IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 September 2023, accepted 3 November 2023, date of publication 9 November 2023, date of current version 15 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3331950

RESEARCH ARTICLE

Introducing Reinforcement Learning in the Wi-Fi MAC Layer to Support Sustainable Communications in e-Health Scenarios

GOLSHAN FAMITAFRESHI[1], M. SHAHWAIZ AFAQUI[2], (Member, IEEE), AND JOAN MELIÀ-SEGUÍ[1,3], (Senior Member, IEEE)

[1]Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya (UOC), 08018 Barcelona, Spain
[2]School of Computing, Engineering and Digital Technologies, Teesside University, TS1 3BX Middlesbrough, U.K.
[3]Faculty of Computer Science, Multimedia and Telecommunication, Universitat Oberta de Catalunya (UOC), 08018 Barcelona, Spain

Corresponding author: Golshan Famitafreshi (gfamitafreshi@uoc.edu)

This work was supported in part by the Spanish Ministry of Science, Innovation and Universities through the Project RF-VOLUTION under Grant PID2021-122247OB-I00; in part by the Spanish Ministry of Culture and Sports, and the European Funds for the Recovery, Transformation and Resilience Plan, through the HydraSport Project under Grant EXP_75087; in part by Generalitat de Catalunya through SGR Funds under Grant 2021-SGR-00174; and in part by the Support of a Universitat Oberta de Catalunya Doctoral Grant.

ABSTRACT The crisis of energy supplies has led to the need for sustainability in technology, especially in the Internet of Things (IoT) paradigm. One solution is the integration of passive technologies like Energy Harvesting (EH) into IoT systems, which reduces the amount of battery replacement. However, integrating EH technologies within IoT systems is challenging, and it requires adaptations at different layers of the IoT protocol stack, especially at the Medium Access Control (MAC) layer due to its energy-hungry features. Since Wi-Fi is a widely used wireless technology in IoT systems, in this paper, we perform an extensive set of simulations in a dense solar-based energy-harvesting Wi-Fi network in an e-Health environment. We introduce optimization algorithms, which benefit from the Reinforcement Learning (RL) methods to efficiently adjust to the complexity and dynamic behaviour of the network. We assume the concept of Access Point (AP) coordination to demonstrate the feasibility of the upcoming Wi-Fi amendment IEEE 802.11bn (Wi-Fi 8). This paper shows that the proposed algorithms reduce the network's energy consumption by up to 25% compared to legacy Wi-Fi while maintaining the required Quality of Service (QoS) for e-Health applications. Moreover, by considering the specific adjustment of MAC layer parameters, up to 37% of the energy of the network can be conserved, which illustrates the viability of reducing the dimensions of solar cells, while concurrently augmenting the flexibility of this EH technique for deployment within the IoT devices. We anticipate this research will shed light on new possibilities for IoT energy harvesting integration, particularly in contexts with restricted QoS environments such as passive sensing and e-Healthcare.

INDEX TERMS Medical Internet of Things, access point coordination, sleep/wake-up, machine learning, reinforcement learning, energy harvesting technologies, passive communications.

I. INTRODUCTION

According to Cisco, approximately 30 billion connected Internet of Things (IoT) devices will exist by the end of 2023 [1], this fast development and deployment of IoT ecosystems, from smart cities to smart agriculture, have a negative impact on the environment and planetary resources [2]. One crucial factor in mitigating these harmful effects is sustainability, which can be conducted through various passive technologies such as Energy Harvesting (EH) techniques. EH technologies are environment-friendly and reliable approaches that have the ability to expand the lifespan of IoT devices, enable multifunctional wireless networks, while also diminishing the disadvantages of conventional batteries. The wind and solar photovoltaic capacity experiences a threefold growth, surging from approximately 75 GW in

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

2020 to 230 GW by 2030 [3]. In addition, EH technologies are the leading part of the Net Zero 2050 project [4],[1] in which the trade-off between emitted greenhouse gases to the atmosphere and the amount of removed greenhouse gases from the atmosphere is balanced. On the one hand, Medical IoT (MIoT) allocates 20% of the global IoT systems [5], and on the other hand, healthcare is responsible for 4-5% of the emissions of greenhouse gases. Thus, using these technologies in MIoT delivers a threefold advantage: reducing the amount of greenhouse gas emissions [6], lowering maintenance costs, and improving human well-being [7].

An illustrative example of the importance of the EH in MIoT can be presented in a pandemic situation, such as the COVID-19 crisis in 2020. Hospitals' total capacity was nearly occupied by patients needing special medical care, and thus, field hospitals or mobile medical units had to be quickly assembled under critical circumstances. Since one essential issue in establishing mobile medical units is to provide reliable and adequate energy sources, cooperating energy harvesting technologies (solar cells, piezoelectric, and thermoelectric harvesters) might assist in offering sustainable communications and energy sources, especially for medical devices and monitoring systems.

Pursuing sustainability in Information and Communication Technology (ICT), specifically wireless communications, is a concerning issue where carbon-based energy carriers fuel systems. It has been estimated that around 80% of greenhouse gas emission is due to carbon-based energy carriers (fossil fuels such as coal, oil, natural gas, gasoline, and diesel fuel) [8]. As stated by the authors in [9], wireless communication constitutes 75% of ICT, with wireless being the predominant mode of communication in IoT systems. Wireless communications presently contribute to 4% of the total global CO_2 emissions, which is projected to rise due to the growing number of connected devices. To reduce the emission of greenhouse gases, it is necessary to understand the energy requirement of the systems and minimize the usage of these energy carriers, whether by introducing EH technologies and renewable energy carriers or applying energy-efficient methods. In this regard, selecting an appropriate wireless technology and integrating it with a proper EH technique in terms of power density and form factor is essential for successful integration. Optimization algorithms such as channel adaptation or energy-aware routing algorithms have been used to reduce energy consumption at different layers of the IoT protocol stack. However, since Medium Access Control (MAC) layer operations consume most of the wireless communication's energy budget, this layer can benefit more from optimization algorithms such as channel access optimization methods. Given that advanced energy optimization methods involve the consideration of increasingly complex features, the significance of Machine Learning (ML) algorithms in this context cannot be underestimated.

IEEE 802.11, commonly referred to as Wi-Fi, is the dominant wireless communication technology in indoor IoT systems (it is reported that 51% of the wireless communication in 2022 belongs to Wi-Fi communication [10]). The channel access method of legacy IEEE 802.11 includes the contention-based Enhanced Distributed Channel Access (EDCA) mechanism, which defines four Access Categories (AC) for provisioning Quality of Service (QoS) for different traffic types based on the MAC layer parameters. The aforementioned EDCA mechanism is used to support service differentiation by assigning different Contention Window (CW) sizes, transmit opportunity (TXOP), Arbitration Inter-Frame Space (AIFS), and retransmission limit value. However, it faces inherent issues of using static parameters assignment of CW size AIFSN, TXOP limit, and retransmission limit without taking into consideration the current status of ACs as well as the number of stations competing to gain access to the shared channel. In addition, there might be another issue, where stations can act selfishly and choose a very small CW in order to increase their channel access [11]. This results in a decrease in channel access opportunities for well-behaved stations. Lastly, EDCA does not differentiate between applications with the same traffic type but different levels of QoS requirements. This means that applications with high QoS requirements may experience higher latency or packet loss than applications with low QoS requirements. Consequently, these issues affect the manner in which the stations contend to access the shared channel, which leads to more collisions and thus impacts the overall performance of the network. These issues become a complex problem in dense deployments, which are characterized by frequent collisions [12]. The increased collision rate can impact energy consumption as they need retransmissions and activating collision avoidance mechanisms, leading to increased energy usage. Therefore, integrating EH techniques with a dense Wi-Fi network in a medical environment poses significant challenges due to collisions and increased energy consumption, making it a complex problem. Moreover, the existing IEEE 802.11 channel access mechanism with predefined MAC layer parameter configurations is unable to meet the specific QoS requirements mandated in medical settings. Despite these obstacles, exploring the implementation of EH techniques in such networks remains essential to unlock their potential benefits.

In recent years, ML algorithms have demonstrated a powerful capability to improve and evolve optimization problems from classical optimization methods in wireless networks [13]. For instance, features such as the Access Point (AP) coordination mechanism in Wi-Fi 7 and beyond can reduce the network's energy consumption by coordinating the schedules of the transmission time between APs, and reducing the overall delay of the network [14]. To meet this feature, complex configurations, and non-linear optimization

[1]The the United Nations (UN) and the Intergovernmental Panel on Climate Change (IPCC) lead in promoting net zero emissions. The 2015 Paris Agreement, under the UNFCCC, urges nations to achieve net zero emissions by the latter half of the 21st century.

are required that can be fulfilled through ML-based algorithms. Thus, ML algorithms are necessary, particularly in the MAC layer operations and mechanisms, to make the configuration dynamic, flexible, and energy-efficient for dense and heterogeneous networks while provisioning QoS requirements. Furthermore, these algorithms can play a crucial role in optimizing the MAC layer to support EH techniques. By leveraging ML, the MAC layer can adaptively adjust parameters based on real-time network conditions and harvested energy availability. This integration of ML and EH enables intelligent resource management, reducing collisions, improving energy efficiency, and maximizing the benefits of energy harvesting in wireless networks. Through efficient resource allocation and intelligent energy management, ML-enhanced MAC layer operations help in enabling a more sustainable approach to wireless communication by minimizing energy waste and prolonging the lifespan of battery-powered devices.

To the best of our knowledge, this is one of the first papers that tries to achieve sustainability in IoT-based QoS-restricted[2] dense MIoT scenarios. Three ML-based algorithms are proposed which intend to guarantee the QoS requirements - End-to-End delay (E2E delay) and Packet Loss Ratio (PLR) - while maximizing the total remaining energy of the network and consequently reducing the emission of the greenhouse gases. This article is recapitulated in the following contributions:

- We propose novel RL-based optimization algorithms for a solar-based Wi-Fi system in a medical IoT scenario.
- We assume the AP coordination concept from the upcoming Wi-Fi amendment (IEEE 802.11bn) while supporting backward compatibility with the IEEE 802.11 standard.
- We present an objective function to maximize remaining energy and minimize E2E delay and PLR for medical-grade QoS criteria.
- We accomplish an extensive set of simulations on Network Simulator 3 (ns-3) to evaluate the suitability of our proposals.

The remainder of this article is structured as follows. In Section II, we highlight the relevant studies in the literature. The existing problem is elaborated in Section III. Sections IV and V explain the methodology and simulations setup, respectively. Section VI is devoted to the performance evaluation of the proposals' analytical discussions. Finally, in Section VII, we provide final remarks and future work.

II. RELATED WORK

Modifying and optimizing the MAC layer operations of IEEE 802.11 has undergone a lot of investigations and research studies. However, these optimization studies do not address the dense EH-based QoS-restricted environments, which require dynamic changes, such as medical IoT systems where real-time, emergency, and multimedia applications are employed.

A set of techniques described in the literature consists of modifying the initialization of the MAC layer parameters, such as CW, AIFS, and TXOP, to adjust the channel access scheduling in the EDCA mechanism. Dynamic initialization of AIFS is presented in [15] to support QoS requirements for real-time and non-real-time medical applications. Nevertheless, this algorithm is not able to meet QoS in a saturated condition. Other works intend to address QoS-restricted environments by defining fixed or dynamic CW values adaptation algorithms. In [16], the authors explain an algorithm that doubles the CW values when the channel is busy, or collision occurs, whereas in [17], the authors define a dynamic selection of CW values based on the traffic load of the network to reduce the collision rate and transmission delay in the network. However, these optimization algorithms modify the fundamental of the EDCA mechanism, and they may not be compatible with the IEEE 802.11 standardization. In one of the most recent MAC layer modifications [18], the authors proposed an enhanced Preliminary Channel Access (PCA) method, which allows the transmission for Real-Time Application (RTA) access to the channel faster than other transmissions compared to PCA and EDCA mechanisms. The aim of this study aligns with Wi-Fi 7, where low delay and high reliability are the two requirements for the RTA use cases. The proposed mechanism provides backward compatibility in Wi-Fi scenarios. However, this work does not include any energy-related analysis.

There are a few energy-harvesting MAC layer protocols in the literature which consider the integration of the energy harvesting technologies with WLAN communication. Some of these works try to achieve an energy-efficient MAC protocol even by only increasing the energy budget of the network or reaching an optimal energy consumption point. In [19], the authors propose algorithms to reduce the energy consumption of a Wi-Fi solar-based network. Whereas in [20], an optional energy-saving mode feature in IEEE802.11ah is evaluated along with the EH technique deployment. However, the authors do not consider QoS restricted environment in these works.

Another set of techniques used in the literature is based on ML. Specifically, RL is able to provide reliable solutions to complex decision-making problems. Recently, these techniques have attracted more attention among researchers in the wireless communication networks area. Proposing new features in the IEEE 802.11be and beyond to meet the IoT systems requirements (distributed management and deployment) may increase the Wi-Fi network's density, dynamic condition, and complexity. Thus, deploying RL-based algorithms becomes more influential in this scenario [21]. The RL-based algorithms can be deployed in IEEE 802.11 standard to optimize the existing techniques and the defined parameters, which lead to reducing the

[2]QoS-restricted scenario refers to a situation where the available network resources, such as bandwidth, capacity, energy, and processing power, are limited due to the high number of connected stations. This limitation creates challenges in delivering the desired level of service to all applications or users simultaneously.

collision probability, increased throughput, and optimized frame length [22], [23], [24]. Recent studies demonstrate that Deep RL (DRL) algorithm can improve the handovers in mmWave communications [25], optimize the resource unit allocation for multi-user scenarios [26], configure the channel bonding [27], or address the channel allocation and AP clustering issues in MIMO networks [28]. Other studies on RL-based algorithms focus on Wi-Fi management, such as the works presented in [29], [30] for channel and band selection or management architecture [31].

An RL-based optimization algorithm for updating the CW value based on the collision probabilities is proposed in [32], in which the throughput increases while the delay maintains a certain level. Although this mechanism introduces an optimal point between the collision rate probability and CW increase or decrease, it does not consider the network's total energy consumption and provision of QoS for applications with different traffic types.

As indicated in Table 1, most of the aforementioned studies benefit from the integration of the traditional RL with deep neural networks, such as Deep RL (DRL) to handle complex tasks, Deep Q-Network (DQN) to approximate the Q-value function, Deep Deterministic Policy Gradient (DDPG) to approximate both the policy and Q-value functions and offering better performance in continuous action domains, and Deep Q-Learning (DQL) to handle high-dimensional state spaces more effectively. In addition, some other approaches focus on maximizing the reward, such as Multi-Armed Bandit (MAB) or the ones that consider interaction between agents, such as Multi-Agent RL (MARL).

As stated in the introduction, while EH integration can help IoT systems attain sustainability by lowering the demand for traditional batteries, it also introduces new obstacles regarding physical dimensions, communication protocols, and user privacy. Thus, it is necessary to reduce the energy budget of the IoT system to tackle energy-related issues, which can be addressed through energy model optimizations. One energy model optimization, proposed in [33], was the first work to present the MAC layer operation modifications (optimal selection of CW initialization) based on the type of medical applications and AP coordination concept. However, the authors conclude that an ML algorithm would benefit the performance of such systems.

Table 1 demonstrates how our research contributes to the field by comparing the features of the current study with the relevant existing literature. According to the extensive study on the existing literature, we believe that several outstanding throughput and fairness optimization studies propose innovative solutions that need a comprehensive transformation and re-imagining of the IEEE 802.11 standard. Nevertheless, ensuring backward compatibility poses a complex challenge, as existing and old equipment cannot be changed. It is essential to highlight that our effort is to strongly achieve backward compatibility in the proposed RL-based algorithms that focus on optimizing the energy (remaining) of the network.

To the best of our knowledge, no RL-based MAC layer optimization assesses the feasibility of integrating energy harvesting technologies aligned with provisioning QoS for medical applications. Thus, this article enhances the previous work and improves the algorithm's flexibility to the dynamic behavior of dense networks by introducing RL-based optimization algorithms. These algorithms are able to reduce the energy consumption of the MAC layer operations in a solar-based Wi-Fi network while meeting specific QoS medical-grade parameters for medical applications like PLR and delay. Thus, this novel integration provides a fresh perspective and yields significant advancements in sustainable IoT-based Wi-Fi communication.

III. PROBLEM STATEMENT

The IEEE 802.11 MAC layer is a contention-based distributed layer that consumes most of the energy budget of the network due to associated collisions, retransmissions, and back-off mechanisms [34]. Figure 1 demonstrates the default access technique known as the two-way handshaking scheme [35].

To assess channel conditions prior to transmission, the traditional IEEE 802.11 protocol employs a technique known as Physical Clear Channel Assessment (PHYCCA) at each network node. By measuring the energy level in the channel, the node can determine whether it is above a predetermined threshold, called the carrier sensing threshold (CST). If the measured energy level exceeds the CST, indicating that the channel is occupied, the node will defer its transmission and wait for the channel to become available. This mechanism allows nodes to avoid transmitting when the channel is already in use, reducing the likelihood of collisions and improving overall network efficiency.

Once a station detects the channel to be busy, it initiates a random back-off process by generating a random back-off time within a CW size. The CW defines the range of possible back-off values. In this process, a slotted binary exponential back-off random interval is selected from the range of [0, CW], where CW initially starts with a minimum value of CW_{min}.

If a transmission attempt is unsuccessful, the CW value is doubled, resulting in a larger range of possible back-off values, up to a maximum value of CW_{max}. This doubling process continues with subsequent unsuccessful transmission attempts, allowing for increased back-off times and reducing the chances of collisions. On the contrary, a successful transmission results in the CW value being reset to its minimum value, CW_{min}. This reset aims to exploit the clear channel and ensure quick access to the medium for the station that successfully transmitted its data.

In other words, the random back-off procedure is defined to reduce the collision probability, energy consumption, and to allow fair access to the medium in a distributed manner, which can be controlled by initializing and changing the MAC layer parameters value. The selection of appropriate MAC layer values is crucial, not only for efficient channel access but

TABLE 1. Features comparison of related work and our proposal.

Properties / Studies	Modification	AP coordination	QoS support	Energy harvesting	Energy optimization	Throughput optimization	ML approach	Compatibility with legacy	Year
Son et al. [15]	Dynamic AIFSN	✗	✓	✗	✗	✓	✗	✗	2016
Tian et al. [16]	BEB[1]	✗	✓	✗	✗	✓	✗	✗	2016
Syed et al. [17]	BEB	✗	✓	✗	✗	✓	✗	✗	2017
Chemrov et al. [18]	Improved PCA	✗	✓	✗	✗	✓	✗	✓	2022
Lin et al. [19]	Energy-based DCF	✗	✗	✓	✓[2]	✗	✗	✓	2015
Zhao et al. [20]	Random Sleeping Window	✗	✗	✓	✓	✗	✗	✓	2015
Kumar et al. [22]	Dynamic CW_{min}	✗	✗	✗	✗	✗	DRL	✓	2021
Guo et al. [23]	Adaptive time slot	✗	✓	✗	✗	✓	MARL	✓	2022
Cho [24]	Adaptive transmission rate	✗	✗	✗	✗	✓	RL	✓	2021
Koda et al. [25]	Beamforming	✗	✗	✗	✗	✗	DRL	✓	2019
Kotagiri et al. [26]	Distributed Resource Unit	✗	✗	✗	✗	✓	DQN	✓	2020
Luo et al. [27]	Channel bonding	✗	✗	✗	✗	✓	DQN	✓	2020
Krishnan et al. [28]	Distributed MIMO	✗	✗	✗	✗	✓	DDPG	✓	2019
Carrascosa et al. [29]	AP association	✗	✗	✗	✗	✓	MAB	✓	2019
Han et al. [30]	Handoff management	✗	✗	✗	✗	✓	DQN	N/A	2019
Bast et al. [31]	Slicing	✗	✗	✗	✗	✓	DQL	✓	2019
Ali et al. [32]	Scaled back-off	✗	✗	✗	✗	✓	QL	✗	2018
Famitafreshi et al. [33]	Optimal application-based CW	✓	✓	✓	✓	✗	✗	✓	2021
Our proposal	Dynamic MAC parameters	✓	✓	✓	✓	✗	RL	✓	2023

[1]Binary Exponential Back-off.
[2]The energy analysis is based on the energy consumption per successful transmission, but not total energy consumption.

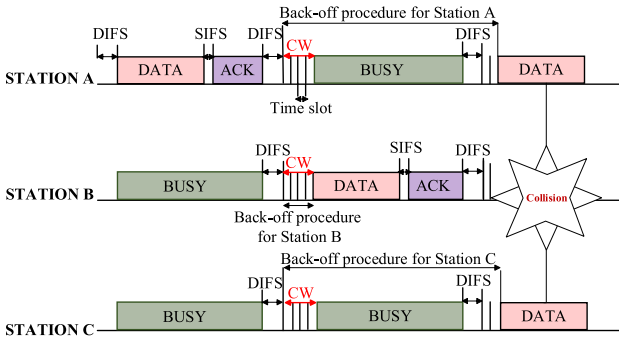


FIGURE 1. CSMA/CA back-off procedure.

also to minimize energy consumption related to collisions. For instance, a larger CW value reduces collisions and minimizes energy wastage associated with collision events, which is beneficial for energy efficiency. However, a larger CW value can also introduce increased delays in accessing the channel, potentially affecting QoS requirements, particularly for time-sensitive applications. On the other hand, a smaller CW value reduces delay but increases the probability of collisions and energy consumption. In addition, regarding other MAC layer parameters, such as AIFSN and TXOP, the system performance can be improved by adjusting these values. Reducing AIFSN values grants higher-priority frames a briefer back-off time, facilitating quicker transmission after the channel becomes idle. Modifying the TXOP permits consecutive transmission of multiple frames, minimizing the back-off and contention overhead. Thus, finding the right balance for the MAC layer parameters is crucial to optimize network performance, energy efficiency, and meeting the specific QoS restrictions imposed by the applications running on the network. Since IoT-based networks have a high level of collision probability, there is a need to apply modifications

TABLE 2. Quality of service requirements for e-Health applications.

Application Type	QoS Parameters			
	E2E Delay (ms)	Packet Loss Ratio (%)	Jitter (ms)	Sensitivity to Context
ECG [15], [36], [37]	<30	<10	25	✓
EEG [15], [36], [38]	<30	<10	25	✓
EMR [39], [40], [41]	< 100	<10	30	✗

to the selection of these parameters and adjust them to the condition of the network and application types. However, these modifications have to be aligned with the standard and address the restricted QoS e-health environments.

Although ML algorithms have been introduced to improve the performance of Wi-Fi communication effectively, they mainly focus on increasing the throughput without considering the QoS requirement and energy efficiency of the network, which are necessary for the successful integration of EH techniques within Wi-Fi-based IoT systems. To fill the existing gap in the literature, in this article, we formulate the optimization problem based on three RL-based algorithms that are able to initialize different MAC layer parameters (i.e., initialization of CW values) dynamically based on the network condition and application traffic type while maintaining the two critical QoS parameters for QoS-restricted e-health applications known as E2E delay and PLR. The medical-grade QoS restrictions that are considered in this study are listed in Table 2.

Along with the high collision probability, unfair access to the medium, and increasing throughput requirement to meet the emerging applications (i.e., health-tracking wearable, 4k and 8k video streaming, VR or AR and gaming) provisions, such as ultra-reliable low-latency communication requirements and extremely high throughput, the channel access mechanism of the Wi-Fi MAC layer faces another concern for prioritizing and coordinating the transmissions. The extremely high throughput and

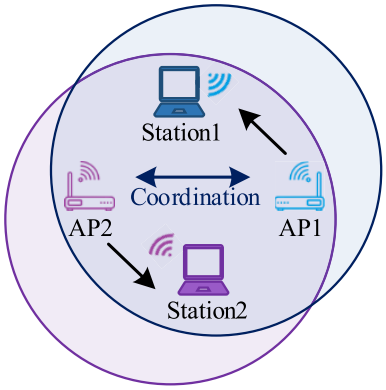


FIGURE 2. AP coordination concept.

ultra-low latency requirements listed above are beyond the capabilities of IEEE 802.11ax, even though the recently released IEEE 802.11ax emphasizes network performance and user experience in high-density deployment scenarios. IEEE 802.11ax only supports communication from a single AP and executes spatial reuse between APs and nodes without coordination among nearby APs. As a result, its ability to efficiently use time, frequency, and spatial resources is severely limited. In contrast, Wi-Fi 7 and Beyond with EHT capabilities improve this capability by allowing APs to share data and control information, increasing the throughput, reducing latency, and improving spectrum efficiency. Multi-AP coordination, which includes coordinated spatial reuse, coordinated orthogonal frequency-division multiple access, coordinated beamforming, and joint transmission, is one of the main differences between Wi-Fi 7 and Beyond and IEEE 802.11ax [14]. Therefore, to address these requirements, the upcoming amendment IEEE 802.11bn defines key concepts such as distributed multi-link operation, integrated mmWave operations, Physical (PHY) and MAC layer enhancement, and Multi AP coordination [42]. Among all these new features, controlling the delay of the system is possible based on AP coordination, which is a critical point for MIoT applications. The AP coordination technique states that to improve the performance of their associated non-AP stations, the so-called master APs can interact with other APs (slave APs) within their broadcast range. The master AP receives the beacon frames from the slave APs. The master AP might employ this technique to dynamically request the slave APs to rearrange the resources depending on the channel conditions (see Figure 2). While the need for uncoordinated systems is the main emphasis of this method, it is significant to emphasize that this strategy may also be used to coordinate systems [43].

In contrast to the aforementioned advantages of the concept of the AP coordination, it faces several issues when it comes to backward compatibility. The interoperability, influence on the network performance, implementation complexity, resource allocation, and overhead management are just a few of the new difficulties and challenges that come with AP coordination with backward compatibility. The functioning of AP coordination may cause interoperability issues when older devices find it difficult to interact with newer APs, resulting in decreased overall network performance, reliability, and efficiency, which can lessen the advantages of optimized performance of the current devices. The network's design and implementation may become more difficult, and additional management and configuration for overhead is needed if it supports both backward compatibility and advanced coordinating features. Furthermore, resource allocation for newer devices capable of handling more sophisticated coordination methods may be wasteful due to coordinating APs allocating resources based on the capabilities of old devices [44].

In this article, each group of nodes (i.e., non-AP stations) is associated with one respective AP in the assessed scenario. The main purpose of the proposed algorithms is to reduce the collision probability by differentiating the initialization of CW once per node and then per cell (i.e., all the nodes associated with their corresponding AP). However, per-node analysis increases the level of processing in the nodes, which can cause an increase in energy consumption. For this reason, centralized techniques are introduced to reduce the level of node processing. For example, the AP receives information about radio resource measurements of nodes [45]; therefore, the nodes with the low E2E delay increase their CW values to delay the data frame transmission.

In contrast, the nodes that exceed the medical-grade QoS threshold will reduce the CW to access the medium faster and immediately start the transmission. Furthermore, the agent makes decisions based on the remaining energy and E2E delay, which allows the node to harvest energy while maintaining medical-grade QoS requirements. In addition, the effect of the AP coordination concept is assumed in the proposed algorithms, in which APs are able to reschedule the resources based on the medium access conditions. Finally, a sleep/wake-up method is applied to obtain a higher level of energy reduction in the network.

IV. METHODOLOGY

This section utilizes RL-based optimization algorithms derived from Markov Decision Processes (MDP) to meet the proposed objectives of this paper.

In this study, MDP depicts the interaction between APs and nodes within a solar-based Wi-Fi, and it updates the decision-making in which an agent (AP) interacts with the environment. This framework models the optimization problem sequentially and is simplified as the following tuple.

$$(\Lambda, \Delta, \Gamma_a, \Phi_a) \tag{1}$$

where Λ is the representative of the group of states, which include the variables that define the environment or observation (in this paper, it corresponds to remaining energy, E2E delay, and PLR). Λ is updated at each step of the execution of the algorithms. Δ is the group of actions the agent executes in the AP, which is responsible for dynamic

Algorithm 1 Delay-Based Algorithm

Initialization: $\mathrm{CW_{min}} = 31, \mathrm{CW_{max}} = 1023$ Application-based QoS threshold
Input: Delay_node
Output: $\mathrm{CW_{new}}$_node
if Delay <= QoS Threshold **then**
    $\mathrm{CW_{new}} = ((\mathrm{CW_{current}} + 1) \times 2) - 1$
else if then
    $\mathrm{CW_{new}} = \left(\frac{\mathrm{CW_{current}}+1}{2}\right) - 1$
end if
return $\mathrm{CW_{new}}$
end procedure

changes of the MAC layer parameters (corresponding to the main operation of the proposed algorithms). $\Gamma_a\left(\Lambda^{t+1}|\Lambda^t, a\right)$ is the probability transition function depicting the probability of the action a (belongs to Δ) takes place in state t to reach state $(t+1)$ in the environment. Finally, Φ_a is the reward function, calculated after the execution of action a and updated at each step of the algorithm's execution to provide feedback for the decision-making in the next state [46], [47].

Each proposed algorithm highlights different decision-making methods for initializing MAC layer parameters. The first, second, and third sections address the initialization of MAC layer parameters for each node in the assessed scenario, where the CW values consider the primary parameter due to the random back-off procedure. Nevertheless, in Subsection IV-E, decisions are made based on the sleep/wake-up mode. Additionally, in Subsection VI-D, we explore the effects of AIFSN and TXOP (considered as key MAC layer parameters) adjustments, alongside an evaluation of the CW adjustments. The difference between the proposed algorithms can be explained based on the level of the information that is fed to the APs to make decisions (i.e., only considering E2E delay or considering E2E delay and remaining energy). It is expected that the more information is considered, the level of optimization will be higher. In addition, the way that each algorithm selects the associated nodes to apply the dynamic changes of MAC layer parameters is different, which can be performed cell-wise or node-wise. In particular, in the case of CW values, it is expected that the differentiation in CW initialization reduces collision probability, addresses the medical-grade QoS requirement, and gives nodes with lower energy levels more opportunity to harvest energy. Aligning with the CW value changes the impact of the AIFSN and TXOP dynamic adjustments, as other MAC layer parameters need to be investigated.

A. DELAY-BASED ALGORITHM

The first proposed algorithm is the delay-based algorithm, which aims to reduce the collision probability of each node by delaying or accelerating the data transmissions in each node. This differentiation in initiating and selecting CW helps to avoid extra collisions due to the simultaneous transmissions.

Algorithm 2 Extremum-based AI algorithm.

Initialization: $\mathrm{CW_{min}} = 31, \mathrm{CW_{max}} = 1023$ Application-based QoS threshold
Input: Delay_node, RemainingEnergy_node
Output: $\mathrm{CW_{new}}$_node
for `<nodes associated to the same AP>` **do**
    `<Sorting based on the node's remaining energy>`
end for
if node with maximum remaining energy **then**
    if Delay <= QoS Threshold **then**
        $\mathrm{CW_{new}} = ((\mathrm{CW_{current}} + \alpha_{min}) - 1)$
    else if then
        $\mathrm{CW_{new}} = ((\mathrm{CW_{current}} - \alpha_{max}) - 1)$
    end if
end if
if node with minimum remaining energy **then**
    if Delay <= QoS Threshold **then**
        $\mathrm{CW_{new}} = ((\mathrm{CW_{current}} + \alpha_{\mathrm{max}}) - 1)$
    else if then
        $\mathrm{CW_{new}} = ((\mathrm{CW_{current}} - \alpha_{\mathrm{min}}) - 1)$
    end if
end if
return $\mathrm{CW_{new}}$
end procedure

As explained in Algorithm 1, the agent (AP) checks the delay value for each node individually and makes decisions based on comparing the obtained E2E delay values with the defined corresponding medical-grade QoS threshold for each node at each time that the algorithm runs.

B. RANK-BASED EXTREMUM NODES ALGORITHM

In the second algorithm, the decision-making is made per cell to reduce the level of complexity and increase the residual energy at the node level. Then, within each cell, the algorithm selects the nodes with the extremum (i.e., the nodes with the minimum and maximum remaining energy) value of remaining energy, while the rest of the nodes store their current CW values and do not enter the following steps. According to Algorithm 2, the CW values are reduced by the factor of α ($\{2, 4, 8, 16\}$) if the delay of the extremum node exceeds the QoS requirement threshold. Otherwise, the α value is added to the current CW values. It is worth mentioning that the selection of the α value is in line with the IEEE 802.11 standard initialization of the CW values. This algorithm aims to assign different values of CW to the node to give the ones with the minimum energy the opportunity to harvest more energy and those with the maximum value of remaining energy to start the data frame transmission immediately.

C. RANK-BASED ALL NODES ALGORITHM

In contrast to the second algorithm, the decision-making is made at each node in the third one. In Algorithm 3,

Algorithm 3 Rank-Based All Nodes AI Algorithm

Initialization: $CW_{min} = 31, CW_{max} = 1023$ Application-based QoS threshold
Input: Delay_node, Remaining Energy_node
Output: CW_{new}_node
for <nodes within cell> **do**
    <sorting based on the node's remaining energy and select the median node>
end for
if remaining energy_node > node with median remaining energy **then**
    if Delay <= QoS Threshold **then**
        $CW_{new} = ((CW_{current} + \alpha) - 1)$
    else if then
        $CW_{new} = ((CW_{current} - \alpha) - 1)$
    end if
end if
if remaining energy_node < node with median remaining energy **then**
    if Delay <= QoS Threshold **then**
        $CW_{new} = ((CW_{current} + \alpha) - 1)$
    else if then
        CWnew $= ((CW_{current} - \alpha) - 1)$
    end if
end if
return CW_{new}
end procedure

within the cell, the node with the median value of remaining energy keeps the CW values the same as its previous state. In comparison, the nodes with greater remaining energy than the median value slightly increase their delay duration if their delay lowers the QoS requirement. Otherwise, their CW values are decreased by the largest defined α values ({16,8}) to start the transmissions faster, differentiate the transmission's starting point, and reduce the collision probability.

In the opposite case, the nodes with the remaining energy values less than the median value need to increase their opportunity to harvest more energy for continuing the communication. The nodes with a delay less than QoS requirements increase the delay value by selecting larger CW values ($\alpha = 16,8$). On the contrary, although nodes have to start the transmission immediately to reduce the delay, the algorithm delays the communications for a short duration to allow them to harvest energy and begin the frame transmission while maintaining the QoS restrictions ($\alpha = 4,2$). Thus, in this algorithm, the agent allows all the nodes to differentiate the start of communication by initiating the CW values based on the condition of the channel and their remaining energy.

D. PROPOSED ALGORITHMS FOR AIFSN AND TXOP ADJUSTMENTS

While the RL-based algorithm designed for initializing CW values could be repurposed for adjusting other MAC layer parameters such as AIFSN and TXOP, it is essential to acknowledge the intrinsic differences between the CW and AIFSN, where AIFSN remains constant, whereas CW adjustments rely on random processes. Given this distinction, we have made subtle modifications to the algorithms to accommodate AIFSN adjustments. The node with an E2E delay less than QoS requirements updates its AIFSN value in the next window as follows.

$$AIFSN_{new} = AIFSN_{old} + 1 \tag{2}$$

where the initial value for AIFSN is considered as 2, and the maximum value is set to 15. We propose a strategy to prevent extended idle periods due to the maximum selection of AIFSN for all the nodes or the same value for all the nodes. AIFSN value increases until it is set to its maximum allowable limit and the E2E delay exceeds the QoS threshold. In this case, when the AIFSN reaches this maximum in a given window, it is then reset to its initial value in the following window. By resetting the AIFSN value to the initial value, we increase the level of randomness in AIFSN selection and prevent undesirable delay for all the nodes.

Shifting our focus to TXOP adjustment, the procedures mirror those outlined in Algorithms 1 to 3, ensuring consistent and coherent changes. If the node shows an E2E delay less than the QoS threshold, the transmission is delayed by increasing the idle listening duration.

$$TXOP_{new} = TXOP_{old} - 32 \tag{3}$$

By reducing the TXOP value, the duration a node requires to access the channel for transmission is diminished. This, in turn, allocates greater chances for other nodes to initiate their transmissions. It's worth highlighting that any alterations to the TXOP timing limit must occur in the multiple of 32 microseconds. The permissible range for TXOP values spans from 0 to 7.04 ms. However, in this case, when the E2E delay exceeds the QoS threshold, the TXOP value needs to gradually increase, and reduce the delay value. Selecting a large TXOP value for one node or setting it to the maximum value affects the network's fairness and gives other nodes less opportunity to start transmission. Thus, to prevent this behavior if the E2E delay of other nodes exceeds the limit, each algorithm reduces the TXOP of the node with the highest value of TXOP. In contrast to the AIFSN adjustment, higher TXOP values increase the idle listening duration for other nodes and violate network fairness.

E. SLEEP/WAKE-UP MODE

In this algorithm, to reduce the collision probability and improve the energy efficiency of the network, instead of differentiating the CW values in the proposed RL-based

Algorithm 4 Sleep/Wake-Up Method

1: **Initialization:** $CW_{min} = 31, CW_{max} = 1023$ Application-based QoS threshold
2: **Input:** Delay_node
3: **if** Delay >= QoS Threshold **then**
4: node is forced to sleep
5: **else if then**
6: Trigger the proposed RL-based algorithms
7: **end if**
8: **end procedure**

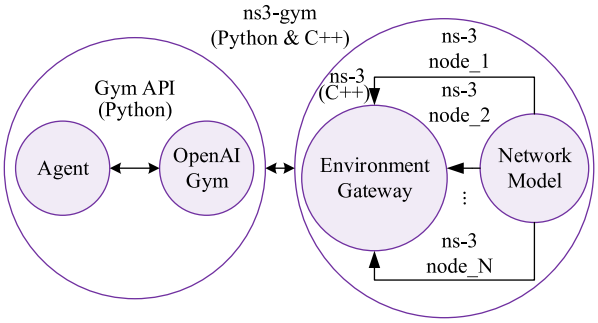


FIGURE 3. The general framework of the integration of RL with ns-3.

optimization algorithms, the selected nodes are forced to sleep mode, which will determine the starting time of the frame transmission of nodes (see Algorithm 4). The sleep/wake-up method procedure is detailed in [35]. This method mimics intermittent communication [48] when the communication is interrupted due to insufficient energy to keep the system powered up or a high level of interference in the channel.

V. SIMULATION SETUP

In this section, the system model and experimental environment to implement a dense solar-based Wi-Fi network in field hospital circumstances in the ns3-gym environment are described in detail. Then, the evaluation metrics are explained. This simulation setup allows us to assess the proposed RL-based optimization algorithms under specified conditions. Section VI will evaluate and discuss these algorithms.

A. SYSTEM MODEL

The proposed RL-based optimization algorithms are deployed in the ns3-gym framework, which consists of three main components, known as ns-3 simulator, OpenAI Gym, and ns3-gym middleware (see Figure 3).

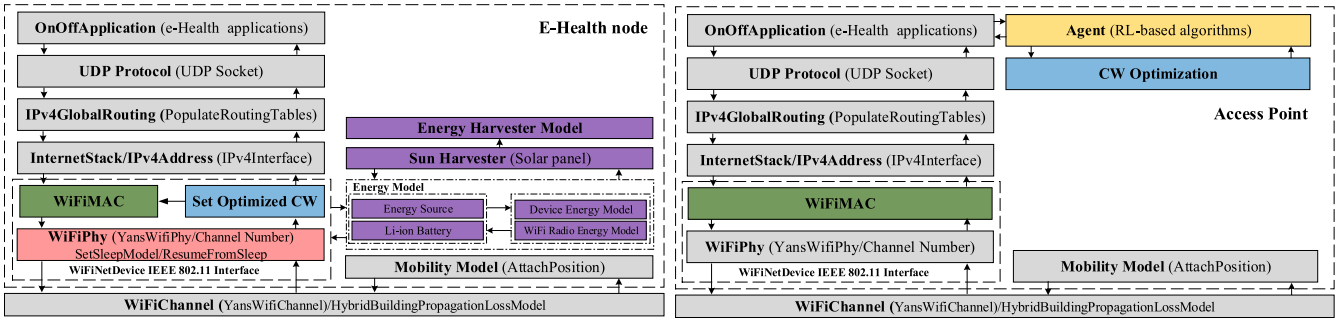
The ns-3 simulator is an open-source network simulator that mimics real-world constraints and provides features that need to be accomplished to meet IoT requirements [49], [50]. According to Figure 3, the ns-3 simulator is considered as an actual deployment of a solar-based dense Wi-Fi network in ns-3 (environment), and the E2E delay, PLR, remaining energy, and consumed energy are extracted from it (observation parameters). Since the reward function returns a numerical value as feedback to the action, the agent is able to make the most optimal decision based on this value. Here, the defined objective function (see Equation 4) is maximized based on the reward function feedback, where the remaining energy is maximized, and the E2E delay and PLR values are minimized.

$$\text{Objective Function} = \frac{\text{Remaining energy}}{\text{E2E delay} \times \text{PLR}} \tag{4}$$

The protocol level implementation of ns-3 Wi-Fi solar-enabled nodes is highlighted in Figure 4. The purple blocks correspond to components of energy-harvesting modules that have been integrated into the node. These modules define the type of source of energy (battery or a capacitor), Wi-Fi radio energy model, and device energy model. The device energy model locates the Wi-Fi radio energy model in each node for calculating the energy consumption of each state of transmission or the total energy consumption of the node [51]. It is worth mentioning that the solar energy harvester model, designed in [52], is appended to the energy-related model since the official version of ns-3 does not include this module. This implementation closely mimics the real behavior of a solar energy harvester that takes into account various aspects of the harvesting process in the ns-3 solar harvesting system. This technology realistically develops a solar panel and mathematically calculates many solar features influencing energy harvesting.

Along with the energy-related modules added to the ns-3 architecture, the proposed algorithms directly impact the initialization of the MAC layer parameters, shown in the green and blue blocks. These blocks describe the Wi-Fi MAC layer protocol implementation we adapt during the simulation. Another module that affects the MAC layer operation is the PHY layer, shown by the pink block. This block defines the different transmission states of the communication, where the sleep/wake-up method is defined to reduce the contention, which directly impacts the MAC layer [35]. The architecture of an AP in ns-3 is illustrated in Figure 4b, where the optimization problem is formulated, and the AP coordination concept is introduced (the blue block). In addition, the MAC layer modification commands and initialization values are sent to the nodes from the APs that the agent controls based on actions (the green block). Moreover, the PHY layer generates the sleep/wake-up commands triggered by the AP. Here, the APs decide when and which node has to go to sleep or wake-up mode. It is noted that no changes or modifications are applied to the gray-scale blocks.

The second part of the n3-gym is the ns3-gym middleware, which sends the gathered information to the environment gateway entity for saving the numerical data in a structured manner and encoding the received actions from the agent to numerical data. In addition, ns3-gym middleware receives the



(a) Architecture of a sensor node in ns-3.

(b) Architecture of an AP in ns-3.

FIGURE 4. Ns-3 IEEE 802.11 architecture to support RL algorithm and EH.

environment information (i.e., observation) and sends it to the agent.

The third part of the ns3-gym framework is the OpenAI Gym, which fundamentally is a toolkit capable of creating new ML algorithms in a range of simulated environments. The Python process, consisting of the agent and the Gym environment, establishes communication with the C++ process, responsible for the ns-3 network simulation, through ZMQ sockets. Readers interested in learning more about the OpenAI Gym are referred to [53].

B. ENERGY MODEL FOR THE CONSIDERED WI-FI SCENARIO

Various components contribute to the overall energy consumption of Wi-Fi systems. Yet, achieving an exact calculation of the total energy consumption proves intricate, owing to the dynamic characteristics of wireless channels, the complexity of the Wi-Fi protocol, and the fluctuations in network traffic patterns. This section delves into a simplified explanation of the total energy consumption within the specific Wi-Fi scenario under consideration.

The total energy consumption of the Wi-Fi system corresponds to the sum of the energy consumption during the fundamental states, such as transmission, reception, transition, idle, and back-off states, which is formulated as follows.

$$E_{\text{Total}} = \sum_{i=1}^{m} \big(E_{\text{Tx(m)}} + E_{\text{Rx(m)}} + E_{\text{Idle(m)}} + E_{\text{Transition(m)}} + E_{\text{Back-off(m)}}\big) \tag{5}$$

where

$$\begin{aligned} E_{\text{Tx}} &= E_{\text{Ack-tx}} + E_{\text{Data-tx}}, \\ E_{\text{Rx}} &= E_{\text{Ack-rx}} + E_{\text{Data-rx}}, \\ E_{\text{Transition}} &= E_{\text{Sleep}} + E_{\text{Wake-up}}, \\ E_{\text{Back-off}} &= E_{\text{BO}} + E_{\text{Collision}} \end{aligned}$$

In this case E_{Tx}, E_{Rx}, $E_{\text{Transition}}$ correspond to the energy consumption in the transmission state, reception state, and changing from sleep mode to wake-up or vice versa. In addition, this formula considers the energy consumption during the back-off procedure ($E_{\text{Back-off}}$), which includes initialization of the CW procedure and collision. Here, m refers to the number of the stations in the network. Furthermore, each component can be decomposed as the multiplication of the power and duration of that state as follows.

$$\begin{aligned} E_{\text{Tx}} &= P_{\text{Ack-tx}} \times T_{\text{Ack-tx}} + P_{\text{Data-tx}} \times T_{\text{Data-tx}}, \\ E_{\text{Rx}} &= P_{\text{Ack-rx}} \times T_{\text{Ack-rx}} + P_{\text{Data-rx}} \times T_{\text{Data-rx}}, \\ E_{\text{Idle}} &= P_{\text{Idle}} \times T_{\text{Idle}}, \\ E_{\text{Transition}} &= P_{\text{Sleep}} \times T_{\text{Sleep}} + P_{\text{Wake-up}} \times T_{\text{Wake-up}}, \\ E_{\text{Back-off}} &= P_{\text{BO}} \times T_{\text{BO}} + P_{\text{Collision}} \times T_{\text{Collision}} \end{aligned}$$

According to the EDCA mechanism, the duration of the acknowledgment, successful data frame transmission, idle [54], back-off procedure, and collision are defined as follows.

$$\begin{aligned} T_{\text{Ack}} &= DIFS + T_{\text{Slot}} + SIFS, \\ T_{\text{Data}} &= DIFS + T_{\text{Slot}} + SIFS + T_{\text{Ack}}, \\ T_{\text{Idle}} &= AIFS, \\ T_{\text{BO}} &= (CW - 1) \times T_{\text{Slot}}, \\ T_{\text{Collision}} &= SIFS + T_{\text{Ack}} + AIFS \end{aligned}$$

It is important to mention that the energy consumption during the frame retransmission is considered in the $E_{\text{Collision}}$, in addition, the sleep and wake-up duration could vary based on the specific AC and power-saving settings that is considered in the algorithm.

In line with the total energy consumption formula and aforementioned analysis, an inverse relationship exists between the CW and total energy consumption ($\frac{1}{\text{CW}} \propto E_{\text{Total}}$). As the CW value decreases, the duration before transmission becomes shorter. This reduction in waiting time leads to increased transmission attempts due to more active states and potential retransmissions, ultimately resulting in elevated energy consumption.

A simulation-based decomposition of the energy consumption of the network is analyzed in [35], where the E_{Tx},

E_{Rx}, E_{Idle}, $E_{Transition}$ and $E_{Back\text{-}off}$ states consumed 0.5 J, 3.42 J, 9.02 J, 1.02 J, and 0.63 J, respectively. According to the analyzed energy model, the network's total energy consumption equals 14.59 J, and each node consumes 0.364 J in the system. In this paper, we examine all the MAC layer parameters outlined in Equation 5, encompassing not only those mentioned but also encompassing wake and sleep states.

C. NETWORK SCENARIO DEFINITION AND EVALUATION METRICS

The simulation environment is a one-floor field hospital (similar to an office-type building) with an area of 3200 m^2 and 3 m of the height of the floor. According to Figure 5, this area is divided into 8 symmetric rooms (each room area is 20 m × 20 m). The rooms are separated through wooden walls (brown lines), and the external walls are defined as concrete walls with windows (black bars). The simulations were carried out using the hybrid building propagation model, which provides the required flexibility to represent the Wi-Fi implementation in a building environment. This model includes several factors, such as the frequency in use, the environments (urban, suburban, or rural), International Telecommunication Union defined path loss model [55], the position of the interfering nodes, external wall penetration loss of different types of buildings (such as windows, without windows, concrete among others), and the internal wall loss. All of the factors mentioned above, among others, are used to derive the indoor path loss, which is used along with the transmit signal power to derive the received signal power. By design choice, we consider a single AP per room, in which one AP is located in the middle of the room, and 5 nodes are placed randomly within the room. The distance between the AP and each node is randomly selected from 1 to 10 m. This article considers the worst-case scenario of Wi-Fi communication, where the 2.4 GHz frequency band is used for communications. The restricted bandwidth of 2.4 GHz, with only three non-overlapping channels, causes high interference, and due to the few non-overlapping channels, the 2.4 GHz frequency can become congested quickly. For this reason, the main issue of contention-based communications in the 2.4 GHz frequency band is the high collision probability.

It is important to note that we have employed IEEE 802.11n in the simulated scenario. This choice ensures that the scenario that has been discussed here aligns precisely with the enterprise model outlined in the IEEE 802.11ax standard and described in [59]. Furthermore, most assessments of Wi-Fi 7 and Beyond follow the operational guidelines established by the IEEE 802.11ax Working Task Group. This is because the development of Wi-Fi 7 builds upon discussions held with Task Group ax(TGax), essentially serving as an extension of those conversations.

To evaluate the performance of the proposed RL-based optimization algorithms in each set of simulations, we consider 8 AP (triangles in Figure 5) and nodes associated with each one (red circles). In addition, the master APs

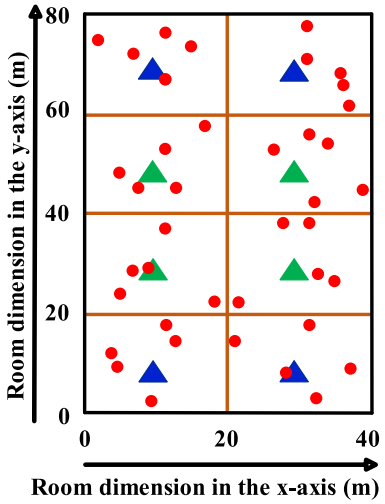


FIGURE 5. Layout of the Wi-Fi deployment in the field hospital.

TABLE 3. PHY layer parameters for the simulation.

Parameter	Value
Wireless Standard	IEEE 802.11n
Frequency band	2.4 GHz
Physical transmission rate	MCS 5 for data frames
Physical transmission rate	MCS 0 for control frame
Propagation loss model	Hybrid building propagation loss
External Wall penetration loss	7 dB
Internal Wall penetration loss	4 dB
Transmission power	16 dBm
Energy detection threshold	−62 dBm
CCA mode1 threshold	−82 dBm
Guard interval	Short
Channel bandwidth	20 MHz
Channel Number	1
Number of the AP	8
Stations per AP	5

TABLE 4. MAC layer parameters for the simulation.

Parameter	Value
Default CW_{min}	31
Default CW_{max}	1023
AIFS	3 μs
TXOP	0 ms
RTS/CTS	Disabled
MSDU aggregation	Disabled
MPDU aggregation	Disabled

are illustrated as green triangles and the slave APs as blue triangles. Each node uses three medical applications in the simulations: ECG, EEG, and EMR (representing medical file transferring). For the final results, each set of simulations is repeated 20 times with different seed values (to add randomness to the implementation) to increase the confidence level.

The parameters of the PHY and MAC layers are listed in Table 3 and Table 4. Since the Request to Send (RTS)/Clear to Send (CTS) mechanism (the four-way handshaking mechanism) increases the transmission time inherently, and a non-optimal frame aggregation can increase the error rate,

TABLE 5. Traffic characteristics in the simulation study.

Traffic Type	ECG	EEG	EMR
Access Category	BE	BE	BE
Traffic model	ON-OFF (0.650–0.350) CBR [56]	ON-OFF (0.29–0.71) CBR [15]	ON-OFF (0.05–0.95) Exponential [15]
Data rate	12 kbps [57]	32 kbps [58]	4.1 Mbps [15]
Packet size (Bytes)	147 [15]	155 [58]	1528 [15]

as is shown in Table 4, frame aggregation and RTS/CTS are disabled. The default MAC layer parameters are selected as the Best Effort (BE) Access Category (AC) corresponding to the traffic model of the three medical applications. Only the CW values change dynamically in this paper, and AIFSN and TXOP are maintained constant. It is noted that during the MAC layer modifications, all the parameters of the PHY layer are kept constant according to Table 3.

Table 5 summarizes the traffic model, data rate, and packet size for ECG, EEG, and EMR applications as the selected medical applications.

The Li-Ion battery and the panel dimension (corresponding to the size of a remote blood oxygen monitoring [60]) are adopted from [35]. The metrics for the evaluation are defined based on their equations in Table 6.

This setup allows us to explore the performance of our proposed RL-based algorithms by automatically varying the CW values as a MAC parameter for nodes based on the current condition of the channel and then with the offered sleep/wake-up method.

VI. PERFORMANCE EVALUATION AND DISCUSSION

In this section, through extensive ns3-gym-based simulations, we evaluate the performance of the proposed RL-based optimization algorithms in the selected environment. The performance evaluation analysis first focuses on CW value changes to define the usefulness and assess the proposed RL-based algorithms. In Section VI-D, we investigate the impact of other MAC layer parameters on the performance of the network. For this reason, we will use a consistent EDCA queue, denoted as EMR, to assess the effects of altering various MAC layer parameters. Subsequently, we will contrast the outcomes with those of a system adhering to default values (Tabel 4). This clarification will be explicitly presented in Section VI-D. As highlighted in Section III, to manage the network resources more efficiently and decrease E2E delay in this scenario, the concept of AP coordination will be introduced to the proposed algorithms and evaluated in this section. In this study, the master and slave cells are distinguished based on their Frame Error Rate (FER).

A. COMPARISON OF LEGACY WITH THE PROPOSED ALGORITHMS FOR CW ADJUSTMENT

Previous studies indicated that there is an optimal value for initializing CW for a node that can retain the collision probability at the lowest possible level and, as a result, reduces energy consumption while maintaining the E2E latency and PLR metrics below the medical-grade QoS requirements [35]. Additionally, this optimal value varies depending on the application and traffic types and the condition of the network. Therefore, it is necessary to dynamically assign an optimal CW value to each node in IoT-based networks to prevent long E2E delays and high collision probability. It is expected that such a scenario will benefit from RL algorithms. In each algorithm initially the CW_{min} and CW_{max} values are set to the default values (see Table 4). Then, these values are selected based on the received information from the environment.

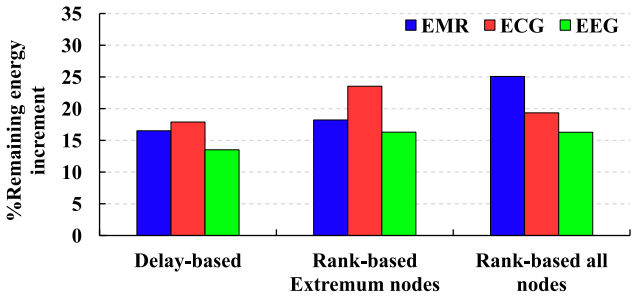


FIGURE 6. Legacy comparison with the proposed algorithms under CW adjustment for medical applications.

Figure 6 illustrates the improvement of remaining energy by applying the RL algorithms in the Wi-Fi-based IoT system for medical applications. While the rank-based all-nodes algorithm, when used for EMR application, increases the energy efficiency up to 26.5% (from 14.68 J to 18.36 J), it reduces the E2E delay, PLR (15.87 ms and 2.4%receptively), ECG, and EEG applications benefit more from the rank-based extremum nodes algorithm. Where the remaining energy improves up to 23.5% and 16.2% (4 and 3 J improvement per application), delay reduces to 9 and 3 ms for ECG and EEG applications, respectively (PLR is negligible in ECG and EEG applications). The results are shown in Table 7 can be explained by differentiating the CW values initialization among nodes based on the traffic model characteristics. As indicated in Figure 7, the selected CW values are relatively large in the delay-based algorithm, which increases the network delay considerably and consumes more energy. However, since the CW values initialization is more distributed among all possible intervals in the rank-based all-nodes algorithm, it has a better adaptation to the network condition. Consequently, it can save more energy for applications with high data rates and large packet sizes, such as EMR applications. In contrast to the high-traffic load applications, the applications with low traffic load levels benefit the rank-based extremum nodes algorithm with a lower level of complexity.

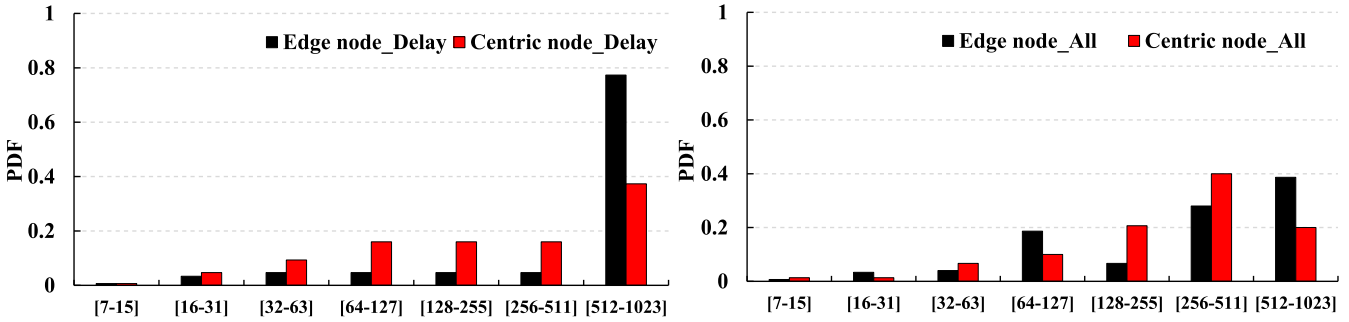
B. COMPARISON OF LEGACY WITH THE PROPOSED ALGORITHMS UNDER AP COORDINATION ASSUMPTION FOR CW ADJUSTMENT

In this set of simulations, we evaluate the impact of the AP coordination concept on the performance of the network with

TABLE 6. Metrics for evaluation.

Metric	Definition	Equation	Reference
E2E delay	Starts by generating the frame from the source until it reaches its destination	$D_{\mathrm{Tx}} + D_{\mathrm{queuing}} + D_{\mathrm{contention}}$	[61]
Remaining energy	Includes initial energy of a Li-Ion battery and total harvested energy subtracted by the consumed energy	$\mathrm{E}_{\mathrm{Initial}} + \mathrm{E}_{\mathrm{Harvested}} - E_{\mathrm{Consumption}}$	[35]
PLR	Number of lost packets in the transmission procedure[3]	$1 - PDR$	[35]
FER	Number of lost frames in the transmission procedure[4]	$1 - FSR$	[11]

[3]Packet Delivery Ratio (PDR) is the number of the packets that are received successfully.
[4]Frame Success Rate (FSR) is the number of the acknowledged frames divided by the total transmitted frame.



(a) Delay-based algorithm. (b) Rank-based all-nodes algorithm.

FIGURE 7. Probability Density Function for CW changes for the proposed RL-based algorithms.

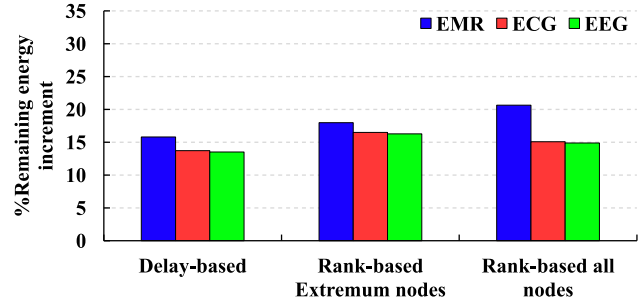


FIGURE 8. Impact of master cells AP coordination on the proposed algorithms under CW adjustment.

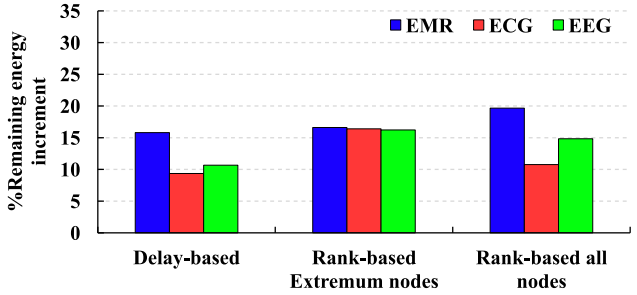


FIGURE 9. Impact of slave cells AP coordination on the proposed algorithms under CW adjustment.

three proposed RL-based optimization algorithms. If the cell has a FER value greater than the average FER of the network, it is considered the master cell, otherwise, it is deemed a slave cell.

For this scenario, the proposed algorithms are deployed only at the master cells. As illustrated in Figure 8 the system faces improvements in the evaluated metrics compared to the legacy in all the cases. However, the improvement percentage is less than the case where we implement three algorithms in all the cells (cf. Table 7). The logic behind the obtained results can be explained through the differentiation of the initialization in the CW values for central and edge nodes, where the master cells (located as central nodes) face more variation in the initialization of the CW values. However, for the nodes in the slave cell, the initialization is fixed to the default values. The results prove that more variation in the initialization of the CW values based on the network condition can improve the energy efficiency of the network while reducing the QoS parameters such as delay and PLR.

In the following scenario, when the proposed algorithms are applied to the slave cells, the obtained results demonstrate a lower improvement level than the previous results. Regardless of the network conditions, the initialization of the CW values for the nodes in the center (which face more collisions) is fixed as the default values. Therefore, the network reaches at most 19.60% remaining energy improvement for EMR application in the case of the rank-based all-nodes algorithm, 16.40%, and 16.23% in the case of ECG and EEG application under the deployment of the rank-based extremum nodes algorithm (see Figure 9).

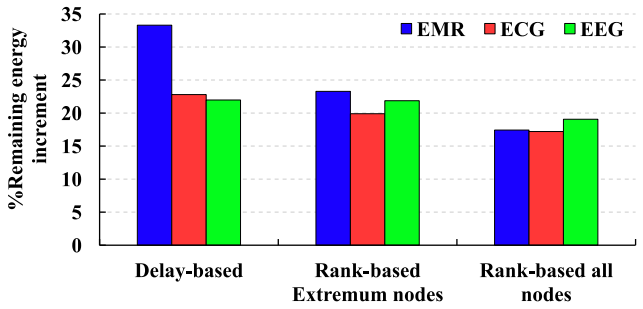


FIGURE 10. Impact of sleep/wake-up method on the proposed algorithms.

It can be concluded that when the proposed algorithms are only applied to a group of cells in the AP coordination, although the E2E delay can be controlled through the AP coordination concept, the fixed CW initialization for the other group of cells can degrade the network's overall performance. For this reason, the overall network improvement is smaller than when the proposed algorithms are applied to all the nodes. It can be perceived from Table 7 improvements are less than 1 J for the remaining energy metric for all the applications regardless of whether the AP coordination is deployed in master or slave cells.

In the following set of simulations, to reduce the collision probability, and consequently, increase the remaining energy, instead of differentiating the CW initialization on the nodes, the algorithms will force the node to go to sleep.

C. COMPARISON OF LEGACY WITH THE PROPOSED ALGORITHMS UNDER SLEEP/WAKE-UP METHOD

In this set of simulations, we introduce the sleep/wake-up method (introduced in [35]) to our proposed algorithms. In this case, the algorithm restricts the node to transmit, with a condition that the E2E delay does not exceed the set QoS threshold. Therefore, the sleeping node has the opportunity to harvest energy and then, in the next time step, follows the procedure of the proposed RL-based optimization algorithm. According to Figure 10, all the proposed algorithms improve the energy efficiency for the three selected applications. Although E2E delay values are still under the QoS threshold, these values increase considerably compared to the previous evaluation. In some cases, such as the rank-based extremum nodes algorithm for EMR, the E2E delay exceeds the threshold. In contrast to the earlier cases, the highest remaining energy level while keeping the QoS values at the desired level corresponds to the delay-based algorithm (cf. Table 7), where the system is able to conserve energy around 4 J for ECG and EEG applications and 5 J for EMR application. The reason is the simplicity of the procedure of the algorithm, which makes decisions faster than the other algorithms. Finally, the node can harvest more energy in a shorter duration.

TABLE 7. Impact of proposed algorithms on network metrics.

		ECG	EEG	EMR
Legacy Wi-Fi	**Remaining energy (J)**	17.89	18.10	14.68
	E2E Delay (ms)	18.75	7.81	36.01
	PLR (%)	3.9	1.5	32.5
	Objective Function	2.5	15.45	0.125
(a) Three algorithms				
Delay-based	**Remaining energy (J)**	21.1	20.54	17.10
	E2E Delay (ms)	17.89	4.86	21.57
	PLR (%)	<0.05	<0.05	29.82
	Objective Function	3.03	26.66	0.265
Rank-based extremum nodes	**Remaining energy (J)**	**21.05**	**22.11**	17.35
	E2E Delay (ms)	**9.64**	**3.06**	34.46
	PLR (%)	<0.05	<0.05	27.6
	Objective Function	**5.05**	**72.43**	0.182
Rank-based all nodes	**Remaining energy (J)**	21.36	21.04	**18.36**
	E2E Delay (ms)	12.19	4.44	**15.87**
	PLR (%)	<0.05	<0.05	**23.92**
	Objective Function	2.79	52.81	**0.483**
(b) AP coordination in master cells				
Delay-based	**Remaining energy (J)**	20.35	20.54	17
	E2E Delay (ms)	17.11	4.74	29.07
	PLR (%)	<0.05	<0.05	29.4
	Objective Function	6.96	22.90	0.198
Rank-based extremum nodes	**Remaining energy (J)**	**20.85**	**21.04**	17.32
	E2E Delay (ms)	**8.81**	**3.48**	29.07
	PLR (%)	<0.05	<0.05	27.4
	Objective Function	**3.41**	**42.88**	0.217
Rank-based all nodes	**Remaining energy (J)**	20.59	20.79	**17.71**
	E2E Delay (ms)	15.85	4.74	**24.45**
	PLR (%)	<0.05	<0.05	**27.6**
	Objective Function	1.12	32.72	**0.262**
(c) AP coordination in slave cells				
Delay-based	**Remaining energy (J)**	19.57	20.03	16.11
	E2E Delay (ms)	17.75	4.74	21.57
	PLR (%)	<0.05	<0.05	29.82
	Objective Function	6.18	18.22	0.25
Rank-based extremum nodes	**Remaining energy (J)**	**20.83**	**21.03**	17.11
	E2E Delay (ms)	**9.64**	5.23	34.46
	PLR (%)	<0.05	<0.05	27.6
	Objective Function	**5.03**	**63.54**	0.179
Rank-based all nodes	**Remaining energy (J)**	19.82	20.78	**17.56**
	E2E Delay (ms)	12.9	**4.44**	**25.85**
	PLR (%)	<0.05	<0.05	**23.92**
	Objective Function	1.14	25.68	**0.702**
(d) Sleep/Wake-up method				
Delay-based	**Remaining energy (J)**	**21.97**	**22.08**	**19.57**
	E2E Delay (ms)	**38.42**	**27.72**	**87.67**
	PLR (%)	<0.05	<0.05	**9.67**
	Objective Function	1.74	12.53	**0.231**
Rank-based extremum nodes	**Remaining energy (J)**	21.45	20.06	18.1
	E2E Delay (ms)	39.99	12.42	103.99
	PLR (%)	<0.05	<0.05	11.55
	Objective Function	**4.43**	**27.41**	0.150
Rank-based all nodes	**Remaining energy (J)**	20.97	21.55	17.24
	E2E Delay (ms)	34.8	15.34	92.63
	PLR (%)	<0.05	<0.05	33.5
	Objective Function	2.24	13.02	0.055

D. COMPARISON OF LEGACY WITH THE PROPOSED ALGORITHMS FOR AIFSN AND TXOP ADJUSTMENTS

As discussed in Section I, the EDCA mechanism prioritizes channel access through three key parameters: CW, AIFSN, and TXOP, aiming to deliver the necessary QoS for diverse applications. This series of simulations showcases the individual impacts of these parameters on network performance. The algorithms under examination dynamically select CW, AIFSN, and TXOP, each addressed separately, to pursue this goal.

According to the EDCA framework, a lower AIFSN value signifies elevated traffic priority for channel access. However, unsuitable AIFSN adjustments can compromise QoS regarding E2E delay, fairness among stations, and diminishing overall network throughput. Reduced AIFSN intensifies contention dynamics, fostering potential collisions, while

TABLE 8. Impact of MAC parameters changes on the proposed algorithms.

		ECG	EEG	EMR
Legacy Wi-Fi	**Remaining energy (J)**	17.89	18.10	14.68
	E2E Delay (ms)	18.75	7.81	36.01
	PLR (%)	3.9	1.5	32.5
(a) Three algorithms_CW				
Delay-based	**Remaining energy (J)**	21.1	20.54	17.10
	E2E Delay (ms)	17.89	4.86	21.57
	PLR (%)	<0.05	<0.05	29.82
Rank-based extremum nodes	**Remaining energy (J)**	**21.05**	**22.11**	17.35
	E2E Delay (ms)	**9.64**	**3.06**	34.46
	PLR (%)	<0.05	<0.05	27.6
Rank-based all nodes	**Remaining energy (J)**	21.36	21.04	**18.36**
	E2E Delay (ms)	12.19	4.44	**15.87**
	PLR (%)	<0.05	<0.05	**23.92**
(b) Three algorithms_AIFSN				
Delay-based	**Remaining energy (J)**	16.06	17.28	15.66
	E2E Delay (ms)	27.41	24.76	29.61
	PLR (%)	2.5	2.5	31.02
Rank-based extremum nodes	**Remaining energy (J)**	**17.09**	**18.03**	16.85
	E2E Delay (ms)	**24.6**	**23.14**	35.42
	PLR (%)	2.5	2.5	30.5
Rank-based all nodes	**Remaining energy (J)**	16.11	17.27	**17.58**
	E2E Delay (ms)	26.11	25.74	**19.61**
	PLR (%)	2.5	2.5	**28.61**
(c) Three algorithms_TXOP				
Delay-based	**Remaining energy (J)**	15.07	16.28	12.06
	E2E Delay (ms)	28.61	25.93	58.56
	PLR (%)	2.5	2.5	46
Rank-based extremum nodes	**Remaining energy (J)**	**16.07**	**17.07**	12.67
	E2E Delay (ms)	**25.16**	**24.18**	52.42
	PLR (%)	2.5	2.5	45
Rank-based all nodes	**Remaining energy (J)**	15.9	16.27	**13.46**
	E2E Delay (ms)	28.64	27.75	**49.65**
	PLR (%)	2.5	2.5	**40**
(d) Fine-tuning MAC layer parameters				
MAC fine-tuning	**Remaining energy (J)**	23.02	23.8	20.12
	E2E Delay (ms)	11.08	13.25	32.28
	PLR (%)	<0.05	1.2	30.5

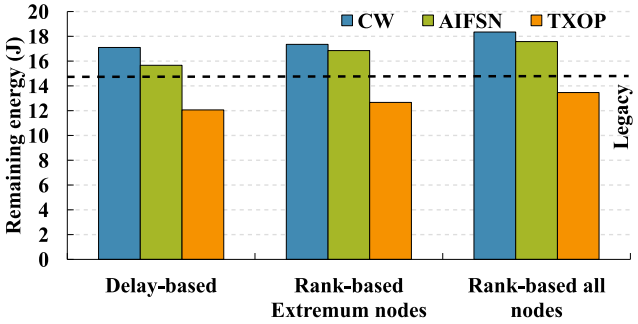


FIGURE 11. Impact of MAC layer dynamic changes on remaining energy for EMR application.

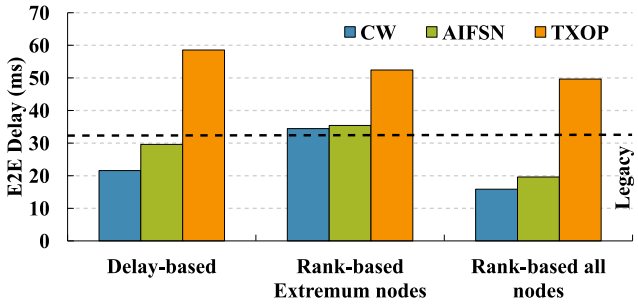


FIGURE 12. Impact of MAC layer dynamic changes on E2E for EMR application.

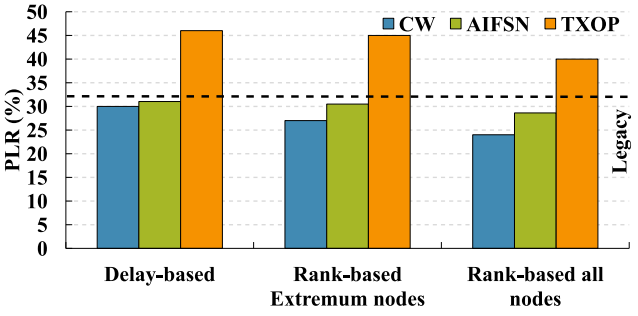


FIGURE 13. Impact of MAC layer dynamic changes on PLR for EMR application.

higher values lead to extended idle periods and channel under-utilization. Thus, as explained in subsection IV-D, it is necessary to set a proper AIFSN value for each individual node. This phenomenon is corroborated by the findings in Table 8, revealing decreased remaining energy due to prolonged idle duration and increased collisions, which is less than 1 J reduction in remaining energy for ECG and EEG applications. In the case of E2E delay, this reduction shows 15 ms for ECG and EEG applications. However, this energy reduction is marginal when compared with legacy values. In the case of the EMR application, a proper AIFSN value conserves the energy of the network around 3 J by reducing the E2E delay for 16.4 ms.

Regarding dynamic TXOP, a station granted a TXOP can transmit frames consecutively without inter-frame channel contention. This enhances throughput efficiency within the station's allocated time window. Nonetheless, this increased throughput for one station can hinder other stations' channel access, resulting in extended idle periods and decreased overall throughput, surpassing the fairness issues attributed to inappropriate TXOP selection. This behavior is depicted in Figure 11 to Figure 13, where dynamic TXOP RL-based algorithms exhibit the lowest remaining energy values, in contrast to the dynamic initialization of CW values, which result in the highest remaining energy values. According to Table 8 the remaining energy values under TXOP RL-based algorithms decrease more than 1 J per application. Here, the legacy value is illustrated as a straight dotted line that indicates the values achieved by the legacy network, in which all devices use the default MAC layer parameters without ML adaptation. These values are listed in Table 4. These outcomes underscore the need for more comprehensive network metrics when making AIFSN and TXOP adjustments. It is important to note that, since the changes are more considerable in the case of the EMR application, we only visualize the impact of the MAC layer parameters in network metrics, and the rest of the metric comparisons can be found in Table 8.

Notably, although TXOP dynamics reduce remaining energy, E2E delay remains within QoS requirements, which means less than 30 ms for ECG and EEG applications and less than 100 ms for EMR application. This implies the efficacy of the proposed algorithms in such scenarios for these three MAC layer parameters. However, these algorithms demand more comprehensive information (more network metrics) for

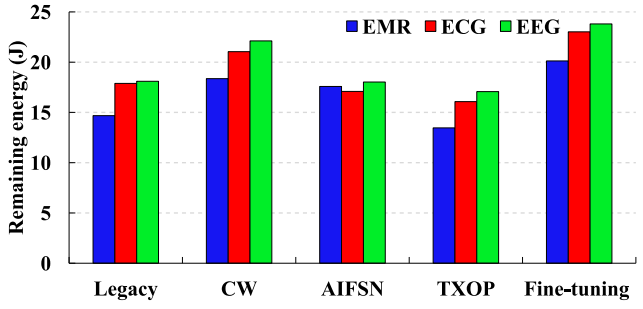


FIGURE 14. Remaining energy actual values comparison per application.

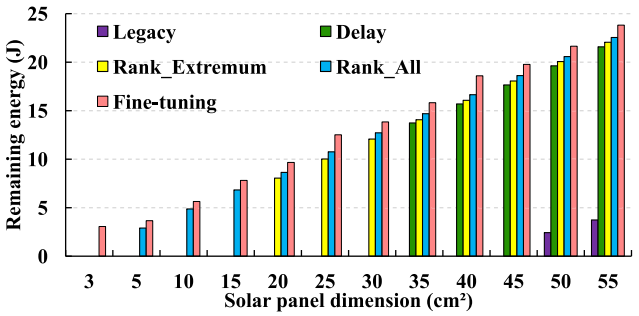


FIGURE 15. Impact of the proposed algorithms on the solar panel form factor.

enhanced decision-making accuracy. Consequently, dynamic CW initiation proves more effective within the current scenario.

Drawing upon the results deriving from the proposed RL-based optimizations at the MAC layer, the finest configuration for individual MAC parameters becomes achievable and tailored to each application's unique demands. Within the simulations, these optimal values are considered when notable enhancements in system performance arise, particularly evident in conserving the network's total energy exceeding 5 J for ECG and EEG while staying below 5 J for EMR scenarios. Nevertheless, the E2E delay increased slightly from the CW adjustment results due to the AIFS duration increase.

To facilitate a more comprehensive comparison of remaining energy among the legacy MAC layer parameters, the proposed RL-based algorithms, and the optimized MAC layer parameters, we visually present the actual remaining energy values per application in Figure 14. As previously noted, the fine-tuned MAC layer parameters outperform other cases through the most energy-efficient configuration.

E. IMPACT OF THE PROPOSED ALGORITHMS ON THE SOLAR CELL FORM FACTOR

According to the literature, since solar cells, piezoelectric, and thermoelectric harvesters provide a more reliable and higher power density level for indoor environments (100 mW/cm^3, 2 W/cm^3, 50 mW/cm^3 for solar cell, piezoelectric, and thermoelectric, respectively), they are widely used in IoT applications, especially in MIoT systems, where the flexibility of the form factor is critical [33]. Thus, in this paper, it is essential to investigate the impact of the proposed algorithm on the form factor of the selected EH technology. The extent of detail included in evaluated simulations could also be used to understand the requirements for the solar panel size. This is important and relevant to the topic under discussion because this study aims to develop a viable system to be implemented in a real-world scenario, such as an e-health environment. As concluded from previous scenarios, in the case of the EMR application, the rank-based algorithm had superiority over other proposed algorithms. Nevertheless, as the proposed RL-based algorithms provide the fine-tuning MAC layer adjustments, utilizing these optimal values significantly influences network performance, consequently leading to a substantial reduction in solar panel dimensions. This conclusion is also demonstrated in Figure 15, which reduces the required solar panel size to 3 cm^2. However, the smallest possible size of the panel in the scenarios with legacy, delay-based, and rank-based extremum nodes are 50 cm^2, 35 cm^2, and 20 cm^2, respectively. For instance, in the case of the legacy, the minimum size of the solar panel needs to be 50 cm^2 to keep powering the MIoT system up. Thus, legacy communications with solar panels smaller than 50 cm^2 are impossible. This comparison supported the advantage of integrating energy harvesting technologies within IoT systems.

Our previous study [35] revealed that the idle state accounts for the highest energy consumption in IEEE 802.11 nodes. This finding aligns with the fact that idle listening can be a significant source of energy drain in wireless systems. The nodes need to continuously monitor the wireless medium for available transmission slots, which consumes energy even when no actual data transmission is taking place. Therefore, longer waiting times and increased idle time would contribute to higher energy consumption by keeping the nodes in an active state for a longer duration. Based on the aforementioned discussion, it is crucial to emphasize that the remaining energy trend for legacy stations does not align with the trend observed in algorithms aimed at increasing the CW to improve the energy consumption of nodes. In other words, the energy consumption patterns of legacy stations do not follow the same trajectory as nodes implementing CW-increasing algorithms in an effort to enhance energy efficiency, and it is important to consider these findings when using the proposed algorithms.

F. DISCUSSION

As explained throughout this article, the fixed initialization of the MAC layer parameters to the default values is unsuitable and inefficient for IoT systems with unpredictable behavior. Our previous studies demonstrated that the initialization of the CW values needs to be adapted optimally to the various applications [35]. In this article, the CW value initialization is updated dynamically based on the network conditions changes to optimize the network's performance and meet

the medical-grade QoS requirements. Generally speaking, the results indicated that the cell-based algorithms perform better than the node-based ones. The reason is that in cell-based algorithms, the AP considers the condition of all the associated nodes to make decisions. In contrast, in the node-wise algorithm, AP makes decisions individually for each associated node regardless of its state, compared to other nodes within the same cell. In addition, it is shown that for applications with high traffic load, the rank-based all-nodes optimization algorithm has superiority over the other proposed algorithms. This is due to the algorithm's high level of flexibility, which converges to the optimal condition very quickly. However, in the case of the low traffic load (ECG and EEG applications), the rank-based extremum nodes optimization algorithm performs better than the other algorithms.

In addition, we demonstrated that adjusting the initialization of the CW values is particularly pronounced, as it directly shapes how stations compete for channel access, consequently influencing collision rates and overall network efficiency. While AIFSN predominantly governs the hierarchy of frame priorities rather than altering the fundamental contention process, and TXOP primarily governs the sequential transmission of multiple frames after channel contention, their effects might carry a different weight than those originating from CW adjustments. Thus, it can be concluded that enhancements to AIFSN and TXOP primarily focus on optimizing and prioritizing frame transmission within an already contented channel. The actual outcomes of optimizing energy utilization, E2E delay, and PLR through RL-driven adjustments within the MAC layer are presented in Table 8. The optimal metrics corresponding to the considered medical applications, achieved via MAC parameter tuning, are distinctly highlighted. It is comprehensible that the dynamic manipulation of AIFSN and TXOP values shows the rank-based all-nodes algorithm to outperform alternative propositions for EMR applications. Conversely, the rank-based extremum nodes exhibit the most productive outcomes regarding ECG and EEG applications. Notably, despite TXOP adjustments falling slightly short of the efficiency gained through CW value optimization, it still satisfactorily fulfills the stringent medical QoS requisites, particularly concerning E2E delay.

Furthermore, we indicate that assuming the AP coordination concept in the proposed algorithms can improve the network's performance. Nevertheless, this improvement is less than the case without applying this concept. For these reasons, we declare that there is no unique initialization for CW values. In addition, introducing the sleep/wake-up mode into our proposals indicates that although the energy efficiency of the network increases, it causes an increment in the E2E delay values, while still maintaining the QoS requirements. It is noted that the most optimal value for evaluated metrics in each set of simulations is highlighted in Table 7 and Table 8, where the assessed PLR values for ECG and EEG applications are negligible. We have demonstrated that the acquired optimal MAC layer parameters yield a discernible enhancement in network performance, with a particular focus on energy consumption. This improvement surpasses the impact of isolated adjustments to individual MAC layer parameters. In the end, we convey that deploying the proposed RL-based optimization algorithms align with fine-tuning of MAC layer parameters making the panel size reduction in IoT systems possible; thus, it takes a crucial step towards achieving sustainability in future IoT systems.

VII. CONCLUSION AND FUTURE WORK

This article presented the possibility of integrating energy harvesting technologies within a Wi-Fi-based MIoT system through extensive simulations on the ns3-gym framework. We proposed three RL-based algorithms for the IEEE 802.11 MAC layer optimization, a novel approach that improves the energy efficiency of the system while provisioning the medical grade QoS requirement. The obtained results demonstrated that in the case of applications with high traffic load, the rank-based all-nodes algorithm can reduce the energy consumption of the system by more than 25%. We also demonstrated that further improvement could be achieved by deploying the sleep/wake-up method to the proposed algorithm, in which the improvement increased to 30%. In addition, the AP coordination concept from the upcoming IEEE 802.11bn amendment was evaluated in this system. We believe this article will shed light on integrating energy harvesting into dense networks such as IoT systems. Moreover, our research highlights that the acquired optimal MAC layer parameters bring about a noticeable enhancement in network performance, particularly regarding energy consumption. This advancement goes beyond the effects achieved by making isolated adjustments to individual MAC layer parameters. The enhancements introduced by the proposed algorithm and the particular fine-tuning of MAC layer parameters demonstrate the feasibility of downsizing solar cells while enhancing their flexibility for integrating this EH technique with IoT devices. These advancements pave the way for integrating them seamlessly into IoT devices, particularly wearable medical devices, leading to a new era of compact, versatile, and energy-efficient technology. This research can be enhanced by introducing a deep learning approach to the proposed algorithms to make the decisions more accurate and flexible to the network changes. Furthermore, we can consider more information (other MAC layer parameters) and different medical grades QoS requirements for the RL-based optimization algorithm.

REFERENCES

[1] Cisco. *Cisco Annual Internet Report (2018–2023) White Paper*. Accessed: Mar. 26, 2021. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/whitepaper-c11-741490.html

[2] S. Nižetić, P. Šolić, D. L.-D.-I. González-de-Artaza, and L. Patrono, "Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future," *J. Cleaner Prod.*, vol. 274, Nov. 2020, Art. no. 122877. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095965262032922X

[3] L. C. Voumik, M. A. Islam, S. Ray, N. Y. M. Yusop, and A. R. Ridzuan, "CO_2 emissions from renewable and non-renewable electricity generation sources in the G7 countries: Static and dynamic panel assessment," *Energies*, vol. 16, no. 3, p. 1044, Jan. 2023.

[4] 2023 European Climate Foundation. *Net-Zero 2050 Project*. Accessed: Aug. 16, 2023. [Online]. Available: https://europeanclimate.org/net-zero-2050/

[5] T. Ahmad and D. Zhang, "Using the Internet of Things in smart energy systems and networks," *Sustain. Cities Soc.*, vol. 68, May 2021, Art. no. 102783.

[6] L. Rodríguez-Jiménez, M. Romero-Martín, T. Spruell, Z. Steley, and J. Gómez-Salgado, "The carbon footprint of healthcare settings: A systematic review," *J. Adv. Nursing*, vol. 79, no. 8, pp. 2830–2844, Aug. 2023.

[7] D. Shin and Y. Hwang, "Integrated acceptance and sustainability evaluation of Internet of Medical Things: A dual-level analysis," *Internet Res.*, vol. 27, no. 5, pp. 1227–1254, Oct. 2017.

[8] P.-P. Pichler, I. S. Jaccard, U. Weisz, and H. Weisz, "International comparison of health care carbon footprints," *Environ. Res. Lett.*, vol. 14, no. 6, Jun. 2019, Art. no. 064004.

[9] E. Udo, L. Oborkhale, and C. Nwaogu, "Analysis and evaluation of energy efficiency of 5G networks in wireless communication," *Arid Zone J. Eng., Technol. Environ.*, vol. 19, no. 2, pp. 175–182, 2023.

[10] K. Pahlavan and P. Krishnamurthy, "Evolution and impact of Wi-Fi technology and applications: A historical perspective," *Int. J. Wireless Inf. Netw.*, vol. 28, no. 1, pp. 3–19, Mar. 2021.

[11] M. S. Afaqui, S. Brown, and R. Farrell, "Detecting MAC misbehavior of IEEE 802.11 devices within ultra dense Wi-Fi networks," in *Proc. 25th Int. Conf. Telecommun. (ICT)*, Jun. 2018, pp. 213–219.

[12] M. S. Afaqui, J. Finnegan, and S. Brown, "Evaluation of HARQ for improved link efficiency within dense IEEE 802.11 networks," *Comput. Commun.*, vol. 191, pp. 217–232, Jul. 2022.

[13] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?" *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.

[14] C. Deng, X. Fang, X. Han, X. Wang, L. Yan, R. He, Y. Long, and Y. Guo, "IEEE 802.11 be Wi-Fi 7: New challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2136–2166, 4th Quart., 2020.

[15] S. Son, K.-J. Park, and E.-C. Park, "Medical-grade channel access and admission control in 802.11e EDCA for healthcare applications," *PLoS ONE*, vol. 11, no. 8, Aug. 2016, Art. no. e0160052.

[16] G. Tian, S. Camtepe, and Y.-C. Tian, "A deadline-constrained 802.11 MAC protocol with QoS differentiation for soft real-time control," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 544–554, Apr. 2016.

[17] I. Syed, S.-H. Shin, B.-H. Roh, and M. Adnan, "Performance improvement of QoS-enabled WLANs using adaptive contention window backoff algorithm," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3260–3270, Dec. 2018.

[18] K. Chemrov, D. Bankov, E. Khorov, and A. Lyakhov, "Smart preliminary channel access to support real-time traffic in Wi-Fi networks," *Future Internet*, vol. 14, no. 10, p. 296, Oct. 2022.

[19] H.-H. Lin, M.-J. Shih, H.-Y. Wei, and R. Vannithamby, "DeepSleep: IEEE 802.11 enhancement for energy-harvesting machine-to-machine communications," *Wireless Netw.*, vol. 21, no. 2, pp. 357–370, Feb. 2015.

[20] Y. Zhao, O. N. C. Yilmaz, and A. Larmo, "Optimizing M2M energy efficiency in IEEE 802.11ah," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2015, pp. 1–6.

[21] M. Gan, D. X. Yang, E. A. AboulMagd, O. M. Montemurro, and S. McCann, *Look Ahead to Next Generation*, document IEEE 802.11-22/0030-01, IEEE 802.11 Wireless Next Generation Standing Committee, Huawei, 2022. Accessed: Sep. 18, 2022. [Online]. Available: https://mentor.ieee.org/802.11/dcn/22/11-22-0030-01-0wng-look-ahead-to-next-generation.pptx

[22] A. Kumar, G. Verma, C. Rao, A. Swami, and S. Segarra, "Adaptive contention window design using deep Q-learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 4950–4954.

[23] Z. Guo, Z. Chen, P. Liu, J. Luo, X. Yang, and X. Sun, "Multi-agent reinforcement learning-based distributed channel access for next generation wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1587–1599, May 2022.

[24] S. Cho, "Reinforcement learning for rate adaptation in CSMA/CA wireless networks," in *Advances in Computer Science and Ubiquitous Computing*. Springer, pp. 175–181, 2021.

[25] Y. Koda, K. Nakashima, K. Yamamoto, T. Nishio, and M. Morikura, "Handover management for mmWave networks with proactive performance prediction using camera images and deep reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 2, pp. 802–816, Jun. 2020.

[26] D. Kotagiri, K. Nihei, and T. Li, "Multi-user distributed spectrum access method for 802.11ax stations," in *Proc. 29th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Aug. 2020, pp. 1–2.

[27] Y. Luo and K.-W. Chin, "Learning to bond in dense WLANs with random traffic demands," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11868–11879, Oct. 2020.

[28] N. N. Krishnan, E. Torkildson, N. B. Mandayam, D. Raychaudhuri, E.-H. Rantala, and K. Doppler, "Optimizing throughput performance in distributed MIMO Wi-Fi networks using deep reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 1, pp. 135–150, Mar. 2020.

[29] M. Carrascosa and B. Bellalta, "Decentralized AP selection using multi-armed bandits: Opportunistic e-Greedy with stickiness," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2019, pp. 1–7.

[30] Z. Han, T. Lei, Z. Lu, X. Wen, W. Zheng, and L. Guo, "Artificial intelligence-based handoff management for dense WLANs: A deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 31688–31701, 2019.

[31] S. de Bast, R. Torrea-Duran, A. Chiumento, S. Pollin, and H. Gacanin, "Deep reinforcement learning for dynamic network slicing in IEEE 802.11 networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2019, pp. 264–269.

[32] R. Ali, N. Shahin, Y. B. Zikria, B.-S. Kim, and S. W. Kim, "Deep reinforcement learning paradigm for performance optimization of channel observation–based MAC protocols in dense WLANs," *IEEE Access*, vol. 7, pp. 3500–3511, 2019.

[33] G. Famitafreshi, M. S. Afaqui, and J. Melià-Seguí, "A comprehensive review on energy harvesting integration in IoT systems from MAC layer perspective: Challenges and opportunities," *Sensors*, vol. 21, no. 9, p. 3097, Apr. 2021.

[34] J. F. Kurose and K. W. Ross, "Computer networking," Top-Down, Tech. Rep., 1986.

[35] G. Famitafreshi, M. S. Afaqui, and J. Melià-Seguí, "Enabling energy harvesting-based Wi-Fi system for an e-health application: A MAC layer perspective," *Sensors*, vol. 22, no. 10, p. 3831, May 2022.

[36] G. Cisotto, E. Casarin, and S. Tomasin, "Requirements and enablers of advanced healthcare services over future cellular systems," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 76–81, Mar. 2020.

[37] S. Shukla, M. F. Hassan, M. K. Khan, L. T. Jung, and A. Awang, "An analytical model to minimize the latency in healthcare Internet-of-Things in fog computing environment," *PLoS ONE*, vol. 14, no. 11, Nov. 2019, Art. no. e0224934.

[38] E. Ibarra, A. Antonopoulos, E. Kartsakli, J. J. P. C. Rodrigues, and C. Verikoukis, "QoS-aware energy management in body sensor nodes powered by human energy harvesting," *IEEE Sensors J.*, vol. 16, no. 2, pp. 542–549, Jan. 2016.

[39] G. Kokkonis, K. Psannis, E. Kostas, M. Roumeliotis, Y. Ishibashi, B. G. Kim, and A. G. Constantinides, "Transferring wireless high update rate supermedia streams over IoT," in *New Advances in the Internet of Things*. Springer, 2018, pp. 93–103.

[40] C. Chakraborty, B. Gupta, and S. K. Ghosh, "A review on telemedicine-based WBAN framework for patient monitoring," *Telemed. e-Health*, vol. 19, no. 8, pp. 619–626, Aug. 2013.

[41] L. A. Al-Tarawneh, "Medical grade QoS improvement using IEEE802.11e WLAN protocol," in *Smart Technologies and Innovation for a Sustainable Future*. Springer, 2019, pp. 229–235.

[42] *Wireless LAN Working Group (C/LAN/MAN/802.11 WG), IEEE 802.11-23/0480r0: UHR Proposed PAR*. Accessed: May 20, 2023. [Online]. Available: https://mentor.ieee.org/802.11/dcn/23/11-23-0480-00-0uhr-uhr-proposed-par.pdf

[43] E. Reshef and C. Cordeiro, "Future directions for Wi-Fi 8 and beyond," *IEEE Commun. Mag.*, vol. 60, no. 10, pp. 50–55, Oct. 2022.

[44] A. Garcia-Rodriguez, D. López-Pérez, L. Galati-Giordano, and G. Geraci, "IEEE 802.11be: Wi-Fi 7 strikes back," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 102–108, Apr. 2021.

[45] *IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 1: Radio Resource Measurement of Wireless LANs*, IEEE Standard 802.11, IEEE Computer Society LAN/MAN Standards Committee, 2008.

[46] C. Farquhar, S. Kafle, K. Hamedani, A. Jagannath, and J. Jagannath, "Marconi–Rosenblatt framework for intelligent networks (MR-iNet Gym): For rapid design and implementation of distributed multi-agent reinforcement learning solutions for wireless networks," *Comput. Netw.*, vol. 222, Feb. 2023, Art. no. 109489.
[47] P. Gawłowicz and A. Zubow, "ns-3 meets OpenAI gym: The playground for machine learning in networking research," in *Proc. 22nd Int. ACM Conf. Modeling, Anal. Simulation Wireless Mobile Syst.*, Nov. 2019, pp. 113–120.
[48] V. Deep, M. L. Wymore, A. A. Aurandt, V. Narayanan, S. Fu, H. Duwe, and D. Qiao, "Experimental study of lifecycle management protocols for batteryless intermittent communication," in *Proc. IEEE 18th Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, Oct. 2021, pp. 355–363.
[49] L. Campanile, M. Gribaudo, M. Iacono, F. Marulli, and M. Mastroianni, "Computer network simulation with ns-3: A systematic literature review," *Electronics*, vol. 9, no. 2, p. 272, Feb. 2020.
[50] ns-3 Network Simulator. *What is ns-3?* Accessed: Aug. 16, 2023. [Online]. Available: https://www.nsnam.org/
[51] K. S. Adu-Manu, N. Adam, C. Tapparello, H. Ayatollahi, and W. Heinzelman, "Energy-harvesting wireless sensor networks (EH-WSNs): A review," *ACM Trans. Sensor Netw.*, vol. 14, no. 2, pp. 1–50, May 2018.
[52] G. Benigno, O. Briante, and G. Ruggeri, "A sun energy harvester model for the network simulator 3 (ns-3)," in *Proc. 12th Annu. IEEE Int. Conf. Sens., Commun., Netw. Workshops (SECON Workshops)*, Jun. 2015, pp. 1–6.
[53] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "OpenAI gym," 2016, *arXiv:1606.01540*.
[54] W. Feng, "Performance analysis of IEEE802.11e EDCA wireless networks under finite load," *Wireless Netw.*, vol. 26, pp. 4431–4457, May 2020.
[55] International Telecommunication Union. *Propagation Data and Prediction Methods for the Planning of Indoor Radio Communication Systems and Radio Local Area Networks in the Frequency Range 300 MHz to 100 GHz*. Accessed: Jun. 10, 2023. [Online]. Available: https://www.itu.int/rec/R-REC-P.1238-8-201507-S/en
[56] M. R. Yuce and J. Khan, *Wireless Body Area Networks: Technology, Implementation, and Applications*. Boca Raton, FL, USA: CRC Press, 2011.
[57] X. Liang and I. Balasingham, "Performance analysis of the IEEE 802.15.4 based ECG monitoring network," *Proc. 7th IASTED Int. Conf. Wireless Opt. Commun.* Princeton, NJ, USA: Citeseer, 2007, pp. 99–104.
[58] S. Fauvel and R. Ward, "An energy efficient compressed sensing framework for the compression of electroencephalogram signals," *Sensors*, vol. 14, no. 1, pp. 1474–1496, Jan. 2014.
[59] S. Merlin and G. Barriac. *TGax Simulation Scenarios*. Accessed: Aug. 24, 2023. [Online]. Available: https://mentor.ieee.org/802.11/dcn/14/11-14-0980-14-00ax-simulationscenarios.docx
[60] Zacurate Company. *Zacurate 500DL Pro Series Fingertip Pulse Oximeter*. Accessed: Apr. 15, 2022. [Online]. Available: https://zacurate.com
[61] G. Carneiro, P. Fortuna, and M. Ricardo, "FlowMonitor: A network monitoring framework for the network simulator 3 (NS-3)," in *Proc. 4th Int. ICST Conf. Perform. Eval. Method. Tools*, 2009, pp. 1–10.

GOLSHAN FAMITAFRESHI received the B.S. degree in electrical and electronics engineering and the M.Sc. degree in electrical engineering from The University of Texas–Pan American, USA, in 2011 and 2015, respectively, and the M.Sc. degree in applied telecommunication and engineering management from Universitat Politècnica de Catalunya (UPC), in 2018. She is currently pursuing the Ph.D. degree in network and information technologies with Universitat Oberta de Catalunya (UOC). Her research interests include energy harvesting technologies integration in wireless communication networks and the Internet of Things, applying machine learning-based optimization in wireless communication networks with a focus on Wi-Fi and 5G and beyond technologies.

M. SHAHWAIZ AFAQUI (Member, IEEE) received the master's degree in electrical engineering (wireless systems) from the KTH Royal Institute of Technology, in 2009, and the Ph.D. degree in networks engineering from the Technical University of Catalonia (UPC), in 2017. He has been a Research Fellow with the Connect Centre Ireland, Maynooth University, from 2016 to 2018, and the Wireless Network Group (Wine), Universitat Oberta de Catalunya, from 2018 to 2021. He is currently a Senior Lecturer in electrical and electronics engineering with the School of Computing, Engineering and Digital Technologies, Teesside University, Middlesbrough, U.K. His research interests include radio resource optimization of WLANs (reliability, frequency, energy, interference management, and energy harvesting solutions support, among others), cross-layer optimization, network security, 3GPP LTE-WLAN aggregation, the Internet of Things (IoT) enabling technologies, and other aspects of 5G.

JOAN MELIÀ-SEGUÍ (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunication engineering from Universitat Politècnica de Catalunya and the Ph.D. degree from Universitat Oberta de Catalunya (UOC). He has been a Researcher with Universitat Pompeu Fabra, a Visiting Researcher with the Palo Alto Research Centre (Xerox PARC), and a Fulbright Visiting Scholar with the Massachusetts Institute of Technology. He is currently an Associate Professor with the Faculty of Computer Science, Multimedia and Engineering, UOC. He has authored over 40 technical peer-reviewed publications. His research interests include low-cost RF-based sensing, the Internet of Things, and sustainability.