

## Article

# TLOA: A Power-Adaptive Algorithm Based on Air–Ground Cooperative Jamming

Wenpeng Wu <sup>1,2</sup>, Zhenhua Wei <sup>1,3,\*</sup>, Haiyang You <sup>1</sup>, Zhaoguang Zhang <sup>1</sup>, Chenxi Li <sup>1</sup>, Jianwei Zhan <sup>1</sup> and Shan Zhao <sup>3</sup>

<sup>1</sup> Rocket Force University of Engineering, Xi'an 710025, China

<sup>2</sup> PLA 967XX Unit, Henan, China

<sup>3</sup> School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China

\* Correspondence: wzh016001@aliyun.com

## Abstract

Air–ground joint jamming enables three-dimensional, distributed jamming configurations, making it effective against air–ground communication networks with complex, dynamically adjustable links. Once the jamming layout is fixed, dynamic jamming power scheduling becomes essential to conserve energy and prolong jamming duration. However, existing methods suffer from poor applicability in such scenarios, primarily due to their sparse deployment and adversarial nature. To address this limitation, this paper develops a set of mathematical models and a dedicated algorithm for air–ground communication countermeasures. Specifically, we (1) randomly select communication nodes to determine the jammer operation sequence; (2) schedule the number of active jammers by sorting transmission path losses in ascending order; and (3) estimate jamming effects using electromagnetic wave propagation characteristics to adjust jamming power dynamically. This approach formally converts the original dynamic, stochastic jamming resource scheduling problem into a static, deterministic one via cognitive certainty of dynamic parameters and deterministic modeling of stochastic factors—enabling rapid adaptation to unknown, dynamic communication power strategies and resolving the coordination challenge in air–ground joint jamming. Experimental results demonstrate that the proposed Transmission Loss Ordering Algorithm (TLOA) extends the system operating duration by up to 41.6% compared to benchmark methods (e.g., genetic algorithm).

**Keywords:** communication jamming; resource scheduling; mathematical model; intelligent algorithm; cognitive jamming strategy

Academic Editors: Wei Zong and Yang-Wai Chow

Received: 17 December 2025

Revised: 19 January 2026

Accepted: 23 January 2026

Published: 2 February 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

Jamming resource scheduling is a core technology in communication countermeasures, aiming to efficiently accomplish jamming tasks while conserving jamming resources [1–3]. In air–ground dynamic communication countermeasures, the number of communication nodes and jammers is large, and their spatial layout is complex. Communication parties exhibit sophisticated wireless links, strong adaptive power adjustment, and robust self-organizing network capabilities, which pose challenges to the system operating duration required for comprehensive communication jamming.

To disrupt air-ground communication network systems, jammers must apply dynamic jamming to all communication-receiving nodes. Unlike centralized, high-power jammers, distributed jammers feature spatially dispersed deployment and limited energy; their power synthesis and cooperative jamming rely more heavily on resource scheduling. Physical-layer security (PLS) techniques have emerged as effective approaches to address security challenges in wireless communication systems, including UAV communications and cooperative jamming [4]. Studies such as Hamamreh and Arslan's comprehensive survey [5] have classified PLS techniques into signal-to-interference-plus-noise ratio-based and complexity-based approaches, providing a foundational framework for designing secure communication and jamming strategies. Specifically, artificial noise injection and channel adaptation techniques proposed in PLS research [5–7] offer valuable insights for optimizing jamming resource allocation, as they leverage wireless channel characteristics to enhance the effectiveness of intended signals (or jamming signals) while suppressing unintended receivers (or communication nodes). Additionally, UAV communication security studies [4,8] highlight the unique challenges of airborne communication scenarios, such as dominant line-of-sight propagation and high mobility, which are highly relevant to air-ground joint jamming and require tailored resource scheduling strategies. Novel physical-layer key generation methods [9,10] that exploit subcarrier indices and channel gain characteristics also inspire innovations in jamming effect estimation and resource coordination for distributed jammers.

Existing research on resource scheduling has predominantly focused on radar jamming [11–19], yet their mathematical models and scheduling strategies are not applicable to communication jamming. Within communication jamming studies, greater attention has been paid to static resource scheduling problems. Progress has been achieved through conventional optimization approaches such as global search algorithms [20], convex optimization theory [21], various genetic algorithms [22–24], intelligent optimization algorithms [25–28], and knowledge-based Bayesian neural network algorithms [29]. These methods establish mathematical models for communication countermeasures by taking communication links, jamming power, and jamming frequency bands as optimization objectives, and seek optimal jamming solutions within the solution space. Most of these studies start from aspects such as jamming patterns, targets, and power, and explore optimal jamming strategies by constructing adversarial models between communicators and jammers. However, when dynamically scheduling jamming resources, the aforementioned methods suffer from shortcomings such as long search time and poor effectiveness. Path-loss-based prioritization has been widely used as a heuristic in wireless communication and cooperative jamming [30]; for example, Deng et al. [30] proposed a resource hopping mechanism in OTFS-SCMA systems that leverages channel characteristics to suppress jamming, demonstrating the effectiveness of path loss-related strategies in dynamic scenarios. A notable innovation of this work is its integration of resource hopping with OTFS's delay-Doppler domain advantages, which provides a new perspective for dynamic jamming resource scheduling. Building on this idea, this paper extends path-loss-based scheduling to the distributed air-ground joint jamming domain, addressing the unique challenges of multi-node coordination and dynamic power adjustment that are not fully covered in existing studies.

The rapid development of machine learning technologies has facilitated research on communication countermeasure algorithms [31–34]. Nevertheless, the difficulty in acquiring communication countermeasure datasets has limited the application of deep learning techniques [35]. Reinforcement learning algorithms can autonomously interact with the environment without prior information to learn optimal strategies, thus being widely used in dynamic resource scheduling [36–43]. However, most of these studies assume that jammers can directly and accurately obtain jamming effects. From a practical scenario

perspective, two limitations exist: first, they overlook implementation approaches (e.g., jamming index estimation and communication quality assessment); second, they fail to consider counter-jamming strategies adopted by communicators, such as power adjustment and channel switching [44,45]. Proactive eavesdropping and monitoring strategies, such as the UAV-based scheme proposed by Mobini et al. [46], have also explored dynamic resource optimization in wireless systems, but their focus on information collection rather than jamming resource scheduling leaves gaps in addressing air-ground joint jamming requirements.

To address such issues, some scholars have modeled the communication counter-measure process as a Markov Decision Process (MDP) based on the decision-making principles of communication jamming and proposed corresponding jamming methods [47]. They use channel alignment as an evaluation metric for effective jamming [48], employ Q-learning algorithms [49] to predict changes in communication channels, and develop jamming effectiveness metrics by integrating fundamental jamming principles and variations in communication target behaviors [50]. However, these jamming effect evaluation methods remain inadequate for air-ground joint jamming. They target independent and deterministic bidirectional communication links, assuming that jammers can quickly decipher communication link information and identify both the transmitting and receiving ends of communications. In complex communication networks, each communication node receives signals from multiple links, making it difficult for jammers to identify specific communication links. Notably, the aforementioned studies do not take the maximum received power of communication nodes as the jamming target.

The optimization objectives of the aforementioned algorithms focus on physical-layer parameters (e.g., jamming patterns and power) for single jammers, and their objective functions or reward functions rarely involve distributed jamming resource scheduling. These methods rely on heuristic iterative search strategies of algorithms or pre-trained neural networks in machine learning, failing to leverage information such as communication jamming scenarios and electromagnetic propagation characteristics to reduce decision dimensionality. By simply migrating relevant algorithms, they lack research on distributed communication jamming strategies and generally suffer from slow convergence and susceptibility to local optima.

References [51–55] have adopted techniques such as multi-agent reinforcement learning and hierarchical reinforcement learning to study a small number of high-power, long-range ground-deployed communication jammers. Reference [56] designed a proximal policy optimization algorithm to address the problem of scheduling jamming resources for a small number of airborne mobile jammers. However, the application scenarios of these algorithms do not involve a large number of communication nodes or jammers, and they pay little attention to the coordination between airborne and ground jammers.

In summary, after the static deployment of jammers, existing studies consider limited environmental information and exhibit poor practicality. They cannot rapidly perform dynamic jamming resource scheduling under conditions of sparse scenarios, multi-dimensionality, and incomplete prior information, leading to waste of jamming resources and degradation of jamming effectiveness. The reasons are as follows:

- The training and actual deployment of algorithms such as reinforcement learning and deep reinforcement learning are separated. When facing unknown and dynamic communication power scheduling strategies, first, the rapid decision-making of algorithms relies on long training durations and known air-ground communication scheduling strategies, making them unable to adapt to flexible and changeable battlefield environments; second, after jammers are deployed, multiple interactions with the environment are required, making it impossible to quickly achieve the goal of

comprehensive communication jamming and rendering them unsuitable for jammers with limited power.

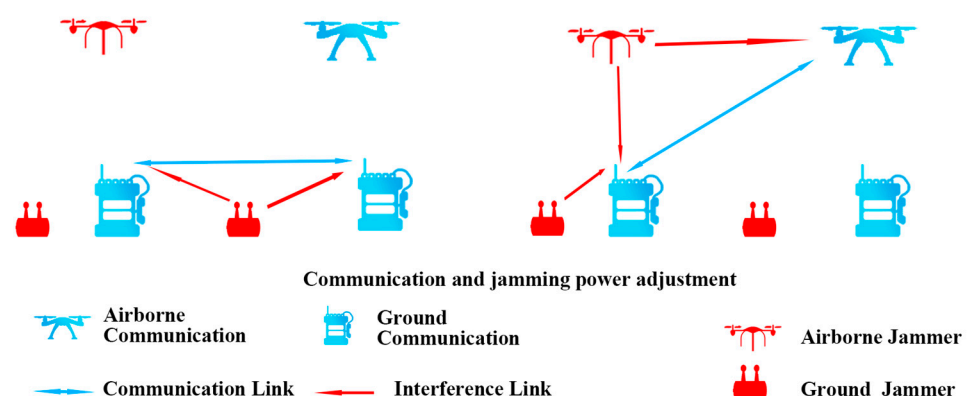
- Communication jammers can identify information about communication nodes and jammers from the electromagnetic environment and locate the source direction of signals. However, there is a lack of specific algorithmic strategies for distributed jammers to use reconnaissance information for environmental cognition and jamming effect evaluation.
- There is a scarcity of mathematical models suitable for air–ground joint jamming, as well as strategies for jamming power superposition and operational timing scheduling.

To address the above issues, the contributions of this paper are as follows:

- Based on the requirements of high-speed communication countermeasures, a deterministic power scheduling strategy is adopted, eliminating the need for complex round-by-round iterative searches in intelligent algorithms and training strategies with poor interpretability in machine learning algorithms.
- Based on communication information reconnoitered by jammers, a jamming effect evaluation strategy is designed by integrating mathematical estimation and changes in communication targets.
- A simulation experiment is designed: based on differentiated electromagnetic propagation models, the traditional method of selecting jammers based on spatial distance is abandoned, and a strategy of selecting jammers by sorting transmission path loss in ascending order is proposed.

## 2. Communication Countermeasure Mode

The air–ground communication network system has self-organizing and encryption functions. It is difficult for the jamming side to obtain the transmitting and receiving ends of the communication link through reconnaissance [57–62]. Therefore, it is necessary to apply jamming to the maximum receiving power of communication nodes. In a complex electromagnetic environment, a single jammer has limited information and cannot obtain overall feedback for decision-making, which makes cooperation difficult, as shown in Figure 1.



**Figure 1.** Schematic diagram of air–ground dynamic resource scheduling.

### 2.1. Mathematical Model of the Air–Ground Communication Network System

The path loss between aerial devices, as well as between aerial and ground devices, follows line-of-sight (LoS) propagation [23]. In Equation (1),  $L_s$  represents the path loss of the LoS propagation model,  $f$  denotes the electromagnetic wave frequency (MHz),

$n_1$  is the environmental factor that varies with propagation conditions, and  $R$  stands for the LoS transmission distance (km).

$$L_s = 32.5 + 20 \lg f + 10n_1 \lg R \quad (1)$$

The path loss between ground devices follows two-ray propagation [28]. In Equation (2),  $L_d$  represents the path loss of the two-ray propagation model,  $n_2$  denotes the terrain influence exponent (which varies with propagation conditions),  $R'$  indicates the two-ray transmission distance (km), and  $h_t$  and  $h_r$  correspond to the transmitter height and receiver height of the electromagnetic wave respectively (m).

$$L_d = 120 + 10n_2 \lg R' - 20 \lg h_t - 20 \lg h_r \quad (2)$$

In Equation (3),  $P_c^t$  represents the transmission power of the communication signal,  $P_c^r$  denotes the received power of the communication signal (dBW),  $G_{tc}$  indicates the antenna gain of the communication transmitter in the direction of the communication receiver,  $G_{rc}$  stands for the antenna gain of the communication receiver in the direction of the communication transmitter,  $L_c$  corresponds to the path loss of the communication transmission, and  $L_{pc}$  represents the cable and connector loss at the communication receiver (all units in dB).

$$P_c^t = P_c^r - G_{tc} - G_{rc} + L_c + L_{pc} \quad (3)$$

The received power at the communication receiver must satisfy the communication link margin requirement, taking into account the environmental noise power. In Equation (4),  $RS$  represents the receiver sensitivity of the communication equipment,  $\sigma^2$  denotes the environmental noise power, and  $SFM$  corresponds to the System Fade Margin.

$$P_c^r - \max(RS, \sigma^2) > SFM \quad (4)$$

## 2.2. Mathematical Model of Air–Ground Joint Communication Jamming

### 2.2.1. Jamming Range

Distributed communication jamming has the effect of power superposition. In order to reduce the calculation of weak jamming power, according to the environment and device performance, an upper limit  $D'_s$  of the calculable jamming spacing is set. The communication jamming signal is line-of-sight communication, and the distance of the jammer should not be greater than the line-of-sight propagation distance  $D_s$ . In Equation (5), the jamming range for calculation between the  $i$ -th jammer and the  $j$ -th communication device is  $D_{ij}$ , which takes the minimum value of the two distance limitations.

$$D_{ij} = \min(D_s, D'_s) \quad (5)$$

### 2.2.2. Quantized Jammer-to-Signal Ratio

Within the computable jamming range, the ratio of the total received jamming power (from both aerial and ground jammers) to the maximum communication received power is denoted as  $k_{jb}$ . In Equation (6), if the  $i$ -th jammer is an aerial jammer, its jamming power is expressed as  $P_{ia}^{rj}$ , with being the total number of jammers  $I$ . If the  $i$ -th jammer is a ground jammer, its jamming power is denoted as  $P_{is}^{rj}$ . Here,  $P_{ja}^{rj}$  represents the maximum received power from aerial communication devices for the  $j$ -th communication receiver, while  $P_{js}^{rj}$  indicates the maximum received power from ground communication devices for the  $j$ -th communication receiver (all units in dBW).

$$k_{jb} = \sum_{i=1}^I (P_{ia}^{ij} + P_{is}^{ij}) x_{ij} + \sigma^2 - \max(P_{ja'}^{ij}, P_{js}^{ij}), j' \in (1, j) \cup (j, J) \quad (6)$$

To intuitively evaluate the impact of jamming signals on communication signals and facilitate calculation of the required jamming power for successful jamming, the jammer-to-signal ratio (JSR) needs to be normalized. If the power ratio of the jamming signal to the received communication signal is not less than the threshold  $k_j$ ; the jamming is considered successful, the normalized JSR<sub>*j*</sub> for the *j*-th communication device equals 1, as shown in Equation (7).

$$JSR_j = \begin{cases} 0, & k_{jb} < k_j \\ 1, & k_{jb} \geq k_j \end{cases} \quad (7)$$

The air-ground communication network employs anti-jamming measures such as multi-hop relaying, requiring the jamming system to effectively disrupt received signals at all communication devices. In Equation (8), *J* represents the total number of communication devices. The equality holds if and only if comprehensive jamming is achieved across the entire system; otherwise, it fails to hold.

$$\sum_{j=1}^J JSR_j = J \quad (8)$$

### 2.2.3. System Operating Duration

If comprehensive communication jamming cannot be achieved during system operation, the operational duration of the jamming system is considered terminated. The total operational duration *T* of the jamming system is the sum of the durations *T<sub>iq</sub>* of individual scheduling instances.

In Equation (9), after the *i*-th jammer undergoes the *q*-th scheduling, *T<sub>iq</sub>* is its operating duration this time, with the unit of hours (H).  $\eta_s$  is the battery power limit required for normal jamming, *U<sub>out</sub>* is the battery operating voltage, with the unit of volt (V), *E<sub>esq</sub>* is the battery power consumed in this scheduling, with the unit of ampere-hours (AH), *P<sub>i</sub>* is the jamming output power of the *i*-th jammer, and *P<sub>x</sub>* is the power consumed by the jammer to maintain other functions, with the unit of watts (W).

$$T_{iq} = \frac{\eta_s U_{out} E_{esq}}{(P_i + P_x)} \quad (9)$$

### 2.3. Strategies of the Air-Ground Communication Network System

In the absence of jamming, the communication system's initial power scheduling strategy employs reduced transmission power to maintain basic transmitting and receiving functions, achieving dual objectives of energy conservation and minimized electromagnetic signature exposure.

During signal transmission, communication devices select the path with minimal transmission loss and employ the lowest power level sufficient to meet communication requirements. In Equation (10),  $P_{cj}^{\min}$  denotes the minimum transmission power for the *j*-th communication device; *L<sub>cj'</sub>* represents the path loss for communication with other devices.

$$P_{cj}^{\min} = SFM + \max(RS, \sigma^2) - G_{tc} - G_{rc} + \min(L_{cj'}) + L_{pc}, j' \in (1, J) \quad (10)$$

When comprehensive communication jamming is achieved, the received power at all communication nodes becomes suppressed. In response, the communication system adaptively adjusts its transmission power based on the current conditions: it may increase power to counteract the jamming, cease transmission to evade jamming, or reduce power to maintain operations while confusing the jammer.

#### 2.4. Air–Ground Communication Jamming Strategy

Intelligent jamming adopts two approaches for effect evaluation. First, based on the information reconnoitered by jammers, the JSR is estimated according to the positional relationship between jammers and communication devices, as well as electromagnetic wave propagation characteristics. Second, changes in communication parameters such as transmission rate, power, and channel are reconnoitered to determine jamming effectiveness. However, existing studies only introduce the above approaches in the background section and do not mention them in the algorithm design. There is a lack of effective strategies for processing reconnaissance information and estimating the JSR in “many-to-many” communication jamming scenarios.

This paper uses historical and real-time communication information reconnoitered by jammers to estimate parameters such as the source of communication signals, transmit power, receive power, and JSR, as detailed in Algorithm 1. The estimation of communication signal sources involves the following process: after jammers detect electromagnetic signals from the electromagnetic environment, they identify the direction of the signal source, compare it with the pre-stored positions of communication nodes and jammers, and determine whether the detected signals are communication signals. On this basis, the transmit power and receive power of communication nodes are estimated according to the spatial positional relationship between communication nodes and electromagnetic propagation characteristics. Jammers maintain unobstructed communication among themselves, and the jamming power of each jammer is known, enabling the estimation of the JSR at communication nodes. Compared with the simple judgment method that only relies on changes in communication state information, the proposed method—after estimating the jamming-to-signal ratio—can identify misleading information (e.g., communication parties deliberately reducing communication power to feign jamming effectiveness), resulting in a more comprehensive cognitive function.  $P_{cj}^t$  denotes transmission power for the  $j$ -th communication device;  $P_c^{r\max}$  is the max received power of the communication signal, with the unit of watts (W).

---

#### Algorithm 1: Cognitive and Jamming Effect Estimation Strategy

---

**Input:** The detected communication information

**Output:** Estimated communication and jamming effects

- 1: Select  $i$  in order, for  $i = 1$  to  $I$
  - 2:     Select  $j$  in order, for  $j = 1$  to  $J$
  - 3:         If the distance between  $i$  and  $j$  is less than  $D_{ij}$
  - 4:             Store the azimuth angle and elevation angle information of  $j$  relative to  $i$
  - 5:             End if
  - 6:     If the distance between  $i$  and  $j$  is less than  $D_{ij}$
  - 7:          $i$  detected the source direction and power of  $j$
  - 8:     End if
  - 9:     Select  $j'$  in order, for  $j' = 1$  to  $J$
  - 10:     Compare  $j'$  and  $j$ , and infer the information of  $j$
-

---

```

11:   Estimate  $P'_{cj}$  (round it to two decimal places using the method of rounding
      down for 0 and rounding up for 1)
12:   Jammer Information Sharing
13:   Select  $j$  in order, for  $j = 1$  to  $J$ 
14:   Update  $P'_c$ 
15:   Select  $j$  in order, for  $j = 1$  to  $J$ 
16:   Update  $P_c^{r\max}$  (with an estimation margin of 0.3 dB)
17:   Select  $i$  in order, for  $i = 1$  to  $I$ 
18:   Estimate  $JSR_j$ 
19:   If  $\sum_{j=1}^J JSR_j = J$ 
20:     The communication jamming is successful
21:   Else:
22:     The communication jamming fails

```

---

After estimating the status of the communication network system, TLOA is designed to leverage the advantages of short distance and power superposition. Assuming that all communication nodes operate at full power, aiming at the maximum received power of all communication devices, all jammers execute full-power jamming to achieve an overall jamming layout, which is a static layout. On this basis, TLOA performs dynamic jamming resource scheduling by randomly selecting communication nodes as jamming targets, preferentially using jammers with smaller transmission path loss based on the transmission path loss sorting to save energy, and using an estimation strategy to schedule the jamming power of jammers, avoiding the inefficiency of random search, as shown in Algorithm 2.

---

**Algorithm 2:** Transmission Loss Order Algorithm

---

**Input:** Estimated communication information and real-time information of the interfering party

**Output:** New real-time information of the interfering party

```

1:   Initialize: The power required of  $j$  for successful jamming is  $P_j^n$ , the path loss
      between  $i$  and  $j$  is  $L_{cij}$ 
2:   Select  $j$  out of order, for  $j = 1$  to  $J$ 
3:   Select  $i$  in order, for  $i = 1$  to  $I$ 
4:   If the distance between  $i$  and  $j$  is less than  $D_{ij}$ 
5:     Calculate  $L_{cij}$ , sort them in ascending order, and form an jamming list
6:   End if
7:   Select  $i$  in order, for  $i = 1$  to  $I$ 
8:    $P_j^n = P_c^r k_j - \sigma^2$ 
9:   Traverse  $i$  within the jamming list.
10:  Estimate  $P_i^n$  (the jamming power requirement of  $i$ )
11:  If the battery power cannot guarantee  $P_i^n$ 
12:     $P_i^t = 0$ 
13:  Else:
14:     $P_i^{new} = P_i^t + P_i^n$ 
15:    If  $P_i^{new} < P_j^{limit}$  (upper limit of the jammer power)
16:       $P_i^t = P_i^{new}$ 
17:       $P_j^n = 0$ 

```

---

---

```

18: Else:
19:    $P_i^t = P_j^{limit}$ 
20:    $P_j^n = P_i^{new} - P_j^{limit}$ 
21:   If  $P_j^n > 0$ 
22:     If  $i = I$ 
23:       The jamming fails, break
24:     Else:
25:        $i = i + 1$ 
26:   Else:
27:     If  $j = J$ 
28:       The jamming is successful, break
29:     Else:
30:        $j = j + 1$ 

```

---

In TLOA, the operations are carried out sequentially. First, traverse the jammers and then the communication nodes. Second, traverse the communication nodes and then the jamming list. Third, traverse the communication nodes twice.

In Algorithm 2, the jamming party adopts real-number encoding, which can represent continuous real values and reduces coding conversion errors when scheduling jamming resources. Transmission path loss ordering refers to sorting the transmission path losses between jammers and communication nodes in ascending order, and selecting jammers one by one according to this order. This strategy effectively solves the problems of heterogeneous jammer decision-making, jammer selection, and continuous power value scheduling, while reducing the decision dimension.

By using the above four algorithms, this thesis designs the confrontation scenarios of the air-ground communication network system and the air-ground joint jamming, as shown in Algorithm 3.

---

#### **Algorithm 3:** Air-ground Communication Countermeasure Simulation

---

**Input:** Initial State of the Air-ground Communication Network System

**Output:**  $T$

```

1: Initial state scheduling of the air-ground communication network system
2: The jammer perceives the environment and schedules jamming resources first
3: Loop
4:   The communication party adjusts its power after being jammed
5:   The jamming party schedules its jamming resources
6:   If the jamming is successful and the remaining battery power is sufficient
7:     Continue the loop
8:   Else:
9:     Break
10: Calculate  $T$ 

```

---

#### *2.5. Conversion from a Dynamic-Stochastic Problem to a Static-Deterministic Problem*

The air-ground joint jamming resource scheduling problem is inherently dynamic and stochastic, stemming from two key factors. On the one hand, communication systems adaptively adjust their transmission power based on actual scenarios. This time-varying power adjustment directly leads to continuous fluctuations in the received power of communication nodes, keeping the signal strength of jamming targets in a dynamic state. On the other hand, communication nodes switch dynamically between active and dormant states. Meanwhile, the uncertainty of environmental factors during electromagnetic propagation (such as environmental impact parameters in line-of-sight propagation scenarios)

results in random characteristics in the scope of jamming target sets and path loss values, further increasing the complexity of the problem.

From the perspective of optimization objectives, the core of this problem is to maximize the effective operating duration of the entire jamming system while satisfying various constraints. These constraints mainly include the following: the output power of each jammer must be controlled within its minimum and maximum rated power range; the total energy consumption of all jammers cannot exceed the total energy limit of their on-board batteries; the jamming effect must meet the preset JSR threshold, meaning the ratio of jamming signal power to communication signal power at each communication node must reach the specified standard. Due to the aforementioned dynamic and stochastic factors, key parameters such as the JSR and effective jammer sets change over time, making the entire problem a time-varying optimization problem with stochastic constraints.

The core logic of TLOA in converting a dynamic-stochastic problem to a static-deterministic one lies in two design concepts: “cognitive certainty” and “decision-time decoupling”.

In terms of cognitive certainty of dynamic parameters, instead of tracking the dynamic changes in the received power of communication nodes in real time, the algorithm estimates the maximum possible received power of communication nodes by fusing historical reconnaissance data and real-time electromagnetic perception information. This approach is fully justified: the transmission power adjustment of communication systems is always limited by their maximum hardware-rated power (25 W for ground communication nodes and 10 W for airborne communication nodes), so the upper limit of their received power is deterministic. The maximum received power estimated based on the maximum transmission power can serve as a stable reference benchmark.

For the handling of random path loss, the algorithm uses the expected value of path loss instead of its random fluctuation value. This expected value is calculated through preset electromagnetic propagation models, where the environmental factor for line-of-sight propagation is fixed at 3, and the terrain influence exponent for two-ray propagation is fixed at 4. The rationality of this simplification is mainly reflected in two aspects: first, in short-term battlefield scenarios, environmental conditions are relatively stable (the power adjustment interval of the communication system is 90 s), so environmental factors will not change drastically; second, the random fluctuation range of path loss is extremely small (the fluctuation value corresponding to the variance is less than 1 dB), which is much smaller than the preset JSR threshold (4.77 dB), and its impact on the final jamming effect is negligible.

In terms of decision-time decoupling, the algorithm no longer performs separate optimization decisions on jamming power and activation status for each time slot. Instead, it designs a set of static jamming resource allocation schemes, including a fixed jammer activation set and a unified jamming power allocation strategy. This scheme must satisfy that the JSR of all communication nodes meets the preset threshold. The calculation of JSR is based on the estimated maximum received power and the expected value of path loss, thereby converting the originally time-varying dynamic optimization problem into a static-deterministic problem that can be solved in one go.

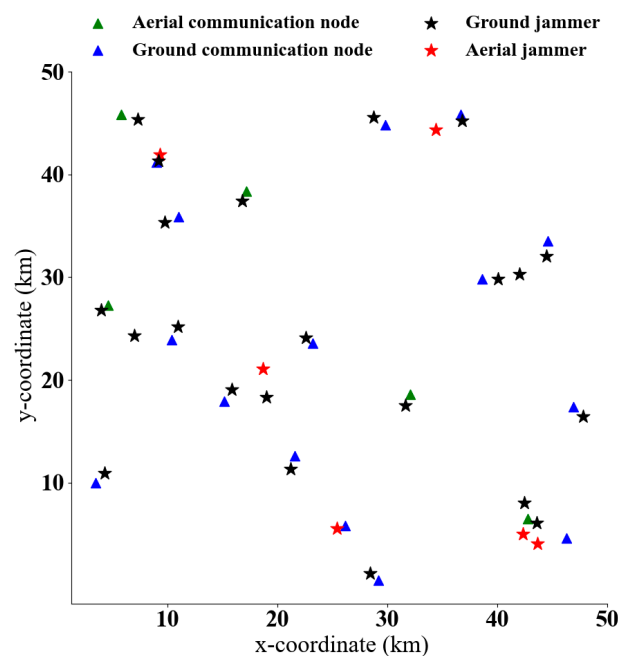
Regarding the proof of the effectiveness of the conversion strategy, first, consider the impact of communication power fluctuations. Since the actual transmission power of the communication system will never exceed its maximum rated power, the actual received power of communication nodes must be less than or equal to the maximum received power estimated by the algorithm. The JSR is inversely proportional to the received power; the smaller the actual received power, the larger the JSR. Therefore, the actual JSR must be greater than or equal to the JSR calculated based on the maximum received power. Second, consider the impact of path loss fluctuations. The greater the path loss, the more

severe the attenuation of the jamming signal, so the most unfavorable scenario is when the path loss takes the maximum value. At this time, the JSR will decrease accordingly. However, since the TLOA reserves a 0.3 dB margin when designing the static solution, even considering the most unfavorable path loss fluctuations, the final JSR can still meet the preset threshold.

### 3. Simulation Experiment

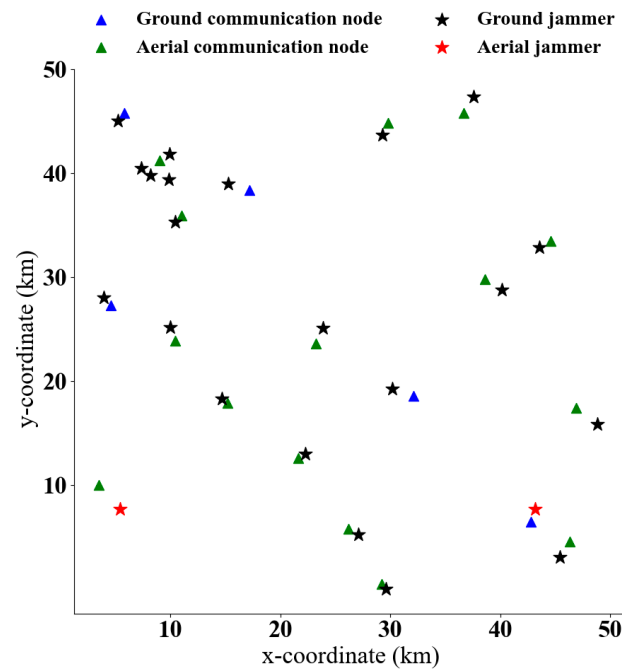
#### 3.1. Parameter Setting of the Simulation Scenario

The differences in the composition structure of the air–ground communication network system and the jammer layout will affect dynamic jamming resource scheduling. In response, this paper designs four simulation scenarios, as shown in Figures 2–5 and Table 1. The scenario design is not arbitrary but is strategically tailored to mimic real-world battlefield characteristics of “dynamic node activation” and cover key environmental variables, ensuring comprehensive verification of the TLOA’s effectiveness, robustness, and scalability.

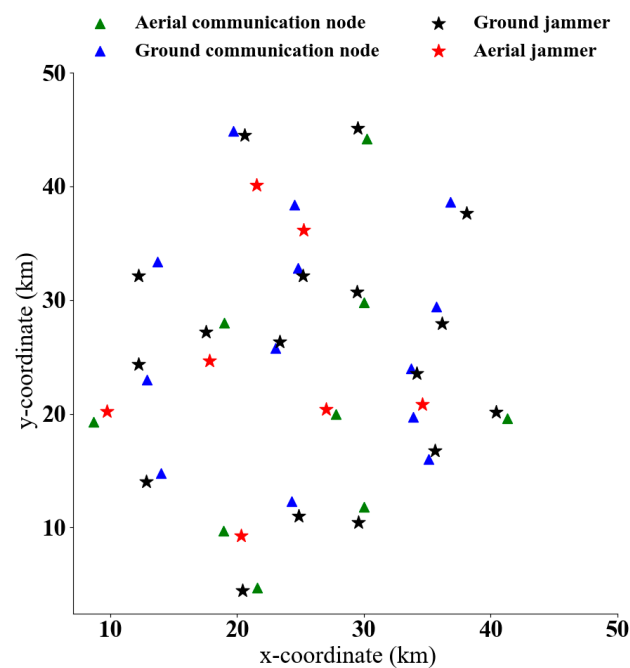


**Figure 2.** Scenario 1: ground-dominated communication network with balanced jammer deployment.

In Figures 2–5, the green triangles represent airborne communication nodes, the blue triangles denote ground communication nodes, the black pentagrams signify ground jammers, and the red pentagrams indicate airborne jammers. The horizontal axis and vertical axis represent spatial distance, with the unit of kilometers (km).



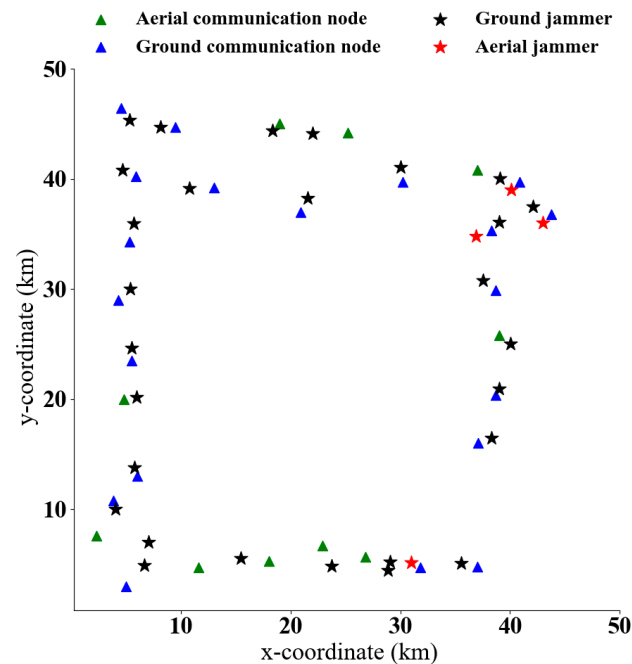
**Figure 3.** Scenario 2: ground-dominated communication network with sparse airborne jammers.



**Figure 4.** Scenario 3: hybrid communication network with moderate jammer configuration.

**Table 1.** Setting of communication countermeasure scenarios.

|                              | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|------------------------------|------------|------------|------------|------------|
| Airborne Communication Nodes | 5          | 5          | 9          | 10         |
| Ground Communication Nodes   | 15         | 15         | 13         | 20         |
| Ground jammers               | 22         | 21         | 17         | 28         |
| Airborne jammer              | 6          | 2          | 7          | 4          |



**Figure 5.** Scenario 4: large-scale dense communication network with heavy ground jammer deployment.

The key design logic of the four scenarios is as follows:

First, dynamic activation of deployed nodes mimics real-world communication networks where different air-ground communication nodes are activated at varying times based on mission requirements. This forms a time-varying air-ground communication network system, which directly alters the network's anti-jamming capabilities (e.g., the number of active nodes influences the effect of jamming signal superposition and the self-organizing recovery capability of the communication network) and the characteristics of electromagnetic propagation paths (e.g., activation of airborne nodes increases line-of-sight links, while ground node activation enhances two-ray propagation dominance).

Second, heterogeneous jammer deployments (Scenarios 1 vs. 2) test the algorithm's optimization performance when airborne jamming resources are adjusted—a common tactical adjustment in air-ground coordinated jamming missions.

Third, scaled node counts and density (Scenarios 3 vs. 4) validate the algorithm's scalability: Scenario 4 increases the total number of communication and ground jammer nodes to simulate large-scale joint operations, introducing higher decision dimensionality and electromagnetic complexity to challenge the algorithm's efficiency and robustness.

In the experiment, an Intel (R) Core (TM) i5-8300H CPU @ 2.30 GHz processor with 16.0 GB of RAM and an NVIDIA GTX1080Ti graphic card were used, and the environment of Anaconda 23.7.4 was selected for verification. The parameters of the simulation experiment are shown in Table 2.

**Table 2.** Table of simulation scenario parameter settings.

| Project                                                        | Parameter   |
|----------------------------------------------------------------|-------------|
| Maximum transmitting power of airborne communication equipment | 10 W        |
| Antenna gain of airborne communication equipment               | 2 dBi       |
| Height of airborne communication equipment                     | 2000–3000 m |
| Maximum transmitting power of ground communication equipment   | 25 W        |
| Antenna gain of ground communication equipment                 | 2.5 dBi     |
| Height of ground communication equipment                       | 5 m         |
| Communication frequency band                                   | 600 MHz     |

|                                                                               |               |
|-------------------------------------------------------------------------------|---------------|
| Receiving sensitivity of communication equipment                              | −133 dBW      |
| Communication link margin                                                     | 12 dB         |
| Ambient electromagnetic noise                                                 | −115 dBW      |
| Antenna gain of jammers                                                       | 2 dBi         |
| Maximum jamming power of airborne jammers                                     | 20 W          |
| Height of airborne jammers                                                    | 300–2000 m    |
| Maximum jamming power of ground jammers                                       | 40 W          |
| Height of ground jammers                                                      | 3 m           |
| Planar area of the simulation scenario                                        | 50 km × 50 km |
| Attenuation of cables and cable connectors at the communication receiving end | 1 dB          |
| Environmental factor for line-of-sight propagation                            | 3             |
| Environmental factor for two-ray propagation                                  | 4             |
| JSR                                                                           | 3             |
| Maximum battery power limit required for normal jamming                       | 0.95          |
| Operating voltage of jammer batteries                                         | 24 V          |
| Battery capacity of airborne jammers                                          | 12 AH         |
| Battery capacity of ground jammers                                            | 20 AH         |
| Energy consumption of other functions of jammers                              | 4 W           |
| Minimum distance between jammers and communication equipment                  | 500 m         |
| Calculated distance for two-ray propagation                                   | 15 km         |
| Calculated distance for line-of-sight propagation                             | 20 km         |
| Population size                                                               | 10            |
| Power levels of intelligent optimization algorithms                           | 1 w           |
| Time limit for the first decision of intelligent optimization algorithms      | 10 s          |
| Power levels of reinforcement learning algorithms                             | 5 w           |
| Time limit for non-first jamming decision                                     | 1 s           |
| Communication state adjustment interval                                       | 90 s          |
| Angle Recognition Accuracy of Communication Nodes                             | 0.3 degrees   |

### 3.2. Results and Analysis of the Comparative Experiment

In this thesis, the TLOA is compared with DQN [63], DDQN [64], the probability mutation artificial bee colony algorithm (PMABCA) [26], the simple random search algorithm (SRSA), and the genetic algorithm (GA) [65]. After each algorithm runs 10 times, the average value is taken.

DQN approximates the action-value function (Q-function) using a deep neural network (DNN) and adopts experience replay to break the correlation between training samples, while DDQN mitigates the overestimation bias of Q-values by separating the target network (for calculating target Q-values) and the evaluation network (for selecting actions) [63,64]. PMABCA is an improved variant of the artificial bee colony (ABC) algorithm; it introduces a probability-based mutation mechanism to enhance global search capability. The algorithm divides bees into employed bees, onlooker bees, and scout bees: employed bees exploit food sources (candidate solutions), onlooker bees select food sources based on fitness values, and scout bees abandon poor-quality food sources (via probability mutation) to avoid local optima. It is widely used in continuous optimization problems such as resource scheduling [26]. SRSA is a search algorithm with minimal computational complexity. It randomly generates candidate solutions within the feasible solution space without leveraging prior information or iterative optimization. Each solution is independently sampled, and the optimal solution is selected based on the fitness function. While simple to implement, it suffers from low search efficiency and is prone to missing global optima in complex high-dimensional scenarios. GA is a population-based

evolutionary algorithm inspired by biological natural selection and genetic variation. It initializes a population of candidate solutions, and iteratively optimizes via three core operations: selection (reserving high-fitness individuals), crossover (combining genetic information of parent individuals), and mutation (randomly altering individual genes to maintain population diversity). It is suitable for complex combinatorial optimization problems but may face slow convergence in dynamic scheduling scenarios [65].

In terms of computational complexity, the key characteristics of each algorithm are analyzed as follows:

- DQN/DDQN: The time complexity is  $O(T \times B \times (D + H))$ , where  $T$  denotes the number of training steps (set to 1000 in this study),  $B$  is the batch size (32),  $D$  represents the state dimension (equal to the number of jammers), and  $H$  is the number of hidden layer neurons (64). The complexity is dominated by the forward and backward propagation of the deep neural network, and DDQN introduces an additional target network update overhead ( $O(U \times (D + H))$ ,  $U = 100$  update frequency), which is negligible compared to the overall complexity.
- PMABCA: Its time complexity is  $O(G \times N \times (D + T_{\text{limit}}))$ , where  $G = 500$  (maximum number of generations),  $N = 10$  (population size),  $D = M$ , and  $T_{\text{limit}} = 100$  (food source trial limit). The complexity is higher than that of GA due to the additional trial count monitoring and probability mutation operations for scout bees.
- SRSA: With a time complexity of  $O(S \times D)$  ( $S = 1000$  random sampling times,  $D = M$ ), it exhibits the lowest complexity among all algorithms. This is because it avoids iterative optimization and only involves random sampling and fitness evaluation of candidate solutions.
- GA: The time complexity is  $O(G \times N \times D)$ , where  $G = 100$  (maximum number of generations),  $N = 10$  (population size), and  $D = M$ . The complexity is linearly related to the number of jammers, dominated by iterative selection, crossover, and mutation operations, as well as fitness evaluation for each individual in the population.
- TLOA: The time complexity is  $O(n \times m + m \log m)$ , where  $n$  is the number of communication nodes and  $m$  is the number of jammers. The  $m \log m$  term originates from the sorting of transmission path losses, and  $n \times m$  comes from traversing the jammer–communication node pairs. The overall complexity is linear with the scale of nodes and jammers, ensuring efficient operation in large-scale scenarios.

Table 3 presents the detailed parameter settings of all comparative algorithms to ensure the reproducibility of the experiment.

**Table 3.** Parameter settings of comparative algorithms.

| Project                             | Parameter                                          |
|-------------------------------------|----------------------------------------------------|
| Population size of GA               | 10                                                 |
| Maximum number of generations of GA | 100                                                |
| Crossover rate of GA                | 1                                                  |
| Mutation rate of GA                 | 1                                                  |
| Population size of PMABCA           | 10                                                 |
| Number of max generation of PMABCA  | 500                                                |
| Mutation rate of PMABCA             | 0.9                                                |
| Food source trial limit of PMABCA   | 100                                                |
| State space of DQN and DDQN         | communication node, jammer and environmental state |
| Action space of DQN and DDQN        | 2                                                  |
| Gamma of DQN and DDQN               | 0.95                                               |
| Epsilon of DQN and DDQN             | initial: 1.0                                       |

|                                             |                                      |
|---------------------------------------------|--------------------------------------|
| Epsilon_min of DQN and DDQN                 | 0.01                                 |
| Epsilon_decay of DQN and DDQN               | 0.995                                |
| Learning rate of DQN and DDQN               | 0.001                                |
| Neural network architecture of DQN and DDQN | 64-dimensional fully connected layer |
| Activation function of DQN and DDQN         | ReLU                                 |
| State_size of DQN and DDQN                  | jammer number                        |
| Batch_size of d = DQN and DDQN              | 32                                   |

The TLOA can schedule continuous jamming power values, the number of jammers, and operational timing sequences—a capability absent in the compared algorithms. Therefore, the comparison algorithms are configured with fixed power levels and utilize all available jammers. Before executing the first jamming operation, distributed jammers have ample computation time, allowing extended decision-making duration for the algorithm's initial deployment. However, since communication systems employ power-adjustment countermeasures and dynamic jamming requires high timeliness while aiming for prolonged comprehensive jamming effects, each subsequent scheduling cycle is constrained by significantly shorter time limits.

Reinforcement learning algorithms have been rarely studied for dynamic distributed communication jamming, lacking corresponding reward functions. Therefore, reward functions for DQN and DDQN are designed as shown in Equation (11).

$$R = \begin{cases} 1 - P_{aj} / P_{mj}, & \text{succceed} \\ -1, & \text{otherwise} \end{cases} \quad (11)$$

In Equation (11),  $P_{aj}$  is the sum of the jamming powers of the current jammers, and  $P_{mj}$  is the sum of the upper limit values of the jamming powers of the jammers. A positive reward is obtained when the jamming is successful; otherwise, a negative reward is obtained.

In Table 4, the comparison of first scheduling time demonstrates that TLOA, as a deterministic algorithm, exhibits robustness in solving problems through a single iteration. Compared with DQN and DDQN algorithms, it has the advantages of eliminating the need for training and iterative search for solutions.

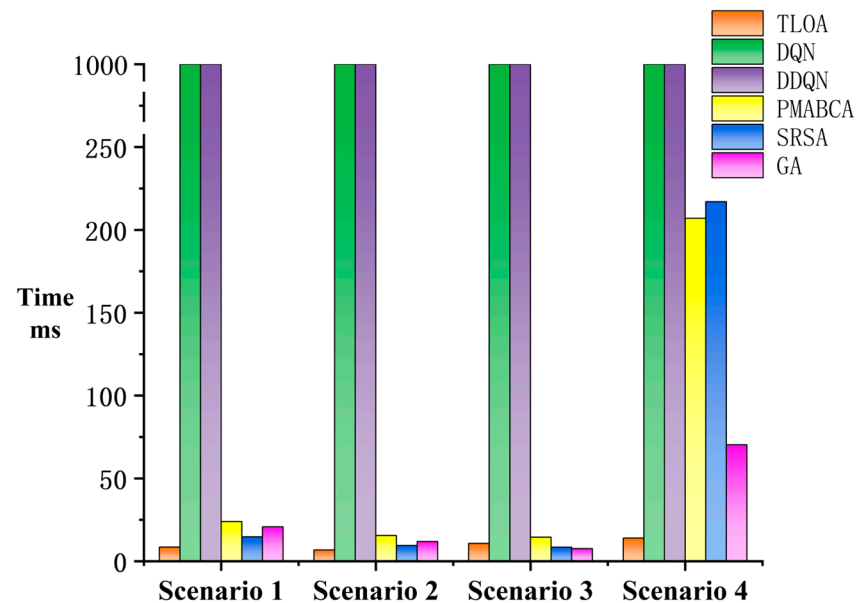
**Table 4.** Comparison of the algorithm results.

| Time Average |                           | Algorithm      |                |                |               |                |                |
|--------------|---------------------------|----------------|----------------|----------------|---------------|----------------|----------------|
|              |                           | TLOA           | DQN            | DDQN           | PMABCA        | SRSA           | GA             |
| Scenario 1   | first decision            | 11.9 ms        | 10 s           | 10 s           | 18.9 ms       | 6.9 ms         | 56.8 ms        |
|              | each decision             | <b>8.6 ms</b>  | 1 s            | 1 s            | 23.9 ms       | <b>14.8 ms</b> | 20.8 ms        |
|              | system operating duration | <b>23.97 H</b> | <b>16.93 H</b> | 16.65 H        | 15.7 H        | 15.65 H        | 15.87 H        |
| Scenario 2   | first decision            | 11.9 ms        | 10 s           | 10 s           | 8.9 ms        | 5.9 ms         | 43.9 ms        |
|              | each decision             | <b>6.8 ms</b>  | 1 s            | 1 s            | 15.5 ms       | <b>9.5 ms</b>  | 11.9 ms        |
|              | system operating duration | <b>41.08 H</b> | <b>20.3 H</b>  | 19.78 H        | 15.53 H       | 17.0 H         | 16.925 H       |
| Scenario 3   | first decision            | 26.9 ms        | 10 s           | 10 s           | 7.9 ms        | 8.9 ms         | 55.9 ms        |
|              | each decision             | <b>10.8 ms</b> | 1 s            | 1 s            | 14.5 ms       | 8.5 ms         | <b>7.6 ms</b>  |
|              | system operating duration | <b>38.95 H</b> | 21.08 H        | <b>21.38 H</b> | 17.03 H       | 17.9 H         | 17.75 H        |
| Scenario 4   | first decision            | 22.9 ms        | 10 s           | 10 s           | 24 ms         | 30.9 ms        | 285 ms         |
|              | each decision             | <b>14 ms</b>   | 1 s            | 1 s            | 207 ms        | 217 ms         | <b>70.3 ms</b> |
|              | system operating duration | <b>16.1 H</b>  | 0.2 H          | 0.53 H         | <b>0.55 H</b> | 0.27 H         | 0.2 H          |

The bold text in the table indicates the comparative data of the algorithms.

Figure 6 shows the scheduling time per iteration. Compared with the benchmark algorithms, our proposed algorithm reduces the scheduling time by 42%, 28.5%, −42%, and

80% in the four scenarios, respectively. The results indicate that the scheduling time of DQN and DDQN reaches the predefined time limit. This is because reinforcement learning algorithms struggle to converge under small-sample and short-time conditions, often requiring multiple iterations to improve solution quality. In most scenarios, TLOA demonstrates significant advantages in decision-making speed, addressing the inefficiency issue of stochastic search strategies.



**Figure 6.** Comparison of the time of each decision.

The “time of each decision” refers to the average time for the algorithm to complete one jamming resource scheduling iteration (including jammer selection, power adjustment, and jamming effect estimation). Its brevity is critical to adapting to the dynamic countermeasures of practical feasibility of distributed jammers. The comparison between Scenario 1 and Scenario 2 reveals that our algorithm maintains superior computational efficiency when dealing with different proportions of airborne and ground communication nodes. In Scenario 4, where the number of air–ground communication nodes increases and the network structure becomes more complex, the solution efficiency of intelligent optimization algorithms declines sharply due to the increased iteration number required to search for optimal solutions. In contrast, TLOA exhibits a much smaller increase in time cost, highlighting its robustness and scalability in complex environments.

Figure 6 illustrates the comparison of system operation duration. Compared with Scenarios 1, 2, 3, and 4, the algorithm in this chapter extends the operation duration by 41.6%, 102.3%, 82.2%, and 2827.3%, respectively. The comparison of the four scenarios indicates that the system operation duration of dynamic jammer scheduling depends on the static layout of jammers that can achieve overall communication jamming. In the complex Scenario 4, the benchmark algorithms fail to solve the problem quickly, resulting in a short system operation duration that cannot meet the continuous, dynamic, and overall jamming requirements of the jamming side, while TLOA operates stably and efficiently.

In air–ground joint jamming, the large number of communication nodes and jammers, wide power range, sparse scenarios, and complex signal propagation links lead to an increase in decision dimensions. In traditional centralized jamming and simple jamming scenarios, heuristic and traversal-based stochastic search strategies exhibit low efficiency. The lack of cognition about distributed communication jamming and power scheduling strategies causes iterative search in genetic algorithms, simple random search, and

probability mutation artificial bee colony algorithms to be inefficient, making it difficult to quickly obtain feasible solutions in complex scenarios.

Reinforcement learning algorithms rely on the setting of reward functions and require training strategies and dynamic jamming strategies adapted to air–ground joint jamming. The existing research lacks such adaptations, leading to long training time, poor training effect, low search efficiency in dynamic scheduling, and low solution quality of reinforcement learning algorithms.

### 3.3. Results and Analysis of Experiments Under Non-Ideal Channel State Information

To address the concerns raised by the reviewers regarding channel state information (CSI) assumptions and simulation validity, we supplement the original experiments with a set of comparative tests under non-ideal CSI scenarios. Specifically, two key modifications are implemented to simulate practical wireless channel characteristics:

To further simulate the channel attenuation fluctuation in non-ideal scenarios, we modify the path loss calculation models by adding a random disturbance term in the range of  $-2$  to  $2$  dB for both the LoS and two-ray propagation loss functions [5]. The modified models are expressed as follows:

$$L_s = 32.5 + 20 \lg f + 10n_1 \lg R + \delta \quad (12)$$

$$L_d = 120 + 10n_2 \lg R' - 20 \lg h_t - 20 \lg h_r + \delta \quad (13)$$

where  $\delta \sim U(-2, 2)$  denotes the uniform random disturbance term for simulating non-ideal channel attenuation fluctuations.

Meanwhile, we adjust the algorithm constraint conditions: the benchmark algorithms (DQN, DDQN, PMABCA, SRSA, GA) are not forced to adopt fixed power levels and full jammer activation, while our proposed TLOA retains the advantages of continuous power control and adaptive jammer selection.

The performance comparison results of six algorithms under non-ideal CSI scenarios are shown in Table 5. The evaluation metrics include three core indicators: first decision time, each decision time, and system operating duration.

**Table 5.** Algorithm performance comparison under non-ideal CSI.

| Algorithm    |                           | TLOA    | DQN    | DDQN   | PMABCA   | SRSA    | GA       |
|--------------|---------------------------|---------|--------|--------|----------|---------|----------|
| Time Average |                           |         |        |        |          |         |          |
| Scenario 1   | first decision            | 13.9 ms | 10 s   | 10 s   | 27.4 ms  | 42.1 ms | 164 ms   |
|              | each decision             | 10 ms   | 1 s    | 1 s    | 23.2 ms  | 38.4 ms | 148.3 ms |
|              | system operating duration | 0.05 H  | 0.17 H | 0.16 H | 0.22 H   | 0.17 H  | 0.16 H   |
| Scenario 2   | first decision            | 12.7 ms | 10 s   | 10 s   | 16.3 ms  | 35.6 ms | 126 ms   |
|              | each decision             | 9.2 ms  | 1 s    | 1 s    | 14 ms    | 26.1 ms | 85.9 ms  |
|              | system operating duration | 0.05 H  | 0.22 H | 0.24 H | 0.47 H   | 0.25 H  | 0.22 H   |
| Scenario 3   | first decision            | 13.6 ms | 10 s   | 10 s   | 17.5 ms  | 37.3 ms | 143 ms   |
|              | each decision             | 13.4 ms | 1 s    | 1 s    | 14.7 ms  | 25.9 ms | 96.6 ms  |
|              | system operating duration | 0.05 H  | 0.25 H | 0.27 H | 0.85 H   | 0.32 H  | 0.26 H   |
| Scenario 4   | first decision            | 21 ms   | 10 s   | 10 s   | 142.4 ms | 184 ms  | 1.7 s    |
|              | each decision             | 18 ms   | 1 s    | 1 s    | 110.6 ms | 170 ms  | 422 ms   |
|              | system operating duration | 0.05 H  | 0.01 H | 0.01 H | 0.18 H   | 0.09 H  | 0.01 H   |

It should be clarified that the original design of all involved algorithms (including TLOA, DQN, DDQN, PMABCA, SRSA, and GA) is based on the assumption of ideal CSI. Under non-ideal CSI scenarios with  $\pm 2$  dB power estimation error and channel attenuation disturbance, the performance of all algorithms shows varying degrees of degradation

compared with the ideal CSI scenario. This indicates that non-ideal CSI is a common challenge for interference optimization algorithms in wireless communication systems, rather than a defect specific to the proposed TLOA. Despite the performance degradation caused by non-ideal CSI, TLOA still maintains significant advantages over all benchmark algorithms in terms of decision latency and system operating stability.

**Decision Latency:** In all four test scenarios, TLOA's first decision time (12.7 ms–21 ms) and each decision time (9.2 ms–18 ms) are 2–3 orders of magnitude lower than those of DQN and DDQN (1 s–10 s), and 1.2–8.1 times lower than those of PMABCA, SRSA, and GA.

**System Operating Duration:** TLOA maintains a stable operating duration of 0.05 H across all scenarios, which is lower than most benchmark algorithms (0.09 H–0.85 H). Even in Scenario 4 where DQN, DDQN, and GA achieve short operating durations (0.01 H), their decision latency is still far higher than TLOA, which means they sacrifice decision efficiency for shorter operating time.

The core reason for TLOA's superiority is its continuous power control and adaptive jammer selection mechanism, which avoids the performance bottleneck caused by fixed power levels and full jammer activation in benchmark algorithms.

#### 4. Conclusions and Future Prospects

This paper has established mathematical models for air–ground communication networks and air–ground joint communication jamming. It has also designed cognition and jamming effect estimation strategies to reduce the dimensionality of jamming decisions and evaluate jamming effects based on jammer reconnaissance information. The TLOA has been proposed to schedule the number of jammers, jamming power, and operation time slots by leveraging cognitive communication node information, eliminating redundant operations of stochastic search and inefficient training processes of reinforcement learning to accelerate algorithmic speed.

Simulation results demonstrate that compared with DQN, DDQN, stochastic search algorithms, genetic algorithms, and improved artificial bee colony algorithms, the TLOA exhibits superior capabilities in finding feasible solutions, faster jamming decision-making, and efficient resource allocation by exploiting propagation path loss differences. This approach conserves jamming power and extends system operation time.

In the future, it is possible to study the estimation strategies of the jamming path loss under complex terrains and the noise in the inhomogeneous electromagnetic environment [66,67]. In addition, in terms of communication countermeasure games, when the communication nodes are unknown, it is important to study the cognitive decision-making of the jammers to gradually detect and estimate the information such as the positions and types of the communication nodes [68], and study the multi-agent strategies when the communication network between the jammers is disconnected. It is possible to combine the distributed communication countermeasure scenarios, design strategies to reduce the decision-making dimension, give full play to the advantages of reinforcement learning algorithms in game confrontation, and introduce intelligent algorithms such as deep learning and image recognition to enrich the methods for scheduling jamming resources [69,70].

This study admits that the performance of the proposed algorithm is affected by non-ideal CSI, which is a key limitation of the current work. In practical applications, the inaccuracy of CSI measurement will directly lead to the deviation of power control and jammer selection decisions. Therefore, the design of interference optimization algorithms robust to non-ideal CSI will be the core direction of our future research. The follow-up work will focus on integrating CSI estimation error compensation mechanisms into the

algorithm framework, and verifying the algorithm performance through hardware-in-the-loop experiments to further improve the practical value of the research.

**Author Contributions:** Z.W. proposed a design plan; W.W. and Z.Z. conducted experimental simulations; Z.W. and W.W. wrote the manuscript; J.Z. provided guidance on the experiments and performed the final revision of the manuscript completed by Z.W. and W.W.; J.Z., C.L., S.Z. and H.Y. provided guidance on the experiments and performed the final revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the General Project of Shaanxi Provincial Natural Science Basic Research Program, grant number 2025JC-YBMS-730.

**Data Availability Statement:** The data that support the findings of this study are available on GitHub via <https://github.com/hellogoodstudents/TLOA> (accessed on 19 January 2026). Other data related to this study are available from the corresponding author.

**Acknowledgments:** The authors would like to thank the reviewers for their valuable comments and suggestions that helped improve the quality of this manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Jornet, J.M.; Knightly, E.W.; Mittleman, D.M. Wireless communications sensing and security above 100 GHz. *Nat. Commun.* **2023**, *14*, 841.
2. Shrestha, R.; Guerboukha, H.; Fang, Z.; Knightly, E.; Mittleman, D.M. Jamming a terahertz wireless link. *Nat. Commun.* **2022**, *13*, 3045.
3. Kong, Z.; Cui, J.; Ding, L.; Huang, T.; Yan, S. Jamming precoding in AF relay-aided PLC systems with multiple eavesdroppers. *Sci. Rep.* **2024**, *14*, 8335.
4. Al-Turjman, F.; Hamamreh, J.M. Security in UAV/Drone Communications. In *Drones in IoT-Enabled Spaces*; CRC Press: Boca Raton, FL, USA, 2019; pp. 189–205.
5. Hamamreh, J.M.; Furqan, H.M.; Arslan, H. Classifications and Applications of Physical Layer Security Techniques for Confidentiality: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 1773–1828.
6. Hamamreh, J.M.; Arslan, H. Joint PHY/MAC Layer Security Design Using ARQ with MRC and Null-Space Independent, PAPR-Aware Artificial Noise in SISO Systems. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6190–6204.
7. Kırık, M.; Hamamreh, J.M. A Novel Interference Signal Superposition Algorithm for Providing Secrecy to Subcarrier Number Modulation-Based Orthogonal Frequency Division Multiplexing Systems. *Trans. Emerg. Telecommun. Technol.* **2022**, *33*, e4678.
8. Hamamreh, J.M.; Madni, F. Adaptable Secure Communication Framework for Automobile Intelligent Transportation Systems. *RS Open J. Innov. Commun. Technol.* **2024**, *4*, 11.
9. Furqan, H.M.; Hamamreh, J.M.; Arslan, H. New Physical Layer Key Generation Dimensions: Subcarrier Indices/Positions-Based Key Generation. *IEEE Commun. Lett.* **2020**, *24*, 2674–2678.
10. Hamamreh, J.M.; Furqan, H.M. A New Scheme for Improving Channel-Based Secret Key Generation Rates. *IEEE Wirel. Commun. Lett.* **2024**, *13*, 3133–3136.
11. Wang, X.; Huang, T.; Liu, Y. Resource allocation for random selection of distributed jammer towards multistatic radar system. *IEEE Access* **2021**, *9*, 29048–29055.
12. Li, S.; Liu, G.; Zhang, K.; Qian, Z.; Ding, S. DRL-Based Joint Path Planning and Jamming Power Allocation Optimization for Suppressing Netted Radar System. *IEEE Signal Process. Lett.* **2023**, *30*, 548–552.
13. Zhang, D.; Sun, J.; Yi, W.; Yang, C.; Wei, Y. Joint Jamming Beam and Power Scheduling for Suppressing Netted Radar System. In *2021 IEEE Radar Conference (RadarConf21)*, Atlanta, GA, USA, 8–14 May 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.
14. Lu, D.J.; Wang, X.; Wu, X.T.; Chen, Y. Adaptive allocation strategy for cooperatively jamming netted radar system based on improved cuckoo search algorithm. *Def. Technol.* **2023**, *24*, 285–297.

15. Xin, Q.; Xin, Z.; Chen, T. Cooperative Jamming Resource Allocation with Joint Multi-Domain Information Using Evolutionary Reinforcement Learning. *Remote Sens.* **2024**, *16*, 1955.
16. Yao, Z.; Tang, C.; Wang, C.; Shi, Q.; Yuan, N. Cooperative jamming resource allocation model and algorithm for netted radar. *Electron. Lett.* **2024**, *58*, 834–836.
17. Jin, W.-C.; Kim, K.; Choi, J.-W. Adaptive Jamming Considering Location Information Inaccuracy for Anti-UAV System. In *2021 International Conference on Information Networking, Jeju Island, Republic of Korea, 13–16 January 2021*; IEEE: New York, NY, USA, 2021; pp. 480–482.
18. Xiong, M.; Zhuo, J.; Dong, Y.; Jing, X. A layout strategy for distributed barrage jamming against underwater acoustic sensor networks. *J. Mar. Sci. Eng.* **2020**, *8*, 252.
19. Wu, Z.; Luo, Y.; Hu, S. Optimization of jamming formation of USV offboard active decoy clusters based on an improved PSO algorithm. *Def. Technol.* **2024**, *32*, 529–540.
20. Wang, H. Research on Anti-Jamming Strategy of IRS Communication Pair under Information Uncertainty. Master's Thesis, Nanjing University of Posts and Telecommunications, Nanjing, China, 2022.
21. Wu, L.; Wang, W.; Ji, Z.; Yang, Y.; Cumanan, K.; Chen, G.; Dobre, O.A. UAV-assisted maritime legitimate surveillance: Joint trajectory design and power allocation. *IEEE Trans. Veh. Technol.* **2023**, *72*, 13701–13705.
22. Wei, Z.; Wu, W.; Zhan, J.; Zhang, Z. Distributed communication interference resource scheduling using the master-slave parallel scheduling genetic algorithm. *Sci. Rep.* **2025**, *15*, 3431.
23. Tang, C.; Ding, J.; Zhang, L. LEO satellite downlink distributed jamming optimization method using a non-dominated sorting genetic algorithm. *Remote Sens.* **2024**, *16*, 1006.
24. Amuru, S.; Buehrer, R.M. Optimal Jamming Against Digital Modulation. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 2212–2224.
25. Yao, Z.; Liu, T.; Wang, C. Cooperative jamming resource allocation model based on the improved firefly algorithm. In *Proceedings of the 6th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 21–23 October 2022*; Association for Computing Machinery: New York, NY, USA, 2023; pp. 1719–1724.
26. Ke, L. Research on Communication and Jamming Integrated Resource Allocation Based on Intelligent Optimization Algorithm. Master's Thesis, Xidian University, Xi'an, China, 2023.
27. Wu, T.; Zou, Q.; Yang, Y.; Zhang, X.; Liu, S. A hierarchical comb interference resource allocation algorithm based on greedy strategy and evolutionary algorithm. In *2022 7th International Conference on Intelligent Computing and Signal Processing, Xi'an, China, 15–17 April 2022*; IEEE: New York, NY, USA; pp. 299–303.
28. Wu, W.; Wei, Z.; You, H.; Li, X.; Zhan, J.; Zhang, Z. EPRSA: Interference resource scheduling algorithms for air-ground communication networks. *Sci. Rep.* **2025**, *15*, 28436.
29. Fan, K. Knowledge Based Communication Jamming Decisions. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2022.
30. Deng, Q.; Ge, Y.; Ding, Z. Jamming Suppression via Resource Hopping in High-Mobility OTFS-SCMA Systems. *IEEE Wirel. Commun. Lett.* **2023**, *12*, 2138–2142.
31. Wang, S.; et al. The Future Is Calling. In *Cognitive Electronic Warfare: The Intelligent Game in the Electromagnetic Space*, 1st ed.; Science Press: Beijing, China, 2024; pp. 89–122. (In Chinese)
32. Sharma, P.; Sarma, K.K.; Mastorakis, N.E. Artificial Intelligence Aided Electronic Warfare Systems—Recent Trends and Evolving Applications. *IEEE Access* **2020**, *8*, 224761–224780.
33. Liu, Z.; Wang, X.; Kang, W.; Chen, Y. Research on multi-UAV collaborative electronic countermeasures effectiveness method based on CRITIC weighting and improved gray correlation analysis. *AIP Adv.* **2024**, *14*, 045123.
34. Xiang, P.; Hua, X.; Lei, J.; Yue, Z.; Ning, R.A. Dynamic Adaptive Jamming Power Allocation Method Based on Deep Reinforcement Learning. *Acta Electron. Sin.* **2023**, *51*, 1223–1234.
35. Zhao, C.; Wang, Q.; Liu, X.; Li, C.; Shi, L. Reinforcement learning based a non-zero-sum game for secure transmission against smart jamming. *Digit. Signal Process.* **2021**, *112*, 103002.
36. Rao, N.; Xu, H.; Jiang, L.; Song, B.; Shi, Y. Allocation Algorithm of Distributed Cooperative Jamming Power Based on Multi-Agent Deep Reinforcement Learning. *Acta Electron. Sin.* **2022**, *50*, 1319–1330.
37. Amuru, S.; Tekin, C.; van der Schaar, M.; Buehrer, R.M. Jamming Bandits—A Novel Learning Method for Optimal Jamming. *IEEE Trans. Wirel. Commun.* **2015**, *15*, 2792–2808.
38. ZhuanSun, S.; Yang, J.-A.; Liu, H.; Huang, K. A novel jamming strategy-greedy bandit. In *2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), Guangzhou, China, 6–8 May 2017*; IEEE: New York, NY, USA, 2017; pp. 1142–1146.

39. ZhuanSun, S.; Yang, J.A.; Liu, H. An algorithm for jamming strategy using OMP and MAB. *EURASIP J. Wirel. Commun. Netw.* **2019**, *2019*, 85.
40. ZhuanSun, S.; Yang, J.; Liu, H.; Huang, K. An Algorithm for Jamming Decision Using Dual Reinforcement Learning. *J. Xi'an Jiaotong Univ.* **2018**, *52*, 63–69.
41. Zhou, C.; Ma, C.; Lin, Q.; Man, X.; Ying, T. Intelligent Bandit Learning for Jamming Strategy Generation. *Wirel. Netw.* **2023**, *29*, 2391–2403.
42. Rao, N.; Xu, H.; Zhang, Y.; Wang, D.; Jiang, L.; Peng, X. Joint Optimization of Jamming Link and Power Control in Communication Countermeasures: A Multiagent Deep Reinforcement Learning Approach. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7962686.
43. Li, X.; Cui, Q.; Zhao, B.; Zhang, X.; Jiang, B.; Tao, X. Distributed Multi-Agent Interference Coordination in Native AI Enabled Multi-Cell Networks for 6G. In *2023 26th International Symposium on Wireless Personal Multimedia Communications (WPMC), Tampa, FL, USA, 16–19 November 2023*; IEEE: New York, NY, USA, 2023; pp. 8–13.
44. Li, F.; Xiong, J.; Zhao, X.; Zhao, H.; Wei, J.; Su, M. Wireless Communications Interference Avoidance Based on Fast Reinforcement Learning. *J. Electron. Inf. Technol.* **2022**, *44*, 3842–3849.
45. Niu, Y.; Wan, B.; Chen, C. A Centralized Multi-User Anti-Composite Intelligent Interference Algorithm Based on Improved Q-Learning. *Electronics* **2023**, *12*, 1735.
46. Mobini, Z.; Chalise, B.K.; Mohammadi, M.; Suraweera, H.A.; Ding, Z. Proactive Eavesdropping Using UAV Systems with Full-Duplex Ground Terminals. In *2018 IEEE International Conference on Communications Workshops, Kansas City, MO, USA, 20–24 May 2018*; IEEE: New York, NY, USA, 2018; pp. 1–6.
47. Tom, V. *Reinforcement Learning: The Markov Decision Process Approach*; MIT Press: Cambridge, MA, USA, 2021; pp. 133–152.
48. Yang, H.; Zhang, J. Research on Intelligent Interference Algorithm Based on Reinforcement Learning. *Electron. Meas. Technol.* **2018**, *41*, 49–54.
49. Martin, A.; Anders, H. *Reinforcement Learning*; Wiley: Hoboken, NJ, USA, 2023; pp. 327–349.
50. Zhou, C.; Lin, X.; Ma, S.; Ying, T.; Man, X. Intelligent Decision-Making for Selection of Communication Jamming Channel and Power. *J. Electron. Inf. Technol.* **2024**, *46*, 3957–3965.
51. Ning, R.; Hua, X.; Zisen, Q.; Song, B.; Yujhao, S. Allocation Method of Communication Interference Resource Based on Deep Reinforcement Learning of Maximum Policy Entropy. *J. Northwestern Polytech. Univ.* **2021**, *39*, 1077–1086.
52. Peng, X.; Xu, H.; Jiang, L.; Rao, N.; Song, B. A Deep Reinforcement Learning Communication Jamming Resource Allocation Algorithm Fused with Noise Network. *J. Electron. Inf. Technol.* **2023**, *45*, 1043–1054.
53. Ning, R.; Hu, X.; Jialin, S.S. Q-Learning Intelligent Jamming Decision Algorithm Based on Efficient Upper Confidence Bound Variance. *J. Harbin Inst. Technol. Engl. Ed.* **2022**, *54*, 162–170.
54. Hua, X.; Bailin, S.; Lei, J.; Ning, R.; Yunhao, S. An Intelligent Decision-Making Algorithm for Communication Countermeasure Jamming Resource Allocation. *J. Electron. Inf. Technol.* **2021**, *43*, 3086–3095.
55. Bailin, S.; Hux, X.; Zisen, Q.; Ning, R.; Peng, P. A Collaborative Communication Jamming Decision Algorithm Based on Deep Reinforcement Learning. *Acta Electron. Sin.* **2022**, *50*, 1301–1309.
56. Liu, Y.F.; Li, X.S.; Yang, J.A.; Yang, D.J.; Wang, J. SANER-PPO Algorithm-Based Jamming Resource Allocation for UAV Swarm. *Control Decis.* **2024**, *39*, 3937–3945.
57. Fu, X.; Yan, H. Neural Network Optimal Control for Tripartite UAV Confrontation Systems Based on Fuzzy Differential Game. *Sci. Rep.* **2024**, *14*, 21547.
58. Wu, Z.; Pan, L.; Yu, M.; Liu, J.; Mei, D. A Game-Based Approach for Designing a Collaborative Evolution Mechanism for Unmanned Swarms on Community Networks. *Sci. Rep.* **2022**, *12*, 18892.
59. Liu, Y.; Mao, H.; Zhu, L.; Xiao, Z.; Han, Z.; Xia, X.-G. Routing and Resource Scheduling for Air-Ground Integrated Mesh Networks. *IEEE Trans. Wirel. Commun.* **2022**, *22*, 4090–4105.
60. Russell, S. AI Weapons: Russia's War in Ukraine Shows Why the World Must Enact a Ban. *Nature* **2023**, *614*, 620–623.
61. Frchetti, M.D.; Berner, J.; Liu, X.; Henry, E.R.; Maksudov, F.; Ju, T. Large-Scale Medieval Urbanism Traced by UAV–LiDAR in Highland Central Asia. *Nature* **2024**, *634*, 1118–1124.
62. Gao, Y.; Liu, M.; Yuan, X.; Hu, Y.; Sun, P.; Schmeink, A. Federated Deep Reinforcement Learning Based Trajectory Design for UAV-Assisted Networks with Mobile Ground Devices. *Sci. Rep.* **2024**, *14*, 22753.
63. Liao, Y.; Gao, G.; Jing, Y. Ultra-Reliable Intelligent Link Scheduling Based on DRL for Manned/Unmanned Aerial Vehicle Cooperative Scenarios. *Phys. Commun.* **2024**, *63*, 102304.

64. Li, Q.; Cheng, H.; Yang, Y.; Tang, H.; Wang, J.; Luo, G.; Sun, W. Deep Reinforcement Learning-Based Resource Allocation for 5G Machine-Type Communication in Active Distribution Networks with Time-Varying Interference. *Mob. Netw. Appl.* **2022**, *27*, 2264–2279.
65. Tossa, F.; Abdou, W.; Ansari, K.; Ezin, E.C.; Gouton, P. Area Coverage Maximization under Connectivity Constraint in Wireless Sensor Networks. *Sensors* **2022**, *22*, 1712.
66. Gu, J.; Ding, G.; Wang, H.; Xu, Y. Integrated Communications and Jamming: Toward Dual-Functional Wireless Networks Under Antagonistic Environment. *IEEE Commun. Mag.* **2023**, *61*, 181–187.
67. Gu, J.; Ding, G.; Wang, H.; Xu, Y. Sensing Assisted Integrated Communication and Jamming Systems with RSMA for Dynamic Suspicious Communications. *IEEE Trans. Veh. Technol.* **2024**, *73*, 5965–5970.
68. Rao, N.; Xu, H.; Wang, D.; Qi, Z.; Zhang, Y.; Gu, W.; Peng, X. Efficient Jamming Resource Allocation Against Frequency-Hopping Spread Spectrum in WSNs with Asynchronous Deep Reinforcement Learning. *IEEE Sens. J.* **2024**, *24*, 13560–13577.
69. Zhenhua, W.; Jianwei, Z.; Siming, H. Introduction. In *Air-Ground Joint Distributed Communication Jamming Technology and Practice*, 1st ed.; National Defense Industry Press: Xi'an, China, 2023; pp. 1–5.
70. Hua, X.; Jun, W.; Lei, J. Communication Interference Technologies and Methods. In *Principles and Applications of Modern Communication Countermeasures*, 1st ed.; National Defense Industry Press: Beijing, China, 2022; pp. 403–411.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.