ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/394324363

Dynamic Slot Allocation in Wireless Body Area Networks: Exploring Q-learning Approaches

Article · August 2025

DOI: 10.12720/jcm.20.4.457-474

CITATIONS
0

READS
39

9 authors, including:

Abdu Ibrahim Adamu
Audu Bako College of Agriculture, Dambatta, Kano, Nigeria
6 PUBLICATIONS **24** CITATIONS

SEE PROFILE

Darmawaty mohd ali
MARA University of Technology
81 PUBLICATIONS **391** CITATIONS

SEE PROFILE

Wan Haszerila Wan Hassan
Technical University of Malaysia Malacca
17 PUBLICATIONS **80** CITATIONS

SEE PROFILE

All content following this page was uploaded by Abdu Ibrahim Adamu on 11 August 2025.

The user has requested enhancement of the downloaded file.

Dynamic Slot Allocation in Wireless Body Area Networks: Exploring Q-learning Approaches

Abdu I. Adamu [1], Darmawaty M. Ali [1, *], Saidatul Izyanie B. Kamarudin [2], Suzi S. Sarnin [1], Wan Haszerila W. Hassan [1], Mansir Abubakar [2], and Alwatben B. Rashed [3]

[1] Wireless Communication Technology Group (WiCOT), Faculty of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UiTM), Selangor, Malaysia

[2] Faculty of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM), Selangor, Malaysia

[3] Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

Email: 2023402878@student.uitm.edu.my (A.I.A.); darma504@uitm.edu.my (D.M.A.); saidatulizyanie@uitm.edu.my (S.I.B.K.); suzis045@uitm.edu.my (S.S.S.); 2022603854@student.uitm.edu.my (W.H.W.H.); mansir@uitm.edu.my (M.A.); batoolawtban@gmail.com (A.B.R.)

*Corresponding author

***Abstract*—Real-time monitoring through wearable and implanted devices is made possible by Wireless Body Area Networks (WBANs), which have emerged as a key component of modern healthcare. These networks provide substantial advantages for patient treatment by enabling ongoing health data collection. The requirement for fast throughput, low packet delay, Packet Delivery Ratio (PDR), and energy efficiency under dynamic network conditions makes creating a Medium Access Control (MAC) protocol crucial. A Q-Learning-Based MAC Protocol (QL-MAC) designed for slot allocation in WBANs is proposed in this paper. QL-MAC improves network performance across important metrics by dynamically optimizing slot allocation using Reinforcement Learning (RL). By adjusting to different network densities and traffic patterns, the protocol guarantees steady gains in communication. QL-MAC Outperforms Adaptive MAC (ADT-MAC), Dynamic Medical Traffic Management MAC (DMTM-MAC), Traffic Aware MAC (TA-MAC), Multi-Constraints MAC (McMAC), and IEEE 802.15.6 MAC protocols. Experimental results show that QL-MAC achieves higher throughput, reduces latency, maintains a better PDR, and has lower energy consumption, even as network density increases. The benefits of QL-MAC make it especially appropriate for applications where reliable communication and energy efficiency are critical, like chronic disease management and remote patient monitoring. This study also reaffirms the role of machine learning in optimizing communication protocols for next-generation healthcare systems. The results highlight the potential of RL-based approaches to address the unique challenges of WBANs, such as dynamic channel conditions and resource contention. QL-MAC ensures dependable and energy-efficient communication by intelligently managing slot allocation, opening the way for advanced healthcare applications.**

***Keywords*—Wireless Body Area Networks (WBANs), medium access control, slot allocation, modern healthcare, energy efficiency, reinforcement learning, Q-learning**

Manuscript received March 12, 2025; revised April 20, 2025, accepted May 12, 2025; published August 1, 2025.

I. INTRODUCTION

Wireless Body Area Networks (WBANs) are specialized wireless sensor networks that improve healthcare efficiency by allowing real-time health monitoring and quick medical treatments. They use sensors implanted on or inside the human body to capture physiological data such as heart rate, temperature, and glucose levels. These are then transferred to a central node known as the coordinator. WBANs are gaining popularity in healthcare research due to their suitability for applications such as remote patient monitoring, fall detection, and chronic illness management [1, 2].

WBAN development heavily focuses on tiny, low-power sensors and wearable devices, such as smartwatches and fitness trackers, that smoothly interface with the WBAN framework. These devices continuously monitor vital signs while resolving power consumption issues, as WBAN systems have limited battery power [3, 4] Consequently, significant research areas are energy-efficient communication protocols, sensor designs, and power management systems.

Reinforcement Learning (RL) [5−7] RL, a subset of Machine Learning (ML), plays a significant role in WBAN MAC protocols. It investigates how software agents should act in each environment to maximize some form of cumulative reward. RL is one of the three main ML paradigms that play a significant role in network optimization. As a result, when an agent responds differently in diverse environments, an appropriate acting strategy must be taught. RL is used in such cases [8].

Q-learning is a model-free RL technique that directs an agent's actions based on its current condition [8]. It updates a Q-function initially set to preset values for each action taken and the associated reward. The Q-value is constantly modified to improve decision-making. QL has applications in various disciplines, such as routing protocol policy formulation, resource allocation, secure network design, and power distribution methods.

doi:10.12720/jcm.20.4.457-474

Our research digs into cross-layer Medium Access Control (MAC) protocol designs that interact with the physical and network levels to improve overall network performance and resource allocation. These MAC protocols are intended to efficiently accommodate a wide range of devices coexisting on the same network, including classic computers, smartphones, and IoT devices, as networks become more diversified [9].

Dynamic slot allocation is a solution that dynamically adjusts time slot assignments based on real-time network conditions. However, developing an adaptive solution that combines energy efficiency, Packet Delivery Ratio (PDR), and low latency is complex, particularly in WBANs where energy resources are limited. ML approaches, notably reinforcement learning, have shown promise in addressing dynamic allocation issues. QL, a model-free RL approach, allows for autonomous decision-making in situations with little prior knowledge, making it ideal for the adaptive allocation of slots in WBAN.

This research presents a QL-based approach to dynamic slot allocation in WBANs that optimizes slot assignments by continually learning from network states and situations. The QL agent monitors critical variables like traffic load and energy levels, then adjusts slot allocations to enhance packet delivery and energy efficiency. Our contributions are as follows.

- We use a Q-Learning Technique (QL-MAC) to formulate dynamic slot allocation in WBANs as a reinforcement learning problem. This entails creating appropriate states, behaviors, and reward systems adapted to the difficulties faced by WBANs. We assess QL-MAC's performance using sensitivity analysis and comprehensive simulations, showing that it successfully optimizes essential performance measures.
- Comparing our results to traditional MAC protocols and static slot allocation schemes, we discover significant gains in PDR, throughput, and energy efficiency. These findings highlight Q-Learning's ability to meet the requirements of dynamic, resource-constrained contexts, such as WBANs.
- We show how various learning parameter values, such as (α, γ, and ϵ), affect important performance metrics in a WBAN
- We also examine potential ways to improve QL-MAC and other AI-driven protocols in the future. We focus on overcoming the constraints of high-dimensional state spaces, which make generalization and adaptability difficult. By resolving these problems, QL-MAC's robustness and usefulness in practical situations may be further enhanced, opening the door for more clever and flexible communication protocols.

The remainder of the paper is organized as follows. Section II discusses previous research on slot allocation techniques and ML applications in WBANs. Section III discusses the proposed QL concept and details its implementation. Section IV summarizes the simulation results and performance analysis. Section V ends with a conclusion.

II. Literature Review

WBANs have received a lot of study attention because of their importance in healthcare and wellness monitoring [1, 2, 10].

The efficiency of medium access in WBANs is critical since it directly influences packet delivery, latency, and energy consumption, all of which are required for reliable health monitoring systems. Existing approaches to medium access in WBANs are broadly classed as static, adaptive, and ML-based solutions [11, 12] .

Traditional MAC protocols, such as IEEE 802.15.6's Time Division Multiple Access (TDMA), use static slot allocation, which assigns a set of time slots to each node [13, 14]. While this strategy is straightforward and predictable, it fails to adapt to changing network conditions, such as fluctuating traffic loads and energy levels. In dynamic WBAN systems, static approaches frequently result in inefficient slot use, more significant delays, and more packet loss [15].

Recognizing the limits of static techniques, numerous slot allocation algorithms have been suggested. These techniques dynamically alter time slots in response to network factors such as traffic load or data priority [16]. Examples include approaches that allot slots depending on real-time data demand or energy availability, increasing network efficiency and dependability. However, many slot allocation approaches are based on heuristics and established thresholds, which may limit their flexibility and scalability in complex, heterogeneous WBAN environments [17].

Recently, ML, particularly RL, has been investigated as a potentially viable method for adaptive slot allocation in [18, 19]. RL allows computers to make data-driven judgments in dynamic contexts by learning the best techniques through trial and error [8]. QL, a prominent model-free RL algorithm, has been used to solve a variety of networking challenges, including traffic scheduling and resource allocation [20]. QL's capacity to learn directly from interactions with the environment without the need for a prior model makes it ideal for slot allocation in areas where network circumstances alter unexpectedly [21]. Research on RL for MAC protocols has yielded encouraging results in terms of improving resource allocation, minimizing packet delays, and increasing network longevity [22].

This research fills these gaps by presenting a QL-based slot allocation approach tailored to WBANs. Our method uses important WBAN indicators like traffic load and energy usage to allot slots while dynamically maximizing packet delivery and energy efficiency. In doing so, we expand the application of QL to a WBAN environment, resulting in a scalable system that adjusts to dynamic network conditions without incurring significant computing overhead.

However, a thorough comparison of several MAC protocols created for WBANs is shown in the table. Channel access techniques, IEEE technological standards, supported traffic kinds, performance, methodology, network simulator use, and ML integration are all considered while evaluating these protocols. Priority-

Based Adaptive MAC (PA-MAC) [23], Dynamic Scheduling Based on Sleeping Slots (DSBS) and Dynamic Scheduling Based on Buffer (DSBB) [24], Traffic-Aware MAC (TA-MAC) [25], Dynamic Medical Traffic Management MAC (DMTM-MAC) [26], Multi-constraints MAC (McMAC) [27], Adaptive MAC (ADT-MAC) [28], and the suggested QL-MAC are among the protocols mentioned. Each protocol uses design techniques to maximize communication and meet the requirements of WBAN applications.

To guarantee dependable and effective communication, the protocols use various channel access techniques, including scheduled, polling, and contention-based methods. While scheduled access and contention are the main features of PA-MAC, TA-MAC, and McMAC, other protocols, such as DMTM-MAC and ADT-MAC, also use polling to accommodate a wider variety of traffic needs. Most of these protocols follow IEEE 802.15.4 or IEEE 802.15.6 standards, which are commonly used in WBANs for short-range, low-power communication. While IEEE 802.15.6 provides more specific traffic kinds like periodic, urgent, and priority-based communications, IEEE 802.15.4 supports emergency, on-demand, and regular traffic.

These protocols support a wide range of traffic, reflecting the different needs of WBAN applications. For example, DMTM-MAC and ADT-MAC favor periodic and urgent traffic with varying levels of priority, whereas PA-MAC and TA-MAC handle emergencies and on-demand, regular, and non-medical traffic. To meet the requirements of individual applications, McMAC further divides traffic into kinds (Type 0 to Type 4). Every protocol in the comparison uses approaches to meet its performance objectives and exhibits notable performance (shown by a checkmark). As the table shows, network simulation tools frequently validate various protocols.

By incorporating machine learning into its operations, the suggested QL-MAC protocol sets itself apart and becomes flexible in dynamic WBAN situations. In contrast to other protocols, QL-MAC successfully supports regular and emergency traffic by combining scheduled, polling, and contention access techniques under the IEEE 802.15.6 standard. It stands out as an inventive solution for WBANs due to its application of machine learning techniques, which improve its capacity to optimize resource allocation and adjust to changing network conditions. This thorough analysis highlights the advantages and disadvantages of current MAC protocols while emphasizing QL-MAC's potential to improve WBAN dependability and performance. Hence. Table I. depicts the summary of some MAC protocols and their trend.

TABLE I. SUMMARY OF MAC PROTOCOLS AND THEIR TREND

MAC	Ref.	Channel Access			IEEE Technology	Traffic Type	Networks Performance	Method	Used ML
		Scheduled	Contention	Polling					
PA-MAC	[23]	√	√	–	IEEE 802.15.4	Emergency, on-demand, regular, and non-medical	Improved	Network Simulator	No
DSA-NOVA	[29]	–	√	–	IEEE 802.15.6	UP(0) – UP(7)	Improved	Numerical Analysis	No
DSBS, DSBB	[24]	√	–	–	IEEE 802.15.6	Normal and emergency	Improved	Network Simulator	No
TA-MAC	[25]	√	√	–	IEEE 802.15.4	Emergency, on-demand, regular, and non-medical	Improved	Network Simulator	No
DSS	[30]	√	–	–	IEEE 802.15.6	Normal and emergency	Improved	Network Simulator	No
DMTM-MAC	[26]	√	√	√	IEEE 802.15.6	Periodic, urgent, and on-demand	Improved	Network Simulator	No
TCA	[31]	√	–	–	IEEE 802.15.6	1 and 0	Improved	Network Simulator	No
McMAC	[27]	√	√	√	IEEE 802.15.4	Type 0 – Type 4	Improved	Network Simulator	No
A-MAC	[32]	√	√	–	IEEE 802.15.6	Emergency, periodic, and audio/video	Improved	MATLAB	No
DDSA	[33]	√	–	–	IEEE 802.15.6	Critical and non-critical	Improved	Network Simulator	No
TA-IEEE 802.15.6	[34]	–	√	–	IEEE 802.15.6	UP(0) – UP(7)	Improved	Numerical Analysis	No
ADT-MAC	[28]	√	√	√	IEEE 802.15.6	Emergency Periodic with priority-based (High, medium, low)	Improved	Network Simulator	No
Proposed QL-MAC	–	√	√	√	IEEE 802.15.6	Normal and emergency	Improved	Network Simulator	Yes

III. Materials and Methods

A. Problem Formulation

The problem of dynamic slot allocation in WBANs can be formulated as a finite-horizon Markov Decision Process (MDP) defined by ($\mathcal{S}$, $\mathcal{A}$, $\mathcal{P}$, $\mathscr{R}$, γ) [35].
where:

State space ($\mathcal{S}$) → The state represents the network condition at a given time. The key parameter is the node traffic load. This variable reflects the WBAN's dynamic condition, allowing the algorithm to respond to real-time changes.

Action space ($\mathcal{A}$) → Action denotes allocating or reallocating time slots to nodes. Actions include assigning a specific slot to a specific node and reallocating existing slots in response to traffic needs and priorities.

Transition probability ($\mathcal{P}$) → The probability of transitioning from one state to another given an action. It represents how the system evolves based on slot allocation decisions.

Reward function ($\mathscr{R}$) → The reward function encourages efficient and dependable data transfer. Rewards are given for improving PDRs, reducing latency, and conserving energy. Penalties are imposed if packet drops or delays exceed predefined thresholds, prompting the learning agent to improve performance.

Discount factor (γ) → A factor that determines the importance of future rewards (typically $0 < \gamma \leq 1$).

1) State representation ($\mathcal{S}$)

Each state $s \in \mathcal{S}$ can be defined as:
where:

q_t → Queue length at time t

e_t → Energy level of the node at the time t

c_t → Channel condition at time t

d_t → Data transmission success status at the time t

2) Action representation ($\mathcal{A}$)

Each action $a \in \mathcal{A}$ corresponds to assigning a time slot to a sensor node. Let a_t denote the action taken at the time t, then:

$$a_t = (n, T_s) \tag{1}$$

where:

n → The node selected for transmission

T_s → The allocated time slot duration

3) Objective function: Maximize reward

The objective is to maximize network performance while considering energy efficiency, delay, and reliability.

$$\sum_{t=0}^{T} \gamma^t . \mathcal{R}(s_t, a_t) \tag{2}$$

where:

$\mathcal{R}(s_t, a_t)$ = Reward function based on system performance

γ = Discount factor

T = Total number of time slots

The reward function can be defined as:

$$\mathcal{R}(s_t, a_t) = w_1 . T(s_t, a_t) - w_2 . L(s_t, a_t) - w_3 . E(s_t, a_t) + w_4 . PDR(s_t, a_t) \tag{3}$$

where:

$T(s_t, a_t)$ = Throughput

$L(s_t, a_t)$ = Latency

$E(s_t, a_t)$ = Energy consumption

$PDR(s_t, a_t)$ = Packet Delivery Ratio

w_1, w_2, w_3, w_4 → Weights are assigned to each parameter based on priority.

4) Constraints

a) Slot allocation constraint

Each node must be assigned at most one slot in a frame:

$$\sum_{n=1}^{N} a_{t,n} \leq 1, \ \forall t \tag{4}$$

where N is the total number of sensor nodes.

b) Energy constraint

A node cannot transmit if its energy is below a threshold E_{th}:

$$e_t \geq E_{th}, \ \forall t \tag{5}$$

c) Queue stability constraint

The queue length should not exceed Q_{max}:

$$q_t \leq Q_{max}, \ \forall t \tag{6}$$

d) Latency constraint

The delay for packets must not exceed the maximum acceptable latency L_{max}:

$$L(s_t, a_t) \leq L_{max}, \ \forall t \tag{7}$$

B. Proposed Q-learning Approach for Dynamic Slot Allocation

This paper presents a QL-based technique for enabling dynamic slot allocation in WBANs. This strategy's primary purpose is to dynamically alter time slots to maximize packet delivery, minimize latency, and increase energy economy, which are crucial in WBAN situations where nodes frequently operate under tight resource restrictions.

If the slot of the TDMA MAC is constant. In high-traffic settings, there is a greater likelihood of a collision during the beacon period [36]. Furthermore, there is a substantial propagation delay in low-traffic settings during the beacon time. As such, we propose a QL-MAC protocol for the WBAN network. The slot during the beacon period regulation is consistent with the MDP formulation. QL was utilized to create a MAC protocol that allows nodes in the network to choose the right length of the beacon period based on experience gained from agent-environment interactions. The proposed QL-MAC protocol includes a QL-based approach to dynamically alter time slots to improve system throughput and PDR while minimizing propagation delay and energy consumption. Fig. 1 represents a diagram of the proposed QL-MAC, MDP with agent-environment interaction.

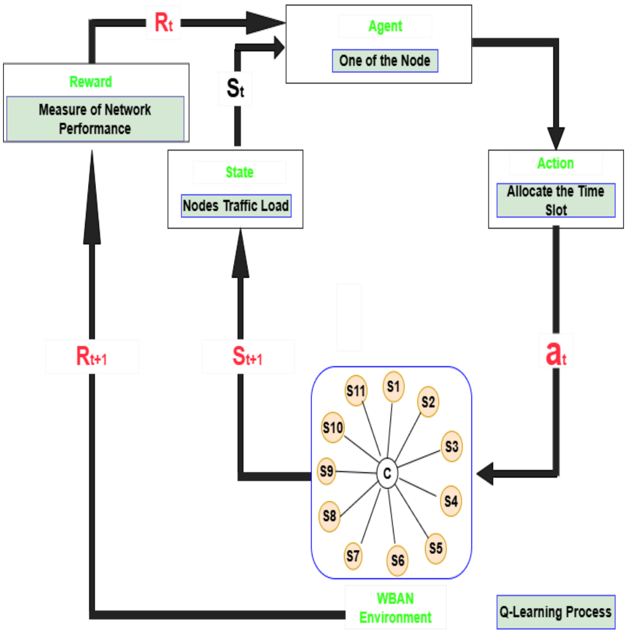

Fig. 1. A diagram of the proposed QL-MAC, markov decision process with agent-environment interaction.

One action is chosen by one environmental state and designated as a policy. If the agent follows policy π at time step t, then $\pi(a|s)$ is the probability of $\mathcal{A}_t = a$, given that $\mathcal{S}_t = s$. If the agent follows the policy π at time t, the probability of $\mathcal{A}_t = a$ under $\mathcal{S}_t = s$ is denoted as $\pi(a|s)$. QL can solve MDPs without complete information by employing an RL technique. A QL system consists of the agent, the environment, a policy, a reward, and a Q-value function. Policy is an action that has been carried out in certain environmental conditions. Generally, a policy is a function or lookup table; it is also the foundation of an RL agent. In Ref. [37], a clear policy is enough to direct an agent's decision-making process in an RL framework, according to the statement "Policy-determined behavior is adequate". A Q-learning agent uses its policy to establish the optimal course of action for slot allocation in WBANs. The policy guarantees that the agent makes the best choices without requiring human involvement because it changes in response to learned experiences [38]. This is especially helpful in WBANs, where adaptation is necessary due to dynamic network factors such as fluctuating traffic loads, energy limits, and channel reliability.

What makes policy-driven behavior adequate is its ability to maximize decision-making, adjust to shifting circumstances, and attain long-term efficiency. After learning a policy, the agent may allocate slots with little computing expense, guaranteeing real-time responsiveness. Furthermore, Q-learning naturally converges with an optimal or nearly optimal approach as it iteratively improves its policy through interactions with the environment [39]. This ensures consistent and efficient slot scheduling, which lowers packet collisions and enhances WBANs' overall Quality of Service (QoS).

Furthermore, stochasticity, where decisions are made based on a probability distribution rather than being completely deterministic, can be incorporated into the policy. This keeps the agent from becoming mired in less-than-ideal behaviors by enabling the investigation of alternate tactics. The policy guarantees that the WBAN achieves robust and adaptable performance by striking a balance between exploration and exploitation, which makes it an appropriate strategy for effective medium access control [38, 39].

The reward signal, a single value supplied by the environment to the RL agent, determines the goal of an RL job. An agent's only goal is to maximize its reward. If a policy selects one action but the reward in the subsequent situation is low, the policy will be updated to select alternative actions in that condition. Unlike the reward function, which provides a signal for a specified time, the Q-value function produces a good signal for the state's end time. As a result, the Q-value indicates a state's cumulative reward, which an agent will use to decide which action to perform in the future [35, 40].

In the WBAN environment, the number of time slots for successful transmission drops while the number of time slots for unsuccessful transmission increases. We calculate the incentive based on the number of successful time slots, energy efficiency, and performance metrics (throughput, latency, PDR, and energy usage). A higher channel traffic probability suggests fewer time slots are available for successful transmission than for failed transmission. As a result, the reward is reduced and may even be negative. In contrast, a reduced chance of channel traffic leads to a higher reward. The Q-value estimate seeks an additional reward, and the agent selects a suitable action depending on the Q-value in equation (8). This implies that the agent in the Q-learning algorithm maintains the $Q(\mathcal{S}_t, a_t)$ table. However, for $t = 1, 2, 3, ...$, the agent will observe the state $\mathcal{S}_t$ of the MDP in the WBANs network and selects an action a_t from the actions $(\mathcal{A})$. After action a_t, the agent receives a reward $\mathcal{R}(t)$ and then observes the next state, $\mathcal{S}_{t+1}$. The sequence of events creates the agent's learning experience. The Q-table will be updated by the sequence of events under the $Q(\mathcal{S}_t, a_t)$ pairs according to the QL function (8). In addition to the immediate reward obtained at the current step, the Q-learning update automatically incorporates an estimate of future rewards. As defined in [41]:

$$Q(\mathcal{S}_t, a_t) = Q(s_t, a_t) + \alpha\,[(\mathcal{R}_{t+1}) + \gamma max Q(\mathcal{S}_{t+1}, a) - Q(s_t, a_t)] \tag{8}$$

Here, $Q(\mathcal{S}_t, a_t)$ represents the current estimate of the reward associated with acting a_t in state $\mathcal{S}_t$. The term $\mathcal{R}_t + \gamma max Q(\mathcal{S}_{t+1}, a)$ accounts for both the immediate reward $\mathcal{R}_t$ and the expected future rewards, discounted by the factor γ. The learning rate α determines how much the new message supersedes the old one.

The optimal range for the discount factor γ is $0 < \gamma \leq 1$. The learning rate α ranges from 0 to 1 [42, 43]. Fig. 7 shows the details.

The agent only evaluates the current reward if the discount factor γ is 0. The agent seeks a long-term reward if the discount factor γ is 1 [43].

The agent chooses an action based on the state S_t and the maximum Q-value for the following state S_{t+1} may be calculated using the action a_t. As in Eq. (9).

The greedy policy is the most widely used tactic, in which the agent selects the course of action that maximizes the Q-value:

$$a_t = \text{argmax}\,(s_t, a) \tag{9}$$

This guarantees that, based on previous experience, the agent constantly chooses activities that will result in the maximum expected return. An ε-greedy method, in which the agent periodically investigates alternative behaviors with a tiny probability ϵ to avoid premature convergence to a suboptimal policy, is frequently employed to strike a balance between exploration and exploitation.

The Q-value function is a key component in the agent's decision-making process. Its findings guide the agent in selecting the most appropriate action by maximizing the Q-value function. This process allows the agent to achieve the largest rewards over a succession of actions without necessarily focusing on a single reward [44]. However, Fig. 2 depicts the State Observation, Action Selection, and Q-Table Update in our proposed QL-MAC following the Bellman Equation.

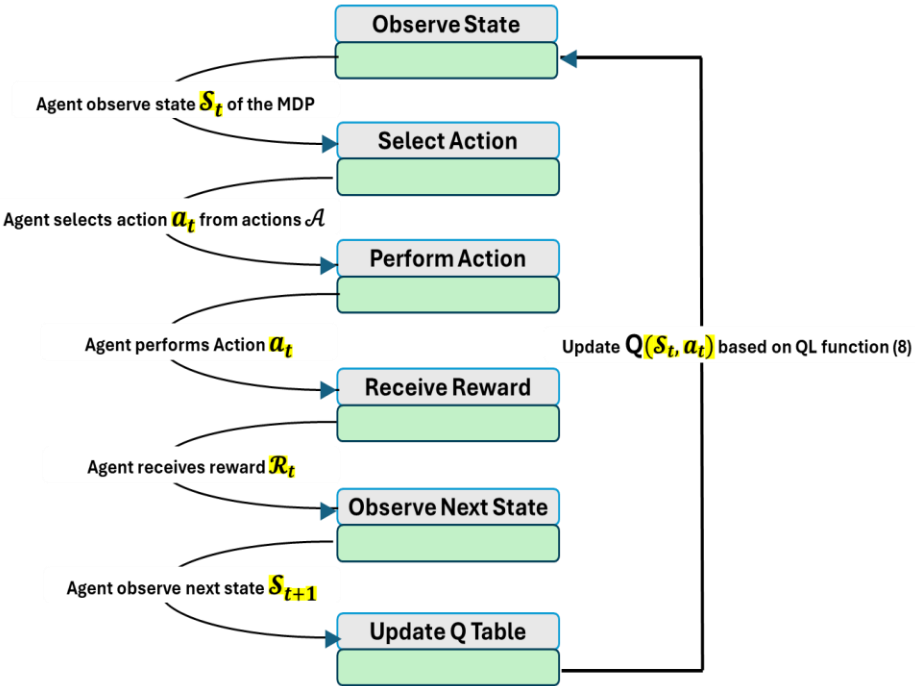

Fig. 2. State observation, action selection, and Q-table update in our proposed QL-MAC following the bellman equation [41].

C. *Performance Metrics*

1) Network throughput

The throughput for a WBAN network is defined as TH, as follows [45, 46]:

$$TH = \frac{\sum D_{successful}}{T_{simulation}} \tag{10}$$

where $D_{successful}$ is the total successfully received data in bits, and the $T_{simulation}$ is the system's total simulation time in (seconds).

2) Energy consumption

The total amount of energy used by a sensor node during its communication time across its operating states is known as its total energy consumption, denoted by E_{total}.

In this research, energy consumption computations can be used for three cases: idling, successful transmission, and collision [47]. Here, we consider only energy consumption during successful transmission and collisions. Assume that all nodes in the WBAN network are powered on.

$$E_{total} = E_{success} + E_{collission} \tag{11}$$

where $E_{success}$ Denote the energy consumption of successful transmission in a WBAN network and $E_{collission}$ Denote the energy consumption during the collision in a TDMA MAC of the WBAN network.

$$Energy\ efficiency = \frac{\sum D_{successful}}{T_{total}} \tag{12}$$

where $D_{successful}$ represent the total successfully received data (bits) and T_{total} represent the total energy consumed (Joules).

3) Packet delay

The average amount of time that passes between a packet being generated at the sensor node and being received at the hub is known as the packet delay of that node. Better performance is indicated by a lower packet latency of the protocol's performance [48, 49]. It is computed as:

$$Delay = \frac{\sum T_{arrival} - T_{sent}}{P_{recieved}} \tag{13}$$

where $T_{arrival}$ is the time packet received T_{sent} The time packet was sent, and $P_{recieved}$ the number of packets successfully received.

D. Packet Delivery Ratio

The PDR is the ratio of the total number of packets received by the receiver node to the number of packets produced by the sender nodes. The PDR assesses network dependability, and higher values indicate better protocol performance. In mathematics, the PDR is represented as:

$$PDR = \frac{\sum P_{Recieved}}{\sum P_{sent}} \times 100\% \quad (14)$$

where the $P_{recieved}$ and P_{sent} represent the total number of packets received and sent, respectively.

E. Computational Complexity of QL-MAC

One important factor influencing the viability of the suggested QL-MAC protocol for real-time deployment in WBANs is its computational complexity. The state-action space exploration and the iterative Q-value table updates are the two main sources of computing overhead in RL [50, 51]. Network parameters, including channel quality, traffic load, and node density, are included in the state space of QL-MAC, whereas slot allocation choices are made in the action space as in Eq. (3).

In that case, α is the learning rate, γ is the discount factor, R_t is the reward received at time t, *St* is the current state, and $\mathcal{A}_t$ is the action executed. This update rule has an overall complexity of $O(|\mathcal{S}|.|\mathcal{A}|)$, which is proportional to the size of the state space ($|\mathcal{S}|$) and the action space ($|\mathcal{A}|$).

To maintain QL-MAC's computational efficiency, we used many optimization strategies. First, we used state abstraction, which organizes related states into higher-level representations, to minimize the dimensionality of the state space. This method lessens the computational load by drastically lowering the number of distinct states the program needs to handle. Second, we employed a lookup table to store Q-values to facilitate constant-time access and runtime modifications [50]. Because of these design decisions, QL-MAC has a far lower computational cost than conventional MAC protocols like IEEE 802.15.6, which depend on rigid scheduling algorithms.

It is crucial to recognize that, despite these optimizations, QL-MAC does have certain computing costs because of its adaptive nature. In particular, the ε-greedy strategy's exploration-exploitation trade-off necessitates a regular assessment of several actions to determine the best slot allocation method [51]. However, the considerable performance improvements in throughput, energy economy, and delay reduction make this overhead acceptable. Additionally, the lightweight RL structure of QL-MAC guarantees compatibility with resource-constrained WBAN devices, which are usually outfitted with low-power microcontrollers.

F. Addressing the Convergence Rate of Q-learning

A crucial component of the suggested QL-MAC protocol is the Q-learning convergence rate, which directly affects the protocol's capacity to adjust to changing network conditions in WBANs. Study in Ref. [52]. The convergence of Q-learning in RL is contingent upon several critical factors, such as the discount factor (γ), exploration rate (ε), and learning rate (α). These parameters determine the algorithm's speed and efficiency in learning the best slot allocation techniques. We used a decaying ε-greedy approach for QL-MAC to balance exploitation and exploration during the learning process. As the system learns more about the environment, this method eventually converges to an exploitation phase, ensuring that the algorithm sufficiently explores the state-action space in the early stages. Our simulations' empirical results show that QL-MAC stabilizes performance in a manageable number of iterations, usually needing 400–500 episodes to converge under various network conditions. In environments with limited states and finite actions, this convergence pattern is consistent with the theoretical predictions of Q-learning [53].

We conducted further tests to assess the effects of various learning parameters on convergence and overall performance to shed more light on how sensitive QL-MAC is to these variations. Fig. 7 depicts the details. The rate at which the Q-values are updated during training is largely determined by the learning rate (α). Higher learning rates speed up convergence, but if the updates are too frequent, they could result in less-than-ideal policies [52, 54] . Using methodical testing, we discovered that a learning rate of α = 0.1 achieves the best possible balance between stability and convergence speed [43]. Likewise, during 500 episodes, the exploration rate (ε) decreased exponentially from a starting value of 0.1 to a final value of 0.01 [55]. This exponential decay allows the algorithm to eventually take advantage of learned policies, guaranteeing adequate exploration in the early training phases. To prioritize long-term rewards, which are especially crucial in WBANs since decisions taken simultaneously can have a lasting impact on network performance, the discount factor (γ) was finally adjusted to 0.9. These parameter selections guarantee effective QL-MAC convergence while preserving resilience to dynamic shifts in the network environment. However, the purpose of using the parameters is depicted in Table II.

G. Simulation Environment for the Proposed Model

We run simulations of the QL algorithm in a WBAN setting, investigating different node layouts and traffic patterns. The network is initially configured with a default slot allocation, which the QL agent adjusts based on feedback from a reward function. For performance testing, baseline protocols such as classic ADT-MAC, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC are included, allowing for a thorough evaluation of the QL approach's effectiveness.

The suggested approach enables WBANs to dynamically respond to network conditions, resulting in superior PDR and increased energy efficiency compared to standard protocols.

TABLE II. PURPOSE FOR USING PARAMETERS

Parameters	Purpose	Selection	Why Selected	Ref.
Discount Factor (γ)	Balances the importance of immediate rewards versus long-term rewards	γ = 0.9	When the discount factor is near 1, the agent can consider future benefits without disregarding present rewards. This is crucial in WBANs, where choices could have long-term effects on resource usage.	[42], [43], [55], [56]
Learning Rate (α)	Controls the step size in updating the Q-table, determining how much the agent learns from new experiences.	α = 0.1	In WBAN circumstances, where dependable learning is critical, a lower learning rate prevents instability by guaranteeing steady and progressive updates to the Q-values.	[43]
Exploration Rate (ε)	Balances exploration (trying new actions) and exploitation (choosing the best-known action).	ε = 0.1	A low exploration rate indicates that the agent occasionally explores but mostly takes advantage of established good behaviors. This is appropriate for WBANs, where too much exploration could impair the dependability of patient monitoring or network performance.	[55]
Simulation Time	Ensures sufficient runtime for the agent to converge towards optimal policies.	200−300 seconds	This duration is sufficient in WBAN simulations to observe steady-state performance, evaluate the learning algorithm's efficacy, and guarantee that the dynamics of the network environment are adequately represented.	[42]

TABLE III. OUR QL-BASED MAC PARAMETERS

Parameters	Values
Slot length	10ms
Simulation Time	300s
Number of Nodes	1 Central Hub and 11
Frequency Band	2.5GHz
Payload	105 Bytes
Transmissions Rate	1,024 Kbps
Superframe Length	32
Transmission Power	0.01J
Learning Rate (α)	0.1
Discount Factor (γ)	0.9
Exploration rate	0.1
Number of Episodes	400-500

To assess critical performance parameters such as PDR, packet delay, network throughput, and energy consumption, we used Castalia and OmNet++ for our simulation and Python libraries such as OpenAI Gym to create the model [57]. Castalia was chosen because it is a discrete event simulator for WBANs and other complicated systems [58]. The simulation results were compared to benchmark protocols under identical network conditions to guarantee a rigorous and unbiased evaluation. For IEEE 802.15.6-MAC, the superframe structure used a fixed slot size, including a beacon, five Exclusive Access Phase (EAP) slots for emergency traffic, five Random Access Phase (RAP) spaces for connection creation, and 15 TDMA-MAP slots for periodic traffic transmission.

Our simulations used a WBAN with a central hub and 11 sensor nodes. Each sensor node sent a 105-byte data packet to the hub. The simulation parameters for the QL-MAC protocol were configured with 32 slots in each beacon period, each with a duration of 10ms, which is like previous studies [59−61]. The transmission power is −10 dBm. A 2.4 GHz Body Area Network (BAN) radio chipset enabled 1,024 Kbps data transmission, ensuring reliable wireless connectivity. Additionally, the discount factor (γ) is set to 0.9. The learning rate α is set to 0.1. The Exploration rate is set at 0.1. The simulation was run for 300 seconds, with three repetitions, to reduce unpredictability and temporal aberrations in packet transfers at the application layer. Tables II and III show the details of the parameters of our proposed QL-based MAC protocol for WBAN networks. and the reason for selecting a range of values for them.

IV. RESULTS AND DISCUSSION

The performance of the Q-Learning MAC (QL-MAC) protocol is compared to the International Standard MAC for WBAN (IEEE 802.15.6-MAC) [62]. Dynamic Medical Traffic Management MAC (DMTM-MAC) [59]. Multi-Constraints MAC (McMAC) [27]. Traffic-Aware MAC (TA-MAC) [23]. and Adaptive Traffic MAC (ADT-MAC) [28]. These protocols were chosen for comparison because they share conceptual parallels with the key characteristics of the proposed approach. DMTM-MAC is a new protocol that prioritizes emergency traffic and uses a dynamic superframe structure based on traffic demand, like the traffic-adaptive approach of ADT-MAC. McMAC, TA-MAC, and ADT-MAC prioritize traffic mitigation to improve network performance, a key component of QL-MAC. IEEE 802.15.6-MAC is a benchmark since the QL-MAC protocol is based on its superframe structure. Furthermore, the unresolved issues with fixed slot allocation in the IEEE 802.15.6 standard for WBAN communication necessitate their inclusion in the performance evaluation of the proposed protocol.

A. Network Throughput vs Number of Nodes

Fig. 3 shows the network throughput (in kbps) as a function of the number of nodes for six MAC protocols: QL-MAC, ADT-MAC, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC. It provides insights into the scalability and efficiency of these protocols in dealing with varying network densities, particularly in WBAN scenarios.

Fig. 3 also shows how throughput, an important statistic for evaluating network performance, grows with the number of nodes when utilizing the QL-MAC protocol. As the number of nodes increases from two to twelve, throughput progressively increases from around 100 kbps to slightly more than 140 kbps. This rising trend demonstrates that the QL-MAC protocol can easily change its slot allocation technique to accommodate more nodes without significantly reducing performance. This behavior is consistent with the strengths of RL, as QL allows the protocol to learn and adjust communication patterns based on network conditions, resulting in better utilization of available slots and higher throughput.

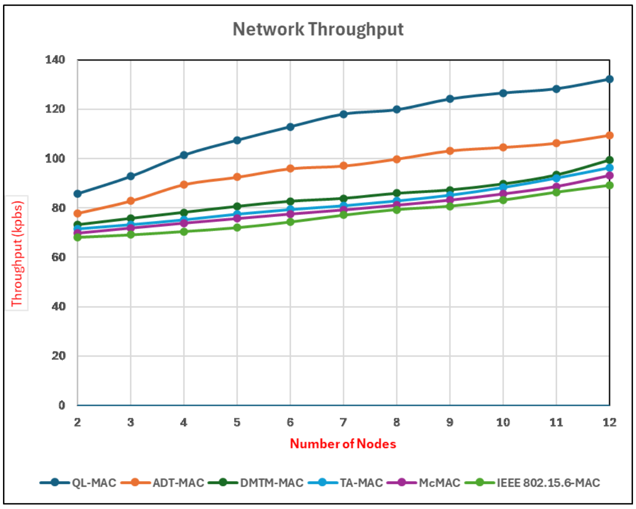

Fig. 3. Network throughput.

This stability trend as nodes rise is significant in situations like WBANs, where devices must work efficiently in dynamic and dense configurations. Traditional adaptive slot allocation algorithms may struggle to sustain high throughput in such settings because they lack the agility to respond rapidly to changing traffic needs. In contrast, QL-MAC, which uses QL, adapts dynamically, making it a promising technique for retaining high performance while scaling networks. This versatility may benefit applications requiring real-time data transmission and reliability, such as health monitoring or emergency response systems.

Furthermore, Fig. 4 indicates that the QL-MAC protocol might manage even bigger networks with minimal performance deterioration. However, more testing is required to prove this idea. The progressive increase in throughput, rather than a dramatic surge or plateau, suggests that the protocol's effectiveness scales steadily with the addition of nodes. This consistency in performance illustrates the protocol's durability and may make it a viable choice for various wireless networks requiring scalability and steady throughput.

In comparison, QL-MAC is the standout performer, achieving much greater throughput than all other protocols at all node densities. This performance demonstrates its strong capacity to dynamically distribute slots with QL, an adaptive reinforcement learning method. Its throughput gradually improves with the number of nodes, displaying good scalability and flexibility to high network traffic. The constant rise also shows efficient resource usage and minimum packet collisions, which are significant issues in WBANs.

ADT-MAC is ranked second in performance, with throughput gradually improving as the number of nodes grows. However, the gap between ADT-MAC and QL-MAC widens when node density increases. This could reflect ADT-MAC's limits in adjusting to increasingly congested networks despite using some adaptive methods.

The remaining protocols, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC, exhibit comparable and lower throughput values. Their flatter curves indicate that these protocols use static or semi-static slot allocation algorithms, which are less adaptable to dynamic network changes. The IEEE 802.15.6-MAC protocol, built explicitly for WBANs, lags, indicating inadequate slot allocation in crowded network conditions.

Fig. 3 also illustrates scalability differences. While QL-MAC thrives in high-density networks by delivering constant performance gains, static or heuristic-based protocols encounter bottlenecks. This demonstrates the advantages of machine learning-based solutions, which can adjust to changing traffic demands and environmental conditions.

In addition, the diagram highlights QL-MAC's promise in WBAN applications by demonstrating its ability to optimize throughput while effectively managing network congestion. The findings call for incorporating adaptive and learning-based approaches, such as QL, into future WBAN MAC protocol designs to improve scalability and dependability.

B. Energy Consumption vs Number of Nodes

Fig. 4 shows the energy consumption findings for several MAC protocols, including QL-MAC, ADT-MAC, TA-MAC, McMAC, DMTM-MAC, and IEEE 802.15.6-MAC, as the number of nodes grows. The QL-MAC protocol consistently has the lowest energy consumption across all network densities, confirming its efficacy in resource-constrained WBAN settings. This is due to the adaptive slot allocation technique QL-MAC uses, which optimizes communication scheduling and lowers wasteful energy use during data transmission.

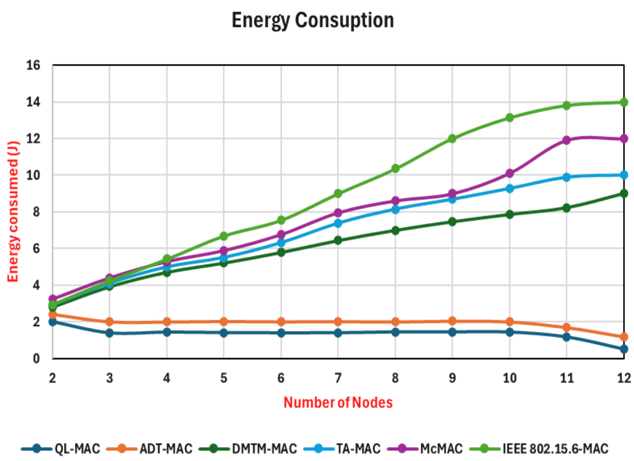

Fig. 4. Energy consumption vs number of nodes.

Fig. 4 shows the energy usage of a QL-MAC protocol as the number of nodes in the network grows. The y-axis displays energy consumption in joules (J), and the x-axis depicts the number of nodes ranging from 2 to 12. The graph shows that energy usage decreases as the number of nodes increases, beginning at around 2.0 J for two nodes and below 0.5 J for twelve. As the network grows from two to four nodes, energy consumption decreases dramatically, suggesting that the QL-MAC protocol quickly optimizes energy efficiency in small networks. This first drop indicates that adding a few nodes improves slot usage and reduces idle or redundant transmissions, resulting in lower total energy consumption.

As the network size increases from 4 to 10 nodes, energy consumption stabilizes at 1.5 J, indicating that the protocol maintains a balance in slot distribution while not increasing energy demands. Interestingly, as the network grows to 10 or more nodes, energy consumption drops further, significantly declining as the number of nodes approaches 12. This pattern indicates that the QL-MAC protocol effectively adapts to higher network sizes by optimizing communication slots more efficiently through learned behaviors that reduce energy consumption. This behavior is advantageous in resource-constrained contexts, such as WBANs, where energy saving is essential for device longevity and continuous service.

The IEEE 802.15.6-MAC and DMTM-MAC protocols consume much more energy, particularly as the network density increases. This increase is primarily due to their reliance on static allocation techniques or inefficient scheduling, which results in increased energy waste from collisions, retransmissions, and idle listening. McMAC, while marginally more efficient than DMTM-MAC, experiences similar difficulties as the number of nodes grows.

The TA-MAC and ADT-MAC protocols outperform McMAC and DMTM-MAC while retaining modest energy consumption levels. However, when the number of nodes increases, their energy efficiency decreases because they lack the robust adaptability that QL-MAC provides. QL-MAC's RL methodology allows dynamic slot management, which reduces idle times and ensures energy-efficient operation even in high-density scenarios.

C. Packet Delay vs Number of Nodes

Fig. 5 shows the packet delay (in milliseconds) as a function of the number of nodes for six different MAC protocols: QL-MAC, ADT-MAC, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC. The results provide insight into the protocols' efficiency in handling delays under increasing network loads, an important performance metric in WBANs.

Fig. 5 depicts the packet delay in a QL-MAC protocol as the number of nodes in the network rises. The y-axis plots packet delay in milliseconds (ms), while the x-axis indicates the number of nodes, ranging from 2 to 12. The graph shows that packet latency increases as the number of nodes increases, beginning around 10ms with two nodes and reaching around 24 ms with twelve nodes. This rise in delay is expected as more nodes compete for available communication slots, resulting in more significant congestion and waiting times for each node to provide data.

The constant delay growth illustrates a typical issue in wireless networks, where increased node density can cause queuing and transmission delays. While QL-MAC effectively manages delays in smaller networks, the progressive increase shows that maintaining low latency may become difficult as the network grows. Despite this, the QL-based strategy still provides benefits by dynamically modifying slot assignments and reducing time compared to previous non-learning systems. In environments like WBANs, a tolerable increase in delay may be acceptable if improvements in other measures, such as throughput and energy efficiency, offset it.

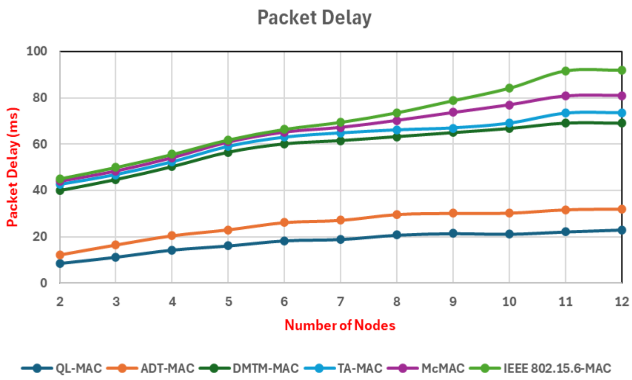

Fig. 5. Packet delay.

QL-MAC consistently delivers the shortest packet delay across all node counts, indicating its excellence in reducing transmission latency. The time remains constant even as network density grows, demonstrating the protocol's efficiency in dynamically distributing slots via QL. QL-MAC's minimal and consistent delay makes it ideal for time-sensitive WBAN applications such as healthcare monitoring, where real-time data transfer is critical.

ADT-MAC has larger delays than QL-MAC but still beats the other protocols. Its progressive increase in delay with node count suggests a decently adaptive process, albeit one with limits in dealing with extreme network congestion.

The other protocols, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC, have much larger latency, which rises sharply as network capacity increases. IEEE 802.15.6-MAC has the worst delay performance, reaching 90ms at high node counts. This demonstrates its failure to adapt effectively to increasing traffic situations due to inflexible slot distribution and increased contention in a more extensive network.

The protocols' scalability is clear from their performance trends. While QL-MAC and ADT-MAC scale well with the number of nodes, the other protocols suffer from increased network density, resulting in delays that may exceed tolerable levels for WBAN applications. This gap emphasizes the importance of adaptive and learning-based approaches to achieve low-latency communication.

D. Packet Delivery Ratio vs Number of Nodes

Fig. 6 shows the PDR as a function of the number of nodes for six MAC protocols: QL-MAC, ADT-MAC, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC. PDR, expressed as a percentage, represents the ratio of successfully delivered packets to total packets sent and is an important indicator of network reliability.

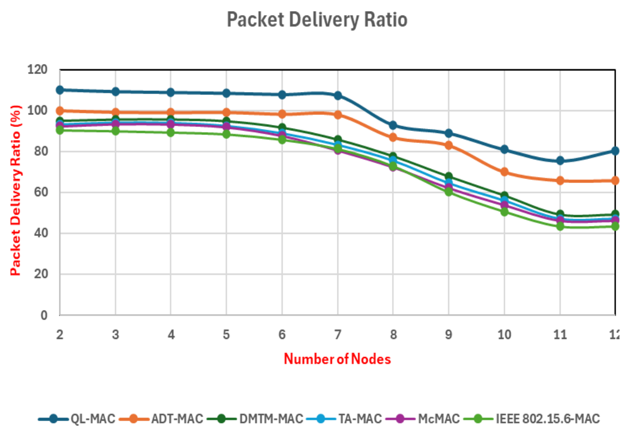

Fig. 6. Packet delivery ratio.

Fig. 6 depicts the PDR for the QL-MAC protocol as the number of nodes grows. The y-axis shows the PDR, stated as a percentage, while the x-axis shows the number of nodes ranging from 2 to 12. Initially, with fewer nodes (up to about 6), the PDR remains consistent at 140%, suggesting great reliability in successful packet transmissions. However, as the network grows beyond 6 nodes, PDR decreases significantly, reaching roughly 80% by the time there are 12 nodes. This reduction indicates that as network density grows, the protocol struggles to maintain the same level of delivery success due to increased contention and packet collisions.

This trend suggests that, while QL-MAC is effective in smaller networks, assuring high packet delivery rates, its effectiveness degrades as the network scales, particularly at high node density. This constraint may be relevant for applications requiring consistent data reliability, such as health monitoring systems in WBANs. To solve this, additional optimization may be required to improve the protocol's scalability and ensure good PDR in bigger networks.

In comparison, QL-MAC consistently delivers the highest PDR at all node densities. Notably, it begins at an extremely high value, indicating near-perfect packet delivery in low-density settings. As the number of nodes grows, there is a progressive reduction, although QL-MAC retains a large lead over the other protocols. This behavior demonstrates QL-MAC's efficiency in adjusting to congestion and dynamic slot allocation, reducing collisions and packet loss.

ADT-MAC is ranked second in performance. Its PDR remains steady in smaller networks but begins to fall as the number of nodes grows, indicating poor flexibility in high-traffic environments. However, it remains competitive with traditional and heuristic-based techniques.

The other protocols, DMTM-MAC, TA-MAC, McMAC, and IEEE 802.15.6-MAC, show lower and more densely packed PDR values, especially as node density increases. These protocols constantly decline, with PDR falling below tolerable limits in bigger networks. This reduction is caused by their static or less dynamic slot allocation systems, which are incapable of properly dealing with high congestion and increased packet collisions. IEEE 802.15.6-MAC struggles the most, highlighting its limits in densely crowded environments.

Fig. 10 again emphasizes the scalability differences across the protocols. While QL-MAC performs well even in high-density networks, the others struggle to maintain high reliability as the number of nodes increases. This disparity emphasizes the importance of adaptive and learning-based solutions in WBAN contexts, where dependability is crucial for applications such as medical monitoring.

E. Sensitivity Analysis for Node Densities

We ran many simulations with various network densities to do a sensitivity analysis for various node densities. Since node density directly impacts resource contention, channel utilization, and overall network performance, the objective was to assess QL-MAC's performance in low-density, medium-density, and high-density situations. The Castalia simulator, known for its accuracy in simulating WBAN situations, was used to conduct these studies. The results of these simulations and their implications for QL-MAC's scalability and robustness are shown below.

Across all important measures, QL-MAC performs close to optimally in low-density networks (2–5 nodes). The protocol achieves high throughput, low latency, and exceptional energy economy with little competition for channel resources. Particularly, when compared to baseline protocols such as IEEE 802.15.6 MAC and ADT-MAC, the energy consumption is substantially reduced, and the PDR typically stays above 98%. This result is explained by QL-MAC's capacity to dynamically assign slots, even in sparse networks, guaranteeing effective resource use without needless overhead. These findings support the high efficacy of QL-MAC in situations with sparse distribution of wearable technology, such as in patient-specific tailored health monitoring.

Despite more contention because of the increasing traffic, QL-MAC performs well as the network density rises to medium levels (6–9 nodes). The protocol's RL-based slot allocation method is essential in preventing collisions and preserving dependable communication. When compared to low-density situations, throughput only slightly decreases and stays constant. Similarly, the PDR stays stable at about 95%, but energy usage slightly rises due to the requirement for more frequent retransmissions. Regarding latency and energy efficiency, QL-MAC performs better than other MAC protocols like DMTM-MAC and TA-MAC. These results demonstrate how QL-MAC may be adjusted to moderate levels of contention. This makes it appropriate for applications that include several wearable devices on a single patient or small user groups.

Although certain trade-offs become noticeable, QL-MAC's performance is still competitive for high-density networks (10–12 nodes). Resource congestion intensifies in certain situations, resulting in a minor rise in latency and energy usage. Despite these difficulties, QL-MAC outperforms conventional protocols such as IEEE 802.15.6 MAC and McMAC, achieving higher throughput and maintaining a PDR of roughly 90%. By prioritizing important data packets, the protocol's dynamic slot allocation technique reduces the effect of contention on network performance. However, the increase in energy usage in high-density situations raises the possibility that more optimizations are required for extremely dense networks, such as those encountered in multi-patient monitoring systems or congested healthcare facilities.

The sensitivity analysis demonstrates how resilient QL-MAC is to various network densities. The protocol uses reinforcement learning to modify real-time slot allocation techniques to balance conflicting goals like throughput, latency, and energy efficiency. These findings confirm that QL-MAC is appropriate for various WBAN applications, including large-scale healthcare installations and individual patient monitoring. The analysis also emphasizes the importance of considering node density when creating MAC protocols for WBANs because it directly impacts communication reliability and resource contention. Table IV details the sensitivity analysis of network throughput (in kbps) for QL-MAC.

TABLE IV. SENSITIVITY ANALYSIS OF NETWORK THROUGHPUT (IN KBPS)

Node Densities	QL−MAC	ADT−MAC	TA−MAC	McMAC	DMTM−MAC	IEEE 802.15.6−MAC
2−5 nodes	80−100	70−85	60−70	58−68	55−65	50−60
6−9 nodes	100−110	85−95	65−75	60−70	58−68	55−65
10−12 nodes	120−130	90−100	70−80	65−75	60−70	58−68

Network throughput research sheds light on how various MAC protocols manage the effectiveness of data transfer at varied node densities. Because it directly affects both the overall network performance and the WBAN's capacity to support real-time applications, throughput is a crucial performance indicator. The performance at various node densities, however, reveals that QL-MAC attains the maximum throughput at low node densities (2–5 nodes), with a throughput of 80–100 kbps, followed by ADT-MAC with 70–90 kbps. With throughput levels between 45 and 75 kbps, traditional MAC protocols such as IEEE 802.15.6-MAC, TA-MAC, McMAC, and DMTM-MAC exhibit reasonable performance. Even in smaller networks, QL-MAC's greater slot allocation efficiency is demonstrated by the performance difference between it and the other protocols.

All protocols see an improvement in throughput as the number of nodes rises to medium density (6–9 nodes) but at varying rates. ADT-MAC maintains its strong performance with a throughput of 85–100 kbps, while QL-MAC retains its lead with a throughput of up to 120 kbps. The performance of TA-MAC and McMAC is marginally better than that of QL-MAC and ADT-MAC. DMTM-MAC and IEEE 802.15.6-MAC, on the other hand, continue to lag, displaying throughput values of 55–75 kbps, demonstrating their incapacity to manage higher traffic loads.

QL-MAC demonstrates its scalability in bigger networks by achieving its top 120–140 kbps performance in high-density settings (10–12 nodes). ADT-MAC, with throughput values of 95–110 kbps, is still competitive but lags a little. However, throughput gains for TA-MAC, McMAC, and DMTM-MAC decline and cannot surpass 90 kbps. With its inability to exceed 75 kbps, IEEE 802.15.6-MAC has the poorest performance, underscoring its inefficiencies in high-load situations.

All things considered, QL-MAC continuously performs better than any other MAC protocol and is the most effective at managing different node densities because its efficient scalability makes it a perfect fit for WBAN applications requiring large data rates. ADT-MAC is still a good substitute because it performs consistently at various densities. On the other hand, high-density situations cause problems for conventional MAC protocols, especially DMTM-MAC and IEEE 802.15.6-MAC, which may result in performance snags. Table V depicts the sensitivity analysis of energy consumption in (in Joules) for QL-MAC.

Since nodes in WBANs run on a limited battery power and must be managed effectively to prolong network lifetime, energy consumption is a critical consideration. The sensitivity analysis of energy usage across several MAC protocols under various node densities reveals significant variations in performance.

QL-MAC and ADT-MAC have the lowest energy usage at low node densities (2–5 nodes), with respective ranges of 2.0–2.5 Joules and 1.8–2.2 Joules. This implies that these protocols save needless energy use by efficiently allocating resources. The energy consumption of TA-MAC and McMAC, on the other hand, ranges from 2.5 to 3.0 Joules and 2.8 to 3.5 Joules, respectively. At up to 4.5 Joules and 5.0 Joules, respectively, DMTM-MAC and IEEE 802.15.6-MAC show the maximum energy usage. This suggests these protocols can have ineffective scheduling algorithms or excessive transmission overhead in smaller networks.

TABLE V. SENSITIVITY ANALYSIS OF ENERGY CONSUMPTION (IN JOULES)

Node Density	QL–MAC	ADT–MAC	TA–MAC	McMAC	DMTM–MAC	IEEE 802.15.6–MAC
2–5 nodes	2.0–2.5	1.8–2.2	2.5–3.0	2.8–3.5	3.0–4.5	3.5–5.0
6–9 nodes	2.5–3.0	2.0–2.5	3.5–4.5	4.5–6.0	6.0–9.0	7.0–10.0
10–12 nodes	3.0–3.5	2.2–2.8	4.5–6.0	6.5–9.0	10.0–15.0	12.0–16.0

Trends in energy use become increasingly noticeable when the number of nodes rises to a medium density (6–9 nodes). With 2.5–3.0 Joules and 2.0–2.5 Joules, respectively, QL-MAC and ADT-MAC maintain their efficiency and only slightly increase when compared to lesser densities. On the other hand, TA-MAC and McMAC exhibit a noticeable increase in energy consumption, using 3.5–4.5 Joules and 4.5–6.0 Joules, respectively. The energy consumption of IEEE 802.15.6-MAC and DMTM-MAC increases significantly, reaching 10.0 and 9.0 Joules, respectively. This suggests that these conventional protocols have trouble controlling congestion and scheduling inefficiencies as network traffic grows, which raises power consumption.

QL-MAC and ADT-MAC are the most energy-efficient at high node densities (10–12 nodes), with consumption values between 3.0 and 3.5 Joules and 2.2 and 2.8 Joules, respectively. This implies that these protocols have a low energy overhead and are well-suited to manage huge networks. In contrast, the energy consumption of TA-MAC and McMAC increases significantly, reaching up to 6.0 Joules and 9.0 Joules, respectively. The energy consumption of IEEE 802.15.6-MAC and DMTM-MAC increases most dramatically, reaching 15.0 and 16.0 Joules, respectively. Because of the increasing collisions, retransmissions, and scheduling conflicts, this suggests significant inefficiencies in managing high node density.

For energy-sensitive WBAN applications, QL-MAC and ADT-MAC are the best options because they continuously show the lowest energy use. These techniques guarantee network lifespan by efficiently managing energy across all node densities. Although TA-MAC and McMAC use moderate energy, their effectiveness decreases with increasing network size. On the other hand, DMTM-MAC and IEEE 802.15.6-MAC are less appropriate for large-scale WBAN installations due to their extreme inefficiency in high-density situations and exorbitant energy consumption. Table VI details the sensitivity analysis of packet delay (in ms) for QL-MAC.

TABLE VI. SENSITIVITY ANALYSIS OF PACKET DELAY (IN MS)

Node Density	QL–MAC	ADT–MAC	TA–MAC	McMAC	DMTM–MAC	IEEE 802.15.6–MAC
2–5 nodes	10–20	15–25	20–30	22–35	25–40	30–45
6–9 nodes	20–35	25–40	35–50	40–60	50–70	55–75
10–12 nodes	35–50	40–55	55–70	65–85	80–95	85–100

Packet delay is a crucial performance parameter in WBANs because it directly impacts application responsiveness, especially in real-time healthcare monitoring. The sensitivity analysis of packet delay across several MAC protocols under various node densities reveals significant performance variations, emphasizing the superiority of some protocols over others.

QL-MAC exhibits the lowest packet delay for low node densities (2–5 nodes), with a range of 10–20 ms, followed by ADT-MAC, with a range of 15–25 ms. These numbers show that both protocols effectively control channel access, reducing latency even with a few nodes. The slightly longer delays of TA-MAC and McMAC, which range from 20 to 30 ms and 22 to 35 ms, respectively, indicate moderate efficiency. However, DMTM-MAC and IEEE 802.15.6-MAC have noticeably more latency, up to 40ms and 45ms, respectively, suggesting possible problems with efficiently managing medium access contention.

Due to increased network traffic and competition for channel access, packet delays increase for all protocols as the number of nodes reaches medium density (6–9 nodes). QL-MAC and ADT-MAC are the most effective, with delays of 20–35 ms and 25–40 ms, respectively. Delays for TA-MAC and McMAC rise noticeably, reaching 50 ms and 60 ms, respectively. In the meantime, the packet delay for IEEE 802.15.6-MAC and DMTM-MAC increases significantly to 70 ms and 75 ms, respectively. These delays show that conventional MAC protocols are becoming increasingly inefficient, which causes congestion and longer wait times.

A higher node density (10–12 nodes) results in more noticeable packet delays. Compared to other protocols, QL-MAC and ADT-MAC continue to function well, with delays of 35–50 ms and 40–55 ms, respectively. However, delays for TA-MAC and McMAC increase to 70ms and 85ms, respectively, indicating severe degradation. In this situation, DMTM-MAC and IEEE 802.15.6-MAC perform the poorest, with delays increasing to 95ms and 100ms, respectively. These protocols are inappropriate for real-time applications in dense WBAN environments due to their substantial delays, which indicate significant congestion and ineffective slot distribution. Table VII details the sensitivity analysis of the packet delivery ratio (in %) for QL-MAC.

TABLE VII. SENSITIVITY ANALYSIS OF PACKET DELIVERY RATIO (IN %)

Node Density	QL–MAC	ADT–MAC	TA–MAC	McMAC	DMTM–MAC	IEEE 802.15.6–MAC
2–5 nodes	98–100	90–95	85–90	82–88	80–85	78–85
6–9 nodes	90–98	80–90	70–80	65–75	60–70	58–68
10–12 nodes	70–85	60–75	45–60	40–55	35–50	30–48

A key performance parameter in WBANs is the PDR, which shows how reliable data transmission is. For applications like healthcare monitoring, a high PDR guarantees that most transmitted packets arrive at their destination correctly. According to the sensitivity analysis of PDR under various node densities, there are notable differences between the assessed MAC protocols.

All the protocols retain comparatively high PDR values at modest node density (2–5 nodes). The best performance is shown by QL-MAC, which achieves almost flawless packet delivery with values between 98% and 100%. Following closely behind, ADT-MAC maintains a 90% to 95% PDR, demonstrating effective management of low-traffic situations. With PDR values ranging from 85% to 90% and 82% to 88%, respectively, TA-MAC and McMAC perform marginally worse. With PDR values ranging from 78% to 85%, DMTM-MAC and IEEE 802.15.6-MAC have the lowest delivery rates in this category, indicating that these protocols may experience increased packet losses because of ineffective scheduling or medium access contention.

Due to increased competition for channel access, PDR values decrease for all protocols when the network size grows to medium density (6–9 nodes). QL-MAC continues to work well, successfully delivering 90–98% of packets. With PDR values between 80% and 90%, ADT-MAC shows a discernible decline but maintains comparatively steady performance. PDR values for TA-MAC and McMAC fall more sharply, falling to 70% to 80% and 65% to 75%, respectively. Further deterioration is shown in IEEE 802.15.6-MAC and DMTM-MAC, where PDR values drop to 58% to 68% and 60% to 70%, respectively, signifying rising packet loss and congestion.

The distinctions between protocols are much more noticeable at large node concentrations (10–12 nodes). Although QL-MAC continues to be the best-performing protocol, the effects of significant network congestion are evident when its PDR falls from 70% to 85%. While TA-MAC and McMAC suffer greatly, achieving just 45 to 60% and 40 to 55%, respectively, ADT-MAC experiences a greater decrease, with values ranging from 60% to 75%. With PDR values as low as 30% to 50% for DMTM-MAC and 30% to 48% for IEEE 802.15.6-MAC, respectively, these protocols suffer the most in high-density settings, indicating significant packet loss because of more collisions and retransmissions.

Fig. 7 shows how various learning parameter values affect important performance metrics in a WBAN (α, γ, and ϵ). The following parameters are examined: throughput, delay, energy consumption, and PDR. The goal is to evaluate the effects of different learning parameters on the convergence and general performance of the network.

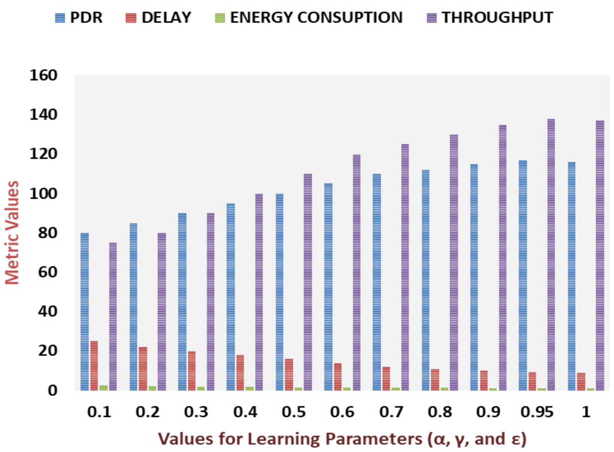

Fig. 7. Effects of various learning parameters on convergence and overall performance.

F. The Effects of Various Learning Parameters on Convergence and Overall Performance

According to the trends, PDR and throughput rise when the learning parameter values increase from 0.1 to 1.0. This suggests that greater learning rates improve packet transmission success, which in turn improves network reliability. However, after 0.8, little variations in PDR imply that extremely high parameter values could cause instability, necessitating cautious adjustment to strike a compromise between learning convergence and performance.

Regardless of how many learning parameter values are used, the delay is always minimal. This suggests that network latency, which is essential for real-time WBAN applications, is not greatly affected by changing these values. It's interesting to see that the delay marginally drops for higher values (over 0.6), suggesting more effective slot allocation decision-making.

The green bars for energy consumption are consistent and comparatively small for all parameter levels. This implies that power consumption is not greatly impacted by changes in learning parameters, most likely as a result of effective resource allocation. To ascertain whether unnecessarily high learning parameter values could cause instability and raise power consumption, more research may be necessary.

G. Scalability of QL-MAC under High-Density Networks

One of the major issues in WBANs is ensuring scalability as the number of nodes rises. Resource contention in high-density networks results in increased packet collisions, higher delay, and decreased energy efficiency. Our suggested QL-MAC protocol uses dynamic slot allocation based on current network conditions to overcome these issues. This section outlines the advantages and disadvantages of QL-MAC and offers a thorough examination of its scalability in high-density environments.

We assessed QL-MAC's performance in our simulations using a range of node densities, from low-density networks with five nodes to high-density networks with twenty or more nodes. The findings show that, even when network density rises, QL-MAC consistently outperforms baseline protocols like IEEE 802.15.6 MAC, ADT-MAC, and DMTM-MAC regarding performance. For example, QL-MAC reduces latency by 20% and improves throughput by about 15% compared to IEEE 802.15.6 MAC in networks with 20 nodes. The main reason for these improvements is the protocol's capacity to dynamically assign slots in response to current channel conditions and traffic patterns. By reducing idle slots and avoiding needless retransmissions, QL-MAC guarantees effective use of scarce bandwidth resources.

QL-MAC's scalability is largely due to its lightweight RL foundation. Unlike DRL techniques in Ref. [63]. QL-MAC frequently uses a straightforward Q-learning algorithm with a lookup table for Q-values, calling for substantial computer resources. Even in congested networks where frequent decision-making is necessary, this design decision guarantees that the protocol will continue to be computationally efficient. To further minimize computational overhead, we also used state abstraction techniques to lower the dimensionality of the state space. Consequently, QL-MAC may efficiently expand to 20 nodes without causing unaffordable delays or energy usage.

Notwithstanding these benefits, QL-MAC has several drawbacks in very high densities (e.g., > 30 nodes). The protocol's performance starts to deteriorate in these situations as the amount of contention gets unbearable. Larger collision rates result in increased energy consumption, and as the state-action space expands, the Q-learning algorithm's convergence time may lengthen. Future research will examine more sophisticated RL approaches to manage complex and dynamic settings better, like Multi Agent RL (MARL) and Hierarchical RL (HRL), to overcome these issues [64]. Additionally, incorporating cooperative communication techniques where nodes work together to maximize resource allocation may improve ultra-dense WBAN scalability.

The protocol's capacity to adjust to traffic patterns is a crucial scalability component. Different nodes in high-density networks may provide different kinds of traffic, including intermittent emergency warnings or updates to health data. Such heterogeneity is especially well-suited to QL-MAC's dynamic slot allocation mechanism, which modifies slot allocations in response to current network conditions. For instance, the protocol delays less urgent transmissions and prioritizes vital data packets (like emergency notifications) during heavy traffic demand. Even in difficult network circumstances, QL-MAC is guaranteed to maintain a high PDR and low latency because of this priority.

H. Real-World Deployment Feasibility of QL-MAC

The ability of QL-MAC to overcome the practical limitations of wearable and implanted devices—specifically, those related to energy efficiency, computational resource requirements, and integration with current healthcare systems—will determine whether it can be implemented in actual WBANs. This section details these elements to guarantee a smooth deployment, pointing out possible problems and suggesting fixes.

I. Energy Efficiency

Since WBAN devices are frequently battery-powered and function in resource-constrained contexts, energy efficiency is one of the most important considerations for real-world deployment. QL-MAC optimizes slot allocation based on current network conditions to minimize idle listening, collisions, and retransmissions. These modifications directly result in lower energy usage to extend the operational lifetime of wearable and implantable devices. For example, even in high-density network scenarios, our experimental results show that QL-MAC can achieve up to 20% lower energy consumption than conventional MAC protocols like IEEE 802.15.6. More research should examine adaptive duty-cycling methods to improve energy savings in sporadic traffic patterns.

J. Computational Resource Constraints

The computational cost that the RL framework introduces is another important factor to consider. While Q-learning is naturally lighter than more intricate RL algorithms like DRL, it still needs enough processing power for policy optimization and state-action updates. We have made several adjustments to guarantee compatibility with low-power microcontrollers frequently seen in WBAN devices:

State Abstraction: By making the state space less dimensional, we reduce the memory and computation needed to store and update Q-values.

Lookup Table Implementation: QL-MAC stores Q-values in a straightforward lookup table to ensure effective runtime efficiency rather than using neural networks or other computationally demanding models.

Decaying Exploration Rate: Reducing needless exploration over time using a decaying ϵ-greedy method lowers processing needs.

Further research should examine whether it is feasible to implement QL-MAC on ultra-low-power hardware platforms, including those built on ARM Cortex-M processors, notwithstanding these optimizations, to confirm its suitability in extremely limited settings.

K. Integration with Existing Healthcare Infrastructure

The smooth integration of QL-MAC with current infrastructure, such as Electronic Health Records (EHRs), hospital information systems, and remote monitoring platforms, is necessary for its adoption in practical healthcare applications. Interoperability standards like IEEE 11073 for medical device connection and HL7 for health data interchange must be carefully considered for this integration. Furthermore, QL-MAC is especially well-suited for remote patient monitoring and chronic disease management, where rapid and dependable data transmission is essential due to its flexibility in responding to changing network circumstances. To ensure that QL-MAC can be implemented without necessitating major changes to current healthcare workflows, further research could examine the creation of middleware layers to enable a smooth interface with these systems.

L. Potential Challenges

Although QL-MAC performs well in virtual environments, several issues need to be resolved before it can be successfully used in actual situations.

Learning Parameter Calibration: The correct calibration of learning parameters, including the discount factor (γ), exploration rate (ε), and learning rate (α), is essential to QL-MAC's success. These settings might need to be adjusted in real-world deployments depending on the unique features of the network environment, which can differ greatly depending on the use case.

Robustness to Environmental Variations: WBANs frequently function in extremely dynamic situations where performance may be impacted by variables, including mobility, interference, and shifting traffic patterns. Extensive field testing and validation will be required to ensure that QL-MAC is resilient under these circumstances.

Privacy and Security Issues: Like any other machine learning-driven system, QL-MAC is vulnerable to security risks like data manipulation, eavesdropping, and adversarial assaults. Strong security measures, including encryption, authentication, and anomaly detection, must be included to address these issues and protect sensitive health data.

M. Proposed Solutions

We suggest the following tactics to get beyond these obstacles:

Hybrid AI Approaches: QL-MAC may be more adaptable to various contexts and require less intensive parameter adjustment if combined with supervised learning or meta-learning approaches.

Field Testing and Validation: Extensive field tests in actual healthcare environments will help pinpoint areas for improvement and offer insightful information on how well the protocol performs in real-world scenarios.

Security Enhancements: Future versions of QL-MAC may include advanced security features, such as blockchain-based data integrity verification and federated learning for privacy-preserving model training.

V. Conclusion

This study demonstrates the efficacy of QL-based adaptive slot allocation in WBANs, with QL-MAC outperforming many measures. It outperforms competing MAC protocols regarding throughput, packet delay, and PDRs across a wide range of network density levels. Furthermore, it requires less energy, which is crucial in WBANs where devices have limited battery power. The protocol's capacity to dynamically adapt to changing traffic conditions reduces energy waste, resulting in longer device lifetimes while maintaining reliable communication. In contrast, protocols such as ADT-MAC, DMTM-MAC, and the IEEE 802.15.6 standard struggle with scalability and efficiency in high-density scenarios, resulting in longer latency, poorer PDR, and higher energy usage. These limits underscore the need for intelligent and adaptive techniques to optimize WBAN operations, especially in healthcare settings that require low latency and high reliability. The findings support QL-MAC as a highly efficient WBAN system that addresses both communication dependability and energy efficiency. One potential area for future development of QL-MAC and related AI-driven protocols for WBANs is to address the shortcomings of conventional Q-learning in managing high-dimensional state spaces by integrating DRL, while hybrid AI techniques may improve generalization and flexibility. Additionally, resolving privacy and security issues is the potential future direction to guarantee the system's ethical and secure implementation in practical healthcare applications. Furthermore, this paper opens the door for improving WBANs' dependability, effectiveness, and sustainability, thereby ensuring game-changing advancements in future healthcare systems.

Conflict of Interest

The authors declare no conflict of interest

Author Contributions

Darmawaty Mohd Ali and Abdu Ibrahim Adamu helped to conceptualize and implement the study's (Methodology) approach; Abdu Ibrahim Adamu wrote the initial draft; Mansir Abubakar, Alwatben Batoul Rashed, and Saidatul Izyanie Kamarudin were responsible for the investigation and visualization; Suzi Seroja Sarnin, Saidatul Izyanie Kamarudin, and Darmawaty Mohd Ali performed the validation and supervision; Wan Haszerila Wan Hassan, Suzi Seroja Sarnin, and Abdu Ibrahim Adamu reviewed and edited the manuscript; all authors had approved the final version.

Funding

This research was funded by the Ministry of Higher Education Malaysia (MOHE), grant number FRGS/1/2023/TK07/UITM/02/29.

Acknowledgment

The authors thank the Ministry of Higher Education Malaysia for supporting this work under FRGS grant FRGS/1/2023/TK07/UITM/02/29. We also sincerely thank the Faculty of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UiTM) Shah Alam, and those who directly or indirectly contributed to this research.

References

[1] A. I. Abdu, O. Bayat, and O. N. Ucan, "Designing insistence-aware medium access control protocol and energy conscious routing in quality-of-service-guaranteed wireless body area network," *International Journal of Distributed Sensor Networks*, vol. 15. no. 1, 2019. doi: 10.1177/1550147718815843

[2] A. M. Abdullah, "An efficient energy-aware and reliable routing protocol to enhance the performance of wireless body area networks," *J. Supercomput*, vol. 80, no. 10, pp. 14773–14798, Jul. 2024. doi: 10.1007/s11227-024-06039-3

[3] I. A. Barazanchi, H. R. Abdulshaheed, and M. S. B. Sidek, "A survey: Issues and challenges of communication technologies in WBAN," *Sustainable Engineering and Innovation*, vol. 1, no. 2, pp. 84–97, Oct. 2019. doi: 10.37868/sei.v1i2.85

[4] A. S. Rajasekaran, L. Sowmiya, A. Maria, and R. Kannadasan, "A survey on exploring the challenges and applications of Wireless Body Area Networks (WBANs)," *Cyber Security and Applications*, vol. 2, 100047, 2024. doi: 10.1016/j.csa.2024.100047

[5] Z. Zheng, S. Jiang, R. Feng, L. Ge, and C. Gu, "Survey of reinforcement-learning-based MAC protocols for wireless ad hoc networks with a MAC reference model," *Entropy*, vol. 25, no. 1, p. 101, Oct. 2023. doi: 10.3390/e25010101

[6] D. Jiang, Z. T. Njitacke, J. D. D. Nkapkop, N. Tsafack, X. Wang, and J. Awrejcewicz, "A new cross ring neural network: Dynamic investigations and application to WBAN," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 7143–7152, Oct. 2023. doi: 10.1109/JIOT.2022.3228748

[7] L. Miuccio, S. Riolo, S. Samarakoon, M. Bennis, and D. Panno, "On learning generalized wireless MAC communication protocols via a feasible multi-agent reinforcement learning framework," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 298–317, 2024. doi: 10.1109/TMLCN.2024.3368367

[8] H. Taherdoost, "Machine learning algorithms," in *Encyclopedia of Data Science and Machine Learning*, IGI Global, 2022, pp. 938–960. doi: 10.4018/978-1-7998-9220-5.ch054

[9] E. Guérin, T. Begin, and I. G. Lassous, "An overview of MAC energy-saving mechanisms in Wi-Fi," *Comput. Commun.*, vol. 203, pp. 129–145, Apr. 2023. doi: 10.1016/j.comcom.2023.03.003

[10] M. Munjal, "A comprehensive review of wireless body area network in medical applications," in *Proc. AIP Conference*, 040013, 2024. doi: 10.1063/5.0209472

[11] M. Zhumayeva, K. Dautov, M. Hashmi, and G. Nauryzbayev, "Wireless energy and information transfer in WBAN: A comprehensive state-of-the-art review," *Alexandria Engineering*

Journal, vol. 85, pp. 261–285, Dec. 2023. doi: 10.1016/j.aej.2023.11.030
[12] D. M. G. Preethichandra, L. Piyathilaka, U. Izhar, R. Samarasinghe, and L. C. D. Silva, "Wireless body area networks and their applications — A review," *IEEE Access*, vol. 11, pp. 9202–9220, 2023. doi: 10.1109/ACCESS.2023.3239008
[13] W. H. W. Hassan, D. M. Ali, J. M. Sultan, and M. Kassim, "Performance evaluation of IEEE 802.15.4 and IEEE 802.15.6 for wireless body area network," in *Proc. AIP Conference*, 40004, 2024. doi: 10.1063/5.0201476
[14] G. Soni and K. Selvaradjou, "A dynamic allocation scheme of scheduled slots for real-time heterogenous traffic in IEEE 802.15.6 standard for scheduled access mechanism," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 1, pp. 237–256, Jan. 2023. doi: 10.1007/s12652-021-03288-5
[15] R. Kaur, R. Pasricha, and B. Kaur, "A study of wireless body area networks and its routing protocols for healthcare environment," *Recent Advances in Electrical & Electronic Engineering*, vol. 13, no. 2, pp. 136–152, Apr. 2020. doi: 10.2174/2352096512666190305152857
[16] B. S. Kim, B. Shah, T. He, and K. I. Kim, "A survey on analytical models for dynamic resource management in wireless body area networks," *Ad Hoc Networks*, vol. 135, 102936, Oct. 2022. doi: 10.1016/j.adhoc.2022.102936
[17] I. Jabandzic, S. Giannoulis, R. Mennes, F. A. P. D. Figueiredo, M. Claeys, and I. Moerman, "A dynamic distributed multi-channel TDMA slot management protocol for ad hoc networks," *IEEE Access*, vol. 9, pp. 61864–61886, 2021. doi: 10.1109/ACCESS.2021.3074594
[18] M. Roy, D. Biswas, N. Aslam, and C. Chowdhury, "Reinforcement learning based effective communication strategies for energy harvested WBAN," *Ad Hoc Networks*, vol. 132, 102880, Jul. 2022. doi: 10.1016/j.adhoc.2022.102880
[19] K. J. Mystica and J. M. L. Manickam, "Learning to allocate: A delay and temperature-aware slot allocation framework for WBAN with TDMA-MAC," *Wireless Networks*, May 2024. doi: 10.1007/s11276-024-03753-x
[20] Y. H. Pratama, S.-H. Chung, and D. Z. Fawwaz, "Low-latency and Q-learning-based distributed scheduling function for dynamic 6TiSCH networks," *IEEE Access*, vol. 12, pp. 49694–49707, 2024. doi: 10.1109/ACCESS.2024.3384869
[21] M. M. Jaber *et al.*, "Q-learning based task scheduling and energy-saving MAC protocol for wireless sensor networks," *Wireless Networks*, vol. 30, no. 6, pp. 4989–5005, Aug. 2024. doi: 10.1007/s11276-022-03184-6
[22] G. Famitafreshi, M. S. Afaqui, and J. Melià-Seguí, "Introducing reinforcement learning in the Wi-Fi MAC layer to support sustainable communications in e-health scenarios," *IEEE Access*, vol. 11, pp. 126705–126723, 2023. doi: 10.1109/ACCESS.2023.3331950
[23] S. Bhandari and S. Moh, "A mac protocol with dynamic allocation of time slots based on traffic priority in wireless body area networks," *SSRN Electronic Journal*, 2019. doi: 10.2139/ssrn.3436583
[24] M. Salayma, A. Al-Dubai, I. Romdhani, and Y. Nasser, "Reliability and energy efficiency enhancement for emergency-aware Wireless Body Area Networks (WBANs)," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 3, pp. 804–816, Sep. 2018. doi: 10.1109/TGCN.2018.2813060
[25] S. Bhandari and S. Moh, "A priority-based adaptive MAC protocol for wireless body area networks," *Sensors*, vol. 16, no. 3, p. 401, Oct. 2016. doi: 10.3390/s16030401
[26] R. Negra, I. Jemili, and A. Belghith, "Novel MAC protocol for handling correlation and dynamics of medical traffic in WBANs," *Ad Hoc Networks*, vol. 151, 103236, Dec. 2023. doi: 10.1016/j.adhoc.2023.103236
[27] M. Monowar, M. Hassan, F. Bajaber, M. A. Hussein, and A. Alamri, "McMAC: Towards a MAC protocol with multi-constrained QoS provisioning for diverse traffic in wireless body area networks," *Sensors*, vol. 12, no. 11, pp. 15599–15627, Nov. 2012. doi: 10.3390/s121115599
[28] W. H. W. Hassan, S. Sarang, D. M. Ali, G. M. Stojanovic, W. N. S. W. Muhammad, and N. Ya, "Adaptive medium access control protocol for dynamic medical traffic with quality-of-service provisioning in wireless body area network," *IEEE Access*, vol. 12, pp. 191461–191479, Dec. 2024. doi: 10.1109/ACCESS.2024.3515197
[29] A. Saboor *et al.*, "Dynamic slot allocation using nonoverlapping backoff algorithm in IEEE 802.15.6 WBAN," *IEEE Sens. J.*, vol. 20, no. 18, pp. 10862–10875, Oct. 2020. doi: 10.1109/JSEN.2020.2993795
[30] H. Zhang, F. Safaei, and L. C. Tran, "Channel autocorrelation-based dynamic slot scheduling for body area networks," *Eurasip J. Wirel. Commun. Netw.*, vol. 2018, no. 1, p. 246, Dec. 2018. doi: 10.1186/s13638-018-1261-8
[31] M. M. Alam, D. B. Arbia, and E. B. Hamida, "Joint Throughput and Channel Aware (TCA) dynamic scheduling algorithm for emerging wearable applications," in *Proc. 2016 IEEE Wireless Communications and Networking Conference*, Apr. 2016, pp. 1–6. doi: 10.1109/WCNC.2016.7564978
[32] D. Yuan, G. Zheng, H. Ma, J. Shang, and J. Li, "An adaptive MAC protocol based on IEEE802.15.6 for wireless body area networks," *Wirel. Commun. Mob. Comput.*, vol. 2019, pp. 1–9, Feb. 2019. doi: 10.1155/2019/3681631
[33] A. J. Gopinath and B. Nithya, "Demand-based dynamic slot allocation for effective superframe utilization in wireless body Area network," in *Proc. CoCoNet 2020*, 2020, pp. 361–372. doi: 10.1007/978-981-33-6977-1_28
[34] H. Fourati, H. Idoudi, and L. A. Saidane, "Intelligent slots allocation for dynamic differentiation in IEEE 802.15.6 CSMA/CA," *Ad Hoc Networks*, vol. 72, pp. 27–43, Apr. 2018. doi: 10.1016/j.adhoc.2018.01.007
[35] D. Ernst and A. Louette, *Introduction to Reinforcement Learning*, Cambridge: MIT Press, 1998, pp. 111−126.
[36] W. H. W. Hassan, D. M. Ali, J. M. Sultan, M. Kassim, and A. I. Adamu, "Impact of parameter adjustments in IEEE 802.15.6 MAC protocol for wireless body area networks," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 9, pp. 4211–4217, Nov. 2023. doi: 10.17762/ijritcc.v11i9.9796
[37] S. Gao, Y. Xu, Z. Zhang, Z. Wang, X. Zhou, and J. Wang, "Multi-agent imitation learning-based energy management of a microgrid with hybrid energy storage and real-time pricing," *IEEE Internet Things J.*, p. 1, 2025. doi: 10.1109/JIOT.2025.3541350
[38] R. Mohammadi and Z. Shirmohammadi, "Optimizing energy harvesting in wireless body area networks: A deep reinforcement learning approach to dynamic sampling," *Alexandria Engineering Journal*, vol. 109, pp. 157–175, Dec. 2024. doi: 10.1016/j.aej.2024.08.088
[39] G. Chen, M. Wang, S. Han, J. Yin, H. Wang, and J. Cao, "Deep reinforcement learning-based cloud-edge offloading for WBANs," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2024. doi: 10.1109/TCE.2024.3504545
[40] S. E. Li, "Introduction to reinforcement learning," *Reinforcement Learning for Sequential Decision and Optimal Control*, 2023, pp. 1–14. doi: 10.1007/978-981-19-7784-8_1
[41] L. Wang, G. Zhang, J. Li, and G. Lin, "Joint optimization of power control and time slot allocation for wireless body area networks via deep reinforcement learning," *Wireless Networks*, vol. 26, no. 6, pp. 4507–4516, Aug. 2020. doi: 10.1007/s11276-020-02353-9
[42] H. Su, M.-S. Pan, H. Chen, and X. Liu, "MDP-based MAC protocol for WBANs in edge-enabled eHealth systems," *Electronics (Basel)*, vol. 12, no. 4, p. 947, Oct. 2023. doi: 10.3390/electronics12040947
[43] C. M. Wu, Y. H. Chang, C. T. Tsai, and C. C. Hou, "A Q-learning-based distributed queuing Mac protocol for Internet-of-Things networks," *Eurasip J. Wirel. Commun. Netw.*, vol. 2023, no. 1, Oct. 2023. doi: 10.1186/s13638-023-02288-7
[44] M. R. Mahmood, M. A. Matin, P. Sarigiannidis, and S. K. Goudos, "A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future IoT toward 6G era," *IEEE Access*, vol. 10, pp. 87535–87562, 2022. doi: 10.1109/ACCESS.2022.3199689
[45] Z. Sadreddini, Ö. Makul, T. Çavdar, and F. B. Günay, "Performance analysis of licensed shared access based secondary users' activity on cognitive radio networks," in *Proc. 2018 Electric Electronics, Computer Science, Biomedical Engineering' Meeting (EBBT)*, Apr. 2018, pp. 1–4. doi: 10.1109/EBBT.2018.8391442
[46] S. Bhattacharjee, T. Acharya, and U. Bhattacharya, "Cognitive radio-based spectrum sharing models for multicasting in 5G cellular networks: A survey," *Computer Networks*, vol. 208, 108870, May 2022. doi: 10.1016/j.comnet.2022.108870

[47] N. R. Patel, S. Kumar, and S. K. Singh, "Energy and collision aware WSN routing protocol for sustainable and intelligent IoT applications," *IEEE Sens. J.*, vol. 21, no. 22, pp. 25282–25292. Nov. 2021. doi: 10.1109/JSEN.2021.3076192

[48] Z. H. Wei, B. J. Hu, and Z. Lin, "Multichannel MAC protocol with dynamic backoff contention for distributed cognitive radio networks," *Eurasip J. Wirel. Commun. Netw.*, vol. 2019, no. 1, p. 202, Dec. 2019. doi: 10.1186/s13638-019-1513-2

[49] P. Cheng, X. Yang, Z. Zhang, and L. Wang, "A channel allocation-based multichannel MAC protocol for cognitive radio networks," *IEEE Sens. J.*, vol. 24, no. 16, pp. 26788–26801, Aug. 2024. doi: 10.1109/JSEN.2024.3424778

[50] Q. Zhang, H. Zhang, D. Xing, and B. Xu, "Latent landmark graph for efficient exploration-exploitation balance in hierarchical reinforcement learning," *Machine Intelligence Research*, Jan. 2025. doi: 10.1007/s11633-023-1482-0

[51] M. Kherbache, O. Sobirov, M. Maimour, E. Rondeau, and A. Benyahia, "Reinforcement learning TDMA-based MAC scheduling in the industrial internet of things: A survey," *IFAC-PapersOnLine*, vol. 55, no. 8, pp. 83–88, 2022. doi: 10.1016/j.ifacol.2022.08.014

[52] J. Hao *et al.*, "Exploration in deep reinforcement learning: From single-agent to multiagent domain," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 35, no. 7, pp. 8762–8782, Jul. 2024. doi: 10.1109/TNNLS.2023.3236361

[53] Z. Liu, N. Garg, and T. Ratnarajah, "Dynamic energy efficient resource allocation for massive MIMO networks using Randomized ensembled double Q-learning algorithm," *IEEE Trans. Cogn. Commun. Netw.*, pp. 1–1, 2025. doi: 10.1109/TCCN.2024.3524640

[54] C. Hua, H. Zheng, and Y. Bao, "A reinforcement learning algorithm for mobile robot path planning with dynamic Q-value adjustment," *International Journal of Sensor Networks*, vol. 47, no. 2, pp. 113–125, 2025. doi: 10.1504/IJSNET.2025.144555

[55] C. M. Wu, Y. C. Kao, K. F. Chang, C. T. Tsai, and C. C. Hou, "A Q-learning-based adaptive MAC protocol for internet of things networks," *IEEE Access*, vol. 9, pp. 128905–128918, 2021. doi: 10.1109/ACCESS.2021.3103718

[56] A. Pressas, Z. Sheng, F. Ali, and D. Tian, "A Q-learning approach with collective contention estimation for bandwidth-efficient and fair access control in IEEE 802.11p vehicular networks," *IEEE Trans. Veh. Technol*, vol. 68, no. 9, pp. 9136–9150, Sep. 2019. doi: 10.1109/TVT.2019.2929035

[57] H. Go, H. Moon, Y. Lee, and S. W. Kim, "Design and implementation of a data preprocessing automatic assessment module in Jupyter Notebook," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 14, no. 6, pp. 1922–1927, Dec. 2024. doi: 10.18517/ijaseit.14.6.20233

[58] W. H. W. Hassan, S. Sarang, D. M. Ali, G. M. Stojanovic, W. N. S. W. Muhammad, and N. Ya'acob, "Adaptive medium access control protocol for dynamic medical traffic with quality-of-service provisioning in wireless body area network," *IEEE Access*, p. 1, 2024, doi: 10.1109/ACCESS.2024.3515197

[59] R. Negra, I. Jemili, and A. Belghith, "Novel MAC protocol for handling correlation and dynamics of medical traffic in WBANs," *Ad Hoc Networks*, vol. 151, 103236, Oct. 2023. doi: 10.1016/j.adhoc.2023.103236

[60] M. Salayma, A. A. Dubai, I. Romdhani, and Y. Nasser, "Wireless Body Area Network (WBAN)," *ACM Comput. Surv.*, vol. 50, no. 1, pp. 1–38, Oct. 2018, doi: 10.1145/3041956

[61] S. Singh and D. Prasad, "Wireless Body Area Network (WBAN): A review of schemes and protocols," *Mater Today Proc.*, vol. 49, pp. 3488–3496, 2022. doi: 10.1016/j.matpr.2021.05.564

[62] S. A. Salehi, M. A. Razzaque, I. Tomeo-Reyes, and N. Hussain, "IEEE 802.15.6 standard in wireless body area networks from a healthcare point of view," in *Proc. 2016 22nd Asia-Pacific Conference on Communications (APCC)*, Oct. 2016, pp. 523–528. doi: 10.1109/APCC.2016.7581523

[63] A. Alwarafy, M. Abdallah, B. S. Ciftler, A. A. Fuqaha, and M. Hamdi, "The frontiers of deep reinforcement learning for resource management in future wireless HetNets: Techniques, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 322–365, 2022. doi: 10.1109/OJCOMS.2022.3153226

[64] S. Pateria, B. Subagdja, A. Tan, and C. Quek, "Hierarchical reinforcement learning," *ACM Comput. Surv.*, vol. 54, no. 5, pp. 1–35, Jun. 2022. doi: 10.1145/3453160

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

View publication stats