size) in numerous independent runs. Given the step sizes are large enough to explore the possible extrema in the search space within an affordable computational effort, this pseudo brute force method gives values of KPIs that can be taken as approximation of minimum and maximum values for the normalization purposes [11].

To be reflective of the real goals of the system, Eq. (7) is designed such that the normalized values of area spectral efficiency (between 0 and 1), energy efficiency (between 0 and 1), and user service rate (between 0 and 1) are multiplied in the numerator to jointly maximize these KPIs while the normalized value of throughput gap (between 1 and 2) is included in the denominator to minimize the difference between throughput obtained and achieved by the users. The problem in Eq. (7) is a mixed-integer nonlinear programming problem with complexity of the order of $\mathcal{O}\big((R_{max} - R_{min} + 1)^{|C|}\big)$. It is computationally difficult to achieve an optimal solution for a non-convex multi-objective problem in a dynamically changing network, which makes its application in real-time optimization systems impossible. To this end, this paper proposes a deep reinforcement learning-based framework that is capable of determining the optimal S-zone size for all QoS categories with the objective of maximizing network KPIs.

V. Proposed Framework

This section discusses the design of the proposed framework. A BBU implements the optimization agent, which collects the network parameters and specifies the S-zone size for each QoS category.

A. State Space

- The average SINR of each QoS category is impacted by the change in S-zone size of QoS categories as divulged in Eq. (1), which has an impact on the area spectral efficiency, energy efficiency, user service rate, and throughput satisfaction. The average SINR of each QoS category can be given as:

$$\varphi_c = \frac{\sum\limits_{x \in N_c} \Gamma_x}{|N_c|}, \forall c \in C. \tag{8}$$

- The UE service rate of each QoS category given in Eq. (5) determines the ratio of UEs from each QoS category that gets served, thus directly impacting the learning objective.
- The throughput satisfaction of each QoS category given in Eq. (6) relates to how well the achieved throughput compares to the throughput demanded by UEs in each QoS category. The high value of throughput satisfaction indicates a network overshooting or undershooting throughput, which requires some adjustment of S-zones.

In conjunction, the state vector of the proposed framework with the cardinality of $3|C|$ is defined as:

$$\mathbf{s}^t = \{\varphi_1^t, ..., \varphi_{|C|}^t, \mathsf{U}_1^t, ..., \mathsf{U}_{|C|}^t, \mathsf{T}_1^t, ..., \mathsf{T}_{|C|}^t\}. \tag{9}$$

B. Action Space

For each QoS category, the action is to either increase or decrease the S-zone radius by d unit (measured in meters) or to keep it the same, that is, $\mathbf{a}_c = \{-d, 0, d\}$. Having a centralized agent responsible for adjusting the S-zone size for all QoS categories in the network will result in a combined action set. The incremental action space has been selected to circumvent the combinatorically large action space that can be obtained by considering each combination of the QoS categories as an individual action, affecting the learning and convergence of the deep reinforcement learning agent greatly. Even with the incremental action space, the size of combined action space is $3^{|C|}$ for all QoS categories, which grows exponentially with QoS categories.

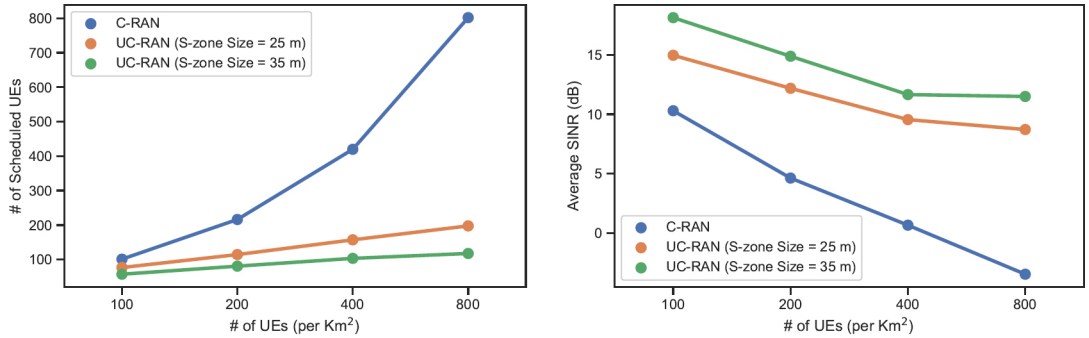
Inspired by [12] to reduce deep reinforcement learning's large action space, the action space of each QoS category is considered as a separate action branch controlling an individual degree of freedom for each QoS category. By allowing individual action dimensions to operate independently, this approach ensures a linear increase in the size of combined action space with the number of QoS categories, of the order of $2|C|+1$. For example, if $|C| = 2$, the action space includes increasing/decreasing R_1 by d meters, increasing/decreasing R_2 by d meters or keeping R_1 and R_2 unchanged. In a similar way, the combined action space dimensionality reduction approach is scalable to networks with a greater number of QoS categories.

C. Reward Function

The reward function in the proposed framework primarily focuses on two aspects for the S-zone size estimation in a dynamic environment: 1) finding the optimal trade-off between system-wide KPIs formulated as a multi-objective function given in Eq. (7), and 2) penalizing the agent for failure to satisfy the S-zone radius constraint given in Eq. (7). The utility function (u^t) at each TTI t is equivalent to the objective function given in Eq. (7). Subsequently, the reward is calculated

TABLE I: Network simulation and training parameters.

Symbol	Parameter Name	Parameter Value
λ_{UE}, λ_{DBS}	UE & DBS average density	$10^3 \backslash km^2$
PLE	Path-loss exponent	3
R_{min}	Minimum S-zone size	$10\,m$
R_{max}	Maximum S-zone size	$80\,m$
R_{init}	Initial S-zone for each QoS category	$(R_{max} + R_{min})/2\,m$
d	Action space stepsize	$3\,m$
α, β	Weightage parameters in Eq. (7)	$0.4, 0.4$
Z	Penalty for wrong action	-1



(a) Number of scheduled UEs.

(b) Average SINR (dB).

Fig. 3: Comparison of user-centric (UC-RAN) and non-user-centric (C-RAN) networks.

as follows:

$$r^t = \begin{cases} e^{\zeta(u^t - 1)} & \text{if constraint given in Eq. (7) is met.} \\ Z & \text{otherwise,} \end{cases} \quad (10)$$

where $\zeta > 1$ in the exponential term is used to amplify the difference between values of the utility function and $-1 < Z < 0$ is a negative constant to punish the agent for choosing an S-zone size that is not within the specified bounds of R_{min} and R_{max}. The exponential shaping of the reward against utility values allows the deep reinforcement learning agent to give a much higher reward when it achieves higher utility values and much lesser when it achieves lesser or mid-range utility values.

D. Agent Training & Testing Procedure

The schematic diagram of the proposed framework is shown in Fig. 2. As part of the training process, the agent stores the experience tuple $\{\mathbf{s}^t, \mathbf{a}^t, r^t, \mathbf{s}^{t+1}\}$ in the experience pool and updates the deep neural network weights by applying the stochastic gradient descent algorithm to a minibatch of data at each epoch t (equivalent to a TTI). In every episode, consisting of T epochs/TTIs, the agent is initialized at R_{init} for all QoS categories, and the environment is initialized with different random seeds to generate different mobility patterns. The deep neural network includes four fully connected layers, and three rectified linear unit activation functions with input layer neurons equal to the number of state variables $3|C|$ and output layer equivalent to the number of actions $2|C| + 1$.

VI. Experimental Evaluation

This section presents the performance of proposed framework with system model presented in Section III. The target coverage area of CBS is 1 square kilometer. The UEs and DBSs are distributed through an homogeneous Poisson point processes within the CBS coverage region with densities of $10^3 \backslash km^2$. The minimum and maximum S-zone size considered in this work are 10 meters and 80 meters, respectively with the action space step size of 3 meters. The number of maximum epochs / TTIs (T) in each training and evaluation episode is set to 1000, where each TTI's duration is set to 1 ms. The weight parameters α and β in Eq. (7) are each set to 0.4. The set of the network parameters are shown in Table I.

A. Comparison of User-centric with Non-user-centric architecture

To compare the performance of the proposed user-centric approach with a non-user-centric approach, we simulate a Cloud Radio Access Network (C-RAN) model which considers similar assumptions as taken for a user-centric architecture to ensure a fair comparison between the two architectures. These assumptions are: (i) the DBSs are deployed in high density, (ii) each UE is allocated the full bandwidth of the system, (iii) there is a one-to-one association between UE and DBS, and (iv) the UE is associated with a DBS providing the maximum channel gain. With these assumptions, the only contrasting factor in C-RAN and UC-RAN architectures is the S-zone parameter. Fig. 3 shows the average SINR and number of scheduled UEs plots for varying UE densities. It

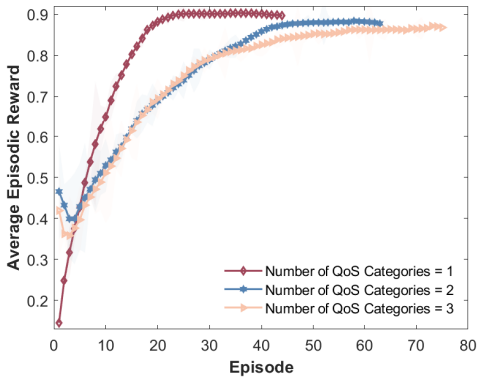


Fig. 4: Convergence of the average episodic reward values for varying number of QoS categories.

can be observed that the average SINR in the case of C-RAN falls drastically with the increase in the density of UEs in the network. At the same time, UC-RAN architecture with the additional degree of freedom (S-zone size) is able to achieve much higher average SINRs at the cost of lesser scheduled UEs. The S-zone size controls the separation between the scheduled UEs, impacting the average SINR and the number of scheduled UEs. From Fig. 3, it can be hypothesized that the C-RAN (traditional Heterogeneous network) architecture will not be able to perform better in a network with dense DBS deployment, which is envisaged for 6G and beyond networks. On the other hand, the UC-RAN architecture can provide an effective solution to this problem by incorporating an additional degree of freedom (S-zone size). Note that the C-RAN architecture is the design of traditional Heterogeneous network adopted in current cellular systems architecture.

B. Convergence Comparison for Varying Number of QoS Categories

The convergence of the proposed framework with dynamicity in the network due to heterogeneous user application demands is shown for different numbers of QoS categories in Fig. 4. The value of the utility function is normalized with the upper and lower limits, determined by the pseudo brute-force solution so that the reward function can have a maximum and minimum value of 1 and -1, respectively. For each of the considered cases in Fig. 4, the learning converges towards higher reward function values after a certain amount of training episodes. The greater the number of QoS categories, the longer it takes to converge due to a larger state space, action space, and search space, requiring more TTIs to explore the environment. Additionally, as the number of QoS categories increases, the reward function tends to converge to a lower reward value. This is mainly due to the expansion of S-zone space and the increase in the minimum required number of TTIs to reach to optimal S-zone (R_c^*) from the initial S-zone (R_c^{init}) for each QoS category.

VII. Conclusion

In this paper, we proposed a deep reinforcement learning-based user-centric RAN optimization framework under dynamic user application demands. Unlike previous cellular network approaches, the proposed framework employs a concept of elasticity within user-centric systems that employ non-uniform virtual cells (also called S-zones) for different QoS categories. The proposed framework introduces a less complex approach than brute-force technique by accurately learning the mapping of environmental conditions to S-zone size of corresponding QoS categories. In general, this research aims to introduce intelligence into user-centric elastic networks to accommodate user applications' non-uniform throughput and latency requirements. With the proposed framework, the paradigm of traditional cellular networks could be transformed into demand-driven, elastic, user-centric systems in future 6G and beyond networks. The complete version of this work has been published as a transaction paper [13].

Acknowledgements

This work is supported by the National Science Foundation under Grant Numbers 1923669 and 1923295. For more details please visit www.AI4networks.com.

References

[1] B. Romanous, N. Bitar, A. Imran, and H. Refai, "Network Densification: Challenges and Opportunities in Enabling 5G," in *2015 IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, 2015, pp. 129–134.

[2] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What Should 6G Be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, jan 2020. [Online]. Available: https://doi.org/10.1038%2Fs41928-019-0355-6

[3] U. S. Hashmi, S. A. R. Zaidi, and A. Imran, "User-Centric Cloud RAN: An Analytical Framework for Optimizing Area Spectral and Energy Efficiency," *IEEE Access*, vol. 6, pp. 19 859–19 875, 2018.

[4] K. Humadi, I. Trigui, W.-P. Zhu, and W. Ajib, "Dynamic Base Station Clustering in User-Centric mmWave Networks: Performance Analysis and Optimization," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4847–4861, 2021.

[5] S. K. Kasi, U. S. Hashmi, M. Nabeel, S. Ekin, and A. Imran, "Analysis of Area Spectral & Energy Efficiency in a CoMP-Enabled User-Centric Cloud RAN," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 4, pp. 1999–2015, 2021.

[6] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Control-Data Separation Architecture for Cellular Radio Access Networks: A Survey and Outlook," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 446–465, 2016.

[7] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning Wireless Cellular Networks of Future: Outlook, Challenges and Opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.

[8] U. S. Hashmi, A. Rudrapatna, Z. Zhao, M. Rozwadowski, J. Kang, R. Wuppalapati, and A. Imran, "Towards Real-Time User QoE Assessment via Machine Learning on LTE Network Data," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, 2019, pp. 1–7.

[9] L. Sboui, Z. Rezki, A. Sultan, and M.-S. Alouini, "A New Relation Between Energy Efficiency and Spectral Efficiency in Wireless Communications Systems," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 168–174, 2019.

[10] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How Much Energy is Needed to Run a Wireless Network?" *IEEE Wireless Communications*, vol. 18, no. 5, pp. 40–49, 2011.

[11] P. Mannion, S. Devlin, J. Duggan, and E. Howley, "Reward Shaping for Knowledge-based Multi-objective Multi-agent Reinforcement Learning," *The Knowledge Engineering Review*, vol. 33, p. e23, 2018.

[12] A. Tavakoli, F. Pardo, and P. Kormushev, "Action Branching Architectures for Deep Reinforcement Learning," 2017. [Online]. Available: https://arxiv.org/abs/1711.08946

[13] S. K. Kasi, U. S. Hashmi, S. Ekin, A. Abu-Dayya, and A. Imran, "D-RAN: A DRL-based Demand-Driven Elastic User-Centric RAN Optimization for 6G & Beyond," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2022.