DADEs: 5G Dual-Adaptive Delay-aware and Energy-saving System with Tandem Learning

Sheng Lin[1,4], Chao Qiu[1,4], Jingchao Tan[1], Xiaofei Wang[1,4], Yajun Yang[1], Ying He[2,4], Jing Jiang[3].
[1]Tianjin University, Tianjin, China
[2]Shenzhen University, Shenzhen, China
[3]XI'AN University of Posts and Telecommunications, Xi'an, China
[4]Guangdong Laboratory of Artificial Intelligence and Digital Economy, Shenzhen, China

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10001035

***Abstract*—Nowadays, numerous primary technologies, like ultra-dense networks (UDNs) and Base Stations (BSs) sleeping state, are developed in fifth-generation (5G) networks. Due to the UDNs, the number of BSs in 5G networks is proliferating, along with the energy consumption. Therefore, it is necessary to cut down the energy attrition in 5G networks under the assurance of delay. Till now, some researchers have proved that the association of users and the sleeping states of BSs have a significant effect on energy consumption and latency in 5G networks. However, the traditional solutions associate users and select states nonadaptively without the dual consideration of energy-saving and delay. In view of this, we propose a dual-adaptive delay-aware and energy-saving system (DADEs) in 5G networks. To further optimize the energy and delay of 5G BSs, the model is split into two tandem problems: user association and BS state selection. Meanwhile, a tandem deep reinforcement learning (T-DRL) algorithm is presented to make decisions in these problems for optimizing and balancing performance between delay and energy adaptively. Additionally, the real datasets of 5G users and BSs are used and trained in this paper. Finally, simulation results show that the DADEs saves more than 50% of energy with an adaptive and satisfying latency.**

***Index Terms*—5G Networks, Energy Saving, Delay-aware, Tandem Deep Reinforcement Learning**

I. Introduction

Nowadays, with the large-scale commercialization, interests in fifth-generation (5G) networks elicit escalating attention [1]. 5G networks could be regarded as the collection of primary technologies, including ultra-dense networks (UDNs), energy-aware communication, etc. [2]. In this context, the performance of 5G networks is skyrocketing, which is expected to support greater network densification, a higher density of users, as well as the lower latency [3].

However, due to the booming increase in Base Stations (BSs) and users in UDNs, energy consumption mounts up explosively [4]. It is reported that 5G networks consume about 4.7% of electricity resources and generate 1.7% of total carbon emissions in the world [5] and 15% of BS energy is wasted due to the fixed user association in 5G Networks [6]. Hence, user association and energy control for 5G BSs couldn't be ignored. On the other hand, the essential applications of 5G networks are low-latency communications. It follows that these applications require a minor-delay web environment [7]. These challenges have spawned the research hotspots, which is how to control BSs adaptively and intelligently [8].

Traditionally, numerous studies tried to solve the above challenges. For example, the authors in [9] switched off BSs by setting the traffic threshold for energy-saving. Without predicting network traffic, such frequent starting and shutting operations overturn the benefit of energy-saving due to the extra delay and energy. The works in [10]–[12] used reinforcement learning to control BS states, where the data were simulated and trained. The authors in [13] used renewable energy in BSs, which aimed to achieve zero carbon emission. The works in [14], [15] saved energy by controlling the frequency of BSs according to different applications. And the works in [16] put forward meta-heuristic algorithms to optimize the green deployment of BSs.

Although the above traditional methods enable saving energy, numerous nontrivial issues in the current methods prevent them from being used as a generic scheme for the realistic scene, including:

1) Traditional methods are unable to operate BSs adaptively according to traffic in 5G Networks. Fixed switching-off strategies obtained by history traffic are ineffective in some emergencies, such as holidays and events.

2) Users association in 5G Networks is ignored in the optimization of energy and delay. Traditionally, the fixed user association strategy is used in 5G networks, e.g., the maximum received signal strength (max-RSS) [17]. However, with the dense distribution of BSs, users have more choices in establishing the data connections.

3) Traditional energy-saving strategies lack a real 5G dataset. Until now, most of the studies are conducted by emulated data, which can not reflect the real networks.

This paper proposes a dual-adaptive delay-aware and energy-saving system (DADEs) with tandem deep reinforcement learning (T-DRL) in 5G networks. Specifically, the T-DRL makes strategies for user association and BS state selection problems. Additionally, the real 5G datasets are used and trained in the simulation. The DADEs puts the

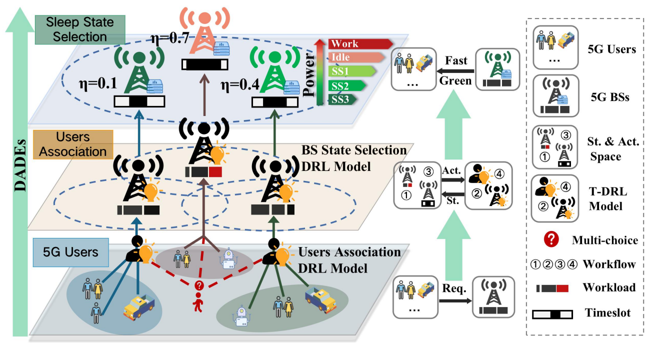


Fig. 1. 5G Dual-Adaptive Energy-saving and QoS-aware System

total requests into the user association model and compute association strategies in each timeslot. After that, entire workloads are aggregated and passed to the BS state selection model. Finally, according to different delay sensitivity factors, DADEs changes BSs states by state-selecting strategies adaptively.

The contributions of this paper are as follows.

- **We propose the DADEs in 5G networks.** We optimize the energy consumption and ensure the low latency of 5G networks from two aspects, the user allocation problem and BS state selection problem adaptively.
- **We first attempt to study energy and delay optimization problems by the real 5G dataset.** The dataset contains the data of 5G BSs over two months, which includes basic information about users and the statistics of uplink and downlink in each 5G BS.
- **We propose the T-DRL to solve user allocation and BS state selection problems.** The simulations show that with the low-latency assurance, the T-DRL saves more than 50% of energy than the base models.

The rest of the paper is organized as follows. The system model is presented in Section II, including *the 5G energy-saving sleep states*, *networking traffic model* and *energy and delay model*. Then we formulate the optimization problem in Section III. Section IV introduces the T-DRL algorithm, containing the user allocation algorithm and BS state selection algorithm. We display the simulation results in Section V and conclude in Section VI.

II. System Model

The DADEs consists of massive 5G BSs, users and a T-DRL module. The workflow of DADEs is described as follows: ① When users request data from 5G networks, the basic information (e.g., GPS location, data size, etc.) will be sent to the T-DRL. And then, the T-DRL collects the information and puts them into the user association model. ② The user association model generates association strategies decide each user establish a unique association to 5G BS. ③ After determining the association, the T-DRL collects the workload from BSs and computes the state strategies according to the delay sensitivity factor in the state selection model. ④ Finally, when BSs have completed the transmission in a timeslot, BSs select a specific sleep state. The details of the DADEs are shown in Fig. 1.

TABLE I
5G Base Station State Information [19].

State	Active		SS1	SS2	SS3	SS4
	Work	Idle				
Power	207W	132W	82.2W	35.5W	13.3W	9.51W
(de)Activation duration			35.5us	0.5ms	5ms	0.5s

A. 5G Energy-saving Sleep States

Besides work and idle state, the authors in [18] advised that different sleep states (SS) in 5G BSs could save energy when the BS is not working. For each state, the (de)activation duration and the power of each state are different. Specific information is shown in Table I.

It can be seen from Table I that when the BS enters a deeper sleep state, more energy is saved. However, when users request data from the BS, they also need to wait longer because the BS takes longer to activate from the sleep state.

B. Networking Traffic Model

A large-scale and ultra-dense 5G cellular networks consist of M 5G BSs, denoted by $\mathcal{M} = \{ m_i \mid i = 1, 2, ..., M\}$, each BS has bandwidth B_{m_i}. Meanwhile, there are N 5G users, denoted by $\mathcal{N} = \{ n_j \mid j = 1, 2, ..., N\}$ distributed in the networks. Users have different request sizes and locations, which are represented as d_{n_j} and x_{n_j}, y_{n_j}, respectively. All BSs work according to the timeslot denoted by $\mathcal{T} = \{ \tau \mid 1, 2, \cdots, T\}$.

In the traffic model, users request data from the designated BS, and each 5G BS serves multiple users simultaneously. Given a bandwidth B_{m_i} and the noise spectral density N_0, the rate of user n_j served by BS m_i (r_{m_i,n_j}) is:

$$r_{m_i,n_j} = B_{m_i} \cdot log_2(1 + \frac{P_{m_i} \cdot h_{m_i,n_j}}{N_0 \cdot B_{m_i} + \sum_{k \neq i} P_{m_k} \cdot h_{m_i,n_j}}) \quad (1)$$

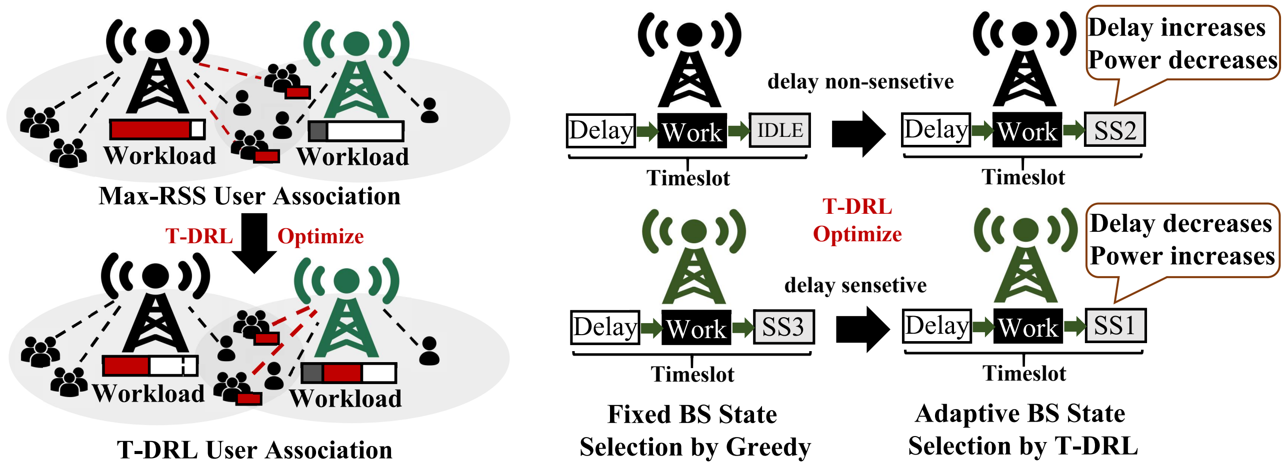
where P_{m_i} is the cell transmit power and h_{m_i,n_j} is the user channel gain served by BS m_i.

After known the rate of user n_j, the request time of user n_j is: $t_{m_i,n_j} = d_{n_j}/r_{m_i,n_j}$, where d_{n_j} represents the request size of user n_j. If the BS completes all the datacomms in a timeslot t, it will select a specific sleep state until the next timeslot. Otherwise, it will be working during the whole timeslot. The working ($t^w_{m_i}$) and sleeping time ($t^s_{m_i}$) of BS m_i are:

$$t^w_{m_i} = t^a_{m_i} + \sum_{n_j \in \mathcal{N}} t_{m_i,n_j}, \quad (2)$$

$$t^s_{m_i} = \begin{cases} t - t^a_{m_i} - \sum_{n_j \in \mathcal{N}} t_{m_i,n_j} & if\ t \geq t^w_{m_i} \\ 0 & otherwise \end{cases} \quad (3)$$

where $t^a_{m_i}$ is activation duration referred from TABLE I.



(a) Optimization of User Association from max-RSS to T-DRL.

(b) Optimization of BS State Selection from Greedy to T-DRL.

Fig. 2. User Association and State Selection Problems.

C. Energy and Delay Model

Since the BS selects a SS after working, it needs to activate from the SS before the next timeslot. According to the TABLE I, the deeper the SS is, the longer the activation time are. Therefore, when calculating the user delay of a BS, both the time for transmission and activation should be considered. Moreover, because all the users in BS m_i need to wait for activation, a total activated delay is the activation duration multiplied by the user number (s_{m_i}). The delay is:

$$D = \sum_{m_i \in \mathcal{M}} D_{m_i} = \sum_{m_i \in \mathcal{M}} (s_{m_i} \cdot t^a_{m_i} + t_{m_i}). \tag{4}$$

The energy consumed by the BSs in a timeslot is the sum of working energy (E^w) and sleeping energy (E^s). According to the equation 3, timeslot can be divided into working and sleeping time for most BSs, which have different power p^w and p^s referred from TABLE I. Specifically, working time includes the activation duration ($t^a_{m_i}$) and request time (t_{m_i}); and the BS sleeps during the remaining time in the timeslot. Therefore the energy consumption E describes as:

$$\begin{aligned} E &= \sum_{m_i \in \mathcal{M}} E_{m_i} = \sum_{m_i \in \mathcal{M}} (E^w_{m_i} + E^s_{m_i}) \\ &= \sum_{m_i \in \mathcal{M}} (\, p^w \cdot t^w_{m_i} + p^s \cdot t^s_{m_i} \,). \end{aligned} \tag{5}$$

III. Problem Formulation

The paper introduces the energy-delay reward (EDR) to optimize energy and delay jointly.

$$EDR = \eta \cdot D + (1 - \eta) \cdot E, \tag{6}$$

where $\eta \in$ [0,1] is the **delay sensitivity factor** to balance the effect between delay and energy. In particular, when η approaches 0, the users are insensitive to delay, and the DADEs will pay all attention to saving energy. On the contrary, when η comes to 1, the system will focus on reducing latency without considering energy consumption.

To optimize the EDR, we divide the optimization problem into the user association problem and the BS state selection problem. These problems will be optimized jointly, achieving the minimization of EDR. These problems will be introduced as follows respectively.

A. User Association Problem

In the user association problem, which is described in the Fig. 2(a), the T-DRL module computes the association strategies $\Pi_u = \{\, \pi_{n_j} \mid n_j \in \mathcal{N} \}$.

In general, in the user association problem, the objective is balancing workload and minimizing the working time for all BSs by optimizing the association from max-RSS to T-DRL algorithm, which is expressed as the following equation:

$$\begin{aligned} \Pi_u &= \arg\min_{\Pi_u} \sum_{m_i \in \mathcal{M}} t^w_{m_i} \\ &= \arg\min_{\Pi_u} \sum_{m_i \in \mathcal{M}} (t^a_{m_i} + \sum_{n_j \in \mathcal{N}} t_{m_i,n_j}) \\ &= \arg\min_{\Pi_u} (\sum_{m_i \in \mathcal{M}} t^a_{m_i} + \sum_{m_i \in \mathcal{M}} \sum_{n_j \in \mathcal{N}} t_{m_i,n_j}). \end{aligned} \tag{7}$$

B. BS State Selection Problem

After associations have been solved, the DADEs enters the BS state selection problem. Energy and delay are optimized jointly here, and the *EDR* is used to evaluate the state selection performance, which represents as follows:

$$\begin{aligned} EDR &= \sum_{m_i \in \mathcal{M}} EDR_{m_i} \\ &= \sum_{m_i \in \mathcal{M}} (\eta \cdot D_{m_i} + (1 - \eta) \cdot E_{m_i}) \\ &= \sum_{m_i \in \mathcal{M}} [\, \eta \cdot (s_{m_i} \cdot t^a_{m_i} + t_{m_i}) \\ &\quad + (1 - \eta) \cdot (p^w \cdot t^w_{m_i} + p^s \cdot t^s_{m_i}) \,]. \end{aligned} \tag{8}$$

The DADEs generates strategies $\Pi_s = \{\, \pi_{m_i} \mid m_i \in \mathcal{M}\}$ and controls 5G BSs entering into specific sleep state

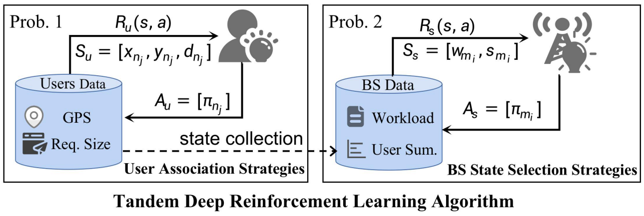


Fig. 3. Tandem Deep Reinforcement Learning Algorithm

adaptively rather than being idle or SS3. The optimization is described in the Fig. 2(b).

Finally, in the BS state selection problem, the optimization objective is expressed as follows:

$$\begin{aligned}\Pi_s &= \arg\min_{\Pi_s} EDR \\ &= \arg\min_{\Pi_s} \sum_{m_i \in \mathcal{M}} EDR_{m_i} \\ &= \arg\min_{\Pi_s} \sum_{m_i \in \mathcal{M}} (\eta \cdot D_{m_i} + (1-\eta) \cdot E_{m_i}) \\ &= \arg\min_{\Pi_s} \sum_{m_i \in \mathcal{M}} [\ \eta \cdot (s_{m_i} \cdot t^a_{m_i} + t_{m_i}) \\ &\qquad + (1-\eta) \cdot (p^w \cdot t^w_{m_i} + p^s \cdot t^s_{m_i})\ ].\end{aligned} \tag{9}$$

IV. Tandem Deep Reinforcement Learning

The T-DRL solves problems in Section III. It contains two DRL models in a tandem form, and each DRL model optimizes the user association problem and the BS state selection problem. The details of the T-DRL are shown in Fig 3, and each DRL model will be introduced as follows.

A. Deep Reinforcement Learning for User Association

In the DRL model of user association, the state space S_u, action space A_u and reward function R_u are expressed by the following equations. The state S_u is represented as:

$$S_u = \{\ [x_{n_j}, y_{n_j}, d_{n_j}] \mid n_j \in \mathcal{N}\}, \tag{10}$$

where $[x_{n_j}, y_{n_j}, d_{n_j}]$ stand for the position and request size of user n_j in a timeslot.

The action A_u is denoted as:

$$A_u = \{\ \pi_{n_j} \mid \pi_{n_j} \in [0, num_u], \pi_{n_j} \in \Pi_u, n_j \in \mathcal{N}\}. \tag{11}$$

π_{n_j} is the association strategy for user n_j, which indicates that user n_j is associated with the π_{n_j} closest BS. Meanwhile, num_u stands for the number of BSs for each user to select.

The DRL model for user association minimizes the sum of working time for all BSs, so R_u is denoted as:

$$R_u(s, a) = \{\ -\sum_{m_i \in \mathcal{M}} t^w_{m_i} \mid s \in S_u, a \in A_u\}. \tag{12}$$

Finally, the DRL model of user association updates by:

$$\mathbb{E}_{\theta_u}[\nabla_{\theta_u} log\theta_u(s, a) R_u^{\theta_u}(s, a)], \tag{13}$$

and θ_u is the DRL weights of user association model.

Algorithm 1: Tandem Deep Reinforcement Learning

Input: 5G user and BS dataset which contain:
1) Two-dimensional position of BSs X_m,Y_m
2) Two-dimensional position of Users X_n,Y_n
3) request size of Users D_n

```
Initialize the T-DRL for user association and BS state
  selection with random weights θ_u and θ_s
for episode=1,2,3,...,N do
    for timeslot τ = 1, 2, ..., T do
        for each user n_j ∈ N do
            User n_j upload position x_{n_j}, y_{n_j} and request
              size d_{n_j} to the T-DRL.
        end
        T-DRL forms users' data into S_u.
        Put S_u into the user association model.
        Compute association strategies A_u.
        BSs establish associations according A_u.
        Calculate R_u according to equation (12)
        Update θ_u according to equation (13)
        for each BS m_i ∈ M do
            BS m_i uploads the sum of users and workload
              to the T-DRL.
        end
        Put S_s into BS state selection model.
        Generate association strategies A_s.
        Change BSs state according A_s.
        Calculate R_s according to equation (16)
        Update θ_s according to equation (17)
    end
end
```

B. Deep Reinforcement Learning for BS State Selection

In BS state selection model, the state space S_s, action space A_s and reward function R_s are expressed as follows.

The state space S_s is described as:

$$S_s = \{\ [w_{m_i}, s_{m_i}] \mid m_i \in \mathcal{M}\}, \tag{14}$$

where w_{m_i} and s_{m_i} are the workloads and sum of users in a timeslot served by BS m_i. And the action A_s is:

$$A_s = \{\ \pi_{m_i} \mid \pi_{m_i} \in [0, num_s], \pi_{m_i} \in \Pi_s, m_i \in \mathcal{M}\}. \tag{15}$$

Since the BS states are defined in [19] and the activation duration must be shorter than the timeslot set as 100ms, the num_s is fixed in the paper. Specifically, when $\pi_{m_i} = 0$, the BS m_i selects idle state during sleep time, when $\pi_{m_i} = 1, 2, 3$, the BS m_i selects SS1, SS2 and SS3 respectively.

The DRL model for BS state selection minimizes the EDR we proposed in equation (9), so the R_s is defined as follows:

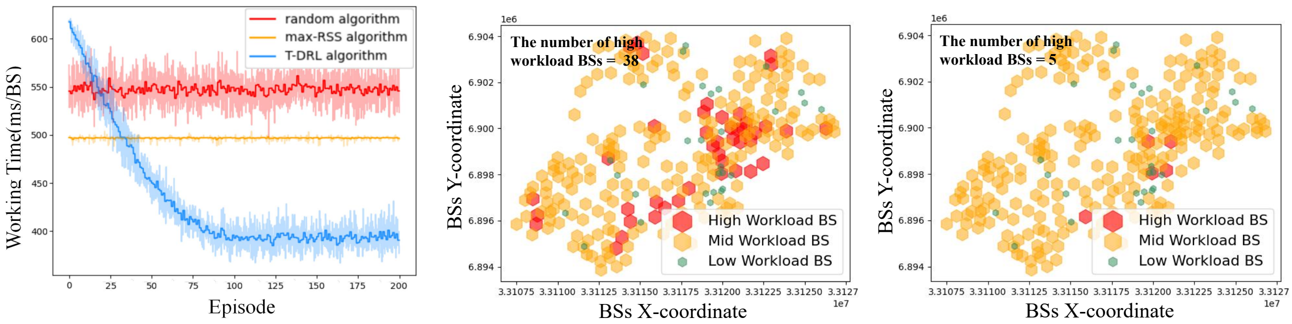
$$R_s(s, a) = \{\ -EDR \mid s \in S_s, a \in A_s\}. \tag{16}$$

Eventually, BS state selection DRL model updates by:

$$\mathbb{E}_{\theta_s}[\nabla_{\theta_s} log\theta_s(s, a) R_s^{\theta_s}(s, a)], \tag{17}$$

where θ_s is the DRL weights of BS state selection model.

In summary, the T-DRL solves the energy consumption and delay problems using the tandem-DRL models. By split into two tandem models, the T-DRL can dually optimize energy consumption and latency in a timeslot. Meanwhile,



(a) Working Time in 5G Networks. (b) User Association by MaxRss. (c) User Association by T-DRL.

Fig. 4. Simulation Results of User Association Model.

TABLE II
STATISTICS FROM 5G BASE STATIONS AND USERS DATA SET

Raw Data Details	
Date	28/03/2021 - 28/05/2021
Raw Dataset Size	2.99 GB
Num. of Users in each timeslot	2500
Total Num.of Users	5500
Total Num.of Requests	4 Million
Total Num.of 5G BSs	278
Total Size of Downlink Data	1 PB

due to a tandem form, the optimization of the first model further increases the effectiveness of the second, which ultimately improves the final optimization of the DADEs. The details of T-DRL refer to Algorithm 1.

V. SIMULATION RESULTS

A. Database Introduction

0.4cm The dataset contains 5G BSs and users data from $28th$ March 2021 to $28th$ May 2021, where the data of 5G BSs includes the geographic location, the number of users served by each BS, and the size of workloads in 5G networks. The dataset of users consists of the geographic location and the request size. The statistics of the dataset are shown in TABLE II, with around 2.99 GB of raw data. The datasets are mainly used for model training and evaluation.

B. Simulation Setup

The DADEs and T-DRL are simulated and trained in a virtual machine. To show the superiority of the T-DRL, the random and the max-RSS algorithm [17] are used as base models in the user association problem, and the greedy [20] and single DRL algorithm [10] is additionally used as baseline models in the BS state selection problem. In the simulation, working time and EDR are used to evaluate the performance of user association and state selection models. Specifically, policy gradient (PG) is used as the DRL model.

In order to simulate a real 5G communication environment, the average data transmission rate and other parameters are calculated by the real 5G dataset. Specifically, the bandwidth B_m for each BS is 10 MHz, the length of timeslot is 100ms [19], η equals 0.7 and num_c is 4 [19].

C. Simulation Result

We compare the working time and the EDR with different algorithms in the user association and BS state selection problem. The strategies generated by the models are saved and visualized through follow figures.

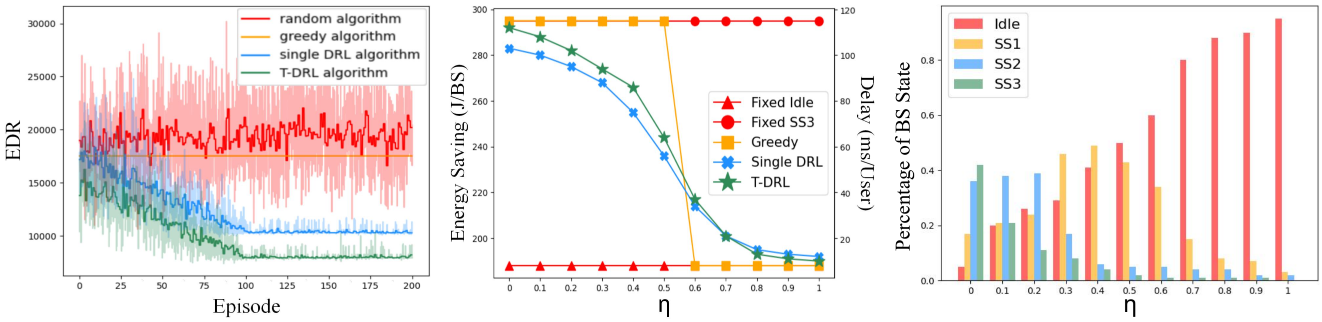
1) *User Association Simulation:*

In the user association simulation, Fig. 4(a) shows that the DRL model converges after 100 episodes and reduces the working time for BSs effectively. Due to the random algorithm making strategies uncertainly, the working time is much greater than others. In the max-RSS algorithm, all users are associated with the fixed BS, which could not reduce the working time effectively when the networks are busy. In the T-DRL, most users are still associated with the max-RSS BS. However, some users will be associated with the other BS adaptively when the traffic increases, which achieves the best performance in minimizing working time.

To reflect the superiority of the T-DRL, the workloads by max-RSS and T-DRL are plotted. The performance is shown in Fig. 4(b) and Fig. 4(c). The red hexagon represents the high workload BS in these figures, which needs a longer time to work. The yellow hexagon represents the mid workload BS, and the green is the low workload BS. These figures clearly show that when the number of users is fixed, the number of high workload BSs in T-DRL is much less than the max-RSS, and the user association is more balanced in 5G networks.

2) *BS State Selection Simulation:*

To evaluate the energy-saving and delay-aware performance between different algorithms, the EDR is drawn in Fig. 5(a). In this figure, the red line is the EDR in the random algorithm where the BS selects states randomly. The yellow stands for the EDR with the greedy algorithm that BSs always choose the idle state after working if the η is more than 0.5; otherwise, they will select SS3. Meanwhile, to show the better performance of T-DRL, the single DRL is also used as the baseline. The blue line describes the EDR by the single DRL, and the green is the T-DRL we proposed. The figure shows that the T-DRL achieves the most incredible performance in EDR for 5G networks.



(a) EDR in 5G Networks. (b) Energy Saving and Delay. (c) Percentage of BS State.

Fig. 5. Simulation Results of BS State Selection Model.

Fig. 5(b) evaluates the adaptivity of energy-saving and delay-aware between the T-DRL and others with different delay sensitivity factors η in 5G networks. Fixed idle algorithms maintain BSs on idle without consideration of η, which achieves the lowest delay and most minuscule energy saving, and the fixed SS3 is the contrary. Compared with the fixed and greedy algorithms, single DRL and T-DRL control BSs more adaptively by the delay sensitivity factor. And the T-DRL performs better in energy-saving and delay than the single DRL. The distributions of BS state selection strategies by T-DRL are plotted in Fig. 5(c). When η approaches 1, DADEs controls more BSs being idle to guarantee the low delay in 5G networks. Conversely, when η comes to 0, major BSs enter the deeper sleep state to save energy.

VI. Conclusion

In paper, we proposed a dual-adaptive delay-aware and energy-saving system in 5G networks. Then we designed the tandem deep reinforcement learning algorithm to optimize the 5G BS energy consumption and delay. We use a real 5G dataset for training and evaluating the T-DRL model in simulation. The simulation results show that with a different delay sensitivity factor, DADEs with the T-DRL saves more than 50% of energy with an adaptive delay in 5G networks than the greedy, random and single-DRL algorithm.

VII. Acknowledgment

This work was supported in part by the Open Research Fund from Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ) under Grant GML-KF-22-03, in part by the National Natural Science Foundation of China under Grant 62072332, and Grant 62002260, in part by the China Postdoctoral Science Foundation under Grant 2020M670654.

References

[1] D. López-Pérez, A. De Domenico, N. Piovesan, G. Xinli, H. Bao, S. Qitao, and M. Debbah, “A survey on 5G radio access network energy efficiency: Massive mimo, lean carrier design, sleep modes, and machine learning,” *IEEE Commun. Surv. & Tutor.*, 2022.

[2] M. Agiwal, A. Roy, and N. Saxena, “Next generation 5G wireless networks: A comprehensive survey,” *IEEE Commun. Surv. & Tutor.*, vol. 18, no. 3, pp. 1617–1655, 2016.

[3] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, “Channel state information prediction for 5G wireless communications: A deep learning approach,” *IEEE Trans. on Netw. Sci. & Eng.*, vol. 7, no. 1, pp. 227–236, 2018.

[4] A. He *et al.*, “Spectrum and energy efficiency in massive mimo enabled hetnets: A stochastic geometry approach,” *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2294–2297, 2015.

[5] E. Gelenbe *et al.*, “The impact of information technology on energy consumption and carbon emissions,” *IEEE Ubiquity*, vol. 2015, no. June, pp. 1–15, 2015.

[6] M. Diamanti, M. Tsampazi, and E. E. Tsiropoulou, “Energy efficient multi-user communications aided by reconfigurable intelligent surfaces and uavs,” in *IEEE SMARTCOMP*, 2021, pp. 371–376.

[7] T. Ji, C. Luo, L. Yu, Q. Wang, S. Chen, A. Thapa, and P. Li, “Energy-efficient computation offloading in mobile edge computing systems with uncertainties,” *IEEE Trans. on Wirel. Commun.*, 2022.

[8] C. Qiu, H. Yao, X. Wang, N. Zhang, F. R. Yu, and D. Niyato, “AI-chain: Blockchain energized edge intelligence for beyond 5G networks,” *IEEE Netw.*, vol. 34, no. 6, pp. 62–69, 2020.

[9] J. Peng, P. Hong, and K. Xue, “Stochastic analysis of optimal base station energy saving in cellular networks with sleep mode,” *IEEE Commun. Letters*, vol. 18, no. 4, pp. 612–615, 2014.

[10] A. El-Amine, M. Iturralde, H. A. H. Hassan, and L. Nuaymi, “A distributed q-learning approach for adaptive sleep modes in 5G networks,” in *IEEE WCNC*, 2019, pp. 1–6.

[11] N. Islam, A. Alazab, and M. Alazab, “A reinforcement learning based algorithm towards energy efficient 5G multi-tier network,” in *IEEE CCC*, 2019, pp. 96–101.

[12] A. El-Amine, H. A. H. Hassan, M. Iturralde, and L. Nuaymi, “Location-aware sleep strategy for energy-delay tradeoffs in 5G with reinforcement learning,” in *IEEE PIMRC*. IEEE, 2019, pp. 1–6.

[13] A. Israr *et al.*, “Renewable energy powered sustainable 5G network infrastructure: Opportunities, challenges and perspectives,” *J. Netw. & Comput. Appl.*, vol. 175, p. 102910, 2021.

[14] J. An, Y. Zhang, X. Gao, and K. Yang, “Energy-efficient base station association and beamforming for multi-cell multiuser systems,” *IEEE Trans. Wirel. Commun.*, vol. 19, no. 4, pp. 2841–2854, 2020.

[15] Y. Zhu and S. Wang, “Joint traffic prediction and base station sleeping for energy saving in cellular networks,” in *IEEE ICC*, 2021, pp. 1–6.

[16] H. Ganame, L. Yingzhuang, and H. Ghazzai, “5G base station deployment perspectives in millimeter wave frequencies using meta-heuristic algorithms,” *Electronics*, vol. 8, no. 11, p. 1318, 2019.

[17] D. Liu *et al.*, “User association in 5G networks: A survey and an outlook,” *IEEE Commun. Surv. & Tutor.*, vol. 18, no. 2, pp. 1018–1044, 2016.

[18] B. Debaillie, C. Desset, and F. Louagie, “A flexible and future-proof power model for cellular base stations,” in *IEEE VTC Spring*. IEEE, 2015, pp. 1–7.

[19] A. El Amine, J.-P. Chaiban, H. A. H. Hassan, P. Dini, and L. Nuaymi, “Energy optimization with multi-sleeping control in 5G heterogeneous networks using reinforcement learning,” *IEEE Trans. on Netw. & Serv. Manag.*, 2022.

[20] P. Lähdekorpi, M. Hronec, P. Jolma, and J. Moilanen, “Energy efficiency of 5G mobile networks with base station sleep modes,” in *IEEE CSCN*, 2017, pp. 163–168.