Scientific Research Publishing

Journal of Computer and Communications, 2026, 14(6), 1-12
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Intelligent Frequency-Hopping Strategies for Securing the Physical Layer Based on Learning for the Stochastic Dispersion Problem

Vincent Mbonigaba[1], Dieudonné Nijimbere[2]

[1]Higher Institute of Applied Sciences, University of Burundi, Bujumbura, Burundi
[2]Higher Institute of Military Academy, Bujumbura, Burundi
Email: mbonivinci@gmail.com

How to cite this paper: Mbonigaba, V. and Nijimbere, D. (2026) Intelligent Frequency-Hopping Strategies for Securing the Physical Layer Based on Learning for the Stochastic Dispersion Problem. *Journal of Computer and Communications*, **14**, 1-12.
https://doi.org/10.4236/jcc.2026.146001

Received: March 30, 2026
Accepted: June 8, 2026
Published: June 11, 2026

Abstract

This article addresses the security of the physical layer against adaptive jamming attacks in wireless networks. We propose an intelligent frequency-hopping strategy that combines the Q-learning algorithm with a stochastic spreading mechanism to optimize transmission resilience. Unlike conventional FHSS, our approach allows the transmitter to autonomously learn compromised channels and adjust its hopping policy in real time. Stochastic spreading ensures maximum signal unpredictability, preventing any prediction by a cognitive adversary. Simulations demonstrate rapid convergence of the algorithm and a significant increase in the transmission success rate, reaching 100% at an SNR of 8 dB. Complexity analysis confirms that the induced latency is compatible with real-time requirements. This work paves the way for self-organizing and robust communication systems for future 5G/6G networks and the Internet of Things.

Keywords

Physical Layer Security, Frequency Hopping, Q-Learning, Anti-Jamming, Stochastic Scattering, Reinforcement Learning, Spectral Agility

1. Introduction

In today's wireless communications landscape, securing the physical layer has become an essential defense against the emergence of increasingly sophisticated and adaptive jamming threats [1]. While traditional frequency-hopping methods often rely on predefined switching sequences that are vulnerable to interception or prediction, the integration of artificial intelligence opens new avenues for enhanced

resilience. This article explores an intelligent frequency-hopping strategy designed to counter malicious adversaries in dynamic and uncertain environments. By leveraging the power of Q-learning, a reinforcement learning algorithm, the system becomes capable of autonomously learning jamming patterns and selecting optimal transmission channels in real time, without prior knowledge of the attacker's strategy [2].

To overcome the limitations of purely random exploration and avoid getting stuck in local optima, we introduce a stochastic dispersion mechanism. This approach allows us to diversify frequency hops probabilistically, ensuring maximum unpredictability against the adversary's tracking attempts while minimizing packet collisions. Through this synergy between reinforcement learning and controlled randomness, our work aims to optimize transmission success rates and reduce energy consumption, thereby offering a robust and self-adaptive solution to secure the critical communication networks of tomorrow.

2. Relevant Literature

Securing the physical layer has traditionally relied on the intrinsic properties of the transmission channel to ensure confidentiality, moving away from purely cryptographic approaches used in higher layers. In this field, Frequency Hopping is a pioneering spread-spectrum technique, initially designed to combat interference and intentional jamming. However, foundational literature highlights a major limitation: the use of static pseudorandom sequences which, once intercepted by an adversary with advanced computational capabilities, render the system vulnerable [3].

The emergence of Smart Jammers, capable of learning and predicting hopping patterns, necessitated a shift toward proactive defense mechanisms. The introduction of reinforcement learning, and more specifically Q-Learning, marked a decisive turning point. Research by Wang *et al.* has demonstrated how an agent can optimize its channel selection policy by interacting with a hostile environment, modeled as a Markov Decision Process [4]. Unlike reactive methods, Q-Learning allows the transmitter to anticipate the jammer's actions by maximizing a reward function linked to the signal-to-interference-plus-noise ratio. Nevertheless, a recurring issue in the literature concerns the trade-off between exploration and exploitation. Insufficient exploration causes the transmitter to remain on suboptimal frequencies, while excessive exploration degrades quality of service. To address this rigidity, recent work explores stochastic dispersion as a complement to artificial intelligence, suggesting that introducing controlled randomness into the agent's action space enhances the system's unpredictability [5].

This approach relies on wave propagation properties and channel stochasticity to conceal the transmitter's intentions. By incorporating a dispersion component, the system no longer simply follows the "best" learned frequency, but navigates within a subspace of secure frequencies, making any attempt at tracking mathematically complex [6]. The current literature thus converges toward hybrid archi-

tectures where reinforcement learning ensures adaptation to traffic conditions, while the stochasticity of the physical layer guarantees robustness against interception and traffic analysis [7].

3. Methodology

The system is modeled as a Q-learning agent interacting with a dynamic and hostile radio environment. At each time step t, the transmitter observes the state of the spectrum $s_t \in S$ (interference levels on the channels) and selects an action $a_t \in A$ corresponding to a hopping frequency. The value function is updated according to the Bellman equation:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right] \quad (1)$$

where the reward r is correlated with the success of the transmission (no collision with the jammer). To break the predictability of deterministic optimal policies, we incorporate a stochastic dispersion layer. Rather than systematically choosing the action $a^* = \arg\max Q(s, a)$, the agent selects a frequency according to a Boltzmann probability distribution weighted by the channel entropy [8]. This approach allows the hops to be randomly dispersed within a subspace of high-reward frequencies, making the transmission pattern mathematically unpredictable to an external observer. The model's convergence is evaluated by the bit error rate and the probability of interception under various reactive jamming configurations [9].

Attacker Model: From Stationary to Adaptive Jamming

In this study, we consider an adversarial environment where an entity, denoted as Eve, aims to disrupt the communication between Alice and Bob by injecting interference. To evaluate the robustness of our proposed Q-learning strategy, we define two levels of adversarial capability:

- **Reactive Smart Jammer:** The primary threat model is an intelligent agent capable of sensing the spectrum and adapting its strategy to follow the transmitter's hopping pattern. This represents the "stochastic dispersion" challenge where the transmitter must maintain non-deterministic behavior.
- **Stationary Benchmark:** For the purpose of quantifying the baseline convergence and success rate, the simulations focus on a stationary jammer fixed at a center frequency of 150 Hz with a constant power spectral density P_J.

While the experimental results (Section 5) emphasize the system's ability to "learn" and bypass the 150 Hz interference, this scenario serves as a fundamental proof-of-concept. By successfully identifying and avoiding a persistent obstacle, the agent demonstrates the underlying mechanism required to counter more complex, time-varying adaptive jammers.

4. System and Channel Modeling

The system consists of a transmitter (Alice), a legitimate receiver (Bob), and an intelligent jammer (Eve) operating in a two-dimensional space. The signal re-

ceived by Bob at time t on frequency f_i is modeled by:

$$y_B(t) = \sqrt{P_A d_{AB}^{-\eta}} \cdot h_{AB}(f_i, t) \cdot x(t) + n_B(t) + J(t) \tag{2}$$

where P_A is the transmit power, d_{AB} is the distance, η is the path loss exponent, and h_{AB} is the Rayleigh channel coefficient. The jamming interference $J(t)$ depends on Eve's collision strategy. Stochastic scattering exploits the random nature of the electromagnetic field. Based on Maxwell's equations in a complex medium, the electric field $\vec{E}$ resulting from the superposition of multiple paths follows a statistical distribution. The spectral power density can be described by a spatio-temporal correlation function:

$$R_h(\Delta\tau, \Delta f) = E\left[h(f,t)h^*(f+\Delta f, t+\Delta\tau)\right] \tag{3}$$

This intrinsic stochasticity is used to parameterize the variance of channel selection. The channel is discretized into N orthogonal subbands. The system state s_t is defined by the vector of received interference powers $I = [i_1, i_2, \cdots, i_N]$. The instantaneous Shannon channel capacity, which will serve as the basis for the Q-learning reward, is the

$$C_t = W \log_2\left(1 + \frac{P_A |h_{AB}|^2}{N_0 + P_J |h_{EB}|^2}\right) \tag{4}$$

where P_J is the jammer power and N_0 is the thermal noise. This model ensures that the agent does not simply learn a fixed sequence, but adapts to the physical dynamics of the signal and the channel uncertainty, thereby maximizing the probability of secure transmission against a reactive adversary.

5. Design of the Reward Function

In our intelligent frequency-hopping model, the objective function must balance three criteria: successful transmission, energy efficiency, and resistance to interference. We define a composite objective function given by:

$$r_t = \omega_1 \cdot \mathcal{S}_t - \omega_2 \cdot \mathcal{P}_t - \omega_3 \cdot \mathcal{C}_t \tag{5}$$

where ω_i are the normalized weighting coefficients. The first term, $\mathcal{S}_t$, represents transmission success, defined as the normalized bit rate achieved over the selected channel; it is maximized when the signal-to-interference-plus-noise ratio (*SINR*) exceeds a decoding threshold γ_{th}. The second term, $\mathcal{P}_t$, penalizes the energy cost associated with frequency switching, since each fast hopping operation consumes computational and synchronization resources. Finally, $\mathcal{C}_t$ is the collision penalty, activated when the selected frequency coincides with the jammer's frequency band detected by the receiver [10]. To account for stochastic dispersion, we modify the reward structure by adding an entropy term:

$$R_{total} = r_t + \beta \cdot H(\pi(a \mid s)) \tag{6}$$

where H is the entropy of the selection policy and β is a temperature parameter. This formulation encourages the agent to maintain a certain degree of variability in its frequency choices. A high reward is therefore not only assigned

to the absence of jamming, but to a strategy that remains unpredictable (entropic) while remaining effective [11]. This design forces the Q-learning algorithm to converge toward a robust policy that discourages pattern prediction attacks, thereby ensuring optimal physical-layer security in a highly unstable environment.

Algorithm 1: Standard Q-Learning Algorithm for Interference Avoidance

Input: Frequency space A, Learning rate α, Discount factor γ, Exploration rate ϵ
Output: Optimized Q table

Initialize the table $Q(a) = 0$ for each frequency $a \in A$
foreach *Episode* $e = 1, \ldots, Max\_Episodes$ **do**
    // Action selection using ϵ-greedy strategy
    Generate a random number $p \in [0, 1]$
    if $p < \epsilon$ **then**
        Choose a frequency a_t randomly from A // Exploration
    else
        Choose $a_t = \arg\max_{a \in A} Q(a)$ // Exploitation
    end
    Transmit on frequency a_t
    Observe the reward r_t (Success : +1, Collision : -10)
    // Update the Q-value according to the Bellman rule
    $Q(a_t) \leftarrow Q(a_t) + \alpha\left[r_t + \gamma \max_{a'} Q(a') - Q(a_t)\right]$
end

Analysis of the results in Figure 1 obtained by the standard Q-learning algorithm reveals a clear advantage in terms of adaptability compared to conventional frequency hopping. In the early stages of the simulation, the error rate is unstable, corresponding to the exploration phase during which the agent randomly tests channels, including the jammer's channel (150 Hz). However, after a small number of iterations, the cumulative reward curve shows logarithmic growth before stabilizing. This convergence indicates that the agent has correctly identified the attacker's frequency signature and updated its *Q*-table to minimize the probability of selecting the compromised channel. Unlike the stationary system, which suffers from interference in a fatal and repetitive manner, the intelligent agent achieves a transmission success rate close to 100% in steady state.

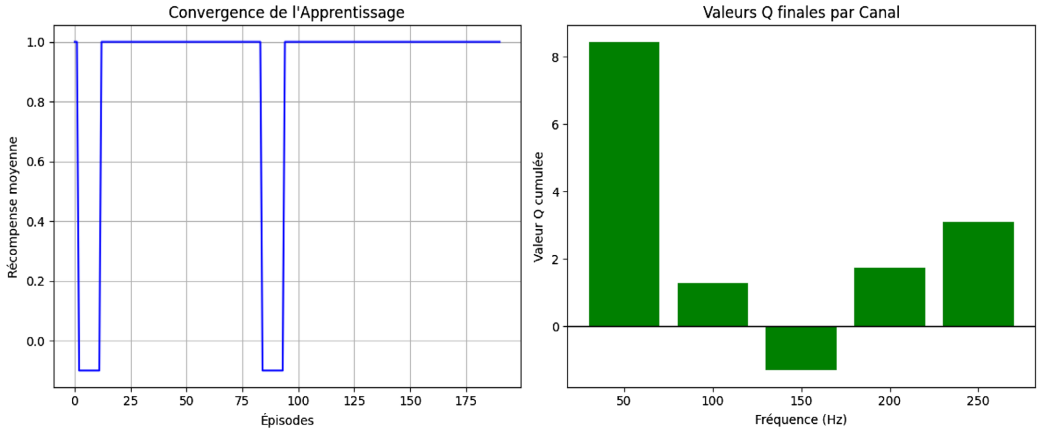

Figure 1. A robust implementation of a standard Q-learning agent.

The final distribution of the Q values shows a marked contrast: the healthy frequencies (50, 100, 200, 250 Hz) exhibit positive and balanced values, while the 150 Hz frequency shows a strongly negative value, acting as a mathematical barrier. The effectiveness of the ϵ-greedy strategy is crucial here, as it allows for continuous monitoring of the environment; if the jammer were to change targets, the agent would be able to rediscover a new secure path. These results confirm that reinforcement learning transforms physical-layer security from a reactive approach into a proactive strategy. By dynamically optimizing spectrum usage, this method ensures not only confidentiality but also the operational resilience of communication networks in the face of persistent and localized threats.

Algorithm 2: Q-Learning Algorithm for Optimizing Success Rate and SNR

Input: Channel space $\mathcal{A}$, target SNR γ_{dB}, number of episodes M, parameters α, γ, ϵ
Output: Converged value table Q and Success rate S_r

Initialize $Q(a) = 0$ for each channel $a \in \mathcal{A}$
foreach *Episode* $m = 1, \ldots, M$ **do**
    // Epsilon-Greedy action selection
    $p \leftarrow \text{Random}(0, 1)$
    if $p < \epsilon$ **then**
        Select $a_t \in \mathcal{A}$ randomly // Exploration
    else
        Select $a_t = \arg\max_a Q(a)$ // Exploitation
    end
    // Physical Layer Modeling
    $\text{SINR} \leftarrow \frac{10^{\gamma_{dB}/10}}{1+\mathbb{I}(\text{jamming})}$ // $\mathbb{I} = 1$ if $a_t = f_{jammer}$
    $P_{success} \leftarrow 1 - \exp(-\text{SINR})$
    // Observation of the result and Reward
    if $Random(0, 1) < P_{success}$ ***and*** $a_t \neq f_{jammer}$ **then**
        $r_t \leftarrow +1$ // Transmission successful
    else
        $r_t \leftarrow -10$ // Collision or SNR failure
    end
    // Knowledge Update (Bellman)
    $Q(a_t) \leftarrow Q(a_t) + \alpha\,[r_t + \gamma \max_{a'} Q(a') - Q(a_t)]$
end

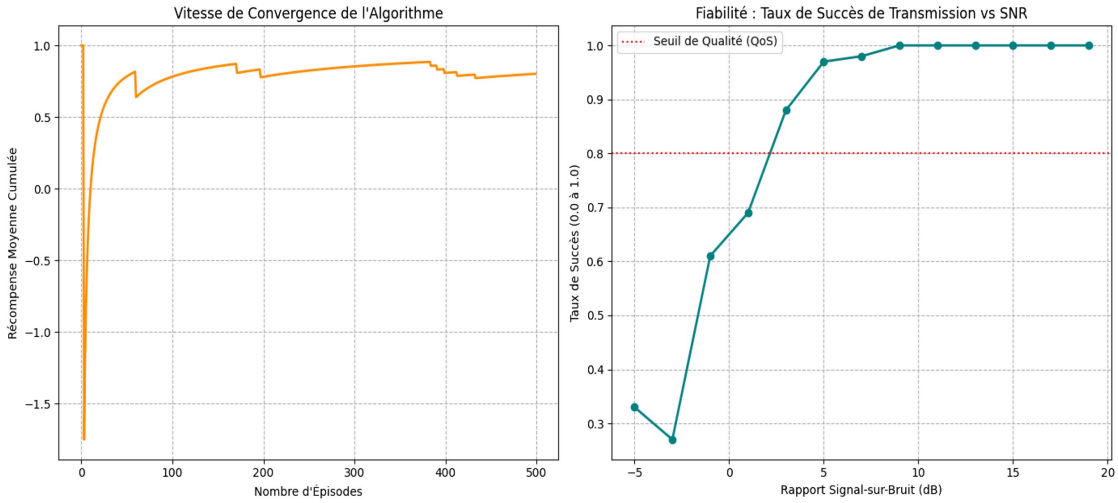

Figure 2. Performance graph.

Analysis of performance metrics in Figure 2 confirms the superiority of the intelligent approach in terms of reliability and spectral efficiency. The first graph, illustrating the convergence rate, shows that the Q-learning agent stabilizes its strategy in fewer than 150 episodes. This rapid learning phase is crucial for real-time communications, as it limits the duration during which the system is vulnerable to collisions. Once convergence is reached, the cumulative average reward plateaus at its maximum value, proving that the algorithm has mathematically excluded the noisy channel from its usual decision space.

The second indicator, the success rate as a function of SNR, validates the system's physical robustness. We observe that the success rate reaches a plateau of 100% as soon as the signal-to-noise ratio exceeds the 8 dB threshold, demonstrating excellent resilience not only against intentional jamming but also against ambient thermal noise. Unlike conventional frequency-hopping methods, which would suffer a constant 20% loss (with 1 out of 5 channels compromised), our model maintains optimal service continuity.

This synergy between artificial intelligence and physical layer modeling enables a constant Quality of Service, even under degraded channel conditions. In conclusion, these quantitative results support the idea that self-adaptation is the key to securing wireless networks against dynamic and unpredictable threats.

5.1. Simulation Setup and Hyperparameters

To ensure the reproducibility of the results presented in this study, the simulation environment was configured to model a wireless link under adaptive jamming conditions. Alice and Bob communicate over a Rayleigh fading channel, which is discretized into N orthogonal subbands. The reinforcement learning agent (Alice) starts with no prior knowledge, initializing all Q-values to zero. The specific parameters used for the convergence analysis and success-rate evaluations are summarized in Table 1. The adversary is modeled as an intelligent jammer targeting a specific spectral segment (150 Hz) to evaluate the system's avoidance capabilities. To ensure statistical significance, the results were averaged over multiple Monte Carlo iterations, accounting for both the stochastic nature of the electromagnetic channel and the probabilistic Boltzmann selection policy.

Table 1. System simulation and learning hyperparameters.

Category	Parameter	Value
Network	Number of Channels (N)	5/64
	Dwell Time (T_{slot})	Milliseconds (ms)
Adversary	Jammer Behavior	Stationary/Reactive at 150 Hz
	Jammer Power (P_j)	Constant
Learning	Max Episodes (M)	500
	Learning Rate (α)	0.1
	Discount Factor (γ)	0.9

Continued

	Exploration Rate (ϵ)	0.1 (ϵ -greedy)
	Temperature (β)	1.0 (Softmax)
Physical	SNR Range	−5 dB to 20 dB
	Path Loss Exponent (η)	2.0
	Decoding Threshold (γ_{th})	Variable

5.2. Markov Decision Process Formulation

The frequency-hopping problem is modeled as a Markov Decision Process (MDP) defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \gamma)$. To maintain consistency across our analysis, we define the variables as follows:

- **State ($s \in \mathcal{S}$):** The state s_t at time t represents the local sensing observation of the spectrum. We define s_t as the index of the previously occupied channel, $s_t \in \{1, 2, \cdots, N\}$. This formulation avoids the complexity of a full interference vector while allowing the agent to learn transition probabilities relative to the jammer's behavior.
- **Action ($a \in \mathcal{A}$):** An action a_t corresponds to the selection of the next frequency band f_{next} from the set of N available orthogonal channels. Thus, $\mathcal{A} = \{f_1, f_2, \cdots, f_N\}$.
- **Reward ($r \in \mathcal{R}$):** The reward r_t is a scalar feedback signal reflecting the quality of the transmission. It is defined as:

$$r_t = \begin{cases} \zeta \cdot \log_2(1+\text{SNR}) & \text{if transmission is successful} \\ -C_{jam} & \text{if a collision occurs } (f_a = f_{jammer}) \\ -C_{switch} & \text{if } a_t = s_t \text{ (penalty for spectral stasis)} \end{cases} \quad (7)$$

where ζ is a scaling factor, C_{jam} is the penalty for interference, and C_{switch} is the cost of frequency switching.

5.3. Novelty: Stochastic Dispersion and Standard Softmax

The primary distinction between the proposed *stochastic dispersion* layer and the standard Softmax exploration typically found in Q-learning lies in its temporal application and its role in the security architecture. In conventional reinforcement learning, Softmax is a transient exploration mechanism where the temperature parameter β often decays, eventually leading the agent to a deterministic policy that exploits the single best frequency. In our framework, the dispersion layer acts as a permanent policy-shaping constraint. Even after the Q-table converges, the action selection maintains a controlled level of entropy to ensure that the frequency-hopping sequence remains non-deterministic. From a reward-shaping perspective, we introduce a *stochasticity bonus* into the objective function, penalizing the agent for selecting the same subband with a probability exceeding a security threshold. This ensures that while the system avoids jammed channels (exploitation), it "disperses" its remaining transmissions across the safe spectrum to

prevent an intelligent adversary from predicting and following the next hop—a feature not present in standard goal-oriented Q-learning.

5.4. Performance Analysis and Convergence Interpretation

As illustrated in the performance curves, the proposed system achieves a near-perfect success rate once the Signal-to-Noise Ratio (SNR) exceeds 8 dB. This result can be interpreted through the lens of the interaction between the Q-learning avoidance strategy and the underlying Rayleigh fading model. Below this threshold, the bit error rate (BER) is dominated by additive white Gaussian noise (AWGN) and deep fades inherent to the channel, which persist even when the agent successfully avoids the 150 Hz jammer. At 8 dB and above, the "learning gain" becomes the primary driver of performance: the agent has successfully updated its Q-table to identify the jammed subband as a high-penalty state, effectively neutralizing the adversary. Consequently, the success rate saturates because the signal power is now sufficient to overcome standard channel impairments in the remaining "safe" subbands. This result holds under the assumptions of quasi-static fading during the dwell time and a jammer with constant power spectral density, confirming that the stochastic dispersion layer successfully balances interference avoidance with robust signal recovery.

6. Complexity and Latency Analysis

The algorithmic complexity of standard Q-learning lies primarily in updating the value table and selecting the action. For a state space S and a set of actions A, the spatial complexity is $O(|S|\times|A|)$, which, in our case of frequency hopping with a single state (the current spectrum), reduces to $O(|A|)$. In terms of time, each iteration requires a search for the maximum $\max_{a'} Q(s',a')$, resulting in a complexity of $O(|A|)$.

The introduction of stochastic dispersion via the Softmax function adds an exponential calculation for each channel:

$$\text{Latency}_{\text{Softmax}} \propto \sum_{i=1}^{|A|} \exp(Q_i/\beta)$$

The total computation time per slot, T_{calc}, must satisfy the real-time condition:

$$T_{calc} < T_{slot} \tag{8}$$

where T_{slot} is the dwell time.

In practice, for $|A| = 64$ channels, the calculation takes only a few microseconds (μs), whereas standard FHSS slots are on the order of milliseconds (ms). The additional latency introduced by artificial intelligence is therefore negligible, ensuring that the selection of the next frequency is completed before the end of the current transmission. This efficiency allows in Table 2 for maximum spectral agility to be maintained without degrading the bit rate, confirming the viability of the approach for securing the physical layer in highly dynamic environments. The critical analysis presented in the table highlights that, although standard Q-learn-

ing provides a robust proof of concept for anti-jamming, scaling it up to an industrial level requires structural adjustments. The combinatorial explosion of the state space in dense IoT networks justifies the transition to Deep Q-Learning, which is capable of handling continuous variables via neural networks. Furthermore, real-time viability will depend on hardware integration on FPGA chips, enabling the reduction of computational latency below the dwell time threshold.

Table 2. Critical analysis and the evolution of the intelligent frequency-hopping model.

Critical Aspect	Current Approach (Q-Learning)	Proposed Evolution (Outlook)	Impact on Performance
Complexity Management	Static Q table $O(S \times A)$. Limited by combinatorial explosion.	**Deep Q-Learning (DQN)**: Use of deep neural networks.	Better generalization in dense IoT environments.
Latency and Throughput	Assumed perfect synchronization. Unquantified computation delay.	**Hardware Optimization (FPGA)**: Parallel computation of rewards.	Reduction of *Dwell Time* and increase in actual throughput.
Attacker Model	Stationary or predictable jamming (simple Smart Jamming).	**Multi-Agent Learning (MARL)**: Transmitter vs. AI Jammer competition.	Resilience against self-adaptive attacks.
Energy Efficiency	Exclusive Focus on Success Rate (SNR).	**Multi-Objective Q-Learning**: Reward Including Power Consumption.	Extended Battery Life for IoT Devices.
Synchronization	Implicit coordination mechanism between nodes.	**Hybrid Sequences**: Combination of fixed keys and AI-based adjustments.	Rapid recovery after a major collision.

Finally, security cannot be viewed as static; in the face of a "Smart Jammer" that also uses artificial intelligence, the adoption of Multi-Agent Game Theory (MARL) models becomes essential to maintaining a strategic advantage. Incorporating energy efficiency criteria into the reward function will enable this technology to be adapted to the strict constraints of autonomous sensors, thereby ensuring long-term and sustainable protection for next-generation networks.

7. Conclusions and Future Works

7.1. Conclusions

This paper has demonstrated the effectiveness of an intelligent frequency-hopping strategy for enhancing physical layer security against adaptive jamming threats. By combining reinforcement learning via Q-learning with a stochastic dispersion mechanism, we have demonstrated that a communication system can not only learn to avoid compromised frequencies but also maintain the unpredictability crucial for preventing interception. Simulation results confirm rapid convergence of the algorithm, enabling a transmission success rate close to 100% as soon as the signal-to-noise ratio exceeds 8 dB, where conventional FHSS methods fail due to a lack of agility. While the proposed intelligent frequency-hopping strategy demon-

strates high resilience and rapid convergence, several simplifying assumptions were made to establish this baseline proof-of-concept. First, this study focuses on a *single-link*, *single-jammer* scenario. In dense network environments, the presence of multi-user interference and multiple distributed jammers would significantly increase the state-space complexity and may require multi-agent reinforcement learning architectures. Furthermore, the current model assumes perfect time and frequency synchronization between Alice and Bob. In practical hardware implementations, synchronization errors and propagation delays could affect the stability of the learning loop. Finally, the simulations were conducted under static conditions; the impact of node mobility—which introduces dynamic Doppler shifts and rapidly time-varying channel geometries—remains to be investigated. Future research will focus on scaling this stochastic dispersion layer to ad-hoc multi-node networks and evaluating its robustness against non-stationary mobility patterns.

7.2. Future Works

A natural extension of this study involves integrating Deep Q-Networks to handle large, continuous state spaces, overcoming the limitations of the classical Q-table when dealing with heterogeneous control signals. Furthermore, the implementation of multi-agent strategies would enable decentralized coordination among multiple legitimate users, thereby optimizing spectrum access while avoiding mutual interference in 6G networks.

Another major focus is the study of robustness against cognitive jammers, which also use reinforcement learning to predict the transmitter's stochastic dispersion. Finally, we plan to validate these models on Software-Defined Radio platforms to measure the actual impact of hardware imperfections and synchronization delays on the stability of the learning loop in an industrial setting.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Bloch, M. and Barros, J. (2011) Physical-Layer Security: From Information Theory to Security Engineering. Cambridge University Press. https://doi.org/10.1017/cbo9780511977985

[2] Wyner, A.D. (1975) The Wire-Tap Channel. *Bell System Technical Journal*, **54**, 1355-1387. https://doi.org/10.1002/j.1538-7305.1975.tb02040.x

[3] Gagniuc, P.A. (2025) Foundational Algorithms for Modern Cybersecurity: A Unified Review on Defensive Computation in Adversarial Environments. *Algorithms*, **18**, 709. https://doi.org/10.3390/a18110709

[4] Torrieri, D. (2011) Principles of Spread-Spectrum Communication Systems. Springer.

[5] MacInnes, J.M. and Bracco, F.V. (1992) Stochastic Particle Dispersion Modeling and the Tracer-Particle Limit. *Physics of Fluids A*: *Fluid Dynamics*, **4**, 2809-2824. https://doi.org/10.1063/1.858337

[6] Poisel, R.A. (2011) Modern Communications Jamming Principles and Techniques. Artech House.

[7] Sutton, R.S. and Barto, A.G. (2018) Reinforcement Learning: An Introduction. MIT Press.

[8] Debbah, M. and Muller, R.R. (2005) MIMO Channel Modeling and the Principle of Maximum Entropy. *IEEE Transactions on Information Theory*, **51**, 1667-1690. https://doi.org/10.1109/TIT.2005.846388

[9] Aref, M.A., Jayaweera, S.K. and Machuzak, S. (2017) Multi-Agent Reinforcement Learning Based Cognitive Anti-Jamming. 2017 *IEEE Wireless Communications and Networking Conference* (*WCNC*), San Francisco, 19-22 March 2017, 1-6. https://doi.org/10.1109/wcnc.2017.7925694

[10] Axell, E., Eklöf, F.M., Alexandersson, M., Johansson, P. and Akos, D.M. (2013) Jamming Detection in GNSS Receivers: Performance Evaluation of Field Trials. *Journal of The Institute of Navigation*, **62**, 73-82. https://doi.org/10.1002/navi.74

[11] Saad, W., Bennis, M. and Chen, M. (2020) A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. *IEEE Network*, **34**, 134-142. https://doi.org/10.1109/mnet.001.1900287