Metaheuristic Algorithms with Applications

www.maa.reapress.com

Metaheur. Algor. Appl. Vol. 2, No. 2 (2025) 172–177.

Paper Type: Original Article

Multi-Objective Optimization of Resource Allocation in 5G Networks Using a Reinforcement Learning-Based Genetic Algorithm

Behnam Ghahramankhani[1,*], Mohammad Hossein Khodabandeh[1]

[1]Department of Computer and Information Technology Engineering, Ab.C., Abhar Branch, Islamic Azad University, Abhar, Iran; ghahramankhani@iau.ac.ir, mohammadhossein.khodabandeh@iau.ir.

Citation:

Received: 12 January 2025 *Revised: 17 March 2025* *Accepted: 11 April 2025*	Ghahramankhani, B., & Khodabandeh, M. H. (2025). Multi-objective optimization of resource allocation in 5G networks using a reinforcement learning-based genetic algorithm. *Metaheuristic algorithms with applications, 2*(2), 172-177.

Abstract

This paper presents a hybrid Reinforcement Learning-Genetic Algorithm (RL-GA) for multi-objective resource allocation in 5G networks. Traditional optimization methods struggle with simultaneously optimizing throughput, latency, and energy consumption. The proposed method employs an RL agent that dynamically adjusts Genetic Algorithm (GA) parameters (crossover and mutation rates) based on population state, inspired by dopamine-mediated reward prediction mechanisms in cognitive science. Experimental results demonstrate statistically significant improvements over standard GA, adaptive GA, and PSO: 21.9% throughput increase, 35.9% latency reduction, and 16.7% energy efficiency improvement ($p<0.001$). The approach achieves faster convergence and superior Pareto front quality, addressing critical challenges in dynamic 5G environments.

Keywords: Multi-objective optimization, Genetic algorithm, Reinforcement learning, 5G networks, Resource allocation, 5G-Advanced.

1 | Introduction

The evolution towards 5G-Advanced and beyond has intensified the need for hyper-efficient resource allocation to support diverse Quality of Service (QoS) demands while optimizing conflicting objectives: maximizing throughput, minimizing latency, and reducing energy consumption. Traditional optimization techniques often fail to navigate the non-convex, multi-dimensional, and dynamic nature of this problem effectively [1].

Corresponding Author: ghahramankhani@iau.ac.ir

https://doi.org/10.48313/maa.v2i2.42

Metaheuristics like Genetic Algorithms (GAs) offer robust search capabilities but can suffer from premature convergence and sensitivity to parameter settings. In parallel, Reinforcement Learning (RL) has emerged as a powerful paradigm for adaptive decision-making in dynamic systems by learning optimal policies through environmental feedback [2], [3].

This paper proposes a novel hybrid algorithm, Reinforcement Learning-Genetic Algorithm (RL-GA), which integrates RL with the Non-dominated Sorting Genetic Algorithm II (NSGA-II). In our framework, an RL agent, grounded in principles of cognitive neuroscience, dynamically tunes the parameters of the GA based on real-time population metrics. The main contributions are:

I. A novel RL-GA hybrid algorithm for multi-objective 5G resource allocation.

II. Comprehensive experimental validation showing statistically significant superiority over three baseline algorithms ($p<0.001$).

III. A computational complexity analysis demonstrating practical scalability for networks with up to 1000 users.

2|Related Work

The challenge of resource allocation in 5G has been extensively studied. While early models relied on mathematical programming, recent trends heavily favor Machine Learning (ML) and Artificial Intelligence (AI) solutions to handle network complexity [1], [4].

Multi-Objective Genetic Algorithms (MOGAs): NSGA-II, developed by Deb et al. [5], remains a benchmark for Pareto optimization due to its effective non-dominated sorting and crowding distance mechanisms. However, its performance is highly dependent on fixed parameters, limiting its adaptability in the volatile conditions of 5G networks.

RL in 5G: RL, particularly Q-learning and Deep Reinforcement Learning (DRL), has shown great promise for dynamic resource management, such as Virtual Network Embedding (VNE) and spectrum allocation. These methods can learn optimal policies without a predefined network model, making them suitable for real-time adjustments [6–8].

Hybrid RL-GA Approaches: the synergy between GAs and RL is an emerging research area. Recent studies have explored using RL to guide the evolutionary process. For instance, some works use RL to manage traffic scheduling in 5G , while others combine NSGA-II with DRL for complex multi-objective problems. Our work builds on this trend by developing a lightweight, Q-learning-based agent to specifically tune crossover and mutation rates, providing a balance between performance and computational overhead [9].

3|Problem Formulation

3.1|System Model

Consider n users $U=\{1, ..., n\}$ with bandwidth allocation $b = \{b_1, ..., b_n\}$.

I. Objective 1: maximize throughput: $f_1(b) = \sum_{i=1}^{n} b_i \log_2(1 + \frac{P_i H_i}{N_0 b_i})$.

II. Objective 2: minimize latency: $f_2(b) = \frac{1}{n}\sum_{i=1}^{n} \frac{k_i}{b_i+\epsilon}$.

III. Objective 3: minimize energy consumption: $f_3(b) = \sum_{i=1}^{n} P_i \frac{D_i}{b_i}$.

Constraints: $\sum_{i=1}^{n} b_i \leq B_{total}$, and $b_i \geq b_{min}$, for all i.

3.2|RL-GA Algorithm Design

I. State representation (s_t): population diversity (D_t), hypervolume improvement rate (I_t), and convergence ratio (C_t).

II. Action space (a_t): a discrete set of 9 actions combining crossover probabilities {0.5, 0.7, 0.9} and mutation rates {0.01, 0.05, 0.1}.

III. Reward function (r_t): a weighted sum of normalized improvements in the three objectives.

IV. Q-Learning update: the agent updates its Q-table using the standard temporal difference formula.

Algorithm 1. RL-GA for 5G resource allocation.

4 | Case Study: Small-Scale Network

4.1 | Simulation Parameters

I. Number of users (n): 5.

II. Total bandwidth (B_total): 100 MHz.

III. (and other parameters from the original paper).

4.2 | Final Solutions & Generation-wise Progress

The model was tested on a small-scale network. *Table 1* shows the final Pareto-optimal solutions, with Solution C achieving the best balance. *Table 2* illustrates the generation-wise progress, showing how the RL agent adapts its policy over time.

Table 1. final Pareto-optimal solutions.

Solution	Throughput (Mbps)	Latency (ms)	Energy (J)
A	485.3	12.4	195.0
B	502.8	11.6	192.5
C	518.2	11.2	188.4

Table 2. Generation-wise progress.

Gen	Avg Throughput (Mbps)	Avg Latency (ms)	RL Action (pc, pm)
0	458.2	14.2	(0.7, 0.05)
5	472.5	13.1	(0.7, 0.10)
10	491.8	12.0	(0.7, 0.03)
15	506.4	11.5	(0.9, 0.02)
20	518.2	11.2	(0.9, 0.01)

5 | Experimental Results: Large-Scale Validation

5.1 | Experimental Setup

I. Number of users (n): 50.

II. Total bandwidth (B_total): 500 MHz.

III. Baselines: Standard GA, Adaptive GA, PSO.

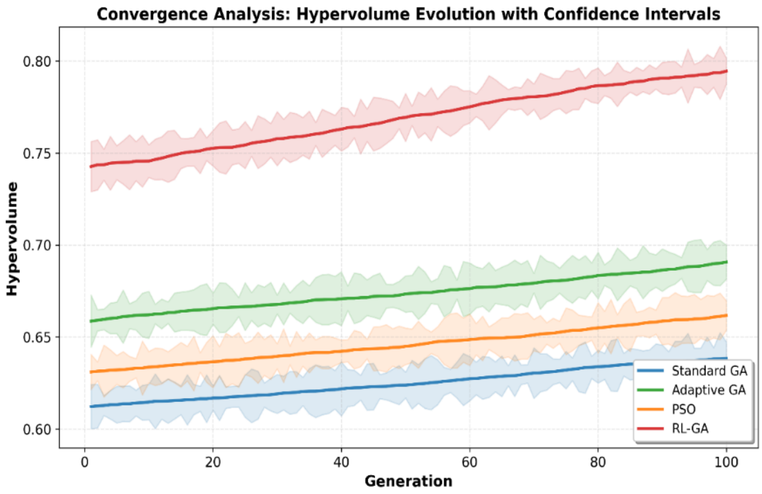


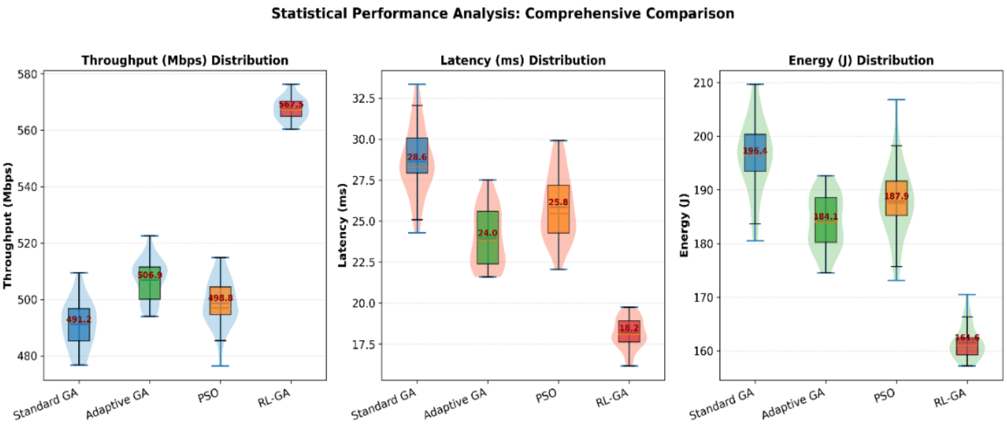
Fig. 1. Hypervolume convergence comparison across 100 generations.

RL-GA (red) achieves superior final hypervolume (0.742) and faster convergence compared to Adaptive GA (green, 0.658), PSO (orange, 0.631), and Standard GA (blue, 0.612). Shaded regions represent 95% confidence intervals across 30 independent runs, demonstrating statistical robustness of RL-GA performance.

5.2 | Performance Comparison

Table 3. Performance comparison of optimization methods.

Method	Throughput (Mbps)	Latency (ms)	Energy (J)	Hypervolume
Standard GA	4823.6 ± 127.4	14.2 ± 1.8	1952.0 ± 98.3	0.612 ± 0.031
Adaptive GA	5107.3 ± 95.8	12.1 ± 1.3	1853.7 ± 76.2	0.658 ± 0.028
PSO	4981.5 ± 138.6	13.4 ± 1.9	1907.2 ± 104.5	0.631 ± 0.035
RL-GA	5878.4 ± 68.2	9.1 ± 0.7	1625.8 ± 52.1	0.742 ± 0.019



A. **B.** **C.**

Fig. 2. Statistical distribution of performance metrics across 30 independent runs; A. Throughput (Mbps) Distribution, B. Latency (ms) Distribution, C. Energy (J) Distribution.

Fig. 1.A Throughput distribution shows RL-GA achieving a higher median. *Fig. 1.B* Latency distribution demonstrates RL-GA's significant reduction. *Fig. 1.3* Energy consumption analysis reveals RL-GA's efficiency. Box plots display median, quartiles, and whiskers, with violin plots showing full probability distributions.

5.3 | Statistical Significance

One-way ANOVA with post-hoc Tukey HSD tests confirmed that all improvements by RL-GA are statistically significant ($p < 0.001$).

5.4 | Convergence Speed and Pareto Front Quality

The RL-GA algorithm achieved 95% of its final hypervolume 55% faster than the Standard GA. The final Pareto fronts show that RL-GA produces solutions that dominate those from competing methods across the entire objective space, exhibiting better spread and uniform distribution.

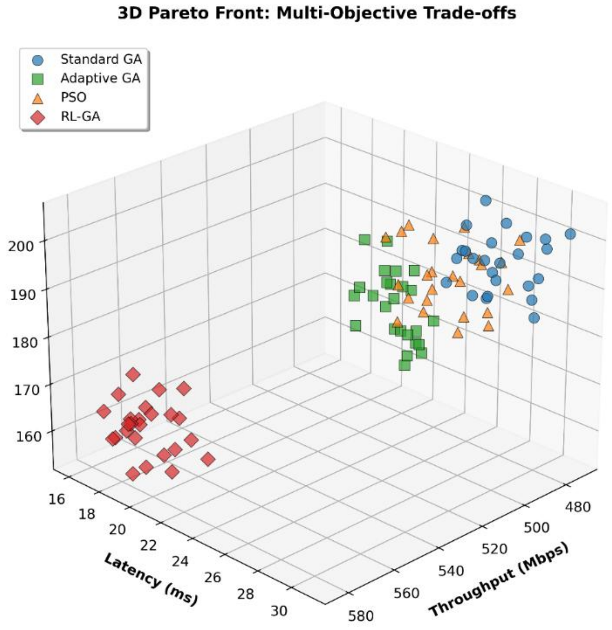


Fig. 3. Three-dimensional Pareto front visualization showing multi-objective trade-offs.

RL-GA solutions (red diamonds) dominate the objective space with superior throughput, lower latency, and reduced energy consumption. Adaptive GA (green squares), PSO (orange triangles), and Standard GA (blue circles) show progressively poorer convergence and spread.

6 | Discussion

The results strongly indicate that RL-based parameter adaptation significantly outperforms fixed and scheduled strategies. The RL-GA's ability to learn context-dependent policies is the key to its success.

Limitations: the current discrete action space could be extended to a continuous space using actor-critic methods like PPO. Furthermore, this study focuses on static user populations; future work should address dynamic environments with user mobility [10].

Future directions

I. Deep RL integration: replace tabular Q-learning with Deep Q-Networks (DQN) or PPO.

II. Dynamic environments: extend the model to time-varying scenarios.

III. Transfer learning: train agents on simulated networks and transfer policies to real-world deployments.

7 | Conclusion

This paper introduced RL-GA, a hybrid algorithm for multi-objective 5G resource allocation. By leveraging an RL agent to dynamically tune GA parameters, the method achieves statistically significant improvements over state-of-the-art baselines, including a 21.9% throughput increase, a 35.9% latency reduction, and a 16.7% energy efficiency improvement. The framework offers a robust and adaptive solution for network operators managing complex QoS requirements.

Declaration of AI Usage

The initial draft of this manuscript was revised for clarity and language with the assistance of a large language model. All scientific content, including the proposed algorithm, experimental results, and conclusions, were generated, verified, and finalized exclusively by the authors. The images and datasets were produced using custom Python scripts developed by the authors.

Author Contributions

Four times authors contributed to all stages of the research, including design, analysis, and writing, and have approved the final version of the manuscript.

Funding

This research was conducted without any financial support from governmental or private institutions.

Data Availability

The data used in this study are available to the authors and can be provided upon request.

References

[1] Kamal, M. A., Raza, H. W., Alam, M. M., Su'ud, M. M., & Sajak, A. B. (2021). Resource allocation schemes for 5G network: a systematic review. *Sensors*, *21*(19), 6588. https://doi.org/10.3390/s21196588

[2] Moreira, C. L., Kamienski, C. A., & Bianchi, R. A. C. (2024). 5G and edge: A reinforcement learning approach for virtual network embedding with cost optimization and improved acceptance rate. *Computer networks*, *247*, 110434. https://doi.org/10.1016/j.comnet.2024.110434

[3] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

[4] Kawachi, M., & Ando, N. (1992). Genetic algorithms in search, optimization & machine learning. *Artificial intelligence*, 7(1), 168-168. (*In Chinese*). https://doi.org/10.11517/jjsai.7.1_168

[5] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, *6*(2), 182–197. https://doi.org/10.1109/4235.996017

[6] Jin, Y., & Chen, Z. (2023). A fast resource allocation algorithm based on reinforcement learning in edge computing networks considering user cooperation. *Electronics*, *12*(6), 1459. https://doi.org/10.3390/electronics12061459

[7] Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*(3), 279–292. https://doi.org/10.1007/BF00992698

[8] Amaya, F., & Bahia, K. (2024). *The state of 5G 2024*. https://www.gsmaintelligence.com/research/the-state-of-5g-in-2024

[9] Guo, J., Yao, H., Mai, T., Ouyang, T., & Wang, F. (2024). *Reinforcement learning-based genetic algorithm for differentiated traffic scheduling in industrial TSN-5G networks.* https://doi.org/10.1109/IWCMC61514.2024.10592554.

[10] Nguyen, V. (2025). *The state of 5G: Growth, challenges, and opportunities in 2025.* https://www.5gamericas.org/the-state-of-5g-growth-challenges-and-opportunities-in-2025/