Security Enhanced Dynamic Bandwidth Allocation-Based Reinforcement Learning

ASHRAF A. ABU-EIN[1,*], WALEED ABDELKARIM ABUAIN[2],
MOHANNAD Q. ALHAFNAWI[1], OBAIDA M. AL-HAZAIMEH[3]
[1]Department of Electrical Engineering,
Al-Balqa Applied University, Irbid-21510,
JORDAN

[2]Department of Computer Science, College of Science and Computer Engineering,
Yanbu Taibah University,
Al-Madinah Al-Munawwarah 42353,
SAUDI ARABIA

[3]Department of Computer Science and Information Technology,
Al-Balqa Applied University,
Al-Huson University College, Irbid-21510,
JORDAN

Corresponding Author

Abstract: - Recently, the problem of allocating bandwidth has arisen due to the limitation of bandwidth resources. Reinforcement learning is a good technique that can be used for improving throughput, and efficiency and minimizing the overall blocking of the network. To optimize performance metrics such as throughput and Quality of Service (e.g., QoS), this research employs Reinforcement Learning (e.g., RL) and models bandwidth allocation in networking as a Markov Decision Process (e.g., MDP). Interacting with the network and modifying rewards-based policies, the agent acquires the ability to allocate bandwidth efficiently using RL techniques like Q-learning. Resource management, quality of service (e.g., QoS), fairness, security, and privacy are among the challenges the approach addresses in Dynamic Bandwidth Allocation (e.g., DBA). This approach illustrates how RL can enhance network performance and decision-making across a variety of applications. The obtained results indicate that RL algorithms are more effective in enhancing network performance, Quality of Service (e.g., QoS), and user fairness.

Key-Words: - Markov Decision Process, Reinforcement Learning, Dynamic Bandwidth Allocation, QoS, Network security, Database Administrator.

Received: March 19, 2024. Revised: August 17, 2024. Accepted: October 2, 2024. Available online: November 5, 2024.

1 Introduction

Throughput and bandwidth are metrics that quantify the rate at which a network can transmit data. Precise bandwidth estimations are crucial for optimizing overall performance, network routing, and file distribution. Precise bandwidth estimates are essential for effective traffic engineering and capacity planning. These predictions can be obtained using a variety of approaches and tools that are now available on the market, [1], [2].

1.1 DBA

Traditional methods of allocating bandwidth statically at a fixed rate caused inefficiency and wasted capacity because users hardly utilized the full allocation. With the emergence of modern optical fiber technology, Dynamic Bandwidth Allocation (DBA) enables you to split your bandwidth in any manner that you like depending on how much you use it. In the DBA system, users get given a base amount of bandwidth which can be expanded when the need arises while unused ones can be redistributed to other users. Many database management systems (e.g., DBS) algorithms and techniques have been devised aimed at efficient utilization of bandwidth through sharing out unutilized capacities to more needy units, [3].

1.2 DBA in Communication Systems

To make efficient utilization of the resources, support scalability, and enabling Quality of Service (e.g., QoS) with cost-effectiveness in communication systems Dynamic Bandwidth Allocation (e.g., DBA) is essential. With the ability to dynamically allocate bandwidth according to current needs, the network performance and capacity are enhanced thus achieving better utilization of resources using DBA. It has also created to share bandwidth in an even manner among clients and change with traffic fluctuations without the usual congestion it being which helps customers be more satisfied. Secondly, DBA helps the system to scale and elastic; hence, you can work on high-demand systems as well too heavy for the server where experimentation with new technologies is performed because lifting it would be an unaffordable change in infrastructure. Additionally, this Database Administrator (e.g., DBA) as a resource can be dynamically allocated to help prevent over-provisioning, and a lot of amount goes into getting resources unnecessary, [3], [4].

1.3 Challenges of DBA

Dynamic Bandwidth Allocation (e.g., DBA) faces several challenges despite its benefits, [4], [5]. Key issues include:

1) Resource Management: Accurately tracking and analyzing user requests, traffic patterns, and resource availability is essential for effective DBA, but it can be a complex and demanding task.
2) QoS and Fairness: To avoid congestion and provide fair bandwidth distribution while achieving QoS requirements across diverse applications, it is important to balance different traffic needs.
3) Bursty Traffic: To prevent under- or over-provisioning of resources, DBA systems need to be able to swiftly adjust to changing traffic patterns, which calls for precise prediction and adaptive approaches.
4) Scalability and Overhead: Scalable DBA systems must be capable of managing an increasing number of users and services without experiencing significant overhead or performance degradation.
5) Security and Privacy: It is imperative to safeguard sensitive information that is transmitted between central systems and users, including traffic profiles and bandwidth requirements, to preserve user privacy and network security.

Improving resource management, developing more accurate DBA algorithms, and securing more efficient bandwidth allocation are some of the ways to tackle these problems.

1.4 Role of RL in DBA

RL can improve DBA considerably because with RL decisions are not made based on the circumstances that surround them and fixed around. In the network systems, there exist RL algorithms whose role is to set the correlation of traffic modes, traffic congestion, and QoS requirements. These algorithms communicate directly with the computer or the communication system so that accurate identification of the required bandwidth can be made. By using history datasets, the RL agents can predict the traffic condition, and thus the agents can alter their behavior depending on the change in the traffic condition. They use feedback structures to adjust the allocation policies in a step-by-step process, to improve the network utility, and thus offering better services and better Quality of Service to the users. This means that it is capable of solving several conflicting objectives such as fairness and priority within dynamic environments characteristic of DBA systems. These models own the utilization of reinforcement learning in finding the bandwidth-sharing arrangements, [5], [6], [7], [8].

2 Related Works

Managing bandwidth bears higher importance due to the distinct service requirements and as the mobile and IoT networks progress. The problem with conventional methods is that they cannot cope with the current network structures because rigid processes are required to fit the model, [9], [10]. Modern development in Reinforcement Learning (e.g., RL) has made it possible to have dynamic and adaptive ways of managing the resources hence coming up with solutions to these challenges, [11], [12], [13], [14], [15], [16], [17]. Exploration of several methods in this body of research on bandwidth allocation looks at applying restricted and self-adaptive reinforcement learning methods. Table 1 provides a summary of key studies from the literature on these approaches, with an the emphasis on noteworthy improvements in throughput, latency, and overall Quality of Service (e.g., QoS):

1) Resource Allocation Method for Network Slicing Using Constrained Reinforcement Learning: Conventional resource orchestration solutions can no longer be effectively used, as the mobile networks' complexity has increased, and the mathematical models required by these solutions are incapable of handling epistemic uncertainty. This paper provides a technique for resource management in the network slicing environment constrained by a small amount of RL. The proposed strategy yielded substantial enhancements in performance indicators, including a 16% to 76% increase in cumulative throughput, a 10% to 3% decrease in discontent, and a reduction in latency from 0.9 ms to 1.6 ms.

2) Self-Adaptive Bandwidth Allocation for Low-Latency Communications Using Reinforcement Learning: Conventional supervised learning methods for bandwidth allocation encounter difficulties because they require a large amount of labeled training data. This study presents a system based on Reinforcement Learning that dynamically calculates incentives for various bandwidth choices in order to optimize network latency. This reinforcement learning (RL) approach achieved a reduction in latency of up to 50% when compared to the baseline schemes.

3) Intelligent Dynamic Bandwidth Allocation Method for Quality of Service in IoT: In order to optimize bandwidth management in IoT networks, an Intelligent Dynamic Bandwidth Allocation (e.g., IDBA) algorithm is suggested. The IDBA method surpasses the traditional Point of Service Activation (e.g., PSA) and Dynamic Bandwidth Allocation (e.g., DBA) methods, resulting in a 97% increase in throughput and an improvement in Quality of Service (e.g., QoS) metrics, including bandwidth usage, resulting in an average latency of 34 ms, and Packet loss of 4%" can be clarified to show under what circumstances this value is achieved packet loss (4%). The method dynamically adjusts to guarantee consistent bandwidth, even in low-bandwidth conditions.

4) Bandwidth Allocation in WiMAX Networks Using Reinforcement Learning: Addressing the challenge of bandwidth allocation in WiMAX networks, this investigation concentrates on Quality of Service (e.g., QoS) and packet scheduling. Traditional methods are outperformed by the proposed RL-based scheduler, which can adapt to dynamic traffic patterns and various QoS requirements. Simulation outcomes demonstrate that the RL-based scheduler effectively manages various traffic classes and enhances latency for both real-time and non-real-time services.

5) Dynamic Bandwidth Allocation for Quality-of-Service Over Ethernet PONs: The paper explores bandwidth allocation in Ethernet Passive Optical Networks (e.g., EPONs) and suggests enhancements to existing algorithms to support differentiated services. The study highlights the "light-load penalty" issue and proposes queue management techniques with priority scheduling. The results indicate that early bandwidth allocation for lightly loaded Optical Network Units (e.g., ONUs) improves average and maximum packet delays and throughput. Due to its early allocation strategy, the enhanced DBA-2 algorithm achieves a throughput of 95% compared to 88% with DBA-1.

Table 1. Literature review

Reference	Algorithm	Description and results
[18]	CRL-based network slicing resource allocation	Improving RL network slicing resource distribution with preset limitations. Improvements were made to cumulative throughput (76%), satisfaction (3%), and latency (1.6ms).
[19]	Reinforcement learning-based self-adaptive bandwidth allocation in low-latency communications	Using RL, reduce latency by 50% relative to the baseline.
[20]	WiMAX Bandwidth Allocation using Reinforcement Learning	This study compares RL against RR, Schedule mSIR, WRR, TRS+ RR, and TRS+mSIR using delay settings and simulation. RR accounts for less than half of TRS's rtPS and nrtPS delays. The simulation shows that the recommended Scheduler optimizes nrtPS and rtPS traffic and bandwidth distribution among applications.
[21]	Smart Dynamic Bandwidth Allocation for IoT Quality of Service	Compared to PSA and DBA methods, the IDBA method outperforms them in terms of enhancing the quality of service, including ideal bandwidth usage, low latency (34ms), reduced packet loss (4%), and increased throughput (97%), even in the low bandwidth range.
[22]	Ethernet PON QoS Dynamic Bandwidth Allocation	In this research, they examine how DBA2 affects throughput improvement. Because of its early allocation property, DBA2 can reach a throughput of 95%, whereas DBA1 only managed 88%.

3 The Proposed Rl-Learning Approach and System Model

This research was implemented at University College (i.e., “Al-Balqa Applied University”), which utilizes three main servers with varying bandwidth needs depending on user load and time of day. We employ the MDP components to present our system model, which includes the state space (network conditions and traffic requirements), the action space (e.g., Available bandwidth for allocation), the transition probabilities (system changes from agent actions), and the reward function (feedback on decisions). To optimize bandwidth allocation, Reinforcement Learning (e.g., RL) will be used in this paper to determine the most efficient distribution of resources. In general, Figure 1 provides an overview diagram of reinforcement learning, [23], [24], [25], [26].

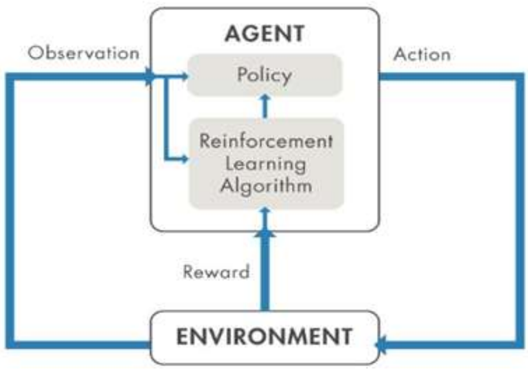

Fig. 1: An overview diagram of reinforcement learning

The system will set bandwidth values daily through a central unit to maximize user service, compare these values against actual server demands, and adjust based on a reward function. This approach aims to continually refine bandwidth settings to improve system performance, as illustrated in the accompanying Figure 2.

This research employs State–Action–Reward–State–Action (e.g., SARSA) learning to predict action-value functions for different state-action pairs and uses a convergence-based exploration algorithm to address uncertainties in server demands, [27], [28], [29]. The SARSA algorithm updates the action-value function Q (s, a) as follows:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\, [r(s,a,s') + \gamma\, Q(s',a') - Q(s,a)]$$

Where α (0< α <1) is the learning rate. The exploration algorithm balances exploration and exploitation with an exploration time threshold τ. The system explores actions for τ time units and then exploits the best policy for the remaining time T – τ, [30], [31], [32].

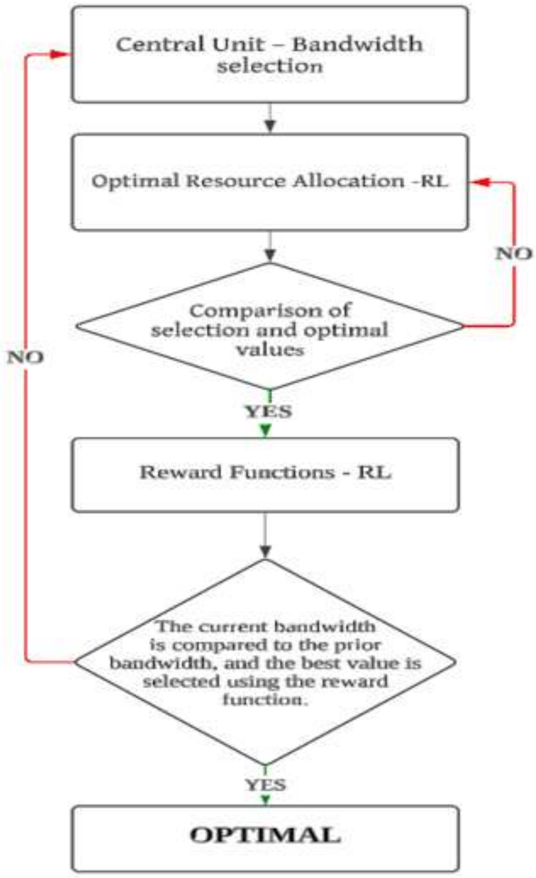

Fig. 2: RL to determine optimal bandwidth allocation

4 Results and Discussion

In this study, we compare two bandwidth allocation algorithms: equal allocation and Reinforcement Learning. The base unit of distribution for bandwidth is 500 Mbps, with possible allocation values of (500, 1000, 1500, 2000)F Mbps. Over 30 days, the maximum required bandwidth for the servers totals 6000 Mbps. The following are the real bandwidth requirements of every server. Table 2 displays the user service percentage when the available bandwidth is 3000 and the maximum required bandwidth is 6000. Similarly, Table 3 shows the user service percentage when the available bandwidth is 4500 and the maximum required bandwidth is 6000. The comparison for each table is provided in Figure 3 and Figure 4, respectively.

Table 2. User service percentage when available BW=3000 and Max Required BW=6000

	Server 1	Server 2	Server 3	Total
Reinforcement Learning	75%	68%	73%	72%
Equal Allocation	57%	59%	81%	65.7%

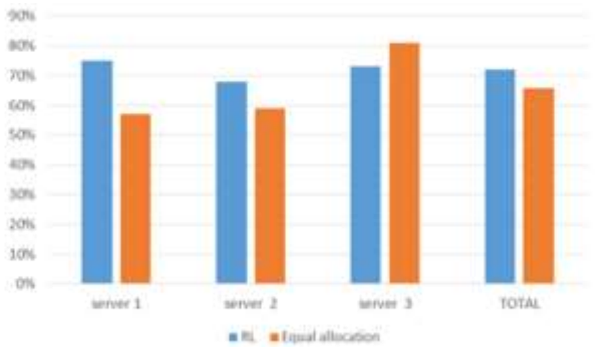

Fig 3. Compare available BW=3000 Mbps and Max Required BW=6000 Mbps

Table 3. User service percentage when available BW=4500 and Max Required BW=6000

	Server 1	Server 2	Server 3	Total
Reinforcement Learning	94%	94%	90.6%	93.5%
Equal Allocation	83%	84.4%	86.6%	84.8%

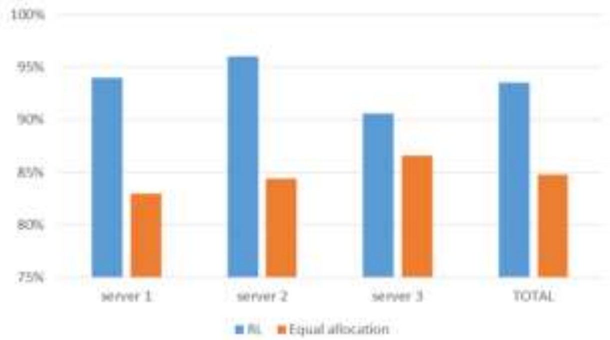

Fig. 4: Compare available BW=4500 Mbps and Max Required BW=6000 Mbps

5 Conclusion and Future Work

This research aimed to address the challenge of bandwidth allocation in networks using Reinforcement Learning (e.g., RL) techniques, specifically modeled as a Markov Decision Process (MDP). We developed a system model and problem formulation for adaptive bandwidth allocation, with a focus on three servers at University College. The study compared traditional equal allocation with RL-based dynamic bandwidth allocation (e.g., DBA), demonstrating that RL algorithms are more effective in improving network performance, Quality of Service (e.g., QoS), and user fairness. The findings indicate that RL-based DBA can substantially enhance network performance and resource utilization. Future research could explore several promising areas, including Autonomous Vehicles: Implementing RL-based DBA to optimize bandwidth allocation and communication in autonomous vehicle networks, smart Networks: Developing networks that autonomously adapt bandwidth allocation through experiential learning and real-time data, and Network Security. Utilizing RL for more effective detection and response to security threats, thereby strengthening network defenses.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the authors used ChatGPT in order to speed up the process of identifying gaps in existing algorithms. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References:

[1] K. Lai and M. Baker, "Measuring bandwidth," in *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies*. The Future is Now (Cat. No. 99CH36320), 1999, pp. 235-245. http://doi.org/10.1109/INFCOM.1999.749288

[2] J. W. Roberts, "A survey on statistical bandwidth sharing," *Computer networks*, vol. 45, pp. 319-332, 2004. http://doi.org/:10.1016/j.comnet.2004.03.010

[3] R. A. Butt, S. M. Idrus, K. N. Qureshi, N. Zulkifli, and S. H. Mohammad, "Improved dynamic bandwidth allocation algorithm for XGPON," *Journal of Optical Communications and Networking*, vol. 9, pp. 87-97, 2016. http://doi.org/10.22214/ijraset.2023.57744.

[4] A. Esmailpour and N. Nasser, "Dynamic QoS-based bandwidth allocation framework for broadband wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 2690-2700, 2011.

[5] S. Ma, X. Hu, X. Liao, and W. Wang, "Deep reinforcement learning for dynamic

bandwidth allocation in multi-beam satellite systems," in 2021 *IEEE 6th International Conference on Computer and Communication Systems (ICCCS)*, 2021, pp. 955-959.

[6] B. Skubic, J. Chen, J. Ahmed, L. Wosinska, and B. Mukherjee, "A comparison of dynamic bandwidth allocation for EPON, GPON, and next-generation TDM PON," *IEEE Communications Magazine*, vol. 47, pp. S40-S48, 2009. http://doi.org/10.1109/MCOM.2009.4804388.

[7] B. Ayyoub, A. Abu-Ein, B. Zahran, J. Nader, and O. Al-Hazaimeh, "Enhance Linux Security Server Misconfigurations and hardening Methods", *Information Sciences Letters Info*, Vol. 12, Issue 3, 2023.

[8] M. Lim, A. Abdullah, N. Jhanjhi, M. K. Khan, and M. Supramaniam, "Link prediction in time-evolving criminal network with deep reinforcement learning technique," *IEEE Access*, vol. 7, pp. 184797-184807, 2019. http://doi.org/10.1109/ACCESS.2019.2958873.

[9] A. Quran, S. Troia, O. Ayoub, N. Di Cicco, and M. Tornatore, "A Reinforcement Learning-based Dynamic Bandwidth Allocation for XGS-PON Networks," in 26th *International Conference on Optical Network Design and Modeling*, 2022, pp. 1-3.

[10] R. M. Perera, B. Oetomo, B. I. Rubinstein, and R. Borovica-Gajic, "DBA bandits: Self-driving index tuning under ad-hoc, analytical workloads with safety guarantees," in *2021 IEEE 37th International Conference on Data Engineering (ICDE),* 2021, pp. 600-611.

[11] B. Cao, X. Zheng, K. Yuan, D. Qin, and Y. Hong, "Dynamic bandwidth allocation based on adaptive predictive for low latency communications in changing passive optical networks environment," *Optical Fiber Technology,* vol. 64, p. 102556, 2021.

[12] Z. Peng, "A novel dynamic bandwidth allocation algorithm for Ethernet PON," *M. Eng. Thesis, School of Electrical and Computer Engineering RMIT University*, Australia, 2011.

[13] I. Chakour, C. Daoui, M. Baslam, B. Sainz-de-Abajo, and B. Garcia-Zapirain, "Strategic Bandwidth Allocation for QoS in IoT Gateway: Predicting Future Needs Based on IoT Device Habits," *IEEE Access*, 2024. http://doi.org/10.21833/IJAAS.2017.010.013.

[14] S. Taha and M. Kavehrad, "Dynamic bandwidth allocation in multi-class connection-oriented networks," *Computer communications*, vol. 27, pp. 13-26, 2004. http://doi.org/ 10.1016/S0140-3664(03)00145.

[15] O. Al-Hazaimeh, M. Al-Jamal, M. Bawaneh, N. Alhindawi, and B. Hamdoni, "A new image encryption scheme using dual chaotic map synchronization," *International Arab Journal of Information Technology*, vol. 18, pp. 95-102, 2021. http://doi.org/10.34028/iajit/18/1/11.

[16] I. S. Al-Qasrawi and O. M. Al-Hazaimeh, "A Pair-Wise Key Establishment Scheme for Ad Hoc Networks," *International Journal of Computer Networks & Communications*, vol. 5, p. 125, 2013.

[17] M. Alaroud, N. Tahat, A. Alomari, and O. M. Al-hazaimeh, "A novel chaotic map partially blind signature scheme based on quadratic residue problems", *Journal of Discrete Mathematical Sciences and Cryptography*, Issue 3, Volume 27, 2024, https://doi.org/10.47974/JDMSC-1823.

[18] S. M. Jafari, M. Taghipour, and M. Meybodi, "Bandwidth allocation in WiMAX networks using learning automaton," *World Applied Sciences Journal*, vol. 15, pp. 576-583, 2011.

[19] Y. Liu, J. Ding, and X. Liu, "Resource allocation method for network slicing using constrained reinforcement learning," in 2021 *IFIP Networking Conference (IFIP Networking)*, 2021, pp. 1-3.

[20] R. Mohandas and D. J. Aravindhar, "An intelligent dynamic bandwidth allocation method to support quality of service in internet of things," *International Journal of Computing*, vol. 20, pp. 254-261, 2021.

[21] C. M. Assi, Y. Ye, S. Dixit, and M. A. Ali, "Dynamic bandwidth allocation for quality-of-service over Ethernet PONs," *IEEE Journal on selected Areas in Communications*, vol. 21, pp. 1467-1477, 2003.

[22] S. Rai and A. K. Garg, "Dynamic Bandwidth Allocation in optical network using machine learning," Jilin Daxue Xuebao (Gongxueban*)/Journal of Jilin University*, vol. 4, pp. 12-23, 2021.

[23] Y. Li, "Deep reinforcement learning: An overview," arXiv preprint *arXiv:*1701.07274, 2017, https://doi.org/10.48550/arXiv.1701.07274.

[24] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Foundations and Trends® in Machine Learning*, vol. 11, pp. 219-354, 2018.

[25] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for distributed dynamic spectrum access," *IEEE Transactions on Wireless Communications,* vol. 18, pp. 310-323, 2018. http://doi.org/10.1109/TWC.2018.2879433.

[26] S. Deng, Z. Xiang, P. Zhao, J. Taheri, H. Gao, J. Yin*, et al.*, "Dynamical resource allocation in edge for trustable internet-of-things systems: A reinforcement learning method," *IEEE Transactions on Industrial Informatics,* vol. 16, pp. 6103-6113, 2020. http://doi.org/10.1109/TII.2020.2974875.

[27] X. Chen, Z. Li, W. Ni, X. Wang, S. Zhang, Y. Sun, *et al.*, "Towards Dynamic Resource Allocation and Client Scheduling in Hierarchical Federated Learning: A Two-Phase Deep Reinforcement Learning Approach," *arXiv:2406.14910*, https://doi.org/10.48550/arXiv.2406.14910.

[28] X. Guo, H. Lin, Z. Li, and M. Peng, "Deep-reinforcement-learning-based QoS-aware secure routing for SDN-IoT," *IEEE Internet of things journal,* vol. 7, pp. 6242-6251, 2019. http://doi.org/10.1109/JIOT.2019.2960033.

[29] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, *et al.*, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Communications Surveys & Tutorials,* vol. 21, pp. 3133-3174, 2019. https://doi.org/10.1002/j.1538-7305.1949.tb00928.x.

[30] L. Xiao, X. Wan, C. Dai, X. Du, X. Chen, and M. Guizani, "Security in mobile edge caching with reinforcement learning," *IEEE Wireless Communications,* vol. 25, pp. 116-122, 2018.

[31] X. Liao, J. Shi, Z. Li, L. Zhang, and B. Xia, "A model-driven deep reinforcement learning heuristic algorithm for resource allocation in ultra-dense cellular networks," *IEEE Transactions on Vehicular Technology,* vol. 69, pp. 983-997, 2019. http://doi.org/10.1038/s41598-023-41082-9.

[32] C. Fang, H. Xu, Y. Yang, Z. Hu, S. Tu, K. Ota, *et al.*, "Deep-reinforcement-learning-based resource allocation for content distribution in fog radio access networks," *IEEE Internet of Things Journal,* vol. 9, pp. 16874-16883, 2022. http://doi.org/10.3390/math10173037.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed to the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.