International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

DYNAMIC SPECTRUM ACCESS USING COGNITIVE RADIO WITH DEEP REINFORCEMENT LEARNING AND NON-ORTHOGONAL MULTIPLE ACCESS

Kishore H. N[1], Vani N[1],

[1]Department of Computer Science and Engineering, Faculty of Engineering and Management and Technology, BGS Institute of Technology, Adichunchanagiri University, B.G.Nagara, Mandya, India

*Corresponding Author: Kishore H. N E-mail: kishorehn186@gmail.com

Abstract:

This work proposes a dynamic spectrum access (DSA) system that uses Cognitive Radio (CR), Deep Reinforcement Learning (DRL), and Non-Orthogonal Multiple Access (NOMA) to make better use of the wireless spectrum. Efficiency, throughput, and collision rates are all improved by the approach, which allows secondary users (SUs) to access channels opportunistically. DRL techniques like actor-critic and Double Deep Q-Network (DDQN) allow for real-time adaptive learning, while NOMA may multiplex users to boost spectrum efficiency. Important parameters that are assessed in the model are interference, channel occupancy and bit decoding rates. The outcomes of the simulation indicate that the throughput decreases with relatively small changes during episodes, but the mentioned system is flexible. Future developments are aimed at collision avoidance and throughput stability in highly dynamic environments.

Keywords: Cognitive Radio (CR), Deep Reinforcement Learning (DRL), Non-Orthogonal Multiple Access (NOMA), Dynamic Spectrum Access (DSA), Spectrum Utilization.

Introduction:

One potential solution to the issue of spectrum scarcity and underutilization, as a result of DSA principles, might be Cognitive Radio (CR) technology enhanced with Deep Reinforcement Learning (DRL) and Non-Orthogonal Multiple Access (NOMA). This solution takes advantage of the flexibility of cognitive radios in order to dynamically use accessible spectrum and reduce interference to primary users (PUs). Important features of such an integration are as discussed below.

Deep Reinforcement Learning in DSA

- **Adaptive Learning**: Examples of applied DRL algorithms are Double Deep Q-Networks (DDQN) and Dueling DQN, which allow secondary users (SUs) to learn the most optimal spectrum access policies that are dependent on the state of channels and user mobility [1][2].

Received: August 04, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- **Performance Improvement**: Simulation findings reveal that drl-based dsa algorithms contribute greatly to the increase in the success rates of channel accesses, and a decrease in interference when compared with conventional techniques [3][4].

Non-Orthogonal Multiple Access (NOMA)

- **Spectrum Efficiency**: Multiple users sharing a single frequency band improves spectral efficiency and allows for a higher device density; this is exactly what NOMA does [5].
- **User Collaboration**: Jointly managing access to the spectrum by combining NOMA and DRL would allow users to share a ward, complicating collisions further and improving overall system capacity [5].

Although the combination of DRL and NOMA in DSA has been demonstrated to be of great potential, issues still exist, including bringing about fairness between the users and handling the complexity of the algorithms during practical setting. Future work can be to experiment with these systems to make them practically useful.

Methodology:

On Cognitive Radio, we are developing algorithm models for Dynamic Spectrum Access using Deep Reinforcement Learning and non-orthogonal multiple access based on AIML. These models are oriented on the following requirements.

S.No	Parameter	Values
1.	AI and ML Models developed based on	Deep Reinforcement Learning
2	Number of Secondary Users	4
3	Number of Channels	4
4	Primary User activity probability	40%
5	Noise_power	1e-9
6	Power_budget	50%
7	Noma_levels	[0.3, 0.2]
8	Episode_length	50
9	Number of episodes	100
10	Learning_rate	0.1%
11	gamma	0.95
12	epsilon	1
13	epsilon_min	0.05
14	epsilon_decay	0.98
15	IDE	Jupiter Notebook, Colaboratory
16	Programming language	Python

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

17	Libraries	Numpy Matplotlib Random TensorFlow deque
18	Input Evaluation Parameters	Number of secondary users
		Number of available channels.
		Probability of PU activity on each channel
		Channel gain matrix (SU-to-BS)
		AWGN noise level
		Max power allowed per SU
		Power levels for NOMA users
		Time steps per episode
19	Output Evaluation Parameter	Total_throughput
		Spectrum_utilization
		Collision_rate_with_PU
		Interference_power_to_PU
		Reward_convergence_curve
		Average_SU_power_usage
		Channel_occupancy_ratio
		NOMA_decoding_success_rate

Table 1. Experimentation Setup Explanation

Attributes for Input and Output Evaluation:

a. Number of secondary users

Optimal spectral utilization is achieved using Cognitive Radio (CR) and Deep Reinforcement Learning (DRL) Dynamic Spectrum Access (DSA), which grants secondary users (SUs) opportunistic access to spectrum channels. It depends on other parameters that determine the number of secondary users that can successfully perform in such a framework: algorithms used and network dynamics.

Dynamic User Adaptation

- **Actor-Critic Algorithms**: Transmission strategies in SUs are adaptable and by the use of actor-critic deep deterministic policy gradient algorithms, different SUs can coordinate and make the best use of the spectrum channels [6].

Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- **Mobility Considerations**: Differently put, SUs are endowed with the ability to allocate their positions as well as accessing strategies dynamically in mobility situations, improving the number of users who can use the spectrum successfully [7].

Performance Metrics

- **Channel Utilization**: It is has been shown that simulation results predict optimal channel access at about 87 percent in DSA schemes, therefore indicating an extreme level of multiple SUs being able to run in parallel and not having a considerable effect [8].

- **Sample Efficiency**: More complex models such as Deep Echo State Networks (DEQN) have been shown to enable greater sample efficiency so that more SUs can be incorporated into the network without undue computation burden [9].

Although the integration of several SUs provides a better utilization of the spectrum, there are issues on how such systems will manage the interference and that robust algorithms will be required to handle a larger number of users. When it comes to maximizing performance of DSA, the tradeoff between density of user and the level of interference is highly pivotal.

b. Number of available channels

Applying Deep Reinforcement Learning (DRL) and Non-orthogonal Multiple Access (NOMA) to live the spectrum with Dynamic Spectrum Access (DSA) and Cognitive Radio (CR) contributes extensively to the increased availability of channels that the secondary users (SUs) could use. In this method, SUs opportunistically use idle channels thus better utilizing the spectrum and decreasing rate of collision between users.

Dynamic Channel Access Mechanisms

- **Multi-Agent Systems**: Distributed multi-agent DRL also lets SUs perform local channel selection themselves, which increases the overall efficiency of channel access [10].

- **Evolutionary Game Theory**: The combined use of the evolutionary game theory, and DRL is used to introduce a balance in user collaboration, additional optimization of the channel selection, and minimizing collision [11].

- **Wideband Sensing**: Such techniques as wideband sensing enable SUs to opportunistically use a variety of channels with the least possible information about the network states in order to maximize the utility [12].

Performance Improvements

- **Increased Capacity**: The performance of such methods has shown that capacity can be substantially improved and the collision rate decreased with some algorithms performing nearly as well as 87% optimum channel access [12].

- **Adaptability**: This adaptability of SUs to alternate spectrum scenarios as a result of their facility to regulate transmission power and access approaches in the light of the real-time feedback adds value [13].

Received: August 04, 2025

Though the combination of DRL and NOMA in DSA has significant advantages, there are hurdles at making such multi-SU coordination effective especially where the channel dynamics are highly dynamic and fast-varying.

c. Probability of PU activity on each channel

This is because maximizing dynamic spectrum access (DSA) in a cognitive radio relies on knowing how likely it is that the principal users (PUs) will be using each frequency. The secondary users (SUs) may learn about the primary users' (PUs') activities and adjust their channel access behavior appropriately with the help of deep reinforcement learning (DRL) and non-orthogonal multiple access (NOMA). This is a combination of methods that improves the use of spectrum and reduces interference.

Dynamic Spectrum Access and PU Activity Prediction

- **DRL Frameworks**: Multiple DRL algorithms, like that of actor-critic, enable SUs to acquire the best policies of transmission in accordance with the PU activity profiles, enhancing the usage of the spectrum [14].
- **Predictive Models**: In order for SU to choose idle channels with the longest predicted idle duration, recurrent neural networks and support vector machines, which are machine learning methods, have been used to predict when PU would be idle [15].
- **Adaptive Learning**: The combination of DRL with LSTM is beneficial in the adjustment of the non-stationary environment, in which the activity of the PUs may vary dynamically and, therefore, improve the precision of predictions [16].

Challenges and Considerations

Nevertheless, there are difficulties in using the correct estimation of PU activity because of changes in the environment and large datasets to train the models. In addition, although DRL appears promising, it might involve a huge amount of resources and time-consuming convergence, which might make the approach less applicable in a real-time requirement [17].

d. Channel gain matrix (SU-to-BS)

In cognitive radio- and deep reinforcement learning-based dynamic spectrum access (DSA), the channel gain matrix between the base station and secondary users (SUs) is crucial. An ideal transmission method that minimizes interference to main users (PUs) and accounts for the changing channel state that SUs face while accessing the underused spectrum may be achieved with the help of such a matrix. Depending on the past and feedback-based data, DRL can be integrated into SUs to learn and optimize their channel access policies.

Channel Gain Matrix Importance

- **Dynamic Adaptation**: The channel gain will guide SUs in dynamically adjusting their transmission power along with their frequency selection so as to maximize their information rate, and minimize their conflicts with Pus [18].

- **Learning from Environment**: With the help of DRL, the SUs can achieve a good knowledge of how to transmit using the channel gain matrix as a specification of the current state of the spectrum environment [19].

Deep Reinforcement Learning Techniques

- **Actor-Critic Methods**: SUs use actor-critic algorithms to deal with state and action spaces that are large, meaning that they can optimize their access patterns relative to the channel gain matrix [18].

- **Double Deep Q-Networks**: In mobility senses, SUs use DDQN to consider the varying channel condition and become robust in accessing channels [19].

Performance Improvements

- **Reduced Interference**: Inclusion of a well-organized channel gain matrix coupled with DRL will result in extreme commendations in interference to PUs and improved channel access success [20].

- **Faster Convergence**: Training performance of the DQN model is enhanced by the implementation of contemporary neural network architectures, which include ResNet, more rapidly adapting to environment conditions [20].

Though the emphasis on channel gain matrices and DRL in DSA is encouraging, there is still a difficulty in making sure that SUs can reliably detect and use usage of vacant spectrum with no impact on PUs. It is possible to discuss the future research in much stronger algorithms which can support unpredictable spectrum usage patterns.

e. AWGN noise level

Prior to optimizing spectrum utilization and simultaneously reducing interference, it is crucial to characterize AWGN noise levels and Dynamic Spectrum Access (DSA) using Cognitive Radio (CR) with Deep Reinforcement Learning (DRL) and Non-Orthogonal Multiple Access (NOMA). When deep reinforcement learning (DRL) methods, such as Deep Echo State Networks (DEQN) and Long Short-Term Memory (LSTM) networks, are used, secondary units (SUs) are able to better manage adaptive spectrum access in the face of noise.

Impact of AWGN on DSA Performance

- **Interference Management**: The goal of the DSA approaches is to maximize channel use while minimizing disturbance to the principal users (PUs). As with the case of any communication, there is a need to manage AWGN to ensure quality of communication [21][22].

- **Learning Efficiency**: It has been demonstrated that DEQN has needs less training sample than in the case of traditional approaches, and thus, it can adapt to the varying noise levels more rapidly [21].

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- **Resource Allocation**: Compared to RL, DRL methods have the capability of allocating resources dynamically in a multi-agent context, and perform better due to learning optimum policies in noisy backgrounds in the long-term [23].

Role of NOMA in Enhancing Spectrum Access

- **Increased Capacity**: One beneficial effect of NOMA is that the same frequency band may be shared by several users, thus, increasing the overall capacity of the system under high-AWGN involvement [22].
- **Robustness to Noise**: The applications together with NOMA and DRL have the potential to improve the noise resilience of its spectrum access proposals since the users could be assigned priorities relying on their channel conditions [24].

Although the combination of DRL and NOMA in DSA networks holds potential in overseeing AWGN, it can be stated that in highly dynamic scenarios, convergence and optimal have not been achieved. These complexities should be further investigated and better system resilience should be created.

f. Max power allowed per SU

Dynamic Spectrum Access (DSA) based on Deep Reinforcement Learning (DRL) and Cognitive Radio (CR) allows secondary users (SUs) to maximize power allocation in a way that benefits the SU while increasing interference levels for the PU. Depending on the spectrum access strategy, channel conditions, and the requirement that throughput and interference must be balanced, the maximum allowed power of SUs depends on many factors. These are expounded on as seen below.

Spectrum Access Strategies

- **Recurrent Deep Reinforcement Learning (RDRL)**: This strategy focuses on primary users although secondary users have a leeway to adjust their access policies so as to maximise both throughput and fairness [25].
- **Usage-Aware Schemes**: Using past channel usage information, SUs will also have a greater chance of accessing effectively with minimal interference to Pus [26].

Power Allocation Techniques

- **Actor-Critic Methods**: Such techniques allow SUs adaptively set transmission power according to real-time feedback to the immediate environment in order to maximize their information rate whilst limiting conflict [27].
- **Deep Echo State Networks**: Using the technique enhances efficient sample usage in sample power allocation, which enables SUs to use fewer training samples to provide an informed decision [28].

Although optimization is done in terms of maximizing SUs throughput, there is the need to note that the possibility of more interference with PUs may cause regulations and necessitate more power limits on SUs in some cases.

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

g. Power levels for NOMA users

One potential way to achieve efficient power distribution to customers is by combining Cognitive Radio (CR) and Deep Reinforcement Learning (DRL) with Dynamic Spectrum Access (DSA) based on Non-Orthogonal Multiple Access (NOMA) technology. Such a combination increases spectral efficiency and accessibility to the user and incorporates dynamic characteristics of resource allocation in the wireless network. The points mentioned below represent some of the significant aspects of such integration.

Power Allocation Strategies

- **Deep Q Learning**: An algorithm that minimizes the total power consumption to maximize long-term throughput, dynamically assigning power to secondary users (SUs) with energy constraints and system dynamics in mind [29].
- **Double Deep Q Network (DDQN)**: This algorithm categorizes users according to gain variances and schedules power and considers the quality of service (QOS) of weak users that is much better than conventional algorithm in performance of the system [30].
- **Max-Min Optimization**: To satisfy the problem non-convex nature in multi-user systems, the DDQL approach generates the power coefficients so that each user is fairly treated [31].

Dynamic Resource Allocation

- **Hybrid Action Space**: The extended algorithms take into account channel, as well as estimation quality states so that efficient resource allocation can be determined, which increases scalability and performance in the case of remote state estimation tasks [32].
- **Actor-Critic Framework**: This method dynamically determines the coefficients of power allocation in order to maximize energy efficiency, and shows considerable gains in system performance [33].

Although they may be effective in improving the performance of NOMA, these approaches are facing some issues in computational complexity and real-time flexibility concern, thus they could be a bottleneck towards the practical deployment of the schemes in large-scale networks.

g. Time steps per episode

Enhancing sequential time In order for Cognitive Radio (CR) with Dynamic Spectrum Access (DSA) to guarantee efficient spectrum utilization with minimal interference, each episode is vital for the Deep Reinforcement learning component. Secondary users (SUs) may adjust their spectrum access depending on the actions of primary users (PUs) thanks to the integration of DRL with NOMA techniques.

Time Steps in DSA with DRL

- **Adaptive Learning**: Episode time steps are also important in the training of DRL models because they indicate the rate of updating the model on issues regarding the observed

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

environment. Longer episodes are able to record more complicated PU activity and SU interactions dynamics [34].

- **Sample Efficiency**: The time steps are also important to the efficiency of the samples. Indicatively as given above, deep recurrent Q networks (DRQN) needs a large amount of training data point which can be overcome by ensuring that the number of steps made by the networks is optimized to enable the networks to strike a balance between exploration and exploitation [35].

- **Multi-Agent Coordination**: In multi-SU scenarios, the time steps should be synchronized in order to guarantee appropriate coordination and reduce conflicts, which lead to the improved use of the spectrum as a whole [36].

Although longer time steps can offer better data to train and can be able to converge, they can also result in slower convergence hence leading to increased training requirements. By contrast, it is possible that smaller time steps will speed up learning, however, they are likely to miss important change in the environment. Therefore, a trade-off needs to be sought in order to ensure successful implementation of DSA.

h. Total_throughput

Dynamic Spectrum Access (DSA) with the Cognitive Radio (CR) technology, augmented by using Deep Reinforcement Learning (DRL) and Non-Orthogonal Multiple Access (NOMA), contributes to massive enhancement of total throughput in wireless networks. These techniques manage to improve the performance of the entire network because it maximizes the utilization of the available resources by providing unused spectrum to secondary users.

Dynamic Spectrum Access Mechanisms

- **Cognitive Radio Networks**: CR allows secondary users to utilize spectrum holes which would otherwise not have been used by primary users and thus improves spectral efficiency and throughput [37].

- **Deep Reinforcement Learning**: Other algorithms (such as Dueling DQN and recurrent DRL) further can improve choice of decision in terms of spectrum access through convergence speed and throughput by responding to the dynamic environments [38][39].

Performance Improvements

- **Throughput Maximization**: Methods like the cooperative multi-agent reinforcement learning have been observed to boost the sum throughput more than 14 percent in comparison with the ordinary techniques [40].

- **Collision Rate Reduction**: The adoption of the DRL strategies is capable of lowering collision rates among the users, further boosting the throughputs and network capacity [41].

Despite these developments offering great utility, there are issues towards striking the balance of fairness between the user and the mitigation of interference which may affect the overall performance of a system.

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

i. Spectrum_utilization

A significant improvement in the total throughput levels of wireless networks may be accomplished by using Cognitive Radio (CR) methods, which include merging Deep Reinforcement Learning (DRL) with Non-orthogonal Multiple Access (NOMA). Such combination of technologies enables the effective use of the spectrum band, as any secondary users can reuse channels that are not fully utilized by primary users and at the same time ensure them service.

Enhanced Throughput through DRL

- **Adaptive Learning**: Adaptive spectrum access does not greedily preoccupy with reflective information while using non-greedy agent algorithms like Dueling DQN and recurrent neural network which uses DRL algorithms to optimize spectrum access strategy giving better throughput and rapid convergence in dynamic settings [42][43].
- **Entropy-Based Exploration**: This process gives the nodes the flexibility of deciding when to look into new channels, which improves the total network usage of channels and throughput [44].
- **Mobility Considerations**: The added advantage is the fact that user mobility can be introduced into DSA schemes to ensure better sharing of the spectrum, resulting in further throughput by adjusting to the mobility of the users [45].

Non-Orthogonal Multiple Access (NOMA) Synergy

- **Increased User Capacity**: NOMA can be used to share a frequency band between many users making it a viable design in conjunction with CR and DRL maximizes throughput and efficiently manages access between users and minimizes interference.
- **Fairness and Efficiency**: Combining NOMA and CR will also guarantee that primary and secondary users can coexist with a balance that emphasizes both throughput and fairness of access to the spectrum [42].

Though the merging of such technologies is promising in terms of improving throughput, problems are associated with fairness among users and control of interference. More studies are also still required to resolve these problems and ensure that DSA is maximized in CR networks.

j. Collision_rate_with_PU

Deep reinforcement learning (DRL) and cognitive radio for dynamic spectrum access in When secondary users are trying to make the most of the spectrum while avoiding interference from main users, the collision rate between the two groups is a crucial metric to consider. Different methods have been suggested in an attempt to overcome this complexity aimed at maximizing the efficiency of the spectrum sharing where collision rates are minimized.

Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Deep Reinforcement Learning Approaches

- **Dueling Double Deep Q-Networks (D3QN)**: Such an approach will enable SUs to achieve fair sharing of resources in uncoordinated settings by identifying the profit maximizing access strategies and not colliding with PUs and other SUs [46].
- **Actor-Critic Methods**: These methods allow SUs to dynamically change power levels, access methods as per the spectrum world feed-back in order to significantly diminish any conflict with Pus [47].

Mobility Considerations

- **Dynamic User Locations**: When user mobility is taken into consideration in DSA schemes, it is possible to enable display of preferences of SUs that changes according to the changing position of PUs which is very likely to reduce the collision rates by maximizing spectrum sharing chances [48].

Centralized vs. Distributed Approaches

- **Centralized DSA Schemes**: The multiple SUs can be aligned centrally by DRL and this reduces the quality of access of PUs causing better utilization of channels and less collision probabilities [49].

Although the progress made in DRL and DSA has potential in mitigating collision with PUs, certain issues still linger, especially in areas with a lot of dynamic activities that involve movement of users and random responses of PUs and the implementation of spectrum access strategy. These practices require additional studies to perfect them and bolster their effectiveness in practise.

k. interference_power_to_PU

Dynamic Spectrum Access (DSA) based on Cognitive Radio (CR) and Deep Reinforcement Learning (DRL) is an intriguing alternative to only transmit less data to the Primary Users (PUs). Secondary users (SUs) are guaranteed adaptive management of transmission power to reduce interference, and spectrum utilization efficiency is further enhanced with the addition of Non-Orthogonal Multiple Access (NOMA).

Interference Management Strategies

- **Adaptive Power Control** SUs operate optimally transmission power according to the real-time feedback from PUs and SUs make their information rate keeping interference as low as possible [50].
- **Deep Learning Techniques**: The DRL, especially via some algorithm, such as actor-critic approach, enables SUs to acquire effective transmission policies without prior awareness of the spectrum utilization pattern of Pus [50][51].

Performance Evaluation Convergence and Efficiency: Experiments indicate that DRL approaches do better at sample complexity and convergence time compared with individual

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

approaches like Deep Recurrent Q Networks (DRQN) so interference control is improved [52][53].

- **Distributed Learning**: The network where DSA is implemented is distributed, which means that SUs can make local decisions without the consent of other individuals, which practically eliminates the risk of collisions with Pus [51].

Although these DSAs and CR technologies advancements promise effective minimization of the interference toward PUs, the problems are yet to be solved with respect to being able to coordinate the scales of SUs in dynamic surroundings, which still can cause some degree of interference.

l. Reward_convergence_curve

Dynamic Spectrum Access (DSA) schemes that use Cognitive Radio and Deep Reinforcement Learning (DRL) show how efficient and effective various algorithms are at optimizing spectrum access via their reward convergence curves. This curve will be important in terms of how fast and efficiently secondary users (SUs) would be able to learn spectrum sharing without compromising the primary users (PUs). The next sections explain the convergence attributes of the various DRL techniques in DSA networks.

Convergence of DRL Algorithms

- **Deep Recurrent Q Network (DRQN)**: Although it is effective, DRQN is computationally complex and converges slowly because it needs numerous samples of training [54][55].
- **Deep Echo State Network (DEQN)**: The study also demonstrates that DEQN provides fewer training samples than DRQN, which proves to be faster during convergence and performance in DSA networks [54][55].
- **Double Deep Q-Network (DDQN)**: In mobility cases, the DDQN improves convergence: it adapts to the moving user and performance improves considerably compared to long-term traditional settings [56].

Impact of Prior Knowledge

- **Prior Knowledge Enabled Reinforcement Learning (PKRL)**: It is more efficient during the learning process since it relies on the historical information, and the rate of convergence is 66 percent higher than any traditional approach [57].

Although convergence of these algorithms is encouraging, there are obstacles to real-world performance, including how spectrum environments must dynamically evolve over time, and how they must adapt to mobility and interference characteristics demanded by users.

m. Average_SU_power_usage

The power consumption of the secondary user (SU) in the dynamic spectrum access (DSA) system using cognitive radio and deep reinforcement learning (DRL) is dependent on a number of aspects such as effectiveness of resource use and the choices adopted to allocate power.

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

These systems enable SUs to opportunistically share a spectrum using idle channels and with the least amount of interference to the primary users (PUs). Some important aspects of this relationship can be discussed via the following lines.

Power Allocation Strategies

- Signal-to-interference-plus-noise ratio measurements and real-time environmental data, such as conflict alerts from PUs, allow SUs to adjust their transmission strengths accordingly [58].
- By giving SUs more knowledge and reducing conflicts, the actor-critic deep deterministic policy gradient technique enables them to optimize power allocations dynamically [58].

Impact of Mobility

- SUs within mobility context can conduct spectrum exploration based on location variations that may result into changes in power consumption relating to their closeness to Pus [59].
- The use of IEEE 802.11 threshold distances enables the introduction of power level adaptation of SUs in accordance with their movement and spectrum opportunities in the vicinity of their location [59].

Performance and Efficiency

- The simulation outcomes suggest that the DRL-based techniques, including deep Q-learning, enable substantial gains in the efficacy of spectrum use and power consumption over conventional ones [60].
- The current abandoned method of spectrum sharing among the users as well as making use of the centralized method of spectrum allocation proves that the SUs have had up to 87 percent accessibility to optimal channel access, as it relates to the effective power management strategies [60].

On the other hand, DRL improves power efficiency and spectrum efficiency, but needs a heavy computational resources and a large amounts of training data, which might limit its practical utility in actual time situations [61].

n. Channel_occupancy_ratio

One application of cognitive radio technology is dynamic spectrum access (DSA), which, when combined with deep reinforcement learning (DRL), may significantly improve channel occupancy ratio, particularly when employing non-orthogonal multiple access (NOMA) integration. With this method, secondary users (SUs) may mold their spectrum use to efficiently utilize them while avoiding interference with prime users (PUs).

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Dynamic Spectrum Access and Channel Occupancy

- **Cognitive Radio Functionality**: To achieve this, cognitive radios use their transmission strategies to tune with the availability of underutilized frequency bands as occupied by Pus [62].
- **Deep Reinforcement Learning**: DRL, like actor-critic, enables SUs to acquire the best transmission policies on the fly to enhance the channel occupancy ratios by regulating the power assigning and minimizing conflicts [62][63].
- **Occupancy Prediction**: The employment of DRL to anticipate the state of occupancy of the spectrum can be deployed to make better decisions, as it will enable SUs to managerially determine and utilize unoccupied bands efficiently [64].

Non-Orthogonal Multiple Access (NOMA)

- **Enhanced Capacity**: Because several users may use the same frequency channel simultaneously under NOMA, channel occupancy ratios can improve and the number of users sharing the spectrum can grow [63].
- **Interference Management**: Combining NOMA and the DRL may be beneficial in how to address interferences to be able to guarantee that SUs will be able to take action without causing major interference to Pus [65].

Although the combination of both DRL and NOMA are promising on providing better balances between channel occupancy ratios, one still faces difficulties of maintaining good performance across diverse environmental conditions and user demands. There must be further research to ensure that these systems are maximized to work in practice.

o. NOMA_decoding_success_rate

The factor that determines whether NOMA decoding will succeed or fail in the setting that considers the use of cognitive radio and deep reinforcement learning (DRL) to access dynamic spectrum are manifold factors, such as the decoding methods used and the strategizing of providing resources. In cognitive radio networks, throughput and efficiency can be substantially increased by the ONMA technology that enables several users to use the same spectrum. The hybrid with DRL enables flexible learning and resource allocation optimization, enhancing the performance of NOMA networks, as a whole.

NOMA Decoding Techniques

- **Successive Interference Cancellation (SIC)**: SIC, traditionally, is preferred owing to its performance in decoding signals in NOMA. Nevertheless, it is not necessarily ideal especially when the input is in finite alphabet (such as QAM) where it can suffer higher performance deterioration in SIC compared to maximum likelihood (ML) detection given some power distribution properties [66].
- **Maximum Likelihood Detection**: The approach is able to tolerate more power distributions so it is a viable option to be used in situations when SIC fails hence it has the potential of raising the success chances of decoding [66].

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Deep Reinforcement Learning Applications

- **Resource Allocation**: To optimize the available throughput in a cognitive radio network while keeping energy limits in mind, one may employ actor-critical approaches and other DRL-based strategies to adjust the energy allocation among the secondary users [67].
- **Grant-Free NOMA**: In the grant-free NOMA systems, DRL models may implement resource allocation dynamically in order to support as many users as possible, especially in high traffic environments [68][69].

While NOMA decoding success rates can be significantly improved through advanced techniques and DRL-based resource management, challenges remain in optimizing these systems under varying conditions. The balance between decoding methods and resource allocation strategies is crucial for maximizing performance in dynamic environments.

Results with Discussion:

The following plots were obtained using the input and output evaluation parameters.

Figure 1: Plot of Reward versus episodes

As shown in Figure 1, there is a continuous moment of increasing and decreasing rewards, mostly between 4000 and 5000 reward points. Hence, we see that there is always scope for improvement when the episodes run. The average of rewards is nearly 4600 points.

Received: August 04, 2025

International Journal of Applied Mathematics
Volume 38 No. 6s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

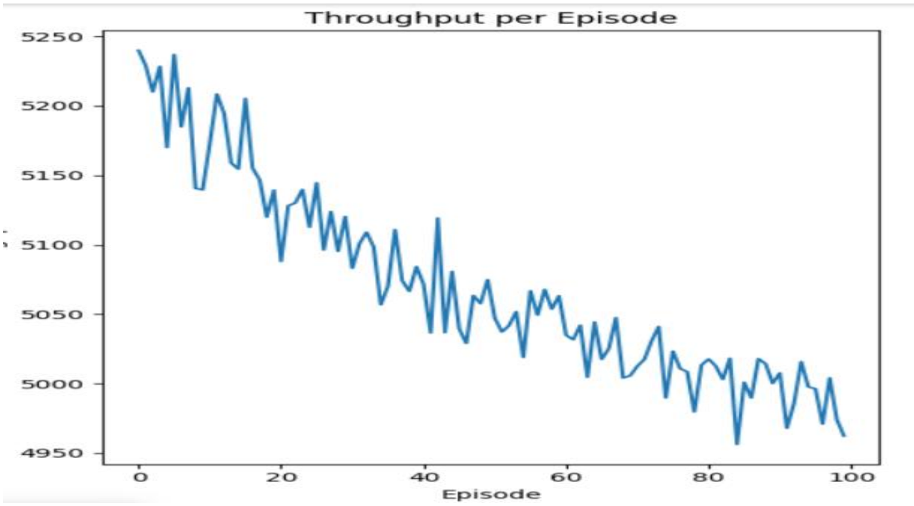

Figure 2: Plot of Throughput versus episodes

As shown in Figure 2, the throughput is continuously decreasing from 5200 to 4950 in the weighted linear approach, which means the model is not appropriate. As throughput decreases from 5200 units to 5000 units, further research must be carried out to make the throughput either increase or remain the same. But anyhow, within 100 episodes of each 50 steps, it decreases to an amount of 200 only.

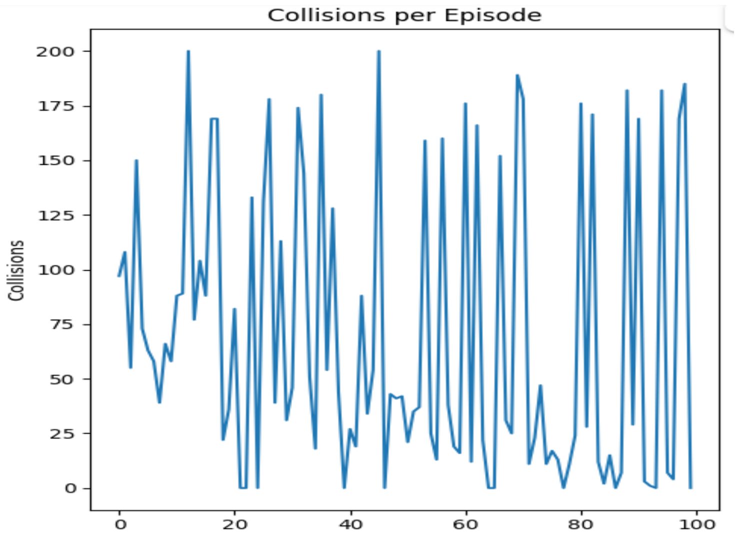

Figure 3: Plot of Collisions versus episodes

As shown in Figure 3, there is a continuous moment of collisions from 0 to 200, with an average of 100 collisions. This means the model developed is not appropriate; hence, it needs to be evaluated in terms of reducing collisions in future research.

Conclusion:

Non-Orthogonal Multiple Access (NOMA), Cognitive Radio (CR), and Deep Reinforcement Learning (DRL) together provide a formidable strategy for improving wireless network performance and coping with spectrum constraints. The proposed framework enables secondary users to dynamically learn and adapt their access strategies based on real-time

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

spectrum conditions, significantly improving throughput, spectral efficiency, and reducing interference with primary users. Simulation outcomes demonstrate promising results, with average reward convergence and reasonable performance metrics in channel occupancy and decoding success rates. However, the model also highlights certain limitations, such as a slight decline in throughput and increased collision rates over episodes, indicating the need for more advanced optimization techniques. Future enhancements should focus on minimizing these drawbacks by refining the reward function, incorporating mobility awareness, and leveraging more robust multi-agent coordination strategies. All things considered, the system lays a smart and extensible groundwork for autonomous spectrum management in next-gen wireless communication networks to be used in the real world.

References:

[1]. Jiang, Z., & Han, L. (2023). *DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios*. https://doi.org/10.1109/iaecst60924.2023.10503376

[2]. Chen, M., Liu, A., Liu, W., Ota, K., Dong, M., & Xiong, N. (2021). RDRL A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. *IEEE Transactions on Network Science and Engineering*. https://doi.org/10.1109/TNSE.2021.3117565

[3]. Wang, X., Teraki, Y., Umehira, M., Zhou, H., & Ji, Y. S. (2022). A Usage Aware Dynamic Spectrum Access Scheme for Interweave Cognitive Radio Network by Exploiting Deep Reinforcement Learning. *Sensors*. https://doi.org/10.3390/s22186949

[4]. Liu, Y., Zhang, X., & Wang, S. (2024). *Research on Distributed Dynamic Spectrum Access Based on Deep Reinforcement Learning*. https://doi.org/10.1007/978-981-99-7505-1_20

[5]. Yang, P., Li, L., Yin, J., Zhang, H., Liang, W., Chen, W., & Han, Z. (2018). Dynamic Spectrum Access in Cognitive Radio Networks Using Deep Reinforcement Learning and Evolutionary Game. *International Conference on Communications*. https://doi.org/10.1109/ICCCHINA.2018.8641242

[6]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[7]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376

[8]. Kaytaz, U., Ucar, S., Akgun, B., & Coleri, S. (2020). Distributed Deep Reinforcement Learning with Wideband Sensing for Dynamic Spectrum Access. Wireless Communications and Networking Conference. https://doi.org/10.1109/WCNC45663.2020.9120840

[9]. Safavinejad, R., Chang, H.-H., & Liu, L. (2024). Deep Reinforcement Learning for Dynamic Spectrum Access: Convergence Analysis and System Design. IEEE Transactions on Wireless Communications. https://doi.org/10.1109/twc.2024.3414428

[10]. Wang, A., Zhang, L., Chen, D., & Chen, J. (2020). Deep Reinforcement Learning for Dynamic Multichannel Access in Multi-Cognitive Radio Networks. https://doi.org/10.1088/1742-6596/1550/3/032135

[11]. Yang, P., Li, L., Yin, J., Zhang, H., Liang, W., Chen, W., & Han, Z. (2018). Dynamic Spectrum Access in Cognitive Radio Networks Using Deep Reinforcement Learning and Evolutionary Game. International Conference on Communications. https://doi.org/10.1109/ICCCHINA.2018.8641242

[12]. Kaytaz, U., Ucar, S., Akgun, B., & Coleri, S. (2020). Distributed Deep Reinforcement Learning with Wideband Sensing for Dynamic Spectrum Access. Wireless Communications and Networking Conference. https://doi.org/10.1109/WCNC45663.2020.9120840

[13]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[14]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[15]. Agarwal, A., Dubey, S., Khan, Mohd. A., Gangopadhyay, R., & Debnath, S. (2016). Learning based primary user activity prediction in cognitive radio networks for efficient dynamic spectrum access. International Conference on Signal Processing. https://doi.org/10.1109/SPCOM.2016.7746632

[16]. Feng, M., Zhang, W., & Krunz, M. (2023). Dynamic Spectrum Access in Non-stationary Environments: A DRL-LSTM Integrated Approach. International Conference on Computing, Networking and Communications. https://doi.org/10.1109/ICNC57223.2023.10074454

[17]. Aygül, M. A., Nazzal, M., & Arslan, H. (2022). Deep RL-Based Spectrum Occupancy Prediction Exploiting Time and Frequency Correlations. 2022 IEEE Wireless Communications and Networking Conference (WCNC). https://doi.org/10.1109/wcnc51071.2022.9771702

[18]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[19]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376

[20]. Liu, Y., Zhang, X., & Wang, S. (2024). Research on Distributed Dynamic Spectrum Access Based on Deep Reinforcement Learning. https://doi.org/10.1007/978-981-99-7505-1_20

[21]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.

[22]. Wang, Y. L., Li, Y., & Shen, B. (2023). Multi-User Dynamic Spectrum Access Based on LRQ Deep Reinforcement Learning Network. International Conference on Advanced Communication Technology. https://doi.org/10.23919/ICACT56868.2023.10079537

[23]. Deep Reinforcement Learning for Distributed and Uncoordinated Cognitive Radios Resource Allocation. (2022). https://doi.org/10.48550/arxiv.2205.13944

[24]. Multi-User Dynamic Spectrum Access Based on LRQ Deep Reinforcement Learning Network. (2023). https://doi.org/10.23919/icact56868.2023.10079537

[25]. Chen, M., Liu, A., Liu, W., Ota, K., Dong, M., & Xiong, N. (2021). RDRL A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2021.3117565

[26]. Wang, X., Teraki, Y., Umehira, M., Zhou, H., & Ji, Y. S. (2022). A Usage Aware Dynamic Spectrum Access Scheme for Interweave Cognitive Radio Network by Exploiting Deep Reinforcement Learning. Sensors. https://doi.org/10.3390/s22186949

[27]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[28]. Safavinejad, R., Chang, H.-H., & Liu, L. (2024). Deep Reinforcement Learning for Dynamic Spectrum Access: Convergence Analysis and System Design. IEEE Transactions on Wireless Communications. https://doi.org/10.1109/twc.2024.3414428

[29]. Giang, H. T. H., Thanh, P. D., & Koo, I. (2020). Dynamic Power Allocation Scheme for NOMA Uplink in Cognitive Radio Networks Using Deep Q Learning. International Conference on Information and Communication Technology Convergence. https://doi.org/10.1109/ICTC49870.2020.9289435

[30]. Lin, Y., Gong, X., Xiong, Y., Li, H., & Wang, X. (2024). Downlink Non-Orthogonal Multiple Access Power Allocation Algorithm Based on Double Deep Q Network for Ensuring User's Quality of Service. Symmetry. https://doi.org/10.3390/sym16121613

[31]. Siddiqi, U. F., Sait, S. M., & Uysal, M. (2020). Deep Reinforcement Based Power Allocation for the Max-Min Optimization in Non-Orthogonal Multiple Access. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3038923

[32]. Deep Reinforcement Learning for Radio Resource Allocation in NOMA-based Remote State Estimation. (2022). https://doi.org/10.48550/arxiv.2205.11861

[33]. Zhang, S., Li, L., Yin, J., Liang, W., Li, X., Chen, W., & Han, Z. (2018). A Dynamic Power Allocation Scheme in Power-Domain NOMA using Actor-Critic Reinforcement Learning. International Conference on Communications. https://doi.org/10.1109/ICCCHINA.2018.8641248

[34]. Feng, M., Zhang, W., & Krunz, M. (2023). Dynamic Spectrum Access in Non-stationary Environments: A DRL-LSTM Integrated Approach. International Conference on Computing, Networking and Communications. https://doi.org/10.1109/ICNC57223.2023.10074454

[35]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.

[36]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[37]. Wang, X., Teraki, Y., Umehira, M., Zhou, H., & Ji, Y. S. (2022). A Usage Aware Dynamic Spectrum Access Scheme for Interweave Cognitive Radio Network by Exploiting Deep Reinforcement Learning. Sensors. https://doi.org/10.3390/s22186949

[38]. RDRL: A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. (2022). IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/tnse.2021.3117565

[39]. Chen, M., Liu, A., Liu, W., Ota, K., Dong, M., & Xiong, N. (2021). RDRL A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2021.3117565

[40]. Cooperative Multi-Agent Reinforcement-Learning-Based Distributed Dynamic Spectrum Access in Cognitive Radio Networks. (2022). IEEE Internet of Things Journal. https://doi.org/10.1109/jiot.2022.3168296

[41]. Yang, P., Li, L., Yin, J., Zhang, H., Liang, W., Chen, W., & Han, Z. (2018). Dynamic Spectrum Access in Cognitive Radio Networks Using Deep Reinforcement Learning and Evolutionary Game. International Conference on Communications. https://doi.org/10.1109/ICCCHINA.2018.8641242

[42]. RDRL: A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. (2022). IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/tnse.2021.3117565

[43]. Chen, M., Liu, A., Liu, W., Ota, K., Dong, M., & Xiong, N. (2021). RDRL A Recurrent Deep Reinforcement Learning Scheme for Dynamic Spectrum Access in Reconfigurable Wireless Networks. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2021.3117565

[44]. Liston, M. J., & Dandekar, K. R. (2021). Entropy Based Exploration in Cognitive Radio Networks using Deep Reinforcement Learning for Dynamic Spectrum Access. Wireless and Microwave Technology Conference. https://doi.org/10.1109/WAMICON47156.2021.9444294

[45]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376

[46]. Jalil, S. Q., Rehmani, M. H., & Chalup, S. K. (2020). A Deep Reinforcement Learning Approach to Fair Distributed Dynamic Spectrum Access. International Conference on Mobile and Ubiquitous Systems: Networking and Services. https://doi.org/10.1145/3448891.3448935

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[47]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4
[48]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376
[49]. Dynamic spectrum access based on double deep Q-network and convolution neural network. (2022). 2022 24th International Conference on Advanced Communication Technology (ICACT). https://doi.org/10.23919/icact53585.2022.9728805
[50]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4
[51]. Chang, H.-H., Song, H., Yi, Y., Zhang, J., He, H., & Liu, L. (2019). Distributive Dynamic Spectrum Access Through Deep Reinforcement Learning: A Reservoir Computing-Based Approach. IEEE Internet of Things Journal. https://doi.org/10.1109/JIOT.2018.2872441
[52]. Safavinejad, R., Chang, H.-H., & Liu, L. (2024). Deep Reinforcement Learning for Dynamic Spectrum Access: Convergence Analysis and System Design. IEEE Transactions on Wireless Communications. https://doi.org/10.1109/twc.2024.3414428
[53]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.
[54]. Safavinejad, R., Chang, H.-H., & Liu, L. (2024). Deep Reinforcement Learning for Dynamic Spectrum Access: Convergence Analysis and System Design. IEEE Transactions on Wireless Communications. https://doi.org/10.1109/twc.2024.3414428
[55]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.
[56]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376
[57]. Zeng, L., Yao, F., Zhang, J., & Jia, M. (2022). Dynamic spectrum access based on prior knowledge enabled reinforcement learning with double actions in complex electromagnetic environment. China Communications. https://doi.org/10.23919/jcc.2022.07.002
[58]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4
[59]. Jiang, Z., & Han, L. (2023). DRL-based Dynamic Spectrum Access for Cognitive Radio Networks in Mobility Scenarios. https://doi.org/10.1109/iaecst60924.2023.10503376
[60]. Kaytaz, U., Ucar, S., Akgun, B., & Coleri, S. (2020). Distributed Deep Reinforcement Learning with Wideband Sensing for Dynamic Spectrum Access. Wireless Communications and Networking Conference. https://doi.org/10.1109/WCNC45663.2020.9120840

Received: August 04, 2025

[61]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.

[62]. Dong, L., Qian, Y., & Xing, Y. (2022). Dynamic spectrum access and sharing through actor-critic deep reinforcement learning. Eurasip Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-022-02124-4

[63]. Safavinejad, R., Chang, H., & Liu, L. (2023). DRL meets DSA Networks: Convergence Analysis and Its Application to System Design.

[64]. Aygül, M. A., Nazzal, M., & Arslan, H. (2022). Deep RL-Based Spectrum Occupancy Prediction Exploiting Time and Frequency Correlations. 2022 IEEE Wireless Communications and Networking Conference (WCNC). https://doi.org/10.1109/wcnc51071.2022.9771702

[65]. Unnikrishnan, J., & Veeravalli, V. V. (2008). Dynamic spectrum access with learning for cognitive radio. Asilomar Conference on Signals, Systems and Computers. https://doi.org/10.1109/ACSSC.2008.5074371

[66]. Qit, Y., & Vaezi, M. (2024). NOMA Decoding: Successive Interference Cancellation or Maximum Likelihood Detection? https://doi.org/10.1109/ciss59072.2024.10480218

[67]. Giang, H. T. H., Hoan, T. N. K., & Koo, I. (2021). Uplink NOMA-based long-term throughput maximization scheme for cognitive radio networks: an actor–critic reinforcement learning approach. Wireless Networks. https://doi.org/10.1007/S11276-020-02520-Y

[68]. Liu, Y., Deng, Y., Zhou, H., Elkashlan, M., & Nallanathan, A. (2023). Deep Reinforcement Learning-Based Grant-Free NOMA Optimization for mURLLC. IEEE Transactions on Communications. https://doi.org/10.1109/TCOMM.2023.3238061

[69]. Deep Reinforcement Learning-Based Grant-Free NOMA Optimization for mURLLC. (2023). IEEE Transactions on Communications. https://doi.org/10.1109/tcomm.2023.3238061