IUST Iran University of Science and Technology

Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Machine Learning-Driven Adaptive Modulation for VLC-Enabled Medical Body Sensor Networks

Reza Bayat Rizi*, Amir R. Forouzan* (C.A.), Farshad Miramirkhani**, and Mohamad F. Sabahi*

Abstract: Visible Light Communication, a key optical wireless technology, offers reliable, high-bandwidth, and secure communication, making it a promising solution for a variety of applications. Despite its many advantages, optical wireless communication faces challenges in medical environments due to fluctuating signal strength caused by patient movement. Smart transmitter structures can improve system performance by adjusting system parameters to the fluctuating channel conditions. The purpose of this research is to examine how adaptive modulation performs in a medical body sensor network system that uses visible light communication. The analysis focuses on various medical situations and investigates machine learning algorithms. The study compares adaptive modulation based on supervised learning with that based on reinforcement learning. The findings indicate that both approaches greatly improve spectral efficiency, emphasizing the significance of implementing link adaptation in visible light communication-based medical body sensor networks. The use of the Q-learning algorithm in adaptive modulation enables real-time training and enables the system to adjust to the changing environment without any prior knowledge about the environment. A remarkable improvement is observed for photodetectors on the shoulder and wrist since they experience more DC gain.

Keywords: VLC-based MBSNs, adaptive modulation, machine learning, reinforcement learning.

1 Introduction

INDOOR visible light communication (VLC) is a developing wireless communication technology that employs light-emitting diodes (LEDs) for transmitting data between devices. VLC functions by modulating the brightness of the light source to transmit digital signals, which can then be detected by a photodetector (PD) in the receiving device. VLC serves as a complement to radio frequency (RF) technology. A significant advantage of indoor VLC is its resistance to electromagnetic interference (EMI). Unlike RF communication, VLC does not rely on the electromagnetic spectrum. As a result, VLC is less susceptible to interference and disruption from other wireless devices. This makes VLC highly suitable for environments where RF interference is a concern, such as hospitals and airplanes. Additionally, indoor VLC has the potential to provide higher data rates due to its wider spectrum compared to the RF spectrum. Moreover, indoor VLC offers enhanced security compared to RF communication since visible light cannot penetrate walls.

Based on the aforementioned statements, VLC holds great promise as a technology for medical body sensor networks (MBSNs). MBSNs involve the use of wearable or implantable devices that monitor various physiological parameters such as heart rate, blood pressure, and body temperature. While VLC offers reliable, secure, and high-bandwidth wireless communication for MBSNs, there are still challenges that must be tackled, particularly the signal weakening

Iranian Journal of Electrical & Electronic Engineering, 2024.
Paper first received 10 August 2024 and accepted 14 December 2024.
* The authors are with the Department of Electrical Engineering, University of Isfahan, Isfahan, IRAN.
E-mail: a.forouzan@eng.ui.ac.ir.
** The author is with the Department of Electrical and Electronics Engineering, Isik University, Istanbul, TURKEY.
Corresponding Author: Amir R. Forouzan.

caused by the dynamic environment. Specifically, the channel DC gain can fluctuate due to the patient's body movements, changes in the distance between the transmitter and receiver, obstructions, and shadowing. These variations result in changes to the received signal strength, which in turn lead to errors in the transmitted data [1-3].

To tackle these challenges, one possible solution is the implementation of adaptive modulation, which allows for dynamic adjustment of the modulation order based on the current channel conditions. By employing adaptive modulation, spectral efficiency (SE) can be enhanced while maintaining the necessary reliability standards for communication in MBSNs. Adaptive modulation involves the flexible adaptation of modulation schemes to strike a balance between data rate and reliability. Various adaptive modulation techniques have been proposed for VLC. However, in this particular study, our focus lies on adaptive modulation utilizing machine learning (ML) algorithms. These techniques leverage the power of data-driven learning and real-time adaptation to changing environments, rendering them a highly effective approach for optimizing system performance. It is important to note that due to the dynamic nature of communication channels, ML techniques may exhibit variability over time.

Previous studies on machine learning applied to link adaptation have predominantly concentrated on other communication technologies, i.e. RF [4-8], and underwater acoustic communication systems [9-12]. Although there have been some investigations regarding learning in VLC, none of these studies have specifically addressed adaptive modulation in the context of VLC-based MBSNs.

The K-nearest neighbor (KNN) method was employed in a multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) system in [4]. To reduce the dimension of the feature space, a sub-carrier ordering technique was proposed. However, [8] introduced a deep convolutional neural network that eliminates the need for pre-processing mentioned in [4]. Both [4] and [5], along with other supervised learning-based studies on link adaptation (LA), rely on offline training algorithms. This limitation restricts their suitability for real-time operations, as discussed in [6], and necessitates a comprehensive training dataset that effectively represents the entire database. To address these limitations, [6] utilized a Q-Learning method for link adaptation in RF systems. Similarly, in [7], a deep Q-learning method was employed, considering the rate region boundaries as states in the reinforcement learning (RL) algorithm. In the context of an indoor RF system, [5] focused on delay propagation and proposed a deep Q-learning method to perform adaptive modulation with outdated channel state information (CSI).

In acoustic underwater communication (AUWC) systems, the significant challenge is the extended propagation delay, which renders the current CSI inaccessible. Addressing this issue, [9] introduced a Dyna-q algorithm that predicts channel state and calculates throughput. In addition, [10] developed a Q-learning algorithm that considers various transmission parameters. It has been established in [11] that there is a weak correlation between signal-to-noise ratio (SNR) and bit error rate (BER) in underwater environments. To tackle the problem of link adaptation in acoustic underwater communication systems, [12] proposed a deep Q-learning method. Summaries of the existing machine learning-based link adaptation studies in RF and AUWC systems are presented in Tables 1 and 2, respectively.

Furthermore, Several studies have explored VLC applications in medical body sensor networks and hospital environments. [13] and [14], explored VLC and infrared data transmission for patient monitoring and medical body sensor networks, while [15] and [16], analyzed the performance of VLC systems for smart patient monitoring and localization in hospital settings, respectively. In addition, [17] reviewed advancements in channel coding and modulation techniques, emphasizing the importance of integrating adaptive technologies to enhance reliability and efficiency in dynamic hospital scenarios. Therefore, this paper proposes a machine learning-based adaptive modulation scheme to address the challenges of dynamic hospital environments and patient mobility. Given this context, the primary contributions of this paper can be summarized as follows:

- In our study conducted in a hospital environment, we employ a realistic ray tracing method to model the channel. This method takes into account user models, man-made objects, and illumination requirements while considering various physical factors such as wavelength-dependent reflection properties, diffuse-specular reflections, measured light sources, and reflection orders up to 10. By incorporating these aspects, we enhance the accuracy of the channel impulse responses (CIRs).
- To tackle the challenges posed by meeting the diverse quality of service (QoS) requirements in evolving VLC-based 6G wireless networks, particularly in medical body sensor networks, we propose a design that utilizes Q-learning for adaptive modulation. Specifically, we focus on the DC-biased optical OFDM (DCO-OFDM)-based adaptive VLC transmission method with intensity modulation and direct detection (IM/DD) scheme. Our approach involves developing precomputed BER expressions and incorporating

adaptive BER-based modulation order switching to optimize the transmission method and improve its performance. This enables us to learn the optimal modulation scheme based on the channel's state information and environmental factors. Through simulation results, we demonstrate that our Q-learning-based approach achieves higher spectral efficiency compared to traditional fixed modulation schemes in various hospital scenarios, highlighting its effectiveness in enhancing system performance.

The remainder of the paper is organized as follows. Section 2 presents a thorough examination of the system model and scenarios. In Section 3, we introduce adaptive learning-driven modulation schemes. The performance of these techniques in both intensive care unit (ICU) ward and family-type patient room (FTPR) scenarios is showcased in Section 4. Lastly, Section 5 contains our concluding remarks.

Table 1 Comparison of Existing Machine Learning-based Link Adaptation Studies in RF systems.

	Year	Method	System Model	Proposed ML Model
[4]	2010	K-nearest neighbor method (SL)	Convolutionally codded MIMO-OFDM wireless system	-Function: mapping between feature sets and MCS -Feature space: SNR of each subcarrier -High feature dimension -Subcarrier ordering to reduce feature dimension -Requires large data set to learn the function
[5]	2021	Deep Q-learning (RL)	Indoor single-input single-output (SISO) wireless system	-Function: current CSI prediction and link adaptation based on outdated CSI -State space: received signal strength (RSS) of τ previous transmitted frames -Action space: various QAM modulation orders -No quantization error -No prior knowledge required about the environment
[6]	2012	Q-learning (RL)	Convolutionally codded MIMO-OFDM wireless system (3GPP-LTE standard)	-Function: learning the best MCS -State space: received SNR averaged over all OFDM subcarriers -Action space: various QAM modulation orders and coding rates -Throughput loss due to the quantization -No prior knowledge required about the environment
[7]	2020	Deep Q-learning (RL)	Wireless system over Rayleigh-faded channel model	-Function: DQN-based AM scheme with a trial strategy -State space: rate region boundaries (divide SNR range into rate regions) -Action space: gray-coded MPSK -No quantization error -No prior knowledge required about the environment
[8]	2019	Deep convolutional neural network (SL)	Convolutionally codded MIMO-OFDM wireless system	-Function: mapping between feature sets and MCS -Feature space: SNR of each subcarrier and noise variance -High feature dimension -No need any preprocessing -Requires large data set to learn the function
[18]	2014	Q-learning (RL)	LTE system	-Function: learning the best MCS with inaccurate channel quality indicator (CQI) -MCS selection based on the knowledge of the effect of previous decisions -CQI is utilized to avoid quantization errors of SNR -No prior knowledge required about the wireless environment

Table 2 Comparison of Existing Machine Learning-based Link Adaptation Studies in AUWC Systems.

	Year	Method	System Model	Proposed ML Model
[9]	2018	Dyna-q algorithm (RL)	Autonomous underwater vehicle (AUV)	-Function: current channel state prediction and adaptive modulation based on predicted current CSI -State space: effective SNR -Action space: BPSK, QPSK, and 8PSK
[10]	2019	Hot-booting Q-learning algorithm (RL)	Underwater acoustic	-Function: adaptive modulation and coding to maximize the quality of service (QoS) with considering multiple transmission factors offline learning stage to increasing convergence speed -State space: various transmission factors of current and previous packets -Action space: MFSK and single carrier coherent modulation

[11]	2021	Multi-layer perceptron (MLP) network (SL)	Acoustic Internet of underwater things (IOUT)	Issue: high propagation loss and drastic channel fluctuation Conventional AMC: uses correlation between SNR and BER -Link quality parameters: BER, SNR, delay spread, and frequency shift -Proving the low correlation between SNR and BER in underwater environment
[12]	2022	LSTM-enhanced DQN-based adaptive modulation (RL)	Underwater acoustic	-Issue: partially observation of acoustic channel -Combining RL with a LSTM neural network -Improvements in underwater system model -Link adaptation based on outdated CSI -State space: effective SNR of τ previous time slots -Action space: BPSK, QPSK, 8PSK and 16QAM -No quantization error -No prior knowledge required about the environment
[9]	2018	Dyna-q algorithm (RL)	Autonomous underwater vehicle (AUV)	-Function: current channel state prediction and adaptive modulation based on predicted current CSI -State space: effective SNR -Action space: BPSK, QPSK, and 8PSK
[10]	2019	Hot-booting Q-learning algorithm (RL)	Underwater acoustic	-Function: adaptive modulation and coding to maximize the quality of service (QoS) with considering multiple transmission factors offline learning stage to increasing convergence speed -State space: various transmission factors of current and previous packets -Action space: MFSK and single carrier coherent modulation

2 System Model

In this section, we present the communication system model and describe the different scenarios in which the proposed model has been assessed. The comprehensive framework for a VLC-based MBSN is depicted in Fig. 1. The diagram showcases the inclusion of a feedback pathway, which enables the transmitter to receive timely updates on the channel's condition. It is assumed that the feedback path exhibits perfection, meaning it provides accurate and error-free information. In this research, the downlink path is considered.

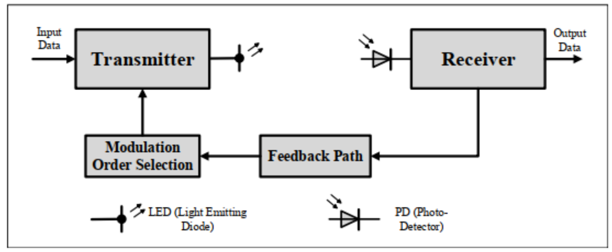


Fig. 1 VLC-based MBSNs system model.

In MBSN systems, on-body sensor nodes need to have low complexity due to power and size restrictions. Hence, in our system model, as illustrated in Fig. 1, the modulation order selection is performed on the transmitter side. M-ary pulse amplitude modulation (PAM) has been utilized for this VLC system with realistic CIR which can be expressed as [2]

$$s(t) = 2P_{avg} \sum_i m_i \, p(t - iT) \tag{1}$$

where $s(t)$ is the modulated signal, P_{avg} denotes the average optical power, $m_i \in \left\{\frac{m}{M-1} \middle| m = 0, 1, \ldots, M-1\right\}$ is the amplitude of the i-th symbol, $p(t)$ is the pulse shape with $T^{-1}\int p(t)\mathrm{d}t = 1$ and $p(t) = 0$ for $t \notin [0, T]$, and T is the symbol duration. The transmitted light is modulated by $s(t)$, which is then transmitted through the channel. The received signal at the PD can be expressed as follows

$$r(t) = s(t) * h(t) + n(t) \tag{2}$$

where $n(t)$ represents the noise including the background interference and the shot noise, which is assumed to be white and Gaussian, and $h(t)$ denotes the channel impulse response (CIR) which is modeled as [20]

$$h(t) = \sum_{k=1}^{M} P_k \, \delta(t - t_k) \tag{3}$$

where P_k represents the gain of the k-th ray, t_k represents the travel duration of the k-th ray, and M represents the total number of collected rays. We assume that $0 \le t_1 < t_2 < \cdots < t_M$ and $\tau + t_M < T$, where τ is the received pulse duration. Therefore, the inter-symbol interference (ISI) is zero at the receiver side. The received photocurrent at the output of the PD is obtained as follows

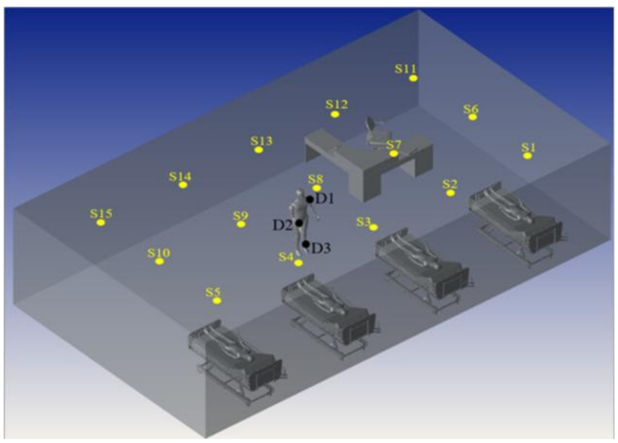
$$I(t) = \sum_{k=1}^{K} P_k s(t - t_k) \; + n(t)$$
$$= \sum_i 2RP_{avg} m_i \sum_{k=1}^{K} P_k p(t - iT - \tau_k) \; + n(t)$$

where R is the responsivity of the PDs. Since there is no established analytical expression for the indoor VLC channel model, it becomes necessary to simulate it under specific conditions. This simulation was carried out in [21] using a site-specific non-sequential ray tracing method in two distinct hospital scenarios. As depicted in Fig. 2, three PDs are placed on the mobile patient's shoulder, wrist, and ankle in both the ICU ward and the

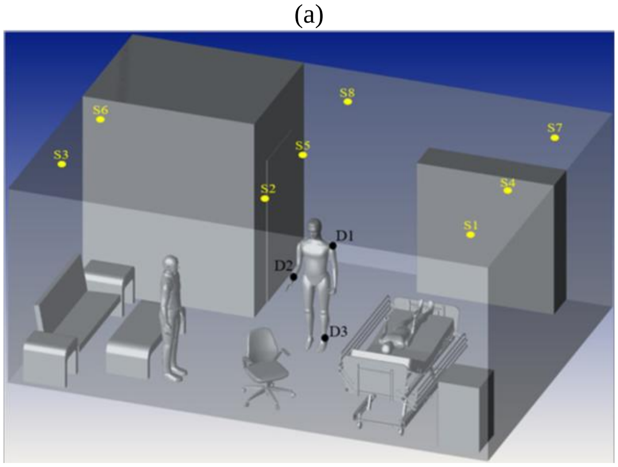
FTPR scenarios. The patient's movements follow random trajectories, and the received CIR for each PD is simulated accordingly. The objective is to maximize the throughput while maintaining a specific constraint on the symbol error rate (SER) along these trajectories. This is achieved by intelligently adjusting the order of PAM. Therefore, the Adaptive modulation optimization problem can be expressed as follows

$$\arg\max_{\mu \in I}\{R_\mu : SER_\mu \leq SER_{tar}\} \tag{4}$$

where R_μ represents the throughput achieved with a specific modulation order. The set I includes all available modulation orders which are denoted by μ. SER_μ refers to the instantaneous symbol error rate which is associated with μ, while SER_{tar} is the maximum acceptable symbol error rate.



(a)



(b)

Fig. 2 Hospital scenarios (a) ICU ward and (b) FTPR. D1, D2, and D3 represent the shoulder, wrist, and ankle sensors, respectively. S1 to S15 denote the LEDs.

3 Proposed Machine Learning-based Adaptive Modulation Scheme

Given the volatile and dynamic nature of the VLC-based MBSNs system, the problem of adaptive modulation can be seen as a complex task within the realm of RL. We begin by offering a concise overview of RL and subsequently explore the adaptive modulation scheme, which relies on Q-learning.

3.1 Reinforcement Learning-based Adaptive Modulation

RL is a subset of machine learning that focuses on the interaction between an agent and its environment. Its primary goal is for the agent to acquire knowledge and make sequential decisions within the environment to maximize a cumulative reward. Unlike supervised learning, which relies on labeled samples representing the entire dataset, RL does not require collecting environmental samples. Instead, the agent learns through a process of trial and error.

Q-learning is a widely used RL algorithm designed to solve Markov decision processes (MDPs). To understand Q-learning, it is important to grasp its fundamental components. The state space, denoted as S, encompasses the observed states s perceived by the agent within the environment. The action space, denoted as A, represents the set of available actions a that the agent can take in each state. The immediate reward function, represented by $r(s,a)$, calculates the reward received immediately after performing a particular action in a specific state. The policy, denoted as $\pi(s)$, determines how the agent maps observed states to corresponding actions. The Q-function, denoted as $Q(s,a)$, estimates the future cumulative discounted reward for an action taken by the agent in a given state, based on a specific policy. In this algorithm, the Q-values are updated according to the following process

$$Q(s,a) \leftarrow Q(s,a) + \alpha \cdot \Big[r(s,a) + \gamma \cdot \max_{a' \in A, s' \in S} Q(s',a') - Q(s,a) \Big] \tag{5}$$

where $\alpha \in [0,1]$, $\gamma \in [0,1]$, s', a' denote the learning rate, discount factor, next state, and possible actions, respectively. The goal of Q-learning is to learn an optimal policy that maximizes the expected cumulative reward over time. The optimal policy is obtained as follows

$$\pi^*(s) = \arg\max_{a \in A} Q(s,a). \tag{6}$$

To balance the exploration and exploitation a common approach is to utilize ϵ-greedy strategy [22].

3.2 Q-Learning-based Adaptive Modulation

In the optimization problem of adaptive modulation, we initially consider the tuple (H_0, ρ) as state space, where ρ represents the quantized received signal-to-noise ratio and the available modulation orders as action space. Therefore, for a specific channel state, when the agent changes the modulation order, it observes another state within the state space. This problem can be formulated as an MDP to solve using the Q-learning algorithm. As can be seen in Fig. 3, the state changes due to patient movements and the agent decisions. In our model, we do not track state changes caused by human movements, and state transitions in the MDP of the Q-learning-based adaptive modulation occur as a result of agent's decisions at the current CIR. It is worth noting that patient movements are slow, allowing the agent to explore sufficiently in each state. Additionally, once the model is trained, the agent determines the modulation order based on primary observations of the channel. The received SNR is expressed as follows

$$\rho = \frac{P}{\sigma_n^2 |H_0|^2} \tag{7}$$

where P is the transmitted optical power, σ_n^2 is the noise power, and H_0 is the channel DC-gain given by

$$H_0 = \int_0^{+\infty} h(t)\mathrm{d}t = \sum_{k=1}^{M} P_k \tag{8}$$

where P_k and M are defined in (3).

The reward function, denoted by $r(s,a)$, is defined as the achieved throughput when performing action, a in state s within the environment and is expressed by

$$r(s,a) = \begin{cases} \log_2(\mu)\left(1 - SER_\mu\right), & \text{if } SER_\mu \le SER_{tar} \\ -SER_\mu, & \text{if } SER_\mu > SER_{tar} \end{cases} \tag{9}$$

where SER_{tar} is the target symbol error rate that must be satisfied. In addition, we incorporate the ϵ-greedy strategy, initially setting a relatively high value for ϵ, to promote exploration of the environment. By choosing random actions in the early stages of learning, the agent gains valuable insights about the environment. As the learning progresses, we gradually decrease the value of ϵ to encourage the agent to rely more on the learned policy. The proposed Q-learning-based adaptive modulation scheme is summarized in Algorithm 1.

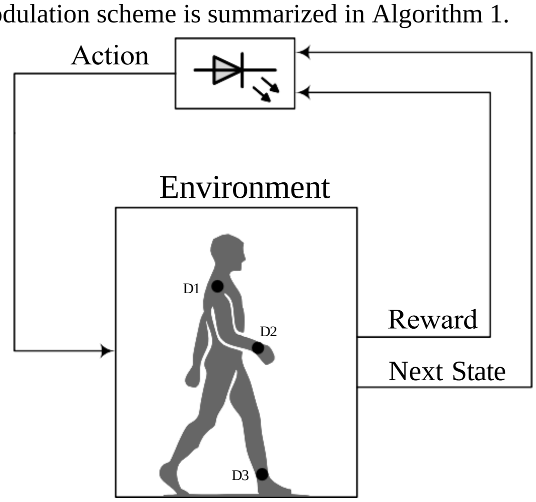


Fig. 3 Reinforcement learning model of VLC-based MBSNs adaptive modulation.

4 Numerical Results

In this research, the obtained CIRs from a previous study [21] were employed. The CIRs were collected from 20 distinct random paths, each containing 10 consecutive points obtained from both the ICU ward and FTPR scenarios. The goal was to compare the SE performance of different schemes: the adaptive modulation scheme with the Q-learning algorithm, the KNN-based adaptive modulation scheme, a non-adaptive scheme, and the optimal achievable SE. Also, a flat fading channel model is employed for all channels due to the low data rates typically encountered in MBSN applications, as it proved to be satisfactory for the study. Table 3 provides an overview of the system model and the parameters of the adaptive modulation algorithm.

Table 3 System Model and Q-Learning Model Parameters

Simulation Parameters	Value
P_{elec}	10 dBm
N_0	6.464×10^{-23}
Responsivity of PDs	1
Modulation Scheme	M-PAM
μ	{2,4,8,16,32,64}
α	0.5
γ	0.5
Min ϵ	0.001
Max Episodes	500
SER_{tar}	10^{-3}

The Q-learning-based modulation scheme operates without the need for acquiring CSI to train the model. Instead, it learns by exploring the environment. While the exploration factor decreases gradually over time, the system consistently engages in exploration to adapt dynamically to changes in the system model and environment. These two features are crucial aspects of this algorithm. For the sake of brevity, we illustrated the training stage of the Q-learning algorithm only in the ICU-ward, as it experiences more fluctuating channel DC-gain and presents a more challenging environment [21]. Fig. 4 demonstrates the operation of the Q-learning-based adaptive modulation in early stages of training process. As can be seen, it starts with exploration, leading to a higher SER than the target initially. However, as time progresses, the SER progressively decreases. Once the Q-table contains sufficient information, the agent shifts towards making more deterministic decisions using a greedy strategy. In Fig. 4, when the system transitions to the greedy decision making, the SER does not drastically drop but fluctuates below the SER_{tar}. This behavior is expected as it aligns with optimizing SE, as excessively low SER values would not be optimal.

Algorithm 1: Q-learning-based Adaptive Modulation for VLC-based MBSN

Data: Indicate $[h_n, \rho_t]$ as s_t and PAM modulation orders as a_t

Input : Environment and Q-learning parameters (α, γ,S,A, min ϵ as le, n)

Output: Optimal Q-values $Q(s_t, a_t)$

Initialize $Q(s_t, a_t)$ with random values or with 0 $\forall s \in S, a \in A$;

while *true* **do**

  Observe the MBSNs channel and calculate and quantize DC-gain as h_n;

  calculate s_t;

  while *n is constant* **do**

    if *t is a multiple of 10* **then**

      Reduce ϵ with coefficient n;

      $\epsilon = \max(\epsilon, \text{le})$;

    if *rand(0,1) $< \epsilon$* **then**

      Choose $a_t \in A$ randomly;

    else

      $a_t = \text{argmax}_a\ Q(s_t, a)$;

    The transmitter sends the modulated signal to the receiver;

    Receive feedback ρ_{t+1} and quantized $\text{SNR}_{r_{t+1}}$;

    $s_{t+1} = [h_n, \rho_{t+1}]$;

    Calculate reward as $r(s_t, a_t)$ based on Eq. 9

    $Q(s_t, a_t) = (1-\alpha)Q(s_t, a_t) + \alpha[r(s_t, a_t) + \gamma \max_{a'}\ Q(s_{t+1}, a')]$;

    Update s_t with s_{t+1};

End

Fig. 5 illustrates the SE performance of different methods. The optimal SE represents the maximum achievable SE while still satisfying the desired SER criterion. The KNN method requires training data, with 60% of the CIRs (equivalent to 12 trajectories) used for training. The value of K, indicating the number of nearest neighbors considered, is set to 3. Both the KNN and Q-learning methods demonstrate significant improvements in SE compared to the non-adaptive method, which consistently employs binary PAM to ensure the desired SER_{tar}. As observed in Figs. 5a, 5c, and 5d, the SE achieved by the KNN method occasionally exceeds the optimal SE, indicating that the desired SER_{tar} is not met in those instances.

In contrast, the Q-learning method consistently meets the desired SER_{tar} in all the figures. However, there are cases where SE falls below the optimal value due to limitations imposed by quantization levels. This occurs particularly when the optimal SER is close to the SER_{tar}, as the method prioritizes meeting the SER_{tar} requirement and acts more cautiously. Increasing the quantization levels can enhance precision but also result in higher complexity. Furthermore, the ongoing exploration process also contributes to this outcome.

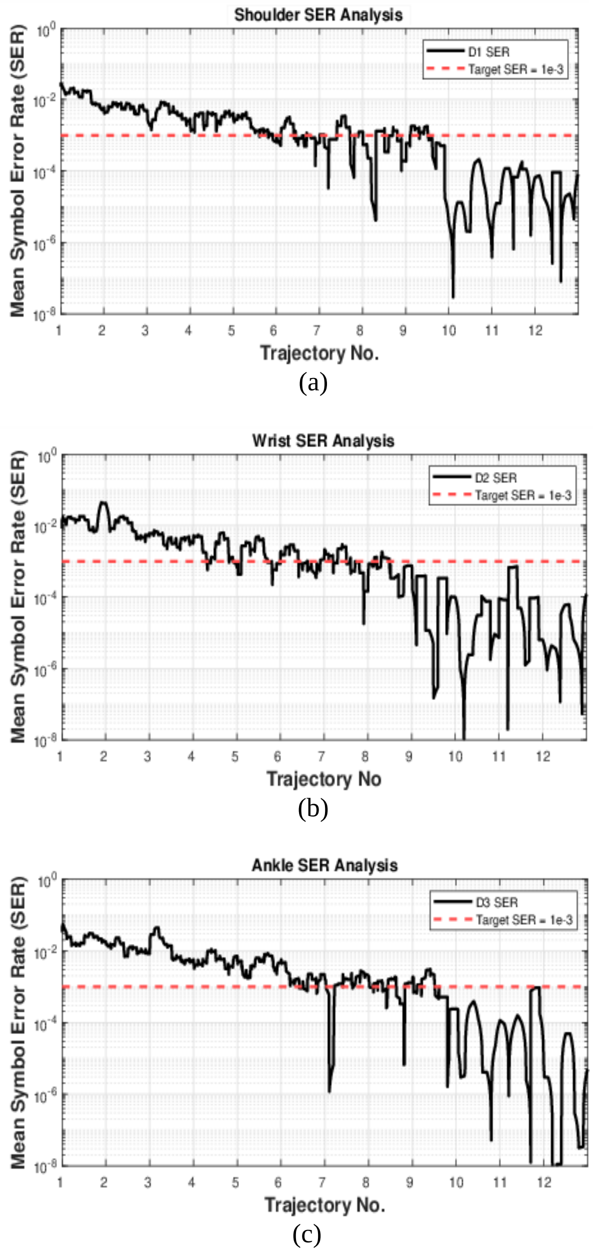


Fig. 4 Training stage of Q-learning based adaptive modulation scheme in ICU ward. (a-c) corresponds to D1-D3, respectively.

As can be seen, the utilization of Q-learning-based adaptive modulation scheme results in significant improvements in SE compared to a non-adaptive scheme. Within the ICU ward, there is an increase of 151% for D1, 178% for D2, and 81% for D3. Furthermore, in the FTPR scenario, our model demonstrated remarkable enhancements in SE. Specifically, there was an improvement of 304% for D1, 303% for D2, and 151% for D3. The greater increase in SE observed in the FTPR scenario compared to the ICU ward suggests that the range of channel DC gain in FTPR is significantly higher than that in the ICU ward. This observation supports the findings mentioned in [21].

Additionally, the application of learning-based adaptive modulation leads to a more significant improvement in SE for PDs located on Shoulder (D1) and Wrist (D2) compared to Ankle (D3) in both scenarios. This can be attributed to the sinusoidal nature of the DC gain in D1 and D2, which occurs due to their line-of-sight (LOS) rays. On the other hand, D3 receives non-line-of-sight (NLOS) rays, resulting in a smoother behavior of the DC gain. Consequently, D3 has a lower range of DC gain, leading to a reduction in SE.

5 Conclusions

In this paper, we proposed an ML-based adaptive modulation scheme for VLC-based MBSNs. The study employed an advanced ray tracing method to obtain CIRs in hospital scenarios. We investigated several modulation schemes, including the adaptive modulation scheme with the Q-learning algorithm, the KNN-based adaptive modulation, and a non-adaptive scheme as a benchmark, to enhance the SE performance. The Q-learning-based modulation scheme demonstrated the ability to dynamically adapt to changes in the system model and environment without explicit CSI requirements.

Through exploration and exploitation, the Q-learning algorithm gradually improved its SER performance and eventually met the desired SER. The KNN method also exhibited improvements in SE compared to the non-adaptive scheme but occasionally failed to achieve the desired SER because of the limitations of the training dataset, since KNN relies heavily on the availability of comprehensive and representative data to make accurate predictions. In contrast, the Q-learning method consistently achieved the desired SER, although in some cases, the SE fell below the optimal value due to limitations imposed by quantization levels and the cautious behavior of the method near the target SER. Increasing quantization levels could improve precision but would lead to higher complexity.

Overall, this research emphasizes the advantages and limitations of the Q-learning-based adaptive modulation approach in VLC-based MBSNs. Future studies could concentrate on optimizing the quantization process or using neural networks to eliminate the use of quantization process and exploring alternative adaptive modulation algorithms to enhance SE performance in VLC-based MBSNs. Furthermore, for higher data rates, where delay becomes crucial in data transmission, more complicated reinforcement learning models can be utilized to track the user's mobility in the environment.

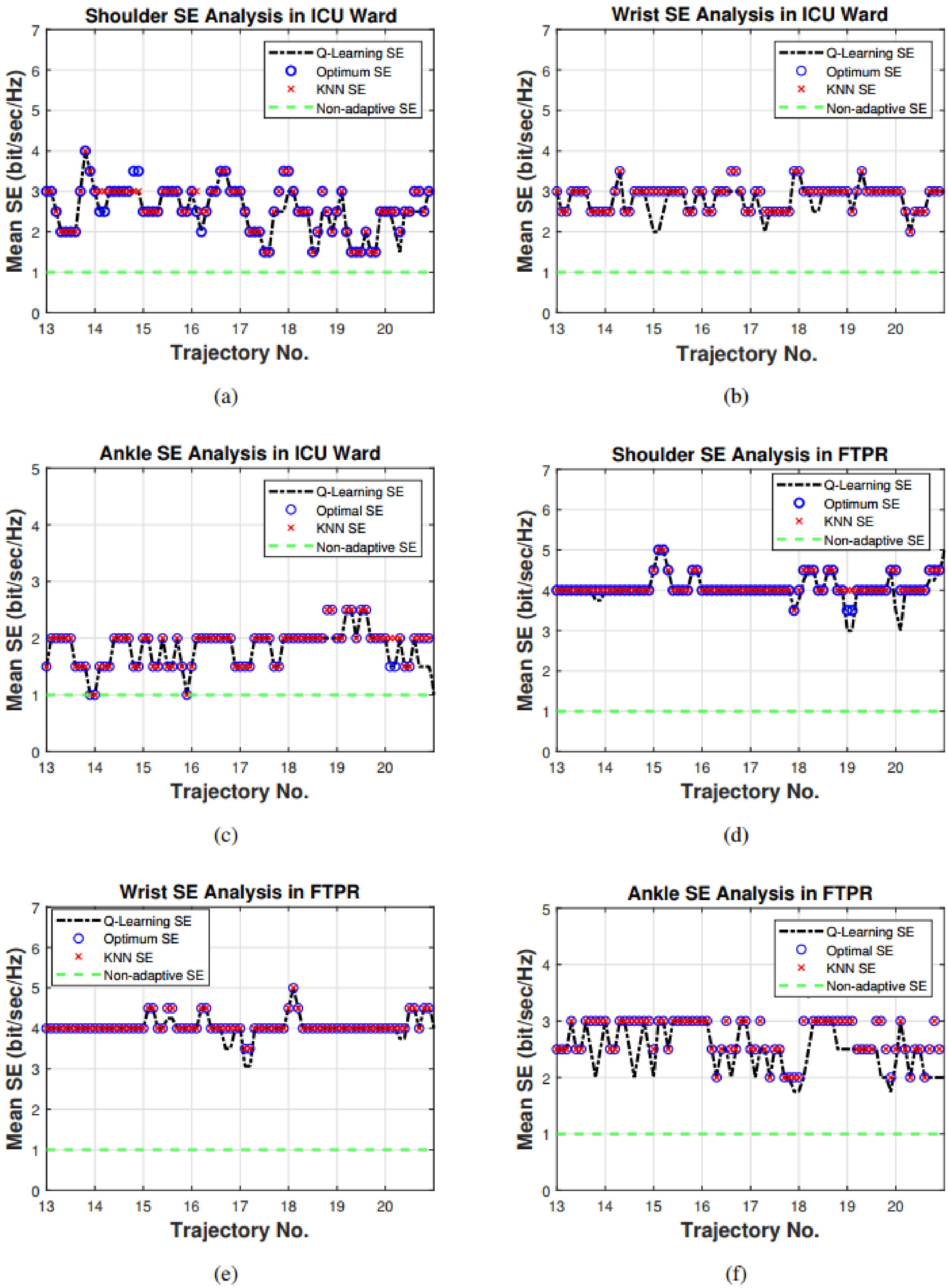


Fig. 5 Spectral efficiency analysis of different modulation schemes in (a-c) ICU ward and (d-f) FTPR.

Author Contributions

RBR performed simulations, interpreted the results, analyzed the data, reviewed the literature, implemented the solution, and wrote the manuscript.

ARF co-supervised RBR, guided him in writing the paper, and reviewed the manuscript.

FM co-supervised RBR, conceptualized and planned the idea, prepared the resources (CIRs), guided him in writing the paper, and reviewed and edited the manuscript.

MFS advised RBR and reviewed the manuscript.

References

[1] S. Arnon et al., *Advanced optical wireless communication systems*, 1st ed. Cambridge University Press, 2012.

[2] Z. Ghassemlooy et al., *Optical wireless communications: System and channel modelling with MATLAB*, 1st ed. CRC Press, 2017.

[3] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2047–2077, 2015.

[4] R. C. Daniels, C. M. Caramanis, and R. W. Heath, "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and snr ordering," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 114–126, 2010.

[5] S. Mashhadi, N. Ghiasi, S. Farahmand, and S. M. Razavizadeh, "Deep reinforcement learning based adaptive modulation with outdated CSI," *IEEE Communications Letters*, vol. PP, pp. 1–1, 07 2021.

[6] J. P. Leite, P. H. P. de Carvalho, and R. D. Vieira, "A flexible framework based on reinforcement learning for adaptive modulation and coding in OFDM wireless systems," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, 2012, pp. 809–814.

[7] D. Lee, Y. G. Sun, S. H. Kim, I. Sim, Y. M. Hwang, Y. Shin, D. I. Kim, and J. Y. Kim, "DQN-based adaptive modulation scheme over wireless communication channels," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1289–1293, 2020.

[8] M. Elwekeil, S. Jiang, T. Wang, and S. Zhang, "Deep convolutional neural networks for link adaptations in MIMO-OFDM wireless systems," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 665–668, 2019.

[9] Q. Fu and A. Song, "Adaptive modulation for underwater acoustic communications based on reinforcement learning," in *OCEANS 2018 MTS/IEEE Charleston*, 2018, pp. 1–8.

[10] W. Su, J. Lin, K. Chen, L. Xiao, and C. En, "Reinforcement learning based adaptive modulation and coding for efficient underwater communications," *IEEE Access*, vol. 7, pp. 67 539–67 550, 2019.

[11] J. Byun, Y.-H. Cho, T. Im, H.-L. Ko, K. Shin, J. Kim, and O. Jo, "Iterative learning for reliable link adaptation in the internet of underwater things," *IEEE Access*, vol. 9, pp. 30 408–30 416, 2021.

[12] Y. Zhang, J. Zhu, H. Wang, X. Shen, B. Wang, and Y. Dong, "Deep reinforcement learning-based adaptive modulation for underwater acoustic communication with outdated channel state information," *Remote Sensing*, vol. 14, no. 16, 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/16/3947.

[13] W. A. Cahyadi, T.-I. Jeong, Y.-H. Kim, Y.-H. Chung, and T. Adiono, "Patient monitoring using visible light uplink data transmission," in *2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2015, pp. 431–434.

[14] C. Lebas, S. Sahuguede, A. Julien-Vergonjanne, P. Combeau, and L. Aveneau, "Infrared and visible links for medical body sensor networks," in *2018 Global LIFI Congress (GLC)*, 2018, pp. 1–6.

[15] B. Fernandez, P. Palacios, C. Azurdia-Meza, N. Boettcher, D. Zabala-Blanco, M. Guti´errez Gaitan, and I. Soto, "Performance analysis of a VLC system applied to a hospital environment for IOT-based smart patient monitoring," in *Proceedings of the 2024 14th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, 2024.

[16] D. A. Candia, P. P. Jativa, C. Azurdia Meza, I. Sanchez, and M. Ijaz, "Performance analysis of the particle swarm optimization algorithm in a VLC system for localization in hospital environments," *Applied Sciences*, vol. 14, no. 6, 2024. [Online]. Available: https://www.mdpi.com/2076-3417/14/6/2514

[17] J. Guana-Moya, M. Roman Canizares, P. Palacios J´ativa, I. Sanchez, D. Ruminot, and F. V. Lobos, "Comprehensive survey on VLC in e-healthcare: Channel coding schemes and modulation techniques," *Applied Sciences*, vol. 14, no. 19, 2024. [Online]. Available: https://www.mdpi.com/2076-3417/14/19/8912

[18] R. Bruno, A. Masaracchia, and A. Passarella, "Robust adaptive modulation and coding (AMC) selection in LTE systems using reinforcement learning," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, 2014, pp. 1–6.

[19] R. Karmakar, S. Chattopadhyay, and S. Chakraborty, "SMARTLA: Reinforcement learning-based link adaptation for high throughput wireless access networks," *Computer Communications*, vol. 110, pp. 1–25, 2017. [Online]. Available:

https://www.sciencedirect.com/science/article/pii/S0140366417306370.

[20] F. Miramirkhani and M. Uysal, "Channel modeling and characterization for visible light communications," *IEEE Photonics Journal*, vol. 7, no. 6, pp. 1–16, 2015.

[21] B. Donmez and F. Miramirkhani, "Channel modeling and characterization for VLC-based MBSNS impaired by 3d user mobility," in *2021 13th International Conference on Electrical and Electronics Engineering (ELECO)*, 2021, pp. 485–489.

[22] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

Reza Bayat Rizi received the M.Sc. degree in Communication Engineering from the Dept. of Electrical Engineering, University of Isfahan, Isfahan, Iran and is currently working in the Power Distribution Unit of Isfahan Mobarakeh steel company. His research interests focus on the application of machine learning and artificial intelligence in electrical engineering and optical wireless communication systems.

Amir R. Forouzan received the B.S. and M.Sc. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1998 and 2000, respectively, and the Ph.D. degree from the University of Tehran, in 2004. From August 1999 to May 2004, he was with the Iran Telecommunication Research Centre (ITRC), as a Research Fellow. From June 2004 to October 2008, he was with the Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand, and from November 2008 to February 2012, he was a Research Scientist with the Electrical Engineering Department (ESAT/SISTA), KU Leuven, Belgium. He joined the Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, Iran, as an Assistant Professor, in February 2010, where he is currently an Associate Professor. His research interests include signal processing and coding in optical communication systems, dynamic spectrum management in DSL, and 5G and beyond mobile communication systems.

Farshad Miramirkhani received the B.Sc. and M.Sc. degrees in Electrical and Electronics, and Communication Engineering from University of Isfahan, Isfahan, Iran, in 2011 and 2014, respectively. He received his Ph.D. degree in Electrical and Electronics Engineering from Ozyegin University, Istanbul, Turkey in 2018. He is currently an assistant professor at Isik University, Department of Electrical and Electronics Engineering. He has contributed to the standardization works of IEEE 802.15.13 (Short Range Optical Wireless Communications) and IEEE 802.11bb (Light Communications for Wireless Local Area Networking), and the LiFi channels developed by him were selected as the LiFi Reference Channel Models. His current research interests include optical wireless communications, VLC-based medical body sensor networks, indoor visible light communications, vehicular visible light communications, underwater visible light communications, and channel modeling. His distinctions include Ozyegin University Best Research Assistant Award in 2018, IEEE Turkey Ph.D. Thesis Award in 2019, Ord. Prof. Bedri Karafakioglu Research Incentive Award in 2019, Isik University Outstanding Scientific Achievement Award in 2020 and IEEE Working Group Award by the leadership of IEEE 802.11bb-2023 in 2023. He currently serves on the editorial board of Wireless Communications for Frontiers in Communications and Networks.

Mohamad F. Sabahi is a full professor of electrical engineering at University of Isfahan, Isfahan, Iran. He received the B.S. and M.S. degrees, in 1998 and 2000, in electronics engineering and communication engineering, respectively, and a Ph.D. degree in electrical engineering in 2008 from the Isfahan University of Technology, Isfahan, Iran. Since 2008, he has been a faculty member with the Department of Electrical Engineering, University of Isfahan, Isfahan, Iran. His research interests include statistical signal processing, detection theory, and wireless communication.