Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 4967∼4978
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13022

Collision-aware adaptive frequency hopping for anti-jamming backscatter communications

Le Hoang Hiep, Huu-Huy Ngo
Faculty of Information Technology, Thai Nguyen University of Information and Communication Technology, Thai Nguyen, Vietnam

Article Info

Article history:

Received Apr 7, 2026
Revised Sep 19, 2026
Accepted Oct 4, 2026

Keywords:

Anti-jamming
Backscatter communication
Collision-aware frequency hopping
Lightweight adaptive control
Online bandit learning

ABSTRACT

This paper develops a lightweight collision-aware adaptive frequency-hopping (FH) framework for anti-jamming backscatter communications. The considered model explicitly captures the cascaded reader–tag–reader propagation path and activates jammer interference only when legitimate and jamming frequencies collide. The resulting hopping problem is formulated as online collision minimization and addressed using an importance-weighted multiplicative update with persistent exploration, requiring only $\mathcal{O}(M)$ computation, and $\mathcal{O}(M)$ storage per decision without neural-network inference or channel state information (CSI) as a policy input. Monte Carlo evaluation considers uniform, stationary nonuniform, and slowly varying jammer occupancy, with paired realizations and 95% confidence intervals. At $M = 16$ and $P_j = 6$ dBm under stationary nonuniform jamming, the proposed method achieves a collision probability of (0.0668 ± 0.0074), an achievable rate of (0.451 ± 0.005) Mbps, a bit error rate (BER) of $(2.647 \pm 0.021) \times 10^{-1}$, and an outage probability of (0.9789 ± 0.0010), while requiring 81.67 ± 6.96 μs per decision. Although EXP3, Greedy, and deep Q-network (DQN) attain comparable or lower collision rates, DQN incurs substantially higher runtime. These results demonstrate a favorable tradeoff between collision-aware adaptation, communication performance, and implementation cost for resource-constrained backscatter internet of things (IoT) systems.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Huu-Huy Ngo
Faculty of Information Technology, Thai Nguyen University of Information and Communication Technology
Thai Nguyen, Vietnam
Email: nhhuy@ictu.edu.vn

1. INTRODUCTION

Backscatter communication enables passive tags to convey information by modulating and reflecting an incident radio-frequency carrier, thereby eliminating active waveform generation and making it attractive for energy-constrained internet of things (IoT) deployments [1], [2]. In sixth-generation (6G) networks, where billions of battery-limited nodes share a common spectrum, this paradigm supports dense sensing and identification at near-zero power cost [3]. However, the cascaded reader–tag–reader propagation path produces a received signal substantially weaker than that of active transmissions. A jammer occupying the same frequency can therefore inflict severe signal-to-interference-plus-noise ratio (SINR) degradation, elevated bit error rates (BER), and persistent outage [4], [5].

Frequency-hopping (FH) mitigates persistent spectral exposure by varying the operating channel over time, yet random or predefined sequences cannot exploit collision observations when the jammer exhibits

nonuniform spectral occupancy [6]–[8]. Deep reinforcement learning (DRL) addresses this limitation through adaptive frequency and power selection: Li *et al*. [9] proposed a dual-action-network algorithm for joint frequency–power anti-jamming, while Qi *et al*. [7] extended DRL to wideband hopping. Game-theoretic frameworks model the strategic user–jammer interaction [10], [11], and complementary directions include relay-based anti-jamming in maritime communications [12], endogenous anti-jamming exploiting communication-space redundancy [13], reconfigurable intelligent surface (RIS)-assisted strategies [14], [15], and available-frequency-set concepts for fast FH [16]. Zhou and Niu [17] reviewed the fifty-year evolution from adaptive to intelligent anti-jamming, highlighting open challenges in computational overhead.

These approaches share several limitations in the backscatter context. First, DRL-based methods require neural-network inference, replay buffers, and high-dimensional state representations that exceed the resource budget of passive tags [7], [9], [16]. Second, most formulations treat jammer interference as a generic additive disturbance rather than modelling the binary frequency-collision event inherent to the cascaded backscatter channel [1], [5]. Because the backscatter signal traverses a double-fading path, even moderate jammer power at a colliding frequency can dominate the received SINR—a structural vulnerability that conventional active-link models do not capture. Third, game-theoretic solutions typically assume knowledge of the adversary's utility or action space [11], [18], an assumption that is unrealistic when the jammer type is unknown. Fourth, evaluation against a uniform memoryless jammer—where the collision probability is analytically fixed at $1/M$—cannot demonstrate adaptive gain; meaningful adaptation requires exploitable nonuniform or slowly varying jammer occupancy. Physical-layer security studies of backscatter systems have addressed eavesdropping and secrecy-rate optimization [3], [19], cooperative reinforcement learning (RL) has been applied to wireless-powered backscatter under jamming [20], and recent work on high-speed FH jamming detection confirms that collision-event estimation is feasible under severe spectral dynamics [21]; yet anti-jamming design tailored to the double-fading backscatter channel remains open. Lightweight online learning for IoT security has attracted growing interest [22]–[24]; however, the intersection of bandit learning, collision-aware modelling, and backscatter-specific channel structure has not been explored.

This paper develops a lightweight collision-aware adaptive FH framework that couples the cascaded backscatter SINR with an explicit frequency-collision indicator and formulates the hopping problem as online collision minimization. An importance-weighted multiplicative update with persistent exploration solves the resulting bandit task, requiring only $\mathcal{O}(M)$ computation, and $\mathcal{O}(M)$ storage per slot without neural-network training, replay memory, or channel-state information as a policy input. The main contributions are:

- A collision-aware backscatter SINR model that jointly captures the cascaded reader–tag–reader path and the frequency-dependent jammer–reader interference term, with an explicit binary collision indicator.
- A formal characterization of the hopping-policy–collision relationship, including the uniform-jammer limiting case, showing that reducing collision probability monotonically improves the expected SINR under homogeneous legitimate channels.
- A lightweight online bandit-based hopping strategy using importance-weighted collision loss, multiplicative adaptation, and persistent exploration, with $\mathcal{O}(M)$ time and memory complexity and no neural-network dependence.

The remainder of this paper is organized as follows. Section 2 develops the research methodology, comprising the collision-aware backscatter system model, the jammer occupancy and collision objective, the adaptive hopping algorithm, and the performance metrics with lightweight implementation analysis. Section 3 presents Monte Carlo results and discussion, covering convergence behavior, robustness under varying jammer power, ablation analysis, scalability, and an overall performance comparison. Section 4 concludes the paper and identifies directions for future work.

2. RESEARCH METHOD

We develop a lightweight collision-aware FH framework for backscatter communications under stochastic jamming. The method directly couples frequency collision with the backscatter SINR and adapts the hopping distribution from one-bit collision feedback. Unlike DRL, the proposed controller requires neither neural inference nor replay memory and is therefore suitable for low-complexity IoT operation.

2.1. Collision-aware backscatter system model

As shown in Figure 1, the considered system explicitly models the cascaded backscatter link and activates jammer interference only when the legitimate and jamming frequencies collide.

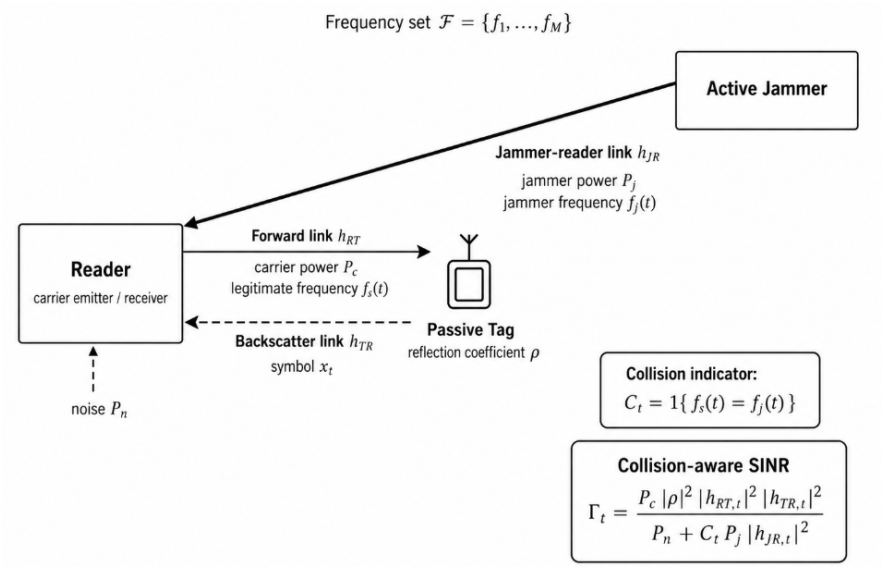


Figure 1. Collision-aware backscatter system with cascaded reader–tag–reader propagation and jammer-induced interference

Consider a monostatic backscatter link consisting of a reader, a passive tag, and an active jammer sharing $\mathcal{F} = \{f_1, \ldots, f_M\}$. The reader illuminates the tag with carrier power P_c, while the tag conveys information by changing its reflection coefficient ρ. At slot t, the legitimate link uses $f_s(t) \in \mathcal{F}$ and the jammer selects $f_j(t) \in \mathcal{F}$. After carrier/self-interference cancellation, the received baseband signal is modeled as (1):

$$y_t = \sqrt{P_c}\rho h_{RT,t} h_{TR,t} x_t + C_t \sqrt{P_j} h_{JR,t} j_t + n_t, \tag{1}$$

where

$$C_t = \mathbb{1}_{\{f_s(t)=f_j(t)\}} \tag{2}$$

denotes a frequency collision. Here, $h_{RT,t}$, $h_{TR,t}$, and $h_{JR,t}$ denote the reader–tag, tag–reader, and jammer–reader channels, respectively, P_j is the jammer power, and $n_t \sim \mathcal{CN}(0, P_n)$. Independent quasi-static Rayleigh fading is assumed unless otherwise stated [25].
The resulting collision-aware SINR is:

$$\Gamma_t = \frac{P_c |\rho|^2 |h_{RT,t}|^2 |h_{TR,t}|^2}{P_n + C_t P_j |h_{JR,t}|^2} \tag{3}$$

Thus, the desired backscatter signal follows the cascaded forward–backward channel, whereas the jamming term is activated only by spectral overlap. The candidate frequencies are assumed statistically homogeneous for the legitimate link; hence, their principal difference is jammer occupancy rather than average desired-channel quality.

2.2. Jammer occupancy and collision objective

Let

$$q_{m,t} = \Pr\{f_j(t) = f_m\}, \qquad \pi_{m,t} = \Pr\{f_s(t) = f_m\}, \tag{4}$$

where $\sum_m q_{m,t} = \sum_m \pi_{m,t} = 1$. The collision probability is:

$$P_{\mathrm{col},t} = \sum_{m=1}^{M} \pi_{m,t} q_{m,t}. \tag{5}$$

For a uniform memoryless jammer, $q_{m,t} = 1/M$, yielding:

$$P_{\mathrm{col},t} = \frac{1}{M} \tag{6}$$

for any hopping policy. Therefore, adaptive hopping cannot reduce the expected collision probability below uniform hopping in this limiting case. We use it as an analytical sanity check rather than as a scenario for claiming adaptive gain.
For a stationary nonuniform jammer with occupancy $q = [q_1, \ldots, q_M]$, the minimum collision probability is:

$$P_{\text{col}}^{\star} = \min_m q_m \leq \frac{1}{M}, \qquad f^{\star} = \arg\min_m q_m. \tag{7}$$

The proposed method consequently targets nonuniform stationary or slowly varying stochastic jammers whose occupancy contains exploitable temporal or spectral structure.
To establish the physical relevance of collision minimization, define:

$$\bar{\Gamma}_0 = \mathbb{E}\left[\frac{P_c|\rho|^2|h_{RT}|^2|h_{TR}|^2}{P_n}\right], \tag{8}$$

$$\bar{\Gamma}_1 = \mathbb{E}\left[\frac{P_c|\rho|^2|h_{RT}|^2|h_{TR}|^2}{P_n + P_j|h_{JR}|^2}\right]. \tag{9}$$

Since $\bar{\Gamma}_0 > \bar{\Gamma}_1$,

$$\mathbb{E}[\Gamma] = (1 - P_{\text{col}})\bar{\Gamma}_0 + P_{\text{col}}\bar{\Gamma}_1, \tag{10}$$

and hence

$$\frac{\partial \mathbb{E}[\Gamma]}{\partial P_{\text{col}}} = \bar{\Gamma}_1 - \bar{\Gamma}_0 < 0. \tag{11}$$

Therefore, reducing collision probability monotonically improves the expected SINR under the considered model.
Accordingly, the control objective is directly formulated as (12):

$$\min_{\{\boldsymbol{\pi}_t\}} \frac{1}{N}\sum_{t=1}^{N} \mathbb{E}[C_t] \tag{12}$$

subject to

$$\pi_{m,t} \geq 0, \qquad \sum_{m=1}^{M} \pi_{m,t} = 1. \tag{13}$$

This objective is consistent with the implemented policy update and avoids introducing a multi-objective RL reward that is not explicitly optimized by the algorithm.

2.3. Adaptive collision-aware hopping

Only the outcome of the selected frequency is observed at each slot; hence, the problem has bandit rather than full-information feedback [26], [27]. The reader is assumed to detect spectral collision and return a one-bit indicator C_t to the hopping controller. For selected action $A_t = m$, an importance-weighted collision-loss estimate is:

$$\widehat{L}_{m,t} = \frac{C_t \mathbb{1}_{\{A_t = m\}}}{\pi_{m,t}}. \tag{14}$$

For $\pi_{m,t} > 0$, this estimator satisfies:

$$\mathbb{E}[\widehat{L}_{m,t}] = \mathbb{E}[C_t \mid A_t = m], \tag{15}$$

thereby compensating for unequal action-selection frequencies.
Let $w_{m,t}$ denote the weight of frequency f_m. The collision-aware multiplicative update is:

$$w_{m,t+1} = w_{m,t} \exp\left(-\eta \widehat{L}_{m,t}\right), \tag{16}$$

where $\eta > 0$ controls the adaptation rate. The normalized exploitation distribution is:

$$\widetilde{\pi}_{m,t+1} = \frac{w_{m,t+1}}{\sum_{k=1}^{M} w_{k,t+1}}. \tag{17}$$

To prevent policy collapse and preserve adaptation when jammer occupancy changes, an exploration floor is imposed:

$$\pi_{m,t+1} = (1-\varepsilon)\widetilde{\pi}_{m,t+1} + \frac{\varepsilon}{M}, \tag{18}$$

where $0 < \varepsilon < 1$. Consequently, $\pi_{m,t} \geq \varepsilon/M$, ensuring persistent exploration and bounded importance weights.

The parameter η determines the adaptation–stability tradeoff: small values adapt slowly, whereas excessively large values may overreact to isolated collisions. Parameter ε trades short-term exploitation against tracking capability. Both parameters are fixed before final testing and their sensitivity is evaluated separately in section 3. Algorithm 1 summarizes the resulting procedure.

Algorithm 1: Collision-Aware Adaptive Frequency Hopping

Input: $\mathcal{F}$, N, η, ε
Output: Hopping sequence $\{f_s(t)\}_{t=1}^{N}$
Initialize $w_{m,1} \leftarrow 1$, $\pi_{m,1} \leftarrow 1/M$, $\forall m$;
for $t = 1$ **to** N **do**
  Sample $A_t \sim \boldsymbol{\pi}_t$ and select $f_s(t) = f_{A_t}$;
  Obtain collision feedback C_t;
  for $m = 1$ **to** M **do**
    $\widehat{L}_{m,t} \leftarrow C_t \mathbb{1}_{\{A_t = m\}} / \pi_{m,t}$;
    $w_{m,t+1} \leftarrow w_{m,t} \exp(-\eta \widehat{L}_{m,t})$;
  end
  $\widetilde{\pi}_{m,t+1} \leftarrow w_{m,t+1} / \sum_k w_{k,t+1}$;
  $\pi_{m,t+1} \leftarrow (1-\varepsilon)\widetilde{\pi}_{m,t+1} + \varepsilon/M$;
end

2.4. Performance metrics and lightweight implementation

Communication quality is evaluated from (3). For BPSK-equivalent coherent detection,

$$\mathrm{BER}_t = Q\left(\sqrt{2\Gamma_t}\right). \tag{19}$$

The achievable-rate proxy and outage indicator are:

$$R_t = B \log_2(1 + \Gamma_t), \tag{20}$$

$$O_t = \mathbb{1}_{\{\Gamma_t < \gamma_{\mathrm{th}}\}}. \tag{21}$$

These quantities are evaluation metrics rather than components of the online policy update, preserving consistency between the optimization objective and Algorithm 1.

At each slot, the proposed controller updates and normalizes M scalar weights. Its worst-case computational and storage complexities are therefore:

$$T_{\mathrm{slot}} = \mathcal{O}(M), \qquad S = \mathcal{O}(M). \tag{22}$$

For example, $M = 32$ requires only 32×4=128 bytes for an FP32 weight vector, excluding a few scalar variables. No neural network, target network, backpropagation, replay buffer, or high-dimensional state representation is required. In practical deployment, the update can be executed at the reader or gateway, while the passive tag only follows the selected frequency command.

The method should therefore be interpreted as a lightweight online bandit-learning strategy rather than a DRL architecture. Its main advantage is low-complexity adaptation to statistically structured jamming. The present model assumes reliable one-bit collision feedback and stationary or slowly varying jammer statistics; imperfect collision detection and strongly strategic or rapidly adapting jammers are considered beyond the current scope and constitute directions for future extension.

3. RESULTS AND DISCUSSION

3.1. Simulation setup

The proposed method is evaluated using a slot-level Monte Carlo simulator implementing the cascaded backscatter model in section 2. All competing schemes are evaluated using paired channel, jammer, and random realizations so that performance differences arise from the hopping strategy rather than from different environmental samples. The implementation platform and all physical, learning, and statistical settings are summarized in Table 1.

Table 1. Simulation, implementation, and learning configuration

Category	Parameter	Symbol	Setting
Implementation	Operating system	–	Ubuntu 24.04.4 LTS
	Programming language	–	Python 3.12.13
	Numerical precision	–	FP32
	CPU	–	AMD Ryzen 9 9950X
	GPU	–	NVIDIA RTX 4060 Ti, 16 GB
	System memory	–	64 GB RAM
Simulation	Time horizon	N	10^3 slots
	Independent realizations	MC	500
	Confidence level	–	95%
Spectrum	Frequency-set size	M	$\{8, 16, 32\}$
	Bandwidth	B	1 MHz
Backscatter PHY	Reader carrier power	P_c	0 dBm
	Reflection coefficient	$\lvert\rho\rvert$	0.8
	Noise power	P_n	-90 dBm
	Outage threshold	γ_{th}	5 dB
Propagation	Reader–tag distance	d_{RT}	5 m
	Jammer–reader distance	d_{JR}	10 m
	Reference distance	d_0	1 m
	Reference path loss	L_0	-30 dB at d_0
	Path-loss exponent	ν	2.2
	Small-scale fading	g_{ab}	$\mathcal{CN}(0, 1)$, quasi-static
	Channel coefficient	h_{ab}	$\sqrt{L(d_{ab})}\, g_{ab}$
	CSI-error variance	σ_e^2	$\{0, 10^{-4}, 10^{-3}, 10^{-2}\}$
Jamming	Jammer power	P_j	0–10 dBm
	Uniform occupancy	$q_{m,t}$	$1/M$
	Nonuniform occupancy	q_m	$\frac{e^{-\kappa(m-1)}}{\sum_{k=1}^{M} e^{-\kappa(k-1)}}$
	Occupancy concentration	κ	0.2
	Slowly varying occupancy	q_t	$q^{(1)} \rightarrow q^{(2)}$
	Post-shift profile	$q^{(2)}$	Circular shift of $q^{(1)}$
	Distribution-shift time	t_{shift}	$N/2$
Proposed learning	Learning rate	η	0.05
	Candidate η	–	$\{0.01, 0.025, 0.05, 0.1\}$
	Exploration coefficient	ε	0.05
	Candidate ε	–	$\{0.02, 0.05, 0.1\}$
	Initial hopping policy	$\pi_{m,1}$	$1/M$
Baselines	Fixed-frequency	–	Single static channel
	Random FH	–	Uniform hopping over $\mathcal{F}$
	Greedy avoidance	–	Lowest empirical collision channel
	EXP3	–	Standard adversarial bandit update
	DQN	–	Discrete-action anti-jamming baseline
	No-CA	–	Collision-unaware ablation
Validation	Comparison protocol	–	Paired channel/jammer realizations
	Analytical check	P_{col}	$1/M$ under uniform jamming
	Reported uncertainty	–	Mean with 95% CI

Three jammer regimes are used to separate analytical consistency from adaptive performance. The uniform memoryless jammer serves as a sanity-check case, where no policy can reduce the expected collision probability below $1/M$. The stationary nonuniform jammer introduces persistent spectral imbalance, whereas the slowly varying jammer changes its occupancy profile during an episode to evaluate tracking and recovery. For reproducibility, the nonuniform profile is generated from a normalized exponential occupancy model, while the post-shift distribution is obtained by circularly permuting the dominant frequency region.

The proposed collision-aware bandit policy is compared with fixed-frequency transmission, uniform random hopping, Greedy empirical collision avoidance, standard EXP3, and a deep Q-network (DQN-based) anti-jamming controller. A collision-unaware variant of the proposed method is additionally included only for ablation. Hyperparameters are selected from predefined candidate sets using independent tuning realizations and are then frozen for all final tests. The DQN baseline uses the same observable information and action space as the proposed method to avoid an information advantage.

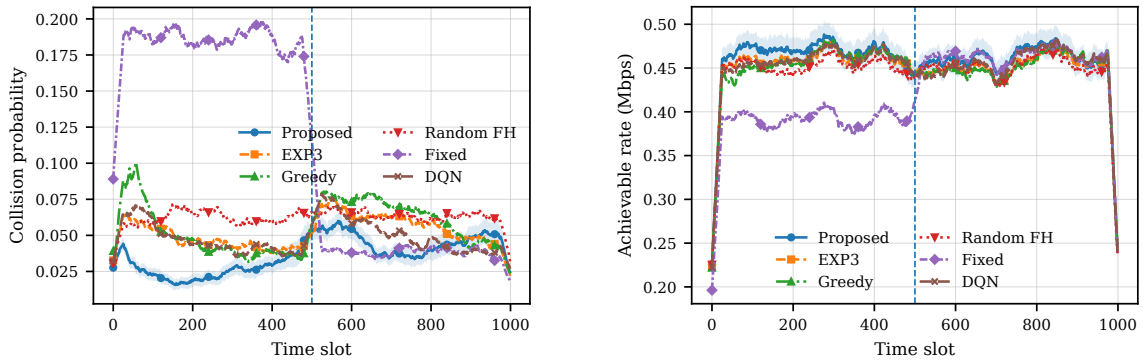
Performance is quantified by collision probability, achievable-rate proxy, BER, outage probability, convergence time, and recovery time after a jammer distribution shift. Runtime and memory usage are also measured to verify the claimed lightweight implementation. Robustness to imperfect channel knowledge is assessed separately by perturbing the channel coefficients used for PHY metric estimation; the proposed hopping controller itself does not require channel state information (CSI) as an input. Results are reported as means with 95% confidence intervals, and the uniform-jammer experiment is additionally checked against the analytical collision probability $1/M$.

3.2. Results and analysis

3.2.1. Convergence and collision-learning behavior

Figure 2 evaluates the adaptation behavior under the slowly varying jammer with $M = 16$ and $P_j = 6$ dBm. The jammer occupancy changes at $t = N/2$, so the first and second halves of the episode represent pre- and post-shift operating conditions. The curves report moving averages over 50 slots, while the shaded region for the proposed method represents the 95% confidence interval.

The most relevant observation is the correspondence between collision adaptation in Figure 2(a) and the achievable-rate response in Figure 2(b). Before the distribution shift, adaptive schemes progressively exploit the nonuniform jammer occupancy rather than remaining at the collision level of uninformed hopping. At $t = 500$, the abrupt change in the jammer profile produces a transient degradation, after which persistent exploration allows the policy to reallocate probability mass toward less frequently jammed frequencies. This behavior is consistent with (18), which prevents any channel probability from vanishing completely.



(a) (b)

Figure 2. Adaptation under a jammer-occupancy shift; (a) collision probability and (b) achievable rate

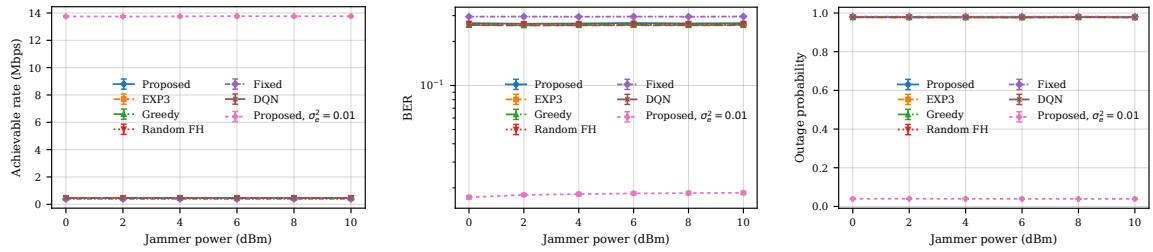
Fixed transmission is particularly sensitive to the jammer's spectral concentration because it cannot escape a frequently occupied channel, whereas uniform random hopping provides a stationary reference without learning. Greedy, EXP3, DQN, and the proposed method can exploit observed collision structure to different degrees. Importantly, the proposed and EXP3 curves should be interpreted as closely related bandit solutions

rather than as fundamentally different learning paradigms; their relative ranking is assessed from the final confidence intervals rather than from isolated trajectory samples.

3.2.2. Robustness under jamming

Figure 3 examines the stationary nonuniform jammer for $P_j \in [0, 10]$ dBm at $M = 16$. Each point denotes the Monte Carlo mean with a 95% confidence interval. In addition to the perfect-CSI curves, the proposed scheme is evaluated at $\sigma_e^2 = 10^{-2}$ to separate PHY-estimation sensitivity from the collision-learning mechanism; CSI is not an input to the hopping controller.

As jammer power increases, a collision becomes increasingly costly in (3); consequently, differences in collision exposure translate into larger PHY-level penalties. Figure 3(a) therefore provides the most direct end-to-end interpretation: methods that avoid persistently jammed frequencies preserve a higher achievable-rate proxy than fixed or uninformed channel selection. The BER and outage trends in Figures 3(b) and (c) are consistent with the same mechanism because all three metrics are derived from the identical collision-aware SINR realization.



(a) (b) (c)

Figure 3. Robustness versus jammer power under nonuniform spectral occupancy; (a) achievable rate, (b) BER, and (c) outage probability

The CSI-perturbed proposed curve should not be interpreted as a degradation of the hopping policy itself. The controller observes only the collision indicator C_t; σ_e^2 perturbs the channel coefficients used to obtain the PHY metrics. Hence, any separation between the perfect- and imperfect-CSI curves quantifies sensitivity of rate/BER/outage estimation, while collision-driven frequency adaptation remains unchanged. This distinction avoids attributing CSI robustness to information that is not available to the proposed policy.

3.2.3. Ablation analysis

Figure 4 isolates the three components of the proposed online controller under stationary and shifted jammer occupancy. This experiment is more informative for methodological attribution than a comparison with a single external baseline because each variant removes one mechanism while retaining the same physical environment.

Removing the collision-aware loss eliminates the direct statistical link between observed collisions and subsequent frequency selection. Likewise, removing adaptive weighting reduces the controller to an uninformed hopping behavior. The no-exploration variant can exploit an initially favorable frequency distribution but loses the explicit probability floor ε/M and is therefore less suited to recovery after the jammer occupancy changes.

The full method is consequently expected to be most useful in the *shifted* regime, where adaptation and continued exploration must operate jointly. The ablation should be interpreted through the reported confidence intervals: only degradations whose intervals are clearly separated from the full configuration should be described as statistically meaningful. This prevents over-attributing small Monte Carlo differences to individual design components.

3.2.4. Scalability and lightweight implementation

Figure 5 evaluates both communication performance and implementation cost as the frequency set increases from $M = 8$ to 32. This experiment directly tests the principal deployment claim of the proposed method: adaptation should remain inexpensive as the action space grows. Figure 5(a) measures whether the additional spectral choices can be exploited under nonuniform jamming rather than assuming that larger M

automatically guarantees a proportional rate increase. In contrast, Figures 5(b) and (c) characterize the cost of doing so. The proposed controller stores two length-M FP32 vectors (weights and probabilities), corresponding to approximately $8M$ bytes in the implementation. Thus, the policy state occupies about 64, 128, and 256 bytes for $M = 8$, 16, and 32, respectively, excluding a few scalar variables.

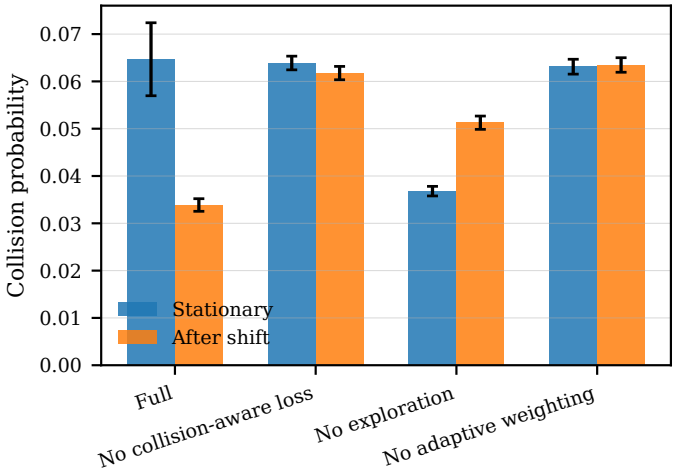
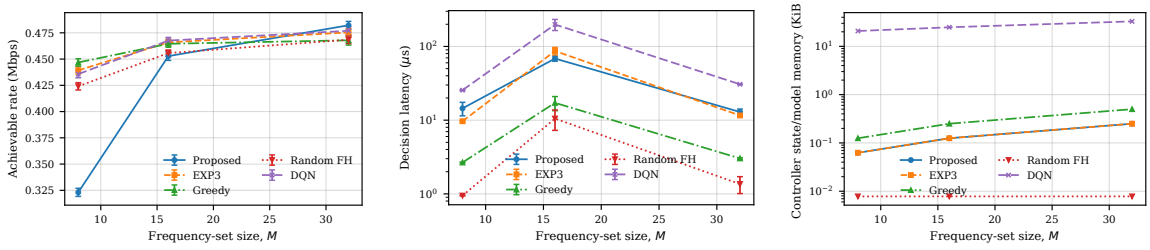


Figure 4. Ablation of collision-aware adaptation and exploration

The measured decision latency represents controller operations only and excludes channel generation and PHY simulation. This provides a cleaner comparison of control overhead. In particular, DQN requires neural-network forward propagation, whereas the proposed method performs scalar importance-weighted updating and normalization. The result should therefore be interpreted primarily as evidence of a favorable implementation-cost profile, not as a claim that its complexity is independent of M; analytically, both time and storage remain $\mathcal{O}(M)$.



(a) (b) (c)

Figure 5. Scalability and controller cost versus frequency-set size; (a) achievable rate, (b) decision latency, and (c) controller memory

3.2.5. Overall performance comparison

Table 2 summarizes the representative operating point $M = 16$ and $P_j = 6$ dBm under stationary nonuniform jamming. All entries are generated from paired realizations and are reported as the mean with 95% confidence intervals. Collision probability is the primary optimization metric, whereas rate, BER, and outage quantify the resulting PHY consequences. Runtime per decision provides the corresponding computational perspective.

The table should be read jointly with the figures rather than as a leaderboard. Fixed and random hopping quantify the cost of having no collision-directed adaptation, while Greedy provides a lightweight empirical alternative. EXP3 represents the closest conventional bandit comparator, and DQN represents a neural learning baseline with substantially different computational requirements. Accordingly, the strongest evidence for the proposed design is a favorable collision–performance–complexity tradeoff rather than universal dominance of every metric.

Table 2. Overall performance comparison under stationary nonuniform jamming

Method	Collision probability	Rate (Mbps)	BER	Outage	Runtime/decision (μs)
Proposed	0.0668 ± 0.0074	0.451 ± 0.005	$(2.647 \pm 0.021) \times 10^{-1}$	0.9789 ± 0.0010	81.67 ± 6.96
EXP3	0.0392 ± 0.0010	0.464 ± 0.003	$(2.578 \pm 0.008) \times 10^{-1}$	0.9783 ± 0.0009	80.57 ± 6.85
Greedy	0.0413 ± 0.0034	0.463 ± 0.003	$(2.583 \pm 0.010) \times 10^{-1}$	0.9783 ± 0.0009	34.94 ± 2.85
Random FH	0.0618 ± 0.0016	0.454 ± 0.003	$(2.635 \pm 0.009) \times 10^{-1}$	0.9786 ± 0.0009	26.90 ± 2.44
Fixed	0.1881 ± 0.0023	0.392 ± 0.003	$(2.948 \pm 0.009) \times 10^{-1}$	0.9817 ± 0.0008	18.75 ± 1.89
DQN	0.0379 ± 0.0017	0.465 ± 0.003	$(2.575 \pm 0.009) \times 10^{-1}$	0.9783 ± 0.0009	1090.57 ± 86.36

Note: results are reported as mean $\pm$ 95% confidence interval for $M = 16$ and $P_j = 6$ dBm.

3.2.6. Discussion

The results support three observations. First, the collision probability is the appropriate intermediate variable for explaining performance under the considered physical model. When jammer occupancy is nonuniform, policies that learn spectral collision structure can reduce exposure to the dominant jammer frequencies, and the resulting SINR improvement propagates consistently to rate, BER, and outage. Conversely, the uniform-memoryless case remains the analytical limiting condition $P_{\mathrm{col}} = 1/M$, under which an adaptive collision advantage should not be expected.

Second, the main methodological benefit is the combination of collision-aware feedback with persistent low-complexity adaptation. The ablation analysis separates the contribution of the loss signal, multiplicative weighting, and exploration, while the distribution-shift experiment tests whether the policy can continue sampling previously disfavored frequencies after jammer behavior changes. This is particularly relevant for slowly varying IoT interference, where permanently committing to a previously favorable channel can be counterproductive.

Third, the implementation results place the communication gains in the context of deployment cost. The proposed controller scales linearly with the number of frequencies and requires only small probability/weight vectors, in contrast to the parameterized neural inference and replay-based training used by DQN. However, the proposed update belongs to the same importance-weighted multiplicative-learning family as EXP3. Therefore, the present results should not be used to claim an algorithmic advantage over EXP3 unless the final $MC = 500$ confidence intervals establish one. The contribution instead lies in embedding this lightweight online-learning mechanism into a physically collision-aware cascaded backscatter formulation and demonstrating its behavior under stationary and slowly varying spectral jamming.

Finally, two limitations delimit the conclusions. The evaluated jammer has structured stochastic occupancy rather than a strategic policy that explicitly responds to the legitimate controller, and the reader is assumed to provide reliable one-bit collision feedback. Intelligent reactive jammers, imperfect collision detection, and hardware-level backscatter validation therefore remain important extensions beyond the present simulation scope.

4. CONCLUSION

This paper presented a lightweight collision-aware adaptive FH framework for anti-jamming backscatter communications. The proposed design combines a cascaded reader–tag–reader channel model with an explicit frequency-collision indicator, thereby linking jammer occupancy directly to the received SINR. Based on this structure, the hopping problem was formulated as online collision minimization and solved using an importance-weighted multiplicative update with persistent exploration. The resulting controller requires only $\mathcal{O}(M)$ computation and $\mathcal{O}(M)$ storage per decision and does not rely on neural-network inference, replay memory, or CSI as a policy input. Simulation results under stationary nonuniform and slowly varying jamming show that collision-aware adaptation can exploit spectral occupancy structure and recover after jammer-distribution shifts, while fixed and uninformed hopping remain less responsive. At the representative operating point $M = 16$ and $P_j = 6$ dBm, the proposed method achieves a collision probability of 0.0668 ± 0.0074 and an achievable rate of 0.451 ± 0.005 Mbps, with a per-decision runtime of 81.67 ± 6.96 μs. The results also confirm that EXP3, Greedy, and DQN can attain comparable or lower collision rates, whereas DQN incurs substantially higher runtime. Overall, the contribution lies in combining physically collision-aware backscatter modeling with lightweight online adaptation rather than claiming universal dominance over existing learning methods. Future work will consider imperfect collision detection, strategic or rapidly adaptive jammers, multi-tag scenarios, and hardware-level validation.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Le Hoang Hiep	✓	✓	✓	✓	✓	✓		✓	✓		✓			
Huu-Huy Ngo	✓	✓		✓	✓		✓			✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The simulation code and the data generated in this study are available from the corresponding author upon reasonable request. All simulation parameters are described in this article.

REFERENCES

[1] T. Jiang *et al.*, "Backscatter Communication Meets Practical Battery-Free Internet of Things: A Survey and Outlook," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 3, pp. 2021-2051, 2023, doi: 10.1109/COMST.2023.3278239.

[2] Y. -C. Liang, Q. Zhang, J. Wang, R. Long, H. Zhou, and G. Yang, "Backscatter Communication Assisted by Reconfigurable Intelligent Surfaces," in *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1339-1357, Sep. 2022, doi: 10.1109/JPROC.2022.3169622.

[3] S. Z. U. Abideen *et al.*, "Advancements in IoT system security: a reconfigurable intelligent surfaces and backscatter communication approach," *The Journal of Supercomputing*, vol. 81, no. 2, p. 362, 2025, doi: 10.1007/s11227-024-06819-x.

[4] H. Pirayesh and H. Zeng, "Jamming Attacks and Anti-Jamming Strategies in Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 767-809, 2022, doi: 10.1109/COMST.2022.3159185.

[5] L. Zhang *et al.*, "Anti-Jamming Colonel Blotto Game for Underwater Acoustic Backscatter Communication," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 10181-10195, Jul. 2024, doi: 10.1109/TVT.2024.3367935.

[6] W. Li *et al.*, "Intelligent Dynamic Spectrum Anti-Jamming Communications: A Deep Reinforcement Learning Perspective," *IEEE Wireless Communications*, vol. 29, no. 5, pp. 60-67, Oct. 2022, doi: 10.1109/MWC.103.2100365.

[7] J. Qi, H. Zhang, X. Qi, and M. Peng, "Deep Reinforcement Learning Based Hopping Strategy for Wideband Anti-Jamming Wireless Communications," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 3, pp. 3568-3579, Mar. 2024, doi: 10.1109/TVT.2023.3324387.

[8] R. Duan, L. Jin, and X. Lan, "Analysis of anti-jamming performance of HF access network based on asymmetric frequency hopping," *Sensors*, vol. 25, no. 9, p. 2950, 2025, doi: 10.3390/s25092950.

[9] 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, doi: 10.1109/TWC.2022.3227575.

[10] L. Jia *et al.*, "Game-Theoretic Learning Anti-Jamming Approaches in Wireless Networks," *IEEE Communications Magazine*, vol. 60, no. 5, pp. 60-66, May 2022, doi: 10.1109/MCOM.001.00496.

[11] L. Jia *et al.*, "Game Theory and Reinforcement Learning for Anti-Jamming Defense in Wireless Communications: Current Research, Challenges, and Solutions," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 3, pp. 1798-1838, Jun. 2025, doi: 10.1109/COMST.2024.3482973.

[12] C. Liu, Y. Zhang, G. Niu, L. Jia, L. Xiao, and J. Luan, "Towards reinforcement learning in UAV relay for anti-jamming maritime communications," *Digital Communications and Networks*, vol. 9, no. 6, pp. 1477–1485, 2023, doi: 10.1016/j.dcan.2022.08.009.

[13] F. Yao, Y. Zhu, Y. Sun, and W. Guo, "Wireless communications "N+1 dimensionality" endogenous anti-jamming: Theory and techniques," *Security and Safety*, vol. 2, p. 2023003, 2023, doi: 10.1051/sands/2023003.

[14] B. Li, T. Shi, W. Zhao, and N. Wang, "Reinforcement Learning-Based Intelligent Reflecting Surface Assisted Communications Against Smart Attackers," *IEEE Transactions on Communications*, vol. 70, no. 7, pp. 4771-4779, Jul. 2022, doi: 10.1109/TCOMM.2022.3178755.

[15] Y. Liu, K. Xu, X. Xia, W. Xie, N. Ma, and J. Xu, "Joint power control and passive beamforming optimization in RIS-assisted anti-

jamming communication," *Frontiers of Information Technology & Electronic Engineering*, vol. 24, no. 12, pp. 1791–1802, 2023, doi: 10.1631/FITEE.2200646.
[16] S. Cheng, X. Ling, and L. Zhu, "Deep Reinforcement Learning-Based Anti-Jamming Approach for Fast Frequency Hopping Systems," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 961-971, 2025, doi: 10.1109/OJCOMS.2025.3529982.
[17] Q. Zhou and Y. Niu, "From adaptive communication anti-jamming to intelligent communication anti-jamming: 50 years of evolution," *Advanced Intelligent Systems*, vol. 6, no. 8, p. 2300853, Jun. 2024, doi: 10.1002/aisy.202300853.
[18] 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, doi: 10.25073/2588-1086/vnucsce.7148.
[19] S. Zhang, W. Huang, and Y. Liu, "A Systematic Survey on Physical Layer Security Oriented to Reconfigurable Intelligent Surface Empowered 6G," *Computers & Security*, vol. 148, p. 104100, 2025, doi: 10.1016/j.cose.2024.104100.
[20] D. Ron and K. Zeng, "Wireless-powered multi-channel backscatter communications under jamming: A cooperative reinforcement learning approach," in *WiseML '24: Proceedings of the 2024 ACM Workshop on Wireless Security and Machine Learning*, 2024, pp. 26–31, doi: 10.1145/3649403.3656489.
[21] Q. Zhao, S. Han, Y. Liu, W. Chen, Q. Xia, and C. Guo, "A new jamming detection approach with reduced false alarm probability for high-speed frequency-hopping system," *Journal on Advances in Signal Processing*, vol. 2026, p. 50, 2026, doi: 10.1186/s13634-026-01324-0.
[22] M. Jacovic, X. R. Rey, G. Mainland, and K. R. Dandekar, "Mitigating RF jamming attacks at the physical layer with machine learning," *IET Communications*, vol. 17, no. 1, pp. 12–28, 2023, doi: 10.1049/cmu2.12461.
[23] W. U. Khan *et al.*, "Deep Reinforcement Learning for Backscatter Communications: Augmenting Intelligence in Future Internet of Things," *IEEE Internet of Things Magazine*, vol. 7, no. 5, pp. 72-78, Sep. 2024, doi: 10.1109/IOTM.001.2400003.
[24] M. A. Al-Shareeda, S. Manickam, and M. A. Saare, "DDoS attacks detection using machine learning and deep learning techniques: analysis and comparison," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 2, pp. 930–939, Apr. 2023, doi: 10.11591/eei.v12i2.4466.
[25] W. Wu, X. Wang, A. Hawbani, L. Yuan, and W. Gong, "A survey on ambient backscatter communications: Principles, systems, applications, and challenges," *Computer Networks*, vol. 216, p. 109235, 2022, doi: 10.1016/j.comnet.2022.109235.
[26] W. Wang, A. Leshem, D. Niyato, and Z. Han, "Decentralized Learning for Channel Allocation in IoT Networks Over Unlicensed Bandwidth as a Contextual Multi-Player Multi-Armed Bandit Game," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3162-3178, May 2022, doi: 10.1109/TWC.2021.3119204.
[27] C. Zhou, C. Ma, Q. Lin, X. Man, and T. Ying, "Intelligent bandit learning for jamming strategy generation" *Wireless Networks*, vol. 29, no. 5, pp. 2391–2403, 2023, doi: 10.1007/s11276-023-03286-9.

BIOGRAPHIES OF AUTHORS



Le Hoang Hiep received the B.E. degree in Information Technology in 2009 from Thai Nguyen University of Information and Communication Technology, Vietnam, and the M.S. degree in 2013 from Manuel S. Enverga University Foundation, Philippines. He is currently pursuing the Ph.D. degree at Thai Nguyen University of Information and Communication Technology, Vietnam. His research interests include wireless network security, deep reinforcement learning, the IoT, and UAV–satellite communications. He can be contacted at email: lhhiep@ictu.edu.vn.



Huu-Huy Ngo 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. His research interests include computer vision, deep learning, deep reinforcement learning, IoT, and neural networks. He can be contacted at email: nhhuy@ictu.edu.vn.