Energy-Efficient Optimization in O-RAN for Intelligent IoT Systems

Mukesh Kumar Maheshwari*†, Alessandro Raschellà*, Michael Mackay*, Ahmed Sabah Mohammed‡,
and Amer Salih‡
*School of Computer Science & Mathematics, Liverpool John Moores University, Liverpool, L3 3AF, UK.
‡School of Engineering, Liverpool John Moores University, Liverpool, L3 3AF, UK.
†Department of Electrical Engineering, Bahria University, Karachi Campus, 75260, Pakistan.
{m.k.maheshwari, a.raschella, m.i.mackay, a.s.mohammed, a.m.salih}@ljmu.ac.uk

***Abstract*—This paper analyzes two crucial aspects towards the development of upcoming massive cellular Internet of Things (IoT) connections: Open Radio Access Network (O-RAN) and Energy Efficiency (EE). O-RAN offers flexibility, interoperability, multi-connectivity and an intelligent architecture. On the other hand, 80% of the energy in mobile networks is consumed in the RAN. Therefore, in this paper we present a method to optimize the O-RAN energy efficiency. Specifically, we model the O-RAN energy consumption as an optimization problem and propose a traffic steering method for maximizing energy efficiency. The performance of the proposed method is analyzed by considering realistic parameters and device locations based on the deployments defined in the Liverpool 5G High Density Deployment (HDD) project. The proposed Algorithm 2 achieves 34.69% higher energy efficiency as compared to Algorithm 1.**

***Index Terms*—O-RAN, Energy Efficiency, xApp, RIC.**

I. INTRODUCTION

In recent years, the Internet of Things (IoT) has continued to roll out around the world. According to Ericsson's forecast, by 2029, there will be 6.7 billion cellular IoT connections, out of which 60% are forecast to be broadband IoT connections [1]. To support these massive connections, a smart, green, and flexible wireless network architecture is required. The recent Open Radio Access Network (O-RAN) specifications, which promote the evolution of RAN architecture that can be leveraged also in IoT systems [2], [3]. On the other hand, the sole use of O-RAN does not guarantee meeting the green requirements for massive connections. For instance, the Beyond 5th Generation (B5G) dense deployment increased the energy demand and it is expected that the telecommunication industry will consume around 30% of the global energy by this year, and in the 6th Generation (6G) era the energy consumption will be even higher [4]. Therefore, enhancing energy efficiency in O-RAN is undoubtedly a crucial aspect for sustainable wireless communication expected in the future IoT society. Building on this foundation, and considering that in wireless networks, due to a distributed radio access network, 80% of the energy is consumed by the Base Station (BS) [5], in this paper, we present an O-RAN-based framework to optimize the energy consumption of the radio access network. The major contributions of this work are as follows:

- We propose an energy-saving technique for O-RAN that considers realistic network parameters obtained from the Liverpool 5G High Density Deployment (HDD) project that is deploying an innovative private 5G network in a number of venues in and around the city of Liverpool [6]. The aim of these deployments is to support large numbers of connections and assess how the network performs under these circumstances.
- We model the energy optimization problem and analyse the energy efficiency gains with a particular focus on scalability. Specifically, we demonstrate that as the number of connections increases, the proposed technique continues to achieve energy savings without significantly compromising network Quality of Service (QoS). This highlights the potential of our solution to support dense deployments in future wireless networks while maintaining sustainable performance.

The rest of the paper is organized as follows. Section II presents the overview of ORAN and literature review together with our novel contributions. Then, Section III introduce the proposed energy optimization model for O-RAN-based network. The performance evaluation is described in Section IV and finally, Section V concludes this paper along with future work.

II. BACKGROUND AND LITERATURE REVIEW

O-RAN is a new way to deploy a wireless network, where the operators can collaborate to share the resources. O-RAN, proposed by O-RAN Alliance has introduced key features such as RAN Intelligence Controller (RIC), an open radio interface to provide interoperability, and the addition of Machine Learning (ML) & Artificial Intelligence (AI) to benefit operators [7], [8]. The O-RAN architecture is shown in Fig. 1. The Service Management and Orchestration (SMO) in the figure is in charge of network orchestration, automation, and optimization, supporting the Non-Real-Time RIC (Non-RT RIC). The Non-RT RIC operates with latency above 1 second, optimizing the network through AI/ML-based rApps and interacts with the Near-Real-Time RIC (Near-RT RIC) via the A1 interface for policy control and AI/ML model updates. The Near-RT RIC, with latency between 10 milliseconds and 1 second, provides real-time optimization through xApps and controls the Centralized Unit (O-CU) and Distributed Unit (O-DU) via the E2 interface. The O-CU handles higher-layer protocols and is divided into O-CU-CP (Control Plane) for signalling and O-CU-UP (User Plane) for user traffic. The O-DU processes lower-layer protocols and is linked to the O-CU via the F1 interface, with F1-C handling control plane signalling and F1-U managing user data, and to the Radio Unit (O-RU) via the Open Fronthaul (OFH) interface to support different functional split options. The O-RU is the physical radio hardware responsible for transmitting and receiving Radio Frequency (RF). The O1 interface connects the SMO to all O-RAN elements for Fault, Configuration, Accounting, Performance, and Security (FCAPS) management. Finally, the O2 interface enables orchestration, management, and automation of cloud-based RAN (O-Cloud) components. This standardized architecture aims to reduce the network operators' CAPEX and OPEX by automating the network functions and management tasks. Moreover, the flexible and dynamic nature of RIC architecture makes the network more responsive, efficient, agile, and easily upgradable [7], [8].

To minimize the energy consumption of the network, the potential energy-saving techniques and ideas for next-phase energy optimization in O-RAN are explored by the O-RAN Alliance Sustainability Focus Group (SUFG) [9]. The features include optimizing the power consumption in O-RU through renewable energy integration, intelligent workload management in the O-Cloud environment, and spatial power-saving methods [9]. Moreover, the integration of 3GPP Release-18 features such as Synchronization Signal block (SSB) less Secondary cell (SCell) operation and Discontinuous Reception(DRX)/Discontinuous Transmission (DTX) mechanism will enhance the network energy efficiency without compromising the Quality of Service (QoS) [9]. The maximization of the energy efficiency of RAN through DRX/DTX has been analyzed in many studies. For instance, authors in [10] analyze the different DRX/DTX configurations for energy saving in O-RAN. The authors considered that radio modules may be separated into transmission & reception. The simulation results show that appropriate cell DRX timers maximize the cell throughput. Moreover, Lian et al, [8] designed two xApps for switching the Radio Cards (RCs) to enhance the energy efficiency of O-RAN. The proposed xApps are simulated using the 'TeraVM' RIC-tester to demonstrate its practical applicability.

Al-Karawi et al., [11] presented a Quantum based load balancing concept to optimize the energy efficiency in O-RAN. The authors solve the problem using a sequential quadratic problem and active set approach. The performance results show that the sequential quadratic problem achieves higher energy efficiency gains as compared to the active set approach. To reduce power consumption and energy efficiency maximization, Motalleb et al., in [12] proposed a joint power and network slicing-based approach. The authors map the physical data center resources to slices, and slices are mapped to services. Authors in [4], [13] published a review on O-RAN energy efficiency. Moreover, the O-RAN architecture, ML models, and estimation of O-RAN energy consumption using ML are also presented [4]. The AI/ML-based methods at RAN & edge infrastructure, and required workflow for cell ON/OFF mechanisms using AI are described in [14]. Recently, Samsung and MediaTek, in collaboration, verified the seamless integration of 5G Reduced Capability (RedCap) features over virtualized RAN (vRAN) and O-RAN [15]. The RedCap technology specified by 3GPP in release 17 to enhance the energy efficiency and battery life for IoT devices [16]. The test was conducted in Samsung's R&D lab in Korea using Samsung's versatile vRAN 3.0 software, O-RAN-compliant radio, and MediaTek's RedCap testing platform equipped with its M60 modem. The test focuses on two energy-saving features for IoT devices: Paging Early Indication (PEI) and extended Discontinuous Reception (eDRX) over vRAN and O-RAN network [15].

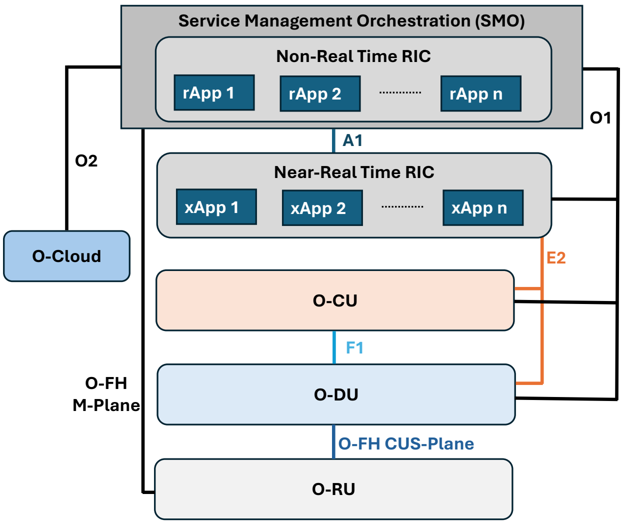

Fig. 1: O-RAN Architecture

The above works propose efficient energy-saving mechanisms for O-RAN and IoT devices. However, network energy optimization that considers realistic network parameters and device mobility is missing. This motivates us to optimize the energy efficiency O-RAN-based radio access network. We (i) mathematically model the energy optimization of O-RAN network; (ii) present algorithms to reduce energy consumption; and (iii) analyze the proposed algorithms using a Matlab-based simulator by considering realistic network parameters and device mobility.

III. SYSTEM MODEL AND POWER CONSUMPTION OPTIMIZATION

In this work, we considered a small 5G network that relies on the O-RAN architecture. The network consists of $\mathcal{K} = \{k_1, k_2, k_3...., k_n\}$ O-RU, which are densely deployed.

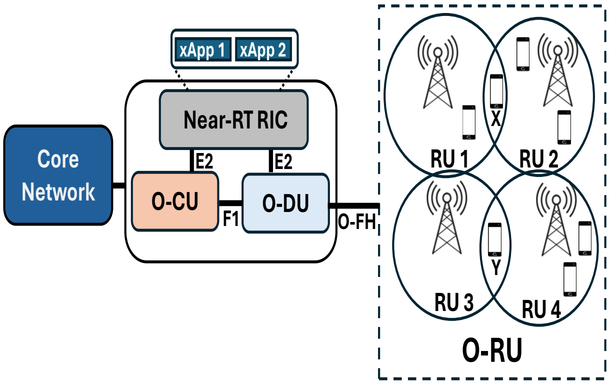

Fig. 2: O-RAN Radio Access Network

In a wireless network, a large amount of power is consumed by the RAN [17], [18] and the power consumed by the RUs consists of fixed and variable power consumption. Let P_{fix} represent the fixed power consumption and ηP_{tra} the variable power consumption, so the total power P_t consumed by a RU is given by:

$$P_t = P_{fix} + \eta P_{tra} \tag{1}$$

where, the variable η is the power amplifier efficiency and P_{tra} is the transmitting power of the RU. Let there are $\mathcal{N} = \{1, 2, 3...., n\}$ number of devices uniformly distributed in the network and associated with the RUs included in $\mathcal{K}$. The power received by a device is given by

$$P_{rec} = P_t - L^{dB}(d) \tag{2}$$

where, $L^{dB}(d)$ is the free space path loss and it is given by

$$L^{dB}(d) = 20log_{10}(f) + 20log_{10}(d) + 20Log_{10}(\frac{4\pi}{c}) \tag{3}$$

To save the power of the network, the effective approach is to switch the RUs between active and sleep mode based on network load. 3GPP introduced the Advanced Sleep Modes (ASMs) technique to reduce the power consumption of the BS. ASM allows the network to switch the BS into sleep mode periodically if the BS is idle [19]. Recently, the O-RAN alliance has introduced ASM for O-RAN [9]. However, switching the RU to sleep mode may affect the Quality of Service (QoS). The objective of this work is to minimize the network energy consumption while maintaining the required QoS. To achieve this, we considered two cases: (i) switch the RU in sleep mode when there is no device; (ii) shift the devices to another RU and switch off the unused capacity. Considering the above cases minimizing the network's power consumption is formulated as:

$$\textit{minimize} \quad P_t \tag{4}$$

$$\textit{subject to:} \quad \sum_{n=1}^{m} RB_n > RB_{req} \tag{5}$$

$$\sum RB \leq RB_t \tag{6}$$

$$\sum_{n=1}^{m} \Gamma_n > \Gamma_{thr} \tag{7}$$

Algorithm 1: Switch the RUs sleep mode to optimize EE

```
Input: number of RUs $l$, device cell association
       matrix $m$, simulation run $s$
for $i = 1$ to $s$ do
    for $j = 1$ to $l$ do
        find the number of devices ($n$) associated
          with $j$ from $m$
        if $n = 0$ then
            Switch the RU to sleep mode
        end
    end
end
```

where, m is the number of devices to be shifted, RB is the number of resource blocks, RB_t is the number of total resource blocks, Γ is the device throughput, Γ_{thr} is the

Algorithm 2: Device RU switching and switching RUs to sleep mode

```
Input: number of RUs $l$, device cell association
       matrix $m$, simulation run $s$, resource blocks
       assigned $r$
for $i = 1$ to $s$ do
    for $j = 1$ to $l$ do
        find the number of devices associated with $j$
          from $m$
        if no device associated with $j$ then
            $j = j + 1$
        else
            find the resource blocks assigned to
              devices from r
            for $k = 1$ to $l$ do
                if $j == k$ or no device in the target
                  RU then
                    $k = k + 1$
                else
                    find the available resource blocks
                      in $k$
                    if required resource blocks are
                      available then
                        shift the devices from RU $j$ to
                          RU $k$
                    else
                        $k = k + 1$
                    end
                end
            end
        end
    end
    if RU without any device then
        Switch the RU to sleep mode
    end
end
```

required throughput by devices. To solve the above optimization problem, we proposed two algorithms implemented on top of a near-RT RIC, which manages the cells using xApps as illustrated in the Fig. 2. Algorithm 1 switches the RU to sleep mode in case of no device. During every simulation run the RIC retrieves the number of devices associated with the RUs through interface E2 and if there is no device in the RU coverage, the RIC triggers the procedure that switches the RU to sleep mode. Therefore, Algorithm 1 only checks the device presence in the RUs coverage. However, there may be a case that a RU has very few devices and it is possible to offload them to another RU. As depicted in Fig. 2, a device “Y” can be in the coverage area of RU 3 and RU 4. That device connected to RU 3 may be shifted to RU 4 and then RU 3 can be switched to sleep mode. On the other hand, if a device “X” is in the coverage area of RU 1 and RU 2, but there is another device in RU 1 and, therefore, it is not possible to turn off any RU. To implement this, we proposed Algorithm 2. In Algorithm 2 the RIC checks if there are devices in a RU coverage (line 4) and if the gNB has the number of resource blocks required by the devices based on their QoS requirements through interface E2 (line 12-line 16). If the number of resource blocks satisfies the QoS requirements in terms of throughput based on the current applications experienced by the devices, they may be offloaded to the other RUs and the previous RU may be switched to sleep mode. Let $\mathcal{L}$ be the number of RUs switched to sleep mode, then the Energy Efficiency (EE) is defined as the ratio of power saved when $\mathcal{L}$ RUs are in sleep mode to the total power consumption of all RUs, and it is given as

$$EE = \frac{\sum_{n=1}^{\mathcal{K}} P_{tn} - \sum_{i=1}^{\mathcal{K}-\mathcal{L}} P_{ti}}{\sum_{n=1}^{\mathcal{K}} P_{tn}} \tag{8}$$

By using Algorithm 1, Algorithm 2, and equation (8), we can obtain the energy efficiency of the network.

IV. PERFORMANCE ANALYSIS

To analyze the performance of the proposed algorithms, we utilized Matlab-based simulator that we implemented in the Liverpool 5G project, which is based on the Vienna 5G System Level (SL) Simulator [20]. The Vienna 5G SL Simulator has the capability to mimic buildings, streets, radio environment conditions, such as path loss and shadowing, developed according to real-world scenarios based on data available in the OpenStreetMap (OSM) database. The Liverpool 5G Simulator extends and enhances this by implementing the deployments identified for the Liverpool 5G project and the gNB based on the project specifications. Moreover, in the Liverpool 5G Simulator, we designed and developed a module that implements a Near-RT RIC able to gather information from the RU through an E2 interface as illustrated in Fig. 2. To evaluate the algorithms proposed in this project, we considered the three gNBs each with three RUs placed according to our real-time deployment in Liverpool’s Salt and Tar area outdoor event venue [6]. The gNB deployment and devices are shown in Fig. 3. The deployment considers 3.8 GHz operating frequency, 100 MHz bandwidth, 4×4 MIMO, 49 dBm equivalent isotropic

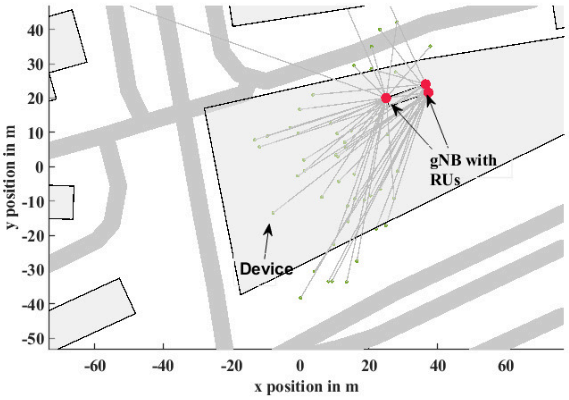

Fig. 3: Network Deployment

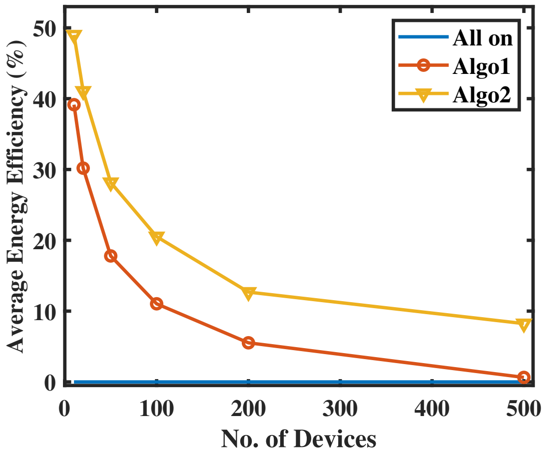

Fig. 4: Energy Efficiency with varying numbers of devices

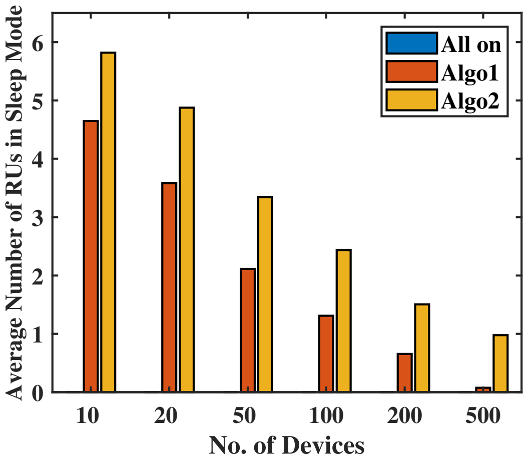

Fig. 5: No. of RUs switched off with varying numbers of devices

radiated power, 400 KHz subcarrier spacing, 12 dBi antenna directivity, and 256 QAM MCS-Table. Each RU consumes fixed power $P_{fix} = 20$ W during the active mode and $P_{fix} = 5$ W during the off mode. Moreover, we considered the power amplifier efficiency $\eta = 1.67\%$, and transmitting power $P_{tra} = 40$ W during the active mode [8], [17]. During the experiments, we considered varying the number of devices, both mobile and stationary, that were held in Liverpool's Salt and Tar venue. The simulator is inputted with the device's location data that can be updated every 1 second. We implemented Algorithm 1 and Algorithm 2 to analyze the energy efficiency of the network, the number of RUs that could be turned off, and the throughput achieved by the device compared to the case when all the RUs are always on. The simulation time slot duration is 1 ms, each simulation is executed 500 times, and we present the averaged results in the following figures.

Specifically, Fig. 4 shows the energy efficiency with increasing numbers of devices. From the figure, we can observe that the energy efficiency decreases with the increasing density of devices, as more devices do not allow the RU to be in sleep mode. The energy efficiency achieved with Algorithm 1 is higher compared to the "all on" case, as in that case all RUs always remain ON even when there is no device. Whereas, Algorithm 1 switches the RU into sleep mode if there is no device. The energy efficiency achieved with Algorithm 2 is 34.69% higher than Algorithm 1 because Algorithm 2 allows the RIC to shift devices to another nearby RU if there are enough resource blocks to accommodate the devices to switch into sleep mode. The number of RUs that can be switched to sleep mode during each simulation run for varying numbers of devices is depicted in Fig. 5. From the figure, we can observe that the RU in sleep mode decreases with the increase in the number of devices. This number is smaller in the case of Algorithm 1 compared to Algorithm 2, as it does not consider shifting devices. Moreover, compared to "all on" case Algorithm 1 allows the networks to switch the RU to sleep mode only in case there is no device for the current simulation run. The number of RU switched to sleep mode achieved by Algorithm 2 is 35% higher than Algorithm 1 at the cost of compromising the device throughput illustrated below.

The network throughput performance for Algorithm 1 and Algorithm 2 as a function of the varying numbers of devices is shown in Fig. 6. From the figure, we can observe that the average throughput per device decreases with the increase of the number of devices, as all devices share the available resource blocks. With a high number of active devices, the number resource blocks received by them is lower, which results in lower throughput. Moreover, the throughput achieved with Algorithm 1 is almost the same as compared to "all on" case, as it only switches RU to sleep mode where there is no device present. For a small number of devices (10 and 20), the throughput achieved by Algorithm 2 is 1.5 Mbps lower compared to Algorithm 1, while the throughput achieved with a higher number of devices is almost similar. This indicates that Algorithm 2 scales efficiently, as it maintains energy efficiency even with a high number of connections, without significantly compromising the throughput. Moreover, the lower values of throughput are due to the fact that Algorithm 2 shifts the device to another RU based on resource block availability, and it does not consider the received signal power or Signal-to-Noise-plus-Interference Ratio (SNIR). Therefore, when a device is associated with another cell, there is a possibility that the power received in the new cell is lower than the previous cell, which impacts the throughput. On the other hand, in this paper, we considered the available resource blocks as a shifting criterion because this information is available in the RIC.

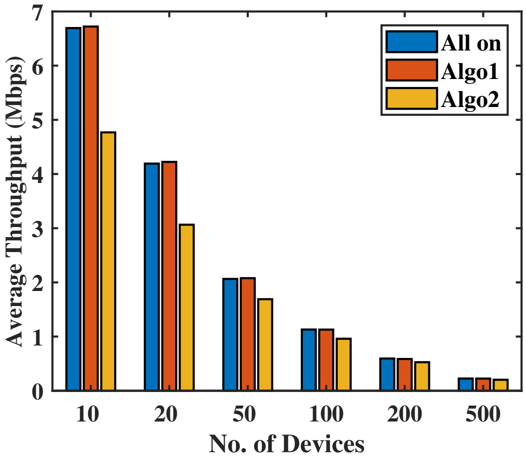

Fig. 6: Average throughput with varying numbers of devices

V. CONCLUSION

In this work, we presented a model to optimize the energy efficiency in an O-RAN-based 5G network. Specifically, we proposed two algorithms to minimize the power consumption of the O-RAN. Algorithm 1 switches RU to sleep mode when there is no device in the coverage. Whereas, Algorithm 2 shifts the devices to another RU based on the available resource blocks and switches the previous RU into sleep mode. The proposed algorithms have been analyzed using the Matlab-based Liverpool 5G Simulator, considering realistic parameters and device locations gathered from the Liverpool 5G HDD project deployments. The simulation results show that RUs that can be turned off with Algorithm 2 are 35% higher as compared to Algorithm 1 at the expense of a reduction of the QoS requirements in terms of throughput. In the future, we will use ML-based algorithms to optimize O-RAN-based network energy efficiency. We will analyze the network performance by considering a more dense deployed network and different scenarios.

ACKNOWLEDGMENT

This work is part of the Liverpool City Region HDD (High Demand Density) project, funded by the Department for Science, Innovation and Technology (DSIT), UK Government.

REFERENCES

[1] "Iot connections outlook – ericsson mobility report," https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/iot-connections-outlook#:~:text=Total%20cellular%20IoT%20connections%20are,20%20percent%20up%20to%202029 [Accessed: 22-04-2025].

[2] L. Wang, J. Zhou, M. Ma, and X. Niu, "Minimizing energy consumption of iot devices for o-ran based iot systems," *Energy Reports*, vol. 9, pp. 379–388, 2023.

[3] R. Firouzi and R. Rahmani, "5g-enabled distributed intelligence based on o-ran for distributed iot systems," *Sensors*, vol. 23, no. 1, p. 133, 2022.

[4] X. Liang, Q. Wang, A. Al-Tahmeesschi, S. Chetty, D. Grace, and H. Ahmadi, "Energy consumption of machine learning enhanced open ran: A comprehensive review," *IEEE Access*, 2024.

[5] P. Lähdekorpi, M. Hronec, P. Jolma, and J. Moilanen, "Energy efficiency of 5g mobile networks with base station sleep modes," in *IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2017, pp. 163–168.

[6] "Success for LCR HDD project first trial - Liverpool5G — liverpool5g.org.uk," https://liverpool5g.org.uk/success-for-lcr-hdd-project-first-trial/, [Accessed 10-02-2025].

[7] M. O-RAN, "O-ran working group 1 (use cases and overall architecture) o-ran architecture description (o-ran. wg1. oad-r003-v12.00): Tech. rep." June 2024.

[8] X. Liang, A. Al-Tahmeesschi, Q. Wang, S. Chetty, C. Sun, and H. Ahmadi, "Enhancing energy efficiency in o-ran through intelligent xapps deployment," in *2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2024, pp. 1–6.

[9] S. F. G. (SUFG), "Potential energy savings features in o-ran," O-RAN ALLIANCE, Buschkauler Weg 27, 53347 Alfter, Germany, Tech. Rep. v01.00, Jan 20, 2025. [Online]. Available: https://www.o-ran.org/o-ran-resources

[10] D. Roh and J.-H. Na, "Performance analysis of different cell drx/dtx configurations for energy saving in o-ran system," in *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2024, pp. 1128–1129.

[11] Y. Al-Karawi, H. Al-Raweshidy, and R. Nilavalan, "Optimizing the energy efficiency using quantum based load balancing in open radio access networks," *IEEE Access*, 2024.

[12] M. K. Motalleb, V. Shah-Mansouri, and S. N. Naghadeh, "Joint power allocation and network slicing in an open ran system," *arXiv preprint arXiv:1911.01904*, 2019.

[13] A. I. Abubakar, O. Onireti, Y. Sambo, L. Zhang, G. K. Ragesh, and M. Ali Imran, "Energy efficiency of open radio access network: A survey," in *IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–7.

[14] M. Ghoraishi, J. O. Sallent, M. Catalan-Cid, G. Bielsa, J.-F. Esteban-Rivas, V. Sark, J. G. Teran, and S. Pryor, "Begreen: beyond 5g energy efficient networking by hardware acceleration and ai-driven management of network functions," in *Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2023, pp. 717–722.

[15] "The potential of 5g redcap and energy efficiency," https://www.samsung.com/global/business/networks/insights/blog/0619-the-potential-of-5g-redcap-and-energy-efficiency/ [Accessed: 29-04-2025].

[16] S. N. K. Veedu, M. Mozaffari, A. Höglund, E. A. Yavuz, T. Tirronen, J. Bergman, and Y.-P. E. Wang, "Toward smaller and lower-cost 5g devices with longer battery life: An overview of 3gpp release 17 redcap," *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 84–90, 2022.

[17] X. Yu, G. Li, and W. Lu, "Power consumption based on 5g communication," in *IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, vol. 5, 2021, pp. 910–914.

[18] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user mimo systems: Is massive mimo the answer?" *IEEE Transactions on wireless communications*, vol. 14, no. 6, pp. 3059–3075, 2015.

[19] F. E. Salem, A. Gati, Z. Altman, and T. Chahed, "Advanced sleep modes and their impact on flow-level performance of 5g networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–7.

[20] M. K. Müller, F. Ademaj, T. Dittrich, A. Fastenbauer, B. R. Elbal, A. Nabavi, L. Nagel, S. Schwarz, and M. Rupp, "Flexible multi-node simulation of cellular mobile communications: the Vienna 5G System Level Simulator," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 17, Sep. 2018.