Comparative Study of Radio Resource Distribution Algorithms

Atef Gharbi
Department of Information Systems, Faculty of Computing and Information Technology, Northern Border University, Saudi Arabia | LISI Laboratory, National Institute of Applied Sciences and Technology (INSAT), University of Carthage, Tunisia
atef.gharbi@nbu.edu.sa (corresponding author)

Abdulsamad Ebrahim Yahya
Faculty of Computing and Information Technology, Northern Border University, Saudi Arabia
abdulsamad.qasem@nbu.edu.sa

Mohamed Ayari
Faculty of Computing and Information Technology, Northern Border University, Saudi Arabia | SYSCOM Laboratory, National Engineering School of Tunis, University of Tunis El-Manar, Tunisia
mohamed.ayari@nbu.edu.sa

Received: 23 December 2023 | Revised: 2 January 2024 | Accepted: 10 January 2024

ABSTRACT

The equitable distribution of radio resources among different users in wireless networks is a difficult problem and has attracted the interest of many studies. This study presents the Proportional Fair Q-Learning Algorithm (PFLA) to enable the equitable distribution of radio resources among diverse users through the integration of Q-learning and proportional fairness principles. The PFLA, Round Robin (RR), and Max Throughput (MaxTP) algorithms were compared to evaluate their effectiveness in resource allocation. Performance was measured in terms of sum-rate throughputs and fairness index. The comparison results showed an improvement in the fairness index metrics for PFLA compared to the other algorithms. PFLA showed gains of 11.62 and 43% in the fairness index compared to RR and MaxTP, respectively. These results show that PFLA is more efficient in utilizing available resources, leading to higher overall system throughput and demonstrating its ability to balance performance metrics between users, especially when the number of users increases.

Keywords-radio resource distribution; proportional fair Q-learning algorithm; round robin; max throughput

I. INTRODUCTION

The allocation of radio resources aims to optimize spectrum efficiency while complying with the predefined fairness criteria. Effortless allocation of radio resources is particularly important in wireless communication networks because the available bandwidth is limited and shared between multiple users [1-2]. To make the best use of these resources, it is necessary to use advanced algorithms to allocate them efficiently and fairly among users. Effective resource allocation algorithms can have a significant impact on the performance of wireless networks. By balancing the competing demands of different users, these algorithms can improve network throughput, reduce latency, and enhance overall user experience [3-4]. Moreover, these algorithms can help network operators reduce costs and improve their return on investment by making more efficient use of their available spectrum.

This paper presents the Proportional Fair Q-Learning Algorithm (PFLA) to optimize radio resource allocation in wireless networks by integrating Q-learning and proportional fairness principles. PFLA dynamically adjusts resource allocation based on network feedback using Q-learning, striving for long-term rewards like throughput and fairness. The proposed algorithm fairly balances the overall system throughput among users by considering factors such as channel conditions, queue sizes, and user priorities. This integrated approach enhances system performance while ensuring equitable treatment for all users, surpassing channel-unaware algorithms in wireless communication scenarios. PFLA can have a significant pragmatic impact on daily wireless communications, as it offers a promise of significantly increased network efficiency and can ensure a more equitable allocation of resources between users. Its implementation results in an improved user experience, reduced latency, and smoother connectivity in daily activities, such as faster downloads and seamless streaming. PFLA was compared to the RR and MaxTP algorithms based on sum-rate throughputs and fairness index. The results showed that RR and MaxTP have limitations with larger user numbers, unlike PFLA, which improves allocation efficiency by considering channel quality and other factors. Incorporating channel conditions and queue sizes, PFLA surpasses RR and MaxTP in wireless network performance and efficiency. This channel- and size-aware strategy is crucial for enhancing wireless data network performance.

II. RELATED WORK

A user pairing technique can be implemented to match users based on their channel conditions and data rate needs and improve fairness among pairs of users by ensuring that those with similar requirements receive equal resource allocation [5-6]. In [7], it was shown that LTE-licensed assisted access achieves proportional fairness with WiFi by adjusting the initial backoff window size or sensing duration and deriving and validating optimal parameter values through simulation experiments. In [8], a deep reinforcement learning-based method was proposed for D2D communication in spectrum sharing, optimizing spectral efficiency and ensuring fairness between network links through resource block scheduling and power control integration. In [9], a decentralized framework was introduced for non-ideal NOMA networks, formulating a clustering algorithm considering cluster sizes and channel gain variations between User Equipment (UE), achieving a balanced α-fair RA framework, and optimizing throughput and fairness through iterative bandwidth and power refinement. In [10], spectrum and power allocation complexities in NOMAs were addressed by formulating principles focusing on Double-Objective Optimization (DOO) to balance the total and minimum rates. Using a power dispersion method, nonconvex DOO problems were converted into Single-Objective Optimization (SOO) ones through global optimal search, prioritizing equal channel gain for Adaptive Proportional Fair (APF) user pairing and optimization restructuring. In [11], a dynamic power allocation scheme was proposed by employing model-free deep reinforcement learning where transmitters adjust transmission power based on collected channel state and quality of service data to maximize a utility function for weighted sum rates, allowing for adaptability for optimal rates or fair scheduling amidst random CSI changes using deep Q-learning. In [12], Federated Learning (FL) was investigated in wireless networks to improve FL model training under limitations such as incomplete CSI and limited local computing resources. This study quantified the training loss gap between the FL client calendar and centralized training, addressing loss reduction through Lyapunov optimization as stochastic optimization. This method integrated a Gauss process regression-based channel prediction method and incorporated client CSI and computing power into planning decisions to mitigate these challenges. In [13], the maximum weight-min fairness was pursued, optimizing LTE-U node retention probability to maximize throughput in both WiFi and LTE-U networks, addressing bandwidth balance with weight factors and solving the optimization challenge through bisection and Laplace transformation inversion methodologies. In [14], non-licensed frequency bands were organized into time slots for LTE frame transmission, and Small Base Stations (SBSs) reserved these slots at designated Access Points (APs) to prevent overlaps and interference between mobile and WiFi networks, ensuring independent resources via the maintenance cycle method by allocating specific slots for mobile systems and other WiFi operations.

The existing advances in wireless network optimization, while promising, confront challenges in adaptability, scalability, and holistic consideration of dynamic network conditions. These methods, ranging from user pairing techniques to decentralized frameworks and deep reinforcement learning, often face limitations in scaling to larger networks, dynamically adapting to diverse network conditions, and ensuring fairness across varying user requirements. A comprehensive solution that integrates adaptive machine learning techniques with an algorithmic framework can offer a more flexible and adaptive approach to address these shortcomings. Combining reinforcement learning with predictive analytics and considering a broader spectrum of network parameters could dynamically optimize resource allocation, ensuring fairness while adapting to real-time changes and catering to a larger and more diverse user base in wireless networks.

III. CLASSIC SCHEDULERS

A. Round Robin (RR) Scheduling

Fair resource allocation scheduling aims to achieve equitable distribution of resources among multiple users in a wireless communication system. A widely used logical strategy to implement resource fairness involves the recursive application of the RR scheduling scheme. This iterative approach is designed to ensure an equitable distribution of resources to all users, with the overarching objective of maximizing the average resource allocation rate for all users while adhering to specific fairness criteria [15-18]. In this context, the term "fair resources" denotes the principle of equitably distributing available resources, such as time slots or frequency bands, among all active users. This strategy aims to prevent any single user from monopolizing resources and ensures that each user receives an equal share. To implement this approach, the system recursively applies the RR scheduling scheme, which is a well-established method of providing equal access to resources. The primary goal of this strategy is to improve the overall average resource allocation rate for all users. It does so while maintaining fairness, which is often characterized by specific fairness criteria that dictate how resources should be allocated to different users. By striking a balance between resource fairness and maximizing resource allocation rates, this approach contributes to efficient and equitable resource utilization in wireless communication systems.

B. Max Throughput (MaxTP) Scheduling

The MaxTP scheduler aims to maximize the overall throughput of the base station by assigning a user to each channel, allowing the maximum data speed in the current transmission time interval. The Max TP scheduler can operate in both time and frequency fields. In the frequency field, UE is assigned to the highest Channel Quality Indicator (CQI) [19].

In the MaxTP scheduling algorithm, channels are allocated to UEs based on their CQI values. UEs with better channel conditions (higher CQI) are given priority in resource allocation. This prioritization aims to maximize data rates for the UEs, resulting in higher overall throughput for the base station. The MaxTP scheduler can adapt to changing channel conditions and dynamically allocate RBs to UEs, ensuring efficient resource utilization and maximizing throughput. The specific implementation and fine-tuning of the algorithm may vary depending on the network requirements and system parameters.

C. Proportional Fair Scheduling (PFS)

PFS strategies have been developed to achieve a balanced trade-off between various performance metrics, primarily focusing on UE throughput and fairness. These schemes rely on evaluating the indicators of both the current and historical channel quality of the UEs. Specifically, they consider the instantaneous and average data rates experienced by UEs over time to determine a priority function. Within PFS, the scheduler, at each time slot, denoted as t, grants the highest priority to the UE that exhibits the maximum priority function. This priority function, denoted as P_k for the k-th UE, is mathematically expressed as:

$$P_k = \text{argmax}\frac{dr_k(t)}{HR_k(t)} \quad (1)$$

where $dr_k(t)$ represents the instantaneous achievable data rate for the k-th UE when connected to the channel at time slot t. Additionally, $HR_k(t)$ signifies the historical average data rate that has been assigned to the k-th UE up to and including the time slot t. In the PFS scheduling scheme, the system executes a two-step process using feedback from channel quality indicators. First, it calculates the priority functions for all UEs and sorts them in descending order. Subsequently, channels are assigned to the UE with the highest priority function. This process ensures a fair distribution of resources among UEs while considering their channel qualities, which in turn aids in optimizing overall network performance. In PFS scheduling algorithm, the priority function P_k is calculated based on the ratio of the current achievable data rate to the past average data rate for each UE. This prioritization ensures that UEs with varying channel conditions and data rate histories are allocated RBs fairly, balancing throughput and fairness.

IV. THE PROPOSED MODEL

A. Definitions and Terminologies

This study considers WiFi networks, where different algorithms can be employed in managing access and resource allocation among devices to ensure fair transmission opportunities. This study considered a wireless network configuration consisting of two primary sets: $N = \{1, 2, \ldots, N\}$, representing a group of agents, and $K = \{1, 2, \ldots, K\}$, representing a collection of available channels. Agents use a random access protocol to send their data over these shared channels. At each discrete time slot, each agent could choose one channel for transmission with an associated transmission probability such as an Aloha-type narrowband transmission scheme. It is assumed that all agents always have pending packets for transmission. After each time slot t, when an agent i attempts to send a packet, it receives binary feedback in the form of an ACK signal denoted as $o_i(t)$. This signal indicates whether their transmission was successful or not. If $o_i(t) = 1$, it means that transmission was successful. On the other hand, $o_i(t) = 0$ signifies a failed transmission or collision.

Let $a_i(\text{t}) \in \{0, 1, \ldots, K\}$ denote the action taken by agent i at the time slot t. Here, $a_i(t) = 0$ corresponds to the scenarios where agent i chooses not to transmit a packet during the time slot t, potentially to mitigate network congestion. On the contrary, $a_i(t) = k$, where $1 \le k \le K$, denotes the i-th agent's decision to transmit a packet on channel k in time slot t. During real-time operation, each agent, denoted as i, makes autonomous decisions independently and in a distributed manner. These decisions aim to acquire efficient spectrum access policies exclusively through ACK signals. Traditional methods for solving this problem become intractable from a mathematical point of view as the network size grows, given its combinatorial nature and the challenge of dealing with partial-state observations. The Q-learning approach was embraced due to its ability to provide satisfactory approximate solutions, even when confronted with large state and action spaces.

B. Q-Learning Approach

The action taken by an agent i at time step t is represented by $a_i(t)$, where $a_i(t) \in \{Transmit, Wait\}$. The *Transmit* action indicates that the agent i initiates transmission, while the *Wait* action indicates that the agent refrains from transmitting. The observation of the channel state at time step t is described as $O_i(t)$ and can have one of the three possible values: $O_i(t) = \{Success, Collision, Idleness\}$. The *Success* observation indicates that a single station has successfully transmitted on the channel, *Collision* indicates that multiple stations have transmitted simultaneously resulting in a collision, and *Idleness* indicates that no station is currently transmitting. The agent analyzes the acknowledgment received from the Access Point (AP) when it transmits and monitors the channel while waiting to determine the value of $O_i(t)$. The channel state at time step t is defined as an action-observation pair called $c_t = (a_i(t), O_i(t))$. The channel state c_t can have five possible combinations: {*Transmit*, *Success*}, {*Transmit*, *Collision*}, {*Wait*, *Success*}, {*Wait*, *Collision*}, and {*Wait*, *Idleness*}. The environmental state s_t contains the history of the action-observation pairs for a specified length of M, which determines the scope of historical data considered by the agent. When the action $a_i(t)$ is executed, the transition from state s_t to state s_{t+1} generates the reward r_t+1, where:

$$r_{t_{+1}} = \begin{cases} 1, \text{ if } z_t = \textit{Success} \\ 0, \text{ if } \text{z}_\text{t} = \textit{Collision} \text{ or } \textit{Idleness} \end{cases} \quad (2)$$

Let $r_i(t)$ denote the reward received by agent i at the beginning of the time slot t. This reward depends on the actions of agent i at the previous time slot, $a_i(t-1)$, and the actions of other agents at the previous time slot, $a_{-i}(t-1)$, which collectively form the elusive network state that agent i is trying to identify. The cumulative discounted reward, denoted by R_i, is calculated as follows:

$$R_i = \sum_{t=1}^{T} \gamma^{t-1} r_i\,(t) \quad (3)$$

where $0 \leq \gamma \leq 1$ serves as a discounting factor, and T represents the time horizon. Typically, γ is set to 1, but it may assume a value less than 1 when T is bounded or unbounded. Following an occurrence of a state-action pair (s_t, a_t) and the subsequent reward r_{t+1}, Q-learning updates the Q-function, represented as (s_t, a_t), through the application of the following update rule:

$$q(s_t, a_t) \leftarrow q(s_t, a_t) + a * (r_{t+1} + \gamma * max_{a'} q(s_{t+1}, a') - q(s_t, a_t) \quad (4)$$

where r_{t+1} is the immediate reward obtained after taking action α_t in state s_t, $max_{a'} q(s_{t+1}, a')$ is the maximum Q value over all possible actions in the next state s_{t+1} and represents the estimated future cumulative reward, $\alpha \in (0, 1]$ signifies the learning rate that governs the magnitude of the update, and γ is a constant between 0 and 1 called the discount factor that determines the importance of future awards. The Q-function is updated iteratively to reflect the acquired reward and the expected future rewards based on the highest Q-value associated with potential future actions. During the period when the Q-function (s, a) is being updated, the decision-making process operates based on this evolving Q-function. A common policy strategy in this context is the ε-greedy policy. Under the ε-greedy policy, the agent selects the action that maximizes the Q-value, denoted as:

$$a_t = argmax_a(s_t, a') \quad (5)$$

In contrast, with a probability of ξ, the agent selects a random action. The rationale behind incorporating randomness through ξ is to strike a balance between exploration and exploitation, enabling the agent to explore new actions while also favoring actions with known high Q-values.

C. Proportional Fair Q-Learning for Dynamic Resource Allocation

Combining Q-learning with the proportional fair algorithm for dynamic resource allocation between users and channels in wireless communication systems can enhance system performance and fairness. The following assumptions were made:

- Multiple users (indexed by i) and multiple channels (indexed by k).
- Each user has a Q-table to learn the Q-values for each state-action pair (state: user's buffer size and channel conditions, action: channel allocation).

Initialize Q-tables for each user, with dimensions (s, a), where s represents the state and a represents the action. Initialize all Q-values to zero. The algorithm shown in Figure 1 uses Q-learning to learn the optimal channel allocation strategy based on observed states and rewards. The fairness factor λ is introduced to ensure proportional fairness in the allocation of resources. Users strike a balance between exploration and exploitation through the ε-greedy policy.

V. PERFORMANCE EVALUATION

Numerical experiments were carried out in Matlab to determine the performance attributes of PFLA. The simulated wireless network was designed to encompass a variable number of users, denoted as N, ranging between 5 and 50, and K channels, varying between 2 and 23. A consistent channel bandwidth, denoted as B = 25MHz, was maintained to calculate data rates.

```
Input: α, γ, ε, λ, M (Maximum number of
iterations).
Output: Learning values, Qm*n

Begin
Initialize Qm*n(st, at) to 0 for all state-action
pairs (st, at).
Generate a random number ξ in [0, 1]
Set st to a random state from the state set:
S={Success, Collision, Idleness}.
Set actt to a random action from the action set:
S={Transmit, Wait}.
Generate a random number x in [0, 1].
for iteration m = 1, …, M do
  for time-slot t = 1, …, T do
    for agent i = 1, …,N do
      if agent i has still data to send do
        Determine the state of the robot si(t).
        If x < ξ then
          Choose ai(t) based on (4)
        else
          Choose the best ai(t)∈ {0, …,K} using Qm*n.
        End if
        Perform action ai(t) and receive reward
          ri(t).
        Update Q-values using Q-learning:
        Calculate the TD error:
        δ = r + γ * max(s', α')- max(s, a)
        Update Q-value: (s, α) ← (s, a) + α * δ
        si(t) ← si(t+1)
      End if
    End for
  End for
End for
Return Qm*n
End
```

Fig. 1. PFL algorithm for dynamic resource allocation.

A. Sum-Rate Throughput

The average rate of each UE in the T schedule interval is defined by:

$$AR_i = \frac{1}{T} \sum_{t=1}^{T} dr_i\,(t) \quad (6)$$

The sum rate is established as a specific metric to measure the optimization. It quantifies the average cumulative throughput of the entire network during the period of the schedule interval T:

$$R_{sum} = \sum_{i=1}^{N} AR_i \quad (7)$$

Based on Figure 1, the results show the sum rate throughput achieved by different scheduling methods for different users. In PFLA, with increasing user numbers, total sum rates tend to increase, demonstrating effective resource allocation based on the Q-learning approach. RR scheduling shows a less consistent trend and does not match PFLA performance for more users. The MaxTP scheduler starts with a competitive sum rate throughput, but as the number of users increases, the performance plateaus. MaxTP does not adapt well to the growing number of users, indicating limitations in resource allocation efficiency. In general, PFLA is more effective in resource allocation in this scenario, with a higher number of users resulting in better average throughput than other methods. RR and MaxTP schedulers have limitations in handling large user numbers, which leads to less efficient resource utilization.

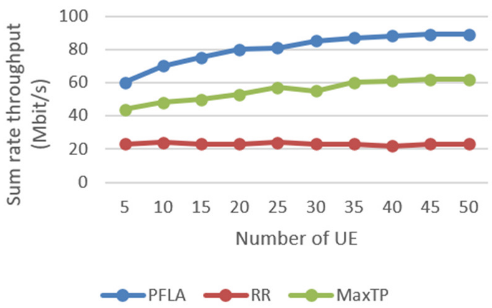

Fig. 2. Sum-rate throughputs for PFLA, RR, and MaxTP.

B. *Fairness Index*

Fairness can be evaluated using Jain's fairness index. In the following equation, n represents the total number of users and x_i is the total amount of throughput obtained by each user:

$$J(x_1, x_2, \dots, x_n) = \frac{\left(\sum_{k=1}^{N} T_k\right)^2}{N\left(\sum_{k=1}^{N} T^2{}_k\right)} \quad (8)$$

Figure 3 shows the fairness index achieved by different scheduling methods for different numbers of users.

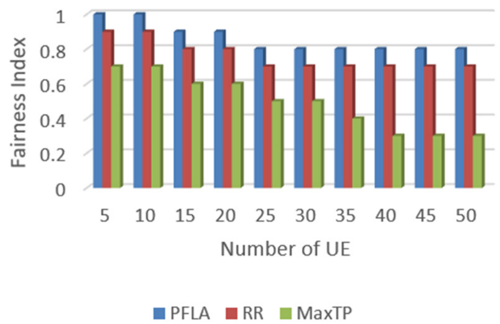

Fig. 3. Comparative analysis of fairness metrics.

Fairness is often measured by a fairness index, and in this case, higher values indicate better fairness. PFLA consistently shows a high level of fairness for all users. With the increase in the number of users, the fairness index is still relatively stable, showing that PFLA maintains equality between users while efficiently allocating resources. The RR scheduler shows relatively low fairness indexes, which decrease as the number of users increases. MaxTP scheduling starts with moderately competitive fairness indexes, but as user numbers increase, fairness decreases. Like RR scheduling, it is difficult to maintain fairness when more users are needed to serve. In summary, PFLA consistently achieved the highest fairness index, demonstrating its ability to balance performance metrics between users, especially if the system has more users. The RR and MaxTP schedulers provide a certain degree of fairness, but their flexibility decreases as users increase.

The following formulas were used to determine the relative advantages of PFLA over RR and MaxTP:

$$Gain_{PFLA-RR} = \frac{PFLA\ average\ fairness\ index - RR\ average\ fairness\ index}{PFLAaverage\ fairness\ index} \times 100$$

$$Gain_{PFLA-RR} = \frac{0.86-0.76}{0.86} \times 100 \approx 11.62\%$$

$$Gain_{PFLA-MaxTp} = \frac{PFLA\ average\ fairness\ index - MaxTP\ average\ fairness\ index}{PFLAaverage\ fairness\ index} \times 100$$

$$Gain_{PFLA-MaxTp} = \frac{(0.86-0.49)}{0.86} \times 100 \approx 43\%$$

VI. CONCLUSION

The choice of an algorithm for radio resource allocation depends on the specific requirements and constraints of the wireless communication network. PFLA aims to maximize system throughput by considering both channel conditions and queue sizes. By incorporating channel conditions and queue sizes into the resource allocation decision-making process, PFLA outperforms the RR and MaxTP algorithms in terms of performance and efficiency in wireless networks. The empirical findings showed an improvement in the fairness index metrics when comparing PFLA with the RR and MaxTP algorithms. The percentage gain between PFLA and RR was determined to be 11.62%, which represents a notable improvement in the fairness index achieved by PFLA over RR. Similarly, when PFLA was compared to MaxTP, a substantial gain of 43% was achieved in the fairness index, indicating the superior fairness achieved by the proposed algorithm in all the experimental scenarios considered. By extending PFLA to support the unique requirements of next-generation wireless technologies, the proposed allocation algorithm can remain relevant and effective in the ever-evolving landscape of wireless communication. Future work involves further development and adaptation of PFLA to align with the specific demands of emerging wireless technologies, such as 6G and IoT networks, ensuring its applicability and optimization for evolving communication paradigms.

ACKNOWLEDGMENT

The authors gratefully acknowledge the approval and the support of this research study by grant no. CSCR-2023-12-2077 from the Deanship of Scientific Research at Northern Border University, Arar, K.S.A.

REFERENCES

[1] G. I. Tsiropoulos, O. A. Dobre, M. H. Ahmed, and K. E. Baddour, "Radio Resource Allocation Techniques for Efficient Spectrum Access in Cognitive Radio Networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 824–847, 2016, https://doi.org/10.1109/COMST.2014.2362796.

[2] A. Ahmad, S. Ahmad, M. H. Rehmani, and N. U. Hassan, "A Survey on Radio Resource Allocation in Cognitive Radio Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 888–917, 2015, https://doi.org/10.1109/COMST.2015.2401597.

[3] H. Beshley, M. Beshley, M. Medvetskyi, and J. Pyrih, "QoS-Aware Optimal Radio Resource Allocation Method for Machine-Type

Communications in 5G LTE and beyond Cellular Networks," *Wireless Communications and Mobile Computing*, vol. 2021, May 2021, Art. no. e9966366, https://doi.org/10.1155/2021/9966366.

[4] V. F. Monteiro, D. A. Sousa, T. F. Maciel, F. R. M. Lima, E. B. Rodrigues, and F. R. P. Cavalcanti, "Radio resource allocation framework for quality of experience optimization in wireless networks," *IEEE Network*, vol. 29, no. 6, pp. 33–39, Aug. 2015, https://doi.org/10.1109/MNET.2015.7340422.

[5] Z. Wang, H. Hu, B. Jia, and T. Xu, "A windowing-based joint user pairing and resource allocation algorithm for V-MIMO systems," *Computer Communications*, vol. 144, pp. 1–7, Aug. 2019, https://doi.org/10.1016/j.comcom.2019.05.004.

[6] M. A. Ruder, A. Moldovan, and W. H. Gerstacker, "SC-FDMA user pairing and frequency allocation with imperfect channel state information," *Physical Communication*, vol. 13, pp. 88–98, Dec. 2014, https://doi.org/10.1016/j.phycom.2014.10.001.

[7] Y. Gao and S. Roy, "Achieving Proportional Fairness for LTE-LAA and Wi-Fi Coexistence in Unlicensed Spectrum," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3390–3404, Feb. 2020, https://doi.org/10.1109/TWC.2020.2973145.

[8] I. Budhiraja, N. Kumar, and S. Tyagi, "Deep-Reinforcement-Learning-Based Proportional Fair Scheduling Control Scheme for Underlay D2D Communication," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3143–3156, Mar. 2021, https://doi.org/10.1109/JIOT.2020.3014926.

[9] A. Celik, M.-C. Tsai, R. M. Radaydeh, F. S. Al-Qahtani, and M.-S. Alouini, "Distributed User Clustering and Resource Allocation for Imperfect NOMA in Heterogeneous Networks," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7211–7227, Jul. 2019, https://doi.org/10.1109/TCOMM.2019.2927561.

[10] K. Long, P. Wang, W. Li, and D. Chen, "Spectrum Resource and Power Allocation With Adaptive Proportional Fair User Pairing for NOMA Systems," *IEEE Access*, vol. 7, pp. 80043–80057, 2019, https://doi.org/10.1109/ACCESS.2019.2908673.

[11] Y. S. Nasir and D. Guo, "Multi-Agent Deep Reinforcement Learning for Dynamic Power Allocation in Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2239–2250, Jul. 2019, https://doi.org/10.1109/JSAC.2019.2933973.

[12] M. M. Wadu, S. Samarakoon, and M. Bennis, "Joint Client Scheduling and Resource Allocation Under Channel Uncertainty in Federated Learning," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5962–5974, Sep. 2021, https://doi.org/10.1109/TCOMM.2021.3088528.

[13] X. Wang, T. Q. S. Quek, M. Sheng, and J. Li, "Throughput and Fairness Analysis of Wi-Fi and LTE-U in Unlicensed Band," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 63–78, Jan. 2017, https://doi.org/10.1109/JSAC.2016.2632629.

[14] U. Challita, L. Dong, and W. Saad, "Proactive Resource Management for LTE in Unlicensed Spectrum: A Deep Learning Perspective," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4674–4689, Jul. 2018, https://doi.org/10.1109/TWC.2018.2829773.

[15] B. Damera, P. C. Babu, and J. S. Mohamed, "Optimized MCE scheduling algorithm to allocate radio resources using evolved Round Robin scheduling," in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Bangalore, India, Jul. 2016, pp. 770–775, https://doi.org/10.1109/ICATCCT.2016.7912103.

[16] S. Alotaibi, "A Fairness-based Cell Selection Mechanism for Ultra-Dense Networks (UDNs)," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11524–11532, Oct. 2023, https://doi.org/10.48084/etasr.6106.

[17] F. O. Ombongi, H. O. Absaloms, and P. L. Kibet, "Energy Efficient Resource Allocation in Millimeter-Wave D2D Enabled 5G Cellular Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6152–6160, Aug. 2020, https://doi.org/10.48084/etasr.3727.

[18] S. Panbude, B. Iyer, A. B. Nandgaonkar, and P. S. Deshpande, "DFPC: Dynamic Fuzzy-based Primary User Aware clustering for Cognitive Radio Wireless Sensor Networks," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12058–12067, Dec. 2023, https://doi.org/10.48084/etasr.6279.

[19] T. Hori and T. Ohtsuki, "QoE and throughput aware radio resource allocation algorithm in LTE network with users using different applications," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–6, https://doi.org/10.1109/PIMRC.2016.7794802.