A Survey on DQN-Based MAC Protocols for Wireless Networks

Junyoung Park
dept. of Software
Kongju National University
Cheonan 31080, South Korea
E-mail: jjunyoung99@gmail.com

Woongsoo Na
dept. of Software
Kongju National University
Cheonan 31080, South Korea
E-mail: wsna@kongju.ac.kr

Abstract—The rapid growth of wireless and mobile traffic has heightened the importance of designing efficient Medium Access Control (MAC) protocols capable of allocating limited wireless resources in dense network environments. Recently, deep reinforcement learning techniques—most notably Deep Q-Networks (DQN)—have received significant attention as a means to optimize MAC operations in highly dynamic environments without relying on prior channel models. This paper provides a systematic survey of DQN-based MAC protocol designs, classifying existing studies into three domains: WLAN & Wi-Fi, IoT & WSN, and heterogeneous wireless networks.

Index Terms—Reinforcement Learning, Deep Q-Network, MAC Protocol, Wireless Networks

I. Introduction

With the continuous increase in the number of wireless and mobile device users, the overall traffic in wireless networks has also been growing rapidly [1]. As the number of nodes attempting to access the network increases, it becomes crucial to allocate limited wireless resources efficiently. In dense network environments, issues such as throughput degradation caused by packet collisions [2], increased transmission delay [3], [4], and unfair channel access leading to imbalanced transmission opportunities [5] can arise, all of which significantly deteriorate network performance. Therefore, the design of a MAC (Medium Access Control) protocol capable of managing resources efficiently and distributing them fairly in dynamic network conditions is essential.

The goal of MAC protocols is to enhance overall network performance by efficiently managing shared wireless resources to achieve high throughput and low latency. Reinforcement learning (RL) is well suited to dynamic wireless networks in which traffic patterns and user demands vary over time, since an agent can improve its policy from interactions without requiring an explicit model of the environment.

This paper focuses on Deep Q-Networks (DQN), one of the most widely used RL techniques, and surveys existing studies that apply DQN to enhance MAC protocol performance. We compare how different works formulate and apply DQN to MAC decision-making and analyze the strengths and limitations of each approach.

The remainder of this paper is organized as follows. Section II introduces the fundamentals of reinforcement learning and the core mechanisms of the DQN algorithm. Section III provides a comparative survey of existing studies and summarizes how DQN has been utilized for MAC protocol optimization in various wireless environments. Section IV concludes the paper by discussing research implications, open challenges, and future directions for reinforcement learning–based MAC protocol design.

II. RL Algorithm

This section briefly summarizes the reinforcement learning notations and the core mechanisms of DQN that will be used throughout the survey.

A. Reinforcement Learning

Reinforcement learning is a framework in which an agent learns a policy that maximizes cumulative rewards through interactions with an environment. [6] The discounted cumulative reward at time t is defined as follows:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}. \tag{1}$$

Reinforcement learning problems are commonly modeled as an MDP $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, [7] where the objective is to find the optimal policy π^* that maximizes the expected cumulative reward. [8]

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[G_t \mid \pi\right]. \tag{2}$$

Since the MAC layer experiences rapidly changing traffic, channel, and collision conditions, RL is well suited because it can update access policies based on observations and rewards without requiring a prior model.

B. Deep Q-Network (DQN)

Q-learning relies on a Q-table, but its scalability is limited in environments with large state spaces, such as wireless networks. DQN uses a deep neural network to approximate the optimal action-value function $Q^*(s,a)$. [9], [10]

$$Q(s,a;\theta) \approx Q^*(s,a). \tag{3}$$

To stabilize learning, DQN stores transitions in an experience replay buffer $\mathcal{D}$ and updates the network using randomly sampled experiences.

$$(s, a, r, s') \sim \text{Uniform}(\mathcal{D}). \tag{4}$$

In addition, DQN uses a target network with parameters θ^- to compute the target value as follows:

$$y = r + \gamma \max_{a'} Q(s', a'; \theta^-). \tag{5}$$

DQN is trained by minimizing the mean squared error between the target y and the predicted value $Q(s, a; \theta)$:

$$L(\theta) = \mathbb{E}_{(s,a,r,s')\sim\mathcal{D}} \left[(y - Q(s, a; \theta))^2 \right]. \tag{6}$$

With its ability to approximate value functions in high-dimensional state spaces and adapt online, DQN can be applied to MAC decision-making problems such as backoff/CW adjustment and transmission timing selection.

III. Survey of DQN-Based MAC Protocol Designs

The performance of MAC protocols in wireless networks is highly dependent on factors such as network topology, traffic characteristics, and resource constraints. In recent years, deep reinforcement learning (DRL) techniques have been actively explored to alleviate channel contention, reduce collisions, and improve overall resource efficiency across diverse wireless environments. In this paper, existing studies are categorized into three domains—WLAN & Wi-Fi, IoT & WSN, and heterogeneous wireless networks—and we examine how DRL approaches, particularly Deep Q-Networks (DQN), address MAC-layer challenges in each domain.

A. WLAN & Wi-Fi

IEEE 802.11-based WLAN and Wi-Fi systems rely on a CSMA/CA contention mechanism, and consequently suffer from performance degradation due to backoff dynamics, hidden-node effects, and channel congestion. These issues typically manifest as:

- A rapid increase in collision probability as network density grows, resulting in notable throughput degradation.
- Limited adaptability to rapidly changing wireless environments and traffic conditions due to the use of fixed backoff parameters.

To mitigate these limitations, recent work in WLAN/Wi-Fi systems has focused on reinforcement learning–based MAC optimization, wherein CSMA/CA parameters are dynamically adjusted in response to environmental conditions.

In [11], the gateway collects information such as per-node traffic load, channel utilization, and congestion levels, and employs a DQN agent to determine key MAC parameters including the contention window (CW) and the use of RTS/CTS. This centralized approach incorporates application-layer QoE requirements (e.g., adaptive bitrate streaming) into MAC decision-making, and experimental results demonstrate significant improvements in throughput and QoE stability compared to conventional CSMA/CA.

The study in [12] proposes DeepMAC, a framework that decomposes MAC operations into functional building blocks (e.g., ACK, carrier sensing, retransmission) and uses DQN to select the optimal combination of these blocks based on current network conditions. Under low traffic load, for example, DeepMAC selectively disables ACK and carrier sensing functions to reduce protocol overhead, achieving higher throughput. This work is distinctive in that it restructures MAC functionality itself through learning-based optimization.

In [13], the authors develop a QoS-aware MAC mechanism for IEEE 802.11ax OFDMA uplink scenarios, where DQN is used to dynamically adjust the CW according to user traffic priority (high-priority vs. low-priority). Experimental results show that, in complex OFDMA multi-user settings, the DQN-based approach achieves higher throughput and lower latency than both Q-learning and baseline schemes.

The work in [14] introduces a distributed DRL-based MAC framework in which each station independently trains its own DQN agent, while the access point (AP) aggregates local models through federated learning (FL) to ensure network-wide fairness. This collaborative learning approach avoids the need for raw data sharing and achieves approximately 20% throughput improvement over traditional DCF, while also enhancing fairness in channel access.

Overall, these studies indicate that DQN-based approaches can overcome key limitations of CSMA/CA by enabling adaptive CW tuning, RTS/CTS control, MAC functional block selection, and distributed backoff optimization, leading to improvements in throughput, delay, and fairness.

B. IoT & WSN

In IoT and WSN environments, MAC protocols face stringent constraints such as ultra-low-power operation, dense node deployment, multi-hop communication structures, and high collision rates. These challenges make it difficult for conventional CSMA/CA-based schemes to ensure reliable link performance. Moreover, IoT sensors, backscatter tags, and energy-harvesting devices often operate under severe energy limitations and exhibit highly irregular traffic patterns, necessitating reinforcement learning–based approaches that can dynamically adjust MAC parameters or optimize access strategies.

Against this backdrop, numerous studies have proposed RL- and DQN-based MAC optimization techniques tailored to the unique characteristics of IoT and WSN systems.

In [15], the authors introduce a DQN-based framework that jointly optimizes directional beam selection and MAC access decisions in IoT environments employing directional communications. The state representation includes surrounding interference levels, link quality metrics, and the orientations of candidate beams, while the action space consists of transmission direction and access decisions. The reward function is designed to reflect both successful transmission probability and interference avoidance capability. Compared with Q-learning, the proposed approach achieves higher throughput and more stable link performance, demonstrating one of the early attempts to address structural challenges of dense IoT networks—such as directional collision and blockage—through learning-based adaptation.

The study in [16] proposes an RL-based multi-parameter optimization method to address CSMA/CA performance degradation in dense multi-hop WSNs. The sink node gathers network state information at each superframe and employs a centralized DQN agent to jointly optimize several MAC parameters, including CW, MaxBE, and MaxCSMABackoffs. Experimental results show that the approach improves throughput by 10–14% while reducing packet loss rate and end-to-end latency, particularly in high-density WSN environments. This study is regarded as a representative example that frames IoT/WSN MAC control as a multi-dimensional decision-making problem rather than a single-parameter tuning task.

In [17], a distributed DRL-MAC mechanism is proposed to intelligently manage the reservation procedure of backscatter IoT devices operating over Wi-Fi infrastructure. Each backscatter tag employs an LSTM-based Double DQN to predict channel occupancy patterns, while the Wi-Fi AP uses DRL to minimize reservation collisions between tags. The framework considers both the ultra-low-power constraints and the irregular traffic behavior inherent to backscatter communication. Experimental results confirm reductions in tag-to-tag collisions, improved link stability, and minimized interference with mainstream Wi-Fi traffic. This work is significant as it addresses asynchronous access challenges common in IoT/WSN systems through learning-based control.

Research on RL- and DQN-based MAC optimization in IoT and WSN environments generally aims to (i) enable intelligent MAC decisions under energy and traffic constraints, (ii) optimize key MAC parameters (e.g., backoff/CW/reservation), and (iii) mitigate CSMA/CA limitations in dense, multi-hop, or directional settings.

C. Heterogeneous Wireless Networks

In heterogeneous wireless networks (HetNets), different MAC protocols—such as TDMA, ALOHA, and CSMA/CA—share the same channel, leading to frequent collisions and unpredictable interference patterns. Traditional rule-based MAC schemes lack the ability to infer or adapt to the behavioral patterns of coexisting protocols. Therefore, model-free approaches that rely solely on observable channel dynamics and learn effective coexistence strategies using DQN are particularly appealing.

The work in [18] proposes CS-DLMA, a Reward-Backpropagation DQN (RB-DQN) approach designed to mitigate repeated collisions in environments where Wi-Fi, TDMA, and ALOHA coexist. The state representation consists of sequential observations of channel activity (Idle/Busy/Collision), the action is defined as transmit or wait, and the reward reflects both successful transmissions and collision penalties. RB-DQN improves upon standard DQN by propagating rewards backward across prior actions, thereby addressing delayed credit assignment in n-step learning. Experimental results show that RB-DQN significantly outperforms traditional CSMA in terms of higher sum throughput and lower collision rates and achieves performance close to that of model-aware optimal strategies.

In [19], the authors propose a DLMA framework based on a ResNet-enhanced Deep Q-Network to jointly consider sum throughput and α-fairness objectives in heterogeneous networks. The residual block architecture enables the agent to better model nonlinear MAC interactions, while the state representation includes slot-level histories and traffic patterns. Experimental analysis shows that ResNet-DQN converges more rapidly than vanilla DQN and Q-learning, and achieves performance nearly equivalent to model-aware optimal policies when coexisting with TDMA and q-ALOHA.

The study in [20] extends single-channel DQN-MAC approaches to multi-channel heterogeneous environments by introducing MC-DLMA. The state is defined using occupancy patterns (Idle/Busy/Collision) across multiple channels, and the action space involves selecting one of several candidate channels for transmission. Simulation results demonstrate that MC-DLMA outperforms Whittle index–based and random access schemes, highlighting the capability of DQN to address high-dimensional channel selection problems in practical HetNet scenarios.

Collectively, these studies demonstrate that DQN and its variants can learn coexistence strategies without protocol models, providing fast convergence (RB-DQN), stronger function approximation (ResNet-DQN), and scalability to multi-channel settings (MC-DLMA).

IV. Conclusion

This paper presented a comprehensive survey of existing studies that leverage Deep Q-Networks (DQN) to enhance MAC protocol performance in wireless networks. We first summarized the fundamental principles of reinforcement learning and the core mechanisms of DQN, and then categorized prior DQN-based MAC research into three domains: WLAN & Wi-Fi, IoT & WSN, and Heterogeneous Wireless Networks.

In WLAN/Wi-Fi, DQN-based schemes improve contention control via adaptive CW tuning, RTS/CTS decisions, MAC functional block selection, and federated/distributed learning, leading to gains in throughput, latency, and fairness. In IoT/WSN, DQN/DRL methods address ultra-low-power constraints and irregular traffic by optimizing multiple MAC parameters and incorporating directional/backscatter characteristics, improving energy efficiency and link reliability. In HetNets, DQN variants (RB-DQN, ResNet-DQN, and multi-channel DQN) enable near-optimal coexistence among TDMA, ALOHA, and CSMA/CA without model knowledge.

Despite these advancements, many studies still rely on simplified simulations and often assume full observability or single-agent formulations. Future research should investigate more realistic large-scale deployments, partial observability, multi-agent learning, and multi-objective optimization that jointly considers throughput, latency, fairness, and energy efficiency. Prototype or testbed-based validation will also be important to bridge the gap between simulation and real-world deployment.

Acknowledgment

This research was supported by the IITP(Institute of Information & Communications Technology Planning & Evaluation)-ITRC(Information Technology Research Center)(IITP-2025-RS-2022-00156353) grant funded by the Korea government(Ministry of Science and ICT) and Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(RS-2019-NR040074)

References

[1] T. S. Rappaport et al., ”Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond,” in IEEE Access, vol. 7, pp. 78729-78757, 2019.

[2] C. Kai, S. Zhang and L. Wang, “Impacts of packet collisions on link throughput in CSMA wireless networks,” in China Commun., vol. 15, no. 3, pp. 1-14, Mar. 2018.

[3] G. Wang, Y. Shu, L. Zhang, and O. W. W. Yang, “Delay analysis of the IEEE 802.11 DCF,” 14th IEEE Proc. PIMRC 2003, vol. 2, pp. 1737-1741, Beijing, China, 2003.

[4] T.-S. Ho and K.-C. Chen, “Performance analysis of IEEE 802.11 CSMA/CA medium access control protocol,” in Proc. PIMRC ’96 - 7th Int. Symp. Personal, Indoor, and Mobile Commun., vol. 2, pp. 407-411, Taipei, Taiwan, 1996.

[5] S. Lu, V. Bharghavan, and R. Srikant, “Fair scheduling in wireless packet networks,” in IEEE/ACM Trans. Netw., vol. 7, no. 4, pp. 473-489, Aug. 1999.

[6] R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction,” in IEEE Transactions on Neural Networks, vol. 9, no. 5, pp. 1054-1054, Sept. 1998.

[7] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” Journal of Artificial Intelligence Research, vol. 4, pp. 237–285, 1996.

[8] M. van Otterlo and M. Wiering, “Reinforcement learning and Markov decision,” in Reinforcement Learning: State-of-the-Art, 2012, pp. 3–42.

[9] V. Mnih et al., “Playing Atari with deep reinforcement learning,” arXiv preprint arXiv:1312.5602, 2013.

[10] V. Mnih et al., “Human-level control through deep reinforcement learning,” Nature, vol. 518, no. 7540, pp. 529–533, 2015.

[11] K. Aruga and T. Fujii, “MAC protocol adaptation method in coordination with application,” 2024 ICAIIC, pp. 463-466, Osaka, Japan, 2024.

[12] H. B. Pasandi and T. Nadeem, “MAC protocol design optimization using deep learning,” 2020 ICAIIC, pp. 709-715, Fukuoka, Japan, 2020.

[13] J. Lei, L. Li, and Y. Wang, “QoS-oriented media access control using reinforcement learning for next-generation WLANs,” Comput. Netw., vol. 219, p. 109426, 2022.

[14] L. Zhang, H. Yin, Z. Zhou, S. Roy, and Y. Sun, “Enhancing wifi multiple access performance with federated deep reinforcement learning,” 2020 IEEE 92nd VTC2020-Fall, pp. 1-6, Victoria, BC, Canada, 2020.

[15] N. Kim, W. Na, D. S. Lakew, N.-N. Dao, and S. Cho, “DQN-based directional MAC protocol in wireless ad hoc network in internet of things,” in IEEE Int. Things J., vol. 11, no. 7, pp. 12918-12928, Apr. 2024.

[16] J. Lei, D. Tan, X. Ma, and Y. Wang, “Reinforcement learning based multiparameter joint optimization in dense multi-hop wireless networks,” Ad Hoc Netw., vol. 154, no. 103357, Mar. 2024.

[17] X. Cao, Z. Song, B. Yang, X. Du, L. Qian, and Z. Han, “Deep reinforcement learning MAC for backscatter communications relying on wi-fi architecture,” 2019 IEEE GLOBECOM, pp. 1-6, Waikoloa, HI, USA, 2019.

[18] Y. Yu, S. C. Liew, and T. Wang, “Carrier-sense multiple access for heterogeneous wireless networks using deep reinforcement learning,” 2019 IEEE WCNCW, pp. 1-7, Marrakech, Morocco, 2019.

[19] Y. Yu, T. Wang, and S. C. Liew, “Deep- reinforcement learning multiple access for heterogeneous wireless networks,” in IEEE J. Sel. Areas in Commun., vol. 37, no. 6, pp. 1277-1290, Jun. 2019.

[20] X. Ye, Y. Yu, and L. Fu, “MAC protocol for mMulti-channel heterogeneous networks based on deep reinforcement learning,” GLOBECOM 2020, pp. 1-6, Taipei, Taiwan, 2020.