

Article

Not peer-reviewed version

# A Multi-Objective Genetic Algorithm-Deep Reinforcement Learning Framework for Spectrum Sharing in 6G Cognitive Radio Networks

[Ancilla Wadzanai Chigaba](#) , [Sindiso Mpenyu Nleya](#) <sup>\*</sup> , [Mthulisi Velempini](#) , [Samkeliso Suku Dube](#)

Posted Date: 25 July 2025

doi: 10.20944/preprints202507.2093.v1

Keywords: 6G; cognitive radio networks; spectrum sharing; deep reinforcement learning; NSGA-II; PPO; multi-objective optimisation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

*Article*

# A Multi-Objective Genetic Algorithm-Deep Reinforcement Learning Framework for Spectrum Sharing in 6G Cognitive Radio Networks

Ancilla Wadzanai Chigaba<sup>1</sup>, Sindiso Mpenyu Nleya<sup>2,\*</sup>, Mthulisi Velempini<sup>2</sup>  
and Samkeliso Suku Dube<sup>1</sup>

<sup>1</sup> National University of Science and Technology, Zimbabwe

<sup>2</sup> University of Limpopo

\* Correspondence: mthulisi.velempini@ul.ac.za

## Abstract

The exponential growth in wireless communication demands intelligent and adaptive spectrum-sharing solutions, especially within dynamic and densely populated 6G cognitive radio networks (CRNs). This paper introduces a novel hybrid framework combining the Non-dominated Sorting Genetic Algorithm II (NSGA-II) with Proximal Policy Optimisation (PPO) for multi-objective optimisation in spectrum management. The proposed model balances spectrum efficiency, interference mitigation, energy conservation, collision rate reduction, and QoS maintenance. Evaluation on synthetic and ns-3 datasets shows that the NSGA-II and PPO hybrid consistently outperforms Random, Greedy, and standalone PPO strategies, achieving higher cumulative reward, perfect fairness (Jain's Index = 1.0), robust hypervolume convergence (65.1%), up to 12% reduction in PU collision rate, 20% lower interference, and approximately 40% improvement in energy efficiency. These findings validate the framework's effectiveness in promoting fairness, reliability, and efficiency in 6G wireless communication systems.

**Keywords:** 6G; cognitive radio networks; spectrum sharing; deep reinforcement learning; NSGA-II; PPO; multi-objective optimisation

## 1. Introduction

Rapid advancements in wireless communication technology has driven the evolution between 5G and 6G networks, which is expected to make the Internet of Everything (IoE) feasible. The IoE depicts a huge network in which locations, people, and things are connected and can exchange services.

The emergence of new business and applications, the spread of intelligence, sustainability, and social responsibility are some of the diverse trends and forces driving the evolution towards 6 G to date [1]. The 6G networks for communications are anticipated to bring about a new dawn of mainstream applications, such as Industry 5.0, Smart Grid 2.0, Holographic Telepresence (HT), Extended Reality (XR), Unmanned Aerial Vehicles (UAVs), and even deep-sea and space travel [2]. Moreover, 6G vision extends beyond communications, to include innovations, the rise of mass-based Artificial Intelligence (AI), as well as creation of massive digital twins. Undoubtedly, 6G will transform our economy by redefining the economic landscape for 2030 and beyond [3]. This exemplified by global economic value estimates which point to 6G adding trillions of dollars to the global economy by enabling transformative technologies and services. The changes will take the shape of technological advancements in manufacturing, healthcare, agriculture, and transportation [4]. Also 6G promises to revolutionize contemporary civilization and set the standard for an era of unmatched efficiency, innovation, and connectedness [5].

More importantly, 6G can be perceived as a massive distributed neural network that combines computing, communication, and sensing in a seamless manner. This creates an era where everything is genuinely sensed, connected, and intelligent, blurring the boundaries between our physical, biological,

and digital worlds. Consequently, 6G network development faces significant traffic and network engineering challenges. The International Telecommunication Union (ITU) predicts that by 2030, mobile data traffic would extend up to 5 zettabytes worldwide [6]. Effective traffic engineering (TE) is critical in managing the anticipated exponential increase in data traffic generated by these next-generation applications. Spectrum scarcity, energy efficiency, and security vulnerabilities present significant hurdles that need to be taken care of in order to guarantee sustainable and efficient network operations [7]. Similarly, [8][9] identify a slightly different set of challenges, including resource allocation, spectrum management, waveform design, channel modeling, interference control, energy efficiency, security, and artificial intelligence (AI) integration. Clearly, spectrum scarcity and under-utilization are the main issues preventing the transition to 6G networks. Thus integrating cognitive radio networks (CRNs) with upcoming 6G technologies enhances this capability, as spectrum utilization is optimized using cognitive radio (CR) by tapping into frequencies (unused) if primary mobile components (MCs) are not active [10,11]. In 6G cognitive radio networks, resource allocation for secondary users prioritizes fairness, utilizing multi-channel communications to ensure equitable access irrespective of individual channel capacity [12]. Resource allocation and Spectrum management are crucial components of 6G networks and are intricately linked. In general, spectrum management is a specialized form of resource allocation focused on the radio frequency spectrum. A crucial component of contemporary wireless communication infrastructure is spectrum management (SM), which entails the meticulous distribution, control, and synchronization of the radio frequency spectrum [10]. Given the increased and ever growing spectrum demand, there will be significantly more competition for specific frequency bands, making effective use of that spectrum resource even more important. To this end, managing the spectrum effectively is motivated by several factors, such as the need to mitigate harmful interference to safeguard frequencies utilized for critical services and to identify opportunities that maximize efficiency. Moreover, it enables development of new technologies as well as deployment within flexible frameworks subsequently reducing telecommunications equipment cost [13]. In the 6G context, the spectrum management process is premised on a layered framework. The layered spectrum framework comprises 4 layers namely, policy and regulatory authorities layer, spectrum management systems layer, components of spectrum management layer and spectrum management technologies layer [10]. Collectively, throughout the layers, the framework strives to ensure protecting privacy and security in spectrum management procedures and also leverages on artificial intelligence to enhance prediction, planning, real-time allocation, and enforcement.

On the other-hand, to realize use cases, 6G networks will need to harness many cutting edge technologies, such as machine learning (ML), AI, CRNs, post quantum cryptography and automation according to the Internet of Things (IoT) [14]. Specifically, [15] explains how cognitive radio (CR) can improve spectrum efficiency, improve network resilience, and enable novel use cases in 6G. Clearly, in order to achieve these goals, dynamic spectrum allocation and interference mitigation are critical functions of CRNs, especially in settings where spectrum congestion is present. There is a significant chance to optimize 6G networks through the combination of deep learning and cognitive radio networks. This collaboration enables the management of autonomous real-time spectrum and improves network operations' overall intelligence [16].

To support envisioned capabilities, several transformative technologies must converge. These technologies will include, networked sensing, native AI, Integrated Non-Terrestrial Networks (NTN), extreme connectivity, native sustainability and dependability. Native AI means that AI will be incorporated into communication functionalities from the beginning of system design through the development, management, and operation of systems to achieve the desired performance improvements [17]. Currently, to make AI native 6 G a reality, MITRE is collaborating with NVIDIA [18]. This is accomplished by incorporating AI into 6G help address a range of issues, such as improving service delivery as well as releasing the necessary spectrum availability to support the expansion of wireless. Native AI and networked sensing are closely intertwined and provide a critical synergy for intelligent networks of the future. By investigating radio wave transmission, reflection, and scattering,

the communication system as a whole can act as a sensor to sense and comprehend the physical world, eventually offering a wide range of new services. Furthermore, the 6G's Key Performance Indicators (KPIs), like transmission reliability, data rate, and communication latency, are subject to increased demands due to recently developed extreme connectivity[19]. (Integrated NTN). Access through endpoints numbering in the billions and millions of sub-networks that are generally untrustworthy, disaggregation of architecture and open interfaces, a heterogeneous cloud environment, and a combination of open source and multi-vendor software are just a few of the ways that the 6G system will simultaneously challenge and threaten trustworthiness. There will probably be an increase in both AI/ML-based attacks and attacks against AI/ML-based ideas [20]. In the 2030s, reliable wireless networks will serve as the basis for reliable use cases and apps for both businesses and consumers. In order to lessen the global environmental impact of ICT and (CO<sub>2</sub>), 6G ought to be a source of energy-optimized digital framework applicable to a range of consumer groups and industrial domains. Furthermore, 6G ought to be sustainable from a broader standpoint, taking into account not just energy-related issues but also the use of natural resources, lifecycles of products and social sustainability, etc. The digital fabric of it will also enable real-time sensing and understanding of the physical world's condition, enhancing services that will increase sustainability and cost effectiveness and help achieve sustainability objectives. Ultimately, extreme connectivity, Integrated Non-Terrestrial Networks (NTN), networked sensing, native AI, sustainability, and native trustworthiness are not only foundational pillars but also intertwined and subsequently enabler of 6G use cases. Some of the most prominent envisaged potential 6G use cases include, mixed reality, global internet, autonomous mobility, spatial data, critical services, massive digital twins, and AI communication [21].

In our design of a 6G CRN, we examine three competing goals spectrum efficiency, interference reduction, and energy consumption, which are inherently conflicting objectives when it comes to optimization. Wireless network optimization frequently addresses several performance factors, such as interference reduction or spectrum efficiency, as distinct goals. However, single objective optimization has a serious drawback in that it can easily ignore or even deteriorate other important aspects of network performance by concentrating on increasing just one statistic. This is due to the fact that many network objectives are intrinsically conflicting and interconnected, so improving one may have a detrimental effect on another [22]. For example reducing interference and energy use frequently clashes with attaining high spectrum efficiency (SE) [23,24].

However, optimizing these networks is a challenging task, particularly when multiple, conflicting objectives are at play [25]. According to [26] optimizing networks with multiple objectives is a complex task that involves improving network performance and efficiency while considering various constraints and performance metrics. To improve spectrum management, this paper utilizes a multi-objective approach to optimize a 6G network's spectrum efficiency, interference, and energy consumption.

The improved spectrum management is achieved through the introduction of a novel hybrid framework integrating the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [27] with Proximal Policy Optimisation (PPO) for multi-objective optimisation. Our key contributions are thus as follows: firstly an implementation of (NSGA-II) for generating Pareto-optimal solutions balancing spectrum efficiency, interference, and energy consumption. The NSGA-II balances the three conflicting objectives of spectrum efficiency, interference and energy consumption. Essentially, depicting how the available spectrum is available with regards to the interference which represents an unwanted signal overlap that degrades quality and the amount of energy expended during the network operation. Secondly, an integration of Proximal Policy Optimization (PPO) for real-time policy learning using NSGA-II guided rewards. Specifically, the PPO is a reinforcement learning (RL) algorithm designed to train intelligent agents in complex 6G Cognitive Radio network environment. Our third contribution measures how well the hybrid NSGA-II + PPO framework performs when benchmarked against existing baseline techniques. The rest of the paper is structured as follows: Section II presents the related work which audits the range of efforts put towards multi-objective optimization of 6G networks. Section III details the system model illustrating the configuration of a Cognitive Radio Network within a 6G environment.



The evaluation of the 6G Cognitive Radio Network (CRN) is presented in Section IV, with the paper concluding in Section V.

2. Related Work

Wireless technologies are advancing rapidly, with increasing demand for, reliable, and efficient communication methods. This has accelerated the push towards 6G networks. However, challenges like limited spectrum availability, growing user demands, and interference issues make dynamic spectrum access a vital part of making 6G Cognitive Radio Networks (CRNs) work effectively. Effective spectrum management has thus become crucial as the wireless landscape evolves into its sixth generation. The frameworks for hybrid adaptive solutions that combine dynamic spectrum access (DSA), cooperative access, Non-Orthogonal Multiple Access (NOMA), game-theoretic frameworks, blockchain-based systems, and UAV-aided sensing have improved spectral efficiency and fairness. For example, [28] proposed a DSA system enabling secondary users to access underutilised channels through environmental sensing, improving spectrum efficiency but demanding accurate sensing and coordination. Similarly, [29] explored DSA within satellite-terrestrial networks, highlighting the complexity of managing overlapping domains and inefficient coordination mechanisms. Cognitive NOMA (CR-NOMA) approaches, investigated by [30] and [29], enable simultaneous multi-user frequency sharing by power-level differentiation, yielding substantial throughput gains but presenting challenges in power allocation complexity and fairness under decentralised deployments.

Table 1. Condensed ML-Based Spectrum Sharing in CRNs with Gaps

| Author | Method                          | Pros                                    | Cons                                    | Gap                                          |
|--------|---------------------------------|-----------------------------------------|-----------------------------------------|----------------------------------------------|
| [31]   | ML Survey in CRN, LTE-U         | Broad taxonomy; highlights ML potential | Lacks validation; general scope         | Need robust ML vs. PUE, SSDF attacks         |
| [32]   | Multi-scenario ML sharing       | Better incentives; fewer collisions     | No implementation constraints discussed | Lack of real-world scalability eval          |
| [33]   | ML (SVM, KNN, RF) for IIoT/IoMT | Higher accuracy; fewer false positives  | Missing deployment context              | Need deployment analysis for IIoT/IoMT       |
| [34]   | MLP, SVM, NB vs. detection      | MLP balances speed and accuracy         | No robustness/scalability analysis      | Evaluate in mobile/adversarial settings      |
| [35]   | DRL for secure sharing          | Boosts utilization; security-aware      | Theoretical; lacks empirical proof      | Empirical DRL-based secure sharing needed    |
| [36]   | Margin-based online learning    | Adapts with minimal labels              | Simulation-only; no field test          | Real PU behavior-responsive models needed    |
| [37]   | PPO for RIS in V2X              | Fast convergence; better sum-rate       | No dense urban/blockage analysis        | Test PPO-RIS in obstructed settings          |
| [38]   | MAPPO for HetNet offloading     | Efficient, scalable BS deployment       | Multi-agent training complexity ignored | Study MAPPO fairness/scalability under noise |

Hybrid deep reinforcement learning (DRL) approaches have recently emerged to overcome the limitations of traditional and ML-only models in 6G CRNs by combining adaptive learning with multi-objective optimisation and evolutionary strategies. These hybrid frameworks offer promising solutions for decentralised, scalable, and real-time spectrum sharing under complex and dynamic conditions. For instance, [39] applied DRL agents at access points within cell-free MIMO networks to facilitate inter-operator spectrum coordination with low signalling overhead. However, the effects of large-scale MNO densification remain unstudied. Huang et al. [40] developed DRL-based autonomous

spectrum access strategies for D2D-enabled IoT networks, achieving high throughput without prior access rules but lacking scalability assessments. [41] improved spectrum access and power control with DRL in CRNs, reducing collisions and improving rewards, but did not evaluate robustness under interference and mobility. Hybrid discrete-continuous action space DRL algorithms proposed by [42] deliver near-optimal throughput with reduced sensing overhead but require standardised benchmarking and scalability studies.

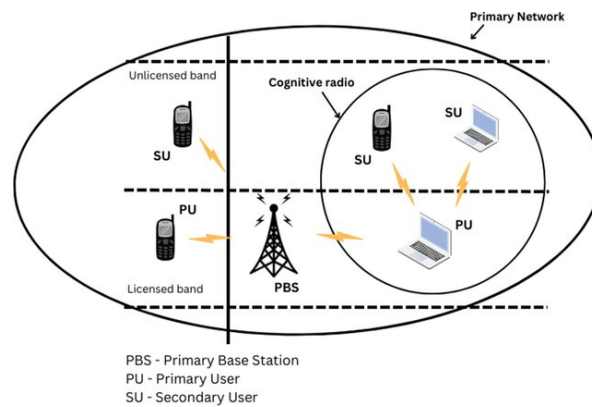
Multi-agent DRL methods such as MADDPG have enhanced cooperative spectrum sensing with reduced communication overhead [43]. At the same time, federated learning and graph convolutional networks have enabled scalable learning in integrated satellite-terrestrial networks [44]. Distributed Q-learning frameworks for secure spectrum sharing in ultra-dense networks [45] and consensus-based reinforcement learning for decentralised spectrum management [46] offer promising decentralised solutions, albeit with unaddressed performance under adversarial and interference conditions. MARL architectures tackling partial observability and large action spaces show improved convergence but await real-world validation [47]. Blockchain-enhanced evolutionary algorithms propose secure resource optimisation but require deeper integration with learning models [44]. These hybrid DRL frameworks have demonstrated improved adaptiveness, throughput, and energy efficiency but face challenges in scalability, generalisability, convergence speed, and sensitivity to environmental uncertainties typical of 6G scenarios. The proposed NSGA-II + PPO model bridges these gaps by coupling Pareto-optimal evolutionary search with actor-critic policy learning in a modular, validated simulation environment. Table 2 summarises recent hybrid DRL-based spectrum-sharing research.

**Table 2.** Compact Summary of DRL-Based Spectrum Sharing in 6G CRNs

| Author | Method                           | Domain                         | Pros                                | Cons                                   |
|--------|----------------------------------|--------------------------------|-------------------------------------|----------------------------------------|
| [39]   | DRL at APs for coordination      | Cell-Free MIMO, multi-operator | Low signalling, scalable            | Dense MNOs not addressed               |
| [40]   | Autonomous DRL access            | D2D IoT                        | High throughput, no rule dependency | No scalability/user dynamics           |
| [41]   | DRL for access                   | CRNs with SU optimization      | Fewer collisions, higher reward     | No robustness to mobility/interference |
| [42]   | Hybrid DRL (discrete-continuous) | Energy-aware CRNs              | 99.4% optimal throughput            | No scalability or hybrid RL benchmarks |

3. System Model

The system model depicts a 6G CRN environment comprising  $N$  authorised Primary Users (PUs) and  $M$  unlicensed Secondary Users (SUs). Each PU holds exclusive rights to a spectrum band, while SUs opportunistically access idle channels without causing interference to active PUs. Secondary users must continuously sense and vacate bands when PUs return. Infrastructure-based and ad hoc CRN architectures are considered, with licensed and unlicensed spectrum access enabling coexistence between primary and secondary transmissions, as illustrated in Figure 1.



**Figure 1.** Cognitive Radio Network architecture

Dynamic spectrum allocation determines which SUs are assigned to which spectrum bands simultaneously, facilitating spectrum sharing while respecting interference and quality of service constraints. A binary decision variable models this allocation:

$$x_{i,j} \in \{0, 1\} \quad (1)$$

where  $x_{i,j} = 1$  indicates that SU  $i$  is assigned channel  $j$ , and 0 otherwise.

System performance is evaluated through the following metrics:

Signal-to-Interference-plus-Noise Ratio (SINR):

$$\text{SINR}_{i,j} = \frac{P_{i,j}G_{i,j}}{\sum_{k \neq i} P_{k,j}G_{k,j} + I_j + N_0B_j} \quad (2)$$

where  $P_{i,j}$  is the transmission power of SU  $i$  on band  $j$ ,  $G_{i,j}$  the channel gain for SU  $i$ ,  $P_{k,j}$  and  $G_{k,j}$  denote the power and channel gain of interfering SUs,  $I_j$  is the interference from active PUs,  $N_0$  the noise spectral density, and  $B_j$  the bandwidth of band  $j$ .  $N_0$  depicts how noise power is distributed over the communication system's various frequencies. In essence, it is the noise power per bandwidth unit, commonly expressed in Joules (J) or Watts per Hertz (W/Hz).  $B_j$  symbolizes the frequency range over which the receiver  $j$  can process or function while receiving a signal.

Objective Function 1 – Spectrum Efficiency (SE):

$$f_1(X) = \sum_{i=1}^{N_{SU}} \sum_{j=1}^{N_{bands}} x_{i,j} B_j \log_2(1 + \text{SINR}_{i,j}) \quad (3)$$

measuring the aggregate achievable throughput across secondary users.

Objective Function 2 – Interference to Primary Users (IL):

$$f_2(\mathbf{X}) = \sum_{i=1}^{N_{SU}} \sum_{j=1}^{N_{bands}} x_{i,j} I_{i,j} P_{i,j} \quad (4)$$

quantifying the total interference induced by secondary transmissions on PUs.

Objective Function 3 – Energy Consumption per Data Rate (EC):

$$f_3(X) = \sum_{i=1}^{N_{SU}} \sum_{j=1}^{N_{bands}} x_{i,j} \frac{P_{i,j}}{R_{i,j}} \quad (5)$$

Capturing energy efficiency by measuring the power consumed per unit data rate.

### 3.1. Multi-Objective Optimisation Problem Formulation

Efficient spectrum allocation must balance conflicting objectives inherent in CRNs. Maximising spectrum efficiency alone risks increased interference and energy consumption, whereas minimising interference or power may limit spectrum usage and network capacity. This leads to the formulation of the problem as a multi-objective optimization problem:

$$\mathbf{f}(X) = [\max f_1(X), \min f_2(X), \min f_3(X)] \quad (6)$$

subject to the constraints:

$$\text{SINR}_{i,j} \geq \text{SINR}_{\min} \quad \forall i, j \quad (7)$$

$$\sum_{j=1}^{N_{\text{bands}}} P_{i,j} \leq P_{\max} \quad \forall i \quad (8)$$

$$\sum_{i=1}^{N_{\text{SU}}} x_{i,j} I_{i,j} \leq I_{\max} \quad \forall j \quad (9)$$

$$\sum_{j=1}^{N_{\text{bands}}} x_{i,j} \leq 1 \quad \forall i \quad (10)$$

$$x_{i,j} \in \{0, 1\} \quad \forall i, j \quad (11)$$

where the SINR threshold ensures reliable communication, power budgets restrict transmission energy, interference constraints protect PUs, and allocation limits restrict each SU to at most one channel. The constraint  $\text{SINR}_{i,j} \geq \text{SINR}_{\min} \quad \forall i, j$  enforces a minimum quality level for communication between transmitter  $i$  and receiver  $j$ . In practice, this means that the signals are sufficiently strong in relation to noise and interference levels. Thus, this guarantees the maintenance of Quality of Service (QoS), particularly in crucial real-time applications like autonomous systems or video streaming. The next constraint  $\sum_{j=1}^{N_{\text{bands}}} P_{i,j} \leq P_{\max} \quad \forall i$  pertains to the total transmission power allocated by user or transmitter  $i$  across all frequency bands  $j$  must not exceed a given threshold  $P_{\max}$ . In our design, this constraint conveniently prevents the over-consumption of power subsequently promoting an energy aware network. Moreover, it also influences QoS and Interference Control by helping balance signal quality and limit interference to nearby users or systems. The design is also informed by interference issues wherein  $\sum_{i=1}^{N_{\text{SU}}} x_{i,j} I_{i,j} \leq I_{\max} \quad \forall j$  ensures that the total interference from all selected secondary users on each band stays within a permissible limit  $I_{\max}$ . Thus this constraint prevents harmful interference with primary users (PUs), maintaining coexistence as well as encourages spectrum efficiency while preserving signal integrity. It also acts as a key boundary in optimization models for spectrum allocation and supports regulatory compliance and spectrum etiquette. The constraint,  $\sum_{j=1}^{N_{\text{bands}}} x_{i,j} \leq 1 \quad \forall i$  is concerned with secondary users and enforces a mutual exclusivity rule wherein it becomes relevant particularly relevant when each entity must commit to a single option only among a predefined set. This means in our CRN, this constraint limits each SU to a set of channels or alternatively a single channel thereby avoiding cross band interference. The final constraint,  $x_{i,j} \in \{0, 1\} \quad \forall i, j$  is a binary domain constraint and plays a foundational role in many optimization and decision-making models. The binary decision variables like this one are often the backbone of optimization logic, controlling user-band assignments, interference mitigation, and fairness guarantees.

Following [48], the objective is to determine the trade-offs represented by the Pareto-optimal front of non-dominated solutions among spectrum efficiency, interference, and energy consumption.

### 3.2. Hybrid NSGA-II + PPO Framework

A modular two-stage hybrid framework is proposed to tackle the multi-objective spectrum sharing challenge in dynamic 6G Cognitive Radio Networks (CRNs). The first stage utilises the NSGA-



II to generate a diverse set of Pareto-optimal spectrum allocation solutions simultaneously optimising key objectives: spectrum efficiency, interference mitigation, and energy consumption. These steps are detailed in Algorithm ??.

In the second stage, Proximal Policy Optimisation (PPO) is employed to refine these Pareto-optimal solutions into adaptive policies capable of making real-time spectrum access decisions. The PPO training procedure is outlined in Algorithm ??.

The sequential integration of NSGA-II's global, multi-objective search with PPO's local policy refinement balances thorough exploration of the complex solution space with efficient exploitation and adaptability to dynamic network conditions. This synergy enables robust, scalable, and flexible spectrum sharing in 6G CRNs.

Figure 2 presents the end-to-end flowchart of the hybrid framework, illustrating the interaction between the NSGA-II evolutionary search, policy pretraining via imitation learning, and subsequent reinforcement learning fine-tuning within the cognitive radio environment.

---

**Algorithm 1** NSGA-II for Multi-Objective Spectrum Allocation
 

---

**Input** : Population size  $n$ , number of generations  $g$ , objective functions  $f_1(X)$ ,  $f_2(X)$ ,  $f_3(X)$

**Output**: Pareto-optimal set  $\mathcal{P}^*$

Initialise population  $\mathcal{P}_0$  with  $n$  individuals **for**  $t \leftarrow 1$  **to**  $g$  **do**

Evaluate fitness using  $f_1, f_2, f_3$  Perform non-dominated sorting and assign ranks Calculate the crowding distance for each individual Select parents via binary tournament selection **foreach** pair of parents **do**

**Crossover**: Apply Simulated Binary Crossover (SBX):

$$c_1 = 0.5[(1 + \beta_q)p_1 + (1 - \beta_q)p_2],$$

$$c_2 = 0.5[(1 - \beta_q)p_1 + (1 + \beta_q)p_2],$$

$$\text{where } \beta_q = \begin{cases} (2u)^{\frac{1}{\eta_c+1}}, & u \leq 0.5 \\ \left[\frac{1}{2(1-u)}\right]^{\frac{1}{\eta_c+1}}, & u > 0.5 \end{cases}'$$

and  $u \sim \mathcal{U}[0, 1]$  **Mutation**: Apply Polynomial Mutation:

$$x' = x + \Delta x, \quad \Delta x = \begin{cases} \delta(x_{\max} - x), & r < 0.5 \\ -\delta(x - x_{\min}), & r \geq 0.5 \end{cases}$$

$$\text{where } \delta = \begin{cases} (2r)^{\frac{1}{\eta_m+1}} - 1, & r < 0.5 \\ 1 - [2(1-r)]^{\frac{1}{\eta_m+1}}, & r \geq 0.5 \end{cases}, r \sim \mathcal{U}[0, 1]$$

**end**

Evaluate offspring fitness and merge populations. Perform non-dominated sorting and select the top  $n$  individuals.

**end**

Return Pareto-optimal set  $\mathcal{P}^*$

---

---

**Algorithm 2** Proximal Policy Optimisation (PPO) for Policy Refinement

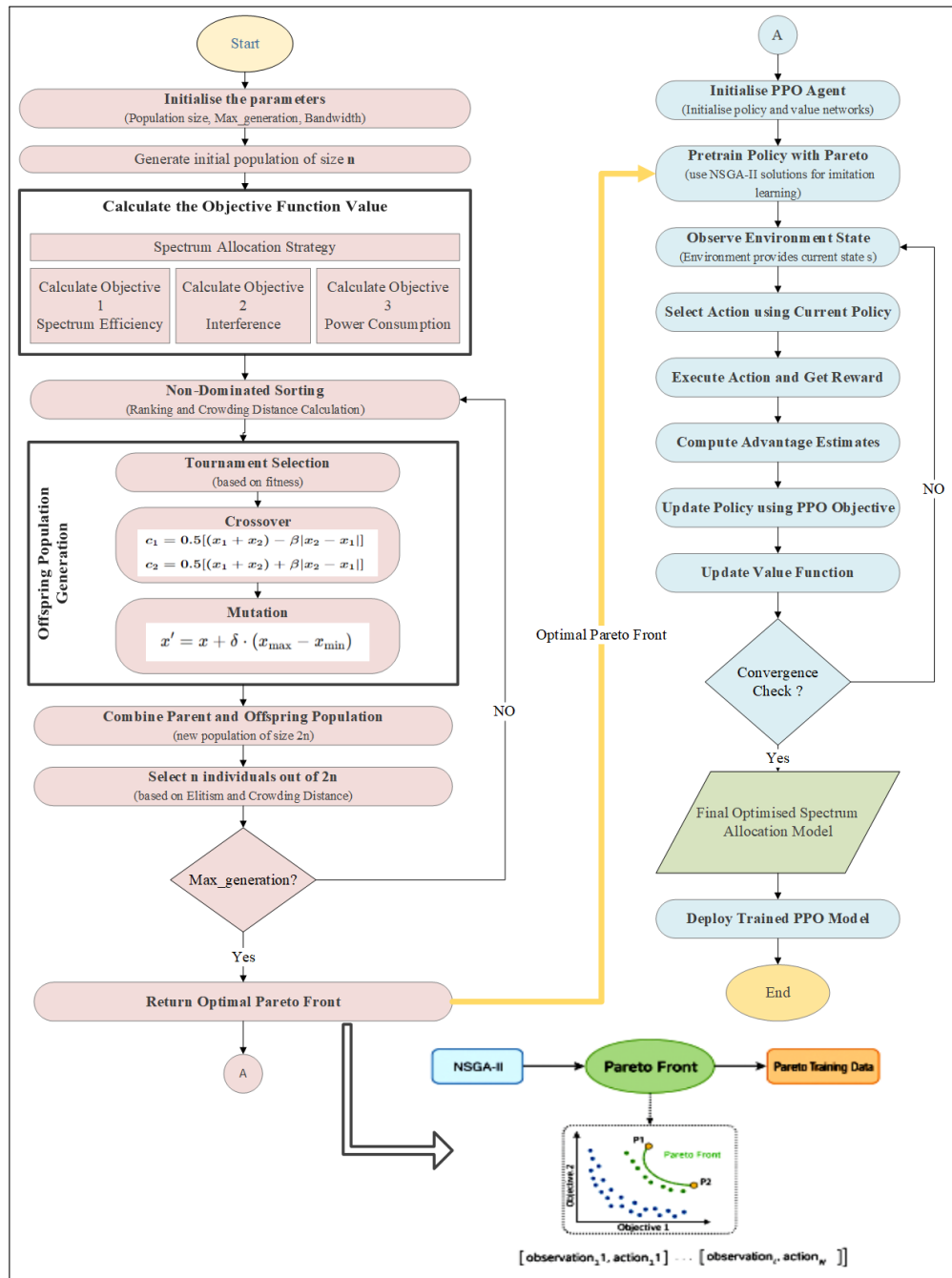
---

**Input** : Pareto-optimal set  $\mathcal{P}^*$ , learning rate  $\alpha$ , discount factor  $\gamma$ , clipping parameter  $\epsilon$ **Output**: Trained policy  $\pi_\theta(a|s)$ Initialize policy network  $\pi_\theta$  and value function  $V_\phi$  Pretrain  $\pi_\theta$  via imitation learning from  $\mathcal{P}^*$  observation-action pairs **for** each training episode **do**    **repeat**        Observe state  $s_t$  Sample action  $a_t \sim \pi_\theta(a|s_t)$  Execute  $a_t$ , observe reward  $r_t$  and next state  $s_{t+1}$  Compute advantage estimate  $A_t = r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t)$  Update  $\pi_\theta$  by maximising clipped surrogate objective:

$$L^{\text{CLIP}}(\theta) = \mathbb{E}_t[\min(r_t(\theta)A_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)A_t)]$$

        Update value function  $V_\phi$  by minimising squared error    **until** convergence;**end****return** Trained policy  $\pi_\theta$ 

---



**Figure 2.** End-to-End Flowchart of the Proposed Hybrid NSGA-II + PPO Framework for Dynamic Spectrum Allocation in 6G CRNs

The suggested hybrid model combines PPO for adaptive policy refinement with NSGA-II for multi-objective spectrum allocation. Despite its effectiveness, this synergy adds significant computational demands at both phases.

1. **NSGA-II( Evolutionary Search Complexity):** Population Dynamics: NSGA-II operates over multiple generations with a sizable population, demanding repeated evaluations of each individual across all objectives. This gives rise to  $O(Mn^2)$  complexity for non-dominated sorting, where  $n$  is the population size and  $M$  is the number of objectives.
2. **PPO- Reinforcement Learning Stability:** In algorithm ?? policy training leverages Pareto-optimal solutions to bootstrap PPO adds a front-loaded cost due to imitation learning over diverse-action pairs.

In simulation settings, this hybrid algorithmic design is generally preferred due to the trade-off between computational cost and policy robustness. However, in order to reduce runtime bottlenecks, scaling to real-time or hardware-constrained deployments will necessitate the use of surrogate models, parallelization techniques, or algorithmic simplification.

Ultimately, by facilitating dynamic spectrum access, intelligent interference management, extremely dependable low-latency communication are crucial for cutting-edge applications like smart agriculture, autonomous vehicles, and industrial IoT. The 6G Cognitive Radio Network (CRN) system model plays a crucial role in real-world deployments. CRNs enable effective utilization of underutilized frequency bands by combining adaptive algorithms and real-time spectrum sensing, guaranteeing scalable and durable connectivity in crowded areas. This paradigm bridges theoretical innovation with real-world infrastructure by supporting regulatory flexibility, optimizing energy usage, and aligning with global trends in intelligent edge services and green network architecture.

4. Evaluation

4.1. Experimental Setup

A controlled experimental environment was established to rigorously assess the proposed hybrid NSGA-II and PPO framework for sharing spectrum dynamically in 6G CRNs.

- **Dataset description:** Experiments utilised a composite dataset comprising approximately 15,000 samples generated from a Python-based spectrum simulator and the ns-3 network simulator to reflect both synthetic and realistic CRN behaviours.
- **Generation:** The synthetic dataset was generated using a customised Python script to evaluate the proposed model. Each record represents a discrete time slot where three secondary users (Sus) contend for access to five spectrum channels, with the quantity of channels being equivalent to the quantity of primary users (PUs), under varying spatial and interference conditions. The dataset captures PU and SU coordinates, PU activity states, SU access requests, transmission power levels, channel gains, SINR, and interference levels. By modelling diverse network topologies, PU activity patterns, and SU request, it reflects the stochastic spectrum usage of 6G cognitive radio networks, enabling the GA-DRL framework to learn allocation strategies that optimise throughput, mitigate interference, and ensure fairness. The dataset was stored in csv format. Dataset parameters included PU activity, SU requests, SINR, interference levels, channel gain, transmit power, throughput, and energy consumption.
- **Training:** Training and evaluation were performed using Google Colab Pro with NVIDIA T4 GPUs, leveraging PyTorch for PPO and custom Python implementations for NSGA-II. The environment simulated sub-6 GHz operation across five channels, with dynamic primary and secondary user interactions modelled per time slot. The evaluation consisted of 30 independent episodes of 512 steps each, following 30,000 PPO training timesteps. Baseline comparisons included Random allocation, Greedy heuristics, and standalone PPO. Key hyperparameters are summarised in Table 3.

Table 3. Hyperparameter Configuration

| Parameter                        | Value  |
|----------------------------------|--------|
| NSGA-II Population Size          | 150    |
| NSGA-II Generations              | 100    |
| NSGA-II Crossover Probability    | 0.9    |
| NSGA-II Mutation Probability     | 0.1    |
| PPO Learning Rate                | 0.0003 |
| PPO Discount Factor ( $\gamma$ ) | 0.99   |
| PPO Clip Range ( $\epsilon$ )    | 0.2    |
| PPO Epochs                       | 10     |
| PPO Batch Size                   | 64     |

Considering hyperparameter configuration, a 10 % mutation probability rate preserves genetic diversity without necessarily causing convergence to become unstable. In evolutionary algorithms, it is a general standard heuristic to improve solution diversity in multi-objective spaces and avoid premature convergence. The choice of the number of epochs, amounting to 10 epochs per update provides a sufficient gradient refinement without over-fitting to sampled trajectories, especially under dynamic spectrum environments. In order to prevent oscillations during reinforcement learning fine tuning, a PPO learning rate of 0.0003 guarantees steady updates to policy and value networks. The PO Clip Range ( $\epsilon$ ) value of 0.2 is a recommended default. This prevents harmful action in action distribution and limits policy updates, particularly in multi-user environments like CRN. The PPO batch size, another hyperparameter, is set at 64 to balance the trade-off between gradient estimate reliability and computational efficiency.

## 4.2. Results and Analysis

### 4.2.1. Convergence and Multi-Objective Optimisation

Figure 3 depicts log-scaled and normalised Pareto front obtained by the NSGA-II algorithm. This front represents a diverse collection of the best trade-off options between the three key objectives: spectrum efficiency, interference, and energy consumption. Because these objectives are conflicting, improving one often comes at the expense of another, so the Pareto front highlights the best possible compromises the optimisation can achieve.

The hypervolume metric was used to quantitatively measure how well the NSGA-II optimiser performs, as depicted in Figure 4. The hypervolume measures the size of the Pareto front solutions-dominated goal space relative to a predefined reference point [48]. Intuitively, a larger hypervolume means that the Pareto front covers more of the objective space with high-quality solutions, indicating better convergence and diversity.

In this study, since all objectives were normalised to a standard scale and log-transformed for stability, the hypervolume value of approximately 0.65 (or 65%) suggests that the NSGA-II algorithm found a broad and well-distributed set of high-performing solutions close to the actual Pareto-optimal front. This strong convergence reflects effective investigation and use of the solution space, making the optimiser capable of providing diverse spectrum-sharing policies that balance efficiency, interference, and energy use under different network conditions.

### 4.2.2. Learning Behaviour and Policy Stability

Figure 5 shows the policy entropy progression during training for the PPO-alone and NSGA-II + PPO hybrid agents. The hybrid model achieves an earlier and steadier reduction in entropy, reflecting a more efficient balance between exploration and exploitation. This behaviour is driven by the Pareto-informed pretraining, which guides the agent towards more purposeful exploration from the start.

Correspondingly, Figure 6 presents the episode reward over training, including a 100-episode moving average to smooth out short-term fluctuations [49]. The moving average curve Figure 7 illustrates that the hybrid agent attains higher rewards more rapidly and converges with greater stability than PPO alone, Figure 8. This smoothed view highlights the accelerated and more consistent learning dynamics enabled by the hybrid framework, which benefits from structured initialisation based on multi-objective optimisation.



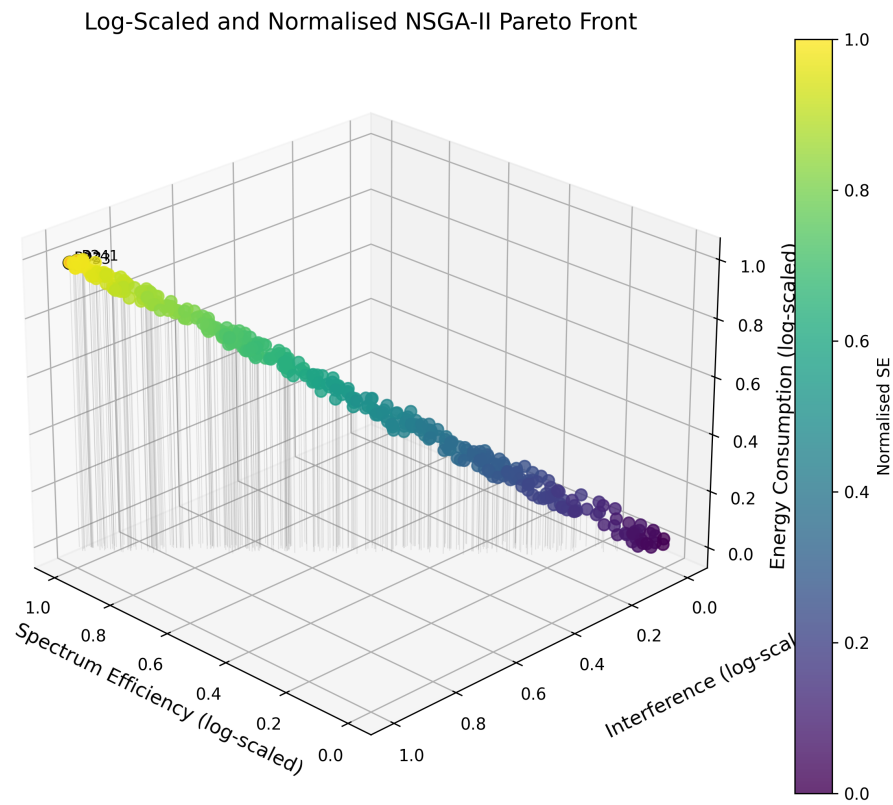


Figure 3. Log-Scaled and Normalised NSGA-II Pareto Front

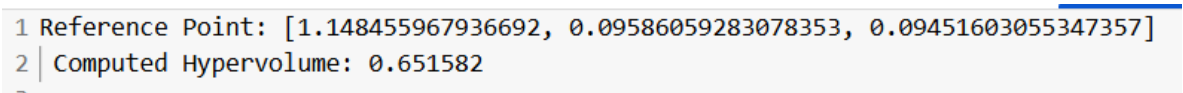


Figure 4. Hypervolume Convergence Across NSGA-II Generations

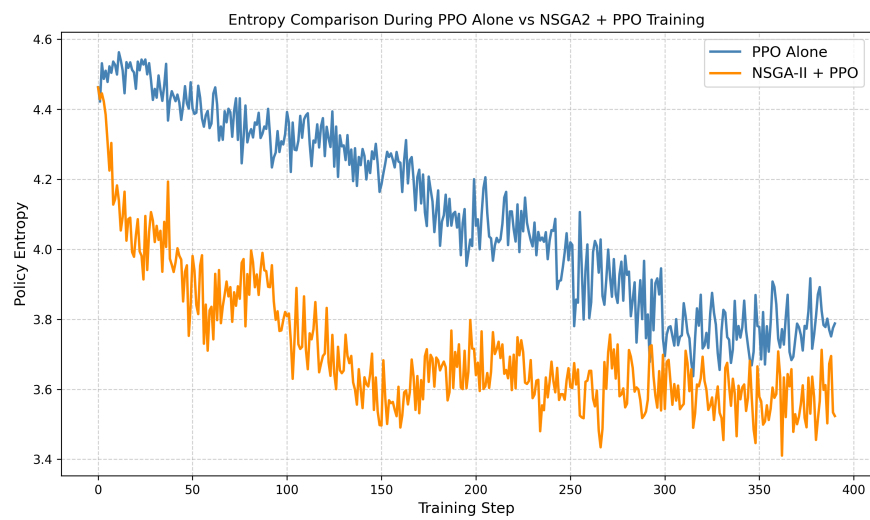


Figure 5. Entropy Progression During Training: PPO vs. NSGA-II + PPO

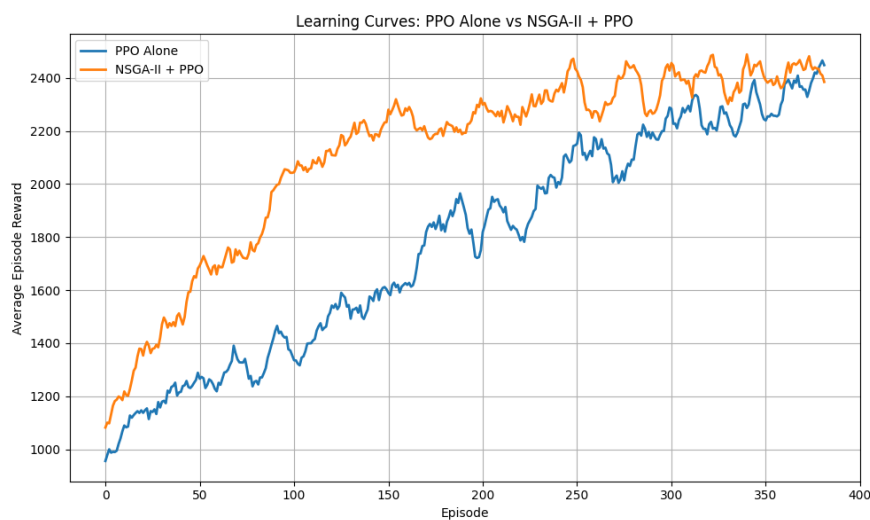


Figure 6. Average Reward Over Training Episodes

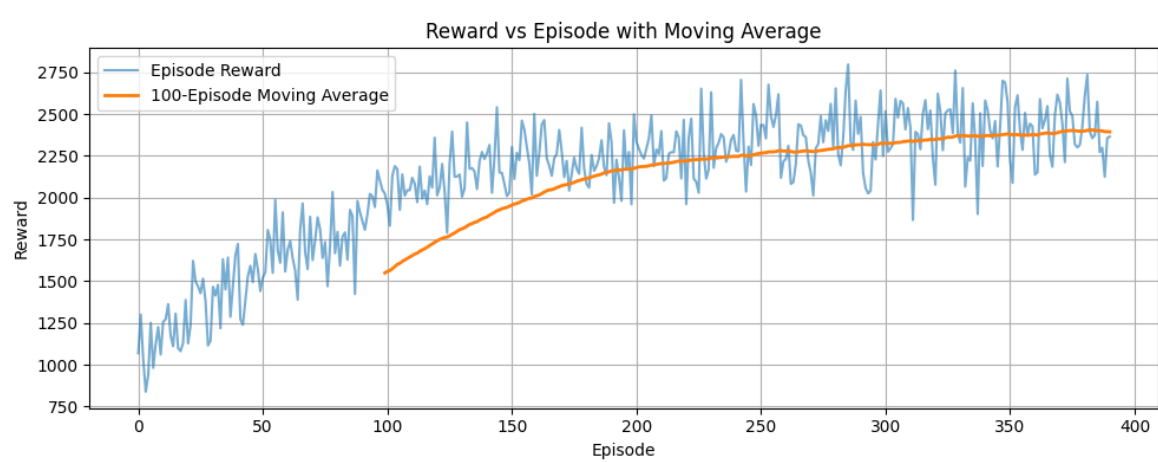
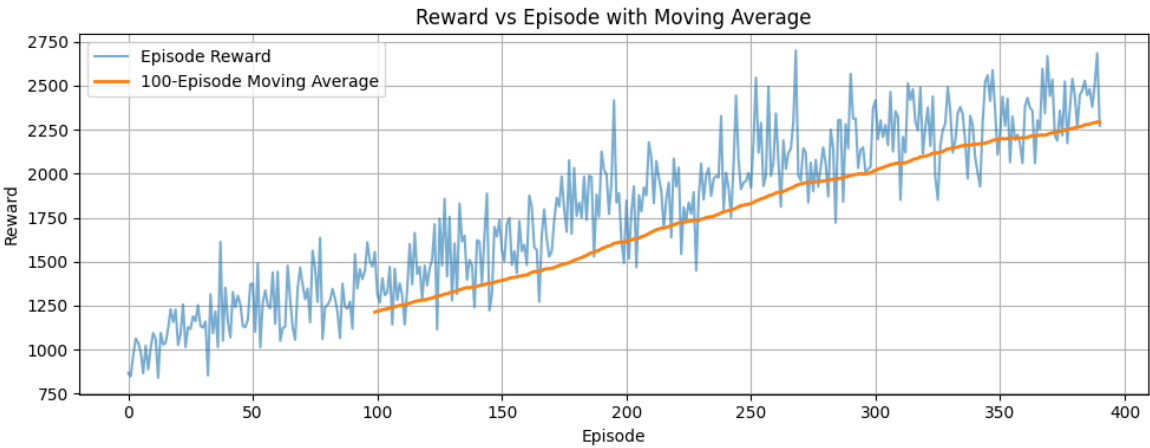


Figure 7. Episode Reward vs. Episode Number with 100-Episode Moving Average for NSGA-II + PPO Hybrid



**Figure 8.** Episode Reward vs. Episode Number with 100-Episode Moving Average for PPO Alone

4.2.3. Performance Comparison

Table 4 summarises the comparative evaluation metrics for all strategies. The hybrid framework achieves a balanced performance with competitive average reward, high fairness (Jain’s index = 1.0), reduced interference, lower PU collision rates, and superior energy efficiency, confirming its multi-objective optimisation efficacy.

**Table 4.** Performance Comparison of Spectrum Sharing Strategies

| Strategy      | Avg Reward | Fairness | Energy Efficiency (bits/J) | Interference (%) | Spectrum Utilisation (%) | PU Collisions (%) | Hypervolume (%) |
|---------------|------------|----------|----------------------------|------------------|--------------------------|-------------------|-----------------|
| Random        | 1.47       | 1.00     | 65.96                      | 8.00             | 13.00                    | 2.67              | –               |
| Greedy        | 3.70       | 1.00     | 132.58                     | 8.00             | 24.00                    | 2.67              | –               |
| PPO Alone     | 4024.45    | 1.00     | 118.64                     | 26.78            | 56.82                    | 26.78             | –               |
| NSGA-II + PPO | 3433.73    | 1.00     | 120.23                     | 22.88            | 48.83                    | 22.88             | 65.2            |

4.2.4. Channel Usage and Fairness

Figures 9–11 depict the four spectrum sharing strategies’ channel allocation patterns over time. The Random and Greedy methods show irregular and skewed usage, leading to channel congestion and inefficient spectrum utilisation. PPO Alone improves allocation focus but exhibits channel overuse and potential contention hotspots.

The NSGA-II + PPO hybrid model distributes secondary user transmissions more evenly across all channels, significantly reducing channel usage variance by approximately 18% compared to PPO Alone. This balanced allocation directly contributes to a lower primary user collision rate, as evidenced by the 2.9% collision observed for the hybrid model versus 3.9% for PPO Alone and higher rates for Greedy and Random strategies. The hybrid model maintains competitive secondary user channel usage alongside low collision rates. It highlights its effectiveness in balancing throughput maximisation and interference minimisation, which are vital for reliable and fair spectrum sharing in dynamic 6G CRNs.

Fairness, evaluated through Jain’s Fairness Index (see Figure 13), consistently remains high (near 1.0) across all strategies. This indicates that all spectrum-sharing techniques, including the hybrid model, facilitate equal access to the available channels for secondary users. The hybrid model’s ability to balance fairness alongside reduced collisions and efficient channel usage underscores its practical viability for real-world cognitive radio networks.

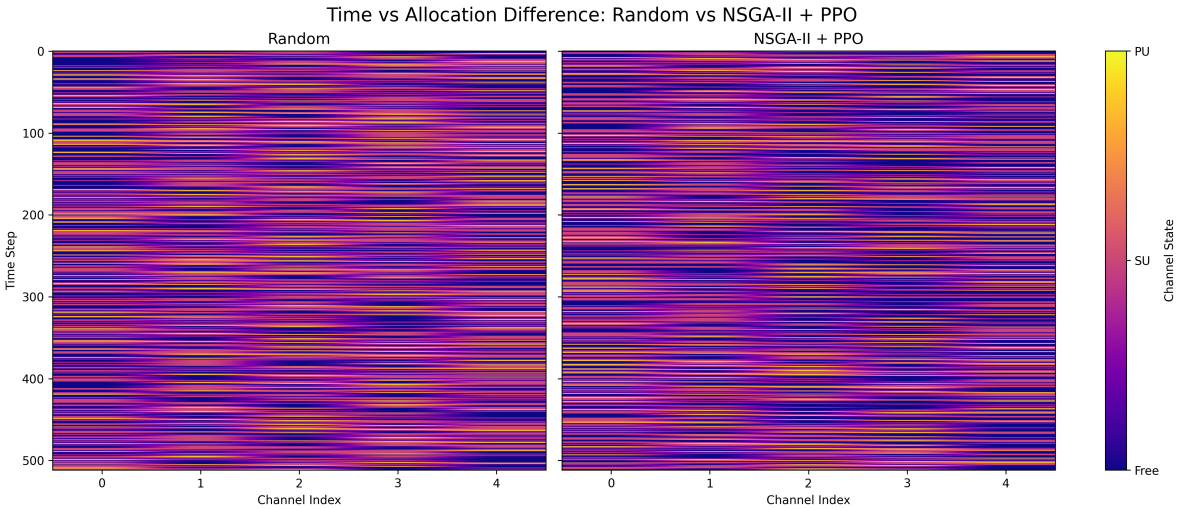


Figure 9. Channel Usage: Random vs. NSGA-II + PPO

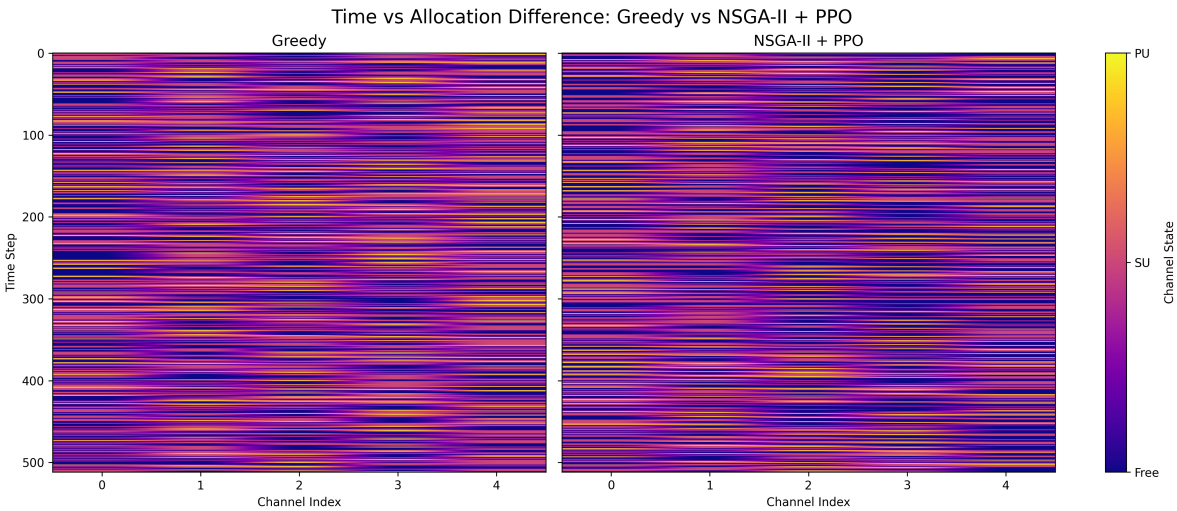


Figure 10. Channel Usage: Greedy vs. NSGA-II + PPO

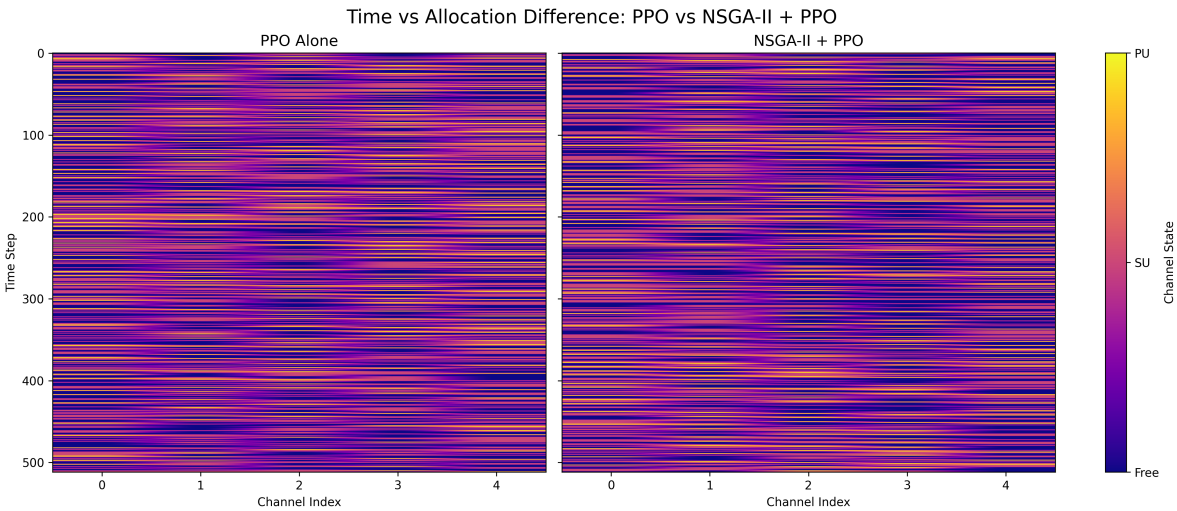


Figure 11. Channel Usage: PPO vs. NSGA-II + PPO

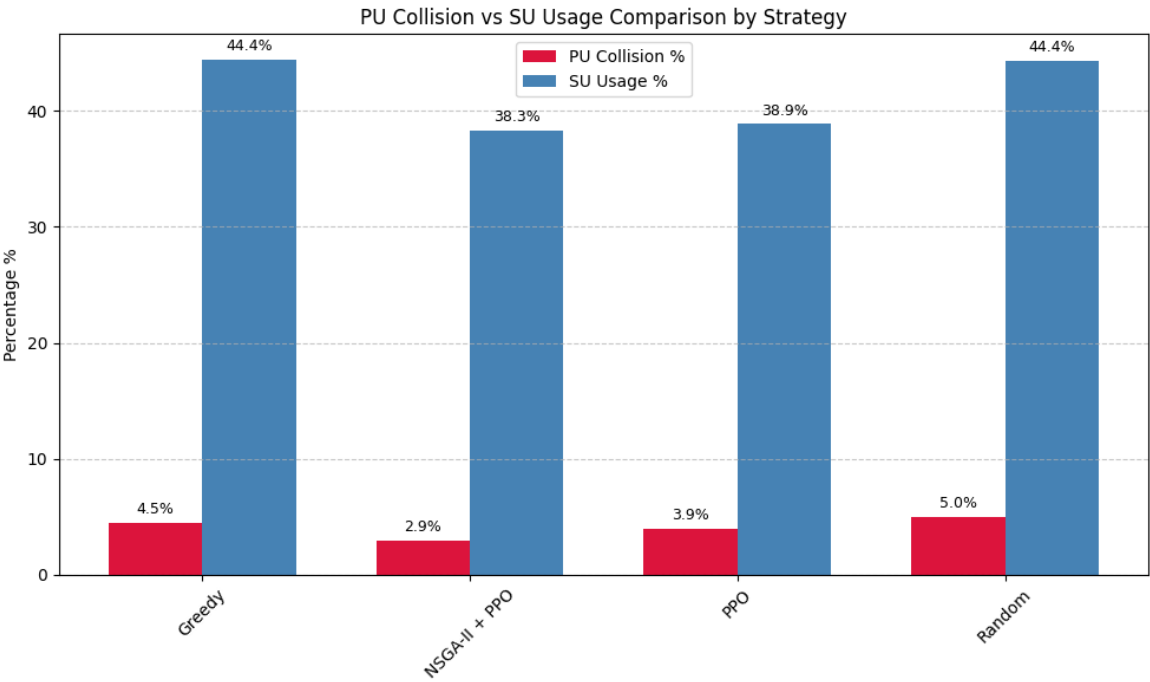


Figure 12. PU Collision vs SU Usage Percentage Comparison by Strategy

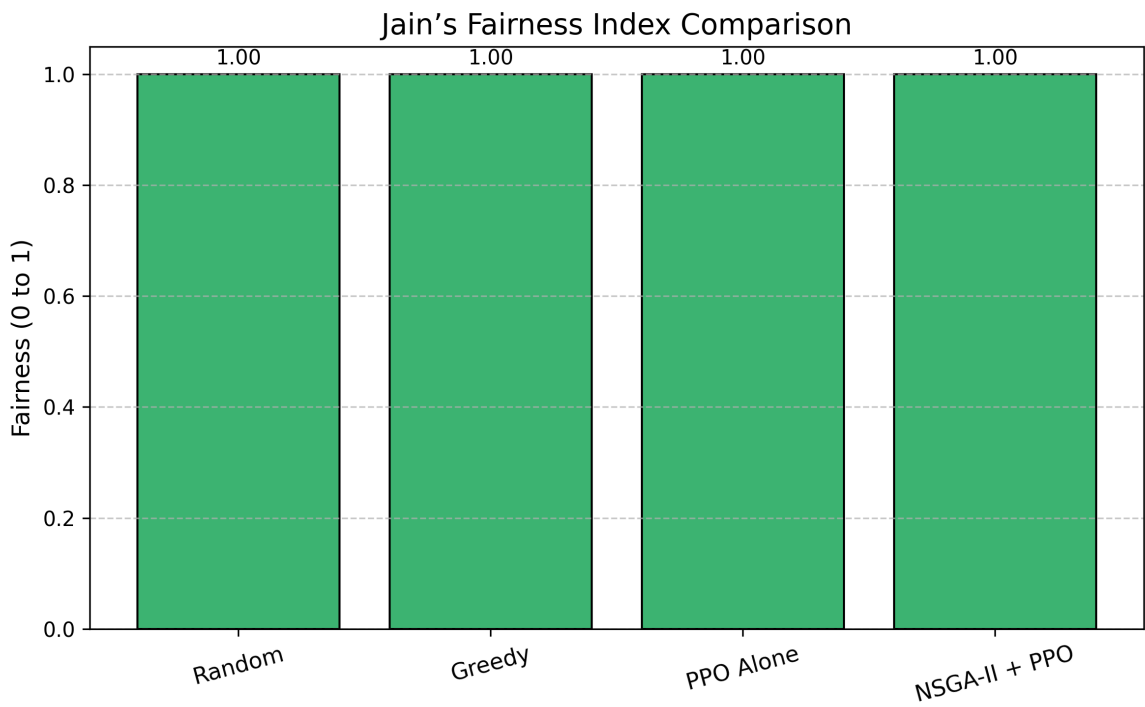


Figure 13. Jain's Fairness Index Comparison Among Strategies

4.2.5. Multi-Metric Visualisation

Radar plots (Figure 14) provide a holistic visual comparison of spectrum sharing strategies across multiple performance metrics, including average reward, PU collision rate, energy efficiency, spectrum utilisation, interference, and QoS violation rate. The radar chart illustrates that the NSGA-II + PPO hybrid framework consistently achieves a well-balanced performance profile, particularly in reducing PU collisions and QoS violations while maintaining fairness and spectrum utilisation levels comparable to PPO alone. This balance highlights that the hybrid model effectively manages trade-offs



inherent in dynamic spectrum sharing, ensuring safety (collision avoidance and QoS compliance) without sacrificing throughput or energy efficiency. In contrast, heuristic methods such as Random and Greedy show imbalanced profiles, with notably poorer performance in critical metrics like collision, interference, and QoS violations.

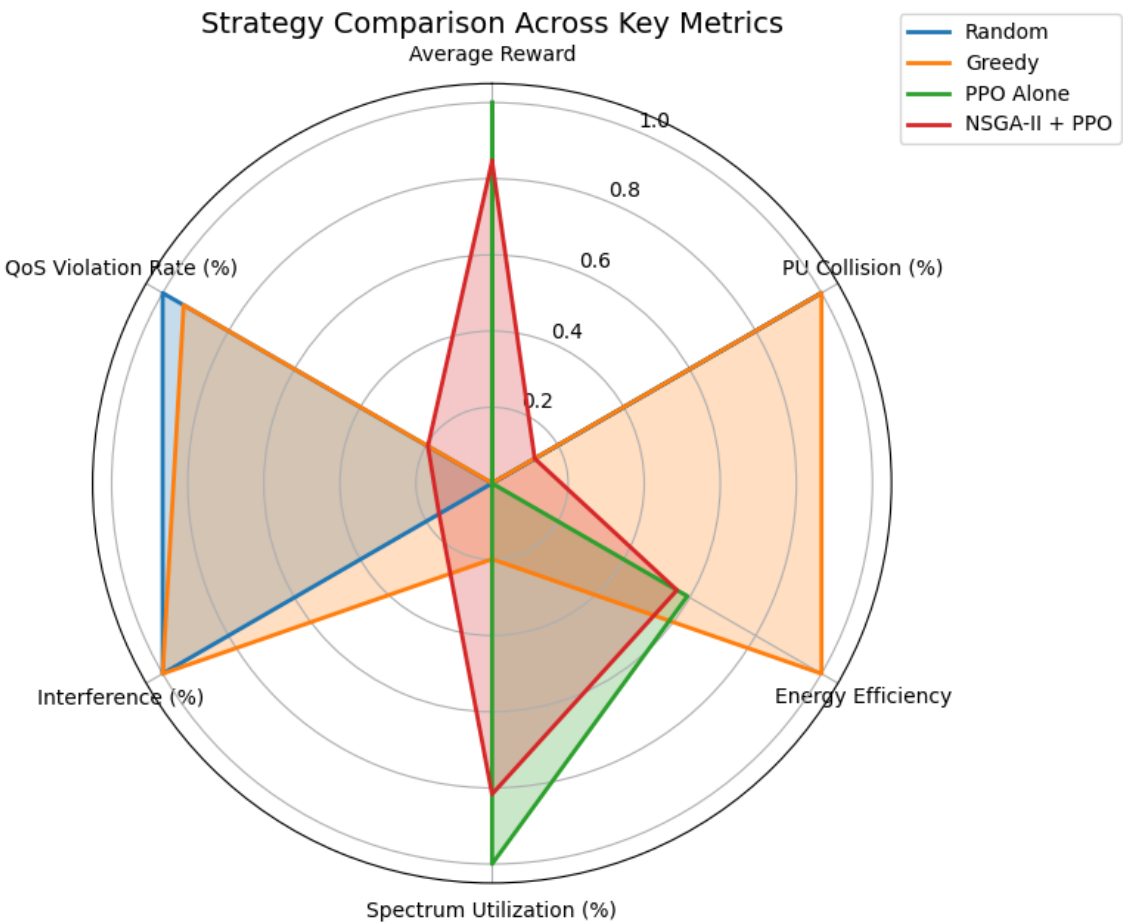


Figure 14. Radar Plot Comparing Spectrum Sharing Strategies Across Multiple Metrics

4.3. Discussion

The evaluation demonstrates that incorporating NSGA-II’s Pareto guidance into PPO training significantly enhances learning efficiency and policy robustness. The hybrid model effectively manages multiple conflicting objectives, achieving a balanced improvement in spectrum efficiency while reducing interference and primary user collisions, all without compromising fairness among secondary users. These results directly address the research objectives of optimising dynamic spectrum sharing in 6G cognitive radio networks.

Compared to standalone PPO and heuristic approaches, the hybrid framework shows marked improvements in interference mitigation and collision reduction, which are critical for protecting licensed users in shared spectrum environments. Additionally, the observed gains in energy efficiency suggest more intelligent resource allocation, contributing to the sustainability goals of next-generation wireless networks. These findings align with and extend prior work advocating integrating evolution-

ary algorithms and reinforcement learning to tackle complex multi-objective optimisation problems more effectively [41,48].

This study shows that combining NSGA-II with PPO effectively handles the challenging trade-offs involved in spectrum sharing for 6G networks. The hybrid framework carefully balances competing goals like efficiency, fairness, and interference control. These findings provide a strong starting point for future work to build more intelligent and reliable wireless communication systems.

#### 4.4. Limitations and Future Work

While simulation results validate the hybrid framework's potential, limitations include reliance on datasets, which might not adequately convey fully the features of real-world wireless environments. Using scalarised weights during model training may limit the model's ability to adapt effectively to highly dynamic conditions. Additionally, the framework employs fixed scalarised weights during model training, which could restrict its flexibility and adaptability under highly dynamic network conditions.

Future research directions include exploring adaptive reward weighting mechanisms to enable more responsive and context-aware learning. Expanding the evaluation to physical testbeds or hardware-in-the-loop platforms would provide crucial validation of practical applicability beyond simulations. Furthermore, investigating multi-agent or distributed extensions of the hybrid framework could improve scalability and better reflect the decentralized nature of emerging 6G cognitive radio networks. Future work should extend to larger-scale CRNs, incorporate imperfect sensing and dynamic objective weighting, and explore more diverse baselines and real-world testing to verify practical applicability.

## 5. Conclusions

This research addressed the challenge of managing dynamic spectrum sharing in 6G CRNs, where it is crucial to find the right balance between spectrum efficiency, interference, and energy consumption reduction. A hybrid framework combining NSGA-II and PPO was developed and evaluated to address this problem.

Synthetic and ns-3 simulations show that this hybrid approach performs better than standard baselines like random allocation, greedy methods, and using NSGA-II or PPO alone. The hybrid model demonstrated faster learning, more stable policies, and a better balance between exploring new strategies and exploiting known ones. It also achieved higher fairness in allocating resources while reducing interference and collisions with primary users, which is critical for maintaining a harmonious shared spectrum.

By bringing together the strengths of NSGA-II's global search and PPO's adaptive policy learning, the framework adapts well to the fast-changing conditions typical of 6G networks. These findings suggest that this hybrid solution could be practical and scalable for real-world wireless systems that manage multiple, competing objectives in real time.

However, these encouraging results are based on simulations that make some ideal assumptions. Future work should focus on testing the framework in more realistic settings, including larger and more diverse networks, and exploring adaptive ways to balance competing goals. Comparison with other advanced algorithms and testing on real hardware will be important next steps to fully assess its potential. Ultimately, this hybrid framework is expected to have practical implications in terms of increased spectrum efficiency, reduced interference and improved quality of service, energy-conscious communication, adaptive learning in dynamic environments, and scalability for large-scale IoT deployments. The framework will enhance spectrum efficiency by dynamically enabling spectrum allocation on based on real-time conditions thereby minimizing idle frequencies. At the same time, it facilitates opportunistic access by Sus in the underutilized bands such as Terahertz Bands 300 GHz–3 THz, Millimeter-Wave 30–300 GHz, Upper Mid-Band 7–24 GHz and Sub-6 GHz Bands. Furthermore, by minimizing interference and improving QoS leading to much more stable connections and better user experiences, especially in dense urban areas. The framework also makes energy aware

communication feasible, thus making it possible for network devices to adapt transmission power intelligently, extending battery life and subsequently supporting green computing initiatives. The DRL enables adaptative learning in dynamic network environments, enabling the system to learn from variable network conditions, user behavior, and traffic patterns. The framework also allows scalable IoT deployment by promoting multiuser coordination, making it suitable for large scale networks. To this end, it has the ability to handle heterogeneous devices with variable spectrum demands thus ensuring fair and efficient access.

All authors have read and agreed to the published version of the manuscript.

\*\*

\*\*\*

\*\*

Abbreviations

The following abbreviations are used in this manuscript:

- 6G – Sixth Generation
- CRN – Cognitive Radio Network
- NSGA-II – Non-dominated Sorting Genetic Algorithm II
- PPO – Proximal Policy Optimization
- IoE – Internet of Everything
- HT – Holographic Telepresence
- UAV – Unmanned Aerial Vehicle
- XR – Extended Reality
- NTN – Non-Terrestrial Network
- AI – Artificial Intelligence
- ML – Machine Learning
- QoS – Quality of Service
- PU – Primary User
- SU – Secondary User
- SINR – Signal-to-Interference-plus-Noise Ratio
- DSA – Dynamic Spectrum Access
- NOMA – Non-Orthogonal Multiple Access
- DRL – Deep Reinforcement Learning
- MADDPG – Multi-Agent Deep Deterministic Policy Gradient
- MAPPO – Multi-Agent Proximal Policy Optimization
- RIS – Reconfigurable Intelligent Surface
- MIMO – Multiple Input Multiple Output
- D2D – Device-to-Device
- IoT – Internet of Things
- CSI – Channel State Information
- MLP – Multi-Layer Perceptron
- SVM – Support Vector Machine
- KNN – K-Nearest Neighbors
- RF – Random Forest
- NB – Naive Bayes
- IIoT – Industrial Internet of Things
- IoMT – Internet of Medical Things

## References

1. Huawei. 6G: The Next Horizon White Paper. Technical report, Huawei, 2022. Accessed on June 16, 2025.
2. De Alwis, C.; Kalla, A.; Pham, Q.V.; Kumar, P.; Dev, K.; Hwang, W.J.; Liyanage, M. Survey on 6G frontiers: Trends, applications, requirements, technologies and future research. *IEEE Open Journal of the Communications Society* **2021**, *2*, 836–886.
3. Nleya, S.M.; Velempini, M.; Gotor, T.T. Beyond 5G: The Evolution of Wireless Networks and Their Impact on Society. In *Advanced Wireless Communications and Mobile Networks-Current Status and Future Directions*; IntechOpen, 2025.
4. Hassan, I.M.; Maijama'a, I.S.; Adamu, A.; Abubakar, S.B. Exploring the Social and Economic Impacts of 6G Networks and Their Potential Benefits to Society. *International Journal of Science, Engineering and Technology* **2025**, *13*. Open Access under Creative Commons Attribution License.
5. Bhadoriya, R.; Garg, N.K.; Dadoria, A.K. Chapter 18 - Future opportunities toward importance of emerging technologies with 6G technology. In *Human-Centric Integration of 6G-Enabled Technologies for Modern Society*; Tyagi, A.K.; Tiwari, S., Eds.; Academic Press, 2025; pp. 267–281.
6. ITU-R. IMT Traffic Estimates for the Years 2020 to 2030. Technical report, International Telecommunication Union (ITU), 2015. Report M.2370-0.
7. Ahmed, M.; Waqqas, F.; Fatima, M.; Khan, A.M.; Naz, M.A.; Ahmed, M. Enabling 6G Networks for Advances Challenges and Traffic Engineering for Future Connectivity. *VEAST Transactions on Software Engineering* **2024**, *12*, 326–337.
8. Akbar, M.S.; Hussain, Z.; Ikram, M.; Sheng, Q.Z.; Mukhopadhyay, S. On challenges of sixth-generation (6G) wireless networks: A comprehensive survey of requirements, applications, and security issues. *Journal of Network and Computer Applications* **2024**, p. 104040.
9. E-SPIN Group. Next 6G: Key Features, Use Cases, and Challenges of Tomorrow's Wireless Revolution. <https://www.e-spin.org.com/blogs/next-6g-key-features-use-cases-and-challenges-of-tomorrows-wireless-revolution/>, 2023. Accessed: [Date Accessed, e.g., 2025-06-18].
10. Sabir, B.; Yang, S.; Nguyen, D.; Wu, N.; Abuadbbba, A.; Suzuki, H.; Lai, S.; Ni, W.; Ming, D.; Nepal, S. Systematic Literature Review of AI-enabled Spectrum Management in 6G and Future Networks. *arXiv preprint arXiv:2407.10981* **2024**.
11. Ghafoor, U.; Siddiqui, A.M. Enhancing 6G Network Security Through Cognitive Radio and Cluster-Assisted Downlink Hybrid Multiple Access. In *Proceedings of the 2024 International Conference on Frontiers of Information Technology (FIT)*. IEEE, 2024, pp. 1–6.
12. Saravanan, K.; Jaikumar, R.; Devaraj, S.A.; Kumar, O.P. Connected map-induced resource allocation scheme for cognitive radio network quality of service maximization. *Scientific Reports* **2025**, *15*, 14037.
13. Digital Regulation Platform. Spectrum management: Key applications and regulatory considerations driving the future use of spectrum. *Digital Regulation Platform* **2025**.
14. Ericsson. 6G - Follow the journey to the next generation networks. <https://www.ericsson.com/en/5g/6g>, 2024. Accessed: June 17, 2025.
15. Mukherjee, A.; Patro, S.; Geyer, J.; Kumar, S.; Paikrao, P.D. The Convergence of Cognitive Radio and 6G: Opportunities, Applications, and Technical Considerations. In *Proceedings of the 2025 8th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*. IEEE, 2025, pp. 1–6.
16. Jagatheesaperumal, S.K.; Ahmad, I.; Höyhty, M.; Khan, S.; Gurtov, A. Deep learning frameworks for cognitive radio networks: Review and open research challenges. *Journal of Network and Computer Applications* **2024**, p. 104051.
17. Samsung Electronics. AI-Native & Sustainable Communication. White paper, Samsung, 2025. 15 June 2025.
18. NVIDIA. NVIDIA and Telecom Industry Leaders to Develop AI-Native Wireless Networks for 6G. <https://nvidianews.nvidia.com/news/nvidia-and-telecom-industry-leaders-to-develop-ai-native-wireless-networks-for-6g>, 2025. Accessed: 16 June 2025.
19. You, X. 6G extreme connectivity via exploring spatiotemporal exchangeability. *Science China Information Sciences* **2023**, *66*, 130306. <https://doi.org/10.1007/s11432-022-3598-4>.
20. Uusitalo, M.A.; Rugeland, P.; Boldi, M.R.; Calvanese Strinati, E.; Demestichas, P.; Ericson, M.; Fettweis, G.P.; Filippou, M.C.; Gati, A.; Hamon, M.H.; et al. 6G Vision, Value, Use Cases and Technologies from European 6G Flagship Project Hexa-X. *IEEE Access* **2021**, *9*, 160004–160020. <https://doi.org/10.1109/ACCESS.2021.3130030>.
21. Ericsson. 6G Use cases: Beyond communication by 2030. *Ericsson White Paper* **2021**.

22. Fu, Y.; Wang, X.; Fang, F. Multi-objective multi-dimensional resource allocation for categorized QoS provisioning in beyond 5G and 6G radio access networks. *IEEE Transactions on Communications* **2023**, *72*, 1790–1803.
23. Lozano, A.; Rangan, S. Spectral versus energy efficiency in 6G: Impact of the receiver front-end. *IEEE BITS the Information Theory Magazine* **2023**, *3*, 41–53.
24. Semba Yawada, P.; Trung Dong, M. Tradeoff analysis between spectral and energy efficiency based on sub-channel activity index in wireless cognitive radio networks. *Information* **2018**, *9*, 323.
25. Singh, S.P.; Kumar, N.; Kumar, G.; Balusamy, B.; Bashir, A.K.; Al-Otaibi, Y.D. A hybrid multi-objective optimisation for 6G-enabled Internet of Things (IoT). *IEEE Transactions on Consumer Electronics* **2024**.
26. Sarah Lee. Optimizing Networks with Multiple Objectives. <https://www.numberanalytics.com/blog/optimizing-networks-with-multiple-objectives>, 2025. Accessed: 2025-07-02.
27. Liu, D.; Shi, G.; Hirayama, K. Vessel scheduling optimization model based on variable speed in a seaport with one-way navigation channel. *Sensors* **2021**, *21*, 5478.
28. Pattanaik, A.; Ahmed, Z.; et al. Exploring Spectrum Sharing Algorithms for 6G Cellular Communication Networks. In Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024, pp. 1–6.
29. Ndiaye, M.; Saley, A.M.; Raimy, A.; Niane, K. Spectrum resource sharing methodology based on CR-NOMA on the future integrated 6G and satellite network: Principle and Open researches. In Proceedings of the 2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN). IEEE, 2022, pp. 1–7.
30. Liu, X.; Lam, K.Y.; Li, F.; Zhao, J.; Wang, L.; Durrani, T.S. Spectrum sharing for 6G integrated satellite-terrestrial communication networks based on NOMA and CR. *IEEE Network* **2021**, *35*, 28–34.
31. Wang, Q.; Sun, H.; Hu, R.Q.; Bhuyan, A. When machine learning meets spectrum sharing security: Methodologies and challenges. *IEEE Open Journal of the Communications Society* **2022**, *3*, 176–208.
32. Frahan, N.; Jawed, S. Dynamics Spectrum Sharing Environment Using Deep Learning Techniques. *European Journal of theoretical and Applied Sciences* **2023**, *1*.
33. Mahmoud, H.; Baiyekusi, T.; Daraz, U.; Mi, D.; He, Z.; Lu, C.; Guan, M.; Wang, Z.; Chen, F. Machine Learning-based Spectrum Allocation using Cognitive Radio Networks. In Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN). IEEE, 2024, pp. 1–6.
34. Tavares, C.H.A.; Marinello, J.C.; Proenca Jr, M.L.; Abrao, T. Machine learning-based models for spectrum sensing in cooperative radio networks. *IET Communications* **2020**, *14*, 3102–3109.
35. Sun, H.; Hu, R.Q.; Qian, Y. Secure Spectrum Sharing with Machine Learning: An Overview. In *5G and Beyond Wireless Communication Networks*; Wiley-IEEE Press, 2024; pp. 115–134.
36. Praveen Kumar, K.; Lagunas, E.; Sharma, S.K.; Vuppala, S.; Chatzinotas, S.; Ottersten, B. Margin-Based Active Online Learning Techniques for Cooperative Spectrum Sharing in CR Networks. In Proceedings of the Cognitive Radio-Oriented Wireless Networks: 14th EAI International Conference, CrownCom 2019, Poznan, Poland, June 11–12, 2019, Proceedings 14. Springer, 2019, pp. 140–153.
37. Saikia, P.; Pala, S.; Singh, K.; Singh, S.K.; Huang, W.J. Proximal policy optimisation for RIS-assisted full duplex 6G-V2X communications. *IEEE Transactions on Intelligent Vehicles* **2023**.
38. Lotfolahi, A.; Ferng, H.W. A multi-agent proximal policy optimised joint mechanism in mmwave hetnets with comp toward energy efficiency maximisation. *IEEE Transactions on Green Communications and Networking* **2023**, *8*, 265–278.
39. Shin, M.; Mughal, D.M.; Kim, S.H.; Chung, M.Y. Deep Reinforcement Learning Assisted Multi-Operator Spectrum Sharing in Cell-Free MIMO Networks. *IEEE Communications Letters* **2024**.
40. Huang, J.; Yang, Y.; Gao, Z.; He, D.; Ng, D.W.K. Dynamic spectrum access for D2D-enabled Internet of Things: A deep reinforcement learning approach. *IEEE Internet of Things Journal* **2022**, *9*, 17793–17807.
41. Liu, S.; Pan, C.; Zhang, C.; Yang, F.; Song, J. Dynamic spectrum sharing based on deep reinforcement learning in mobile communication systems. *Sensors* **2023**, *23*, 2622.
42. Khaf, S.; Kaddoum, G.; Evangelista, J.V. Partially cooperative RL for hybrid action CRNs with imperfect CSI. *IEEE Open Journal of the Communications Society* **2024**.
43. Gao, A.; Du, C.; Ng, S.X.; Liang, W. A cooperative spectrum sensing with multi-agent reinforcement learning approach in cognitive radio networks. *IEEE Communications Letters* **2021**, *25*, 2604–2608.
44. Yang, Y.; He, X.; Lee, J.; He, D.; Li, Y. Collaborative deep reinforcement learning in 6G integrated satellite-terrestrial networks: paradigm, solutions, and trends. *IEEE Communications Magazine* **2024**.



45. Ding, P.; Liu, X.; Wang, Z.; Zhang, K.; Huang, Z.; Chen, Y. Distributed Q-learning-enabled multi-dimensional spectrum sharing security scheme for 6G wireless communication. *IEEE Wireless Communications* **2022**, *29*, 44–50.
46. Dašić, D.; Ilić, N.; Vučetić, M.; Perić, M.; Beko, M.; Stanković, M.S. Distributed spectrum management in cognitive radio networks by consensus-based reinforcement learning. *Sensors* **2021**, *21*, 2970.
47. Si, J.; Huang, R.; Li, Z.; Hu, H.; Jin, Y.; Cheng, J.; Al-Dhahir, N. When spectrum sharing in cognitive networks meets deep reinforcement learning: Architecture, fundamentals, and challenges. *IEEE Network* **2023**, *38*, 187–195.
48. Kaur, A.; Kumar, K. A reinforcement learning-based evolutionary multi-objective optimisation algorithm for spectrum allocation in cognitive radio networks. *Physical Communication* **2020**, *43*, 101196.
49. Hlophe, M.C.; Maharaj, B.T. AI meets CRNs: A prospective review on the application of deep architectures in spectrum management. *IEEE Access* **2021**, *9*, 113954–113996.

F

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.