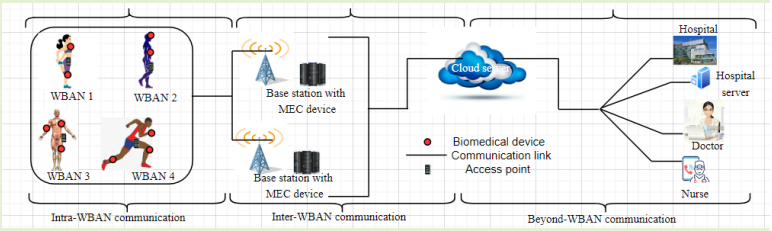
Markov Decision Process Based Energy Aware MAC Protocol for IoT WBAN Systems

Damilola D. Olatinwo, Adnan M. Abu-Mahfouz, *Senior Member, IEEE*, Gerhard P. Hancke, *Fellow, IEEE*, and Hermanus C. Myburgh, *Member, IEEE*

***Abstract*— Internet of Things (IoT) enabled WBAN systems play a crucial role in healthcare monitoring, enhancing patients' well-being at an affordable cost. However, their primary challenge lies in energy constraints. To optimize the limited energy resources and extend the lifetime of devices within the network, a multi-channel MAC protocol with Markov decision process (MDP-HYMAC) is proposed. This protocol improves energy efficiency, throughput, and network lifetime, and minimizes delays by using separate channels for communication between biomedical devices and access points (APs). In addition, a Markov decision process is employed to stochastically model the systems' transition states and explore optimal communication strategies between biomedical devices and APs. Furthermore, an adaptive power allocation scheme, a time-slot allocation scheme, and a back-off strategy are designed to minimize time-slot wastage, energy consumption, and delays. The proposed protocol outperforms the baseline methods by achieving significant improvements in energy efficiency from 4% to 22%. The findings of this research strongly indicate that the proposed protocol has the potential to significantly improve the performance of WBAN systems, particularly in the context of sustainable healthcare monitoring.**



***Index Terms*— healthcare monitoring, wireless body area network, Internet of Things, multi-channel MAC protocol design, reinforcement learning, energy efficiency, channel efficiency**

I. Introduction

INTERNET of Things (IoT) offers promising opportunities to build smart systems. Wireless networks are important technologies of IoT and are widely used in different fields including healthcare, structural health, agriculture, transportation, and environmental monitoring. In the healthcare domain, wireless body area networks (WBANs) are used to achieve remote healthcare monitoring [1]. An IoT WBAN is a special network dedicated to healthcare applications for managing communications between different devices such as biomedical devices, smartphones, and computers. Biomedical devices are positioned in, on, and around a patient's body. They are small, lightweight, and smart devices that are responsible for sensing and collecting vital health signals from the patient's body. Various WBAN biomedical devices have been developed for specific purposes. For example, an electroencephalogram (EEG) sensor is used to measure brain waves to detect abnormalities such as blood clot (cerebral venous sinus thrombosis (CVST)), which is a possible sign of stroke or measure electrical activity of the brain. An electrocardiogram (ECG) sensor is used to measure heart rate (that is, PQRST wave) for possible detection of different heart diseases. Accelerometer sensor is used to measure patients' walking pattern in a three-axis motion to calculate the risk of falling. glucose sensor (glucometer) measures the level of glucose in a patient's body to manage diabetes mellitus. SpO2 readings on a pulse oximeter show the oxygen saturation in a patient's blood [2]. However, despite the unique capabilities of these WBAN biomedical devices, they face many challenges owing to their size as they have limited resources such as battery power, memory, and bandwidth. Among the scarce resources of biomedical devices, the energy resource is a crucial resource on which most activities of other biomedical devices are dependent [3]. For instance, during data transmissions, the communication module consumes more energy than the other biomedical devices modules. Moreover, the WBAN devices share a common channel for communication which is managed or regulated by medium access control (MAC) protocol. One way to manage these limited

Manuscript received May, 2023. "This research work was supported by the Council for Scientific and Industrial Research, Pretoria, South Africa through the Smart Networks collaboration initiative and IoT-Factory Program (Funded by the Department of Science and Innovation (DSI), South Africa)."

D.D Olatinwo is with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria 0002, Pretoria, South Africa (Corresponding author: damibaola@gmail.com).

A.M. Abu-Mahfouz is with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria 0002, Pretoria, South Africa, and also with the Council for Scientific and Industrial Research (CSIR), Pretoria, 0184, South Africa (e-mail: a.abumahfouz@ieee.org).

G.P. Hancke is with the Department of Computer Science, City University of Hong Kong, Hong Kong SAR (e-mail: gp.hancke@cityu.edu.hk).

Hermanu C. Myburgh is with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria 0002, Pretoria, South Africa. (e-mail: herman.myburgh@up.ac.za).

resources and prolong the lifespan of the biomedical devices is by designing efficient MAC protocols to prevent energy wastage, time-slot wastage, and improve channel utilization [4], [5], and [6]. To achieve this, researchers leverage the optimization of MAC protocols to meet WBANs quality-of-service (QoS) requirements such as energy efficiency, delay, reliability, throughput, and a prolonged network lifetime.

In recent times, through research efforts, different solutions have been proposed to address energy consumption problems by using MAC protocols [7] - [11]. However, existing solutions are yet to fully address these issues. In the literature, most classic WBAN MAC protocols are designed to operate on a single channel, i.e., biomedical devices only have one single channel available for all their communications [12] - [15]. In such systems, determining the mode in which the devices access the channel becomes difficult and could greatly affect the performance of the WBAN system in terms of efficiency and reliability resulting in energy wastage, time-slot wastage, and channel utilization issues.

Various wireless communication standards such as IEEE 802.15.6 [16], ESTI Smart-BAN [17], and IEEE 802.15.4 [18] have been employed to meet the diverse requirements and applications of WBANs. These standards were specifically designed to address the unique needs of wireless communication systems. IEEE 802.15.6 focuses on WBANs, while ETSI SmartBAN is tailored for medical device communication.IEEE 802.15.4, on the other hand, was developed for devices that require low-cost and low-data-rate connectivity. Among these three standards, IEEE 802.15.4 is considered the most fully developed short-range standard with a wide range of applications in healthcare WBANs. To overcome the constraints imposed by the energy scarcity and processing power of biomedical devices, IEEE 802.15.4 incorporates appropriate physical and MAC layers for battery-operated devices. It also supports the design of mechanisms such as time-slotted access and multi-channel communication to enhance WBAN performance [19]. Based on the IEEE 802.15.4 standard, this study proposes a hybrid multi-channel MAC protocol with Markov decision process (MDP-HYMAC) to improve energy efficiency, throughput, network lifetime, and minimize delay. The proposed protocol aims to reduce data collisions and beacon between the biomedical devices and APs by analyzing collision patterns and efficiently allocating frequencies over time to prevent neighboring nodes from colliding with each other. The proposed MDP-HYMAC architecture consists of different WBAN scenarios integrated with edge AI architecture composed of multi-access edge computing (MEC) devices. For each WBAN, a heterogeneous scenario that consists of various types of devices with different roles and capabilities was considered. Some of these devices act as ordinary biomedical sensor nodes that can only collect and send health data to an access point (AP) which can perform an edge AI task. Consequently, the energy consumption and all the computational overhead that would have been imposed on each biomedical device by edge AI based on its complex tasks related to data collection, processing, analysis, and decision-making were shifted to the AP side.

In the proposed MDP-HYMAC protocol, separate channels are used for data transmission and control commands. One channel is used to send control commands and the remaining channels are used by biomedical devices for data transmission. Using a separate channel helps to minimize collisions among devices, reduce delay, improve energy efficiency, throughput, and thereby improve the overall system performance. However, the problem of channel utilization is a big challenge. To address this, channel utilization was modeled using Markov decision process and employed dynamic programming method to find optimal policy to ensure efficient channel allocation leading to improved data transmission, minimized packet drop-off or loss, and prolonged the devices lifetime. In addition, the problems of time-slot wastage, energy wastage, and delay were addressed by introducing adaptive power allocation, time-slot, and back-off period schemes to enhance the energy efficiency and reliability of WBAN networks. The following are the contributions of the proposed hybrid multi-channel MAC protocol with Markov decision process (MDP-HYMAC):

- Design and development of a hybrid multi-channel MAC protocol with Markov decision process to improve energy efficiency, throughput, network lifetime, and minimize delay. Integration of edge AI with IoT-enabled WBAN system to facilitate near real-time communication.
- Adaptive power allocation, time-slot, and back-off methods were proposed to reduce time-slot wastage, energy wastage, and delay.
- A Markov decision process (MDP) was used to determine the traffic arrival pattern, model the transition states of biomedical devices, the channel status and the buffer status to prevent congestion, improve energy efficiency, and the lifetime of the network.
- A dynamic programming method was used to solve the channel utilization problem by finding the optimal policy for channel allocation.
- Novel strategies, such as employing wake-up radio, introducing a queuing state, and shifting major overhead transmissions to the AP, were proposed to minimize delay, packet drop ratio, increase the network lifetime, and enhance energy efficiency without compromising the throughput of IoT-enabled WBAN systems.
- Lastly, the proposed MDP-HYMAC protocol outperformed other protocols such as MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 protocols using similar methods.

The structure of this paper is as follows. Section I presents the introduction of the paper. Section II presents a review of the literature of some existing MAC protocols. Section III presents a discussion on the description of the proposed MDP-HYMAC protocol. In Section IV, the proposed method is presented. Section V presents the Markov decision model, while section VI presents the dynamic programming method. The performance evaluation of the proposed protocol while section VIII presents the results and simulation. Section IX concludes the paper.

II. Related Studies

This section begins with an overview of several research articles on energy-efficient WBAN systems that utilize single-

channel MAC protocols. Then it delves into the design of multi-channel MAC protocols for WBAN systems. For example, a single channel coordinated superframe duty cycle hybrid MAC protocol was proposed to improve energy efficiency and prolong the lifetime of biomedical devices [6].

Similarly, Olatinwo et al. [12] proposed a MAC protocol based on a single channel concept. The protocol utilized a transmission scheduling mechanism to improve network energy efficiency by duty-cycling device operations.

Thirumoorthy et al. [20] proposed an energy-efficient distributed queueing MAC protocol for WBANs. The protocol employed a distributed queuing technique to enhance radio channel utilization. However, the paper focused on a homogeneous-based WBAN system.

Sun et al. [21] proposed a MAC protocol based on device priority to enhance WBAN efficiency. These devices were prioritized according to their degree of importance, timeout condition, remaining energy, and sampling rate. Adjustments were made to the total number of device time-slots and conflicting time-slots to enhance the average packet delivery rate. Also, a time-slot allocation algorithm was proposed to enhance network performance.

Various energy conservation strategies were employed to improve the performance of WBAN systems [22]. These strategies included shifting major overhead transmissions to a personal server, introducing a waiting order state, and enabling the retransmission process at the end of a cycle after all transmissions are completed.

Mkongwa et al. [23] proposed a single channel MAC protocol for WBANs to improve the energy efficiency of the network. The protocol used a clear channel assessment (CCA) algorithm and a back-off mechanism to minimize the device contention period, which resulted in a reduction in delay.

In contrast to [3], [12], and [20] - [23], a hybrid multi-channel MAC protocol with Markov decision process was proposed to improve energy efficiency, system throughput, network lifetime, and minimize delay. To enhance channel utilization efficiency, the transition states of the system were stochastically modeled using MDP to explore optimal strategies for communication between devices and APs. In addition, to minimize time-slot, energy wastage, and delay, an adaptive time-slot allocation scheme and a back-off strategy were proposed.

It is important to mention that using a single-channel MAC protocol might not be efficient in resolving concerns such as collisions, energy wastage, and delay, WBANs. Consequently, a comprehensive review of relevant research endeavors that employed multi-channel MAC protocols to enhance the WBAN systems was conducted. For instance, Olatinwo et al. [24] proposed a multi-channel protocol to address issues related to energy consumption, time-slot management, delay, and channel utilization in WBAN systems. For this to be achieved, a time-slot management scheme, channel mapping, channel selection mechanism, and a back-off time policy was proposed to improve energy efficiency, channel utilization, packet delivery ratio, devices lifetime, and reduce delay.

Samal et al. [25] proposed a priority-based traffic scheduling MAC protocol to minimize packet drop and improve system throughput. To address energy wastage in WBAN systems, Li et al. [26] proposed an energy-efficient interference-aware multi-channel MAC (EI-MAC) protocol that employed a channel mapping scheme to analyze channel states and mitigate interference using a collision avoidance technique. However, power control and allocation schemes, time-slot management schemes, and re-transmission mechanisms were not put in place.

To address interference and minimize delay in WBANs, Li et al. [27] proposed a multi-channel MAC protocol. The protocol employed a channel mapping technique to check the availability of the channel to enhance the system's performance. Rasheed et al. [28] proposed a modified superframe structure (MSS-IEEE 802.15.4) to address energy consumption and delay issues in the WBAN. The structure employed a priority-based CSMA/CA mechanism to allocate different priorities to nodes by adjusting their data size and type.

Lastly, Le and Moh [29] improved energy efficiency, throughput, and delay of WBAN systems by designing a hybrid multi-channel MAC protocol that employed a channel selection scheme to prevent collision.

Unlike previous studies such as [23] - [29], this study proposes a heterogeneous based hybrid multi-channel MAC protocol with MDP to reduce collisions, energy and time-slot wastage, delay, and improve the lifetime of the devices. It utilizes an MDP to determine the traffic arrival pattern, model the transition states of biomedical devices, channel status, and buffer status to prevent congestion and improve energy efficiency and network lifetime. A dynamic programming method was used to solve the channel utilization problem and find the optimal policy. In addition, adaptive power and time-slot allocation schemes were proposed. Table I presents a summary of the existing protocols, their strengths, and limitations.

III. Description of the Proposed MDP-HYMAC Protocol

The IEEE 802.15.4 standard has two modes of operation: beacon and non-beacon. The proposed protocol adopts the IEEE 802.15.4 beacon mode, in which beacons are used between the APs and the devices for synchronization. It also uses a superframe structure that has two phases: active and inactive. As shown in Fig. 1, BI represents the beacon interval between two consecutive beacons and is computed as $BI = aBaseSuperframDuration \cdot 2BO$ where BO is the beacon order. The superframe duration is computed as $SD = BaseSuperframDuration \cdot 2SO$. The SD consists of 16 slots and is divided into two phases including the contention access phase (CAP) and the contention-free phase (CFP).

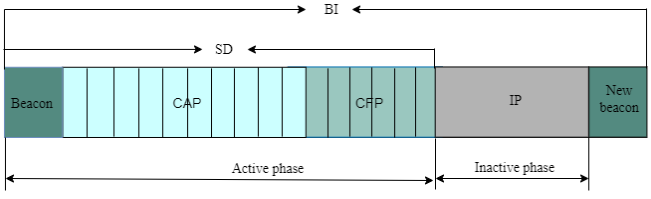


Fig. 1. MDP-HYMAC superframe structure

TABLE I
COMPARISON OF EXISTING METHODS

Reference	Method	Type of network	Performance metric	Strength	Limitation
[3]	Coordinated superframe duty cycle scheme based on devices' traffic information and priority	Single channel	Energy efficiency, delay, packet drop ratio, and devices' lifetime	Enhanced energy efficiency and prolong the devices' lifetime	Multiple re-transmissions
[12]	Transmission scheduling mechanism	Single channel	Energy efficiency, devices' lifetime, and convergence speed	Energy conservation	Packet loss and multiple re-transmissions
[20]	Sleep–wake mechanism with priority queuing model	Single channel	Delay and energy efficiency	Less delay	Resource allocation mechanism not considered
[21]	Time-slot allocation scheme	Single channel	Delay, delivery ratio, and energy efficiency	Enhanced data transmission	Energy conservation and channel utilization mechanisms not considered
[22]	Energy conservation strategies	Single channel	Energy efficiency and devices' lifetime	Minimized energy wastage	Time-slot wastage, multiple re-transmission, and packet
[23]	CCA algorithms and back-off method	Single channel	Energy consumption, throughput, packet delivery ratio, and delay	Enhanced energy efficiency, throughput and minimized delay	Considered single channel
[24]	Heuristic-based power control scheme and channel selection mechanism	Multiple channel	Energy-efficiency, throughput, delay, packet delivery ratio, and network lifetime	Improved energy efficiency and packet delivery ratio	Channel mapping and selection mechanisms may not be efficient enough to address channel utilization issue
[25]	Traffic prioritized load balanced scheduling scheme	Multiple channel	Delay, throughput, and energy efficiency	Minimized packet drop ratio and enhanced throughput	Multiple re-transmission, inefficient channel utilization
[26]	Channel selection mechanism and low energy conservation mechanism	Multiple channel	Delay, throughput, and energy consumption	Improved Energy efficiency	Inefficient channel selection mechanism
[27]	Channel mapping technique	Multiple channel	Delay, throughput, packet error rate and frame error rate	Minimized delay and enhanced throughput	Energy conservation mechanism not considered.
[28]	Wake-up radio-based mechanism	Multiple channel	Energy consumption, throughput, packet drop probability, and average delay	Improved energy efficiency and minimized delay	Channel selection mechanism not considered
[29]	Channel mapping mechanism	Multiple channel	Energy consumption, throughput, packet delivery ratio, and delay	Prevent collisions, reduced delay	Channel mapping mechanism may not be efficient for heterogeneous WBAN system.

At the start of each cycle, the superframe structure of the system begins with a beacon message from the AP which includes the address of the AP and devices, as well as the beginning and end of each phase. The devices in the network are presumed to be in sleep mode awaiting a ready-to-receive (RTR) beacon message from the AP. A wake-up radio is used to turn the devices on and off [30], [31] and this enhance energy efficiency. The wake-up radio operates by switching on the main radio of a device when an incoming signal is sensed and promptly switches the device to an active mode.

In CAP, $\mathcal{N}$ and $\mathcal{Q}$ devices contend to transmit their H-Info by applying CSMA/CA scheme using Algorithm 1. Each successfully contended H-Info contains unique information, including the device ID that is used during transmission. After the AP receives the H-Info, a total acknowledgment (T-ack) message is sent at the end of the CAP instead of after each received health packet. To reduce device waiting time delay and save energy, the T-ack message is sent at the end of CAP instead of after each health packet received. To enable efficient transmission and energy conservation, a queuing order q_{order} state is introduced. Only the devices' synchronous clock is enabled to operate in the q_{order} state, and all other operations are disabled. The devices enter the q_{order} state based on priority. Therefore, in case of critical event occurrence (i.e., $\mathcal{E}_t = 2$), devices with critical data are allowed to transmit first. An M/M/1 queuing method is adopted using the first-in-first-out (F/I/F/O) strategy to model the arrival pattern and the service pattern of the devices.

Devices in the q_{order} state are only activated to other active states by using an active beacon and the ID of the device. After each transmission process (TP), the devices employ the TDMA scheme to send their packets and send an end beacon to the AP. On the other hand, the AP sends a beacon message to the devices after receiving their health packets. The beacon message has an order ack (O-ack) to activate the next device. The next device in q_{order} starts its transmission as soon as an

O-ack message is received. Conversely, if transmission fails, then an O-ack message will not be received. Failed health packet transmission can happen due to packet loss, interference, or congestion in the network. Therefore, to conserve energy, devices with failed packet transmission remain in the q_{order} state until after all transmission processes are completed before transmitting a retransmission (reT) beacon to get the AP ready for the re-transmission process. The operation of the proposed MDP-HYMAC protocol is presented in Fig. 2.

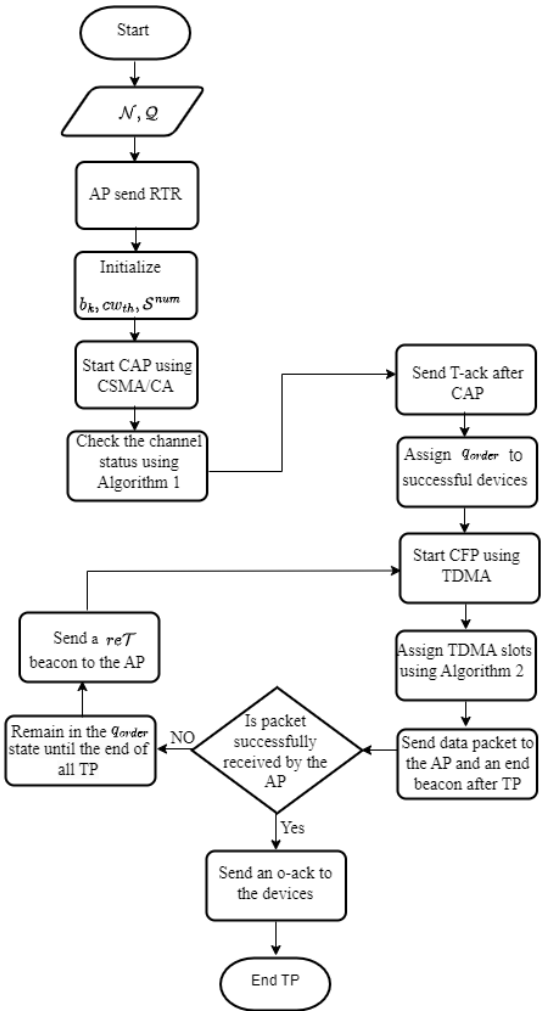


Fig. 2. Flowchart illustrating the operation of the proposed MDP-HYMAC protocol

communication processes in each WBAN are divided into three basic tiers: intra-WBAN, inter-WBAN, and beyond-WBAN. Intra-WBAN communication is established between the biomedical devices and the AP. The biomedical devices are positioned in, on, and around the patient's body. They perform sensing tasks, gather sensed health data, and send them to the AP. The AP coordinates the operation in this tier and could act as a gateway or a local processor. To minimize delays associated with the time that the AP spends providing services to devices and considering the time sensitive nature of these devices, the MEC technology was introduced. The MEC technology brings the functions of the cloud server to the edge network. It is closer to users and helps minimize transmission delays and energy consumption. The major tasks of the AP are shifted to the MEC technology which is in the next communication tier, i.e., inter-WBAN communication. The MEC is very efficient, handles computations faster because it has more computational facilities than the AP. The MEC devices receive data from the AP, analyze and process it, and then forward processed data to the cloud server, which is beyond the WBAN communication tier. Patients' processed data are stored in the cloud database, from which medical experts get patient information for decision making purposes, diagnosis, and treatment administration. The proposed MDP-HYMAC protocol architecture is shown in Fig. 3.

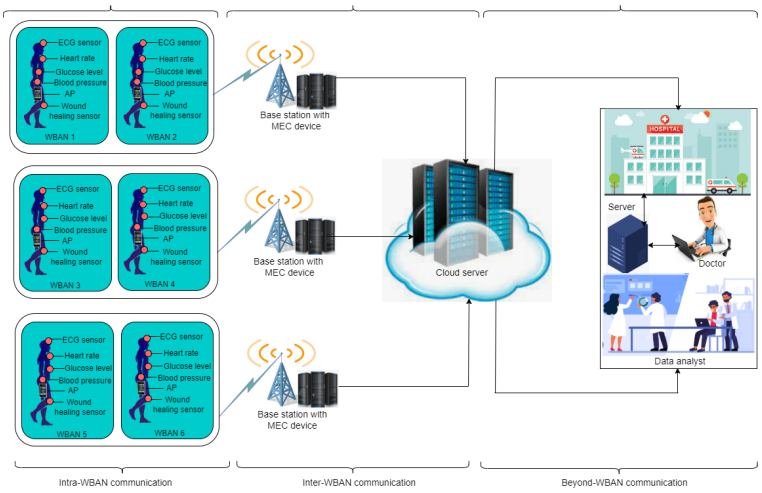


Fig. 3. Proposed MDP-HYMAC system architecture with MEC

IV. Proposed Method

This section presents the proposed MDP-HYMAC architecture, system and mathematical modelling, event occurrence, channel status, buffer status, time-slot management scheme, back-off strategy, and adaptive power allocation scheme in the subsequent subsections.

A. System Architecture

The proposed MDP-HYMAC architecture in this study consists of $\mathcal{K}$ WBANs, $\mathcal{G}$ APs, and $\mathcal{D}$ biomedical devices that transmit their health packets in $\mathcal{M}$ different channels. The

B. System and Mathematical Modeling

The MDP-HYMAC system consists of a total number of WBANs modeled as a set of $\mathcal{K} = \{k_1, k_2, k_3, ..., k_K\}$. There are also a total number of $\mathcal{G} = \{g_1, g_2, g_3, ..., g_G\}$ APs and $\mathcal{D} = \{d_1, d_2, d_3, ..., d_D\}$ biomedical devices. The biomedical devices in all the $\mathcal{K}$ WBANs send their health packets using different channels modeled as $\mathcal{M} = \{m_1, m_2, m_3, ..., m_M\}$. Consequently, for all the WBANs in the proposed MDP-HYMAC, the first channel in a set of $\mathcal{M}$ is used by all the APs to transmit the control commands from the MEC to the biomedical devices. The biomedical devices are assigned the remaining $\mathcal{M}-1$ channels for communication. Assuming that

$\mathcal{W}$ of the total $\mathcal{M}$ channels are dedicated to a WBAN, the AP uses the first channel $\mathcal{W}$ as the control channel to send control commands to devices. The devices are assigned the remaining $\mathcal{W}-1$ channels as data channels to send their health packets to the AP. In each WBAN, biomedical devices are dynamically categorized based on priority and payload size into two classes such as class 1 $(\mathcal{C}_1)$ and class 2 $(\mathcal{C}_2)$. The $\mathcal{C}_1$ devices are assumed to have less-critical health packets and are modeled as a set of $\mathcal{X} = \{x_1, x_2, x_3, ..., x_X\}$, $\forall\ x \in \mathcal{D}$ while the $\mathcal{C}_2$ devices are assumed to have critical health packets and are modeled as a set of $\mathcal{Y} = \{y_1, y_2, y_3, ..., y_Y\}$, $\forall\ y \in \mathcal{D}$. The less-critical health packets are normal data that are delay tolerant, whereas the critical health packets are emergency data that are delay intolerant.

In addition, not all devices on the network have packets to transmit. Only devices that require channel access to transmit data are assigned channels, while other devices enter a low-power sleep mode to conserve energy. It was assumed that devices with critical health packets do not always have data packets to transmit. In addition, channel resources such as energy, time-slots, and bandwidth are allocated to devices based on their priority levels (Φ) expressed using Eqn. 1.

$$\Phi = \frac{\mathcal{D}^T}{\lambda \mathcal{L}} \tag{1}$$

where $\mathcal{D}^T$ represents the type of data, λ represents the traffic rate, and $\mathcal{L}$ represents the packet length. Consequently, in each WBAN, all C_1 and C_2 devices that have health packets to transmit apply the CSMA / CA scheme to contend for transmission opportunities to send their health information (H-Info) to the AP. The H-Info does not contain the actual intended health data, i.e., the payload. Afterward, the successfully contended devices are assigned transmission slots using the TDMA scheme. It is important to mention that if there are no available data channels and the control channel is free then, the C_2 devices can use the control channel to transmit their health packets.

C. Fundamental Requirements of MDP and Assumptions

The fundamental requirements for applying MDP in this study are as follows:

1) States: MDPs require well-defined states representing the system's conditions. In the context of WBANs, these states could represent channel conditions, energy levels, or other relevant parameters.
2) Actions: MDPs involve sequential decision-making, where actions are taken based on the current state of the system. In WBANs, these actions could include channel selection, data communication, and sleep/wake decisions.
3) Transition Probabilities: These probabilities represent the likelihood of transitioning from one state to another when an action is taken. In a WBAN system, transitions depend on various factors including channel conditions, interference, and energy constraints.
4) Rewards: MDPs use rewards to quantify desirability of different actions. In a WBAN system, rewards could represent energy efficiency, successful transmission delivery, collision avoidance, back-off strategy, and the overall system performance.
5) Discount Factor: The discount factor plays a crucial role in balancing immediate rewards with long-term goals. In WBAN context, discount factor could help to achieve a trade-off between immediate rewards (e.g., high throughput) and long-term objectives (e.g., energy efficiency).

Based on these insights, the proposed WBAN MAC protocol and MDP assumptions are as follows:

- The proposed protocol leverages MDP principles and considers channel status when allocating channels to biomedical sensor.
- The reward value guides decisions, e.g., optimizing energy efficiency or minimizing collisions.
- By interacting with the environment, the protocol generates an optimal channel allocation strategy.
- The devices are expected to perform two types of operations: transmitting health packets to the AP and receiving control commands from the AP.
- The network devices are presumed to maintain a constant power level for a specific state, but they employ varying power levels across distinct states.
- The devices utilized a sense-and-transmit approach.
- The traffic arrival pattern is modeled as a Poisson process.

D. Event Occurrence

As shown in Figs. 4 and 5, the generation of less-critical or critical data packets for transmission is not certain, as this depends on a patient's physical condition. Assume a device in the network in each time-slot t has a health packet of a given priority to transmit then this implies an event occurrence denoted as $\mathcal{E}_t$ such that $\mathcal{E}_t \in \{0, 1, 2\}$, where $\mathcal{E}_t = 0$, $\mathcal{E}_t = 1$, $\mathcal{E}_t = 2$, denotes an event occurrence, a less-critical event occurrence, and a critical event occurrence, respectively. Each device transmits at most one health packet at every time interval. Consequently, the system checks periodically for an event (i.e., $\mathcal{E}_t = 1$, $\mathcal{E}_t = 2$), but there is a temporal relationship between the generation of such data packets. If $\mathcal{E}_t = 1$ in a current time-slot t, then the occurrence of $\mathcal{E}_t = 1$ in the next slot t + 1 is uncertain. This can be modeled by the conditional probability $\mathcal{P}_r$, which indicates the likelihood that another $\mathcal{E}_t = 1$ would occur in the next time slot t+1 given that $\mathcal{E}_t = 1$ occurred in the current time-slot t. However, the probability of generating either $\mathcal{E}_t = 0$ or $\mathcal{E}_t = 2$ (that is, there is no occurrence of $\mathcal{E}_t = 1$) in the next time-slot is expressed as $1- \mathcal{P}_r$. Since both occurrence and non-occurrence outcomes of a normal event can be equally likely in the next time-slot $t + 1$ given that it takes place in the current time-slot t, then the conditional probability is expressed as $0.5 < \mathcal{P}_r < 1$.

Additionally, if $\mathcal{E}_t = 1$ does not occur at the current time-slot t, then the probability that such event would occur in the next time-slot $t + 1$ is expressed as $1- \mathcal{P}_c$, where $0.5 < \mathcal{P}_c < 1$

is the conditional probability and $\mathcal{P}_c$ indicates that no $\mathcal{E}_t = 1$ occurred in the current time-slot t and also the next time-slot $t+1$. If a less-critical event occurs in successive time-slots, the probability of it occurring i times is expressed using Eqn. 2.

$$\mathcal{P}[X' = i] = (\mathcal{P}_r^{i-1})(1 - \mathcal{P}_r) \tag{2}$$

Based on Eqn. 2, the average period duration (A_T) for a continuous less-critical event (i.e., $\mathcal{E}_t = 1$) is expressed using Eqn. 3.

$$A_T[X'] = \sum_{i=1}^{\infty} (\mathcal{P}_r^{i-1})(1 - \mathcal{P}_r) = \frac{1}{1 - \mathcal{P}_r} \tag{3}$$

while the average period duration without a less-critical event is expressed using Eqn. 4.

$$A_T[X'] = \frac{1}{1 - \mathcal{P}_c} \tag{4}$$

Similarly, the occurrence of critical events ($\mathcal{E}_t = 2$) is modeled as a two-state process with probabilities $\mathcal{P}_a$ and $\mathcal{P}_b$ where $0.5 < \mathcal{P}_a, \mathcal{P}_b < 1$. Therefore, if a less-critical event occurs in a current time-slot t but then a critical event occurs in the next time-slot $t+1$, the input parameters will be $(1-\mathcal{P}_r)+(1+\mathcal{P}_b)$.

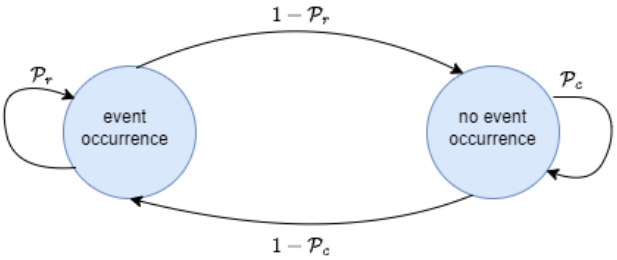


Fig. 4. Conditional probability for less-critical event occurrence .

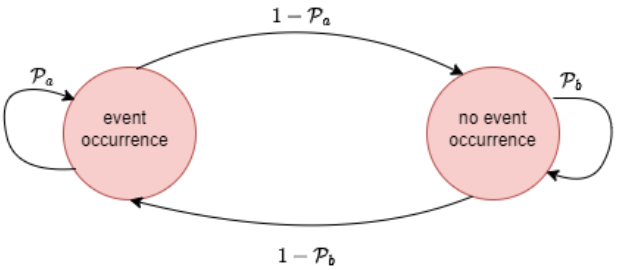


Fig. 5. Conditional probability for critical event occurrence .

E. Channel Status

The channel status of the proposed system is determined using a channel mapping strategy such that all biomedical devices in the network use the strategy to access the channel. For each WBAN, the first $\mathcal{M}$ channel is used by the AP as a control channel and the $\mathcal{M}_1$ remaining channels are used for data transmission by the biomedical devices. Consequently, $\mathcal{M}$ channels are allocated based on a sequence number ranging from 1 to $\mathcal{M}$. The AP checks the channel status before the commencement of any communication to find out if the channel is free or busy. Following this, the AP makes a list of the channel status. In this list, it was assumed that 1 and 0 represent that the reference channel is free or busy, respectively, as shown in Fig. 6.

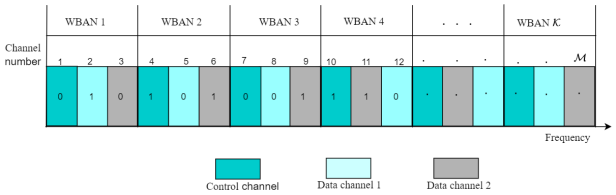


Fig. 6. MDP-HYMAC channel plane

To avoid collisions, the AP updates the channel list after allocating channels to the devices [32]. The mapping matrix $\mathcal{CH}$ used to access the channel is represented in Eqn. 5.

$$\mathcal{CH}_t = \begin{cases} 0 & if\ channel\ is\ busy, \\ 1 & if\ channel\ is\ free \end{cases} \quad \forall\ \mathcal{CH}_t \in \mathcal{M} \tag{5}$$

If $\mathcal{CH}_t$ is 1 at a given time-slot t, then the probability of having a free channel at the next time-slot $\mathcal{CH}_{t+1}$ is denoted by $\mathcal{CH}_{free}$ such that $0.5 < \mathcal{CH}_{free} < 1$ and the probability of having a busy channel at the next time-slot $\mathcal{CH}_{t+1} = 1 - \mathcal{CH}_{t+1}$. Similarly, if $\mathcal{CH}_t$ is 0 in a current time interval t then the probability of having a busy channel in the next time interval $\mathcal{CH}_{t+1}$ is denoted by $\mathcal{CH}_{busy}$ while the probability of having a free channel in the next time interval $\mathcal{CH}_{t+1} = 1 - \mathcal{CH}_{busy}$ where $0.5 < \mathcal{CH}_{busy} < 1$.

F. Buffer Overflow Analysis

APs generally have limited buffer space. In the proposed MDP-HYMAC system, the buffer status of the AP is considered before transmission to prevent congestion, which could lead to packet loss or drop-off issues. To model the buffer status of the AP, we set a threshold denoted by $\mathcal{B}_{th}$. Therefore, at a given time-slot t the buffer status denoted by $\mathcal{B}^s$ and modeled as a two-state process in Eqn. 22.

$$\mathcal{B}_t^s = \begin{cases} 1 & if\ \mathcal{B}^s < \mathcal{B}_{th},\ \ no\ buffer\ space \\ 0 & if\ \mathcal{B}^s \geq \mathcal{B}_{th},\ \ buffer\ has\ space \end{cases} \quad \forall\ \mathcal{B}_t^s \in \mathcal{G} \tag{6}$$

Therefore, if $\mathcal{B}_t^s = 1$ in the current time-slot t, then the probability that the buffer does not have space in the next time-slot $t+1$ is denoted by ω and the probability of having space in the buffer is $1 - \omega$ where $0.5 < \omega < 1$. Conversely, if $\mathcal{B}_t^s = 0$ in the current time-slot t, then the probability of having space in the buffer in the next time-slot $t+1$ is denoted by ϖ and the probability of not having space in the buffer is $1 - \varpi$ where $0.5 < \varpi < 1$.

G. Back-off Strategy

The proposed MDP-HYMAC protocol is designed based on the IEEE 802.15.4 standard with four major phases such as the contention access phase (CAP), contention free phase (CFP), active period, and inactive period (IP). The CSMA/CA protocol is employed as a collision avoidance scheme to prevent repeated periodic collisions and the TDMA protocol is used to allocate time-slots to the devices. When a collision occur in the CAP, the devices perform a random back-off and contend to access the channel again. The back-off period determines the probability of devices accessing the channel. A shorter back-off time increases the probability of contending for a channel. In contrast, it also increases the number of retransmissions. Therefore, it is important to efficiently design a back-off period scheme. It is important to mention that the traditional back-off schemes adopt an exponential method and this increases the collision probability as explained in [33]. Therefore, a new back-off period strategy is proposed for the biomedical devices by introducing a contention threshold value (cw_{th}). Assuming a bit error rate of 0, then, cw_{th} is expressed in Eqn. 7.

$$cw_{th} = \frac{1}{2}\left(cw_{min} + cw_{max}\right) \tag{7}$$

Based on the IEEE 802.15.4 standard, a WBAN is configured with a $\delta = 5$ where δ is a random value ranging from $0 \geq 5$ which enables five number of back-off slots ranging from 0 to 31, i.e., $0-1$, $0-15$, $0-31$, $0-31$, and $0-31$ [33]. This indicates that the proposed back-off scheme enables the devices to contend for a channel up to five times. Assuming the back-off period is uniformly distributed, then, the optimal time for accessing the channel can be calculated. Consequently, the back-off time-slot as well as the mean value of the back-off period (C_o) is modeled in Eqn. 8 and Eqn. 9, respectively.

$$cw = 2^{\delta} - 1 \tag{8}$$

$$E|C_o| = \frac{1}{e+1}\sum_{n=0}^{e} n \tag{9}$$

Assume that three collisions are averaged in a WBAN, then e is determined using Eqn. 10

$$e = 2^{\delta} - 1 = 2^3 - 1 = 7 \tag{10}$$

Substitute Eqn. 10 in Eqn. 9 to give Eqn. 11 and Eqn. 12.

$$E|C_o| = \frac{1}{e+1}\sum_{n=0}^{e} n = \frac{1}{7+1}\sum_{n=0}^{7}(0+1+...+7) \tag{11}$$

$$\therefore\ E|C_o| = E(3) \approx 3 \tag{12}$$

Therefore, back-off time-slot cw_{th} corresponds to the optimal value given in (4). Algorithm 1 presents the back-off period strategy. Note, CCA denotes clear channel assessment.

Algorithm 1 Proposed MDP-HYMAC Back-off Period Strategy

Require: $\Rightarrow \mathcal{X}$ and $\mathcal{Y}$ that have data packets to transmit, back-off time-slot, δ, cw, cw_{max}, cw_{min}, cw_{th}
Ensure: $min_{\delta} = 3$, $max_{\delta} = 5$
locate the boundary slot
check channel status
if channel = $\mathcal{CH}_{free}$ **then**
    perform CCA
else channel = $\mathcal{CH}_{busy}$
    allocate $\delta = 3$
    back-off with cw_{min}
    check channel status
end if
if channel = $\mathcal{CH}_{busy}$ **then**
    back-off with cw_{max}
    allocate $\delta = 5$
    assign a default value = 2 to reset cw
    back-off
    wait for beacon ack
end if
if channel = $\mathcal{CH}_{free}$ **then**
    decrease CW by 1 until it reaches 0
end if
if either x_1 or y_1 access the channel $\mathcal{CH}_t$ successfully **then**
    assign 1 using Eqn. 5
else set back-off time as cw_{min}
end if
if either x_1 or y_1 failed to access $\mathcal{CH}_t$ **then**
    assign 0 using Eqn. 5
    go to step 2 and step 3
    repeat until a successful channel contention
end if

H. Time-Slot Management Scheme

In this study, the proposed MDP-HYMAC is a heterogeneous-based WBAN system where the devices have different data types, priorities, and data rates. These data rates varies from one device to another. For example, an electromyography (EMG) sensor has 1536 kbps data rate, a heart rate sensor has 2.4 Kbps data rate, a temperature sensor has 1 kbps data rate, and an electrocardiography (ECG) sensor has 192 kbps data rate. Following this, using Eqns. 13, 14, and 15, the MEC calculates the number of slots to allocate to each device according to their data rate [34] and sends the computation details to the AP to prevent time-slot wastage.

$$\mathcal{L}^s = \frac{\mathcal{D}^r}{\mathcal{N}^s} \tag{13}$$

$$\mathcal{L}^s/\mathcal{F}^r = \frac{\mathcal{L}^s}{50\ \mathcal{F}^r/sec} \tag{14}$$

$$\mathcal{S}^{num} = \left\lceil \frac{\mathcal{L}^s}{\sigma} \right\rceil \tag{15}$$

where $\mathcal{D}^r$, $\mathcal{L}^s$, $\mathcal{N}^s$, $\mathcal{L}^s/\mathcal{F}^r$, S^{num}, and σ represents the data rate, symbol number, symbol length, number of symbols per frame, slot number, and number of symbols per slot, respectively. Consequently, based on Eqns. 13, 14, and 15, 32 slots are allocated to the EMG sensor, ECG sensor is allocated 4 slots while heart rate and temperature sensors are allocated 1 slot, respectively. Then, the values of the time-slot are stored as an array by the AP for each device. In case of critical and less-critical event occurrences i.e., $\mathcal{E}_t = 2$ and $\mathcal{E}_t = 1$, the AP assigned a high priority and a low priority, respectively, using (1). The time-slot allocation scheme is presented in Algorithm 2.

Algorithm 2 Time-Slot Allocation Scheme

Require: $\mathcal{X} = x_1, x_2, x_3, ..., x_X$ and $\mathcal{Y} = y_1, y_2, y_3, ..., y_Y$
Ensure: $\mathcal{D}^r$, $\mathcal{L}^s$, $\mathcal{N}^s$, $\mathcal{L}^s/\mathcal{F}^r$, S^{num}, σ
for each x **do**:
    determine the priority Φ
    calculate $\mathcal{L}^s$, && $\mathcal{L}^s/\mathcal{F}^r$, using Eqns. 13 and 14
    use Eqn. 15 to compute an optimal time-slot
    store the time-slot values $\forall\ x \in \mathcal{X}$
end for
for each y **do**
    determine the priority Φ
    calculate $\mathcal{L}^s$, && $\mathcal{L}^s/\mathcal{F}^r$, using Eqns. 13 and 14
    use Eqn. 15 to compute an optimal time-slot
    store the time-slot values $\forall\ y \in \mathcal{Y}$
end for

I. Power Allocation Scheme

The amount of power allocated to each biomedical sensor device based on the action $a'_t \in \{s'_0, s'_1, s'_2, s'_3, s'_4\}$ and the time spent by the devices to perform the action a'_t is optimized by setting a time constraint (T_{th}) to assign different times to biomedical devices in different states using Eqn. 16.

$$\mathcal{T}_{th} = \mathcal{T}_{s'_0} + \mathcal{T}_{s'_1} + \mathcal{T}_{s'_2} + \mathcal{T}_{s'_3} + \mathcal{T}_{s'_4} = 1 \tag{16}$$

The power spent by the devices for each action performed is expressed using 17.

$$a'_t = \begin{cases} 0 \leq \mathcal{T}_{th} \leq 1 \\ \\ \mathcal{P}_{min} \leq \varrho < \mathcal{P}_{max} \end{cases} \quad \forall\ a'_t\ \in \mathcal{X}, \mathcal{Y} \tag{17}$$

V. Markov Decision Model Formulation

Markov decision process (MDP) is an optimization model for making decision under uncertainty. It is used to model a stochastic decision making process where an agent, i.e., the biomedical devices interacts with a system or an environment i.e., the MDP-HYMAC. The system remains in a particular state $\mathcal{S}'_t$ while the agent chooses an action $\mathcal{A}'_t$ per decision time. For each WBAN, MDP was employed to model the interaction between a biomedical device and the MDP-HYMAC system. The system process is modeled as a five-tuple: $(\mathcal{S}'_t, \mathcal{A}'_t, \mathcal{P}', \mathcal{R}', \gamma)$ where

- $\mathcal{S}'_t$ represents a set of finite state s'_t at any time instant t such that $s'_t \in \mathcal{S}'_t$.
- $\mathcal{A}'_t$ is a set of finite actions a'_t such that $a'_t \in \mathcal{A}'_t$.
- $\mathcal{P}'$ is the transition probability from one state to another state.
- $\mathcal{R}'$ is the reward obtained after performing an action a'_t such that $r' \in \mathcal{R}'$.
- γ is the discount factor.

A. Modeling the System States and Actions

The main goal of the MDP is to determine an optimal strategy that either minimizes or maximizes a specific objective function. In this study, the proposed system is modeled with five states, and the current state of the devices can be any of these states. These states are represented as $\mathcal{S}'_t = \{s'_{t_0}, s'_{t_1}, s'_{t_2}, s'_{t_3}, s'_{t_4}\}$ where s'_{t_0}, s'_{t_1}, s'_{t_2}, s'_{t_3}, s'_{t_4} denotes the sleep state, the idle state, the sensing state, the receiving state and the transmission state. Consequently, the biomedical devices states and the different actions performed are shown in Fig 7.

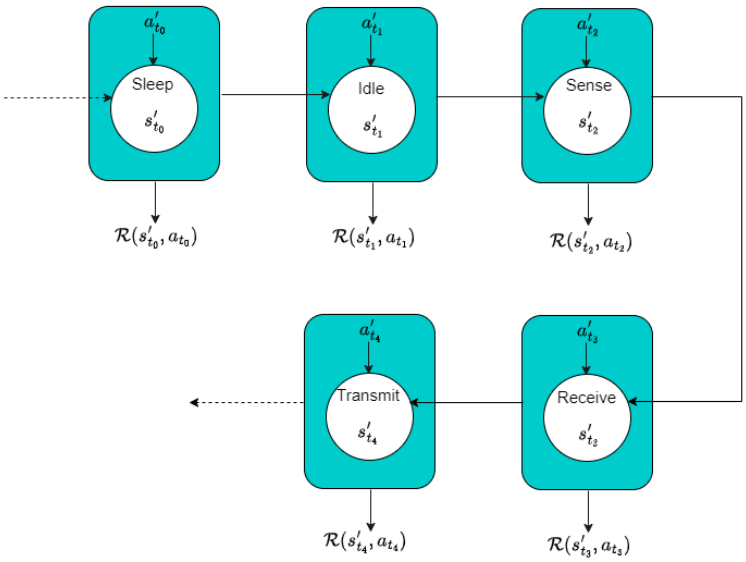


Fig. 7. Markov decision process of the proposed system illustrating the five states and their associated actions

where the major actions performed by the agents are:

- Sensing action: This action instructs the biomedical device to sense the channel.
- Receiving action: This action instructs the biomedical devices to receive control commands from the AP.
- Transmission action: This action instructs the biomedical devices to transmit data to the AP.

The state of the system at any time t is represented as a combination of three state parameters modeled in Eqn. 18.

$$\mathcal{S}'_t = \{\mathcal{E}_t, \mathcal{CH}_t, \mathcal{B}^s_t\} \tag{18}$$

where $\mathcal{E}_t$, $\mathcal{CH}_t$, $\mathcal{B}_t^s$, represents the occurrence of an event, the channel status, and the buffer overflow status. Moreover, the transition of a system's state from s_t to s_{t+1} is represented using Eqns. 19, 20, 21, and 22.

$$\mathcal{S}'_{t+1} = \{\mathcal{E}_{t+1}, \mathcal{CH}_{t+1}, \mathcal{B}^s_{t+1}\} \tag{19}$$

where $\mathcal{E}_{t+1}$ is predicted as:

$$\mathcal{E}_{t+1} = \begin{cases} 2 & w.p.\ [(\mathcal{E}_t \cdot \mathcal{P}_a \cdot 0.5) + (1-\mathcal{E}_t)(1-\mathcal{P}_b)], \\ 1 & w.p.\ [(\mathcal{E}_t \cdot \mathcal{P}_r \cdot 0.5) + (1-\mathcal{E}_t)(1-\mathcal{P}_c)] \\ 0 & otherwise \end{cases} \quad \forall\ \mathcal{X}, \mathcal{Y} \in \mathcal{D} \tag{20}$$

And $\mathcal{CH}_{t+1}$ can be predicted as:

$$\mathcal{CH}_{t+1} = \begin{cases} 1 & w.p.\ [(\mathcal{CH}_t \cdot \mathcal{CH}_{free}) + (1-\mathcal{CH}_t) \\ & \quad (1-\mathcal{CH}_{busy})] \\ 0 & otherwise \end{cases} \quad \forall\ \mathcal{X}, \mathcal{Y} \in \mathcal{D} \tag{21}$$

In addition, $\mathcal{B}^s_{t+1}$ is predicted as:

$$\mathcal{B}^s_{t+1} = \begin{cases} 1 & w.p.\ [(\mathcal{B}^s_t \cdot \omega) + (1-\mathcal{B}^s_t)(1-\varpi)] \\ 0 & otherwise \end{cases} \quad \forall\ \mathcal{X}, \mathcal{Y} \in \mathcal{D} \tag{22}$$

Note, w.p. represents with probability.

The state transition probability matrix values recorded from each current state to the next state based on the action a'_t is stored as $[\mathcal{P}']_{m\times m}$ and $[\mathcal{R}']_{m\times m}$ is the reward matrix where the generated corresponding reward r' is stored while the matrices represented in dimensions $m \times m$ denotes all possible state changes [33]. The $[\mathcal{P}']_{m\times m}$ and $[\mathcal{R}']_{m\times m}$ matrices are modeled in Eqns. 23 and 24, respectively.

$$[\mathcal{P}']_{m\times m} = \begin{array}{c} \\ s'_{t_0} \\ s'_{t_1} \\ s'_{t_2} \\ s'_{t_3} \\ s'_{t_4} \end{array} \begin{array}{c} \begin{array}{ccccc} s'_{t_0} & s'_{t_1} & s'_{t_2} & s'_{t_3} & s'_{t_4} \end{array} \\ \begin{pmatrix} p'_{00} & p'_{01} & p'_{02} & p'_{03} & p'_{04} \\ p'_{10} & p'_{11} & p'_{12} & p'_{13} & p'_{14} \\ p'_{20} & p'_{21} & p'_{22} & p'_{23} & p'_{24} \\ p'_{30} & p'_{31} & p'_{32} & r'_{33} & p'_{34} \\ p'_{40} & p'_{41} & p'_{42} & p'_{43} & p'_{44} \end{pmatrix} \end{array} \tag{23}$$

$$[\mathcal{R}']_{m\times m} = \begin{array}{c} \\ s'_{t_0} \\ s'_{t_1} \\ s'_{t_2} \\ s'_{t_3} \\ s'_{t_4} \end{array} \begin{array}{c} \begin{array}{ccccc} s'_{t_0} & s'_{t_1} & s'_{t_2} & s'_{t_3} & s'_{t_4} \end{array} \\ \begin{pmatrix} r'_{00} & r'_{01} & r'_{02} & r'_{03} & r'_{04} \\ r'_{10} & r'_{11} & r'_{12} & r'_{13} & r'_{14} \\ r'_{20} & r'_{21} & r'_{22} & r'_{23} & r'_{24} \\ r'_{30} & r'_{31} & r'_{32} & r'_{33} & r'_{34} \\ r'_{40} & r'_{41} & r'_{42} & r'_{43} & r'_{44} \end{pmatrix} \end{array} \tag{24}$$

For every action a'_t, a $[\mathcal{P}']_{m\times m}$ is generated as modeled in Eqn. 23. As a consequence, the transition probability from a current state s'_t to the next state s'_{t+1} is represented as $p(s'_t, s'_{t+1})$ such that s'_t, $s'_{t+1} \in \mathcal{S}'_t$. The state transition probability of a system can be evaluated by combining the probabilities of individual state variables. Furthermore, for every action a'_t performed, the generated corresponding reward matrix $[r]'_{m\times m}$ is modeled using Eqn. 25. In Eqns. 23 and 24 the rows, i.e., $s'_{t_0} s'_{t_1}, s'_{t_2}, s'_{t_3}, s'_{t_4}$ is the system state s'_{t_0} at a current time-slot t, respectively, and the columns represent the next state s_{t+1} in time-slot $t+1$. For example, $p_{\mathcal{E}_t}(0,1)$, $p_{\mathcal{CH}_t}(0,1)$, and $p_{\mathcal{B}^s_t}(0,1)$ are the probabilities of transition of each of the system variables to transit from the current state s'_{t_0} to the next state s'_{t_1} with action a'_t.

B. Reward Evaluation

The reward for each action performed by an agent is quantified in the context of the probability of the successful delivery of the health packets (ψ) and the generated reward $r'(s'_t, a'_t)$ assigned to carry out the action a'_t in state s'_t is expressed using [35] Eqn. 25.

$$r(s'_t, a'_t) = \begin{cases} \mathcal{CH}_{busy}(\mathcal{P}_r + (1-\mathcal{P}_c)) \cdot \psi \\ If\ \mathcal{CH}_t = 0, \mathcal{B}^s_t = 0, \mathcal{E}_t) = 1, \\ \mathcal{CH}_{busy}(\mathcal{P}_a + (1-\mathcal{P}_b)) \cdot \psi \\ If\ \mathcal{CH}_t = 0, \mathcal{B}^s_t \in (0,1), \mathcal{E}_t) = 2 \\ \mathcal{CH}_{free}(\mathcal{P}_a + (1-\mathcal{P}_b)) \cdot \psi \\ If\ \mathcal{CH}_t = 1, \mathcal{B}^s_t \in (0,1), \mathcal{E}_t) = 2 \\ \mathcal{CH}_{free}(\mathcal{P}_a + (1-\mathcal{P}_b)) \cdot \psi \\ If\ \mathcal{CH}_t = 1, \mathcal{B}^s_t \in (0,1), \mathcal{E}_t) = 2 \\ \frac{\varpi(1-\omega)}{\mathcal{CH}_{free}(\mathcal{P}_r + (1-\mathcal{P}_c))} \cdot \psi \\ If\ \mathcal{CH}_t = 0, \mathcal{B}^s_t = 1, \mathcal{E}_t) = 1 \\ \frac{\omega(1-\varpi)}{\mathcal{CH}_{free}(\mathcal{P}_r + (1-\mathcal{P}_c))} \cdot \psi \\ If\ \mathcal{CH}_t = 0, \mathcal{B}^s_t = 1, \mathcal{E}_t) = 1 \\ \frac{\varpi(1-\omega)}{\mathcal{CH}_{free}(\mathcal{P}_r + (1-\mathcal{P}_c))} \cdot \psi \\ If\ \mathcal{CH}_t = 1, \mathcal{B}^s_t = 1, \mathcal{E}_t) = 1 \end{cases} \tag{25}$$

VI. Dynamic Programming Method

Dynamic programming is a promising optimization technique that can be used to find optimal solutions to problems such as routing and multi-channel problems. In this study, the dynamic programming method is used to find an optimal policy for channel utilization issues by enhancing system performance and minimizing energy costs. A policy determines the sequence of actions that a device occupying a specific state should carry out. An optimal policy (π^*) provides appropriate actions for each state, given a specific combination of input probability values. For each action, a pair of $[\mathcal{P}']m \times m$ matrix and $[\mathcal{R}']m \times m$ matrix are provided as input to the value iteration process, subject to a given input condition that is determined by the probability values and discount factor γ. The optimal policy is obtained as the outcome. Note that the value iteration method uses the backward induction mechanism [36] to calculate the sum of discounted rewards $\mathcal{V}(s'_t)$ earned by state s'_t while following the policy in time-slot t. For each action $a'_t \in \mathcal{A}'_t$ the state value function for state s'_t satisfies the Bellman equation [37] and the stationary policy denoted as $\pi = (\pi_0, \pi_1, \pi_2, ...)$ are expressed using Eqns. 26 and 27, respectively.

$$\mathcal{V}^*(s') = \mathcal{V}^{\pi}(s'_t) = r'(s'_t, \pi s'_t) + \gamma' \sum_{s'_{t+1}} p\pi_{s'_t}(s'_t, s'_{t+1}) + \gamma' \mathcal{V}^{\pi}(s'_{t+1}) \tag{26}$$

$$\pi(s'_t) = {}^*arg \max_{a'_t \in \mathcal{A}'_t} \left[\sum_{s'_{t+1}} p(s'_t, s'_{t+1})(r'\pi_{s'_t}(s'_t, s'_{t+1}) + \gamma' \mathcal{V}(s'_{t+1})) \right] \tag{27}$$

At the beginning of the value iteration process, an arbitrary value $\mathcal{V}_0$ is assigned to each state s'_t during initialization. The process then iterates through all the states s'_t until convergence is achieved and Bellman equation is used to compute the next iteration on the state s'_t. The iteration continues until convergence $\max_{s'} |\mathcal{V}_{u+1}(s') - \mathcal{V}_u(s')| < \epsilon$. The value iteration method result to the number of iterations as well as the discounted utility values expressed using Eqn. 28.

$$\mathcal{F}\left[\gamma, (r'_0, r'_1, r'_2, r'_3, ...)\right] = r'_0 + \gamma r'_1 + \gamma^2 r'_2 + \gamma^3 r'_3 +, ... \tag{28}$$

The discount factor γ is responsible for determining how much future rewards impact the current reward estimation. Additionally, after taking the number of iterations, $\mathcal{P}'$, $\mathcal{R}'$ and γ as input, the finite-horizon method halts after a fixed number of executions. The method also ensures that a terminal state is reached for every policy [38]. The optimal policy scheme is presented in Algorithm 3. The objective of this algorithm is to obtain optimal channel resource utilization by minimizing the total energy cost while enhancing the system throughput.

Algorithm 3 Optimal Policy Scheme

Require: $\mathcal{CH}_{free}$, $\mathcal{CH}_{busy}$, $\mathcal{P}_r$, $\mathcal{P}_c$, $\mathcal{P}_a$, $\mathcal{P}_b$, ω, ϖ
Ensure: $[\mathcal{P}']$ && $[\mathcal{R}']$ matrices
for each $\mathcal{X}$ && each $\mathcal{Y}$ **do**:
  using Eqns. 20, 21, and 22, evaluate $[\mathcal{P}']$
  using Eqn.25, evaluate $[\mathcal{R}']$
  initialize $\mathcal{V}_0$
  configure $\gamma = h$, such that $h \in (0, 1)$
  initialize ϵ && configure $u = 0$
  repeat
  for each $s' \in \mathcal{S}'$ **do**:
    using Eqn. 26, evaluate $\mathcal{V}^*(s')$
    compute $u = u + 1$
    until convergence
    return $\pi^*(s')$
  end for
end for

VII. Performance Evaluation

This section evaluates the proposed system's performance in terms time and energy analysis, delivery ratio, throughput, delay, and QoE analysis. It also presents the adaptive power allocation scheme. The following subsections provide a detailed performance analysis of the system.

A. Time and Energy Analysis of the Proposed System

Assuming there are $\mathcal{N}$ and $\mathcal{Q}$ number of $\mathcal{C}_1$ and $\mathcal{C}_2$ devices, respectively, that have data packets to transmit in each WBAN, the average packet inter-arrival time for both $\mathcal{C}_1$ and $\mathcal{C}_2$ are expressed as $\mathcal{A}_T^1 = \frac{1}{\lambda \mathcal{T}}$ and $\mathcal{A}_T^2 = \frac{1}{\lambda \mathcal{T}}$, respectively. During the first CCA, it was assumed that the channel is busy when other devices transmit their health packets. Therefore, the total number of health packets served during $\mathcal{CH}_{busy}$ is represented by $\mathcal{H}_{\mathcal{N}}$ and $\mathcal{H}_{\mathcal{Q}}$, which are modeled in Eqns. 29 and 30, respectively.

$$\mathcal{H}_{\mathcal{N}} = \frac{1}{1 - \mathcal{CH}_{busy}^1} \tag{29}$$

$$\mathcal{H}_{\mathcal{Q}} = \frac{1}{1 - \mathcal{CH}_{busy}^2} \tag{30}$$

where the total time spent when the channel is busy denoted by $\mathcal{CH}_{busy}^1$ and $\mathcal{CH}_{busy}^2$ for both $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, are modeled using Eqns. 31 and 32, respectively.

$$\mathcal{CH}_{busy}^1 = \lambda \mathcal{T} - (T_{sp} + T_{be} + 2T_{rt} + 2T_{con} + T_{CCA} + T_{data} + T_{ack} + T_{prop} + T_{wt}(\Omega) \tag{31}$$

$$\mathcal{CH}_{busy}^2 = \frac{T_{wp} + T_{bk} + T_{sp} + 2T_{rt} + 2T_{con} + T_{data}}{\lambda T} \tag{32}$$

where T_{prop}, T_{wp}, T_{bk}, Ω, T_{sp}, T_{be}, T_{wt}, T_{rt}, T_{con}, and T_{data} denotes propagation time, wake-up time, back-off period, probability that channel is busy, startup time from s'_1 to s'_4, average time between two beacons' arrivals, random waiting time to receive an acknowledgment (ack) message, control packet transmission time, and the data transmission time. The time interval in which $Q - 1$ devices spent in the channel is modeled using Eqn. 33.

$$T_{\mathcal{Q}-1}^2 = (\mathcal{Q} - 1)T_Q (T_{wp} + T_{CCA} + T_{sp} + 2T_{rt} + 2T_{con} + T_{data}) (1 - \alpha) \tag{33}$$

In Eqn. 33, α and T_Q denotes the packet loss probability and total time spent in the channel, respectively. Therefore, Eqn. 34 computes the total time spent by the $\mathcal{C}_1$ and $\mathcal{C}_2$ traffic denoted as $\mathcal{T}_{total}^{1,2}$ in each state.

$$\mathcal{T}_{total}^{1,2} = \mathcal{T}_{s'_0} + \mathcal{T}_{s'_1} + \mathcal{T}_{s'_2} + \mathcal{T}_{s'_3} + \mathcal{T}_{s'_4} \quad \forall\, \mathcal{X}, \mathcal{Y} \in \mathcal{D} \tag{34}$$

In 34, the time spent in the sleep state, which includes the wake-up time is denoted by $\mathcal{T}_{s'_0}$; the time spent in the idle state, which includes the random waiting time, as $\mathcal{T}_{s'_1}$; the time spent in the active state, which includes the CCA and the back-off period, as $\mathcal{T}_{s'_2}$; the time spent in the receive state for

receiving control commands through the AP from the MEC, as $\mathcal{T}_{s'_3}$; and the time spent in the transmit state, which includes the startup time, beacon time, data transmission time, and acknowledgement time as $\mathcal{T}_{s'_4}$. In each state, the total power spent by a device $\mathcal{C}_1$ and a $\mathcal{C}_2$ device is modeled Eqn. 35 as:

$$\begin{aligned}\varrho_{s'}^{1,2} = \\ &\varrho s'_0(T_{s'_0}) + \varrho_{s'_1}(T_{s'_1}) + \\ &\varrho_{s'_2}(T_{s'_2}) + \varrho_{s'_3}(T_{s'_3}) + \varrho_{s'_4}(T_{s'_4}) \quad \forall\ \mathcal{X}, \mathcal{Y}\ \in\ \mathcal{D}\end{aligned} \tag{35}$$

where $\varrho s'_0$, $\varrho s'_1$, $\varrho s'_2$, $\varrho s'_3$, $\varrho s'_4$, denotes the power spent in the sleep, idle, sense, receive, and transmit states, respectively. Therefore, the average energy consumed by a $\mathcal{C}_1$ device during transmission, reception, control packets, and ack is modeled using Eqn. 36.

$$\zeta^1 = \zeta_{\mathcal{E}} + \zeta_{bk} + \zeta_{s'_3} + \zeta_{s'_4} \tag{36}$$

where $\zeta_{\mathcal{E}}$ is the total energy spent when the channel is busy and in idle state, ζ_{bk} is the energy consumed during back-off period, $\zeta_{s'_3}$ is the energy consumed during reception, and $\zeta_{s'_4}$ is the energy spent during transmission. Moreover, $\zeta_{\mathcal{E}}$, ζ_{bk}, $\zeta_{s'_3}$, and $\zeta_{s'_4}$ are further expressed in Eqns. 37, 38, 39, and 40, respectively.

$$\zeta_{\mathcal{E}} = \varrho_{s'_1} \cdot (\mathcal{CH}^1_{busy}) \tag{37}$$

$$\zeta_{bk} = \varrho_{s'_3} \cdot (\varrho \mathcal{T}_{wt}) \tag{38}$$

$$\zeta_{s'_3} = \varrho_{s'_3} \cdot (\mathcal{T}_{sp} + 2T_{con} + 2T_{prop}) \tag{39}$$

$$\zeta_{s'_4} = \varrho_{s'_4} \cdot (\mathcal{T}_{data}) \tag{40}$$

The average energy consumed by the C_1 devices is modeled in Eqn. 41 from Eqns. 37, 38, 39, and 40.

$$\begin{aligned}\zeta^1 = \\ &(\varrho_{s'_1}(\lambda\mathcal{T} - (T_{sp} + T_{be} + T_{CCA} + \\ &T_{data} + T_{ack} + T_{wt} \cdot \Omega)) + (\varrho_{s'_3} \cdot T_{wt}) + \\ &(\varrho_{s'_3} \cdot T_{be} \cdot \Omega) + (\varrho_{s'_3} \cdot T_{sp} + 2T_{prop} + 2T_{con}) + \\ &(\varrho_{s'_4} \cdot T_{data}))/\lambda\mathcal{T}\end{aligned} \tag{41}$$

The average energy consumed by the $\mathcal{C}_2$ device is modeled Eqn. 42.

$$\begin{aligned}\zeta^2 = \\ &(\varrho_{s'_1}(\lambda\mathcal{T} - (T_wp + T_{sp} + 2T_{rt} + 2T_{con} + T_{CCA} + \\ &T_{data} + T_{be} \cdot \Omega)) + \varrho s'_4 \cdot T_wp + \varrho s'_3 \cdot T_{be} \cdot \Omega + \\ &\varrho s'_3(T_{sp} + 2T_{rt} + 2T_{con}) + \varrho CCA \cdot T_{CCA} \cdot \varepsilon_{CCA})/\lambda\mathcal{T} \\ &+ T_{\mathcal{Q}}(\varphi T_{be} + T_{bk})\end{aligned} \tag{42}$$

where T_{CCA} and ε_{CCA} represents the CCA transmission time and the total number of CCA. Note, the devices performs two CCA to check the channel status.

B. Packet Delivery ratio

Here, the delivery ratio of the proposed MDP-HYMAC is analyzed. For a WBAN system, the delivery ratio (D_{ratio}) is defined as the ratio of the number of transmitted data packets that are successfully received by the AP $(\mathcal{T}_{success})$ to the total number of transmitted data packets $(\mathcal{T}_{trans})$. The delivery ratio is expressed in Eqn. 43

$$\mathcal{D}_{ratio} = \frac{\sum_{i=1}^{w} \mathcal{T}_{success}(i)}{\sum_{1=1}^{w} \mathcal{T}_{trans}(i)} \tag{43}$$

C. System Throughput

The system throughput is determined based on the total number of data packets that are successfully received and total time. The system throughput is expressed using Eqn. 44.

$$\mathcal{T}h = \frac{\sum_{i=1}^{w} \mathcal{T}_{success}(i)}{\mathcal{T}_{total}} \tag{44}$$

D. Delay Analysis

The system delay was analyzed using a M / M / 1 queuing model [39]. The average delay ϑ^1_{av}, ϑ^2_{av} experienced during $\mathcal{C}_1$ and $\mathcal{C}_2$ transmissions are modeled using Eqns. 45 and 46, respectively.

$$\begin{aligned}\vartheta^1_{av} = \\ &\varphi(T_{bk} + T_{CAP} + T_{CFP}) + \\ &(T_{data} + 2T_{rt} + 2T_{con})\end{aligned} \tag{45}$$

$$\begin{aligned}\vartheta^2_{av} = \\ &\varphi T_{bk} + \frac{\lambda_r E(\mathcal{S}^2)}{2(1-\rho)} + (T_{be} \\ &+ T_{wp} + (T_{data} + 2T_{rt} + 2T_{con}))\end{aligned} \tag{46}$$

In Eqns. 45 and 46, the utilization is $\rho = \frac{\lambda_r}{\mu}$, the mean service time distribution (S) is $\mu = \frac{1}{S}$, and $E(\mathcal{S}^2)$ denotes the variance of the service time.

E. Quality of Experience

Quality of experience (QoE) is a subjective measure of the degree of users' perception based on the QoS parameters [40], while QoS is an objective measure of the network's performance. In WBANs, QoE and QoS are two important metrics used to evaluate the network's performance. The relationship between QoE and QoS in WBANs is heterogeneous for various users, as it depends on different factors that includes the user's application type, physiological state, and the network's topology. Additionally, user's QoE is related to the data transmission demands and is influenced by throughput. The QoE is defined [41] using Eqn. 47

$$q_i = 5 - 5 \cdot e^{\frac{c_i \mathcal{T}h(i)}{\mathcal{T}h^i_{max}}} \tag{47}$$

where the system QoE is denoted by q, is defined as the average of all users' q_i, where q_i denotes user $i's$ QoE, $\mathcal{T}h(i)$ is the user i's throughput, $\mathcal{T}h^i_{max}$ is the maximum throughput that user i needs, which reflects the heterogeneous transmission requirements of different users, and ci is the sensitive parameter to the throughput of user i [40].

VIII. Simulation Results

This section presents the simulation results of the proposed MDP-HYMAC protocol. The system is composed of different WBANs, APs, and devices denoted as $\mathcal{K}$, $\mathcal{G}$, and $\mathcal{D}$, respectively. A total of six WBANs and twelve channels were assumed. Additionally, ten devices with one AP were considered in each WBAN. A single-hop topology was used for intra-WBAN communication, and MATLAB simulation tool was employed for simulation purposes. Based on the IEEE 802.15.4 standard, a unit back-off duration of 20 symbols was used, which is equivalent to 320μ for 2.4 GHz. The proposed protocol was compared with baseline protocols like MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, IEEE 802.15.4, MG-HYMAC, and HYMAC using standard performance metrics like energy efficiency, throughput in terms of total packet received by the AP, packet drop-off, delay, and the devices' lifetime in MATLAB. Table II presents the simulation settings employed in this study.

TABLE II
Simulation Parameters [23], [28]

Parameter	Setting
SD	7860 symbols
BI	15360 symbols
σ	60 symbols
cw_{min}	32
cw_{max}	256
DIFS	$40 \times 16e^{-6}\mu$
SIFS	$12 \times 16e^{-6}\mu$
CCA	8 symbols
Payload	624 bits
Distance	2 - 10 m
Data rate	250 Kbps
Beacon order	4
back-off period	20 symbols
ack packet size	104 bits
Receiving power	1.8 W
Receiving voltage	0.9 V
Transmission voltage	1.5 V
Transmission power	131.5 W
Number of devices	10
Superframe order	3

A. Impact of Transmission Probability on Energy Consumption

The performance of the proposed MDP-HYMAC protocol was compared with that of baseline protocols. The experiment investigated the impact of energy consumption based on the transmission probability of devices in a WBAN system. The MDP-HYMAC protocol, along with baseline protocols such as MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 protocol, was configured using different numbers of devices that varied from 1 to 10. The system's transmission probability was set to 0.8, while $\mathcal{X}$ and $\mathcal{Y}$ were set to 7 and 3, respectively. Various simulation experiments were conducted, as illustrated in Fig. 8. During the simulation experiments, the proposed MDP-HYMAC algorithms were enabled and disabled for the baseline protocols. The results indicated that the higher the transmission probability, the more energy is consumed. Conversely, the MDP-HYMAC protocol has an advantage over other protocols due to the different efficient schemes, such as an adaptive power allocation scheme and time-slot allocation scheme, that were considered. Therefore, when the transmission probability was set to 0.7, a significant improvement of about 5%, 12%, 15%, 17%, and 19% in terms of energy reduction was achieved for the proposed MDP-HYMAC protocol over MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4, respectively. This improvement was due to the Markov decision process that was employed to model the traffic arrival pattern, state transition of the devices, channel status, and buffer status which helped to improve energy efficiency and prevent congestion. Additionally, the dynamic programming method that was employed helped to find an optimal solution to the channel utilization problem and thereby minimized energy cost.

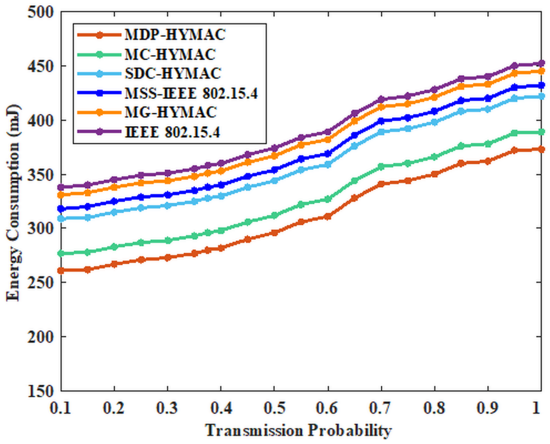


Fig. 8. Comparing the proposed MDP-HYMAC with the baseline MAC protocols to demonstrate its superiority in energy utilization efficiency based on the proposed adaptive power allocation scheme

B. Impact of Number of Devices on Energy Consumption

In this section, the number of devices from 1 to 10 are varied and investigated based on how energy consumption impact the proposed MDP-HYMAC and baseline protocols such as MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4. In a WBAN, a transmission probability of 0.8 was assumed for the MDP-HYMAC protocol and the devices were set as $\mathcal{X}$ = 7, while $\mathcal{Y}$ = 3. Following this, different simulation experiments were performed, and the proposed algorithms were enabled only for the MDP-HYMAC protocol and disabled for the other protocols. The outcome of the simulation experiments is presented in Fig. 9. The experiment results show that more

energy was consumed as the number of devices increased. However, the proposed MDP-HYMAC protocol consumed less energy compared to the baseline protocols. For instance, when the devices were 10 (i.e., $\mathcal{X} = 7$ and $\mathcal{Y} = 3$), the proposed MDP-HYMAC outperformed MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 protocols with a significant improvement of about 4%, 13%, 16%, 19%, and 22% in terms of energy efficiency, respectively. This improvement was due to the proposed optimal channel resource allocation scheme and the modeling of device transition states, channel status, and buffer status to improve energy efficiency and prevent congestion. Additionally, the proposed adaptive power allocation scheme efficiently allocated power to the devices based on time spent in each state to minimize energy consumption. The C_1 and C_2 devices are compared as shown in Fig. 10. Notably, the energy consumption of the C_1 devices is significantly higher than that of the C_2 devices. This difference could be attributed to the fact that, during each cycle, the C_1 devices are assumed to have more data to transmit compared to the C_2 devices. For example, in case 4, when the total number of devices in a WBAN network was configured to 10, with $C_1 = 6$ and $C_2 = 4$ devices, the amount of energy consumed by the C_1 devices and C_2 was approximately 150 mJ and 110 mJ, respectively. The observation depicted in Fig. 10 aligns with the understanding that the more devices in the network, the higher the energy consumption. Furthermore, the impact of different priority traffic (high-priority traffic and low-priority traffic) on network performance was investigated in the context of energy consumption. For example, as shown in Cases 1, 2, 3, and 4 in Fig 9 where C_1 represents less-critical health packets (i.e., low-priority traffic) and C_2 represents critical health packets (i.e., high- priority traffic), it was observed that the high- traffic priority class consumed less energy due to the fewer devices participating in data transmission. While, the low traffic priority class consumed more energy because of the higher number of devices involved in data transmission.

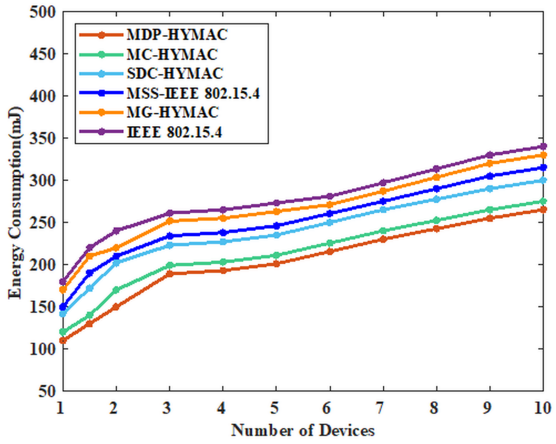


Fig. 9. Comparing the proposed MDP-HYMAC with the baseline MAC protocols to demonstrate its superiority in terms of energy utilization efficiency attributed to the proposed adaptive power allocation and time-slot management schemes

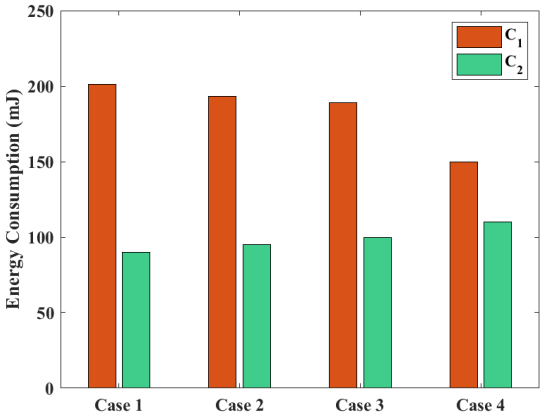


Fig. 10. Comparison of $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ based on energy consumption. For cases 1, 2, 3, and 4, it was assumed that $\boldsymbol{C_1 = 9, 8, 7,}$ and $\boldsymbol{6}$ while $\boldsymbol{C_2 = 1, 2, 3,}$ and $\boldsymbol{4}$, respectively

C. Impact of Different Back-off Attempts on Energy Consumption

The simulation experiments here investigate the impact of energy consumption based on different back-off attempts for the proposed MDP-HYMAC protocol. The devices were varied from 1 to 10 for all protocols, and for the MDP-HYMAC protocol, $\mathcal{X}$ devices were set to 6 and $\mathcal{Y}$ devices to 4. The results of the experiments are presented in Fig. 11. Based on the results, it was observed that the energy consumption of the system increased with the number of devices. However, a slight energy reduction was noticed in the fifth back-off attempt compared to the first, second, third, and fourth attempts. For instance, when the number of devices was set to 5, 4.35 MJ, 4.30 MJ, 4.26 MJ, and 4.24 MJ of energy were consumed in the first, second, third, and fourth attempts, respectively, while about 4.21 MJ energy was consumed in the fifth attempt. Consequently, the proposed back-off period strategy, the optimal channel resource allocation scheme, time-slot allocation, and adaptive power allocation schemes led to this improvement.

D. Impact of Different Number of Devices on Packet Delivery ratio

In this experiment, the proposed MDP-HYMAC protocol and baseline protocols including MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 were compared in terms of packet delivery ratio against the number of devices. To achieve this, the devices were set to $\mathcal{X} = 6$ and $\mathcal{Y} = 4$. Following this, the proposed algorithms were disabled for the baseline protocols and enabled for the proposed protocol. Different simulation experiments were conducted, and the results of the experiment are shown in Fig. VIII-C. The results in Fig. VIII-C indicate that as the number of devices increases, the delivery ratio of the system gradually decreases. However, it was noticed that the delivery ratio of the MDP-based MAC protocol is higher compared to that of MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4,

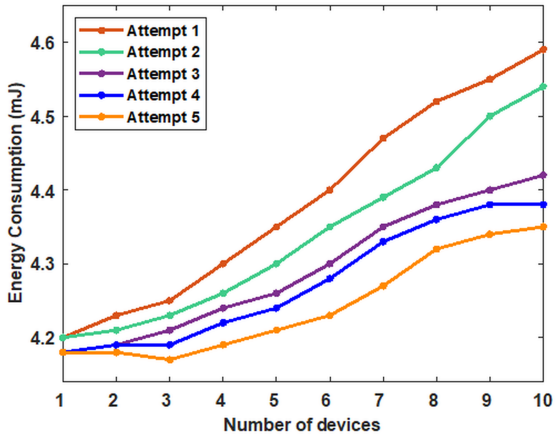
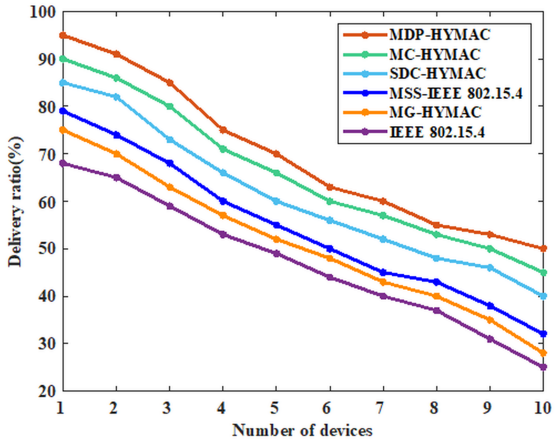


Fig. 11. Evaluating the proposed MDP-HYMAC protocol in terms of energy utilization efficiency based on different back-off attempts to investigate the efficiency of the proposed back-off and adaptive power allocation schemes

MG-HYMAC, and IEEE 802.15.4. For instance, when the devices were increased to 8, the MDP-HYMAC protocol achieved a significant improvement of about 4%, 13%, 22%, 27%, and 32% over MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4, respectively. Therefore, the proposed MDP-HYMAC protocol outperformed the baseline protocols. This improvement was a result of the proposed optimal policy scheme which helped to reduce collisions and enhance packet delivery ratio.



Evaluating the proposed MDP-HYMAC and the baseline MAC protocols based on successfully delivered packets. The evident improvement of the MDP-HYMAC can be attributed to the dynamic programming method employed for optimal channel allocation

E. Impact of Different Number of Cycles on System Throughput

In this section, the performance of the proposed MDP-HYMAC protocol is examined in terms of system throughput against different numbers of cycles, along with baseline protocols that include MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4. The experiments were conducted for about 50 cycles, and the results are presented in Fig. 12. From Fig. 12, it was observed that the proposed MDP-HYMAC protocol performed better than the baseline protocols. For example, at cycle 10, the MDP-HYMAC achieved a significant improvement of about 15%, 33%, 36%, 70%, and 82% over MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4, respectively. This improvement was as a result of the proposed optimal channel resource allocation scheme (dynamic programming method), along with the modeling of the device transition states, channel status, and buffer status using MDP, which helped to determine channel status to prevent collisions.

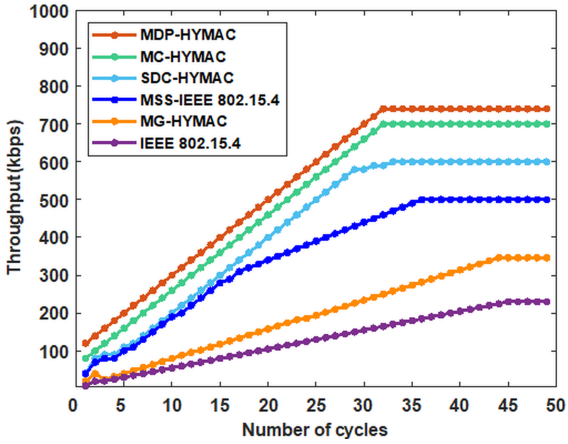


Fig. 12. Comparing the proposed MDP-HYMAC and the baseline MAC protocols based on the successfully received packets. The MDP-HYMAC outperformed the baseline protocol and this can be attributed to the proposed channel allocation scheme

F. Impact of Different Number of Cycles on Health Packets Loss

Generally, packet loss refers to a situation where data packets are lost before reaching their destination. In this study, the performance of the proposed MDP-HYMAC protocol and baseline protocols (MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC and IEEE 802.15.4) was investigated based on the number of packets that were transmitted but did not receive successfully in the AP. The total number of packets lost during transmission was also analyzed. Various simulation experiments were conducted for about 50 different cycles, and the results are presented in Fig. 13. Based on the results, it was observed that at cycle 20, the MDP-HYMAC protocol had fewer packet losses compared to MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 protocols. Therefore, a significant improvement of about 30%, 39%, 57%, 67%, and 71%, respectively. This improvement could be attributed to the optimal channel selection scheme and MDP employed to model the transition states of devices, channel status, and buffer status.

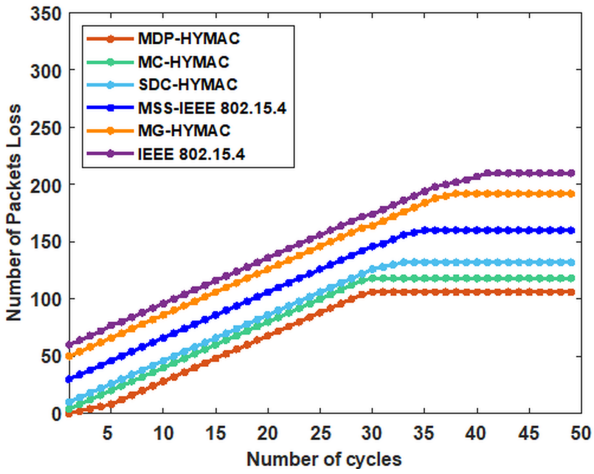


Fig. 13. Evaluation of the proposed MDP-HYMAC with the baseline protocols to investigate packet loss ratio. The evident packet loss reduction of the proposed MDP-HYMAC can be attributed to the proposed channel allocation scheme

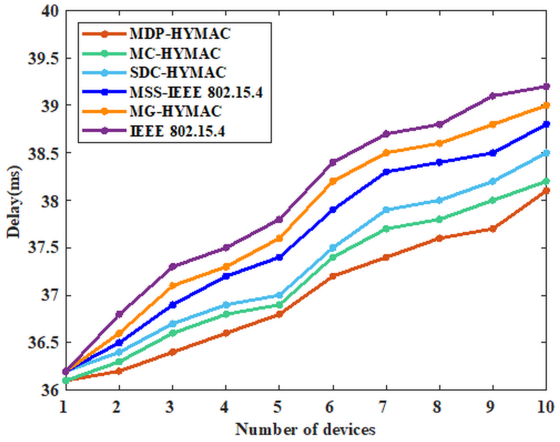


Fig. 14. Comparing the proposed MDP-HYMAC and baseline MAC protocols based on delay, the reduced delay of the proposed protocol is evident. This reduction in delay results from the proposed back-off scheme, which adjusts the contention window based on network conditions to minimize collisions and retransmissions, ultimately leading to reduced delay

G. Impact of the Number of Devices on Delay

In this section, the average packet delivery delay is examined based on the number of devices in a WBAN system. The proposed MDP-HYMAC protocol and baseline protocols (MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4) were configured by varying the number of devices from 1 to 10. For the MDP-HYMAC protocol, it was assumed that $\mathcal{X}$ devices = 7 and $\mathcal{Y}$ devices = 3. The average delay of the system was determined by calculating the time interval between packet generation and successful reception at the AP. Following this, different experiments were conducted, and the results are presented in Fig. 14. The results indicate that as the number of devices increases, so does the delay. Conversely, the proposed MDP-HYMAC protocol has a lesser delay compared to the baseline protocols. For instance, when the network was configured with 5 devices (i.e., $\mathcal{X}$ = 2 and $\mathcal{X}$ =3), a delay reduction of about 0.2%, 0.5%, 1.6%, 2%, and 3% was achieved over MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4 protocols, respectively. This improvement is due to efficient modeling of channel status, transition states, and buffer status using MDP as well as the channel utilization optimal policy that was proposed.

H. Investigation of Lifetime of the Devices Based on Transmission Probability

The impact of transmission probability on the devices' lifetime is investigated in this section by the experiments conducted on five protocols: the proposed MDP-HYMAC, MC-HYMAC, SDC-HYMAC, MSS-IEEE 802.15.4, MG-HYMAC, and IEEE 802.15.4. The number of devices was varied from 1 to 10 and a battery power of 1200 J was employed. For the MDP-HYMAC protocol, the $\mathcal{X}$ devices were set to 7 and the $\mathcal{Y}$ devices to 3. The results of the different experiments conducted are presented in Fig. 15. The transmission probability was varied from 0.1 to 0.9 and it was observed that the higher the transmission probability the lower the lifetime of the devices. Conversely, the MDP-HYMAC protocol outperformed the baseline protocols. For example, when the transmission probability was set to 0.5, a significant improvement of about 7%, 23%, 31%, 43%, and 46% was achieved over the MC-HYMAC, SDC-HYMAC, MSS-IEEE 805.15.4, MG-HYMAC, and IEEE 802.15.4 protocols, respectively. The achieved improvement was due to the different energy efficient strategies such as the adaptive power allocation scheme, time-slot allocation scheme, back-off strategy, and the optimal channel utilization scheme.

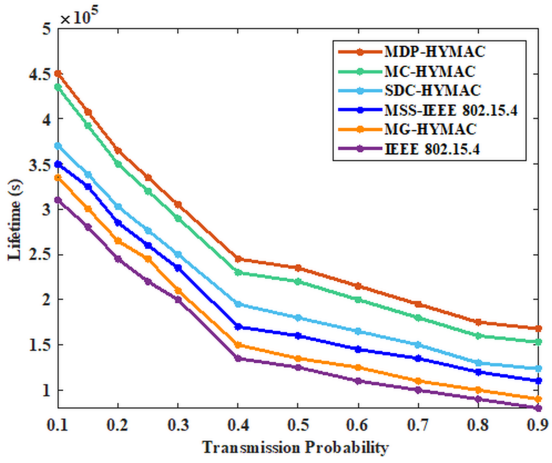


Fig. 15. Evaluating the proposed MDP-HYMAC and baseline MAC protocols based on the devices lifetime. This suggests that MDP-HYMAC achieved optimal device lifespan by managing the balance between data transmission and energy efficiency which impacts device longevity

I. Investigation of Quality of Experience Based on Different Number of Devices

QoE reflects the average level of user satisfaction with data transmission performance, while QoS is an objective mea-

sure of network performance. The simulation experiment conducted here investigates the QoE against the number of devices. To achieve this, each WBAN (i.e., user) was configured with 10 devices, and the proposed schemes were enabled for the MDP-HYMAC and disabled for MC-HYMAC, SDC-HYMAC, MSS-IEEE 805.15.4, MG-HYMAC, and IEEE 802.15.4 protocols. For MDP-HYMAC, it was assumed that $\mathcal{X} = 7$ and $\mathcal{Y} = 3$ devices. The results of the simulation experiments presented in Fig. 16 indicates that as the number of devices increased, the throughput of a user decreased and caused the QoE to gradually decline. However, it was observed that the QoE of the proposed MDP-HYMAC is higher compared to the baseline protocols when the network has 9 devices. Therefore, MDP-HYMAC outperformed the MC-HYMAC, SDC-HYMAC, MSS-IEEE 805.15.4, MG-HYMAC, and IEEE 802.15.4 protocols with a significant improvement of about 9%, 20%, 22%, 27% and 33%, respectively. This improvement is due to the proposed optimal policy, which enables efficient allocation of channel resources to reduce the probabilities of collision and congestion. As a result of this, time wasted on retransmission is reduced while throughput is enhanced.

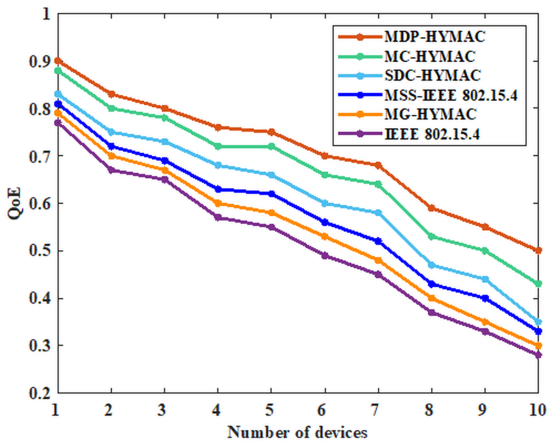


Fig. 16. Comparing the proposed and baseline protocols to understand how the number of devices influences the quality of user experience in terms of throughput

IX. Conclusion and Future Work

This study addressed critical challenges related to energy consumption, time-slot wastage, delay, and channel utilization in WBAN systems. To achieve this, a multi-channel hybrid MAC protocol based on Markov decision process was proposed. The proposed MDP-HYMAC leveraged a Markov decision process to optimize traffic arrival pattern, channel usage, and buffer management to prevent congestion, improve energy-efficiency, and prolong network lifetime. The multi-channel approach enabled the AP and devices to communicate on separate channels. This reduced collisions and enhanced system throughput. To prevent time-slot wastage, a back-off period strategy and a time slot allocation scheme was designed. Additionally, an adaptive power allocation scheme was designed to minimize energy wastage. The proposed MDP-HYMAC protocol demonstrated superior energy efficiency, reduced delay, improved packet delivery ratio, enhanced system throughput, minimized packet loss, extended device lifetime, and enhanced QoE. Future work will focus on incorporating additional resource optimization methods such game theory [42] and machine learning [43], [44] to further improve WBAN system efficiency.

References

[1] D. D. Olatinwo, A. M. Abu-Mahfouz, and G. P. Hancke, "Towards achieving efficient mac protocols for wban-enabled iot technology: a review," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, pp. 1–47, 2021.

[2] S. Kumar, R. Srivastava, S. Pathak, and B. Kumar, "Cloud-based computer-assisted diagnostic solutions for e-health," *InIntelligent Data Security Solutions for e-Health Applications. Elsevier*, pp. 219-235, Jan. 2020.

[3] D.D. Olatinwo, A. M. Abu-Mahfouz, and G.P. Hancke, "A survey on LPWAN technologies in WBAN for remote health-care monitoring," *Sensors*, vol.19, no. 23, pp. 1 - 26, May, 2023.

[4] D.D. Olatinwo, A. M. Abu-Mahfouz, and G. P. Hancke, "Towards achieving efficient MAC protocols for WBAN-enabled IoT technology: A review," *EURASIP Journal on Wireless Communications and Networking*, vol.60, no. 1, pp. 1-47, Mar. 2021.

[5] D. D. Olatinwo, A. M. Abu-Mahfouz, and G. P. Hancke, "A hybrid multi-class mac protocol for iot-enabled wban systems," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6761–6774, 2020.

[6] D.D. Olatinwo, A. M. Abu-Mahfouz, G.P. Hancke, and H.C. Myburgh, "Energy efficient priority-based hybrid MAC protocol for IoT enabled WBAN systems," *IEEE Sensors Journal*, vol.23, no. 12, pp. 13524 - 13538, May, 2023.

[7] S. Chatterjee, S. Chatterjee, S. Choudhury, S. Basak, S. Dey, S. Sain, K. S. Ghosal, N. Dalmia, and S. Sircar, "Internet of things and body area network-an integrated future,"in *Proc. IEEE 8th international conference Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 2017, pp. 396-400.

[8] R. Jaramillo, A. Quintero, and S. Chamberland, "Energy- efficient mac protocol for wireless body area networks,"in *Proc. IEEE 8th international conference Aand Workshop on Computing and Communication (IEMCON)*, 2015, pp. 1-5.

[9] M. Jung and J. Han, "Adaptive Packet Tuning for Energy Efficient Communication in Underlay CSMA/CA Networks," *IEEE Access*, vol. 11, no. 31, Aug. 2023, pp. 95989 - 95998.

[10] 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, 2018.

[11] S. Bhandari and S. Moh, "A priority-based adaptive mac protocol for wireless body area networks," *Sensors*, vol. 16, no. 3, pp. 401, 2016.

[12] D. D. Olatinwo, A. M. Abu-Mahfouz, and G. P. Hancke, "Energy-aware hybrid mac protocol for iot enabled wban systems," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2685-2699, 2021.

[13] L. Aissaoui Ferhi, K. Sethom, and F. Choubani, "Energy efficiency optimization for wireless body area networks under 802.15. 6 standard," *Wireless Personal Communications*, vol. 109, no. 3, pp. 1769-1779, 2019.

[14] T. Samal and M. R. Kabat "A prioritized traffic scheduling with load balancing in wireless body area networks," *Journal of King Saud university-computer and information sciences*, vol. 34, no. 8, pp. 5448–5455, 2022.

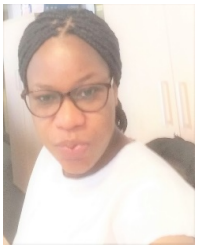
[15] M. Ambigavathi and D. Sridharan, "Traffic priority based channel assignment technique for critical data transmission in wireless body area network," *Journal of King Saud university-computer and information sciences*, vol. 42, no. 8, pp. 1-19, 2018.

[16] A. Saboor, R. Ahmad, W. Ahmed, A. K. Kiani, Y. Le Moullec, and M. M. Alam, "On research challenges in hybrid medium-access control protocols for ieee 802.15. 6 wbans," *IEEE Sensors Journal*, vol. 19, no. 19, pp. 1-19, 2018.

[17] R. Khan, M. M. Alam, and M. Guizani, "A flexible enhanced throughput and reduced overhead (fetro) mac protocol for etsi smartban," *IEEE Transactions on Mobile Computing*, vol. 21, no. 8, pp. 2671-2686, 2020.

[18] C.A. Tokognon, B. Gao, G.Y. Tian, Y. Yan, "Structural health monitoring framework based on Internet of Things: A survey," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 619-635, Feb. 2017.

[19] S. Henna, and M. A. Sarwar, "An adaptive backoff mechanism for IEEE 802.15. 4 beacon-enabled wireless body area networks," *Wireless Communications and Mobile Computing*, vol. 26, pp.1-15, Jun 2018.
[20] P. Thirumoorthy, P. Kalyanasundaram, R. Maheswar, P. Jayarajan, G. Kanagachidambaresan, and I. S. Amiri, "Time-critical energy minimization protocol using pqm (tcem-pqm) for wireless body sensor network," *The Journal of Supercomputing*, vol. 76, pp. 5862–5872, 2020.
[21] G. Sun, K. Wang, H. Yu, X. Du, and M. Guizani, "Priority-based medium access control for wireless body area networks with high-performance design," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5363–5375, 2019.
[22] X. Yang, L. Wang, and Z. Zhang, "Wireless body area networks mac protocol for energy efficiency and extending lifetime," *IEEE sensors letters*, vol. 2, no. 1, pp. 1–4, 2018.
[23] K.G. Mkongwa, Q. Liu, and S. Wang, "An adaptive backoff and dynamic clear channel assessment mechanisms in IEEE 802.15. 4 MAC for wireless body area networks," *Ad Hoc Networks*, vol. 120, pp. 1-15, Sep. 2021.
[24] D. D. Olatinwo, A. M. Abu-Mahfouz, G. P. Hancke, and H. C. Myburgh, "Energy efficient multi-channel hybrid mac protocol for iot enabled wban systems," *IEEE sensors letters*, vol. 23, no. 22, pp. 27967 - 27983, Oct. 2023.
[25] T. Samal and M. R. Kabat, "A prioritized traffic scheduling with load balancing in wireless body area networks," *Journal of King Saud university computer and information sciences*, vol. 34, no. 8, pp. 5448–5455, 2022.
[26] N. Li, X. Cai, X. Yuan, Y. Zhang, B. Zhang, and C. Li, "Eimac: a multi-channel mac protocol towards energy efficiency and low interference for wbans," *IET Communications*, vol. 12, no. 16, pp. 1954–1962, 2018.
[27] C. Li, B. Zhang, X. Yuan, S. Ullah, and A. V. Vasilakos, "Mc-mac: a multi-channel based mac scheme for interference mitigation in wbans," *Wireless Networks*, vol. 24, pp. 719–733, 2018.
[28] M. B. Rasheed, N. Javaid, M. Imran, Z. A. Khan, U. Qasim, and A. Vasilakos, "Delay and energy consumption analysis of priority guaranteed mac protocol for wireless body area networks," *Wireless networks*, vol. 23, pp. 1249–1266, 2017.
[29] T. T. Le, and S. Moh, "Hybrid multi-channel MAC protocol for WBANs with inter-WBAN interference mitigation," *Sensors*, vol. 15, no. 5, 2018, pp. 1–19.
[30] R. Piyare, A. L. Murphy, C. Kiraly, P. Tosato, and D. Brunelli, "Ultra low power wake-up radios: A hardware and networking survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2117–2157, 2017.
[31] A.D. Shoaei, M. Derakhshani, S. Parsaeefard, T. Le-Ngoc, "Learning-based hybrid TDMA-CSMA MAC protocol for virtualized 802.11 WLANs," in *Proc. 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)* IEEE, pp. 1861-1866.
[32] R. Zhang, H. Moungla, J. Yu, and A. Mehaoua, "Medium access for concurrent traffic in wireless body area networks: Protocol design and analysis," *IEEE Transactions on vehicular technology*, vol. 66, no. 3, pp. 2586–2599, 2016.
[33] K. G. Mkongwa, C. Zhang, and Q. Liu, "A reliable data transmission mechanism in coexisting ieee 802.15. 4-beacon enabled wireless body area networks," *Wireless Personal Communications*, pp. 1–22, 2022.
[34] H. Alshaheen and H. Takruri-Rizk, "Energy saving and reliability for wireless body sensor networks (wbsn)," *IEEE Access*, vol. 6, pp. 16 678–16 695, 2018.
[35] M. Roy, C. Chowdhury, and N. Aslam "Designing transmission strategies for enhancing communications in medical iot using markov decision process," *Sensors*, vol. 18, no. 12, pp. 4450, 2018.
[36] M. A. Alsheikh, D. T. Hoang, D. Niyato, H.-P. Tan, and S. Lin, "Markov decision processes with applications in wireless sensor networks: A survey," *A survey," IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. pp. 1239–1267, 2015.
[37] M. L. Puterman, "Markov decision processes: discrete stochastic dynamic programming," *John Wiley & Sons*, 2014.
[38] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *A survey," IEEE Communications Surveys & Tutorials*, vol. 58, no. 3, pp. 2116–2126, 2010.
[39] L. Decreusefond, P. Moyal, "Stochastic modeling and analysis of telecom networks," *John Wiley & Sons*, Dec. 2012.
[40] H. Su, M.-S. Pan, H. Chen, and X. Liu, "Mdp-based mac protocol for wbans in edge-enabled ehealth systems," *Electronics*, vol. 12, no. 4, pp. 947, Jul. 2023.
[41] C. Yao, Y. Jia, and L. Wang, "The qoe driven transmission optimization based on cognitive air interface match for self-organized wireless body area network," *IEEE Access*, vol. 7, pp. 138-203, 2019.
[42] S. O. Olatinwo and T.H. Joubert, "A bibliometric analysis and review of resource management in internet of water things: The use of game theory," *Water*, vol. 10, no. 14, pp. 1-29, May 2022.
[43] S. O. Olatinwo and T. H. Joubert, "Deep learning for resource management in Internet of Things networks: A bibliometric analysis and comprehensive review," *IEEE Access*, vol. 10, pp. 94691-94717, Aug. 2022.
[44] S. O. Olatinwo, T. H. Joubert, and D. D. Olatinwo "Water Quality Assessment Tool for On-site Water Quality Monitoring," *IEEE Sensors Journal*, vol. 24, no. 10, pp. 16450 - 16466, Apr. 2024.



Damilola D. Olatinwo is a researcher with expertise in mathematical modeling. Her focus lies in addressing real-life problems across various research areas, including the Internet of Things, healthcare monitoring, wireless body area networks, machine-to-machine communications, low-power and long-range communication systems. Additionally, her research encompasses designing resource management strategies, communication protocols, and the application of artificial intelligence techniques such as search and optimization methods, decision-making algorithms, and machine learning.

Adnan M. Abu-Mahfouz (Senior Member, IEEE) received the M.Eng. and Ph.D. degrees in computer engineering from the University of Pretoria, Pretoria, South Africa, in 2005 and 2011, respectively. He is currently a Chief Researcher and the Centre Manager of the Emerging Digital Technologies for 4IR (EDT4IR) Research Centre, Council for Scientific and Industrial Research, Pretoria; an Extraordinary Professor with University of Pretoria; a Professor Extraordinaire with the Tshwane University of Technology, Pretoria; and a Visiting Professor with the University of Johannesburg, Johannesburg, South Africa. His research interests are wireless sensor and actuator network, low power wide area networks, software-defined wireless sensor network, cognitive radio, network security, network management, and sensor/actuator node development. Prof Abu-Mahfouz is a Section Editor-in-Chief with the Journal of Sensor and Actuator Networks, an Associate Editor at IEEE INTERNET OF THINGS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON CYBERNETICS, IEEE ACCESS and FRONTIERS IN PLANT SCIENCE, and a member of many IEEE technical communities.

Gerhard P. Hancke (Fellow Member, IEEE) is currently a Professor with City University of Hong Kong, Hong Kong. He received B.Eng and M.Eng degrees in Computer Engineering from the University of Pretoria, South Africa, in 2002 and 2003, and a PhD in Computer Science from the University of Cambridge, United Kingdom, in 2009. Previously he worked as researcher with the Smart Card and IoT Security Centre and as teaching fellow with the Department of Information Security, both located at Royal Holloway, University of London. He also volunteers as Extraordinary Professor at University of Pretoria. His research interests are system security, reliable communication and distributed sensing for the industrial Internet-of-Things.

Hermanus C. Myburgh is an Associate Professor in the Department of Electrical, Electronic, and Computer Engineering at the University of Pretoria, South Africa. He received his B.Eng, M.Eng, and PhD in Computer and Electronic Engineering from the same institution in 2007, 2009 and 2013 respectively. He is the head of the Advanced Sensor Networks (ASN) research group and his research interests are in wireless communication, sensor fusion, machine learning, and mobile health. He is the inventor of a number of smartphone-based hearing assessment solutions and he is a co-founder of and scientific advisor to a South African digital health company, hearX Group (Pty) Ltd.