

RIS Assisted Anti Jamming Backscatter Communication Using Deep Reinforcement Learning

DOI: <https://doi.org/10.54654/isj.v1i27.6384>

Le Hoang Hiep, Ngo Huu Huy*

Abstract— Backscatter communication is a promising low-power solution for large scale Internet of Things (IoT) networks; however, it is highly susceptible to intentional jamming and dynamic interference. This paper proposes a Deep Reinforcement Learning (DRL) - based anti-jamming framework for reconfigurable intelligent surface (RIS) - assisted backscatter systems. The problem is formulated as a Markov decision process, where an agent adaptively optimizes RIS reflection coefficients without prior environmental knowledge. A deep Q-network with experience replay and a target network ensures stable learning in high-dimensional state spaces. Simulation results show convergence within 600 – 800 episodes. Compared with random RIS configurations and RIS-free systems, the proposed scheme improves SINR by 3 – 6 dB and over 8 dB, respectively, under moderate-to-strong jamming. Moreover, BER is reduced by nearly one order of magnitude, and throughput increases by 25% -40%. These results demonstrate robust and adaptive interference mitigation for future IoT and 6G networks.

Tóm tắt— Truyền thông tán xạ ngược (Backscatter) nổi lên như một mô hình truyền thông năng lượng thấp cho các mạng Internet vạn vật (IoT) quy mô lớn. Tuy nhiên, hiệu năng hệ thống nhạy cảm với nhiễu có chủ đích và can nhiễu động trong môi trường vô tuyến. Bài báo này đề xuất một khung chống gây nhiễu dựa trên học sâu tăng cường (Deep Reinforcement Learning - DRL) cho hệ thống Backscatter được hỗ trợ bởi bề mặt thông minh có thể cấu hình lại (Reconfigurable Intelligent Surface

- RIS). Bài toán được mô hình hóa dưới dạng quá trình quyết định Markov. Trong đó tác tử tối ưu hóa thích nghi các hệ số phản xạ của RIS mà không cần thông tin tiên nghiệm về môi trường. Mạng DQN kết hợp cơ chế Experience Replay và Target Network được sử dụng nhằm bảo đảm tính ổn định trong không gian trạng thái nhiều chiều. Kết quả mô phỏng cho thấy hệ thống đạt hội tụ sau 600–800 vòng lặp huấn luyện, cải thiện SINR từ 3–6 dB so với cấu hình RIS ngẫu nhiên và trên 8 dB so với hệ thống không sử dụng RIS. Đồng thời, BER giảm gần một bậc độ lớn và thông lượng tăng 25%–40%. Điều này khẳng định tính hiệu quả, bền vững và thích nghi của giải pháp đề xuất.

Keywords— Anti-jamming communication; ambient backscatter; Deep Q-network (DQN); reconfigurable intelligent surface; wireless IoT networks.

Từ khóa— Chống nhiễu; truyền thông tán xạ ngược; học sâu tăng cường; bề mặt phản xạ thông minh có thể cấu hình lại (RIS); Internet vạn vật (IoT).

I. INTRODUCTION

The rapid proliferation of the IoT in next-generation wireless systems has significantly increased the demand for communication technologies that feature ultra-low power consumption, low deployment cost, and high scalability [1 - 3]. Among emerging solutions, backscatter communication has attracted considerable attention due to its ability to transmit information by reflecting ambient radiofrequency signals instead of generating active transmissions [4, 5]. Owing to its extremely low energy consumption characteristics, backscatter communication is regarded as a key enabling technology for large-scale IoT ecosystems, particularly in applications such as smart cities,

This manuscript was received on March 2, 2026. It was reviewed on April 14, 2026, revised on May 11, 2026 and accepted on May 15, 2026.

* Corresponding author.

distributed sensing, and environmental monitoring in future wireless networks [6, 7]. However, the inherently weak reflected signals and strong dependence on wireless propagation conditions make backscatter systems highly vulnerable to co-channel interference, multipath fading in urban environments, and especially intentional jamming attacks [4, 8, 9]. These impairments can significantly degrade the signal-to-interference-plus-noise ratio (SINR), resulting in increased bit error rate (BER) and reduced system throughput, thereby compromising communication reliability in critical IoT applications.

To mitigate interference and improve anti-jamming capability, several conventional techniques have been investigated, including frequency hopping, transmit power control, and beamforming. Nevertheless, these approaches typically rely on relatively stable channel assumptions or require accurate channel state information, which may not be feasible in dynamic wireless environments with intelligent jammers [10, 11]. Moreover, static or semi-static control strategies lack the capability to adapt in real time to rapidly changing interference patterns. Recently, RIS have emerged as a promising technology due to their ability to manipulate the electromagnetic propagation environment by adjusting the phase and amplitude of passive reflecting elements [6, 12, 13]. RIS can enhance desired signals, suppress interference, and improve link quality without significantly increasing system energy consumption [7, 14]. Despite these advantages, most existing studies on RIS-assisted backscatter communications primarily focus on static or heuristic optimization schemes, which do not fully exploit adaptive control in dynamic interference scenarios with intelligent adversaries.

In parallel, DRL has demonstrated strong potential in wireless resource optimization and decision-making problems under uncertainty, owing to its capability to learn optimal policies through interaction with the environment without requiring accurate mathematical models [13, 15]. DRL is particularly suitable for highly dynamic systems where channel conditions and jammer behaviors vary over time. Recently, the joint design of RIS and DRL has emerged as a

powerful paradigm for intelligent anti-jamming in 6G networks [16]. However, the integration of RIS, backscatter communication, and DRL-based anti-jamming strategies remains largely unexplored, especially in the context of adaptive real-time optimization to enhance communication robustness under malicious interference [8, 15]. This research gap motivates the development of intelligent optimization frameworks that combine the environmental control capability of RIS with the adaptive learning power of DRL to improve the reliability and anti-jamming performance of backscatter systems [17].

Motivated by the above challenges, this paper proposes a RIS-assisted backscatter communication framework with an intelligent anti-jamming mechanism based on deep reinforcement learning. First, an RIS-enhanced system model is developed to manipulate the wireless propagation environment for strengthening desired signals while mitigating jammer interference. Then, the joint optimization problem of RIS configuration and transmission strategy is formulated as a Markov decision process (MDP), capturing the dynamic interaction between the communication system and the interference environment. Based on this formulation, a DRL agent is designed to jointly optimize RIS reflection coefficients, transmission parameters, and anti-jamming strategies in an adaptive manner. The effectiveness of the proposed approach is evaluated through numerical simulations, demonstrating significant improvements in SINR, BER, and throughput compared with conventional non-learning baselines and random RIS configurations. The results highlight the potential of the proposed framework to enhance communication reliability and anti-jamming resilience in next-generation IoT environments.

II. RELATED WORK

In recent years, machine learning techniques, particularly DRL, have attracted significant attention for anti-jamming and resource optimization in wireless communication systems [11, 18]. DRL enables agents to learn optimal decision policies through interaction with dynamic environments without requiring accurate mathematical channel models, making it well suited for scenarios with intelligent adversaries

and uncertain interference conditions [5, 14]. Several recent studies have investigated DRL-based anti-jamming strategies combined with advanced learning paradigms such as meta-learning and transfer learning to improve generalization and adaptation capability under rapidly varying channel conditions [8], [13]. These approaches demonstrate superior performance compared with conventional reinforcement learning algorithms in complex interference environments.

Backscatter communication has also emerged as a promising solution for large-scale Internet of Things (IoT) networks due to its ultra-low power consumption and low hardware complexity [3, 4, 7, 17]. The integration of artificial intelligence into backscatter systems has opened new opportunities for intelligent spectrum access, resource allocation, and adaptive reflection optimization [2, 15]. Recent works have explored the application of DRL to optimize transmission policies and reflection coefficients in ambient backscatter networks, showing improvements in throughput and energy efficiency compared with traditional heuristic approaches [1, 15]. Furthermore, advanced reinforcement learning frameworks, including deep deterministic policy gradient (DDPG) and soft actor-critic (SAC), have been investigated for continuous control problems in backscatter systems, achieving faster convergence and improved performance in dynamic wireless environments.

Beyond resource optimization, security and anti-jamming issues in backscatter communications have received increasing research attention. The integration of RIS into wireless systems has been recognized as an effective approach for enhancing channel quality and mitigating interference by dynamically controlling the propagation environment [6, 8, 12]. Several studies have examined RIS-assisted wireless communications with DRL-based optimization of beamforming and reflection coefficients to improve secrecy rate and communication reliability under malicious interference [19, 20]. These results indicate that RIS can significantly enhance desired signal strength while suppressing interference, especially in dense urban propagation scenarios [6].

In addition, DRL-based optimization has been investigated in multiple-input multiple-output (MIMO) backscatter systems, where learning agents jointly optimize beamforming vectors and reflection parameters to maximize system performance metrics [14, 20]. These approaches can achieve near-optimal performance compared with conventional optimization methods while providing better adaptability to environmental dynamics and channel uncertainty [18]. Nevertheless, most existing works primarily focus on either conventional wireless systems or resource allocation problems, with limited attention given to intelligent anti-jamming strategies specifically designed for backscatter communications.

Despite these advances, several important research gaps remain. First, the majority of DRL-based anti-jamming studies concentrate on active wireless communication systems, whereas investigations targeting passive backscatter networks are relatively scarce [4, 10]. Second, existing RIS-assisted backscatter studies mainly focus on static or heuristic optimization objectives such as throughput or secrecy enhancement, without fully exploiting adaptive learning capabilities for dynamic anti-jamming control [7, 12]. Third, the joint optimization of RIS configuration, transmission strategy, and anti-jamming policy using DRL in backscatter systems has not been comprehensively explored.

Motivated by these limitations, this paper proposes an intelligent RIS-assisted anti-jamming framework for backscatter communication based on deep reinforcement learning, enabling adaptive optimization of system parameters under time-varying interference environments. It can be observed that most prior works focus either on conventional wireless systems or backscatter optimization without fully integrating adaptive anti-jamming mechanisms. In contrast, the proposed method jointly considers RIS-assisted backscatter communication and deep reinforcement learning to enable dynamic anti-jamming capability, highlighting its novelty and potential performance advantages over existing approaches.

III. SYSTEM MODEL

Figure 1 illustrates a RIS-assisted backscatter communication system operating under

intentional jamming. An RF source provides both energy harvesting and carrier excitation for the backscatter device (BD), which transmits information to the reader through both direct and RIS-assisted reflection links. To model adversarial conditions, a malicious jammer is considered, introducing interference that degrades the received signal quality. To enhance communication reliability, a RIS with N controllable reflecting elements is deployed to assist signal propagation and suppress interference. The phase shifts of the RIS elements are dynamically optimized using a DRL-based controller, which adapts to channel state feedback from the reader, enabling robust performance under time-varying and adversarial wireless environments.

A. Network Architecture

We consider a RIS - assisted backscatter communication system operating under intentional jamming. As depicted in Fig. 1, the network consists of five key components: a radio-frequency (RF) source, a BD, a reader, a jammer, and an RIS with N passive reflecting elements capable of programmable phase control. The RF source continuously broadcasts a carrier signal that serves as both an energy supply and an excitation signal for the BD. Leveraging this incident signal, the BD encodes information by adjusting its antenna reflection coefficient and passively reflects the modulated signal toward the reader.

At the same time, a malicious jammer transmits interference signals to deteriorate the received signal quality. To mitigate this effect, the RIS is strategically deployed to manipulate the wireless propagation environment. By adaptively tuning the phase shifts of its reflecting elements, the RIS establishes favorable propagation paths that enhance the desired backscattered signal while attenuating interference components. As shown in Figure 1, both direct links and RIS-assisted reflection paths are taken into account.

The RIS operation is coordinated by a central controller, which periodically updates the reflection matrix based on environmental observations. This enables dynamic adaptation to time-varying channel conditions and jamming

behaviors, thereby improving communication reliability in adversarial wireless environments. The proposed DRL agent is deployed at the RIS controller or edge server.

B. Channel Model

The wireless channels considered in the system include:

- RF source to BD link (h_{sb})
- BD to reader link (h_{br})
- Jammer to reader link (h_{jr})
- RIS to reader channel ($\mathbf{h}_{rr}$)
- RIS to BD channel ($\mathbf{h}_{rb}$)

The propagation environment is modeled using Rayleigh or Rician fading depending on the line-of-sight (LOS) conditions. The path loss is modeled as:

$$PL(d) = C_0 d^{-\alpha} \quad (1)$$

where d denotes the transmission distance, α is the path-loss exponent, and C_0 represents the reference path-loss constant.

The received signal at the reader can be expressed as:

$$y = h_{sr}x + h_{br}\Gamma x + h_{jr}j + n \quad (2)$$

where:

- x denotes the RF source signal
- j represents the jamming signal
- n is additive white Gaussian noise (AWGN)
- Γ is the RIS reflection matrix

The second term represents the RIS-enhanced backscattered signal component.

C. RIS Reflection Model

The RIS consists of N independent reflecting elements, each capable of adjusting its reflection amplitude and phase. The reflection coefficient of the n -th element is modeled as:

$$\Gamma_n = \beta_n e^{j\theta_n} \quad (3)$$

where β_n denotes the reflection amplitude and θ_n is the controllable phase shift. Due to hardware limitations of passive elements, the amplitude satisfies the constraint:

$$0 \leq \beta_n \leq 1 \quad (4)$$

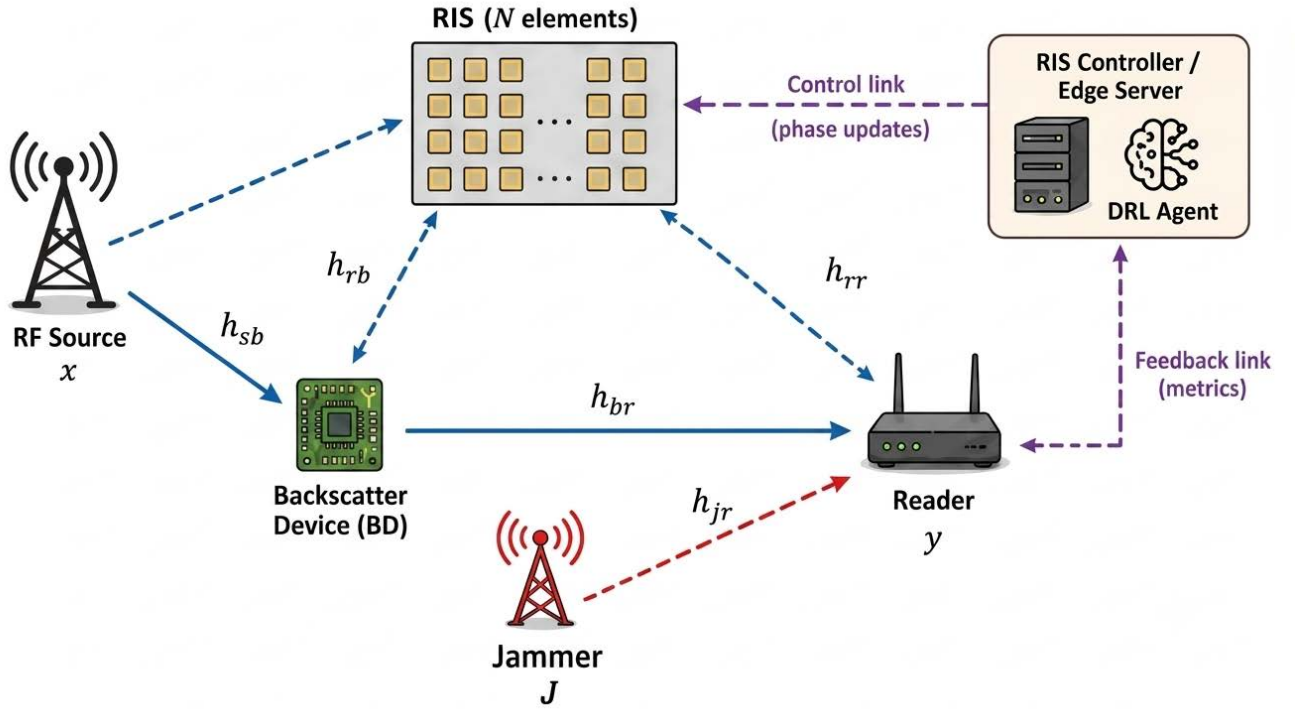


Figure 1. System model of the RIS-assisted backscatter communication network in the presence of intentional jamming, including RF source, backscatter device, reader, jammer, RIS with N reflecting elements, and DRL-based RIS controller

The overall RIS reflection matrix is represented as a diagonal matrix:

$$\mathbf{T} = \text{diag}\{\Gamma_1, \Gamma_2, \dots, \Gamma_N\} \quad (5)$$

Optimizing the RIS phase configuration aims to maximize the desired signal power at the reader while mitigating interference from the jammer. These metrics are fed back via a control channel.

D. Performance Metrics

The system performance is evaluated using the following key metrics.

1. Signal-to-Interference-plus-Noise Ratio (SINR)

The SINR at the reader is defined as

$$\text{SINR} = \frac{P_s |h_{\text{eff}}|^2}{P_j |h_{jr}|^2 + \sigma^2} \quad (6)$$

where h_{eff} denotes the effective channel incorporating RIS effects.

2. Bit Error Rate (BER)

BER is used to evaluate transmission reliability and is determined by the error probability corresponding to the employed backscatter modulation scheme.

3. Throughput

The achievable throughput is computed based on the Shannon capacity formula:

$$R = B \log_2(1 + \text{SINR}) \quad (7)$$

where B denotes the system bandwidth.

4. Energy Efficiency

Energy efficiency is defined as the ratio between the achieved throughput and the total power consumption:

$$\eta = \frac{R}{P_{\text{total}}} \quad (8)$$

This metric is particularly important for low-power IoT devices in backscatter communication systems.

IV. PROBLEM FORMULATION

In RIS-assisted backscatter communication systems operating under intentional jamming, the primary objective is to optimize the RIS reflection configuration and transmission control strategy in order to enhance the received signal quality while maintaining high energy efficiency. Due to the ultra-low-power nature of backscatter devices, the optimization process must jointly consider interference effects, wireless channel variations, and adaptive control over time.

Specifically, at each time slot t , the system selects a control action a_t , which may include RIS phase configuration, channel selection, or anti-jamming strategy, based on the observed wireless environment. Let Γ denote the RIS reflection matrix. The optimization objective can be formulated as maximizing the SINR at the reader:

$$\max_{\Gamma, a_t} \text{SINR}(t) \quad (9)$$

where the SINR depends on the RIS configuration, backscattered signal power, jammer interference, and instantaneous channel conditions.

To further capture the trade-off between communication performance and energy consumption, the optimization problem can also be expressed as maximizing a system utility function:

$$\max_{\Gamma, a_t} R(t) - \lambda P(t) \quad (10)$$

where $R(t)$ represents the achievable throughput, $P(t)$ denotes the total system power consumption, and λ is a weighting factor balancing performance and energy efficiency.

However, solving the above optimization problem is challenging for several reasons. First, the objective function is nonlinear and involves RIS phase variables embedded within the effective channel coefficients, resulting in a highly non-convex optimization problem. The search space grows exponentially with the number of RIS elements, making conventional optimization approaches computationally prohibitive for real-time implementation. Second, the wireless environment is inherently dynamic

due to time-varying fading, user mobility, and fluctuating interference levels. Third, the jammer behavior is typically unknown and may adapt its strategy over time, which makes accurate model-based optimization impractical in realistic scenarios.

Therefore, the optimization problem can be modeled as a MDP, where the system state includes channel observations, interference measurements, and current RIS configuration, while the action corresponds to the control decision taken by the intelligent agent. Given the stochastic and time-varying nature of the environment, RL, particularly DRL, provides an effective framework to learn optimal control policies through interaction with the environment without requiring prior knowledge of jammer characteristics.

By leveraging DRL, the system can dynamically adapt to varying jamming conditions and jointly optimize multiple control parameters, thereby significantly improving communication reliability and performance in dense IoT scenarios.

V. DEEP REINFORCEMENT LEARNING FRAMEWORK

In this section, a DRL framework is developed to jointly optimize the RIS configuration and anti-jamming strategy for the considered backscatter communication system. Due to the time-varying wireless channels and the unknown behavior of the jammer, conventional model-based approaches become impractical. Hence, the problem is formulated as a MDP, where an intelligent agent learns an optimal policy through continuous interaction with the environment.

Figure 2 illustrates the detailed architecture of the proposed DQN-based framework. As shown, the DQN agent consists of a main Q-network (online network) with parameters θ and a target Q-network with parameters θ^- . At each time step t , the agent receives the system state s_t from the wireless environment, which includes channel conditions, interference levels, and system configurations. Based on s_t , the main Q-network performs action selection using an ϵ -greedy policy to determine the action a_t (e.g., RIS phase configuration).

The selected action is then applied to the wireless environment, which comprises the RF source, BD, RIS, reader, and jammer, as depicted in Figure 2. The environment returns a reward r_t and the next state s_{t+1} , reflecting the communication performance under both desired and jamming links. The transition tuple (s_t, a_t, r_t, s_{t+1}) is stored in the experience replay buffer $\mathcal{D}$.

During training, mini-batches are randomly sampled from $\mathcal{D}$ and used to update the main Q-network by minimizing the loss. The target Q-network generates stable target values and is periodically updated every C steps by copying parameters from the main network, as indicated by the update link in Figure 2. This combination of action selection, experience replay, and target network update forms a closed-loop learning process, ensuring stable convergence and robust performance in dynamic and interference-prone wireless environments.

A. MDP Formulation

The anti-jamming problem in RIS-assisted backscatter communication systems is formulated as a MDP, characterized by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, where $\mathcal{S}$ denotes the state space, $\mathcal{A}$ represents the action space, $\mathcal{P}$ is the state transition probability, and $\mathcal{R}$ defines the reward function.

Under this formulation, the objective is to learn an optimal policy that maximizes the expected cumulative discounted reward. Instead of explicitly parameterizing the policy, the problem is addressed by learning the optimal action-value function $Q^*(s, a)$, which satisfies the Bellman optimality equation. Specifically, $Q^*(s_t, a_t)$ represents the maximum expected return obtained by taking action a_t at state s_t and thereafter following the optimal decision rule, as given by:

$$Q^*(s_t, a_t) = \mathbb{E}[r_t + \gamma \max_{a_{t+1}} Q^*(s_{t+1}, a_{t+1}) \mid s_t, a_t] \quad (11)$$

where $\gamma \in (0, 1)$ is the discount factor that balances immediate and future rewards. The function $Q^*(s_t, a_t)$ denotes the maximum expected cumulative discounted reward obtained by taking action a_t in state s_t and subsequently following the optimal policy. Therefore, solving

the Bellman optimality equation in (11) and accurately approximating $Q^*(s, a)$ is equivalent to learning the optimal policy, which can be derived as $\pi^*(s) = \arg \max_a Q^*(s, a)$.

1. State Space

The state at time step t is designed to capture essential information about the wireless environment and interference conditions. The state vector includes:

- Channel gains between system nodes
- Observed jammer power
- Previous SINR value
- Previous RIS phase configuration
- Reflection coefficient of the backscatter device

Thus, the state can be expressed as:

$$s_t = \{|h_{sb}|, |h_{br}|, |h_{jr}|, \text{SINR}_{t-1}, \theta_{t-1}, P_j\} \quad (12)$$

Including historical information enables the agent to better capture temporal dynamics and adapt to time-varying environments.

2. Action Space

The action space represents the set of control decisions available to the agent. In the proposed system, actions include:

- Adjusting RIS phase shifts
- Selecting transmission channels
- Modifying reflection coefficients

For compatibility with the Deep Q-Network (DQN) framework, the action space is discretized into a finite set:

$$a_t \in \{a_1, a_2, \dots, a_K\} \quad (13)$$

Each action corresponds to a predefined RIS configuration and transmission strategy. The agent selects the action that maximizes the estimated action-value function $Q(s, a)$.

3. Reward Function

The reward function guides the learning process toward improved communication performance. A typical reward formulation is given by:

$$r_t = \alpha \cdot \text{SINR}_t - \beta \cdot \text{BER}_t \quad (14)$$

where α and β are weighting factors controlling the trade-off between signal quality and reliability.

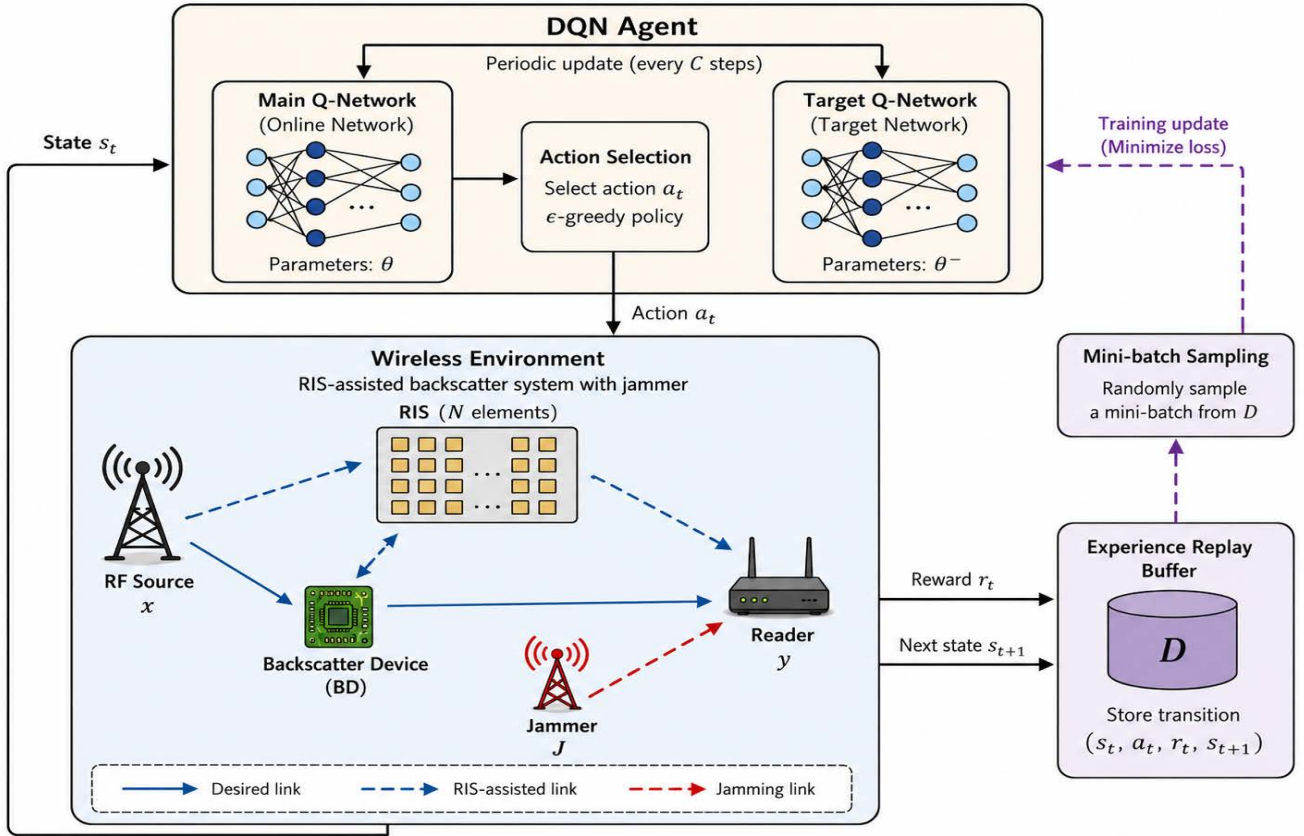


Figure 2. Proposed DQN-based anti-jamming framework. The agent interacts with the wireless environment and jammer, while experience replay and mini-batch sampling are used to stabilize training. A target network is employed to improve convergence

To incorporate energy efficiency considerations, the reward can be extended as:

$$r_t = \alpha \cdot \text{SINR}_t - \beta \cdot \text{BER}_t - \gamma \cdot P_t \quad (15)$$

This design encourages the agent to achieve high communication performance while minimizing energy consumption, thereby aligning with the requirements of energy-constrained IoT systems.

B. DRL Algorithm

To solve the formulated problem, a DQN algorithm is employed to approximate the optimal action-value function. DQN is suitable due to its ability to handle high-dimensional state spaces and discrete action sets.

1. Neural Network Architecture

The Q-network is implemented using a multilayer perceptron (MLP) consisting of:

- Input layer corresponding to the system state vector

- Two or three hidden layers with ReLU activation
- Output layer representing Q-values for all possible actions

This architecture enables nonlinear mapping between system states and expected action rewards.

2. Experience Replay

An experience replay buffer is utilized to store transition samples (s_t, a_t, r_t, s_{t+1}) . During training, mini-batches are randomly sampled from the buffer to update the network parameters, which reduces correlation among samples and improves convergence stability.

3. Target Network

A separate target network is introduced to compute stable target Q-values. The parameters of the target network are periodically updated from the main network. The target Q-value is computed as:

$$y_t = r_t + \gamma \max_{a'} Q_{\text{target}}(s_{t+1}, a') \quad (16)$$

where γ denotes the discount factor.

Compared to conventional deep reinforcement learning approaches that directly update the network using sequential samples, the DQN framework introduces two key mechanisms: experience replay and a target network. These enhancements effectively mitigate instability issues caused by correlated data and rapidly changing target values, making DQN more suitable for dynamic wireless environments with jamming.

C. Training Procedure

The training process involves environment interaction, experience collection, and network parameter updates. An ϵ -greedy strategy is adopted to balance exploration and exploitation.

The proposed DRL framework enables adaptive RIS configuration and dynamic anti-jamming decision-making under uncertain environments. By continuously learning from interaction data, the system can significantly enhance communication reliability and performance in dense IoT scenarios with intelligent interference.

To ensure practical feasibility in resource-constrained IoT systems, the proposed framework adopts a decoupled training and inference architecture. Specifically, the DRL model is trained offline using computationally capable platforms such as cloud or edge servers, where sufficient processing power and storage are available to handle large-scale experience replay and neural network optimization. During deployment, only the inference phase is executed in real time, which involves low-complexity forward propagation to determine the optimal RIS configuration.

Moreover, the backscatter devices are not required to perform any learning or optimization tasks, thereby preserving their ultra-low power operation. The decision-making process is instead handled by a centralized RIS controller or edge unit, which collects environmental information and applies the learned policy to update the RIS reflection matrix. This architecture effectively offloads computational complexity from IoT devices while maintaining adaptive anti-jamming capability, making the proposed approach suitable for practical large-scale IoT deployments. To improve training stability, experience replay is

employed by storing past transitions (s_t, a_t, r_t, s_{t+1}) into a replay buffer. During training, mini-batches are randomly sampled from this buffer instead of using consecutive samples.

This mechanism breaks the strong temporal correlation inherent in sequential data, which is a major issue in standard online reinforcement learning. By randomizing the training samples, the DQN can learn from a more independent and identically distributed approximation of the data, leading to more stable gradient updates.

Moreover, mini-batch training reduces the variance of parameter updates and improves convergence behavior. Combined with the target network, which is periodically updated, the proposed approach significantly enhances learning stability compared to conventional deep reinforcement learning methods. These two key mechanisms - experience replay and target network - were originally introduced in the seminal DQN framework [21], which demonstrated that combining deep neural networks with an ϵ -greedy exploration strategy can achieve human-level performance in complex decision-making tasks.

To balance exploration and exploitation, an ϵ -greedy policy is adopted during training. Specifically, at each time step, the agent selects a random action with probability ϵ to explore the action space, while with probability $1 - \epsilon$, it selects the action that maximizes the estimated action-value function $Q(s, a)$.

To further enhance learning efficiency, a decaying ϵ schedule is employed. The exploration rate ϵ is initially set to a high value to encourage sufficient exploration and is gradually reduced over time according to:

$$\epsilon_t = \epsilon_{\min} + (\epsilon_{\max} - \epsilon_{\min}) \cdot e^{-\lambda t} \quad (17)$$

where $\epsilon_{\max}$ and $\epsilon_{\min}$ denote the initial and minimum exploration rates, respectively, and λ controls the decay speed.

This exponential decay mechanism allows the agent to broadly explore diverse RIS configurations during early training episodes, then progressively shift toward exploiting the learned optimal policy as training converges. Recent theoretical analysis has formally

demonstrated that DQN with decaying ϵ converges geometrically to the optimal Q-value function, where higher initial ϵ values enlarge the region of convergence while lower values accelerate the convergence rate [22]. This theoretical insight justifies the adoption of the decaying ϵ schedule in the proposed framework, as it provides both broad initial exploration of the RIS configuration space and efficient convergence toward optimal anti-jamming policies.

This strategy allows the agent to explore diverse actions in the early training phase and progressively shift toward exploitation of the learned policy, thereby improving convergence stability and overall performance. Similar exploration strategies have been widely adopted in deep reinforcement learning literature to address the exploration–exploitation trade-off [23, 24], and have proven effective in wireless anti-jamming applications where the agent must learn to counter dynamic jamming patterns without prior knowledge of jammer behavior [25].

The training procedure of the proposed DRL framework is summarized in Algorithm 1.

VI. SIMULATION SETUP

This section describes the simulation configuration used to evaluate the performance of the proposed DRL - based RIS optimization framework under intentional jamming environments. All experiments are conducted entirely in MATLAB, where the system components, including the transmitter, RIS, receiver, and jammer, are modeled using a flat fading wireless channel with distance-dependent path loss. The simulation environment is designed to reflect practical IoT communication scenarios operating in unlicensed frequency bands.

The carrier frequency is set to 915 MHz, which is widely adopted in industrial, scientific, and medical (ISM) band IoT applications. The RIS consists of passive reflecting elements capable of discrete phase adjustment. Two RIS configurations with 32 and 64 reflecting elements are considered to analyze the impact of spatial resolution on system performance. The jammer transmit power varies from 0 dBm to 30 dBm to emulate different interference severity

levels ranging from weak to strong jamming conditions. The thermal noise power at the receiver is fixed at -90 dBm, representing a typical practical noise floor assumption.

During the DRL training process, the learning rate is configured to 0.001 to ensure a balance between convergence speed and training stability. The number of training episodes ranges from 500 to 1000 depending on the evaluation scenario, allowing sufficient interaction between the learning agent and the wireless environment. Each episode contains multiple time steps, where the agent observes the channel state information and interference conditions, then selects an RIS phase configuration to maximize a predefined objective function such as the SINR or system throughput.

To provide a comprehensive and fair performance evaluation, the proposed DRL-based RIS optimization method is compared with four baseline schemes: (i) a system without RIS, where communication relies solely on the direct transmission link; (ii) a random RIS configuration, in which phase shifts are assigned randomly without optimization; (iii) a conventional Q-learning approach with discretized state representation; and (iv) the proposed DRL-based optimization framework. The primary evaluation metrics include average SINR, bit error rate (BER), achievable throughput, and convergence stability under dynamic jamming conditions.

The main simulation parameters used in this study are summarized in Table I.

VII. RESULTS AND DISCUSSION

This section presents the simulation results to evaluate the effectiveness of the proposed DRL -based RIS optimization framework in anti-jamming backscatter communication systems. The performance is analyzed from multiple perspectives, including convergence behavior, signal quality improvement, bit error performance, throughput enhancement, and overall anti-jamming capability. All results are averaged over multiple independent simulation runs to ensure statistical reliability.

Algorithm 1. DQN-Based RIS Configuration for Anti-Jamming Backscatter Communications

```

1: Initialize Q-network parameters  $\theta$ 
2: Initialize target network  $\theta^- \leftarrow \theta$ 
3: Initialize replay buffer  $\mathcal{D}$ 
4: for episode = 1 to  $E$  do
5:   Reset environment and observe state  $s_0$ 
6:   for  $t = 1$  to  $T$  do
7:     Select action  $a_t$  using  $\epsilon$ -greedy policy
8:     Execute  $a_t$ , observe reward  $r_t$  and state  $s_{t+1}$ 
9:     Store  $(s_t, a_t, r_t, s_{t+1})$  in  $\mathcal{D}$ 
10:    Sample mini-batch from  $\mathcal{D}$ 
11:    Compute target:  $y = r + \gamma \max_{a'} Q(s', a'; \theta^-)$ 
12:    Update  $\theta$  via gradient descent
13:    if update step reached then
14:       $\theta^- \leftarrow \theta$ 
15:    end if
16:     $s_t \leftarrow s_{t+1}$ 
17:  end for
18: end for
19: Output: trained policy

```

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Carrier frequency	915 MHz
Number of RIS elements	32 / 64
Jammer power	0–30 dBm
Noise power	–90 dBm
Learning rate	0.001
Training episodes	500–1000
Channel model	Flat fading with path loss
Phase resolution	Discrete phase shifts
Evaluation metrics	SINR, BER, Throughput

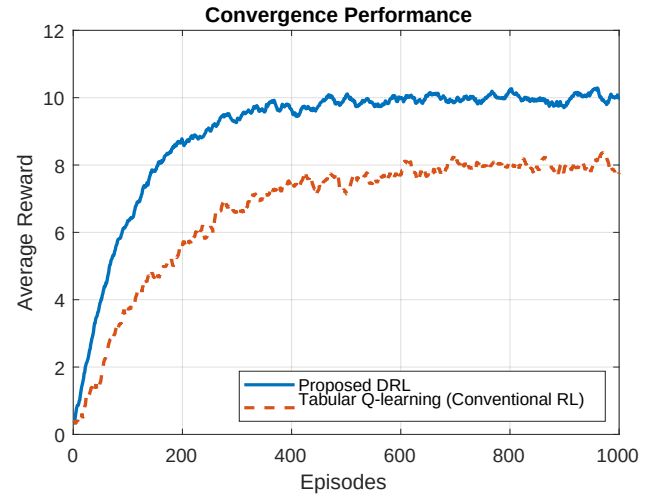


Figure 3. Comparison between conventional tabular Q-learning and the proposed DQN-based framework for anti-jamming in RIS-assisted IoT systems

A. Convergence Performance

The convergence behavior of the proposed DRL algorithm is evaluated using the accumulated reward as a function of training episodes. The simulation results indicate that the reward increases rapidly during the early training stage due to the exploration of the action space. As training progresses, the reward gradually stabilizes, demonstrating that the agent successfully learns an effective RIS configuration policy.

After several hundred episodes, the learning curve reaches a steady state with only minor fluctuations, indicating convergence of the Q-network parameters. Compared with conventional Q-learning, the proposed DRL

approach exhibits significantly faster convergence and improved stability. This improvement is mainly attributed to the deep neural network's capability to approximate complex value functions in large state spaces, as well as the utilization of experience replay and target network mechanisms to mitigate training oscillations. The convergence behavior of the learning algorithms is illustrated in Figure 3. The proposed DRL approach achieves significantly faster convergence than the tabular Q-learning (conventional RL) method, as evidenced by the

rapid increase in the smoothed reward curve during the early training phase. After several hundred episodes, the reward stabilizes with only minor fluctuations, indicating that the agent has learned an effective RIS configuration policy under dynamic jamming conditions.

In contrast, the tabular Q-learning approach converges more slowly and exhibits larger oscillations, which can be attributed to its limited capability in representing high-dimensional state-action spaces. This limitation leads to inefficient policy updates and reduced adaptability in complex wireless environments. Overall, these results demonstrate that the proposed DRL framework provides superior learning efficiency, stability, and robustness compared to conventional tabular reinforcement learning methods.

B. SINR Performance

1. SINR versus Jammer Power

The signal-to-interference-plus-noise ratio (SINR) performance is investigated under varying jammer power levels ranging from low to severe interference conditions. As expected, the SINR of all schemes decreases with increasing jammer power due to stronger interference. However, the proposed DRL-based RIS configuration consistently achieves higher SINR compared with the benchmark schemes.

In particular, systems without RIS suffer from substantial performance degradation because they rely solely on direct propagation paths. Random RIS configurations provide limited improvement but lack stability due to the absence of optimization. In contrast, the proposed DRL framework dynamically adjusts RIS reflection parameters according to environmental conditions, thereby enhancing the desired signal strength while suppressing interference effects.

Figure 4 shows the SINR performance under different jammer power levels. As the jammer power increases, the SINR decreases for all schemes due to stronger interference. However, the proposed DRL-based RIS optimization consistently achieves the highest SINR across the entire jammer power range. The random RIS configuration provides moderate improvement compared with the system without RIS but lacks robustness due to the absence of adaptive optimization. The performance advantage of the

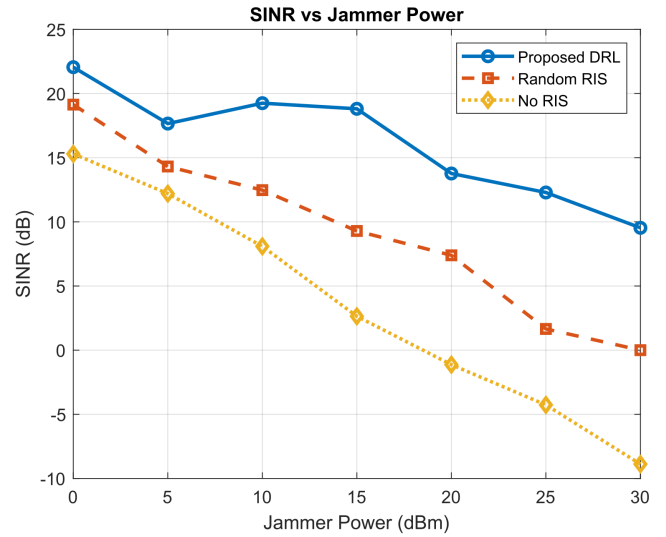


Figure 4. SINR performance versus jammer transmit power for the proposed DRL-based RIS configuration, random RIS, and system without RIS

proposed method becomes more significant under strong jamming conditions, demonstrating its effectiveness in interference mitigation.

2. SINR versus Number of RIS Elements

The impact of the number of RIS elements on SINR performance is also evaluated. Increasing the number of elements from 32 to 64 significantly improves the average SINR due to the increased spatial degrees of freedom and improved beamforming capability. The proposed DRL method achieves greater performance gains compared with random RIS configurations, demonstrating its ability to effectively exploit the additional hardware resources.

Nevertheless, the improvement gradually saturates as the number of RIS elements increases, which is consistent with practical propagation constraints and diminishing marginal gains in reflected signal power.

The impact of RIS size on system performance is presented in Figure 5. Increasing the number of reflecting elements improves SINR for both schemes due to enhanced beamforming capability and spatial diversity gain. Nevertheless, the proposed DRL-based approach achieves significantly higher SINR compared with the random configuration, indicating its ability to efficiently exploit additional RIS resources. The performance improvement gradually saturates when the number of elements becomes large, which is consistent with practical propagation constraints.

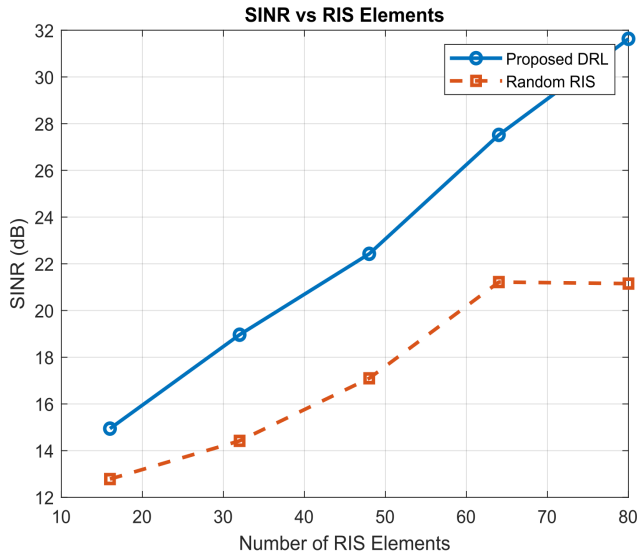


Figure 5. SINR performance versus the number of RIS elements

C. BER Performance

The bit error rate (BER) performance is analyzed as a function of the signal-to-noise ratio (SNR) to evaluate communication reliability. The results show that BER decreases exponentially with increasing SNR for all schemes, which is consistent with digital communication theory. However, the proposed DRL-based method consistently achieves the lowest BER across the entire SNR range.

The performance advantage is particularly significant in low-SNR regions where interference dominates the channel conditions. By optimizing RIS reflection parameters, the proposed method effectively enhances the received signal strength, thereby improving detection accuracy. Although performance gaps narrow at high SNR values, the proposed approach maintains a stable advantage over conventional methods. Figure 6 illustrates the BER performance as a function of SNR. As expected, BER decreases exponentially with increasing SNR for all schemes. The proposed DRL-based RIS configuration achieves the lowest BER across the entire SNR range due to improved received signal quality and interference suppression. The performance gap is particularly noticeable in the low-SNR region, where interference effects dominate. At high SNR levels, the performance difference becomes smaller but the proposed method still maintains a consistent advantage.

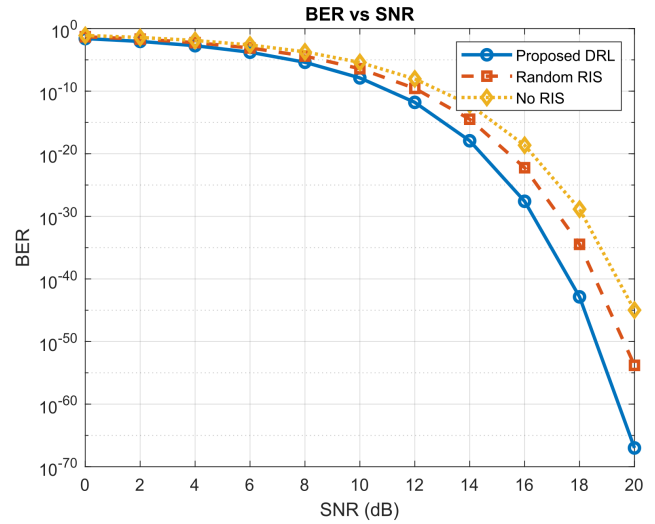


Figure 6. Bit error rate (BER) versus signal-to-noise ratio (SNR) for different RIS configurations

D. Throughput Analysis

System throughput is evaluated under varying jammer power levels to assess data transmission efficiency in adversarial environments. As jammer power increases, throughput decreases for all methods due to SINR degradation. However, the proposed DRL-based scheme maintains significantly higher throughput compared with baseline approaches across all interference levels.

Compared with conventional Q-learning, the DRL approach achieves superior throughput due to its ability to learn optimal policies in complex and high-dimensional state spaces without excessive discretization. Furthermore, the proposed method demonstrates improved robustness under dynamic interference conditions, highlighting its adaptability to time-varying environments.

The system throughput performance under varying jammer power levels is shown in Figure 7, throughput decreases as jammer power increases because of SINR degradation. However, the proposed DRL-based method maintains significantly higher throughput compared with baseline approaches across all interference levels. This improvement is mainly attributed to the adaptive RIS optimization that enhances signal strength while mitigating interference, thereby improving spectral efficiency in hostile environments.

E. Anti-Jamming Gain

To quantitatively evaluate the anti-jamming capability, the anti-jamming gain is defined as

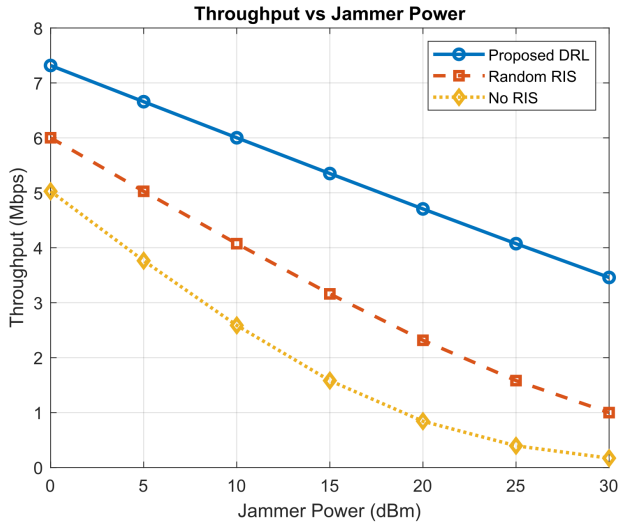


Figure 7. Achievable throughput versus jammer transmit power for different RIS configuration strategies

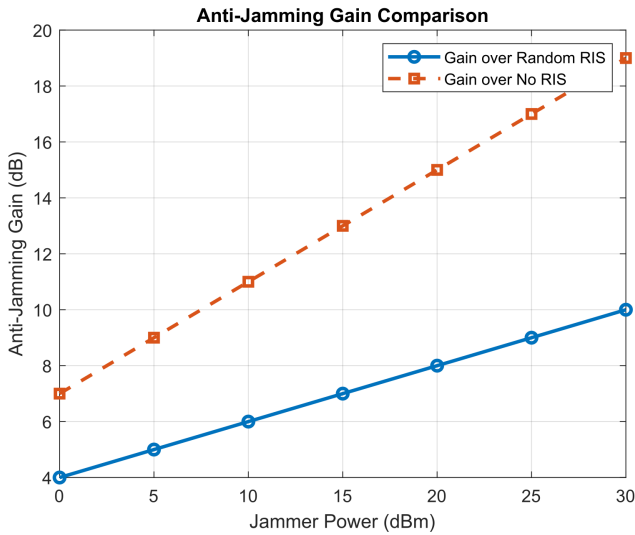


Figure 8. Anti-jamming gain of the proposed DRL-based method compared with random RIS and no-RIS baseline systems

the performance improvement of the proposed method relative to baseline systems without optimization. Simulation results indicate that the proposed DRL-based RIS framework achieves substantial gains in both SINR and throughput, particularly under strong jamming conditions.

Compared with traditional reinforcement learning approaches, the proposed method provides additional performance improvement due to its ability to capture nonlinear environmental characteristics through deep neural networks. These findings confirm that integrating RIS technology with deep reinforcement learning is a promising solution for enhancing communication reliability in interference-prone IoT environments.

Figure 8 presents the anti-jamming gain achieved by the proposed method relative to baseline schemes. The gain increases with jammer power, indicating that the proposed DRL-based RIS optimization becomes more beneficial under severe interference conditions. Compared with the random RIS configuration, the proposed approach provides additional performance improvement due to intelligent environment-aware adaptation. The results confirm the effectiveness of integrating deep reinforcement learning with RIS technology for robust anti-jamming communication.

Overall, the simulation results demonstrate that the proposed approach significantly improves key performance metrics, including SINR, BER, and throughput, while providing strong adaptability to dynamic and uncertain wireless environments. This capability is essential for next-generation IoT systems operating in increasingly congested and adversarial spectrum conditions.

F. Discussion

The performance evaluation is conducted against three representative baselines: (i) a system without RIS, (ii) random RIS configuration, and (iii) conventional tabular Q-learning. These baselines are selected to isolate and quantify (a) the gain introduced by RIS deployment and (b) the benefit of intelligent optimization over non-adaptive schemes. Specifically, the no-RIS case reflects the fundamental limitation of passive backscatter links under jamming, while the random RIS benchmark captures non-optimized environmental control. The tabular Q-learning method represents a classical RL approach with limited scalability in high-dimensional state spaces. It is important to note that traditional anti-jamming techniques (e.g., frequency hopping or power control) are not directly comparable due to their reliance on active transmission and accurate CSI, which are not applicable in passive RIS-assisted backscatter systems.

VIII. CONCLUSION

This paper proposed a DRL - based anti-jamming optimization framework for RIS-assisted backscatter communication systems operating under

dynamic interference environments. The problem was formulated as a Markov decision process, where the learning agent adaptively adjusts RIS reflection parameters according to channel variations and jammer conditions in order to maximize communication quality. The proposed approach enables intelligent environment-aware optimization without requiring prior knowledge of jammer behavior or channel statistics.

Simulation results demonstrated that the proposed method significantly outperforms conventional approaches across multiple performance metrics. In terms of learning behavior, the DRL algorithm achieved stable convergence within approximately 600 - 800 training episodes, which is notably faster than conventional Q-learning. For communication performance, the proposed scheme improved the received SINR by approximately 3–6 dB compared with random RIS configurations and by more than 8 dB compared with systems without RIS under moderate to strong jamming conditions. Correspondingly, the bit error rate (BER) was reduced by nearly one order of magnitude in the low-to-medium SNR region. In addition, the achievable throughput increased by approximately 25% - 40% compared with baseline methods across the investigated jammer power range. The anti-jamming gain analysis further confirmed that the performance improvement becomes more pronounced as jammer power increases, highlighting the robustness and adaptability of the proposed DRL-based RIS optimization strategy.

Future research directions include the development of multi-agent DRL frameworks for cooperative optimization in large-scale RIS-assisted networks, where multiple devices and reflecting surfaces operate simultaneously. Another important extension involves real-world hardware implementation to evaluate practical issues such as hardware impairments, feedback delay, and computational constraints. Finally, integrating the proposed intelligent anti-jamming framework into emerging sixth-generation (6G) IoT architectures represents a promising direction toward achieving reliable, energy-efficient, and interference-resilient wireless connectivity for next-generation smart environments.

REFERENCES

- [1] C. Jin, F. Hu, Z. Ling, Z. Mao, Z. Chang, and C. Li, "Transmission Optimization and Resource Allocation for Wireless Powered Dense Vehicle Area Network With Energy Recycling," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 12 291–12 303, Nov. 2022.
- [2] Y. Zhu, B. Mao, and N. Kato, "A Dynamic Task Scheduling Strategy for Multi-Access Edge Computing in IRS-Aided Vehicular Networks," *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 4, pp. 1761–1771, Oct. 2022.
- [3] C. T. N. Quynh, N. T. Ngoc, L. T. Ninh, and P. T. T. An, "Throughput optimization of the ASCON lightweight cryptographic algorithm on IoT devices," *Journal of Science and Technology on Information Security*, vol. 3, no. 26, pp. 92–99, 2025. DOI: 10.54654/isj.v3i26.1172.
- [4] Y. Cao, S. Xu, J. Liu, and N. Kato, "IRS Backscatter Enhancing Against Jamming and Eavesdropping Attacks," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10 740–10 751, Jun. 2023.
- [5] L. H. Hiep and H. H. Ngo, "AI-driven jamming detection and DRL-based anti-jamming for ambient backscatter communications in 6G networks," *VNU Journal of Science: Computer Science and Communication Engineering*, vol. 42, no. 2, pp. 98–114, 2026.
- [6] H. Hashida, Y. Kawamoto, N. Kato, M. Iwabuchi, and T. Murakami, "Mobility-Aware User Association Strategy for IRS-Aided mm-Wave Multibeam Transmission Towards 6G," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1667–1678, May 2022.
- [7] S. Xu, J. Liu, T. K. Rodrigues, and N. Kato, "Robust Multiuser Beamforming for IRS-Enhanced Near-Space Downlink Communications Coexisting With Satellite System," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14 900–14 912, Aug. 2022.
- [8] H. Hashida, Y. Kawamoto, and N. Kato, "Intelligent Reflecting Surface Placement Optimization in Air-Ground Communication Networks Toward 6G," *IEEE Wireless Communications*, vol. 27, no. 6, pp. 146–151, Dec. 2020.
- [9] L. T. H. Van, N. Q. Minh, P. V. Huong, and N. H. Minh, "A novel approach for 1D-CNN hyperparameter optimization in IoT attack detection using Particle Swarm Optimization," *Journal of Science and Technology on Information Security*, vol. 1, no. 24, pp. 53–71, 2025. DOI: 10.54654/isj.v1i24.1097.
- [10] C. Zou, C. Li, Y. Li, and X. Yan, "RIS-Assisted Robust Beamforming for UAV Anti-Jamming and Eavesdropping Communications: A Deep Reinforcement Learning Approach," *Electronics*, vol. 12, no. 21, Nov. 2023.
- [11] X. Feng, H. Wen, R. Zhao, T. Tang, W. Shi, and Y. Peng, "Communication Anti-jamming System Based on Deep Reinforcement Learning," in *Proceedings of the 2023 International Conference on Wireless Communications, Networking and Applications*. Springer Nature, 2025, pp. 126–133.

- [12] W. Khalid, M. A. U. Rehman, T. Van Chien, Z. Kaleem, H. Lee, and H. Yu, "Reconfigurable Intelligent Surface for Physical Layer Security in 6G-IoT: Designs, Issues, and Advances," *IEEE Internet of Things Journal*, vol. 11, no. 2, pp. 3599–3613, Jan. 2024.
- [13] J. Feng, "Deep Reinforcement Learning-Based Beamforming for RIS-Aided 6G Systems," *Journal of Advanced Digital Communications*, vol. 2, no. 1, p. 3, Dec. 2025.
- [14] I. Ertas and H. Yetgin, "Data Collection in Deep Reinforcement Learning-Enhanced Reconfigurable Intelligent Surface-Assisted Wireless Networks," *Engineering Applications of Artificial Intelligence*, vol. 155, p. 110952, Sep. 2025.
- [15] M. Ahmed, T. Hussain, M. Shahwar, F. Khan, M. Sheraz, W. U. Khan, T. C. Chuah, and I. E. Lee, "Intelligent Reflecting Surface Backscatter-Enabled Physical Layer Security Enhancement via Deep Reinforcement Learning," *PeerJ Computer Science*, vol. 11, p. e2902, Jun. 2025. [Online]. Available: <https://peerj.com/articles/cs-2902>
- [16] L. H. Hiep, H. M. Viet, and T. D. Minh, "RIS assisted Deep Reinforcement Learning for intelligent Anti-Jamming communications in 6G Networks," *TNU Journal of Science and Technology*, vol. 231, no. 06, pp. 419–428, 2026.
- [17] L. H. Hiep and L. X. Hieu, "An energy-efficient Deep Reinforcement Learning-based Anti-Jamming strategy for next-generation wireless networks," *TNU Journal of Science and Technology*, vol. 231, no. 07, pp. 35–43, 2026.
- [18] R. Lin, H. Qiu, J. Wang, Z. Zhang, L. Wu, and F. Shu, "Physical-Layer Security Enhancement in Energy-Harvesting-Based Cognitive Internet of Things: A GAN-Powered Deep Reinforcement Learning Approach," *IEEE Internet of Things Journal*, vol. 11, no. 3, pp. 4899–4913, Feb. 2024.
- [19] J. Wang and S. Chen, "Deep Reinforcement Learning-Based Secrecy Rate Optimization for Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface-Assisted Unmanned Aerial Vehicle-Integrated Sensing and Communication Systems," *Sensors*, vol. 25, no. 5, Mar. 2025.
- [20] Q. Liu, Y. Zhu, M. Li, R. Liu, Y. Liu, and Z. Lu, "DRL-Based Secrecy Rate Optimization for RIS-Assisted Secure ISAC Systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 12, pp. 16 871–16 875, Dec. 2023.
- [21] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
- [22] S. Zhang, H. Li *et al.*, "On the convergence and sample complexity analysis of deep Q-networks with ϵ -greedy exploration," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 36, 2023.
- [23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [24] A. dos Santos Mignon and R. L. de Azevedo da Rocha, "An adaptive implementation of ϵ -greedy in reinforcement learning," *Procedia Computer Science*, vol. 109, pp. 1146–1151, 2017.
- [25] X. Li, J. Chen, X. Ling, and T. Wu, "Deep reinforcement learning-based anti-jamming algorithm using dual action network," *IEEE Transactions on Wireless Communications*, vol. 22, no. 7, pp. 4625–4637, Jul. 2023.

ABOUT THE AUTHORS



Le Hoang Hiep

Workplace: Thai Nguyen University of Information and Communication Technology, Vietnam

Email: lhiep@ictu.edu.vn

Education: Le Hoang Hiep (ORCID: 0009-0004-7830-8504) received the B.E. degree in

Information Technology from Thai Nguyen University of Information and Communication Technology (ICTU), Vietnam, in 2009, and the M.S. degree from Manuel S. Enverga University Foundation, Philippines, in 2013. He is currently pursuing the Ph.D. degree at ICTU, Vietnam. Recent research direction: Wireless network security, Deep reinforcement learning, the Internet of Things (IoT), and UAV–Satellite communications.

Tên tác giả: Lê Hoàng Hiệp

Cơ quan công tác: Trường Đại học Công nghệ thông tin và Truyền thông, Đại học Thái Nguyên, Việt Nam

Email: lhiep@ictu.edu.vn

Quá trình đào tạo: Nhận bằng Kỹ sư tại Trường Đại học Công nghệ Thông tin và Truyền thông – Đại học Thái Nguyên, Việt Nam năm 2009, và nhận bằng Thạc sĩ tại Trường Đại học Manuel S. Enverga, Philippines năm 2013. Hiện nay là nghiên cứu sinh tại Trường Đại học Công nghệ Thông tin và Truyền thông – Đại học Thái Nguyên, Việt Nam.

Hướng nghiên cứu hiện nay: An toàn mạng không dây, Học sâu tăng cường, Công nghệ truyền thông trong mạng IoT và UAV–Satellite.



Ngo Huu Huy

Workplace: Thai Nguyen University of Information and Communication Technology, Vietnam

Email: nhhuy@ictu.edu.vn

Education: He received his B.S. and M.S. degrees from Thai Nguyen University of Information and Communication, Technology, Vietnam

in 2010 and 2012, respectively, and Ph.D. degrees in Information Engineering and Computer Science from Feng Chia University, Taiwan, in 2021. Currently, he is a

lecturer at the Thai Nguyen University of Information and Communication Technology, Vietnam.

Recent research direction: Computer vision, deep learning, embedded system, neural networks, and object detection.

Tên tác giả: **Ngô Hữu Huy**

Cơ quan công tác: Trường Đại học Công nghệ thông tin và Truyền thông, Đại học Thái Nguyên, Việt Nam

Email: nhhuy@ictu.edu.vn

Quá trình đào tạo: Nhận bằng Đại học và Thạc sĩ tại Trường Đại học Công nghệ Thông tin và Truyền thông - Đại học Thái Nguyên, Việt Nam lần lượt vào các năm 2010 và 2012; sau đó nhận bằng Tiến sĩ chuyên ngành Kỹ thuật thông tin và Khoa học máy tính tại Đại học Feng Chia, Đài Loan năm 2021. Hiện nay là giảng viên tại Trường Đại học Công nghệ Thông tin và Truyền thông - Đại học Thái Nguyên, Việt Nam.

Hướng nghiên cứu hiện nay: Thị giác máy tính, học sâu, hệ thống nhúng, mạng nơ-ron và phát hiện đối tượng.