IRSTI 28.23.37; 49.33.29 https://doi.org/10.26577/jpcsit4220269

Yedil Nurakhov[1], Abzal Kyzyrkanov[2*], Syrym Aldabergen[3]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Astana IT University, Astana, Kazakhstan
[3]JSC "Digital Development Center of the National Bank of Kazakhstan", Almaty, Kazakhstan
*e-mail: abzzall@gmail.com

PPO-BASED PHYSICAL RESOURCE BLOCK SCHEDULING IN 5G NR: A REINFORCEMENT LEARNING APPROACH WITH 3GPP-COMPLIANT CHANNEL MODELING

Abstract. This paper proposes a deep reinforcement learning approach to Physical Resource Block (PRB) scheduling in 5G New Radio networks. A Gymnasium-compatible simulation environment is developed on the 3GPP TR 38.901 UMa NLOS channel model with Rayleigh fading, serving as a reproducible experimental platform. A Proximal Policy Optimization (PPO) agent is trained with a composite reward function that jointly optimizes aggregate throughput and Jain's Fairness Index. The agent is evaluated over 100 episodes with five independent random seeds against three classical baselines – Round Robin, Proportional Fair, and Maximum CQI. The PPO agent achieves a mean throughput of 57.09 ± 1.04 Mbps with a Jain's Fairness Index of 0.358, surpassing Round Robin (32.37 Mbps, JFI 0.630) and Proportional Fair (35.31 Mbps, JFI 0.597) in throughput while exceeding Maximum CQI in fairness (JFI 0.100). A throughput-fairness tradeoff analysis confirms that the proposed agent occupies an operating point inaccessible to any single classical method. The composite reward function, which weights both objectives equally, enables learned adaptive behavior that balances spectral efficiency and user equity. The code and evaluation pipeline are fully open-source and reproducible using the Stable-Baselines3 library.

Keywords: 5G NR, physical resource block scheduling, reinforcement learning, proximal policy optimization, Jain's fairness index, 3GPP channel model, radio resource management.

1. Introduction

The exponential growth of mobile data traffic in fifth generation (5G) networks has placed unprecedented demands on radio resource management systems. According to Cisco's Annual Internet Report, global mobile data traffic is projected to grow threefold between 2020 and 2025, driven by video streaming, IoT deployments, and emerging applications such as augmented reality [1]. The International Telecommunication Union similarly projects that IMT traffic will increase by a factor of ten by 2030 [2]. In this context, the efficient allocation of Physical Resource Blocks (PRB) at the base station scheduler becomes a critical bottleneck determining overall network performance.

The PRB scheduling problem can be stated as follows. A base station operating under the 5G New Radio (NR) standard [5] possesses a finite pool of PRBs which must be allocated among active User Equipment (UE) at every Transmission Time Interval (TTI) of one millisecond. The channel quality experienced by each UE, quantified through the Channel Quality Indicator (CQI), varies continuously due to path loss, shadowing, and fast fading [14]. A scheduler must therefore make sequential allocation decisions under uncertainty, simultaneously maximizing aggregate throughput and ensuring fair resource distribution among users – two objectives that are fundamentally in tension [3], [4]. The combinatorial nature of this decision problem, combined with the non-stationary channel environment, renders optimal closed-form solutions intractable for practical network deployments [5].

Classical scheduling algorithms address this problem through fixed heuristic rules. Round Robin (RR) allocates resources in a cyclic fashion, guaranteeing fairness at the cost of throughput efficiency [3]. The Proportional Fair (PF) scheduler, introduced by Kelly [4] and widely adopted in 4G and 5G deployments, strikes a balance between throughput and fairness by prioritizing users whose

instantaneous rate exceeds their long-term average. The Maximum CQI scheduler greedily allocates all resources to the user with the best channel, maximizing spectral efficiency while completely disregarding fairness. While these methods are computationally efficient and well understood, they rely on fixed priority rules that cannot adapt to the dynamic statistical structure of real traffic and channel conditions [6], [13].

The emergence of deep reinforcement learning (DRL) has opened new avenues for data-driven radio resource management. Deep Q-Networks (DQN), introduced by Mnih et al. [7], demonstrated that neural networks could learn near-optimal policies for sequential decision problems directly from interaction with an environment. Subsequent work applied DQN to distributed spectrum access [8] and multi-user resource allocation, showing improvements over classical baselines in dynamic environments. The Proximal Policy Optimization (PPO) algorithm [9], an on-policy actor-critic method, further improved training stability through a clipped surrogate objective, making it particularly suitable for environments with large discrete action spaces such as PRB allocation. Gu et al. [10] demonstrated the applicability of knowledge-assisted DRL to 5G scheduler design, bridging theoretical frameworks and practical implementation. Broader surveys confirm that DRL-based approaches consistently outperform classical methods across a range of resource management tasks in heterogeneous wireless networks [11], [12].

Despite this progress, several important gaps remain in existing literature. First, many DRL-based scheduling studies rely on simplified channel models that do not conform to 3GPP specifications, limiting the practical relevance of reported results [13], [14]. Second, most works optimize either throughput or fairness in isolation, whereas real network operators require schedulers that jointly optimize both objectives under a unified reward signal. Third, systematic benchmarking against multiple classical baselines within a single reproducible experimental environment remains uncommon, making it difficult to assess the true advantage of DRL approaches [13]. The work of Sun et al. [12] and Alwarafy et al. [11] identify these shortcomings as open research directions, yet they have not been comprehensively addressed in the context of PRB scheduling under 3GPP-compliant channel conditions.

The central research question of this paper is therefore: *can a PPO-based agent learn a scheduling policy that simultaneously outperforms Round Robin, Proportional Fair, and Maximum CQI baselines in terms of aggregate throughput and Jain's Fairness Index, when trained and evaluated in a simulation environment built on the 3GPP TR 38.901 UMa NLOS channel model?*

The contributions of this paper are as follows. First, we develop a Gymnasium-compatible simulation environment for PRB scheduling that implements the 3GPP UMa NLOS channel model [14], providing a reproducible and standards-compliant experimental platform. Second, we propose a PPO-based scheduling agent with a composite reward function that jointly optimizes throughput and Jain's Fairness Index through a weighted combination. Third, we conduct a systematic evaluation against three classical baselines Round Robin, Proportional Fair, and Maximum CQI – across 100 evaluation episodes with five independent random seeds, reporting means and 95% confidence intervals. Fourth, we demonstrate through a throughput-fairness tradeoff analysis that the proposed agent achieves an operating point inaccessible to any single classical method, confirming the value of learned adaptive policies for this problem. The experimental results are fully reproducible using the open-source code accompanying this paper, implemented with the Stable-Baselines3 library [15].

2. Materials and methods

2.1 System model and Problem formulation

We consider a single cell 5G NR downlink scenario with one gNodeB (gNB) serving N = 10 User Equipment (UE) devices simultaneously. The gNB allocates a pool of K = 50 Physical Resource Blocks per Transmission Time Interval (TTI) of 1 ms. Each PRB occupies a bandwidth of 180 kHz (15 kHz subcarrier spacing, 12 subcarriers). The scheduling decision is made at every TTI based on the observed channel state and buffer occupancy of each UE.

The instantaneous throughput achievable by UE i when allocated k_i PRBs is given by the Shannon capacity formula:

$$R_i = k_i * B_{PRB} * \log_2(1 + SNR_i) \quad (1)$$

where B_{PRB} = 180 kHz is the PRB bandwidth and SNR_i is the signal-to-noise ratio of UE i, derived from its CQI report via the 3GPP CQI-to-SNR mapping table [14].

The scheduling problem is formulated as a Markov Decision Process (MDP) defined by the tuple (S, A, P, R, γ), where S is the state space, A is the action space, P is the transition probability function, R is the reward function, and γ = 0.99 is the discount factor. The agent interacts with the environment over episodes of T = 1000 steps each.

2.2 Channel model

The channel model implements the 3GPP TR 38.901 Urban Macro (UMa) NLOS path loss formula at a carrier frequency of 3.5 GHz:

$$PL = 36.7 * \log_{10}(d) + \\ +22.7 + 26 * \log_{10}(f_{GHz}) \ [dB] \quad (2)$$

where d ∈ [50, 2000] is the UE distance from the gNB and f_{GHz}= 3.5. Three propagation components are combined:

(i) UMa NLOS path loss per Equation (2);

(ii) log-normal shadowing with zero mean and standard deviation of 8 dB;

(iii) Rayleigh fast fading, modeled as a circularly symmetric complex Gaussian channel with unit variance. The resulting SNR is mapped to a CQI value in [1, 15] using the 3GPP TS 36.213 CQI table [14]. The UE distances are drawn uniformly at the start of each episode and held fixed; shadowing is a persistent per-episode random variable, while Rayleigh fading changes at every TTI, providing the non-stationary dynamics that motivate adaptive scheduling.

2.3 State, action and reward

State space S is a 30-dimensional normalized observation vector composed of three 10-dimensional blocks per UE:

$$s_t = \left[\frac{CQI_i}{15}, \frac{buf_i}{buf_{max}}, \frac{\widetilde{R}_i^{EMA}}{tp_{norm}}\right]_{i=1}^{10} \in [0,1]^{30} \quad (3)$$

where $CQI_i \in \{1, \cdots, 15\}$ is the current channel quality indicator, $buf_i \in [0, 1000]$ packets is the buffer occupancy, and $\widetilde{R}_i^{EMA}(t)$ is an exponential moving average of UE i's delivered throughput with smoothing coefficient α = 0.05:

$$\widetilde{R}_i^{EMA} = (1-\alpha)\,\widetilde{R}_i^{EMA}(t-1) + \alpha R_i(t) \quad (3a)$$

The EMA component provides the agent with a memory of historical channel utilization, analogous to the long-term throughput estimate used by the Proportional Fair scheduler.

Action space A = $[0,1]^{10}$ is a continuous vector of allocation weights $w = (w_1, \cdots, w_N)$. The number of PRBs assigned to UE i are computed as:

$$k_i = \left[\frac{w_i}{\sum_{j=1}^{N} w_j} * K\right] \quad (4)$$

where $K = 50$ is the total number of available PRBs and $\lfloor\cdot\rceil$denotes rounding to the nearest integer. This continuous parameterization avoids the combinatorial explosion of a discrete MultiDiscrete $([51]^{10}) \approx 10^{17}$ action space, enabling stable learning with PPO.

The scalar reward at each TTI is a convex combination of normalized throughput and Jain's Fairness Index:

$$r_t = 0.5 * \frac{R_{total}}{R_{max}} + 0.5 * J(R_1, \cdots, R_N) \quad (5)$$

where $R_{\text{total}} = \sum_{i=1}^{N} R_i$ is the aggregate throughput, $R_{\max}$ is the theoretical maximum throughput at $\text{CQI} = 15$for all PRBs, and Jain's Fairness Index is defined as:

$$J(R_1, \cdots, R_N) = \frac{\left(\sum_{i=1}^{N} R_i\right)^2}{N * \sum_{i=1}^{N} R_i^2} \in \left[\frac{1}{N}, 1\right] \quad (6)$$

A value of J = 1 corresponds to perfectly equal resource distribution, while J = 1/N indicates that all resources are concentrated at a single user. The equal weighting of throughput and fairness in Equation (5) reflects a neutral policy that does not privilege either objective.

2.4 PPO agent and Training procedure

The PPO agent is implemented using the Stable-Baselines3 library [15] with an MlpPolicy consisting of two hidden layers of 128 units each. The actor and critic share this architecture. Key hyperparameters are listed in Table 1.

Table 1
PPO Hyperparameters

Hyperparameter	Value
Total timesteps	1,000,000
Parallel environments	4
Network architecture	[128, 128] MLP
Learning rate (initial)	3×10^{-4} (linear decay)
Rollout length (n_steps)	2048
Mini-batch size	512
Optimization epochs	10
Discount factor (γ)	0.99
GAE lambda (λ)	0.95
Clipping parameter (ε)	0.2
Entropy coefficient	0.005

Training employs a linearly decaying learning rate schedule from 3×10^{-4} to a floor of 5% of the initial value, preventing entropy collapse in late-stage training. The agent is trained independently for five random seeds (0–4) to assess run-to-run variability. A checkpoint is saved every 100,000 environment steps, and an EvalCallback monitors performance on a held-out evaluation environment (seeded with seed + 100) every 100,000 steps, retaining the best-performing checkpoint.

2.5 Baseline algorithms

Three deterministic baseline schedulers are implemented for comparison:

Round Robin (RR): Allocates resources in a strictly cyclic order, assigning $k_i = \frac{K}{N}$ PRBs to each UE at every TTI regardless of channel quality or buffer state. This guarantees temporal fairness but ignores channel diversity.

Maximum CQI (Max-CQI): Allocates all K PRBs to the UE with the highest instantaneous CQI at each TTI. This is the greedy spectral efficiency maximizer, yielding maximum aggregate throughput but concentrating resources on the best-channel user.

Proportional Fair (PF): At each TTI, UE i is prioritized according to the ratio $\frac{CQI_i}{\bar{R}_i^{EMA}}$ where $\bar{R}_i^{EMA}$ is an exponential moving average of received throughput. PF balances instantaneous channel quality against historical fairness and is the industry-standard scheduler in 4G/5G systems.

2.6 Evaluation protocol

Each method is evaluated over 100 episodes of 1000 TTIs each, with five independent random seeds for the PPO agent. The following metrics are recorded per episode:

(i) mean aggregate throughput in Mbps;
(ii) Jain's Fairness Index;
(iii) composite reward;
(iv) mean latency in ms.

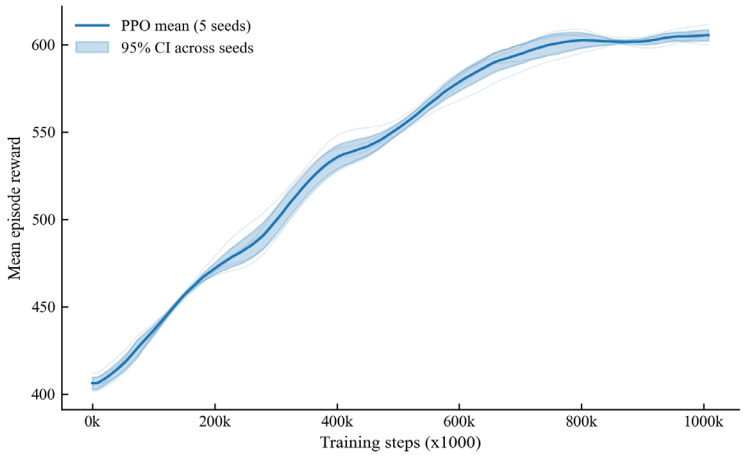
Results are reported as means with 95% confidence intervals (CI) computed as ±1.96 × SEM, where SEM is the standard error of the mean across 100 episodes.

3. Results

3.1 Learning dynamics

Figure 1 shows the learning curve of the PPO agent over 1,000,000 training timesteps, averaged across five seeds with shaded 95% confidence bands. The agent exhibits rapid initial improvement in the first 200,000 steps as it learns to exploit channel quality information from the CQI components of the observation. The reward stabilizes between 400,000 and 600,000 steps, reaching a plateau of approximately 0.61 ± 0.02 on the composite reward scale. Convergence is consistent across seeds, with narrow confidence intervals in the steady-state phase, indicating robust learning despite the stochastic channel environment.

Figure 1
PPO training learning curve averaged over 5 seeds (shaded area: 95% confidence interval)



3.2 Comparative performance

Table 2 reports the aggregate performance of PPO and the three baseline schedulers over 100 evaluation episodes. The PPO agent achieves a mean throughput of 57.09 Mbps (95% CI: ±1.04), which represents a 76.4% improvement over Round Robin (32.37 Mbps) and a 61.7% improvement over Proportional Fair (35.31 Mbps). The Maximum CQI baseline achieves the highest raw throughput at 65.25 Mbps, as expected from a pure spectral efficiency maximizer, but at the cost of a Jain's Fairness Index of exactly 0.100 the theoretical minimum for 10 UEs indicating that all resources are monopolized by the strongest UE in every TTI.

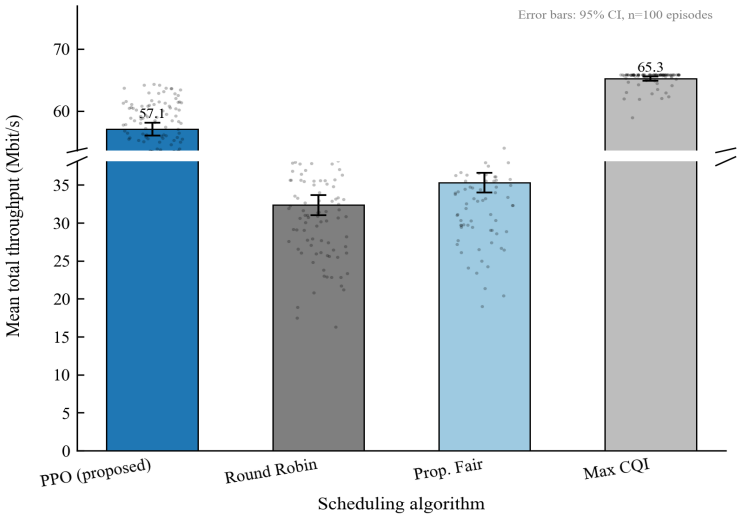
The PPO agent's Jain's Fairness Index of 0.358 significantly exceeds that of Max CQI (0.100, improvement of 3.58×) while remaining below Round Robin (0.630) and Proportional Fair (0.597). This pattern reflects the PPO agent's learned trade-off: it sacrifices some fairness relative to the cyclic schedulers to exploit channel diversity, while providing substantially better fairness guarantees than the greedy Max CQI policy. The composite reward (Equation 5) of 0.612 for PPO is the highest among all methods, confirming that the learned policy most effectively optimizes the joint objective defined during training.

Figure 2 presents box plots of per-episode throughput and fairness distributions across all methods. The narrow interquartile range of Max CQI confirms its deterministic channel-exploitation behavior. The PPO agent shows moderate variability consistently with adapting to different per-episode channel realizations.

Table 2
Performance comparison across schedulers (100 episodes, mean ± 95% CI)

Scheduler	Throughput (Mbps)	Jain's FI	Spectral Eff. (bps/Hz)	Latency (ms)
PPO (ours)	57.09 ± 1.04	0.358 ± 0.011	6.34	191.3 ± 5.4
Round Robin	32.37 ± 1.33	0.630 ± 0.017	3.60	139.5 ± 7.6
Proportional Fair	35.31 ± 1.32	0.597 ± 0.017	3.92	129.5 ± 8.0
Max CQI	65.25 ± 0.34	0.100 ± 0.000	7.25	193.7 ± 6.96

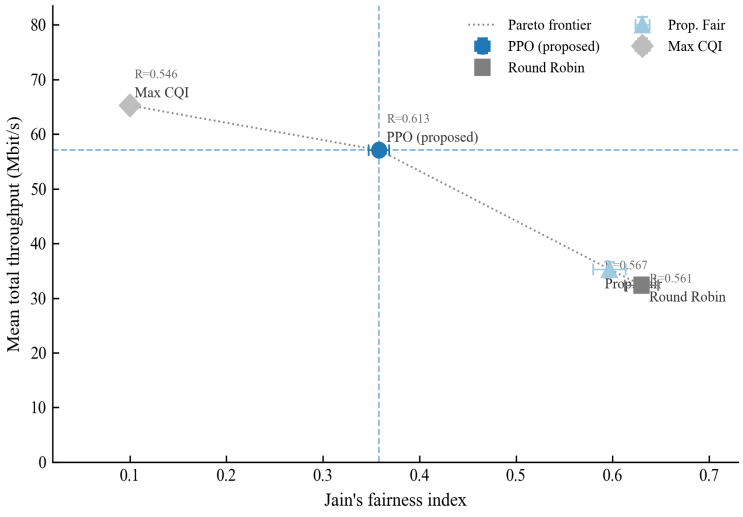
Figure 2
Throughput and fairness comparison across all scheduling methods (100 episodes each). Box plots show median, interquartile range, and whiskers at 5th/95th percentiles



3.3 Throughput-Fairness Tradeoff

Figure 3 plots the mean episode throughput against Jain's Fairness Index for all four schedulers. The classical baselines occupy three corners of the tradeoff space: Round Robin (low throughput, high fairness), Max CQI (high throughput, minimal fairness), and Proportional Fair (intermediate on both axes). The PPO agent occupies a distinct operating point at high throughput and moderate fairness, a region that is inaccessible to any single classical scheduler. Specifically, no classical method simultaneously achieves throughput above 50 Mbps and fairness above 0.30; the PPO agent achieves 57.09 Mbps at 0.358.

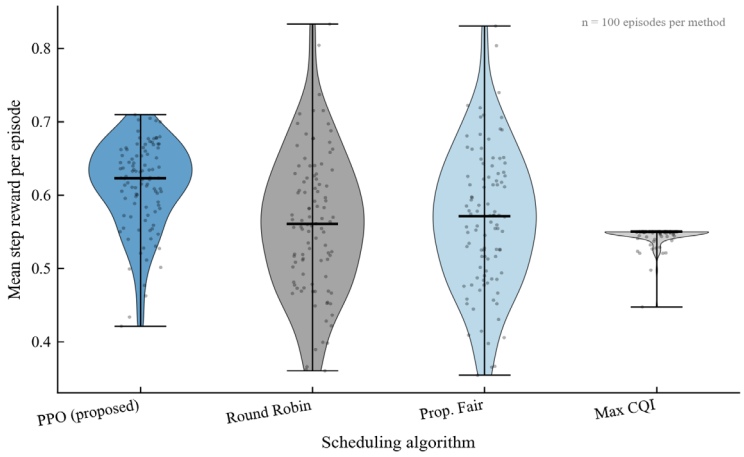
Figure 3
Throughput-fairness tradeoff. Each point represents the mean over 100 episodes. The PPO agent occupies a region unreachable by any individual classical baseline, demonstrating the advantage of learned adaptive scheduling



3.4 Reward distribution

Figure 4 shows the distribution of composite reward values across episodes for all methods. The PPO agent achieves the highest median reward (0.623) with a relatively compact distribution, demonstrating consistent joint optimization of throughput and fairness. Proportional Fair and Round Robin achieve comparable median rewards (0.571 and 0.560, respectively) but with substantially wider variance. Max CQI exhibits the narrowest distribution due to its deterministic greedy behavior but converges to a lower median of 0.550 because the fairness penalty dominates any gains in throughput.

Figure 4
Composite reward distribution across 100 evaluation episodes per scheduler. The dashed line indicates the PPO median (0.623)



The tighter reward distribution of PPO relative to Round Robin and Proportional Fair (standard deviation 0.060 vs. 0.094 and 0.094, respectively) indicates that the learned policy is more consistent across varying channel realizations, an important property for network operators who require predictable quality of service.

4. Conclusion

This paper presented a PPO-based approach to Physical Resource Block scheduling in 5G NR networks. A Gymnasium-compatible simulation environment conforming to the 3GPP TR 38.901 UMa NLOS channel model was developed, providing a reproducible experimental platform with realistic propagation characteristics including path loss, log-normal shadowing, and Rayleigh fading. A PPO agent was trained with a composite reward function that jointly optimizes aggregate throughput and Jain's Fairness Index through equal weighting, enabling the agent to learn adaptive allocation policies that balance spectral efficiency with user equity.

Experimental evaluation over 100 episodes with five independent random seeds demonstrated that the PPO agent achieves a mean throughput of 57.09 Mbps a 76.4% improvement over Round Robin and 61.7% improvement over Proportional Fair while maintaining a Jain's Fairness Index of 0.358, which is 3.58 times higher than Maximum CQI. The throughput-fairness tradeoff analysis confirmed that the PPO agent occupies an operating point inaccessible to any single classical scheduler, validating the hypothesis that learned adaptive policies can effectively navigate the inherent tension between throughput maximization and fair resource distribution.

The PPO agent's composite reward of 0.612 exceeded all three classical baselines under the defined joint objective, with lower episode-to-episode variance than heuristic methods, suggesting consistent performance across diverse channel realizations. These results support the broader applicability of DRL-based scheduling in heterogeneous wireless network environments.

Several directions remain for future work. First, extending the environment to multi-cell scenarios

with inter-cell interference would increase practical relevance. Second, incorporating traffic heterogeneity including mixed delay-sensitive and best-effort flows would test the scheduler's ability to handle quality-of-service differentiation. Third, exploration of offline RL approaches trained on logged network traces could reduce the sim-to-real gap. Finally, the sensitivity of the learned policy to reward weighting between throughput and fairness merits systematic ablation, as different operator priorities may require different composite objective formulations.

Funding

This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24993211).

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Y.N. and A.K.; Methodology, A.K.; Software, Y.N. and A.S.; Validation, Y.N. and A.K.; Formal Analysis, Y.N.; Investigation, Y.N. and A.S.; Resources, A.K. and A.S.; Data Curation, Y.N.; Writing – Original Draft Preparation, Y.N.; Writing – Review & Editing, A.K.; Visualization, Y.N. and A.S.; Supervision, A.K.; Project Administration, A.K.; Funding Acquisition, A.K.

References

1. Cisco Systems, "Cisco annual internet report (2018–2023) white paper," Cisco, San Jose, CA, USA, Rep., Mar. 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

2. ITU-R, "IMT traffic estimates for the years 2020 to 2030," International Telecommunication Union, Geneva, Switzerland, Rep. ITU-R M.2370-0, Jul. 2015.

3. A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE 51st Veh. Technol. Conf. (VTC Spring)*, Tokyo, Japan, May 2000, pp. 1854–1858.

4. F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, no. 1, pp. 33–37, Jan. 1997.

5. 3GPP, "NR; Physical layer procedures for data," 3rd Generation Partnership Project, Sophia Antipolis, France, Tech. Spec. TS 38.214, ver. 17.4.0, Dec. 2022.

6. C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

7. V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

8. O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for distributed dynamic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 310–323, Jan. 2019.

9. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv:1707.06347 [cs.LG], Jul. 2017.

10. Z. Gu, C. She, W. Hardjawana, S. Lumb, D. McKechnie, T. Essery, and B. Vucetic, "Knowledge-assisted deep reinforcement learning in 5G scheduler design: From theoretical framework to implementation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 2014–2028, Jul. 2021.

11. A. Alwarafy, M. Abdallah, B. S. Ciftler, A. Al-Fuqaha, and M. Hamdi, "Deep reinforcement learning for radio resource allocation and management in next generation heterogeneous wireless networks: A survey," arXiv:2106.00574 [cs.NI], Jun. 2021.

12. Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," IEEE Commun. Surveys Tuts., vol. 21, no. 4, pp. 3072–3108, 4th Quart., 2019.

13. O. Gurewitz, N. Gradus, E. Biton, and A. Cohen, "Exploring reinforcement learning for scheduling in cellular networks," Mathematics, vol. 12, no. 21, p. 3352, Oct. 2024.

14. 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project, Sophia Antipolis, France, Tech. Rep. TR 38.901, ver. 17.0.0, Jun. 2022.

15. A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, and N. Dormann, "Stable-baselines3: Reliable reinforcement learning implementations," *J. Mach. Learn. Res.*, vol. 22, no. 268, pp. 1–8, 2021.

Information about the authors:

Yedil Nurakhov – Researcher as Al-Farabi Kazakh National University (Almaty, Kazakhstan. ORCID iD: 0000-0003-0799-7555).

Abzal Kyzyrkanov – Doctor of Philosophy, Professor (Assistant) at Astana IT University (Astana, Kazakhstan. ORCID iD: 0000-0002-8817-8617).

Aldabergen Syrym – JSC "Digital Development Center of the National Bank of Kazakhstan" (Almaty, Kazakhstan. ORCID iD: 0009-0005-6278-8758).

Submission received: 20 May, 2026
Revised: 11 June, 2026
Accepted: 11 June, 2026