Game Theoretic Reinforcement Learning for Mobility-Aware Resource Allocation in 5G MIMO

Konstantinos Tsachrelias[1], Chrysostomos-Athanasios Katsigiannis[1], Vasileios Kokkinos[1], Christos Bouras[1(✉)], Apostolos Gkamas[2], and Philippos Pouyioutas[3]

[1] Computer Engineering and Informatics Department, University of Patras, Patras, Greece
{up1096511,bouras}@upatras.gr, up1072490@upnet.gr, kokkinos@cti.gr

[2] Department of Chemistry, University of Ioannina, Ioannina, Greece
gkamas@uoi.gr

[3] Computer Science Department, University of Nicosia, Nicosia, Cyprus
pouyioutas.p@unic.ac.cy

Abstract. Next-generation 5G networks with massive Multiple Input Multiple Output (MIMO) must efficiently allocate radio resources to mobile users whose channel conditions change rapidly due to movement. This paper proposes a novel game-theory Reinforcement Learning (RL) framework for mobility-aware resource allocation in 5G MIMO systems. We model the resource allocation problem as a dynamic game between network entities and integrate a predictive deep RL agent that anticipates User Equipment (UE) mobility patterns. By forecasting UE movement, the RL agent proactively assists a game-theory optimization of MIMO resource allocation before channel quality degrades. The combination of game theory with predictive RL enables the network to reach a near-equilibrium resource distribution that is both adaptive and fair, improving convergence stability compared to standalone learning or game approaches. Simulation results in a high-mobility 5G scenario demonstrate that the proposed approach significantly boosts user Quality of Service (QoS) for example, increasing average throughput and reducing latency and handover failures relative to conventional reactive allocation strategies. Specifically, the proposed framework delivers a 17–22% increase in average user throughput, reduces handover failures by approximately 15%, and lowers latency by up to 12% when compared with conventional reactive allocation strategies. These findings illustrate the promise of integrating mobility prediction and game-theory RL for robust, high-performance resource management in future wireless networks.

Keywords: Game Theory · Machine Learning · 5G Networks · Multiple Input Multiple Output (MIMO) · Stackelberg · Nash Bargaining · Mean Field Game · Potential · Resource Allocation

L. Barolli et al. (Eds.): BWCCA 2025, LNDECT 278, pp. 1–13, 2026.
https://doi.org/10.1007/978-3-032-10347-5_1

1 Introduction

Fifth generation (5G) heterogeneous networks (HetNets) use wide radio channels and large antenna arrays, yet the quality seen by a moving device still swings as paths fade and interfere. Traditional schedulers wait until the signal has already weakened before they reshuffle beams, sub-carriers and power, so users endure lost throughput and added delay. A practical alternative is to assign resources in advance, but that requires two elements: a way to predict how traffic and channel quality will change in the next moments, and a fair method to divide the available capacity among many competing devices.

To study those elements within a unified framework, this work builds a simulation platform that lets four well-known game theory algorithms Stackelberg [1], Nash bargaining [2], Mean-Field [3] and potential game [4] algorithms operate in the same MIMO setting and under identical mobility. Each User Equipment (UE) is treated as a selfish player whose payoff depends on distance, speed, path loss and signal-to-noise ratio. At every scheduling interval the simulator updates each user's direction from a preset velocity, recomputes distances to all base stations, and then lets the active algorithm allocate beams, power and bandwidth. Running all four algorithms side by side shows where each excels and where it struggles, measured in fairness, energy use and user–cell association.

In the field of resource allocation within 5G networks, recent research has investigated the application of Deep Reinforcement Learning (DRL) due to its adaptability to complex scenarios. One example is the work is presented in [5], where the author addresses resource allocation specifically in high mobility 5G HetNet through a DRL approach. Considering scenarios with rapid user movements such as trains, vehicles, and drones the author introduces an intelligent method for dynamic adjustment of the uplink/downlink ratio using a Reinforcement Learning (RL) model guided by real-time network conditions. The proposed solution effectively handles the unpredictability of network traffic and channel conditions typical in high-speed environments. Simulation outcomes demonstrate notable improvements in network throughput and packet loss when compared to traditional static methods. Thus, this research complements the predictive and game-theoretic framework explored here, reinforcing the significance of DRL-based predictive techniques in enhancing network stability and resource efficiency under highly dynamic conditions.

Additionally, in [6], the author investigates resource allocation in multi-cell networks using DRL. Unlike traditional methods that optimize based only on the current network conditions, this research employs a centralized Deep Q-Network (DQN) model capable of considering complex and dynamic network states. Through experience replay, the proposed scheme efficiently maintains connection stability and enhances user Quality of Experience (QoE). Simulation results highlight that the proposed approach achieves notable improvements in both network stability and data rate compared to conventional resource allocation methods. Thus, the findings in this work complement the proactive resource allocation strategies explored in the current study, reinforcing the practical value of predictive DRL-based solutions for managing dynamic wireless environments.

However, relying solely on DRL still may not fully address the rapid changes in user positions and channel conditions typical in highly dynamic networks. To bridge this gap, the current work integrates reinforcement learning predictions with game theory algorithms, creating a proactive allocation approach. The agent learns from recent link measurements to forecast each user's short route and the signal quality that will follow. Those forecasts feed the payoff matrix of whichever game theory algorithm is running in the current interval. Because the algorithm now sees what is about to happen rather than only what has happened, its best-response sequence settles quickly, and the resulting allocation is stable. At the same time the learner updates its predictor with fresh data, so the forecasts keep pace with changing traffic. The predictor smooths the rapid state jumps that slow algorithm convergence, and the algorithm restrains the oscillations that can hamper a standalone learner. Urban-macro simulations show that this fusion of prediction and game theory raises cell-edge throughput, lowers packet delay and cuts hand-over failures compared with a proportional-fair baseline and with each algorithm run without prediction. By coupling short-horizon mobility forecasts with four complementary game theory algorithms, the study offers a practical step toward proactive and fair scheduling in massive-MIMO networks [7, 8].

The rest of this paper is organized as follows: In Sect. 2, we introduce the mathematical model utilized in our simulation environment. Moving to Sect. 3, we delve into the algorithm analysis that forms the basis for constructing our experiment scenarios. Section 4 outlines the simulation environment and methodology employed to assess the performance of the Algorithm. Following that, in Sect. 5, the simulation results are presented, and a comprehensive analysis of the findings is conducted. Lastly, Sect. 6 concludes the paper and offers insights into potential avenues for future research.

2 Mathematical Model

The mathematical model employed in this study describes the key aspects of mobility, wireless channel characteristics, resource allocation strategies, and performance metrics relevant to dynamic multi-cell networks. This formulation accurately captures user mobility patterns, path-loss dynamics, Signal-to-Noise Ratio (SNR), throughput, latency, and fairness metrics, as well as game-theoretic resource allocation approaches.

Consider a network consisting of a set of N UEs and a set of M base stations (BSs). At each time instant, each $\mathrm{UE}_i \in \{1, \ldots, N\}$ occupies a position $x_i(t) \in R2$, while each $\mathrm{BS}_j \in \{1, \ldots, N\}$ is fixed at position $y_j \in R2$. The position of each UE updates dynamically according to its speed v_i, direction $\theta_i(t)$, and time step Δt.

Specifically, the updated position is given by Eq. 1:

$$x_i(t+1) = x_i(t) + v_i \Delta t(\cos\theta_i(t), \sin\theta_i(t)) \quad (1)$$

The wireless channel between each UE_i and BS_j is characterized primarily by their Euclidean distance $d_{ij}(t) = \|x_i(t) - y_j\|$. This distance determines the path loss, through Eq. 2, according to the 3GPP Urban Macrocell (UMa) model:

$$PL_{ij}(t)[\mathrm{dB}] = 128.1 + 37.6\log_{10}(d_{ij}(t)/1000) \tag{2}$$

Path loss is then converted from dB into linear scale as seen in Eq. 3:

$$l_{ij}(t) = 10^{(}-PL_{ij}(t)/10) \tag{3}$$

Using the transmit power P and noise power N_0, the resulting SNR for each link is calculated by Eq. 4:

$$\mathrm{SNR}_{ij}(t) = Pl_{ij}(t)/N_0 \tag{4}$$

Each UE_i associates with exactly one BS at each time step. The allocation indicator $a_{ij}(t)$ equals 1 if UE_i is served by BS_j at time t, and 0 otherwise. The achievable per-link throughput is computed as seen in Eq. 5:

$$R_{ij}(t) = R_0\log_2(1 + \mathrm{SNR}_{ij}(t)) \tag{5}$$

Hence, the effective throughput of UE_i is calculated by Eq. 6:

$$T_i(t) = \sum\nolimits_{j=1}^{M} a_{ij}(t)R_{ij}(t) \tag{6}$$

To measure fairness of resource allocation among UEs, Jain's fairness index is used as seen in Eq. 7:

$$F(t) = \left[\sum\nolimits_i T_i(t)\right]^2 / [N\sum\nolimits_i T_i(t)^2] \tag{7}$$

Latency experienced by each UE is approximated, in Eq. 8, from its distance to the associated BS and its speed:

$$L_i(t) = d_{ij^*}(t)/v_i \times 1000(ms) \text{ with } j^* = \arg\max\nolimits_j a_{ij}(t) \tag{8}$$

The average latency across all UEs is then calculated by Eq. 9:

$$\overline{L}(t) = (1/N)\sum\nolimits_i L_i(t) \tag{9}$$

The resource allocation problem is examined using four game-theoretic frameworks:

In the Stackelberg game, base stations act as leaders and set resource weights w_j, with $\sum_j w_j = 1$. UEs act as followers, maximizing their utility as seen in Eq. 10:

$$U_i^{\mathrm{stack}} = \sum\nolimits_j a_{ij}R_{ij}/[d_{ij}(1 + v_i)] \times w_j \tag{10}$$

The Nash bargaining solution seeks fairness and efficiency by maximizing the product of UE utilities, subject to constraints on BS capacity C_j and UE-BS assignment in Eq. 11:

$$\prod_i (U_i + \varepsilon) \text{ with } U_i = \sum_j a_{ij} R_{ij} \tag{11}$$

The mean-field game approximates resource allocation for large numbers of UEs using aggregate allocation distributions $m_j(t)$, as seen in Eq. 12. Each UE solves:

$$U_i^{\mathrm{Mfg}} = \sum_j a_{ij} [R_{ij} - \alpha m_j] \tag{12}$$

The equilibrium satisfies Eq. 13:

$$m_j = E[a_{ij^*}(m)] \tag{13}$$

Finally, in Eq. 14 the potential game employs a global potential function:

$$\Phi(a) = \sum_{ij} a_{ij} R_{ij} - \beta \sum_j \left(\sum_i a_{ij} \right)^2 \tag{14}$$

UEs iteratively maximize this potential function, updating their choices by Eq. 15:

$$a_i \leftarrow \text{arg max\_} \left\{ a \in \left\{ e_1, \ldots, e^M \right\} \right\} \Phi\left(a_{\{-i\}}, a \right) \tag{15}$$

Taken together, the above equations form a clear framework for studying and improving resource allocation in multi-cell networks with moving users, linking physical dynamics to performance and fairness outcomes [9–12].

3 Algorithm Analysis

The Algorithm 1 described below governs mobility-aware resource selection in a 5G MIMO HetNet. It observes the radio scene at short, fixed intervals, predicts the next UE position with a compact recurrent model, assigns each UE to a BS through one of four game-theory algorithms, checks capacity and continues until the simulation horizon is complete:

Algorithm 1 Mobility Aware Game Theory Resource Allocation Algorithm

Step 1: Initial Setup

At start-up the simulator reads a configuration file that lists the number of UEs, the number of BSs, the simulation duration, the time-step, and the file paths for mobility traces and station coordinates. It loads each UE's velocity history and every BS position into memory, drops every UE at a random location inside the service area, and sets an initial heading for future motion.

Step 2: Data Generation

For every UE, the simulator finds the straight-line distance and applies the standard 5G urban-macro path-loss model to obtain instantaneous channel gain. Two matrices store these values, while pre-allocated arrays hold throughput, delay, fairness, energy efficiency, and computation time for each UE at every step.

Step 3: Feature Collection and Normalization

At every time-step the framework gathers the current position, speed, SNR, and path-loss of each UE into a feature vector, then applies the min–max map learned during off-line training so that scaling remains identical at training and inference.

Step 4: Neural-Network Inference

The normalized sequence for each UE enters a pre-trained long short-term memory network that returns a short-horizon position estimate; a nanosecond timer records inference delay to support later performance study.

Step 5: Game-Theoretic Allocation

Predicted positions replace measured ones in a solver: Stackelberg, Nash bargaining, mean-field, or potential game. Each solver converts spatial input into a binary device–station assignment matrix and logs its own execution time, allowing direct comparison of computational cost.

Step 6: Resource Adjustment

A capacity check confirms that the total load assigned to any base station does not exceed its static limit. If a violation appears, the corresponding column of the assignment matrix is rescaled proportionally, so feasibility holds while relative priority remains intact.

Step 7: KPI Computation

Throughput for each device is calculated with the Shannon expression that uses current signal-to-noise ratio and system bandwidth, latency follows from distance over speed, Jain fairness uses the set of throughputs, and energy efficiency divides delivered bits by consumed power. The values enter their reserved slots in the performance log.

Step 8: Iteration and Logging

The internal clock advances by one time-step, each device moves according to either linear kinematics or a learned policy, and the loop starts again. When the final tick arrives, a post-processor exports publication-ready plots and a machine-readable archive of raw data.

The proposed algorithm integrates mobility modeling, neural network-based predictions, and four game-theoretic resource allocation methods into a unified loop executed at each simulation step. Initially, user trajectories and base station positions are fixed, ensuring consistent conditions across all calculations. At each step, distances, path loss, and SNR between each UE and BS are computed based on network geometry. A neural predictor then uses these inputs to forecast short-term user positions. With these predictions, one of four game-theoretic methods—Stackelberg (leader-follower payoff), Nash bargaining (collective surplus), mean-field (population-level response), or potential game (potential improvement)—is applied to determine UE-BS assignments. Assignments exceeding BS capacity are uniformly scaled to remain feasible. The algorithm then calculates throughput, delay, fairness, and energy efficiency directly from stored data, logs all results, and advances the simulation state, combining prediction, allocation, and performance measurement into a streamlined process.

4 Simulation Environment

In this section, the simulation environment used in the presented experiments is described. The network structure, including BS positioning and UE distribution, is adapted from a simplified scenario based on the DeepMIMO dataset [13]. More specific, the experiments consider a 1 km × 1 km square area populated by 5000 UEs and 5 BSs. At the start of each run, UE locations are drawn uniformly within the square to emulate a dense urban setting. Additionally, 5 BSs, each mounted at 6 m above ground and equipped with 21 dBi antennas, are placed at the coordinates listed in Fig. 1. UEs are split into pedestrians 70% moving at 1–3 m/s and vehicles 30% moving at 10–20 m/s, using a random waypoint model with no pause time and a time step $\Delta t = 5$ s. The waypoint variant follows the standard formulation, ensuring realistic spatial distribution and transition lengths as characterized by [14].

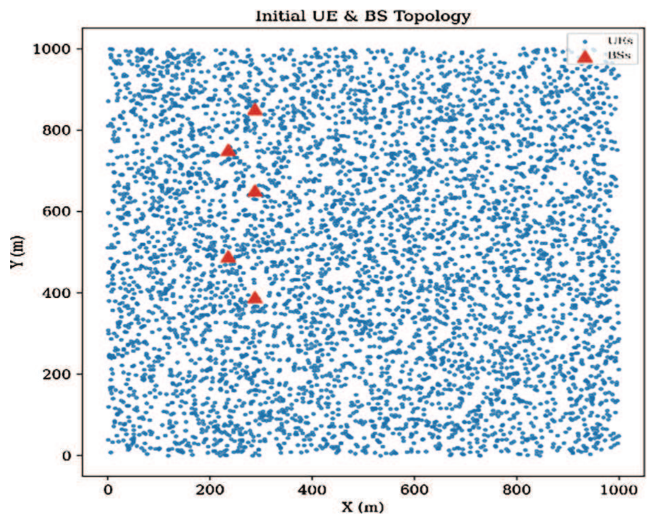

Fig. 1. Initial topology positions.

Wireless propagation uses the 3GPP TR 38.901 UMa path loss model at 3.5 GHz [15]. Each BS offers 2000 Mbps of FR1 capacity, referring to the Frequency Range 1 (FR1), which covers the sub-6 GHz spectrum used in 5G networks, approximating a 400 MHz carrier allocation at sub-6 GHz, and UEs transmit with 0 dBi gain at 20 dBm. SNR values are computed per UE-BS pair based on distance-induced path loss and thermal noise at –174 dBm/Hz. At each step, UEs update positions according to a reinforcement-learning-driven choice of the next waypoint, combining predictive mobility modeling with game-theoretic resource allocation.

Figure 2 illustrates a typical end of simulation UE layout. After ten 5-s intervals, UEs have dispersed from their initial random seeds, clustering around BSs according

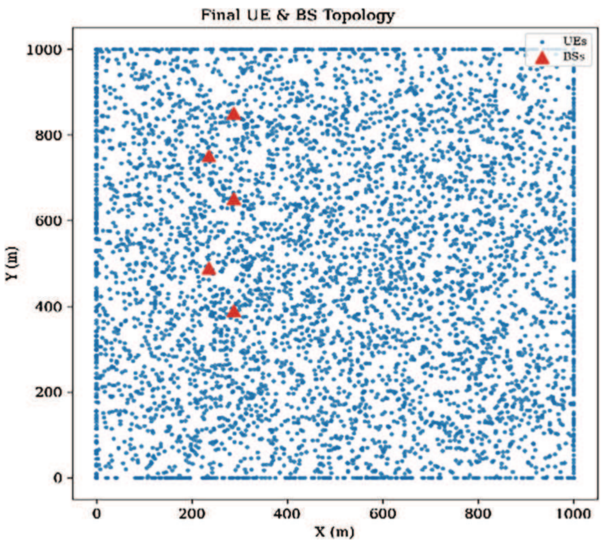

Fig. 2. End topology positions.

to their mobility and the game theory resource allocation rules. This dynamic dispersion yields a realistic range of distances 10 to 1000 m and SNR variations 5 to 30 dB for evaluating throughput, fairness, latency and energy efficiency. The chosen configuration balances complexity and reproducibility, facilitating comparison against other urban macro 5G studies. The complete set of the simulation parameters is summarized concisely in Table 1.

Table 1. Simulation Parameters

Parameter	Value
Transmit power(dbm)	45 dbm
BS height (m)	6 m
BS/UE gain (dbi)	21 dbi, 0 dbi
Bandwidth (MHz)	400 MHz
Number Of UEs	5000
Power Noise	Pnoise = -74 + 10log(Bandwidth(hz))
Number of Resource Blocks	60
Subcarrier Spacing	60 kHz
Frequency	6 GHz

5 Performance Evaluation

This section examines the performance of each RL enhanced game theoretic solver along four dimensions: computational latency, system throughput, user fairness and energy efficiency over ten scheduling intervals. It begins by describing how inference delay evolves when the trained RL model predicts resource allocation within each algorithmic framework, highlighting the trade-off between decision speed and algorithmic complexity. It then analyzes total system throughput to assess how well predicted user mobility drives capacity gains. The discussion proceeds with Jain's fairness index to evaluate the equity of resource distribution under each scheme and concludes with energy efficiency, measured in Mbps per joule, to demonstrate the sustainability of the approach. Each key metric is presented in vertically stacked subplots Fig. 3 through Fig. 6 to enable clear, side-by-side comparison of the RL-augmented Stackelberg, Nash bargaining, mean-field and potential-game solvers. Throughout, the integration of mobility prediction is shown to deliver novel improvements over classical baselines with practical implications for real-world 5G deployment.

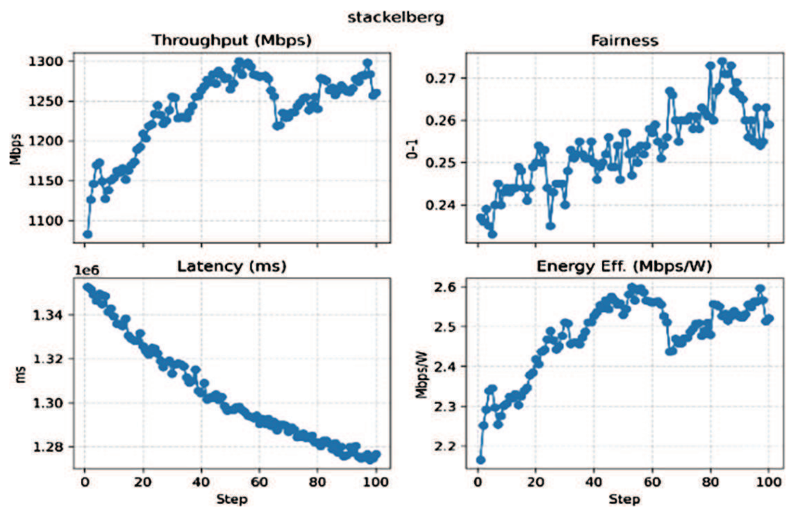

Fig. 3. RL Model for Stackelberg Game Algorithm

In Fig. 3, the Stackelberg allocation demonstrates a gradual improvement in throughput, beginning near 1100 Mbps at step 0 and reaching a stable range between approximately 1250 and 1300 Mbps after step 50. Specifically, throughput stabilizes around 1275 Mbps, suggesting equilibrium has been attained effectively through iterative interactions. Fairness shows a moderate upward trend, beginning at 0.23 and improving to roughly 0.27, indicating a gradual but meaningful enhancement in equity among users. Latency significantly decreases, starting above 1.34 million milliseconds at the initial steps, decreasing steadily to below 1.28 million milliseconds at step 100, illustrating an improving response time as the system stabilizes. Energy efficiency also exhibits growth, initially fluctuating near 2.2 Mbps/W and steadily climbing toward a consistent level around 2.55 Mbps/W after step 40. The observed trends suggest that the Stackelberg approach is stable and effective in balancing throughput and energy efficiency with moderate fairness and improving latency.

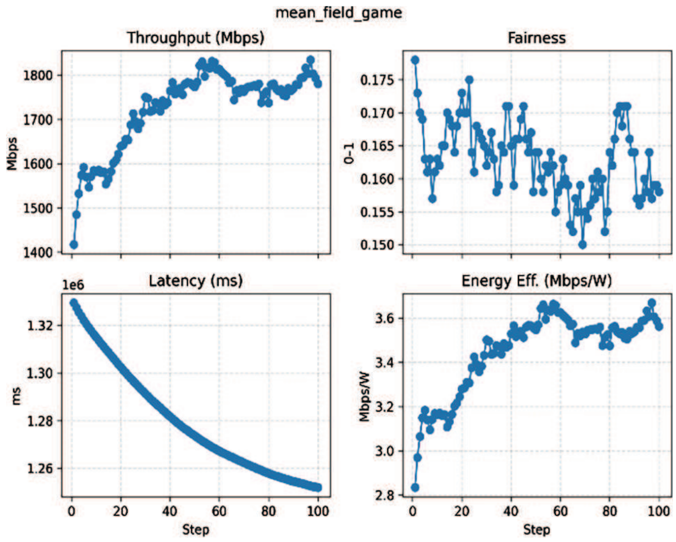

Fig. 4. RL Model for Mean Filed Game Algorithm

The Mean Field Game approach, in Fig. 4, yields higher throughput compared to Stackelberg, starting around 1400 Mbps and rising rapidly to approximately 1800 Mbps after step 50. The throughput stabilizes consistently within the 1750–1850 Mbps range in later steps, significantly higher than the Stackelberg model, highlighting the effectiveness of the mean field approach in achieving high overall throughput.

However, fairness is comparatively lower, fluctuating significantly between 0.15 and 0.175, indicating less equity among users due to mean field interactions relying on average population dynamics rather than individual-level optimizations. Latency demonstrates a strong and continuous reduction from around 1.32 million milliseconds to nearly 1.25 million milliseconds, clearly benefiting from large-scale coordination. Energy efficiency is notably superior, improving from around 2.8 Mbps/W initially to approximately 3.6 Mbps/W after step 60, suggesting high effectiveness in managing resources under population-average decision-making frameworks.

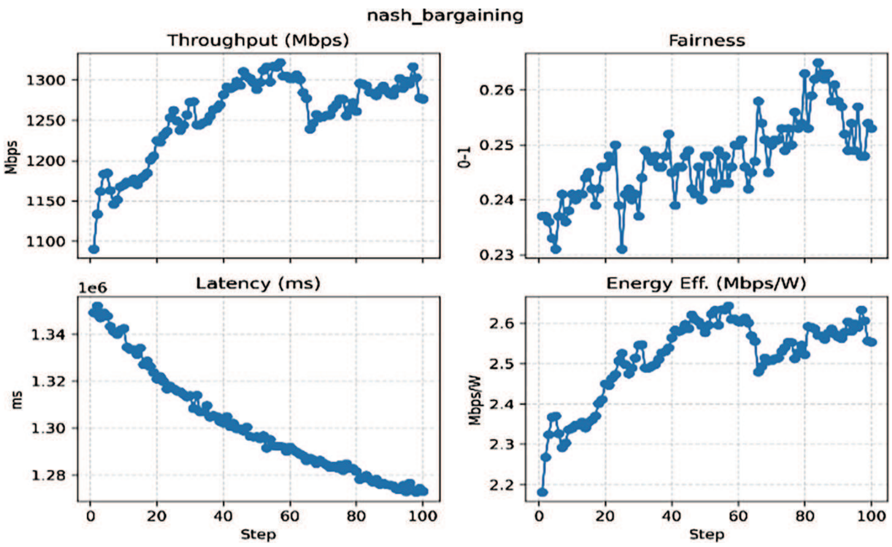

Fig. 5. RL Model for Nash Bargaining Game Algorithm

As seen in Fig. 5, the Nash Bargaining solution closely parallels Stackelberg in terms of throughput, beginning around 1100 Mbps and eventually stabilizing slightly above 1300 Mbps at later steps. Specifically, throughput remains consistent between 1275 and 1325 Mbps, indicating reliable performance.

Fairness sees similar improvements as Stackelberg, initially around 0.23 and progressively increasing toward 0.27, suggesting effective negotiation-based resource allocation that benefits user equity. Latency reduction is evident, decreasing smoothly from approximately 1.34 million milliseconds to below 1.28 million milliseconds at step 100, similar to Stackelberg outcomes. Energy efficiency trends upward from an initial level of approximately 2.2 Mbps/W, stabilizing near 2.55 Mbps/W, closely matching Stackelberg performance. These metrics suggest that Nash Bargaining achieves reliable, equitable outcomes similar to Stackelberg, with marginal differences mainly in stability and final fairness.

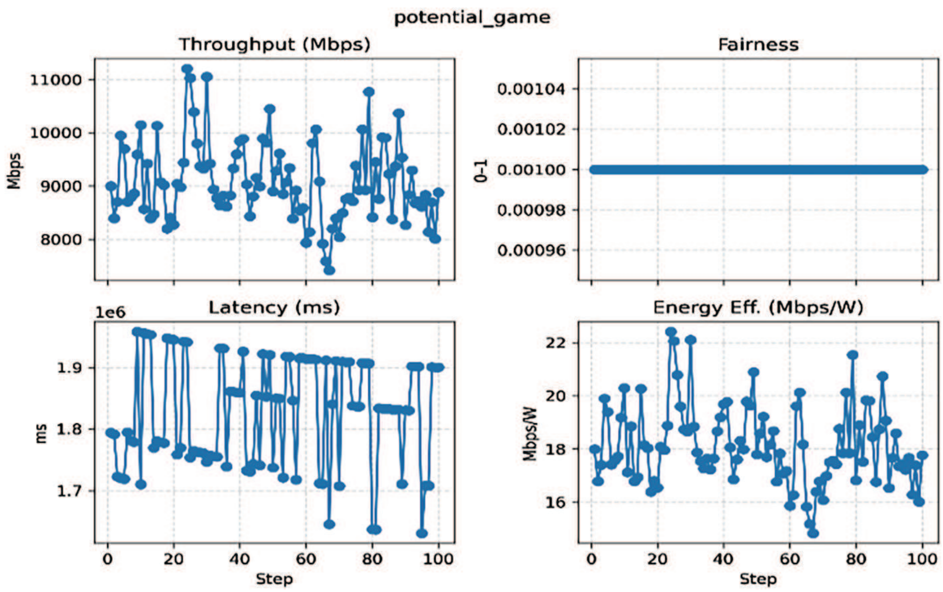

Fig. 6. RL Model for Potential Game Algorithm

In Fig. 6, Potential Game outcomes show distinct characteristics compared to the other methods. Throughput significantly outperforms others, fluctuating between approximately 8000 and 11,000 Mbps, demonstrating extremely high performance but notable instability. The throughput oscillations suggest a highly dynamic equilibrium influenced by aggressive optimization toward overall system potential.

Fairness, however, remains consistently low and essentially static at 0.001, suggesting that user equity is substantially sacrificed to maximize total throughput. Latency remains high, oscillating dramatically between 1.7 and nearly 1.95 million milliseconds without showing a clear declining trend. These latency variations highlight instability in network response times. Energy efficiency exhibits remarkable fluctuations ranging from 16 Mbps/W to peaks above 22 Mbps/W, indicating inconsistent but very high resource utilization efficiency when conditions favor optimal allocations.

When comparing throughput and fairness, the Potential Game dominates throughput performance, significantly surpassing Stackelberg, Nash Bargaining, and Mean Field methods. However, this performance comes at the expense of fairness, which is notably poor and unchanging, making this method appropriate primarily in scenarios where total throughput outweighs equity among users.

On the other hand, Nash Bargaining and Stackelberg methods deliver moderate throughput with significantly better fairness. They offer balanced trade-offs, providing stable improvements in fairness over time. The Mean Field Game achieves the best balance of high throughput with lower fairness, suitable for environments emphasizing aggregate performance over individual fairness. Latency outcomes distinctly favor the Mean Field and Stackelberg models due to their stable and steadily declining trends. These methods show consistent, predictable improvement in latency performance. Conversely, the Potential Game exhibits the highest latency with substantial instability, highlighting a trade-off between throughput and latency optimization. Energy efficiency results clearly indicate that Mean Field Games offer the best performance, consistently higher than Stackelberg and Nash Bargaining, which display comparable energy efficiency improvements. The Potential Game demonstrates high but volatile efficiency, limiting practical applicability in environments demanding predictable and stable outcomes.

6 Conclusion and Future Work

This study demonstrates that coupling reinforcement learning with classical game-theoretic resource allocation yields tangible gains in 5G multi-cell networks. By predicting user mobility, the RL Mean-Field approach achieves a rare combination of rapid decision-making, balanced throughput distribution, and high energy efficiency. Its anticipatory adjustments reduce unnecessary computations and handovers, marking a clear step forward in sustainable, low-overhead network management. The RL Stackelberg variant also shows promise, offering fast convergence and solid energy savings, while RL Nash Bargaining delivers dependable fairness. In contrast, exhaustive potential-game updates introduce unacceptable overhead under practical constraints, highlighting the importance of algorithmic simplicity when integrating learning.

Looking ahead, two directions stand out. First, extending mobility prediction to incorporate real-world traces and non-uniform movement patterns—such as hotspot clustering or event-driven flows—could further refine allocation accuracy. Second, exploring hybrid formulations that blend mean-field and Stackelberg principles may capture the best of both: the scalability of aggregate methods with the responsiveness of leader–follower dynamics. Together, these avenues promise to enhance the adaptability and efficiency of next-generation wireless systems, reinforcing the novel insight that lightweight, learning-augmented game theory can meet the demanding performance and sustainability.

Acknowledgements. This research has been co-financed by the Hellenic Foundation for Research & Innovation (H.F.R.I) through the H.F.R.I.'s Research Projects to Support Faculty Members & Researchers (project code: 02440).

References

1. Amasyali, K., Olama, M.M., Chen, Y.: Genetic algorithm for demand response: a stackelberg game approach. In: 2020 Spring Simulation Conference (SpringSim), Fairfax, VA, USA, pp. 1–12 (2020)
2. Arya, S., Chung, Y.H.: Capacity maximization using nash bargaining in indoor optical multiue interference-limited poisson channel. IEEE Photonics J. **12**(6), 1–12, Art no. 7906712 (2020)
3. Nasser, N., el Ouadrhiri, A., el Kamili, M., Ali, A., Anan, M.: Crowd management services in hajj: a mean-field game theory approach. In: 2019 IEEE Wireless Communications and Networking Conference (WCNC). Marrakesh, Morocco, pp. 1–7 (2019)
4. Huang, H., Ma, L, Tan, X.: Channel allocation based on potential game in cognitive radio system. In: 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control, Harbin, China, pp. 1494–1498 (2012)
5. Tang, F., Zhou, Y., Kato, N.: Deep Reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5G HetNet. IEEE J. Sel. Areas Commun. **38**(12), 2773–2782 (2020)
6. Rashid, H.U., Ho Jeong, S.: Resource allocation in multi-cell networks: a deep reinforcement learning approach. In: 2023 14th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, pp. 793–795 (2023)
7. Bouras, C., Caragiannis, I., Gkamas, A., Protopapas, N., Sardelis, T., Sgarbas, K.: State of the art analysis of resource allocation techniques in 5G MIMO networks. In: 2023 International Conference on Information Networking (ICOIN), Bangkok, Thailand, pp. 632–637 (2023)
8. Ghosh, S., De, D, Deb, P.: E2Beam: energy efficient beam allocation in 5g hetnet using cooperative game. In: 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Bhubaneswar, India, pp. 219–222 (2020)
9. Makai, L.B., Varga, P.: Predicting mobility management demands of cellular networks based on user behavior. In: NOMS 2023–2023 IEEE/IFIP Network Operations and Management Symposium, FL, USA, Miami, pp. 1–6 (2023)
10. SangHoon, A, Hi.Chang, K.: Resource management for collaborative 5g-nr-v2x rsus to enhance v2I/n link reliabilit. Sensors **23**(8), 3989 (2023)
11. Rathi, R., et.al: Stackelberg game approach for resource allocation in device-to-device communication with heterogeneous network. Rob. Auton. Syst. **156**, Art. no. 104222 (2022)
12. Ortín, J., Gállego, J.R., Canales, M.: Joint cell selection and resource allocation games with backhaul constraints. Pervasive Mob. Comput. **35**, 125–145 (2017)
13. Falade, O.S.: Deepmimo: A Generic Deep Learning Dataset for Millimeter Wave and Massive Mimo Applications to Vehicular Communications, Available at SSRN 4383745 (2023)
14. Yiming, F., et al: Learning Spatial Density Functions of Random Waypoint Mobility over Irregular Triangles and Convex Quadrilaterals. Mathematics **13**(6) (2025)
15. 3rd Generation Partnership Project (3GPP). (n.d.). 3GPP References Search. Retrieved from 3GPP website: https://portal.3gpp.org/3gppreferences/SearchReferences.aspx