Comparative Analysis of Scheduling Strategies for RIS-Empowered Wireless Networks with Non-Linear Energy Harvesting

Aditya Guhagarkar, Vimal Bhatia, Sumit Gautam
Department of Electrical Engineering, IIT Indore, 453552, India
Email: {ee210002005, vbhatia, sumit.gautam}@iiti.ac.in

***Abstract*—This paper investigates scheduling schemes for reconfigurable intelligent surface (RIS)-assisted wireless systems with non-linear energy harvesting, comparing time division multiple access (TDMA), random selection, constraint programming (CP), and deep Q-network (DQN) approaches. In scenarios with multiple user equipment (UE), RIS, and base station (BS) antennas, traditional methods such as TDMA and random selection provide simplicity, however struggle to maintain efficiency in dense networks. CP offers optimal scheduling but incurs high computational costs, whereas DQN enables rapid decision-making with near-optimal performance. Results show that DQN achieves approximately 90% of CP's throughput while reducing decision latency by over 1000 times, making it ideal for real-time RIS-assisted wireless networks. This work highlights the trade-off between solution quality and decision latency, emphasizing the potential of learning-based approaches for efficient scheduling in RIS-assisted wireless networks.**

***Index Terms*—Reconfigurable intelligent surfaces, deep Q-network, non-linear energy harvesting, simultaneous wireless information and power transfer.**

I. Introduction

Wireless communication networks are advancing with escalating needs for improved data rates, reduced latency, and increased energy efficiency. Reconfigurable intelligent surface (RIS)s have introduced new avenues for control of electromagnetic (EM) wave propagation to improve wireless communication performance without the need for extra active radio frequency (RF) chains [1]. At the same time, energy harvesting technologies have emerged as leading technologies to increase the operational lifetime of wireless devices and enable sustainable communications [2] [3]. A union of these technologies offers the prospect of building extremely efficient and green wireless systems. Recent developments in wireless communications have emphasized the need for smart resource allocation and scheduling to achieve optimal system performance. In multi-user systems with constrained resources, effective scheduling techniques become essential to optimize network performance overall. The scheduling problem is further complicated by the added dimensions of RIS configurations and energy harvesting limitations [4].

A. Related Works

RIS technology improves wireless communication by structuring propagation environments via (mostly) passive reflection. The authors in [1] examined RIS-based energy harvesting over Nakagami-m channels and showed substantial throughput and outage probability improvements. Their results highlighted the significance of non-linear energy harvesting models since linear models tend to overestimate performance. In [5] the authors, studied RIS-assisted zero-forcing and regularized zero-forcing beamforming for combined information and energy delivery, demonstrating the efficacy of independent optimization of beamforming and RIS phase setting. [6] also studied non-linear energy harvesting in ultra-reliable low-latency networks assisted with RIS, investigating the influence of RIS location and system parameters on overall performance. [7] presented an optimization approach for RIS-aided simultaneous wireless information and power transfer (SWIPT)-internet of things (IoT) with non-linear energy harvesting, further highlighting the contribution of RIS to energy-efficient wireless networks. Further, [8] optimized multi-RIS SWIPT-IoT systems using Karush–Kuhn–Tucker (KKT) conditions, alternating optimization, and Dinkelbach Transform, analyzing rate-energy trade-offs under varying RIS configurations and imperfect channel state information (CSI). In wireless scheduling, conventional methods such as time division multiple access (TDMA) offer simplicity and reliability, but optimization-based solutions, especially constraint programming (CP), are more efficient in terms of resource utilization. Recent advancements have also explored integrating RIS with full-duplex communication. The authors in [9] proposed a Virtual Full-Duplex (VFD) STAR-RIS scheme designed to outperform conventional full-duplex communication in scenarios with high residual self-interference.

Reinforcement learning, particularly deep Q-network (DQN) [10], has been found to be a robust tool for dynamic scheduling in challenging wireless environments. [11] examined federated learning in energy-harvesting networks, where joint energy management and user scheduling was shown to be beneficial, and [12] implemented deep reinforcement learning for distributed networked control systems, where issues with large action space were solved in DQN deployment. [13]. investigated reinforcement learning-based optimization of relay selection and transmission scheduling in unmanned aerial vehicle (UAV)-assisted millimeter wave (mmWave) vehicular networks, where reinforcement learning potential in communication efficiency optimization was proven. Notwithstanding these developments, a complete framework that combines RIS

technology, non-linear energy harvesting, and sophisticated scheduling, is an open research problem.

B. Motivations and Contributions

The integration of RIS with non-linear energy harvesting poses a scheduling challenge due to the need for efficient coordination of energy transfer and data transmission across multiple user equipment (UE)s, RIS units, and base station (BS) antennas. This paper models a system with transmitting BS antennas, RIS units, and multiple UEs. We analyze transmission and inference times under a time-switching energy harvesting model, comparing TDMA, CP, random selection, and DQN. TDMA ensures fair access, CP optimizes resource allocation, random selection provides an unbiased benchmark, and DQN uses reinforcement learning for adaptive scheduling. This analysis identifies the most effective strategy for RIS-assisted energy harvesting while optimizing system performance.

The remainder of this paper is organized as follows: Section II presents the system model, including the physical setup, channel modeling, and non-linear energy harvesting mechanism. Section III details the four scheduling approaches and their implementation. Section IV provides numerical results and comparative analysis. Finally, Section V concludes the paper.

II. System Overview

A. Physical Architecture and RIS Configuration

The considered system comprises of M transmitting antennas at the BS, N RIS, and K UE within a defined coverage area. The transmitting BS antennas are strategically placed at $(0, 0, 10)$, $(0, -5, 10)$, $(-5, 0, 10)$, $(0, 5, 10)$, and $(5, 0, 10)$ to ensure optimal coverage. The RIS units are symmetrically positioned at $(10, 0, 0)$, $(-10, 0, 0)$, $(0, 10, 0)$, and $(0, -10, 0)$. Each RIS is a 64×64 uniform rectangular array (URA) with half-wavelength spacing, providing 4,096 controllable elements per surface. Operating at a carrier frequency F, the system enhances propagation using passive reflection without additional active RF chains. UEs are randomly placed in a circle of radius R with the BS as the center.

B. Channel Modeling and Propagation Characteristics

The wireless channels follow a free-space path loss model, considering three signal paths: direct BS-to-UE, BS-to-RIS, and RIS-to-UE. In this paper, the term 'BS antenna' specifically refers to one of the M transmitting antennas located at the BS. Neither the RIS units nor the UEs have their own transmitting antennas; RIS units only passively reflect signals, and UEs receive them. Path loss is computed using the Friis equation, incorporating distance-dependent attenuation and phase shifts.

The direct path coefficient between BS antenna a and UE u is represented as:

$$h_{a,u} = \sqrt{\beta_0 d_{a,u}^{-\alpha}} e^{j\frac{2\pi}{\lambda} d_{a,u}}, \tag{1}$$

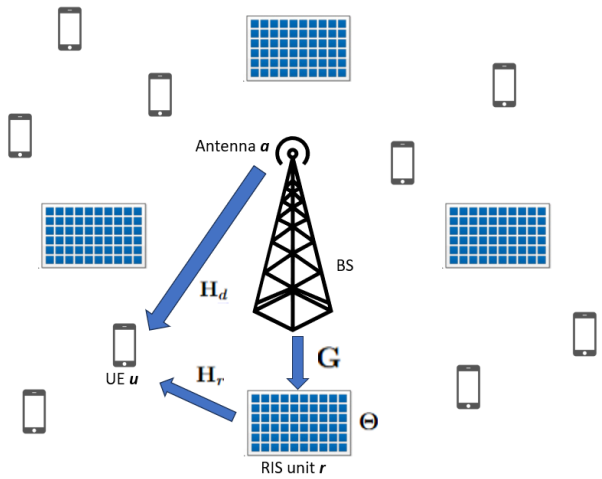


Fig. 1: Physical Architecture.

where β_0 denotes the path loss at the reference distance, $d_{a,u}$ represents the distance between a and u, α is the path loss exponent, and λ is the carrier wavelength.

The complete direct channel matrix for the given BS-UE pair is $\mathbf{H}_d \in \mathbb{C}^{1\times 1}$.

The channel between the a and the RIS unit r is given by:

$$\mathbf{G} \in \mathbb{C}^{J\times 1}, \text{ where } g_{a,n,r} = \sqrt{\beta_0 d_{a,n,r}^{-\alpha}} e^{j\frac{2\pi}{\lambda} d_{a,n,r}}, \tag{2}$$

where $d_{a,n,r}$ is the distance between a and RIS element n of r, and J is the total number of RIS elements.

Similarly, the channel between the RIS and the UE is:

$$\mathbf{H}_r \in \mathbb{C}^{1\times J}, \text{ where } h_{r,u,n} = \sqrt{\beta_0 d_{n,u,r}^{-\alpha}} e^{j\frac{2\pi}{\lambda} d_{n,u,r}}, \tag{3}$$

where $d_{n,k,r}$ is the distance between u and n of r.

The RIS reflection coefficient matrix is defined as:

$$\mathbf{\Theta} = \text{diag}\left(e^{j\phi_1}, e^{j\phi_2}, ..., e^{j\phi_J}\right) \in \mathbb{C}^{J\times J}, \tag{4}$$

where each diagonal element represents a phase shift introduced by the corresponding RIS element.

The effective end-to-end channel, incorporating both direct and RIS-assisted paths, is given by:

$$\mathbf{H}_{\text{eff}} = \mathbf{H}_d + \mathbf{H}_r \mathbf{\Theta} \mathbf{G}. \tag{5}$$

The received signal at u can be expressed as

$$\mathbf{y_{u,a,r}} = \mathbf{H}_{\text{eff}} \mathbf{x}_t \sqrt{\mathbf{P}_t} + \mathbf{n}, \tag{6}$$

where $\mathbf{x}_t$ represents the transmitted BPSK symbols. The transmission is carried out using a phased array transmitter with a peak power of $\mathbf{P}_t$ and a gain of 0 dB. $\mathbf{n} \sim \mathcal{N}(0, N_0 \mathbf{I}_K)$ is the additive white gaussian noise vector with power N_0.

C. Non-Linear Energy Harvesting Model

A distinguishing feature of this system is its implementation of a practical, non-linear energy harvesting model that accurately represents the behavior of real-world energy harvesting

circuits. Unlike conventional linear models that often overestimate harvested energy, particularly at higher power levels, the non-linear model captures the saturation effects inherent in practical rectifier circuits. This system operates in time-switching mode, where energy harvesting and information transmission occur in separate time slots. The harvested energy is modeled as:

$$E_h = \frac{\tau T E'}{1 - \chi(a,b)} \left(\frac{1}{1 + e^{-aP_y + ab}} - \chi(a,b) \right), \quad (7)$$

where $\chi(a,b) = (1 + e^{ab})^{-1}$, τ represents the time-switching factor, a and b are circuit-related parameters derived from experimental measurements [14], E' is the saturation threshold, and P_y is the power of the received signal.

D. Time Switching

At each receiver, a time-switching protocol is used to alternate between energy harvesting and information decoding. The received signal is entirely dedicated to energy harvesting for a fraction τ of the time, and the remaining $(1 - \tau)$ fraction is used for information decoding. The optimal time-switching ratio τ is dynamically determined to ensure that the energy harvesting target $\hat{E}_T$ is met while maximizing the duration available for information decoding. This optimization is performed by solving:

$$E_h(\tau, a, b, P_y, E', T) = \hat{E}_T, \quad (8)$$

for τ, ensuring that sufficient energy is harvested while maintaining adequate time for reliable information decoding.

E. RIS Phase Optimization

The phase configuration of each RIS element significantly impacts system performance. For each BS-RIS-UE combination, the system computes optimal phase configurations that align the phases of the direct path, the RIS-to-UE path, and the BS-to-RIS path to ensure constructive interference at the receiver. The optimal phase coefficient for each RIS element is calculated as:

$$\phi_{\text{opt}} = \angle(h_d) - \angle(h_r) - \angle(g), \quad (9)$$

where $\angle(x)$ is the phase angle of x, h_d represents the channel coefficient of the direct path, h_r represents the channel coefficient from the RIS to the UE, and g represents the channel coefficient from the BS antenna to the RIS. This phase alignment significantly enhances the received signal power compared to random phase configurations, demonstrating the potential of RIS to improve wireless communication performance.

F. Data Rate and Transmission Time Calculation

Following the determination of optimal time-switching ratios and RIS phase configurations, the signal-to-noise ratio (SNR) for each BS-RIS-UE pair is calculated as:

$$\text{SNR}_{u,a,r} = \frac{P_t |(\mathbf{H}_d + \mathbf{H}_r \mathbf{\Theta} \mathbf{G})_{u,a,r}|^2}{N_0}, \quad (10)$$

where N_0 is the noise power.

The achievable data rate for each BS-RIS-UE pair is calculated using Shannon's capacity formula, modified to account for the time-switching factor:

$$R_{u,a,r} = (1 - \tau) \cdot B \cdot \log_2(1 + \text{SNR}_{u,a,r}), \quad (11)$$

where B represents the system bandwidth, and $\text{SNR}_{u,a,r}$ is the signal-to-noise ratio after accounting for all signal paths and interference mitigation. The total number of bits to be transmitted is denoted as n. The time required to transmit this fixed amount of data is then computed as:

$$T_{u,a,r} = \frac{n}{R_{u,a,r}}.$$

These transmission times form the foundation for the scheduling problem, where the objective is to minimize the maximum transmission time across all UEs while ensuring that each UE maintains its required quality of service.

G. Scheduling Problem Formulation

The core challenge addressed in this paper is the optimal scheduling of UEs to BS-RIS pairs. Each UE must be assigned to exactly one BS-RIS pair, with the objective of minimizing the maximum transmission time across all system components. This can be formulated as a mixed-integer linear programming (MILP) problem where:

$$\text{minimize} \quad \max(\hat{T}, \widetilde{T}, \bar{T}) \quad (12)$$

$$\text{subject to} \quad \sum_{a=1}^{M} \sum_{r=1}^{N} X_{u,a,r} = 1, \ \forall u \in \{1, 2, ..., K\} \quad (13)$$

$$\hat{T}(a) = \sum_{u=1}^{K} \sum_{r=1}^{N} X_{u,a,r} \cdot T_{u,a,r}, \ \forall a \in \{1, 2, ..., M\} \quad (14)$$

$$\widetilde{T}(r) = \sum_{u=1}^{K} \sum_{a=1}^{M} X_{u,a,r} \cdot T_{u,a,r}, \ \forall r \in \{1, 2, ..., N\} \quad (15)$$

$$\bar{T}(u) = \sum_{a=1}^{M} \sum_{r=1}^{N} X_{u,a,r} \cdot T_{u,a,r}, \ \forall u \in \{1, 2, ..., K\} \quad (16)$$

$$X_{u,a,r} \in \{0, 1\}, \ \forall u, a, r, \quad (17)$$

where $X_{u,a,r}$ is a binary decision variable indicating whether UE u is assigned to BS antenna a and RIS r, $T_{u,a,r}$ is the transmission time for u when assigned to a and r, and $\hat{T}$, $\widetilde{T}$, and $\bar{T}$ represent the total times for BS antennas, RIS units, and UEs, respectively.

This scheduling problem is addressed using four distinct approaches: TDMA, CP, Random Selection, and DQN, each offering different tradeoffs between computational complexity and performance optimization. The comparative analysis of these approaches provides valuable insights into the most effective scheduling strategies for RIS-empowered wireless systems with non-linear energy harvesting.

III. Scheduling Optimization Approaches

A. Constraint Programming Model Architecture

The CP model formulates the scheduling problem as a mixed-integer linear program that optimally assigns UEs to BS-RIS pairs while minimizing the maximum total transmission time. This section details the mathematical formulation, implementation architecture, and solution approach of the CP model.

1) Decision Variables: The CP model employs several types of decision variables to represent the scheduling problem, summarized in Table I.

Variable	Description
$X_{u,a,r}$	Binary 3D tensor $X \in \{0,1\}^{N\times M\times L}$ indicating whether UE u is assigned to BS antenna a and RIS r.
$\bar{T}_u$	Transmission time for UE u.
$\hat{T}_a$	Transmission time for BS antenna a.
$\widetilde{T}_r$	Transmission time for RIS r.
T_{max}	Maximum transmission time across all components.
$T_{u,a,r}$	Predetermined transmission time for UE u when assigned to BS antenna a and RIS r.
D_u	Data requirement for UE u (fixed at 1 Gb).
$R_{u,a,r}$	Achievable data rate.

TABLE I: Decision Variables in the CP Model

2) Constraints: The CP model incorporates several critical constraints to ensure feasible scheduling:

1) **UE Assignment Constraint:** Each UE must be assigned to exactly one BS-RIS pair:

$$\sum_{a=1}^{M}\sum_{r=1}^{N} X_{u,a,r} = 1, \quad \forall u \in \{1,2,\ldots,K\}. \tag{18}$$

2) **BS antenna and RIS Time Aggregation:** The total time for each BS antenna and RIS is the sum of transmission times for all UEs assigned to that component:

$$\hat{T}_a = \sum_{u=1}^{K}\sum_{r=1}^{N} X_{u,a,r}\cdot T_{u,a,r}, \quad \forall a \in \{1,2,\ldots,M\}, \tag{19}$$

$$\widetilde{T}_r = \sum_{u=1}^{K}\sum_{a=1}^{M} X_{u,a,r}\cdot T_{u,a,r}, \quad \forall r \in \{1,2,\ldots,N\}, \tag{20}$$

where $T_{u,a,r}$ is calculated from the data rates:

$$T_{u,a,r} = \frac{D_u}{R_{u,a,r}} \tag{21}$$

3) **UE Time Calculation:** Each UE's transmission time depends on its assigned BS-RIS pair:

$$\bar{T}_u = \sum_{a=1}^{M}\sum_{r=1}^{N} X_{u,a,r}\cdot T_{u,a,r}, \quad \forall u \in \{1,2,\ldots,K\}. \tag{22}$$

4) **Maximum Time Constraints:** T he maximum time T_{max} must exceed all component times:

$$T_{max} \geq \hat{T}_a, \quad \forall a \in \{1,2,\ldots,M\}, \tag{23}$$

$$T_{max} \geq \widetilde{T}_r, \quad \forall r \in \{1,2,\ldots,N\}, \tag{24}$$

$$T_{max} \geq \bar{T}_u, \quad \forall u \in \{1,2,\ldots,K\}. \tag{25}$$

3) Objective Function: The objective function of the CP model is to minimize the maximum transmission time across all system components, i.e., $\min T_{\max}$. This minimization problem aims to optimize the system's efficiency by balancing the load across all BS antennas and RIS units, thereby minimizing the total time required to serve all UEs.

4) Solution Approach: The CP model is solved using the branch-and-cut algorithm implemented in MATLAB's `intlinprog` solver [15]. This algorithm combines branch-and-bound with cutting planes to efficiently search the solution space of integer programming problems. The computational complexity of this approach is exponential in the worst case, but modern solvers employ various heuristics and preprocessing techniques to achieve reasonable performance for problems of modest size.

B. DQN Model Architecture

The DQN approach formulates the scheduling problem as a sequential decision-making process within a reinforcement learning framework. This section details the reinforcement learning formulation, neural network architecture, and implementation details of the DQN model.

1) Reinforcement Learning Framework: The scheduling problem is cast as a Markov Decision Process (MDP) with the following components:

i) **State Space:** A vector representing the current state of the system, including UE assignments and resource utilization.

ii) **Action Space:** The set of possible BS-RIS pair assignments for the next unassigned UE.

iii) **Reward Function:** A scalar value assessing the quality of each assignment decision.

iv) **Transition Dynamics:** Deterministic rules governing how the system state evolves after each action.

v) **Terminal Condition:** The episode ends when all UEs have been assigned.

2) Environment Design: The DQN training environment is implemented as a custom Gym environment class (`SchedulingEnv`) with the following components:

i) **State Representation:** The state is a vector of length $N + M + L = 19$ (for 10 UE, 5 BS antennas, and 4 RIS units). The first N elements indicate whether each UE has been assigned (1) or not (0). The next M elements represent the cumulative time for each BS antenna, while the final L elements represent the cumulative time for each RIS.

ii) **Action Space:** A discrete space of size $M \times L = 20$ (for 5 BS antennas and 4 RIS units), where each action corresponds to assigning the next unassigned UE to a

specific BS-RIS pair. Action a is mapped to BS antenna $\lfloor a/L \rfloor$ and RIS $a \bmod L$.

iii) **Reward Function:** The negative of the maximum time across all BS antennas and RIS units:

$$R(s, a, s') = -\max\left(\max_{a\in\{1,\ldots,M\}} \hat{T}_a, \max_{r\in\{1,\ldots,N\}} \widetilde{T}_r\right). \tag{26}$$

This reward function encourages the agent to minimize the maximum time, aligning with the objective of the scheduling problem.

iv) **Step Dynamics:** When an action a is taken in state s, the next unassigned UE is identified based on the number of UEs already assigned. The action is decoded to determine the BS-RIS pair, after which the UE is marked as assigned. The cumulative times for the selected BS antenna and RIS are updated accordingly. Finally, the reward is computed, and the episode terminates once all UEs have been assigned.

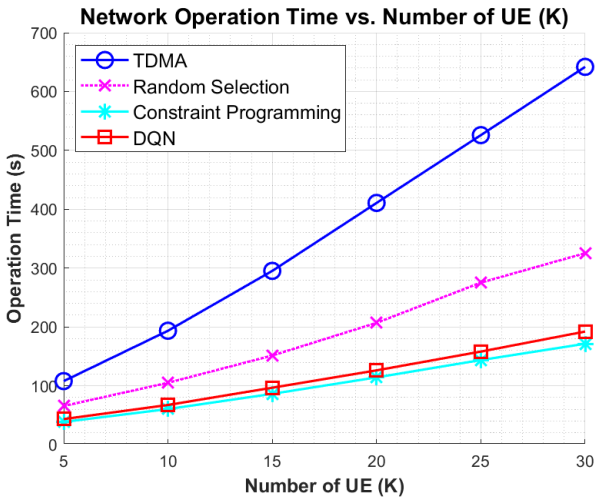
3) DQN Model and Implementation: The DQN model approximates the action-value function $Q(s, a)$ using a neural network and is implemented using the Stable Baselines3 library.

IV. Numerical Results and Analysis

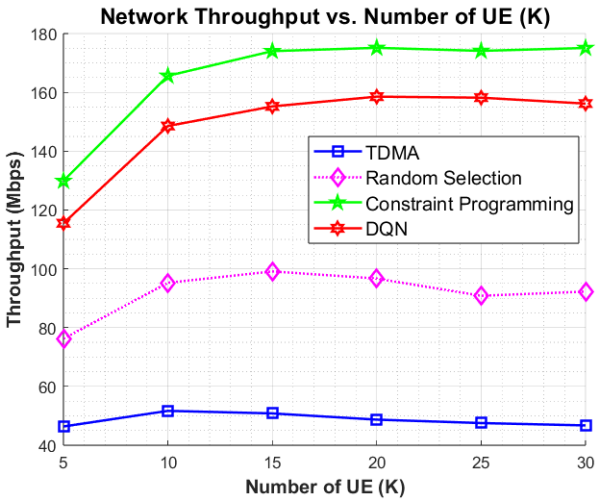
The following parameters are used in the analysis:

Symbol	Value
F	915 MHz
M	5
N	4
R	20 m
P_t	50 mW
N_0	-90 dB m
a	1500
b	0.0022
E'	2.8 mJ
T	15 s
n	1 Gbit
$\hat{E}_T$	1 mJ
B	10 MHz

TABLE II: System Parameters



(a) Operation time comparison



(b) Throughput analysis

Fig. 2: Performance tradeoffs between scheduling methods

TABLE III: Key Performance Metrics

Method	Best Case UEs (Time)	Worst Case UEs (Time)	Avg. Throughput	Inference Time
TDMA	5 (107.84s)	30 (641.98s)	46.7 Mbps	Instant
Random	5 (65.59s)	30 (325.24s)	92.3 Mbps	Instant
CP	5 (38.5s)	30 (171.35s)	173.9 Mbps	$2-42$s
DQN	5 (43.3s)	30 (192.10s)	156.2 Mbps	3ms

A. Performance Comparison

To evaluate the effectiveness of different scheduling methods, we compare their operation time and throughput against the number of users K. Fig. 2 illustrates the trade-offs between these methods, highlighting the differences in execution time (Fig. 2a) and achieved network throughput (Fig. 2b). The inference time metric is particularly critical for real-time applications, where faster decisions enable better adaptability to dynamic network conditions.

B. Key Findings

1) Optimal Method by UE Count: The optimal scheduling method varies based on the number of UEs in the system. For scenarios involving $1-15$ UEs, CP proves to be the most efficient choice, with operation times ranging from 38.5 to 86.2 seconds while maintaining an acceptable computation time between 2 to 10 seconds. As the number of UEs increases to the range of $16-25$ UEs, CP remains faster overall, with operation times between 114.20 and 143.6 seconds. However,

DQN becomes more favorable in scenarios where updates need to be made within milliseconds. When the number of UEs exceeds 25, DQN becomes the preferred approach. While CP achieves an operation time of 171.35 seconds in this range, DQN takes 192.10 seconds. Despite being 12% slower, DQN compensates by making decisions that are 1000 times faster, making it the better option for highly dynamic environments.

2) Why DQN for High Number of UEs?: DQN is the preferred choice for a high number of UEs due to several key advantages. First, it offers exceptional speed, making scheduling decisions in just 3 milliseconds, whereas CP takes 42 seconds for 30 UEs. Second, DQN provides high adaptability, allowing it to instantly adjust to changes in channel conditions, making it ideal for real-time systems. Despite being an approximation, DQN still achieves 90% of CP's throughput, making it a viable alternative in large-scale deployments. Furthermore, DQN exhibits superior scalability, as its decision-making speed remains consistent regardless of the number of UEs, whereas CP's computational time increases significantly with scale.

3) Tradeoff Summary: Each method presents trade-offs that must be carefully considered. CP offers slightly faster operation times, but its decision-making process is relatively slow, making it unsuitable for highly dynamic environments. On the other hand, DQN exhibits slightly slower overall operation times but compensates with near-instantaneous decisions, making it a better choice for real-time applications. Traditional TDMA-based approaches and random selection are simple and easy to implement; however, their efficiency drastically declines as the number of UEs increases, making them impractical for high-density networks.

C. Deployment Recommendations

In factory automation, where the network is stable with 5-20 UEs, CP provides optimal scheduling with minimal complexity. For dynamic environments like smart cities (20+ UEs), where the number of devices fluctuates, DQN is preferred for its adaptability and real-time decision-making. Similarly, in emergency networks, where rapid resource reallocation is crucial, DQN ensures responsiveness under unpredictable conditions. Lastly, CP serves as a reliable benchmark for evaluating scheduling strategies in research setups.

V. Conclusion

In this paper, we explored the use of RIS technology and non-linear energy harvesting in wireless networks to optimize scheduling strategies, comparing TDMA, Random Selection, CP, and DQN. CP performs best for fewer UEs ($5-15$) due to its ability to achieve optimal solutions via mixed-integer linear programming, but its solve time increases significantly with more UEs, limiting real-time applicability. DQN, while less optimal in operation time and throughput, excels in decision latency, making scheduling decisions within milliseconds regardless of UE count, making it ideal for dynamic environments. The choice between CP and DQN depends on application needs—CP suits static networks prioritizing optimality, while DQN is preferable for dynamic networks requiring real-time responsiveness.

References

[1] H. Alakoca, M. Babaei, L. Durak-Ata, and E. Basar, "RIS-Empowered Non-Linear Energy Harvesting Communications Over Nakagami-m Channels," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2215–2219, 2022.

[2] K. Ntontin, A. A. Boulogeorgos, E. Björnson, W. A. Martins, S. Kisseleff, S. Abadal, E. Alarcón, A. Papazafeiropoulos, F. I. Lazarakis, and S. Chatzinotas, "Wireless Energy Harvesting for Autonomous Reconfigurable Intelligent Surfaces," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 114–129, 2023.

[3] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy Harvesting Wireless Communications: A Review of Recent Advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, 2015.

[4] H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wirel. Commun.*, vol. 9, no. 5, pp. 76–83, 2002.

[5] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "RIS-Aided Zero-Forcing and Regularized Zero-Forcing Beamforming in Integrated Information and Energy Delivery," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5500–5513, 2022.

[6] S. Dhok, P. Raut, P. K. Sharma, K. Singh, and C.-P. Li, "Non-Linear Energy Harvesting in RIS-Assisted URLLC Networks for Industry Automation," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7761–7774, 2021.

[7] N. Sharma and S. Gautam, "Optimization Strategy for RIS-Assisted SWIPT-IoTs with Non-Linear Energy Harvesting," in *2024 National Conference on Communications (NCC)*, 2024, pp. 1–6.

[8] N. Sharma, S. Gautam, S. Chatzinotas, and B. Ottersten, "Integrating Multiple RIS for Enhanced SWIPT-IoT Performance: Optimization Strategies," *IEEE Access*, vol. 12, pp. 182 841–182 855, 2024.

[9] J. Jose, P. Shaik, S. Bisen, O. Krejcar, K. Choi, and V. Bhatia, "Virtual full-duplex communication enabled star-ris: Performance analysis and machine learning-assisted optimization," *IEEE Trans. Cogn. Commun. Netw.*, pp. 1–1, 2025.

[10] J. Sun, Z. Qian, X. Wang, and X. Wang, "ES-DQN-Based Vertical Handoff Algorithm for Heterogeneous Wireless Networks," *IEEE Wirel. Communn. Lett.*, vol. 9, no. 8, pp. 1327–1330, 2020.

[11] R. Hamdi, M. Chen, A. B. Said, M. Qaraqe, and H. V. Poor, "Federated Learning Over Energy Harvesting Wireless Networks," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 92–103, 2022.

[12] G. Pang, K. Huang, D. E. Quevedo, B. Vucetic, Y. Li, and W. Liu, "Deep Reinforcement Learning for Wireless Scheduling in Distributed Networked Control," *arXiv preprint*, vol. arXiv:2109.12562, 2024.

[13] A. Guhagarkar, T. Sivalingam, V. Bhatia, N. Rajatheva, and M. Latva-Aho, "Reinforcement Learning-Based Optimization of Relay Selection and Transmission Scheduling for UAV-Aided mmWave Vehicular Networks," in *2024 27th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2024, pp. 1–5.

[14] S. Gautam, E. Lagunas, A. Bandi, S. Chatzinotas, S. K. Sharma, T. X. Vu, S. Kisseleff, and B. Ottersten, "Multigroup Multicast Precoding for Energy Optimization in SWIPT Systems With Heterogeneous Users," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 92–108, 2020.

[15] The MathWorks Inc., "Optimization toolbox version: 9.4 (r2022b)," Natick, Massachusetts, United States, 2022.