drones

MDPI

Article

A Machine Learning-Based Hybrid Approach for Safeguarding VLC-Enabled Drone Systems

Ge Shi [1,*,†], Hongyang Zhou [1,†], Huixin Wu [1,*], Fupeng Wei [1] and Wei Cheng [2]

[1] The School of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450046, China; X20231090799@stu.ncwu.edu.cn (H.Z.); weifupeng@ncwu.edu.cn (F.W.)
[2] The School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China; pupil_119@nwpu.edu.cn
* Correspondence: shige@ncwu.edu.cn (G.S.); wuhuixin@ncwu.edu.cn (H.W.)
† These authors contributed equally to this work.

Abstract

This paper explores the physical layer security performance of collaborative drone fleets enabled by visible light communication (VLC) in a multi-eavesdropper scenario, where multiple drones leverage VLC to serve terrestrial users. To strengthen system security, we formulate a sum worst-case secrecy rate maximization problem. To address the non-convex optimization challenge of this problem, we develop two innovative Q-learning-based position decision algorithms (Q-PDA and Q-PDA-lite) with a dynamic reward mechanism, allowing drones to adaptively optimize their positions. Additionally, we propose an enhanced Tabu Search-based grouping algorithm (TS-GA) to establish the suboptimal user equipment (UE)–drone association by balancing candidate solution exploration and tabu constraint exploitation. Simulation results demonstrate that the proposed Q-PDA and Q-PDA-lite achieve worst-case secrecy rates significantly exceeding those of Random-PDA and K-means-PDA. While Q-PDA-lite exhibits 2% lower performance than Q-PDA, it offers reduced complexity. Additionally, the proposed TS-GA achieves a worst-case secrecy rate that substantially outperforms random grouping, UE-channel-gain-based grouping, and channel-gain-based grouping. Collectively, the hybrid approach integrating Q-PDA and TS-GA achieves 10% near-global optimality with guaranteed convergence, while preserving computational efficiency. Furthermore, this hybrid approach outperforms other combinations in terms of security metrics.

Keywords: unmanned aerial vehicle; physical layer security; Q-learning

check for updates

Academic Editor: Andrey V. Savkin

Received: 23 July 2025
Revised: 7 October 2025
Accepted: 11 October 2025
Published: 16 October 2025

Citation: Shi, G.; Zhou, H.; Wu, H.; Wei, F.; Cheng, W. A Machine Learning-Based Hybrid Approach for Safeguarding VLC-Enabled Drone Systems. *Drones* **2025**, *9*, 721. https://doi.org/10.3390/drones9100721

Copyright: © 2025 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) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Drone fleets have emerged as promising candidates for aerial base stations in future wireless networks, leveraging their flexible deployment and collaborative communication capabilities [1]. With advancements in 5G and 6G technologies, drones can achieve higher data rates and lower latency, significantly enhancing their operational efficiency [2]. This advantage becomes particularly critical in scenarios where traditional terrestrial base stations face network congestion or high traffic demands. However, as drone swarm size increases, conventional radio frequency (RF) communication technologies encounter two major challenges: susceptibility to interference from other devices and vulnerability to signal interception or eavesdropping.

VLC presents a promising solution to address these issues, enabling high-speed data transmission through the high-frequency flickering of LEDs [3]. Drones equipped with

VLC technology exhibit lower sensitivity to RF interference, and their communication is primarily affected by line-of-sight (LoS) fading. In outdoor environments, the impact of non-line-of-sight (NLoS) VLC-channels is largely negligible, thereby minimizing the scattering or reflection through unintended and non-visible routes [4]. Furthermore, VLC-enabled drones can flexibly adjust the coverage of communication cells, thereby delivering secure services to specific groups of UEs [5]. However, if a ground-based eavesdropper infiltrates the user equipments' group, the LoS properties of air-to-ground channel links improves the quality of interception, posing a security threat to VLC-enabled drones.

In recent years, physical layer security (PLS) has emerged as a promising technique to protect drone-aided communications from malicious eavesdroppers [6]. Unlike traditional encryption methods that rely on cryptographic keys, PLS leverages the inherent characteristics of wireless channels to enhance security [7]. Existing PLS strategies primarily utilize friendly jamming [8], UE–drone association [9], trajectory optimization [10], and power allocation [11] to safeguard drone fleets. Specifically, Ref. [8] proposed a deep reinforcement learning-based friendly jammer, which employs deep convolutional neural networks to degrade the interception capability of eavesdroppers. An efficient cooperative data dissemination algorithm was developed in [9] to maximize the minimum amount of data received by all drones through optimizing multi-UE scheduling, association, bandwidth allocation, and drone mobility. Ref. [10] employed a multi-agent deep deterministic policy gradient algorithm to maximize secure capacity by jointly optimizing drone trajectories, transmit power, and jamming power. A suboptimal power strategy for jammers was designed by using successive convex approximation, aiming to maximize the minimum average secrecy rate [11].

Researches on VLC-enabled drones have recently attracted significant attention. Ref. [12] proposed a multi-objective optimization problem that jointly maximizes the sum-rate and rate fairness of UEs using particle swarm optimization algorithm. In [1], an algorithm that combines a machine learning algorithm of gated recurrent units with convolution neural networks was proposed to jointly optimize drone deployment, user association, and power efficiency while meeting the illumination and communication requirements. In [13], a Harris Hawks optimization-based algorithm was proposed to maximize the sum rate of a VLC-enabled drone system using non-orthogonal multiple-access. Ref. [14] jointly optimized the drone's height and peak optical intensity for a weather-dependent covert VLC system. However, there are two key limitations in the aforementioned literature: First, most studies investigate simplified yet impractical models, such as single-drone systems, single-eavesdropper (Eve) systems, or scenarios with a limited number of UEs. Second, few works have explored the security of VLC-enabled drone systems.

To the best of our knowledge, this work is the first attempt to study the joint optimization of drone position and user association against eavesdropping in a VLC-enabled drone systems. In this work, we consider a multi-UE VLC-enabled multi-drone systems, where UEs are overheard by multiple ground eavesdroppers, and downlink transmissions between the drone and UEs follow the TDMA principle. Contributions offered by our work can be summarized as follows:

1. We formulate an optimization problem to maximize the sum of the worst-case secrecy rates of UEs, subject to a UE–drone association constraint. This maximization problem consists of two core subproblems: drone position adjustment and UE–drone association.
2. To address this problem, we first propose a Tabu Search (TS)-based grouping algorithm (TS-GA) to solve the UE–drone association subproblem, and then propose a Q-learning-based position decision algorithm (Q-PDA) to solve the drone position adjustment subproblem. TS-GA first guides the training process of Q-PDA by calcu-

lating initial rewards for drones in each state, and then associates UEs once the drone positions are determined.

3. We evaluate the probability that Q-PDA and TS-GA find the optimal solution for their respective subproblems, respectively.
4. Simulation results demonstrate that the performance of the proposed Q-PDA and Q-PDA-lite exceeds those of Random-PDA and K-means-PDA. Similarly, the proposed TS-GA outperforms random grouping, UE-channel-gain-based grouping, and channel-gain-based grouping. Finally, the proposed hybrid approach achieves near-global optimality compared with other combinations of position decision algorithms and association strategies.

The remainder of this paper is structured as follows: Section 2 presents the system model and problem formulation, Section 3 describes the proposed algorithms, Section 4 discusses the simulation results, and Section 5 concludes the study.

2. System Model

Consider a wireless network consisting of a set $\mathcal{D}$ with D LED-equipped drones. These drones serve a set $\mathcal{U}$ of U UEs that are randomly distributed across a geographical area $\mathcal{A}$. As illustrated in Figure 1, the drones are designed to simultaneously provide downlink communication services and illumination using VLC technique to the ground area. Within this area, multiple eavesdroppers (Eves) attempt to eavesdrop on user equipments (UEs) and intercept confidential data transmitted from drones that associated with the UEs.

In this model, we assume that as long as a UE's position lies within the illumination coverage of the drones, the drones are capable of obtaining the spatial location information of these ground UEs. Moreover, this location information can be shared among the drones through an internal communication mechanism.

When the drones transmit confidential information, a one-to-many relationship exists between the drones and the associated UEs. Specifically, a single drone adopts an equal time allocation to send signals to the associated UEs. Conversely, each UE is restricted to associate only one drone from which to receive confidential information.

We assume that each drone does not commence serving the ground users until it has maneuvered to its optimal position. During the process of wireless transmission, for the purpose of analysis, the drones can be regarded as stationary aerial base stations, simplifying the consideration of communication interactions within the network.

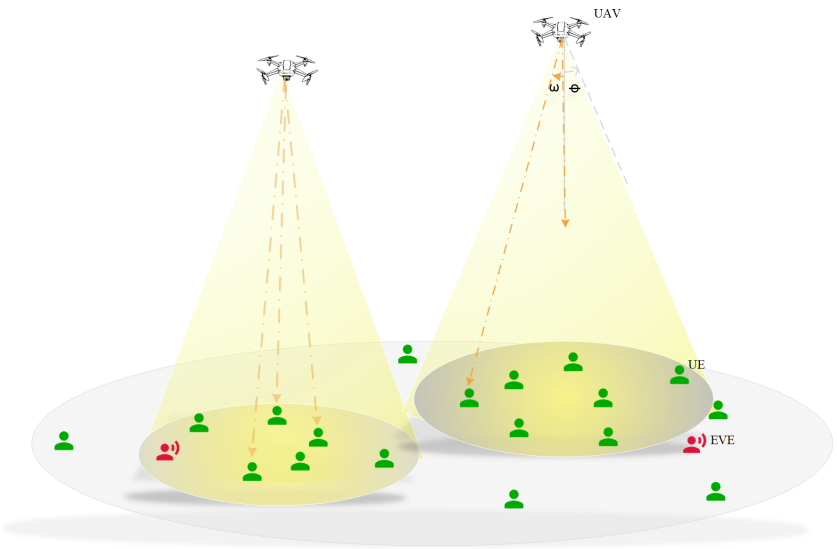


Figure 1. The architecture of a wireless network that consists of drones and users.

2.1. Transmission Model

Given a drone $d \in \mathcal{D}$ located at $\mathbf{q}_d = (x_d, y_d, H_d)$ and a ground UE $u \in \mathcal{U}$ located at $(a_u, b_u, H_u) \in \mathcal{A}$, the probabilistic LoS and NLoS channel model is used to model the VLC link between the drone and the ground UE. A malicious e-th Eve intends to overhear the confidential information of u-th UE, which is located at $(a_e, b_e, H_e) \in \mathcal{A}$.

For simplicity, we do not consider the diffusion of visible light in outdoor environments. Therefore the LoS and NLoS channel gain of the VLC link between the d-th drone to the u-th UE can be expressed as [15]

$$h_{d,u}^{LoS}(\mathbf{q}_d) = \frac{(\gamma+1)A_R}{2\pi v_{d,u}^2(\mathbf{q}_d)} \cos(\phi_{d,u})^\gamma \cos(\varphi_{d,u}) f(\varphi_{d,u}) g_{\text{of}}, \tag{1}$$

$$h_{d,u}^{NLoS}(\mathbf{q}_d) = 0, \tag{2}$$

where $\gamma = \frac{-1}{\log_2 \cos \phi_{1/2}}$ denotes the Lambertian emission order, and $\phi_{1/2}$ and A_R are the half-intensity radiation angle and the area of drone's photodiode (PD), respectively. $v_{d,u}$ is the Euclidean distance from the d-th drone to u-th UE, and $\phi_{d,u}$ and $\varphi_{d,u}$ are the irradiance angle and incidence angle, respectively. These two angles are equal if both the transmitter and receiver are horizontal. g_{of} is the gain of the optical filter, and $f(\varphi_{d,u})$ denotes the gain of the optical concentrator which is given by [16]

$$\mathrm{f}(\varphi_{d,u}) = \begin{cases} \frac{w^2}{\sin^2(\varphi_{d,u})} & 0 \le \varphi_{d,u} \le \Theta, \\ 0 & \varphi_{d,u} \ge \Theta, \end{cases} \tag{3}$$

with w and Θ as the refractive index and the field-of-view (FoV) of the PD used at the u-th UE's side, respectively. According to [1], the probability of the LoS link is $B(h_{d,u}^{\text{LoS}}(\mathbf{q}_d)) = (1 + X \exp(-Y[r_{d,u}(\mathbf{q}_d) - X])^{-1}$, where X and Y are environmental parameters and $r_{d,u}(\mathbf{q}_d) = \sin^{-1}(H/v_{d,u}(\mathbf{q}_d))$ is the elevation angle. The average channel gain from the d-th drone to the u-th UE can be given by

$$h_{d,u}(\mathbf{q}_d) = B(h_{d,u}^{\text{LoS}}(\mathbf{q}_d)) \times h_{d,u}^{\text{LoS}}(\mathbf{q}_d) + B(h_{d,u}^{\text{NLoS}}(\mathbf{q}_d)) \times h_{d,u}^{\text{NLoS}}(\mathbf{q}_d), \tag{4}$$

where $B(h_{d,u}^{\text{NLoS}}(\mathbf{q}_d)) = 1 - B(h_{d,u}^{\text{LoS}}(\mathbf{q}_d))$.

Define the binary association matrix $\mathbf{M}$ between drones and UEs. Its element $m_{d,u} \in \{0, 1\}$ indicates the association status, where $m_{d,u} = 1$ signifies that the d-th drone is paired with the u-th UE , and $m_{d,u} = 0$ otherwise. In our system model, not all UEs will establish connections with drones. This is because providing service to UEs close to Eve or at geographical corners can significantly compromise the system's security. The subset of UEs that are successfully associated with drones is denoted by $\widetilde{\mathcal{U}}$, where $\widetilde{\mathcal{U}} \subseteq \mathcal{U}$. Moreover, each UE within $\widetilde{\mathcal{U}}$ is restricted to be associated with exactly one drone $d \in \mathcal{D}$, so we have the following constraint:

$$\sum_{d \in \mathcal{D}} m_{d,u} = 1, \forall u \in \widetilde{\mathcal{U}}. \tag{5}$$

On the contrary, a drone is allowed to associate with multiple UEs, and the drone employs a uniform time allocation for associated UEs in turn, so that each UE receives its signal during an equal time resource. The number of UEs connected to a drone is not limited. The channel capacity of the u-th UE can be reformulated by

$$C_{d,u}(\mathbf{q}_d) = \frac{1}{2\sum_{u \in \widetilde{\mathcal{U}}} m_{d,u}} \log_2 \left(1 + \frac{|h_{d,u}(\mathbf{q}_d)|^2 P_T}{I_u(\mathbf{q}_d) + \xi^2}\right). \tag{6}$$

where P_T is the transmit power of each drone. The noise source inside the receiver's circuit is mainly dominated by thermal noise and shot noise. These are modeled as additive zero-mean Gaussian noise with the derivation of ξ^2. $I_u(\mathbf{q}_d)$ indicates the interference over the VLC link between the d-th drone and the u-th UE, which is defined as follows:

$$I_u(\mathbf{q}_d) = \sum_{j=1,\, j\neq d}^{D} \left|h_{j,u}(\mathbf{q}_j)\right|^2 P_T. \tag{7}$$

When the j-th drone is also inside u-th UE's reception range but not associated with the u-th UE, the signal transmitted from the j-th drone mixes up with the one from the d-th drone and causes interference to the u-th UE. When the j-th drone is outside the u-th UE's reception range, the channel gain between the j-th drone and the u-th UE is equal to zero, that is, $h_{j,u} = 0$.

For a static Eve $e \in \mathcal{E}$ overhearing the u-th UE, the channel capacity of the e-th Eve can be given by

$$C_{d,e}(\mathbf{q}_d) = \frac{1}{2\sum_{u\in\widetilde{\mathcal{U}}} m_{d,u}} \log_2\left(1 + \frac{\left|g_{d,e}(\mathbf{q}_d)\right|^2 P_T}{I_e(\mathbf{q}_d) + \xi^2}\right), \tag{8}$$

where $g_{d,e}(\mathbf{q}_d)$ denotes the channel gain of the VLC link between the d-th drone and the e-th Eve, ξ^2 denotes the zero-mean Gaussian noise, and $I_e(\mathbf{q}_d)$ indicates the interference over the VLC link between the d-th drone and the e-th Eve, which is defined as follows:

$$I_e(\mathbf{q}_d) = \sum_{j=1,\, j\neq d}^{D} \left|h_{j,e}(\mathbf{q}_j)\right|^2 P_T. \tag{9}$$

When a UE can be overheard by multiple Eves, the Eve with the strongest signal quality should be considered to calculate the worst-case secrecy rate of the UE, while the Eve with less signal quality can be ignored. The worst-case secrecy rate of the u-th UE is given by substituting (6) and (8) as follows:

$$\mathrm{SR}_u(\mathbf{q}_d) = \min[C_{d,u}(\mathbf{q}_d) - C_{d,e}(\mathbf{q}_d)]_+ = [C_{d,u}(\mathbf{q}_d) - \max C_{d,e}(\mathbf{q}_d)]_+. \tag{10}$$

2.2. Problem Formulation

Under the proposed system model, our objective is to identify an optimal strategy that maximizes the sum worst-case secrecy rate across all UEs. This strategy involves two intertwined aspects: drone position adjustment and UE–drone association. Drone position adjustment aims to maximize coverage while minimizing the risk of eavesdropping. Meanwhile, UE–drone association is designed to group drones with UEs while considering the existence of eavesdroppers and the bandwidth utilization. This problem can be formulated as follows:

$$\max_{\mathbf{q}_d, \mathbf{M}} \sum_{u\in\widetilde{\mathcal{U}}} \mathrm{SR}_u(\mathbf{q}_d), \tag{11}$$

$$\text{s.t.} \sum_{d\in\mathcal{D}} m_{d,u} \in \{0,1\}, \quad \forall u \in \widetilde{\mathcal{U}}. \tag{12}$$

The constraint implies that each UE is limited to being associated with exactly one drone. The aforementioned problem constitutes an integer programming problem, presenting two key challenges: its non-convex nature and a large discrete solution space.

3. Methodology

To address this problem, we first propose a Tabu Search (TS)-based grouping algorithm (TS-GA) to solve the UE–drone association subproblem, and then propose a Q-learning-based position decision algorithm (Q-PDA) to solve the drone position adjustment subproblem. Together, the hybrid algorithm TS-GA+Q-PDA solves the problem (11). In this section, we first introduce the proposed TS-GA, followed by the procedure of the Q-PDA and a simplified version of this algorithm.

Q-learning can directly navigate non-convex and discrete solution spaces through iterative learning of optimal policies via state–action value estimation. However, when directly applying Q-learning to scenarios where multiple drones serve a relatively large number of UEs, the UE–drone association subproblem significantly expands the Q-table size and increases training time—making it impractical for high-mobility scenarios. Thus, we first employed TS to solve the UE–drone association subproblem and also provide the UE–drone association results for the training process of Q-learning.

3.1. TS-Based Grouping Algorithm (TS-GA)

In the UE–drone association process, a TS-based grouping algorithm (TS-GA) is proposed to associate UEs and drones from the perspective of the UE. Denote the set of drones within the received signal range of the u-th UE as $\mathcal{U}_u$, where $u = 1, \cdots, U$. For example, $\mathcal{U}_1 = \{1, 3\}$ means that within the received signal range of 1-th UE, there are only two drones indexed 1 and 3, and it can only associate with either the 1-th or 3-th drones. Based on the set of drones, a near-optimal solution can be obtained by applying TS-GA. The basic idea of TS-GA is to use a tabu list to record the solutions that have been searched recently, preventing the algorithm from cycling back to these solutions and thus avoiding getting trapped in local optima. The generation rules of the initial solution, candidate solutions, tabu rules, and aspiration principles are designed as follows:

A. Initial Solution: The initial solution determines the starting point of the search. If the quality of the initial solution is high, the algorithm can conduct in-depth exploration near the high-quality solution from the beginning, accelerating the convergence of the algorithm.

A channel-gain grouping algorithm, which will be introduced in Section 4, is applied to match drones for all UEs according to $\mathcal{U}_i$ to generate a set of solutions. This set of solutions is used as the initial solution $\mathbf{x}(0) = \{d_1, d_2, d_u, \cdots, d_U\}$. d_u represents the index of the drone connected to u-th UE. If $d_u = 0$, it means that the u-th UE is not connected to any drone.

B. Candidate Solution Specification: The candidate solutions are a set of new solutions obtained by making small adjustments and transformations to a known basic solution according to specified rules, and are used to select a solution in a new round of iteration. This scheme inherits the local search ability of the Tabu Search, which is more conducive to in-depth exploration of potential high-quality solutions near the current high-quality solution.

During the l-th iteration, we assume that the grouping solution obtained in this iteration is denoted as $\mathbf{x}(l) = \{d_1, d_2, d_u, \cdots d_U\}$. We take $\mathbf{x}(l)$ as the basic solution for generating candidate solutions in the $(l+1)$ -th iteration. The number of candidate solutions is set in each iteration to be the same as the number of UEs U. The rule of candidate solution generation is designed as follows: switch a different drone for the u -th UE in the solution $\mathbf{x}(l)$ to form the u-th candidate solution. Specifically, for the u-th

candidate solution, we replace the drone connected to the u-th UE in the solution $x(l)$ with a drone randomly selected from the set $\mathcal{U}_u$, as follows:

$$\widetilde{\mathbf{x}}(l+1)_1 = \{d_1 = \text{random}(\mathcal{U}_1), d_2, d_u, \cdots d_U\}, \tag{13}$$

$$\widetilde{\mathbf{x}}(l+1)_2 = \{d_1, d_2 = \text{random}(\mathcal{U}_2), d_u, \cdots d_U\}, \tag{14}$$

$$\widetilde{\mathbf{x}}(l+1)_u = \{d_1, d_2, d_u = \text{random}(\mathcal{U}_u), \cdots d_U\}, \tag{15}$$

$$\cdots \tag{16}$$

$$\widetilde{\mathbf{x}}(l+1)_U = \{d_1, d_2, d_u, \cdots, d_U = \text{random}(\mathcal{U}_U)\}, \tag{17}$$

where $d_u = \text{random}(\mathcal{U}_u)$ means randomly selecting a connectable drone from the set $\mathcal{U}_u$ for the u-th UE. The matrix of the candidate solutions in the $(l+1)$-th iteration is $\widetilde{\mathbf{X}}(l+1) = [\widetilde{\mathbf{x}}(l+1)_1, \widetilde{\mathbf{x}}(l+1)_2, \cdots, \widetilde{\mathbf{x}}(l+1)_U]^{\mathrm{T}}$.

C. Tabu Rules: To prevent the algorithm from repeatedly choosing near the local optimum and falling into an infinite loop, we have set up a tabu list in the algorithm. The tabu list is used to record and store the solutions visited recently. When a set of solutions is visited, all the associations between UEs and drones are put into the tabu list for restriction to prevent from being selected again in the near future. A tabu period in this algorithm is set as T_p, and the tabu period defines the number of iterations during which the solutions in the tabu list are restricted. When the times for a solution to be banned reaches the tabu period, it will be released and can be selected again.

The tabu list $\mathbf{T} \in \mathcal{R}^{U \times D}$ and the tabu period interact with each other, which can effectively guide the search process and are important tools for achieving the global optimization of the tabu algorithm. $\mathbf{T}(i,j)$ represents whether the association between the i-th UE and the j-th drone is tabued. When initialized, $\mathbf{T}$ is a matrix of all 0s. In the process of algorithm, if $\mathbf{T}(i,j) = 0$, it means that this UE–drone association is not tabued. If $\mathbf{T}(i,j) > 0$, it means that this connection has been tabued and has not been released temporarily. A tabu times matrix $\mathbf{T}_{\text{time}} \in \mathcal{R}^{U \times D}$ is also set to record the number of times that each UE–drone connection is tabued. When a group of connections is tabued too many times, the algorithm stops iterating.

(1) Tabu Condition: If the drone and user are repeatedly connected within the tabu period, that connection becomes tabu. Any candidate solution with a tabu connection is prohibited from being selected, unless the Aspiration Criterion is satisfied. This prevents the algorithm from revisiting recently searched areas in the short term (tabu period T_p).

(2) Aspiration Criterion: The Aspiration Criterion serves as an override mechanism to the Tabu Condition. If a candidate solution includes tabu connections and yields a sum worst-case secrecy rate superior to the global best solution, the tabu status of these connections will be disregarded. In such cases, amnesty is granted to these connections, and the resulting candidate solution is selected as the new current solution, thus preventing the algorithm from missing potential improvements solely due to rigid tabu restrictions.

(3) Tabu Operation: In each iteration, when the algorithm selects a candidate solution $\mathbf{x}(l) = \{d_1, d_2, d_u, \cdots, d_U\}$ as the solution, we record the connections in this solution by updating the tabu list $\mathbf{T}$ and set its value as the tabu period. The equation is as follows:

$$\mathbf{T}(u, d_u) = T_p, \;\; \mathbf{T}_{\text{time}}(u, d_u) = \mathbf{T}_{\text{time}}(u, d_u) + 1, \;\; \forall u = 1, \cdots, U. \tag{18}$$

(4) Release Operation: After each iteration, all the banned associations in the tabu list $\mathbf{T}$ need to release one time, that is, decreased by 1. At this time, update the $\mathbf{T}$ matrix operation. The equation is as follows:

$$\mathbf{T}(i,j) = \max\{\mathbf{T}(i,j) - 1, 0\}, \;\; \forall i = 1, \cdots, U, j = 1, \cdots, D. \tag{19}$$

The procedure of the proposed algorithm is described in Algorithm 1.

Algorithm 1: TS-GA.

Input: d-th drone's position $q_d, \forall d = 1, ..., D$
Output: the grouping solution with the best secrecy performance $\mathbf{x}_{\max}$
Initialize $\mathbf{x}(0)$
for each iteration index $l \in [1, \text{max-time}]$ **do**
  Generate the Matrix of candidate solutions $\widetilde{\mathbf{X}}(l)$ according to $\mathbf{x}(l-1)$
  for each candidate solution index $u \in [1, U]$ **do**
    if $\widetilde{\mathbf{x}}(l)_u$ not meet the Tabu Condition **then**
      Execute Release Operation by using Equation (19).
      Execute Tabu Operation by using Equation (18).
      $\mathbf{x}(l) = \widetilde{\mathbf{x}}(l)_u$
      break
    else
      if $\widetilde{\mathbf{x}}(l)_u$ meet the Aspiration Criterion **then**
        Execute Release Operation by using Equation (19).
        Execute Tabu Operation by using Equation (18).
        $\mathbf{x}(l) = \widetilde{\mathbf{x}}(l)_u$
        $\mathbf{x}_{\max} = \widetilde{\mathbf{x}}(l)_u$
        break
      else
        Execute Release Operation by using Equation (19).
      end if
    end if
  end for
  if any element in $\mathbf{T}_{\text{time}}(u, d_u)$> Threshold **then**
    break
  end if
end for
return $\mathbf{x}_{\max}$

3.2. Theoretical Analysis

To quantify the probability of TS-GA finding the optimal solution for the UE–drone association subproblem, we first introduce a realistic assumption: high-quality solutions are locally connected, including the optimal solution. TS-GA search for the optimal solution proceeds in two sequential phases:

(1) Exploration phase: Assuming each iteration has an independent probability δ of transitioning into the high-quality solution set $\mathcal{S}_{\text{high}}$, the probability of not entering $\mathcal{S}_{\text{high}}$ in T_1 iteration is $(1-\delta)^{T_1}$. Thus, the probability of entering $\mathcal{S}_{\text{high}}$ within T_1 iterations is

$$P_1(\text{Enter in } T_1 \leq T) = 1 - (1-\delta)^{T_1}, \tag{20}$$

where T is the total number of iterations.

(2) Stability phase: Each iteration in this phase has a probability η of remaining in $\mathcal{S}_{\text{high}}$ and an average probability ρ of transitioning into the optimal solution from any high-quality solution in $\mathcal{S}_{\text{high}}$. Overall, the probability of finding the optimal solution in $\mathcal{S}_{\text{high}}$ for all T_2 iterations is

$$P_2(\text{Remain for } T_2) = \eta^{T_2}[1 - (1-\rho)^{T_2}], \tag{21}$$

where $\rho \gg \frac{1}{|\mathcal{S}_{\text{high}}|}$.

In conclusion, combining the two phases, the total probability of finding the optimal solution for the UE–drone association subproblem until iteration T is

$$P_{\text{TS}}(T) \geq \left[1-(1-\delta)^{T_1}\right] \cdot \eta^{T_2}[1-(1-\rho)^{T_2}]. \tag{22}$$

By setting the tabu period and the maximum iteration number in the TS-GA, we can have $\eta \to 1$ and $T \approx 1000$. In this case, even if $\delta \ll 1$ and $\rho \ll 1$, we still obtain $P_{\text{TS}}(T) \approx 1$.

3.3. Benchmark Algorithms

The benchmark algorithms for evaluating the performance of TS-GA are defined as follows:

Random grouping: Under random grouping, each UE randomly chooses one drone from those within its signal reception range. The probability of finding the optimal solution for the UE–drone association subproblem is

$$P_{\text{RG}} = \frac{1}{|\mathcal{U}_1|} * \frac{1}{|\mathcal{U}_2|} ... * \frac{1}{|\mathcal{U}_u|} * \frac{1}{|\mathcal{U}_U|}, \tag{23}$$

where $\frac{1}{|\mathcal{U}_u|}$ is the probability of finding the best association for the u-th UE, and $\mathcal{U}_u$ is the set that includes the drones within the reception range of the u-th UE.

UE-channel-gain grouping: Each UE selects the drone with the highest received channel gain as its fixed signal source. The probability of finding the optimal solution for the UE–drone association subproblem is given by

$$P_{\text{UG}} \approx 0. \tag{24}$$

In practice, the optimal solution rarely requires all UEs to choose the drone with the highest channel gain, since the strongest channel gain does not equate to the highest security. For instance, if an Eve is located near a drone, the Eve may also intercept the strongest signal, leading to a significant drop in the overall system security performance.

Channel-gain grouping: Channel-gain grouping is similar to the UE-channel-gain strategy, but only choose a drone for which the UE's received channel gain is greater than Eve's intercepted channel gain. The probability of finding the optimal solution for the UE–drone association subproblem is

$$P_{\text{CG}} = \frac{1}{|\hat{\mathcal{U}}_1|} * \frac{1}{|\hat{\mathcal{U}}_2|} ... * \frac{1}{|\hat{\mathcal{U}}_u|} * \frac{1}{|\hat{\mathcal{U}}_U|}, \tag{25}$$

where $\frac{1}{|\hat{\mathcal{U}}_u|}$ is the probability that the u-th UE selects the optimal drone. Here, $\hat{\mathcal{U}}_u \subseteq \mathcal{U}_u$ denotes the set of drones that can make the associated u-th UE's received channel gain higher than that of any Eve. The optimal association is contained in $\hat{\mathcal{U}}_u$, since associating with drones in $\hat{\mathcal{U}}_u$ typically satisfies the security constraint.

In conclusion, the probability of finding the optimal solution using channel-gain grouping is higher than that using random grouping, i.e., $P_{\text{CG}} \geq P_{\text{RG}}$. For UE-channel-gain grouping, the probability of finding the optimal solution is $P_{\text{UG}} \approx 0$ in most cases.

Compared to these three benchmark algorithms, by properly setting the parameters in TS-GA, the probability of finding the optimal solution is nearly 1, i.e., $P_{\text{TS}}(T) \approx 1$.

3.4. A Q-Learning-Based Position Decision Algorithm (Q-PDA)

This algorithm determines the optimal drone position by identifying the state that occurs most frequently among the end states reached by paths originating from multiple

states. In a complex operational environment with numerous possible states for the drone, the most frequently occurring end state among those reached from multiple starting states is likely to represent a stable and efficient position. This approach allows Q-learning to quickly converge to a near-optimal position, especially in scenarios without specifying the target state in advance. The procedure of the proposed algorithm is described in Algorithm 2. The variables for the subsequent operations in the proposed algorithm are described as follows:

A. State: In this algorithm, we consider the candidate position of a drone as state, $s_t = (x, y) \in \mathcal{S} = \mathcal{X} \times \mathcal{Y}$, and the total number of states is $S = |\mathcal{X}| \times |\mathcal{Y}|$, where $\mathcal{X} = \{v, 2v, 3v, ..., X_{border}\}$, $\mathcal{Y} = \{v, 2v, 3v, ..., Y_{border}\}$, where X_{border} and Y_{border} are the edges of the map, respectively. A minimum separation distance v is maintained to prevent potential collisions among drones. We use v as the degree of discretizations, which determines the size of the divided grid in the predefined service area.

B. Action: We define the movement action space of the drone as

$$\mathcal{A} : \{\text{stay}, \text{north}, \text{south}, \text{west}, \text{east}\}. \tag{26}$$

When the drone is in state s_t, it takes an action a_t, where $a_t \in \mathcal{A}$.

C. Q-table: Similarly as traditional Q-learning, we use Q values to represent the expected cumulative rewards obtained after taking a certain action in a certain state and storing the values of each state–action pair in a matrix.

(1) Initialize Q-table using $\mathbf{r}_i$ vector: $\mathbf{r}_i$ vector provides initial values to $\mathbf{Q}$ matrix. The worst-case secrecy rate is calculated for each candidate position and recorded in the corresponding $\mathbf{r}_i \in \mathcal{R}^{S \times 1}$ vector. By sequentially assuming a drone is positioned at each candidate location, we systematically obtain the worst-case secrecy rate varying with location. when calculating it, the proposed TS-GA (or other grouping algorithms) is needed to associate drones at position s_t with nearby UEs. Furthermore, if $2 \leq i \leq D$, the previously trained drones participate in the grouping process, and the worst-case secrecy rates of the previously deployed positions and their adjacent positions are set to 0 in the $\mathbf{r}_i$ matrix to prevent collisions.

Before training the i-th drone, i-th Q matrix $\mathbf{Q}_i \in \mathbf{R}^{\mathcal{S} \times |\mathcal{A}|}$ is initialized using the $\mathbf{r}_i$ at the beginning, that is

$$\mathbf{Q}_i(s_t, a = \text{stay}) = \mathbf{r}_i(s_t), \tag{27}$$
$$\mathbf{Q}_i(s_t, a = \text{north}) = \mathbf{r}_i(s_t - X_{border}/v), \tag{28}$$
$$\mathbf{Q}_i(s_t, a = \text{south}) = \mathbf{r}_i(s_t + X_{border}/v), \tag{29}$$
$$\mathbf{Q}_i(s_t, a = \text{west}) = \mathbf{r}_i(s_t - 1), \tag{30}$$
$$\mathbf{Q}_i(s_t, a = \text{east}) = \mathbf{r}_i(s_t + 1). \tag{31}$$

(2) Update Q-table: Drones approach the optimal position by continuously choosing the actions and updating the Q values. Through continuous Q-value updates, the optimal strategy is learned, with the aim of finding the most secure positions. At time step t, after performing the action a_t in state s_t, we update the Q-value table $\mathbf{Q}(s_t, a_t)$ following the Bellman optimality equation [17]:

$$\mathbf{Q}(s_t, a_t) \leftarrow \mathbf{Q}(s_t, a_t) + \alpha \left[r_t + \gamma_q \max_{a'} \mathbf{Q}(s_{t+1}, a') - \mathbf{Q}(s_t, a_t) \right], \tag{32}$$

where $\mathbf{Q}(s_t, a_t)$ is the current Q-value when taking action a_t in state s_t, and $\max_{a'} Q(s_{t+1}, a')$ is the maximum Q-value when choosing the best action a' in the next state s_{t+1}. α is the learning rate $0 \leqslant \alpha \leqslant 1$, which controls the influence of new experience on the Q-value

update, r_t is the immediate reward obtained after performing action a_t at time step t, and γ_q is the discount factor $0 \leqslant \gamma_q \leqslant 1$, which controls the influence of future rewards on the current decision.

D. Suboptimal position decision: The suboptimal position of a drone is determined by the end state toward which most starting points eventually converge. In each episode, the i-th drone is placed at a random initial position; after training its Q-table ($\mathbf{Q}_i$), the drone executes actions yielding the highest reward in each step until it reaches a terminal state s_{end}. We then record the frequency of each terminal state across all episodes, with the most frequently encountered terminal state selected as the suboptimal position. This approach leverages the inherent stability of the Q-learning policy, where high-frequency terminal states emerge as "attractors" that balance exploration and exploitation—reflecting positions robust to variations in initial conditions and thus suitable for practical deployment.

Algorithm 2: Q-PDA.

Input: Input state set $\mathcal{S}$, action set $\mathcal{A}$

Training process:
Initialize $\mathbf{r}_i$ (Initialize $\mathbf{r}_i$ as defined in Section 3.2), $\forall s_t = (x_t, y_t) \in \mathcal{S}$, $\forall a_t \in \mathcal{A}$, Flag$= 1$
for Each drone $i \in [1, D]$ **do**
  Generate $\mathbf{Q}_i$, Initialize $\mathbf{s}_{end}(s_t)$
  for Each episode $t \in [1, 2000]$ **do**
    if Flag$= 1$ **then**
      random initial state $s_0 \in \mathcal{S}$, select a random action a_0 from this state s_0, Flag$= 0$.
    end if
    Generate random number $\beta \in [1, 100]$
    if $\beta < 80 - t/25$ **then**
      Random choose an action a_t = random$(\mathcal{A})$
    else
      Choose the action with best reward a_t = argmax$_a \mathbf{Q}_i^t(s_t, a)$
    end if
    Calculate the reward value by using Equation (32) and update $\mathbf{Q}_i^t$.
    if $s_t = s_{t+1}$ or $s_t = 0$ **then**
      $\mathbf{s}_{end}(s_t) = \mathbf{s}_{end}(s_t) + 1$; Flag = 1;
      continue;
    else
      $s_t \leftarrow s_{t+1}$;
      Place i-th drone in s_t;
    end if
    if count$(||\mathbf{Q}_i^t - \mathbf{Q}_i^{t-1}|| \leq 0.0001)$=1000 **then**
      Break;
    end if
  end for
  Determine i-th drone's location by using the most frequently occurred end state, that is $q_i = \text{argmax}_{s_t}(\mathbf{s}_{end}(s_t))$, if the drone's position has already been selected, choose the suboptimal solution.
  return q_i
end for

3.5. A Simplified Q-Learning-Based Position Decision Algorithm (Q-PDA-Lite)

Building upon the previous algorithm, we introduce a simplified version, Q-PDA-Lite, that relies on the $\mathbf{r}_1$ to streamline the training process without grouping UEs with the previously trained drones. By eliminating the need to compute additional $\mathbf{r}_i$ matrices,

this approach significantly reduces computational overhead. During the training for i-th drone, when initializing the i-th Q matrix, the $\mathbf{r}_1$ is preprocessed by setting the values corresponding to previously deployed drones' positions and their adjacent postions to zero.

3.6. Theoretical Analysis

To quantify the probability of Q-PDA finding the optimal solution for the drone position adjustment subproblem, we first introduce two realistic assumptions: Under the ϵ-greedy policy, every state–action pair is visited with positive probability. When Q-values converge to stable values, the greedy policy selects actions corresponding to the optimal policy. Q-PDA searches for the optimal solution proceed in two sequential phases:

(1) Exploration phase: Suppose that starting from any non-optimal state, the probability of jumping directly to the optimal state within one step is

$$p = \frac{4}{(S^2-1)} \frac{\epsilon}{|\mathcal{A}|}, \tag{33}$$

where $\mathcal{A}$ is the action space. The probability of visiting the optimal solution within T_1 iterations is

$$P_3(T_1) = 1 - (1-p)^{T_1}. \tag{34}$$

(2) Convergence and Lock-in phase:

After Q-values converge to a stable value, Q-PDA will stably take optimal actions, thereby stably outputting the optimal solution. By setting learning rate and ergodicity conditions, Q-PDA converges with probability 1, that is

$$\lim_{T \to \infty} P_4(T_2) = 1. \tag{35}$$

In conclusion, combining the properties of the two phases, the probability that Q-PDA finds the optimal solution for the drone position adjustment subproblem within $T = T_2 + T_3$ steps can be expressed as

$$P_{\mathrm{Q-PDA}}(T) \geq \left[1 - (1-p)^{T_1}\right] \cdot P_4(T_2). \tag{36}$$

When T is sufficiently large, $P_{\mathrm{Q-PDA}}(T) \approx 1$.

3.7. Convergence and Complexity Analysis

As shown in Figure 2, the secrecy performance of Q-PDA and Q-PDA-Lite tends to stabilize in the episode of 2000. Therefore, the number of learning episodes is set to 2000 in the following simulations.

The time complexity of TS-GA depends on three steps: candidate solution generation, candidate solution analysis, and tabu-list management. The time cost of the above three steps are $O(U^2)$, $O(U)$, and $O(U^2)$, respectively. The total worst-complexity of TS-GA is $O(T * (2U^2 + U))$, where T represents the maximum number of iterations.

The time complexity of Q-learning needs to be analyzed from two aspects: single update and convergence iteration. The time complexity of a single update is mainly due to the traversal during action selection, that is $O(|\mathcal{A}|)$.The time complexity of convergence iteration is $O(Avgstep * T_Q)$, where T_Q represents the maximum number of episode, and Avgstep is the average number of steps in each episode. In conclusion, the time complexity of Q-PDA and Q-PDA-Lite are $O(D * |\mathcal{A}| * Avgstep * T_Q + D * T * (2U^2 + U) * |\mathcal{A}|)$ and $O(D * |\mathcal{A}| * Avgstep * T_Q + T * (2U^2 + U) * |\mathcal{A}|)$, respectively.

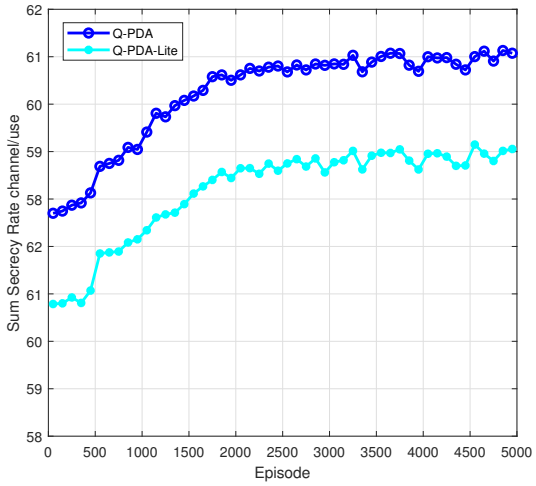


Figure 2. Convergence performance of the proposed algorithm.

3.8. Benchmark Algorithms

The benchmark algorithms for evaluating the performance of Q-PDA and Q-PDA-Lite are defined as follows.

Random-PDA: Within the deployment area, a random two-dimensional coordinate is generated for each drone to be deployed. The drones are directed to these positions, which form the deployment solution of the Random-PDA algorithm. The probability of finding the optimal solution for the drone position adjustment subproblem is

$$P_{\mathrm{RP}} = \frac{1}{S} * \frac{1}{S-1} * \ldots * \frac{1}{S-D+1}, \tag{37}$$

where S is the total number of states.

K-means-PDA: Assuming the system has knowledge of all user locations within the area, the standard K-means algorithm is used to partition the users into a number of clusters equal to the number of drones. The drones are then guided to the centroids of these clusters, which constitute the deployment solution of the K-means-PDA algorithm. This method aims to maximize the channel quality of UEs. The probability of finding the optimal solution for the drone position adjustment subproblem is

$$P_{\mathrm{KM}} \approx 0. \tag{38}$$

In practice, the optimal solution rarely requires drones to select positions based on UE density. Although being surrounded by a large number of UEs (and associating with them) improves UE channel quality, the highest UE channel quality does not equate to the highest system security. For instance, if a drone is deployed near an Eve in a UE-dense cluster, Eve can easily intercept the strong signals intended for UEs—leading to a significant drop in overall security performance.

The probability of Q-PDA finding the optimal solution is approximately 1. For the benchmark algorithms, their probabilities of finding the optimal solution are usually $P_{\mathrm{RP}} \ll 1$ or even $P_{\mathrm{KM}} \approx 0$. This is because the benchmark algorithms only rely on a single selection, making it difficult to hit the optimal solution.

4. Simulation Results

In Section 4.1, we examine the hybrid algorithm TS-GA+Q-PDA by comparing it with other hybrid algorithms. Then, an ablation study is conducted to investigate the impact

of each modular algorithm on the hybrid algorithm TS-GA+Q-PDA. In Section 4.2, we examine the Q-PDA module within the hybrid algorithm TS-GA+Q-PDA. In Section 4.3, we investigate the TS-GA module from the hybrid algorithm TS-GA+Q-PDA. In Section 4.4, the security performance of the hybrid algorithms TS-GA+Q-PDA and TS-GA+Q-PDA-Lite is compared with that of exhaustive search.

All simulations were implemented using MATLAB (version 2016). software. In the simulations, we configured 20 UEs to be distributed within a specified spatial area, all at a uniform height of 0.8 m. To mimic real-world application scenarios more accurately, the UEs' positions were randomly assigned within this area. In the same model, we assumed the presence of four drones, also distributed within the area but at a unified height of 10 m. Since all drones maintain a fixed altitude, our analysis focuses on the two-dimensional plane at that altitude. The parameters used in the following simulations are shown in Table 1.

Table 1. Simulation parameters.

Parameter	Value
Area of outdoor space ($W \times L \times H$)	$(80 \times 80 \times 10)$ m^3
Average electrical ambient noise (ξ)	-98 dBm
Lambertian emission order (β)	1
Half-intensity radiation angle ($\theta_{1/2}$)	60°
PD surface area (A_R)	1 cm^2
Optical filter gain (g_{of})	1
Maximum power of a LED	10–100 W
PD FoV (Θ)	70°
Refractive index (q)	1.5
Learning rate(α)	0.99
Discount factor(γ_q)	0.8

4.1. Comparison of TS-GA+Q-PDA Versus Other 15 Hybrid Algorithms

Based on Sections 3.3 and 3.8, there are four position decision algorithms (including the proposed Q-PDA and Q-PDA-Lite) and four grouping algorithms (including the proposed TS-GA). This yields a total of 16 possible hybrid combinations of position decision and grouping algorithms. In this subsection, we compare the proposed TS-GA+Q-PDA with the other 15 hybrid algorithms.

In Figure 3, the security performance of TS-GA+Q-PDA is presented and compared with other hybrid algorithms at the drone's transmit power of 50 W. Among all the hybrid algorithms, TS-GA+Q-PDA achieves the highest worst-case sum secrecy rate, reaching approximately 81 channel/use.

Additionally, among TS-GA and the benchmark grouping algorithms, all the position decision algorithms have higher performance with the combination of TS-GA. This is because TS-GA focuses its search directly on finding solutions with the best security performance, rather than (like the benchmark algorithms) only identifying solutions that exhibit the characteristics of the optimal solution.

Among Q-PDA and the benchmark position decision algorithms, all the grouping algorithms have higher performance with the combination of Q-PDA. This is because Q-PDA arranges drones in areas with dense UEs while keeping them far from Eves. This deployment enhances the received SINR of UEs and reduces the probability of signal interception by Eves, thereby improving the overall system security performance.

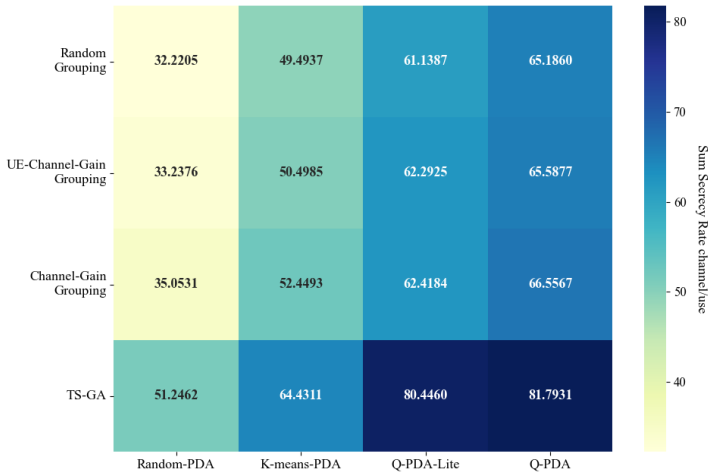
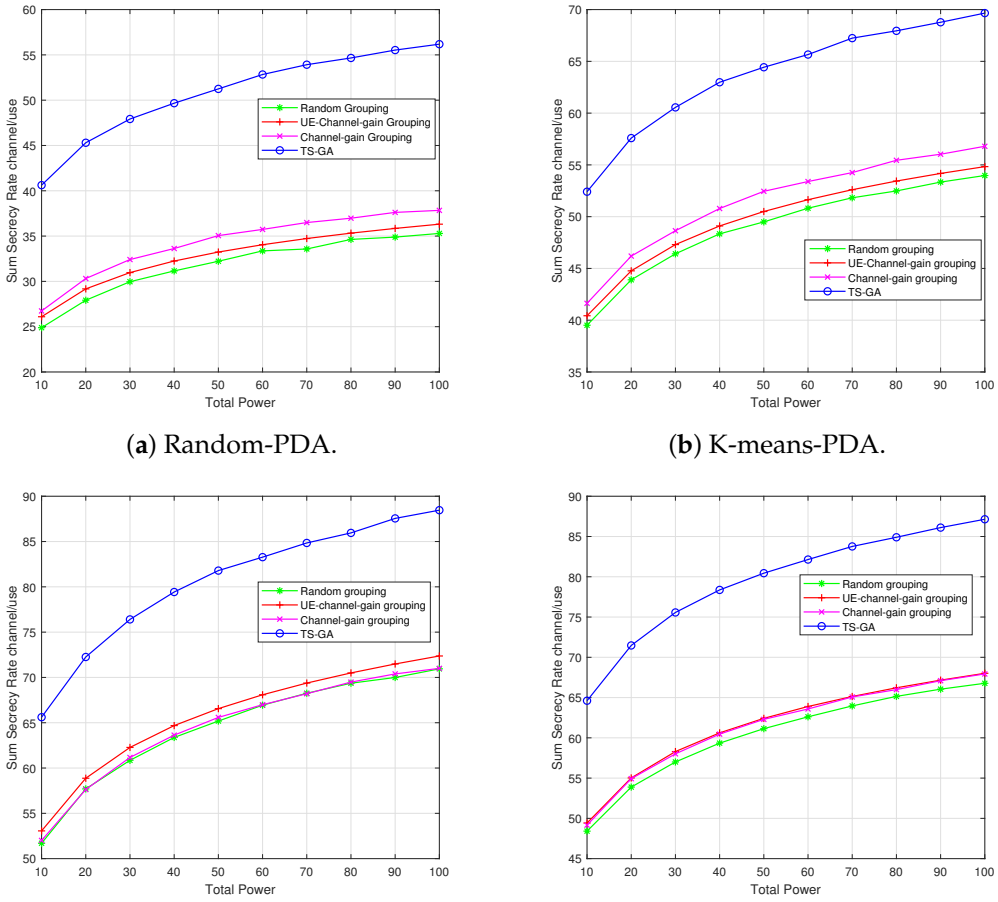


Figure 3. Comparison of secrecy performance of TS-GA+Q-PDA with other hybrid algorithms.

4.2. Comparison of Q-PDA and Q-PDA-Lite Versus Other Benchmark Algorithms

In Figure 4, the security performance of Q-PDA and Q-PDA-Lite variation with the transmit power is presented and compared with random-PDA and K-means-PDA. With an increase in the drones' transmit power, the received SINR of the associated UEs rises, thus improving the system's secrecy performance of Q-PDA, Q-PDA-Lite, and benchmark algorithms.



(**a**) Random-PDA. (**b**) K-means-PDA.

(**c**) Q-PDA. (**d**) Q-PDA-Lite.

Figure 4. Comparison of different PDAs combined with four grouping algorithms.

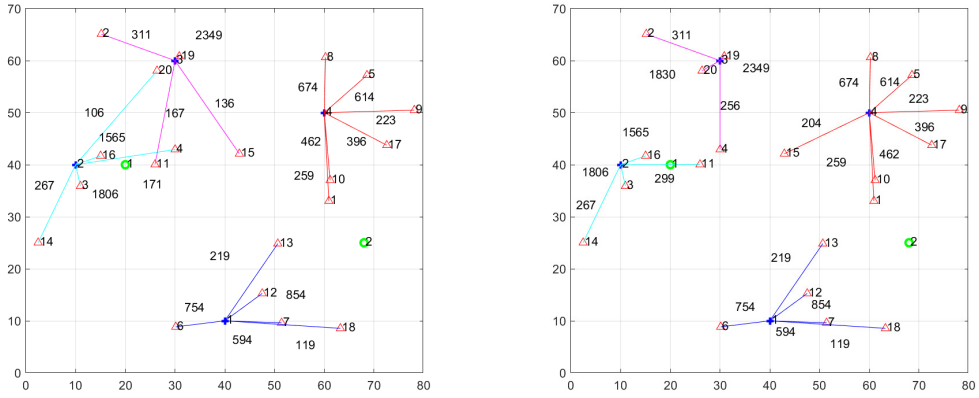
By conducting a horizontal comparison of different subfigures, both Q-PDA and Q-PDA-Lite outperform Random-PDA and K-means-PDA. This is because Q-PDA and Q-PDA-Lite consider both the positions of UEs and Eves. By using Q-PDA and Q-PDA-Lite, drones are arranged in a dense UE area and are far from Eves. In Q-PDA-Lite, we do not implement the grouping of UEs with pre-trained drones in the training, which leads to relatively lower performance compared to Q-PDA. In Figure 4a, when the Random-PDA is employed instead of Q-PDA, drones may end up in sparse regions with a small number of UEs. Moreover, since Eve's presence is disregarded, drones can get close to Eve. This proximity increases the risk of signal interception and decreases the secrecy rate. In Figure 4b, when the K-means-PDA is employed instead of Q-PDA, the K-means-PDA selects centroid positions to maximize user coverage by taking into account the UE distribution. Nevertheless, it still overlooks Eve's positions, and drones may occasionally be deployed in closer proximity to eavesdroppers, which results in degradation of the secrecy rate.

4.3. Comparison of TS-GA Versus Other Benchmark Algorithms

In this subsection, we first compare the performance of TS-GA and its benchmark grouping algorithms, while using Q-PDA as the position decision algorithm. Compared with the benchmark grouping algorithms, TS-GA has fewer associations, yet each association is effective. In the time-sharing system, the reduced number of associations allows UEs with better positions (which are close to drone and far away from Eve) to receive signals for a longer duration, consequently increasing the overall system secrecy rate.

In Figure 5a, the channel gain of most UEs associated to drones is 0, which leads to a zero secrecy rate. This is because random-grouping does not consider the signal strength of the UE or the Eve and randomly selects a drone from those within the UE's signal reception range. Unlike the random-grouping, as shown in Figure 5b, the UE-channel-gain-grouping has each UE selects the drones with the highest channel gain as the association target. However, the secrecy performance of the system may degrade to zero if Eves have better channel gains from the drone than the associated UEs.

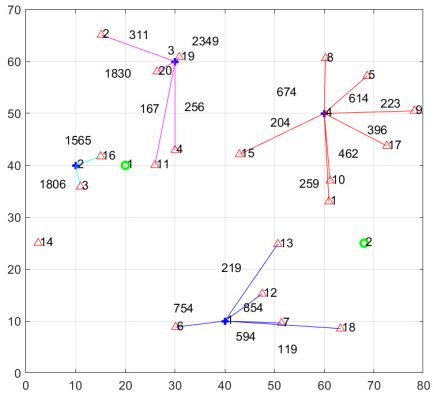
The channel-gain-grouping differs from the above two methods by taking the positions of the Eves into account. As shown in Figure 5c, drones can only associate to UEs that are closer to it than Eve. By selecting a drone with a channel gain greater than the intercepted channel gain for each UE, the number of effective associations increases, thereby better than random-grouping and UE-channel-gain-grouping.



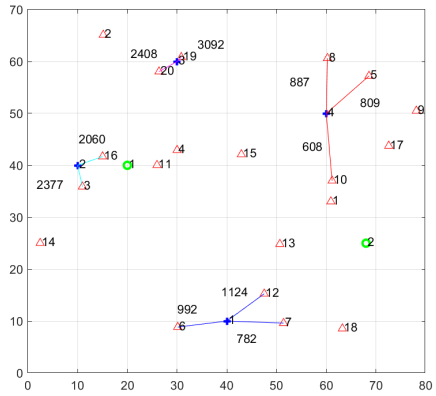
(**a**) Random grouping.

(**b**) UE-channel-gain grouping.

Figure 5. *Cont.*



(**c**) Channel-gain grouping.



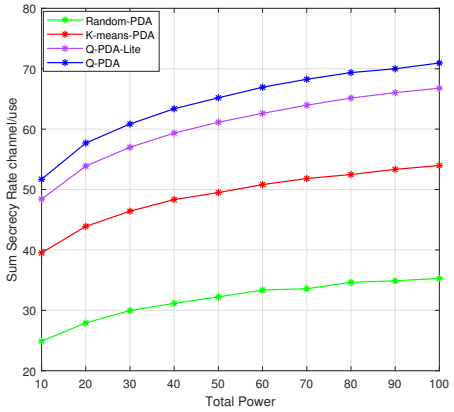
(**d**) TS-based Grouping Algorithm.

Figure 5. UE–drone connections under different grouping methods.

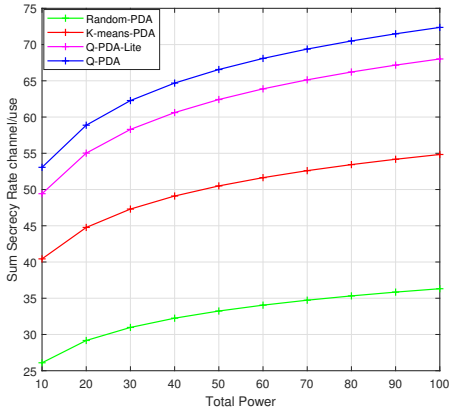
Then, in Figure 6, the security performance of TS-GA varying with the transmit power is presented and compared with random-grouping, UE-channel-gain-grouping, and channel-gain-grouping. With an increase in the drones' transmit power, the received SINR of the associated UEs rises, thus improving the system's secrecy performance of TS-GA and benchmark algorithms.

By conducting a horizontal comparison of different subfigures, TS-GA outperforms random-grouping, UE-channel-gain-grouping, and channel-gain-grouping, reaching up to approximately 88 channel/use. This is because TS-GA focuses its search within the neighborhood of good-quality solutions to explore better candidates, thereby increasing the probability of finding the optimal solution. In Figure 6c, when the channel-gain-grouping is employed instead of TS-GA, the performance improves slightly relative to random-grouping and UE-channel-gain-grouping but still cannot match TS-GA. This is because the channel-gain-grouping only prioritizes the "UE–drone–Eve distance" metric, while TS-GA explores multiple dimensions to generate more optimal groupings.

In Figure 6a, when the random-grouping is employed instead of TS-GA, it merely selects a drone randomly from those within the UE's signal reception range, leading to numerous invalid associations that fail to contribute to the security of the system. In Figure 6b, when the UE-channel-gain-grouping is employed instead of TS-GA, it overlooks the threat of Eves. If an Eve has a better channel gain from the associated drone than the UE itself, the system's security is severely compromised.

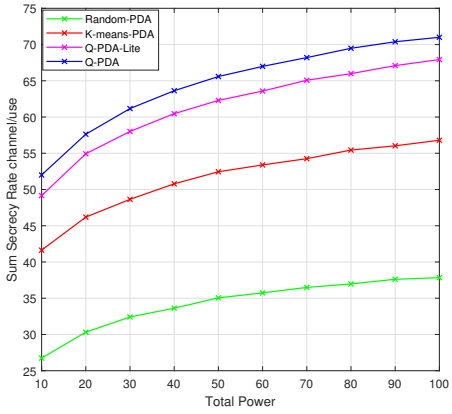


(**a**) Random-Grouping.

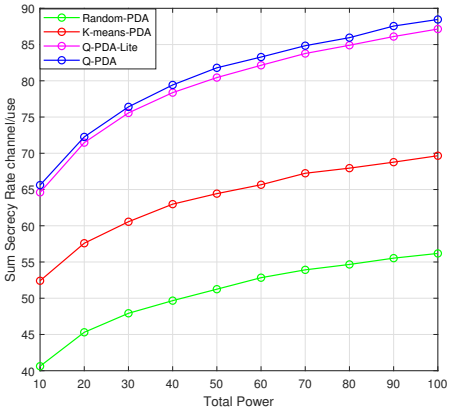


(**b**) UE-Channel-Gain-Grouping.

Figure 6. *Cont.*



(**c**) Channel-Gain-Grouping.



(**d**) TS-GA.

Figure 6. Comparison of different grouping algorithms integrated with four PDAs.

4.4. Compared with Exhaustive Search

As shown in Figure 7, our simulation results show that the performance gap between our proposed algorithms and the global optimum is about 10%. In practical applications, this minor performance gap is often acceptable to gain orders of magnitude in speed and reduce energy consumption, which is critical for extending drone flight times. Q-PDA-Lite combined with TS-GA exhibits performance comparable to Q-PDA paired with TS-GA, while boasting lower complexity—making it well-suited for drone deployment. In this simulation, we employed 2 drones to serve 20 UEs, aiming to reduce the time required for the exhaustive search.

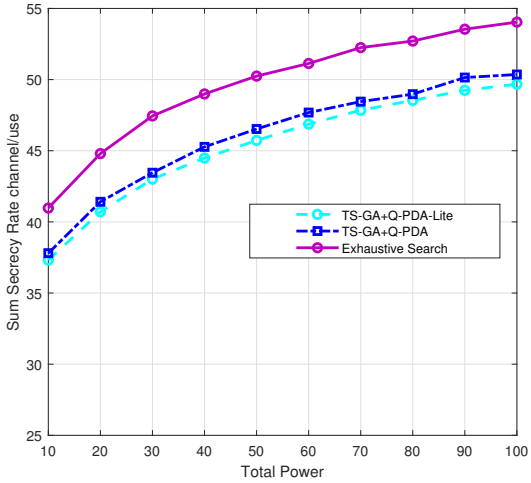


Figure 7. Comparison of secrecy performance with exhaustive search.

5. Conclusions

This paper investigates the security of a multi-drone, multi-UE VLC-enabled system in the presence of multiple Eves. We propose two algorithms: Q-PDA for optimizing drone positions and TS-GA for solving UE–drone association, both aiming to maximize the sum secrecy rate of UEs. Q-PDA-Lite is a low-complexity variant of Q-PDA. Simulation results demonstrate that our proposed algorithm combination TS-GA+Q-PDA outperforms other pairings of position decision and grouping algorithms, with a performance gap of only 10% compared to the global optimal result. In future work, we plan to extend this system to dynamic scenarios.

Author Contributions: Conceptualization, G.S.; methodology, G.S. and H.Z.; software, H.Z.; validation, G.S.; formal analysis, H.Z.; investigation, F.W.; resources, W.C.; data curation, F.W.; writing—original draft preparation, W.C.; writing—review and editing, H.W.; visualization, H.W.; supervision, H.W.; project administration, H.W.; funding acquisition, G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by International Science and Technology Cooperation Project of Henan Province under Grant 242102520040 and the Key Scientific Research Projects of Colleges and Universities in Henan Province (No.23A520031).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict 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. Wang, Y.; Chen, M.; Yang, Z.; Luo, T.; Saad, W. Deep Learning for Optimal Deployment of UAVs with Visible Light Communications. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 7049–7063. [CrossRef]
2. Saad, W.; Bennis, M.; Chen, M. A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. *IEEE Netw.* **2020**, *34*, 134–142. [CrossRef]
3. Aboagye, S.; Ndjiongue, A.R.; Ngatched, T.M.N.; Dobre, O.A.; Poor, H.V. RIS-Assisted Visible Light Communication Systems: A Tutorial. *IEEE Commun. Surv. Tuts.* **2023**, *25*, 251–288. [CrossRef]
4. Shi, G.; Li, Y.; Cheng, W.; Gao, X.; Zhang, W. An Artificial-noise Based Approach for the Secrecy Rate Maximization of MISO VLC Wiretap Channel with Multi-Eves. *IEEE Access* **2020**, *9*, 651–659. [CrossRef]
5. Shi, G.; Aboagye, S.; Ngatched, T.M.N.; Dobre, O.A.; Li, Y.; Cheng, W. Secure Transmission in NOMA-Aided Multiuser Visible Light Communication Broadcasting Network With Cooperative Precoding Design. *IEEE Trans. Inform. Foren. Sec.* **2022**, *17*, 3123–3138. [CrossRef]
6. Zeng, M.; Nguyen, N.; Dobre, O.A.; Poor, H.V. Securing Downlink Massive MIMO-NOMA Networks With Artificial Noise. *IEEE J. Sel. Top. Signal Process.* **2019**, *13*, 685–699. [CrossRef]
7. Dong, L.; Wang, H.M.; Bai, J. Active Reconfigurable Intelligent Surface Aided Secure Transmission. *IEEE Trans. Veh. Technol.* **2022**, *71*, 2181–2186. [CrossRef]
8. Liu, S.; Liu, X.; Du, X.; Guizani, M. Smart Jamming for Secrecy: Deep Reinforcement Learning Enabled Secure Visible Light Communication. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 17915–17928. [CrossRef]
9. Xue, Z.; Wang, J.; Ding, G.; Zhou, H.; Wu, Q. Cooperative Data Dissemination in Air-Ground Integrated Networks. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 209–212. [CrossRef]
10. Zhang, Y.; Mou, Z.; Gao, F.; Jiang, J.; Ding, R.; Han, Z. UAV-Enabled Secure Communications by Multi-Agent Deep Reinforcement Learning. *IEEE Trans. Veh. Technol.* **2020**, *69*, 11599–11611. [CrossRef]
11. Zhou, Y.; Yeoh, P.L.; Pan, C.; Wang, K.; Ma, Z.; Vucetic, B.; Li, Y. Caching and UAV Friendly Jamming for Secure Communications with Active Eavesdropping Attacks. *IEEE Trans. Veh. Technol.* **2022**, *71*, 11251–11256. [CrossRef]
12. Ibraiwish, H.; Eltokhey, M.W.; Alouini, M.S. Energy Efficient Deployment of VLC-Enabled UAV Using Particle Swarm Optimization. *IEEE Open J. Commun. Soc.* **2024**, *5*, 553–565. [CrossRef]
13. Pham, Q.V.; Huynh-The, T.; Alazab, M.; Zhao, J.; Hwang, W.J. Sum-Rate Maximization for UAV-Assisted Visible Light Communications Using NOMA: Swarm Intelligence Meets Machine Learning. *IEEE Internet Things J.* **2020**, *7*, 10375–10387. [CrossRef]
14. Su, D.P.; Lu, D.S.; Yu, Y.C.; Wang, J.Y. Joint Optimization of UAV's Height and Peak Optical Intensity in Weather-Dependent Covert VLC. *IEEE Signal Process. Lett.* **2024**, *31*, 2315–2319. [CrossRef]
15. Cho, S.; Chen, G.; Coon, J.P. Cooperative Beamforming and Jamming for Secure VLC System in the Presence of Active and Passive Eavesdroppers. *IEEE Trans. Green Commun. Netw.* **2021**, *5*, 1988–1998. [CrossRef]

16. Obeed, M.; Salhab, A.M.; Alouini, M.S.; Zummo, S.A. On Optimizing VLC Networks for Downlink Multi-User Transmission: A Survey. *IEEE Commun. Surv. Tutorials* **2019**, *21*, 2947–2976. [CrossRef]
17. Liu, C.; Zhong, Y.; Wu, R.; Ren, S.; Du, S.; Guo, B. Deep Reinforcement Learning Based 3D-Trajectory Design and Task Offloading in UAV-Enabled MEC System. *IEEE Trans. Veh. Technol.* **2025**, *74*, 3185–3195. [CrossRef]

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.