

Article

A Secure Data Collection Method Based on Deep Reinforcement Learning and Lightweight Authentication

Yunlong Wang¹, Jie Zhang^{1,2,*}, Guangjie Han¹  and Dugui Chen³

¹ College of Information Science and Engineering, Hohai University, Changzhou 213200, China; 221320030003@hhu.edu.cn (Y.W.); hanguangjie@hhu.edu.cn (G.H.)

² The Key Laboratory of Ocean Observation Technology, Tianjin 300112, China

³ School of Cultural Tourism, Wenshan Vocational and Technical College, Wenshan 663099, China; cdg102611@gmail.com

* Correspondence: zhangjie9@hhu.edu.cn

Abstract: Cooperative Unmanned Aerial Vehicle (UAV) technology can significantly improve data acquisition in Internet of Things (IoT) environments, which are characterized by wide distribution and limited capacity of ground-based devices. However, due to the open nature of wireless communications, such applications face security threats posed by UAV authentication, especially in scalable IoT environments. To address such challenges, we propose a lightweight chain authentication protocol for scalable IoT environments (LCAP-SIoT), which uses Physical Unclonable Functions (PUFs) and distributed authentication to secure communications, and a secure data collection algorithm, named LS-QMIX, which fuses the LCAP-SIoT and Q-learning Mixer (QMIX) algorithm to optimize the path planning and cooperation efficiency of the multi-UAV system. According to simulation analysis, LCAP-SIoT outperforms existing solutions in terms of computing and communication costs, and LS-QMIX results in superior performance in terms of data collection rate, task completion time, and the success rate of authentication for newly joined UAVs, indicating the feasibility of LS-QMIX in dynamic expansion scenarios.



Academic Editor: Jiangtao Li

Received: 14 April 2025

Revised: 13 May 2025

Accepted: 15 May 2025

Published: 19 May 2025

Citation: Wang, Y.; Zhang, J.; Han, G.; Chen, D. A Secure Data Collection Method Based on Deep Reinforcement Learning and Lightweight Authentication. *World Electr. Veh. J.* **2025**, *16*, 281. <https://doi.org/10.3390/wevj16050281>

Copyright: © 2025 by the authors. Published by MDPI on behalf of the World Electric Vehicle Association. 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/>).

Keywords: data collection; UAV; multi-agent reinforcement learning; lightweight authentication; scalable IoT

1. Introduction

With the rapid development of information technology, Internet of Things technology has been widely used in intelligent transportation, disaster warning, and other fields [1–5]. By perceiving the physical world in real time through IoT devices, IoT provides a basis for data-driven decision-making. However, IoT devices are usually distributed discretely and have limited computing power, which limits the efficiency of data collection [6]. To address this problem, multi-UAV cooperative path planning has become an effective solution to improve data collection efficiency due to its advantages of high mobility, wide area coverage, and flexible deployment [7–13]. However, open wireless communications between UAVs and IoT devices are susceptible to attacks such as eavesdropping, tampering, or forgery, which threaten the integrity of the data and the reliability of decision-making [14], and data transmission security is an important challenge in multi-UAV systems. In addition, complex dynamic environments and the need for collaboration among multiple UAVs significantly increase the difficulty of path planning. Therefore, designing efficient authentication protocols and path planning methods that take into account the security of data collection and UAV path optimization is an important issue in current UAV data collection research.

In the traditional internet environment, commonly used cryptographic protocols such as TLS/DTLS can provide perfect authentication and communication confidentiality [15,16]. However, such common protocols generally suffer from excessive communication messages, computation, and storage overheads, making it more difficult to meet UAV requirements for lightweight and efficient communication. Designing lightweight authentication protocols suitable for UAV environments can meet the need to reduce resource consumption while ensuring security. To fully utilize the potential of multi-UAV systems in collaborative data collection, we also face the challenge of multi-UAV path planning. Traditional UAV path planning, such as heuristics, is difficult to meet real-time and efficiency requirements. Reinforcement learning, which has received widespread attention for its potential in dynamic decision-making and adaptive optimization [17–19], can provide intelligent path planning solutions for multi-UAV collaboration by learning from environmental interactions. System scalability becomes a new yardstick to test the effectiveness of the above path planning and security protocols as monitoring scenarios become more complex and mission scales expand. In path planning, only the identity authentication protocol is used for the authentication of UAVs and IoT devices before data transmission, which cannot meet the requirements in extended scenarios. Therefore, we need to combine simple authentication methods with path planning that uses reinforcement learning, where the authentication methods ensure all devices are verified, and the path planning helps optimize the routes for UAVs while also guiding new UAVs during the authentication process, aiming for a quick, efficient, and secure connection for the new UAVs. It is worth noting that similar issues of collaborative control and secure communication also appear in other multi-agent systems. For example, in a scenario involving connected and autonomous vehicles, vehicles can share state information in real-time via vehicle-to-vehicle and vehicle-to-infrastructure communication to form a multi-agent body system and make collaborative decisions to enhance transportation efficiency and safety [20,21]. The evidence further indicates that the study of multi-agent body collaboration and security protocols has a universal value across scenarios.

This paper proposes a lightweight identity authentication protocol that uses the physical characteristics of the device as an identity identifier. Combining the design concepts of the blockchain, storing tokens using a Merkle tree, and adopting a chain authentication mechanism, distributed secure identity authentication is achieved. At the same time, to optimize trajectory planning in data collection and reduce costs, this paper proposes a multi-UAV data collection method based on Q-learning Mixer (QMIX). This method comprehensively considers obstacle avoidance, flight energy consumption, and the scalability of the number of UAVs. It guides UAVs to collaborate in data collection for IoT devices through reinforcement learning strategies. The main contributions of this paper include the following.

- (1) A lightweight authentication protocol for the multi-UAV system: We propose a lightweight authentication protocol based on chained PUF that aims to reduce the communication cost while supporting the variation in the number of UAVs. This protocol addresses identity security for mobile devices in scalable IoT environments by employing a chained PUF structure for rapid authentication, thereby enabling secure data collection through multi-UAV collaboration in dynamic IoT scenarios.
- (2) A data collection method based on multi-agent reinforcement learning: We transformed the multi-UAV path planning problem into a distributed partially observable Markov decision process (DEC-POMDP) model and designed a QMIX-based multi-UAV data collection method that combines the authentication process with path planning to ensure the efficiency and security of the data collection task in IoT; and a

segmented reward function was designed to support online authentication of new UAV entrants.

The remaining sections of the paper follow this structure: Section 2 reviews the path planning methods and authentication techniques for data collection; Section 3 introduces the system model of the paper; Section 4 elaborates the proposed method; Section 5 evaluates the performance of the proposed method through experimental simulations; and Section 6 summarizes the research work of the paper and proposes future research directions.

2. Literature Review

In multi-UAV collaborative data collection scenarios, path planning determines the efficiency of UAVs in accomplishing their tasks in dynamic environments, while authentication protocols safeguard their security in exchanging data with IoT devices in open wireless channels. Therefore, this section first combs through multi-UAV path planning methods based on reinforcement learning, followed by a review of authentication techniques, and finally summarizes their convergence trends.

Recently, reinforcement learning (RL) has provided new ideas for multi-UAV collaborative path optimization in complex environments by virtue of its end-to-end perception-decision capability. The use of algorithms like Q-learning, Deep Q-Network (DQN), and Actor-Critic has grown significantly. For example, Ejaz et al. [18] suggested a teamwork approach to planning paths using deep Q-learning, which takes into account user needs, risk preferences, and energy use within a reinforcement learning setup, and effectively outlines service quality limits using a Sigmoid-type demand function. However, DQN suffers from the problems that the discrete action space cannot satisfy: the fine manipulation demand, the overestimation of the Q value, and the slow convergence and tendency to fall into the local optimum in the high-dimensional state-action space. To improve the model's temporal modeling capability, Liu et al. [9] reduced the problem of overestimating Q-values and improved how data are processed over time by using average Q-value estimation, long- and short-term memory networks (LSTM), a time-focused attention method, and a layered structure. Zhang et al. [11] proposed a real-time path planning algorithm (RPP-LSTM) that fuses LSTM with deep reinforcement learning (DRL), which utilizes the LSTM as the Q-value network of DQN and superimposes temporal memories to improve the environment perception, but it is more sensitive to the length of the temporal window. To break through the limitations imposed by discrete actions, researchers have proposed algorithms based on policy gradients. For example, Wang et al. [22] used a centralized training and distributed execution strategy to optimize multi-UAV trajectories based on the MADDPG framework and combined it with user offloading decisions to improve energy efficiency and fairness. Silvirianti [19] suggested a DRL algorithm called Q-DDPG and its version with memory, Q-RDDPG, which uses quantum computing to make calculations easier with quantum gates computational complexity, but the quantum simulation still stays at the software layer and the hardware landability is insufficient. Wang et al. [8] designed an Actor-Critic-based distributed decision-making model, which combines local observation and global mean field information to improve the scalability and adaptability of multi-UAV collaboration but requires global mean information broadcasting, which leads to communication overhead and real-time problems. Wu et al. [23] computed the mean Q-value through multiple parallel critic networks to alleviate the Q-value overestimation problem of the traditional DDPG algorithm, which improves the stability but significantly increases the computational burden. Lv et al. [12] proposed an information-theory-based exploration algorithm, Entropy Explorer, which solves the sparse reward problem by calcu-

lating intrinsic reward through state entropy and action entropy and deeply integrates it with the TD3 framework.

The path planning algorithm optimizes the collaboration efficiency and lays an environment-aware foundation for subsequent identity security verification. To realize the security of multi-UAV systems, researchers are also actively exploring the authentication and key negotiation mechanisms to cope with the risks of cyberattacks and privacy leakage. Traditional encryption algorithms such as AES, RSA, and ECC are widely used for UAV communication security. For example, Asghar Khan et al. [24] proposed a certificate-based ring-signing cryptography scheme, which effectively improves the security of UAV-assisted edge computing systems. However, the computational and storage overheads of such algorithms limit their utility in resource-constrained devices. To achieve this, Nyangaresi et al. [25] created a method that uses both symmetric key and ECC to lower the amount of computing power and storage needed by generating session keys on the fly and allowing two-way authentication in environments with limited resources. The introduction of blockchain technology provides new ideas for identity authentication, and its decentralized and tamper-proof features enhance authentication security. Dong et al. [26] proposed a framework based on blockchain with a self-consistent identity, which utilizes decentralized identifiers and verifiable credentials to achieve efficient authentication. Li et al. [27] designed an identity authentication mechanism based on the Merkle tree, which combines the disaster semantic blockchain to write the temporal and spatial associated disaster semantic information into the chain to reduce storage and communication overhead. PUF is a function that utilizes hardware uniqueness to generate non-replicable identifiers, and it is suitable for lightweight authentication of resource-constrained devices due to its low computational overhead and high security. For example, Zhang et al. [28] proposed a two-stage authentication and key negotiation protocol based on PUF, which optimizes the computational efficiency. Tian et al. [29] suggested a lightweight authentication and key negotiation protocol that uses UAVs, which provides unique authentication and protection against tampering for IoT devices by using a PUF.

In short, DRL helps with quick and flexible decision-making for planning paths for multiple UAVs, while technologies like PUF and blockchain offer secure ways to verify identities in limited-resource settings. However, the two lines of research are mostly independent of each other: path optimization algorithms usually assume that a secure channel has been established, while identity protocols lack awareness of flight trajectories. By embedding authentication factors in DRL states and rewards, secure and efficient data collection can be achieved, and fast and secure network entry for additional UAVs can be supported.

3. System Model and Problem Formulation

3.1. Mission Model

In this study, we constructed a secure data collection system for UAV-assisted IoT. As shown in Figure 1, the system consists of $N \triangleq \{1, 2, \dots, N\}$ ground-based IoT device nodes (IoTD) and $M \triangleq \{1, 2, \dots, M\}$ UAVs designed to securely and efficiently collect data from device nodes in the region and offload it to the base station (BS). All IoT device nodes are statically deployed within the mission area, while the UAV is responsible for collecting data from these nodes. After completing the data collection task, the UAV will return to the BS and offload the collected data. The 3D position of UAV i at the time step t is denoted as $P_t^i = (x_t^i, y_t^i, z_t^i)$, where x_t^i and y_t^i denote the coordinates of the UAV's position in the horizontal plane and z_t^i denotes the UAV's flight altitude. It is assumed that each UAV flies at a different altitude and the altitude of the UAV remains constant throughout the mission to simplify flight path planning and avoid collision problems caused by altitude changes.

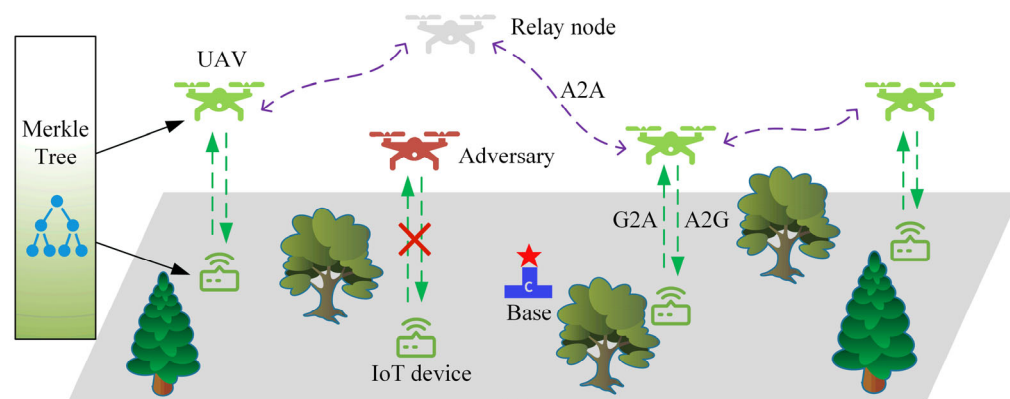


Figure 1. An example of multi-UAV cooperation for secure data collection.

To ensure data security during mission execution, the authentication model considered in this paper consists of legitimate UAVs, ground-based IoT, and a trusted registration authority (RA), as shown in Figure 2. The RA is deployed in the BS and is subject to strict physical and logical security. All entities (including UAVs and IoT) are required to complete registration at the RA before deployment to ensure the legitimacy and trustworthiness of their identities. The objective is to achieve efficient and secure mutual authentication between any two entities (i.e., UAV and UAV or UAV and IoT) within the mission area.

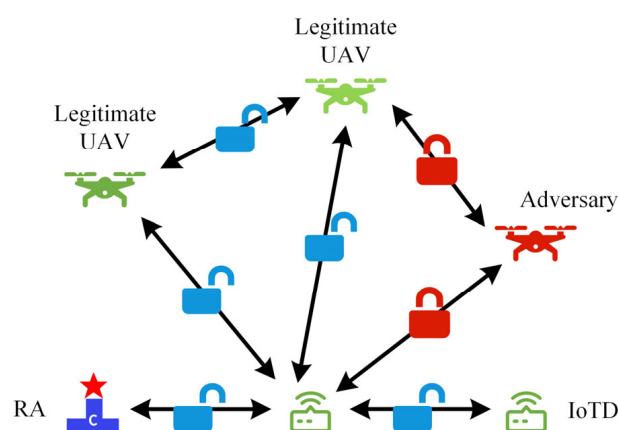


Figure 2. An example of authentication in multi-UAV systems.

3.2. Communications Model

The communication channel model is a key factor for characterizing signal propagation in wireless networks. For the multi-UAV data collection scenario in this paper, to simultaneously satisfy the design goals of real-time path planning, safe and reliable authentication, and UAV number expansion of this system, the channel models are categorized into the following three types: UAV–UAV communication, UAV–IoT communication, and UAV–base communication. The specific models are designed as follows.

- (1) UAV–UAV: The communication between UAVs takes place at high altitude with fewer obstacles, which is suitable for line-of-sight (LOS) links. At this time, the channel propagation characteristics can be modeled using the free-space path loss model [30]:

$$P_r = P_t - 20 \log_{10}(d) - 20 \log_{10}(f) - 20 \log_{10}\left(\frac{4\pi}{c}\right), \quad (1)$$

where P_r is the received power at distance d from the transmitter, P_t is the transmit power, f denotes the communication frequency, and c is the speed of light. The model is formulated

in a concise way for rapid estimation of link power, which satisfies the need for low latency and computational overhead for real-time path planning.

- (2) UAV-IoTD: In the device awareness phase, the UAV has not yet determined the location of the ground sensor, signal propagation may be blocked by obstacles, and the communication link is in non-line-of-sight (NLOS) conditions. The channel model can use a logarithmic distance path loss model [31]:

$$P_r = P_t - L_0 - 10\gamma \log_{10}(d) + X_\sigma, \quad (2)$$

where L_0 denotes the reference path loss, γ is the path loss exponent, X_σ is the random fading factor, and obeys a Gaussian distribution. When the UAV completes device sensing, the UAV approaches the IoTD through path planning, and the communication link turns to LOS conditions. At this time, the free-space path loss model can be used to ensure high reliability and high transmission efficiency of data upload. The two-phase modeling balances the signal uncertainty in the node discovery phase with the high reliability of the session communication to ensure, specifically, better signal quality for secure authentication and data transmission.

- (3) UAV-BS: The communication link between the UAV and the base is mainly a LOS link, and the free-space path loss model is usually applicable. However, when the UAV is far away from the base, the communication quality may degrade. In this case, multi-hop transmissions can be performed by introducing relay UAVs [30], thus extending the communication coverage and reducing the path loss. The total path loss can be expressed as follows:

$$L_{total} = \sum_{i=1}^n L_f(d_i), \quad (3)$$

where $L_f(d_i)$ is the path loss of the i th hop, d_i is the communication distance of the i th hop, and n is the number of relay UAVs. By adjusting the number of relay UAVs, the network can extend the coverage without replacing the model, supporting the future expansion demand for UAV sizes.

The three models mentioned above meet the system's requirements for real-time path planning, safe and reliable authentication, and scale expansion by providing fast estimation at high altitude, switching between NLOS and LOS conditions, and enabling multi-hop scalability. Additionally, all models complete the parameter settings through offline budgeting without adding to the online computation burden.

3.3. Energy Consumption Model

Rotary-wing UAVs are ideal for data collection missions due to their superior hovering capabilities. During a mission, the energy consumption of an UAV mainly comes from two aspects: flight and data communication. Flight consists of both propulsion and hovering states. The energy cost during data communication refers to the energy used for uploading data from IoT devices. Only when the UAV hovers does this part of energy consumption increase during data collection; otherwise, it remains at zero. In addition, the propulsion energy cost is consistent at each time step, assuming that the flight speed of the UAV remains constant in the horizontal direction. According to the literature [32], the flight power consumption of the UAV is calculated as follows:

$$P(V_a) = P_0 \left(1 + \frac{3V_a^2}{U_{tip}^2} \right) + P_1 \sqrt{1 + \frac{V_a^4}{4v_0^4} - \frac{V_a^2}{2v_0^2}} + \frac{1}{2} d_0 \rho s A V_a^3, \quad (4)$$

where P_0 and P_1 denote the UAV blade profile power and the induced power in hover, respectively, U_{tip} is the tip speed of the rotor blade, and v_0 is the average induced speed in hover, d_0 is the fuselage drag coefficient, ρ is the air density, s is the rotor solidity coefficient, A is the rotor disk area, $V_a = V + V_w$ is the velocity of the UAV relative to the air, V_w is the wind speed, and V is the flight speed of the UAV relative to the ground. $P(V_a) = P_0 + P_1$ if it is hovering.

3.4. Optimization Goals

The effectiveness of the data collection task accomplishment is assessed by considering the data collection rate, flight energy consumption and data upload energy consumption. The energy consumption of UAV i at time slot t is defined as follows:

$$e = \sum_{i=1}^M \sum_{t=1}^T e_t^i, \quad (5)$$

where e_t^i denotes the energy consumption size of UAV i at time slot t . The data collection rate is defined as follows:

$$c_t = \frac{\sum_{n=1}^N c_t(n)}{N}, \quad (6)$$

where $c_t(n)$ denotes whether the data in the IoT device is collected or not, and $c_t(n)$ is 1 if it is collected, and 0 otherwise. In order to achieve efficient data collection and energy management, the optimization objective is to jointly maximize data collection rate and minimize total energy consumption, subject to the following constraints:

$$\max \partial e + \beta c, \quad (7)$$

$$\begin{aligned} \text{C1 : } & P_0^i = P_T^i = \text{BS}, \forall i \\ \text{C2 : } & e_i = \sum_{t=1}^T (e_{f,t}^i + e_{c,t}^i) \leq E_{\max}, \forall i \\ \text{s.t. : } & \text{C3 : } \|P_t^i - P_t^j\| \geq d_{\min}, \forall i \neq j, t \\ & \text{C4 : } x_{\min} \leq x_t^i \leq x_{\max}, y_{\min} \leq y_t^i \leq y_{\max}, \forall i, t \\ & \text{C5 : } \text{lic}_{i,n} \leq \psi_{i,n}, \forall i \in M, n \in N \\ & \text{C6 : } c \geq c_{\text{target}} \end{aligned} \quad (8)$$

where BS is the UAV base location, and constraint C1 indicates that the UAVs must depart from the base at the center of the mission area and return to the base after completing the mission. Constraint C2 ensures that the total energy consumption of each UAV cannot exceed its battery capacity, $E_{\max}$ is the maximum energy. Constraint C3 ensures that multiple UAVs do not congregate too much during the mission, $d_{\min}$ denotes the minimum interference distance. Constraint C4 keeps the UAV's flight path within the mission area. Constraint C5 makes sure that the UAV can only connect with the IoT device after it has been verified, where $\text{Auth}_{i,n}$ of 1 means they have successfully verified each other, and zero means they have not, while $\text{Coll}_{i,n}$ of 1 means the UAV is gathering data from the IoT device, and zero means it is not. Finally, constraint C6 sets a target rate for data collection to ensure that the results after the mission is completed do not fall below the predetermined standard.

4. Algorithm Design

In practical mission scenarios, multi-UAV systems are vulnerable to external security threats such as identity forgery and data tampering, which can significantly reduce system reliability and mission execution security. Meanwhile, with the expansion of the task's scale,

the dynamic changes in the number of UAVs place higher requirements on the system's scalability. To address these challenges, we propose a lightweight chain authentication protocol for scalable IoT environments. Based on LCAP-SIoT and QMIX, we propose a multi-UAV cooperative secure data collection algorithm, termed LS-QMIX, to ensure secure communication while optimizing data collection efficiency.

4.1. Scalable Authentication Protocol Based on Chained PUFs

To ensure communication security and adapt to dynamic changes in mission requirements, we design a lightweight chained authentication protocol for scalable IoT environments. The protocol supports entity registration and mutual authentication as well as dynamic device onboarding, making it suitable for data collection scenarios involving multiple UAVs and IoT devices. As shown in Figure 2, the protocol establishes a secure session between any two entities in the system and secures bidirectional communication between UAVs and IoT devices.

Each device is equipped with a PUF to generate unique challenge–response pairs. Prior to deployment, all devices complete identity registration at the RA (deployed at the UAV base). The registration process is divided into three steps, as described below.

- (1) Device i generates its PUF response as:

$$R_{PUF}^i = PUF(C). \quad (9)$$

- (2) The RA calculates the private key of device i based on:

$$PR_i = Hash(R_{PUF}^i). \quad (10)$$

Then, the ECC is used to generate the public key of the device:

$$PU_i = PR_i \times G, \quad (11)$$

where G is a fixed base point on the elliptic curve.

- (3) Token T_i is the identity of device i for the duration of the task and is calculated as:

$$T_i = Hash(R_{PUF}^{i-1} \parallel Timer \parallel PU_i), \quad (12)$$

where is the system clock and is the PUF response of the last device that was previously successfully registered with the RA.

After completion of the above process, device i securely acquires the public key, private key, token, and the unified prefix $PREFIX$ of the session key from the RA via a secure channel and stores them in its local hardware security module, and then RA stores the token into the Merkle tree, which marks the completion of the registration of device i . RA knows $PREFIX$ and the public and private keys and tokens of all registered devices. The relationship between the device tokens in the system presents a chain relationship, as shown in Figure 3, where the serial number indicates the order of device registration, and the serial number 1 indicates that the registration is completed first.

During the execution of the data collection task, the UAV and the IoT device authenticate each other through tokens generated during the enrollment phase. Figure 4 depicts the different messages exchanged in the protocol, where the authentication mechanism ensures that only legitimate devices can establish a secure session. After the successful establishment of the session, the communication is encrypted using a symmetric session key. The specific process is described below.

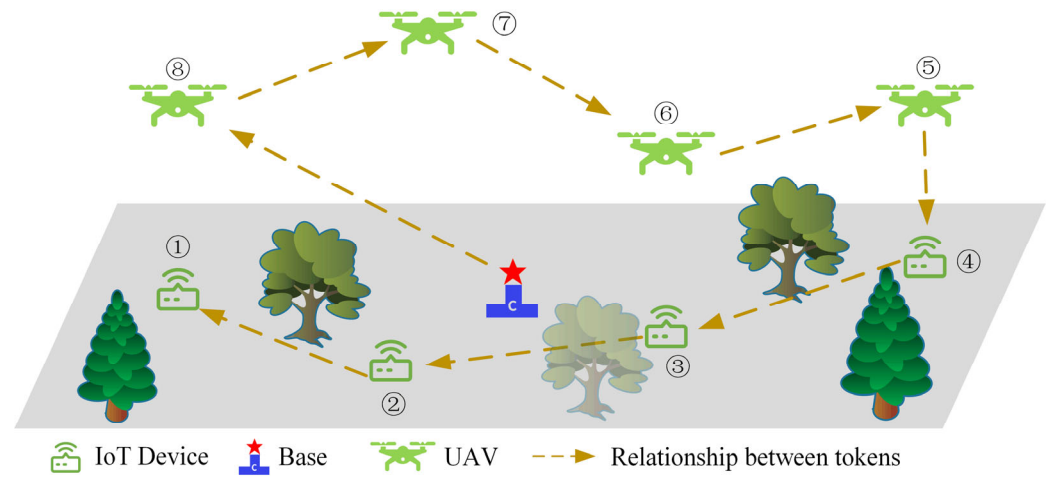


Figure 3. Chained Token Relationship Schematic.

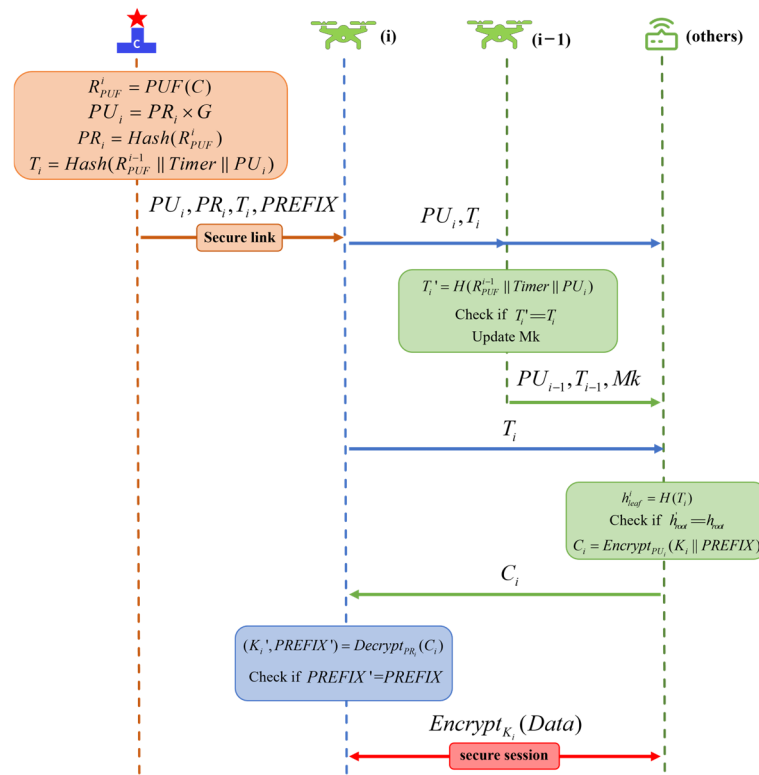


Figure 4. Flowchart of LCAP-SIoT.

- (1) If UAV i is a newly added UAV, broadcast PU_i, T_i to the mission area. otherwise skip to step 3.
- (2) UAV $i - 1$ calculates UAV i 's token based on its own and the received public key:

$$T'_i = H(R_{PUF}^{i-1} || Timer || PU_i). \quad (13)$$

Determine whether T'_i is the same as the received T_i or not. If it is the same, it means that UAV i is a legitimate device, and UAV $i - 1$ as an existing trusted device will add T_i to the Merkle tree, otherwise it is regarded as an illegal device.

- (3) When there is a “to be collected” IoT device in the data collection range, UAV i sends its token T_i and public key PU_i to it.
- (4) The IoT device calculates the hash value of the corresponding Merkle leaf node based on the received T_i :

$$h_{leaf}^i = Hash(T_i). \quad (14)$$

- (5) The IoT device calculates the root hash value h'_{root} accordingly and compares it with the root hash value h_{root} of the locally stored Merkle tree. If they are the same, the authentication passes, otherwise the authentication fails.
- (6) After passing the authentication, the IoT device generates a random session key K_i and encrypts it using PU_i by splicing it with the prefix $PREFIX$, which is subsequently sent to the UAV i to ensure the confidentiality of the key during transmission. The encryption process is as follows:

$$C_i = Encrypt_{PU_i}(K_i || PREFIX), \quad (15)$$

where $Encrypt_{PU_i}$ denotes encryption using the UAV public key.

- (7) The UAV receives and decrypts the data using its private key, and the decryption process is as follows:

$$K_i = Decrypt_{PR_i}(K_i^{enc}), \quad (16)$$

where $Decrypt_{PR_i}$ indicates the decryption operation using the UAV's private key. If $PREFIX'$ is the same as $PREFIX$ in the local security module, the two parties communicate symmetrically, encrypted using session key C to complete the data transmission task.

The protocol fully accounts for the need to easily grow the system and can smoothly add new UAVs, allowing the system to quickly include new equipment for tasks when there are not enough UAV resources, while also ensuring it can expand safely using chained tokens and Merkle tree verification.

4.2. QMIX-Based Algorithm for Secure Multi-UAV Data Collection

4.2.1. Dec-POMDP Model

In the multi-UAV data collection problem, multiple UAVs collaborate to collect data from distributed IoT device nodes in the target area. Due to the instability of wireless communication and the uncertainty of IoT device node states, the UAVs are unable to fully observe the global environment state while performing their tasks. Meanwhile, the limited battery capacity of the UAV requires optimization of its path planning and task assignment to maximize the data collection rate. To this end, this problem is modeled as a decentralized partially observable Markov decision process. According to the system model in this paper, the state space, observation space, action space, and reward function are designed.

State space S : This describes the global state at a certain moment, including the state of each UAV (current position and remaining energy) and the state of the IoT device nodes (location and data collection requirements). For the i th UAV, its local state is defined as follows:

$$s_i = (P_i, E_i, IoTDS_i), \quad (17)$$

where P_i is the 3D coordinates of UAV i , E_i is the residual energy of UAV i , and $IoTDS_i = \{(x_j, y_j, data_j) | j \in \{1, 2, \dots, N\}\}$ denotes the state of IoT device nodes within the communication range of UAV i , including the location coordinates (x_j, y_j) and the data demand state $data_j$ (1 means to be collected and 0 means completed). The joint global state space is

$$S = \{(s_1, s_2, \dots, s_M), S_{IoT}\}, \quad (18)$$

where S_{IoT} denotes the set of states of all IoT device nodes.

Observation space Ω : Due to the partial observability of the system, each UAV can only perceive partial observations from its local environment and cannot directly access the global state. Therefore, the local observation of the i th UAV is

$$o_i = (P_i, E_i, IoTDS_i^{local}, b_i), \quad (19)$$

where b_i denotes the obstacle information in the communication range of UAV i . The joint observation space is defined as follows:

$$\Omega = \{o_1, o_2, \dots, o_M\}. \quad (20)$$

Action Space R : To simplify the problem, we assume that the flight altitude of the UAV remains constant during the mission execution and uses a discretized set of actions for decision-making. For UAV i , its action set is

$$A_i = \{hover, east, south, west, north\}, \quad (21)$$

the elements in the set denote the next optional forward direction of UAV i . The action space is defined as the Cartesian space of all UAV actions. The joint action space is defined as the Cartesian product of the set of actions of all UAVs:

$$A = A_1 \times A_2 \times \dots \times A_M. \quad (22)$$

Joint Action $a \in A$ denotes the choice of actions of all UAVs at a given moment in time, specifically:

$$a = (a_1, a_2, \dots, a_M), \quad a_i \in A_i. \quad (23)$$

Reward function: A reasonable reward function not only needs to incentivize UAVs to improve data collection efficiency but also should comprehensively consider reducing energy consumption, avoiding resource wastage, and promoting collaborative work among agents. In order to realize these goals, we design the reward function from multiple dimensions, specifically including the following aspects:

(1) 2D Distance between UAVs

Vertical separation avoids collisions, but insufficient horizontal spacing may cause communication interference and limit exploration efficiency. We introduce a distance trade-off between UAVs in the reward function. When the unmanned aerial vehicle is smaller than the distance threshold d_{IR} or larger than the maximal distance threshold d_{LOS} , the system will give a penalty ζ_1 to motivate the unmanned aerial vehicle to maintain a reasonable spacing and realize efficient collaboration.

$$r_i^{dist} = \frac{1}{M-1} \sum_{j \neq i} \begin{cases} \zeta_1, & \text{if } d_{ij} < d_{IR} \text{ or } d_{ij} > d_{LOS} \\ \zeta_2, & \text{else} \end{cases}. \quad (24)$$

(2) Data collection

When an uncollected IoT device exists within the UAV data collection communication radius R_{comm} , a positive reward ζ_c is given. If no IoT device exists within the communication range, use the distance $d_{i,nearest}$ between the UAV and the nearest IoT device to measure the reward.

$$r_i^{dev} = \begin{cases} \zeta_3, & \text{if } d_{i,nearest} < R_{comm} \\ -\zeta_3 \frac{1}{\pi} \arctan d_{i,nearest}, & \text{else} \end{cases}. \quad (25)$$

(3) Obstacle avoidance

To ensure the flight safety of the UAV, a fixed penalty ζ_4 is given if the current position is within the range of an obstacle or when passing through an obstacle, and, in general, the penalty value should not be too large to prevent the UAV from being too far away from the obstacle and not choosing to pass through a narrow but safe passage.

$$r_i^{obs} = \zeta_4. \quad (26)$$

(4) Authentication reward for newly joining UAV

The reward is calculated based on the distance $d_{i,i-1}$ between the newly joined UAV i and the trusted UAV $i-1$. When UAV i and UAV $i-1$ are located within the communication range, it means that the authentication process can be carried out, giving a fixed value reward ζ_5 , which indicates that the authentication is successful; otherwise, the smaller $d_{i,i-1}$ is, the bigger the reward is, which encourages UAV i and $i-1$ to get closer as soon as possible. The specific reward function is defined as follows:

$$r^{auth} = \begin{cases} \zeta_5, & \text{if } d_{i,i-1} < R_{comm} \\ \frac{2}{\pi} \zeta_5 \arctan d_{i,i-1}, & \text{else} \end{cases}. \quad (27)$$

Since the newly joined UAV i performs the task, it needs to be guided into the communication range of UAV $i-1$ for authentication before it can collaborate with other UAVs for data collection. The overall reward function needs to be divided into two types: the authentication phase and the collaboration phase. The reward function for the authentication phase is:

$$R_{auth} = \sum_{i=1}^M (\omega_1 r_i^{dist} + \omega_2 r_i^{dev} + \omega_3 r_i^{obs} + \omega_4 r^{auth}), \quad (28)$$

where $\omega_p, p \in \{1, 2, 3, 4\}$ is the weight coefficient of each reward, which is dynamically adjusted according to the task requirements. The reward for the collaboration phase is:

$$R_{collab} = \sum_{m=1}^M (\lambda_1 r_d^i + \lambda_2 r_c^i + \lambda_3 r_{obs}^i), \quad (29)$$

where $\lambda_q, q \in \{1, 2, 3\}$ is the reward coefficient.

A labeled variable phase is added to the state space to indicate the current task phase, phase = 0 for the authentication phase and phase = 1 for the collaborative data collection phase. The final reward function is represented as:

$$R = \begin{cases} \sum_{i=1}^M (\omega_1 r_i^{dist} + \omega_2 r_i^{dev} + \omega_3 r_i^{obs} + \omega_4 r^{auth}), & \text{if phase} = 0 \\ \sum_{m=1}^M (\lambda_1 r_d^i + \lambda_2 r_c^i + \lambda_3 r_{obs}^i), & \text{if phase} = 1 \end{cases}. \quad (30)$$

4.2.2. LS-QMIX Data Collection Algorithm

The LS-QMIX algorithm proposed in this paper realizes the collaborative data collection task for multiple UAVs by introducing the QMIX algorithm. In this algorithm, each UAV is considered an agent and is centrally trained. Each agent estimates its local value function by constructing a local Q-network and employs an action selection mechanism based on a greedy strategy. The agents gradually optimize their data collection paths while exploring the unknown environment. The greedy strategy makes the agent explore new actions with a certain probability to maximize the overall return. During execution, multiple UAVs interact with the environment and accomplish the data collection task by

selecting the best strategy. The experience of the UAVs is stored in an experience playback buffer and used for subsequent network updates.

The overall flow of the LS-QMIX data collection algorithm can be summarized as Algorithm 1. First, the experience playback buffer as well as the local Q-network and the hybrid network for each agent are initialized. Then, in each round of training, the agent selects an action based on the current local observation, receives a global reward for performing the action, and stores the experience into the experience playback buffer. For every certain number of steps, the network parameters are updated based on the experience. Note that if the phase is 0, it means that the newly joining UAV M has not yet been authenticated by the trusted UAV $M + 1$; then R_{auth} needs to be used when performing the joint action to obtain the reward value as well as the network parameter update; otherwise, R_{collab} is used. The final mission execution process is as follows: the existing UAV enters the mission area, senses the IoT device nodes, and triggers the authentication process when it determines the data collection for an IoT device, and all UAVs return to the base after the mission is completed. If a new UAV $M + 1$ joins during the mission execution, the trusted UAV M and UAV $M + 1$ approach each other to complete the authentication. During this process, the Trusted UAV M can continue to collect data from the IoT device, and the authentication and collection proceed in parallel, with the reward function automatically weighing the authentication reward against the collection reward. After the authentication is complete, the newly joined UAV $M + 1$ collaborates with the existing UAVs to complete the remaining data collection tasks.

Algorithm 1: LS-QMIX Algorithm Flow

```

1: Initialization: Experience playback buffer  $D$ , Individual Q network parameter
   ( $\theta, \theta'$ ) and Mixing network parameter  $\phi$ .
2: for episode  $e = 1 \rightarrow E$  do
3:   Initialization environment, phase = 0;
4:   for step  $t = 1 \rightarrow T$  do
5:     for UAV  $m = 1 \rightarrow M + 1$  do
6:       According to  $o_m^t$ , action  $a_m^t$  is chosen with an  $\varepsilon$ -greedy strategy;
7:     end for
8:     Execute action  $a_m = (a_1^t, \dots, a_{M+1}^t)$ ;
9:     The rules for calculating the bonus for UAV  $M$  and UAV  $M + 1$  are decided
       based on phase;
10:    Get the global reward  $r^t$  and the next observation  $o^{t+1}$ ;
11:    Stores  $(o^t, a^t, r^t, o^{t+1})$  in experience buffer  $D$ ;
12:    if  $t$  is Q network update step do
13:      Randomly sample a small batch  $B$  from  $D$ ;
14:      Individual Q values were calculated for each agent;
15:      Calculate the joint Q-value;
16:      Select the reward function according to phase;
17:      Minimize the TD loss, and update the Q-network parameters;
18:    end if
19:    if  $t$  is the target network update step do
20:      Syncs parameters to the target network;
21:    end if
22:    If the UAV completes its mission, end this episode;
23:  end for
24: end for

```

5. Simulation and Analysis

In this section, we consider a 3D simulation environment consisting of ground-based IoT devices, obstacles, and UAVs with different flight altitudes and conduct simulations to evaluate the performance of the proposed secure data collection scheme in a bounded region with ground dimensions of 1000 m $\times$ 1000 m. The UAVs start from a base in the center of the region and collect data from the IoT devices whose path trajectories pass by, assuming that the IoT devices have enough capacity to store data while waiting for the UAVs to collect. It is assumed that all IoT devices have sufficient buffer capacity to store data until collection. The main simulation parameters are shown in Table 1.

Table 1. Simulation parameter setting.

Parameter	Value
Initial energy	300 kJ
Flight altitude	95 m, 100 m, 105 m, 110 m
Flight speed	5 m/s
Wind speed	1 m/s
Blade-tip speed	100 m/s
Hover induced velocity	4.5 m/s
Air density	1.225 kg/m ³
Communication range	150 m
Maximum episodes	18,000
Replay buffer size	100,000
Batch size	256
Learning rate	0.0005
Discount factor	0.99
Target network update frequency	100 steps

5.1. Security Analysis and Comparison

Table 2 compares the security of LCAP-SIoT proposed in this paper with other existing security schemes proposed by Lei et al. [33], Bansal et al. [34], Pu et al. [35], Wazid et al. [36], Srinivas et al. [37], and Ali et al. [38]. In the table, “√” and “X” indicate that the protocol satisfies or does not satisfy the criteria, respectively. Considering legitimate UAVs, IoT devices, and attacker A, LCAP-SIoT is analyzed in detail below based on the security criteria proposed in [35].

Table 2. Comparison of security features.

Scheme	C5.1	C5.2	C5.3	C5.4	C5.5	C5.6	C7	C8	C10	C11	C12
[33]	√	√	√	X	X	√	X	√	X	√	√
[34]	√	√	√	√	√	X	X	X	X	X	X
[35]	√	√	√	√	√	√	√	√	√	X	√
[36]	√	√	√	X	X	√	√	X	√	√	√
[37]	√	√	√	X	X	√	√	X	√	√	√
[38]	√	√	√	X	X	√	√	X	√	√	√
Ours	√	√	√	√	√	√	√	√	√	X	√

[C5] Resistance to known attacks: LCAP-SIoT is resistant to masquerade attacks (C5.1), man-in-the-middle attacks (C5.2), replay attacks (C5.3), node tampering attacks (C5.4), cloning attacks (C5.5), and desynchronizations (C5.6). The token used in the authentication phase is generated by the PUF, which means that attacker A cannot masquerade as a legitimate device and, therefore, cannot launch masquerade attacks and man-in-the-middle attacks. Since the session key negotiation process uses asymmetric encryption, attacker

A cannot access its contents, and replaying previously transmitted messages will not succeed. In the proposed method, the PUF is involved in device identity verification. If an attacker physically captures a legitimate UAV, any tampering with the identity will result in authentication failure, thus resisting node tampering attacks. In addition, PUF is inherently unclonable, so it is able to circumvent cloning attacks.

[C7] Provision of key agreement: In the LCAP-SIoT protocol, UAVs and IoT devices negotiate keys before each communication session to establish a unique session key for each communication session.

[C8] No clock synchronization: In the scheme proposed in this paper, no timestamps are used to participate in the authentication process, so there is no time delay or clock synchronization.

[C10] Mutual Authentication: In LCAP-SIoT, all the devices' tokens are stored in the Merkle tree, and the legitimate devices store the complete Merkle tree, so mutual authentication can be realized.

[C12] Forward secrecy: In the proposed scheme, even if A manages to guess the current session key, it will not affect the security of the next session. Since the key establishment of each session is independent of all other sessions, forward secrecy is satisfied, and backward secrecy is also ensured.

5.2. Cost Analysis of Computing Time

Use T_h , T_e , T_f , and T_a to denote the time required to perform hashing operations, encryption/decryption, fuzzy extractor operations, and analog extractor operations, respectively. The costs associated with these operations have been evaluated in many existing studies [33,39], and in this section $T_h \approx 0.5$ ms, $T_e \approx 8.7$ ms, $T_f \approx 63.075$ ms, and $T_a \approx 2.045$ ms. The computational costs of this paper's methodology and the baseline for the authentication process are presented in Table 3.

Table 3. Comparison of calculation costs.

Scheme	UAV	Ground Station	IoT/Device	Total Cost
[33]	$2T_h$	$4T_h + T_a$	$2T_h \approx 1.0$ ms	$8T_h + T_a \approx 6.045$ ms
[34]	$2T_e$	$2T_e$	-	$4T_e \approx 34.8$ ms
[35]	$4T_e$	$4T_e$	-	$8T_e \approx 69.6$ ms
[36]	$7T_h$	$8T_h$	$16T_h + T_f$	$31T_h + T_f \approx 78.575$ ms
[37]	$7T_h$	$9T_h$	$14T_h + T_e$	$30T_h + T_e \approx 23.7$ ms
[38]	$7T_h$	$7T_h$	$10T_h + T_a$	$24T_h + T_a \approx 14.045$ ms
Ours	$T_h + T_e$	-	$T_h + T_e$	$2T_h + 2T_e \approx 18.4$ ms

From the results, it can be seen that the computational cost of LCAP-SIoT performs the best among other methods except [33,38]. Despite the use of cryptographic operations, we employ a distributed authentication mechanism instead of relying on a centralized terrestrial registry. Doing so allows UAVs and IoT devices to perform only one hash and one encryption/decryption operation during the authentication process, which decentralizes the authentication task and avoids single-point bottlenecks, thus making it easier to support concurrent authentication of large-scale devices. Since the authentication process does not need to rely on the central node, communication during authentication occurs between nodes that are physically close to each other. In a large-scale network environment, this localized communication helps to reduce the overall communication delay and network congestion and to improve the system response speed. Whether the number of IoT device nodes expands or the number of UAVs increases, the system can cope smoothly without performance bottlenecks due to load problems at the authentication center. Through distributed authentication, the system can effectively reduce the data transmission and

communication overhead in the authentication process, and LCAP-SIoT can be more adaptable to scenarios with poor communication conditions than other protocols.

5.3. Communications Cost Analysis

To compute the communication cost, we refer to the method used in [40], which assumes that the hash output (using the secure hash algorithm SHA-1 [41]), the random number, and the identity consist of 160, 128, and 160 bits. In addition, it is assumed that the length of both the PUF query and the response is 128 bits. For the baseline, the timestamp, ECC point, and cipher block are assumed to be 32, 320, and 128 bits, respectively. Table 4 shows the comparison of the communication cost during authentication between the LCAP-SIoT and baseline approaches. In LCAP-SIoT, a total of two messages are exchanged between the UAV and the IoT device during the authentication of the UAV and the IoT device. The first $\langle T_I \rangle$ is 160 bits, and the second $\langle C_i \rangle$ is 128 bits. Therefore, the total communication cost of LCAP-SIoT is $160 + 288 = 448$ bits. The LCAP-SIoT scheme requires only two messages to be exchanged during the authentication period, which reduces the number of messages exchanged as compared to other approaches (e.g., [33], which requires seven messages, and [36,38], which require three messages). The reduction in the number of messages directly reduces the network transmission load and latency, especially in large-scale systems or scenarios with widely distributed nodes.

Table 4. Comparison of communications cost.

Scheme	Number of Messages	Total Cost (bits)
[33]	7	1536
[34]	2	992
[35]	3	1632
[36]	3	1696
[37]	3	1536
[38]	3	1696
Ours	2	448

5.4. Convergence Curves for LS-QMIX

The initial number of UAVs is three, and LS-QMIX expands the number of UAVs to four in the mission to simulate the dynamic addition of new UAVs, all of which travel from the regional center base to collect data from distributed IoT devices. Figure 5 demonstrates the reward changes in LS-QMIX, QMIX, and VDN algorithms under different training rounds. The results show that the reward of LS-QMIX rises slower than that of QMIX at the beginning of training. This is due to the need for LS-QMIX to handle the authentication process of the newly joined UAVs at the beginning of the task execution (phase = 0), at which time the reward function considers both authentication rewards and data collection rewards, resulting in part of the resources being allocated to the authentication phase instead of directly optimizing the data collection efficiency. In contrast, QMIX focuses on collaborative data collection and the reward grows more rapidly. As the training progresses, the reward of LS-QMIX gradually catches up and eventually slightly exceeds that of QMIX. This indicates that after LS-QMIX completes the authentication phase (phase switching to 1), the new UAV is able to efficiently enter the collaborative phase and optimize the data collection efficiency through reasonable task allocation and path planning. The VDN performs poorly in the end due to its simple summation of local Q-values, which cannot adequately capture the interactions between agents.

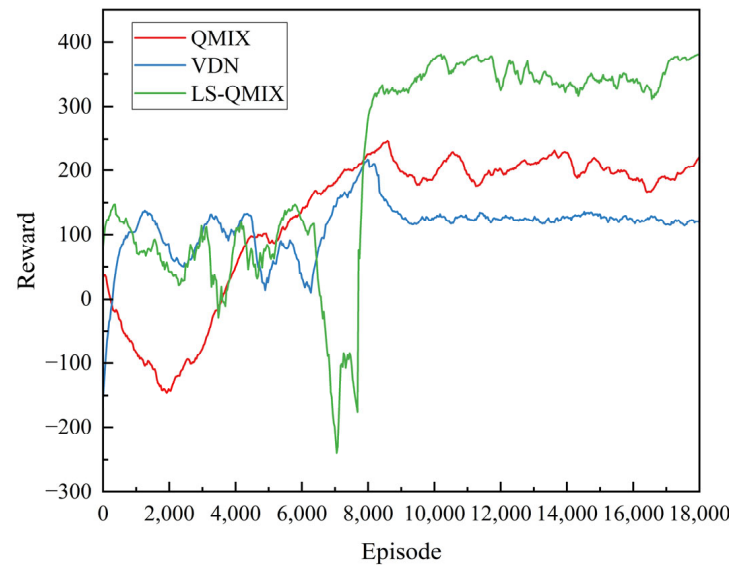


Figure 5. Comparison of rewards.

In a multi-UAV data collection scenario, the Data Collection Ratio is one of the most important metrics to measure the degree of task completion and is used to indicate the proportion of valid data collected by the UAV population to the total amount of data required by the task within a given time or number of rounds. Figure 6 illustrates the data collection rate of each algorithm as a function of the number of training rounds. The collection rate of LS-QMIX rises slower than that of QMIX in the early stages of training for a reason similar to the reward curve: the need for LS-QMIX to prioritize the completion of the authentication of the newly enrolled UAVs results in the trusted UAVs being used primarily for path adjustment rather than for direct data collection. The data collection rate of LS-QMIX ultimately approaches and is slightly higher than that of QMIX. This shows that LS-QMIX not only does not significantly sacrifice collection efficiency after introducing the secure authentication mechanism but also achieves a higher collection rate through collaborative optimization after authentication.

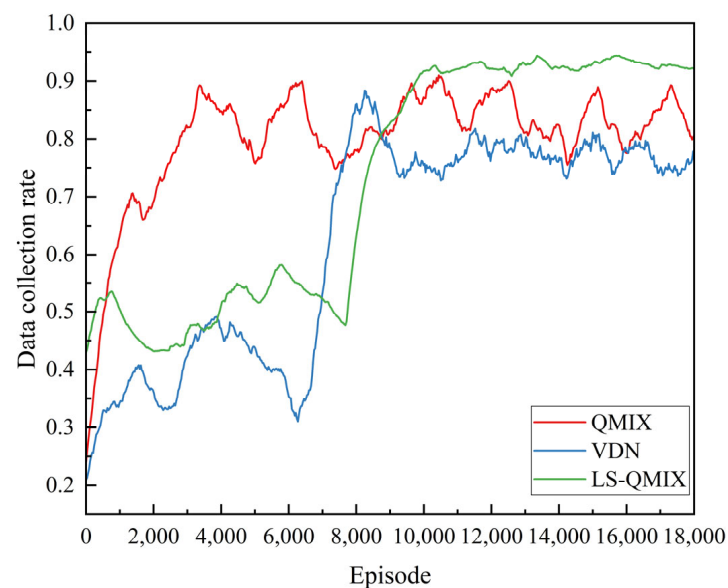


Figure 6. Comparison of data collection rates.

5.5. UAVs Flight Path

To evaluate the performance in the LS-QMIX-based multi-UAV data collection task, Figure 7 analyzes the effect of the trajectories of four UAVs in two different map configurations through the order of IoT device accesses, where UAV1, UAV2, and UAV3 are UAVs that have already been authenticated and participated in the data collection, and UAV4 is the new UAV added to the list. The specificity of the initial paths of UAV4 is demonstrated by its proximity to UAV3 (the previous trusted UAV) and completes the certification process. The red arrow indicates the geographic location where UAV4 was certified by UAV3. Prior to obtaining certification, UAV3 evaluated the flight direction based on a combination of data collection incentives and certification incentives, while UAV4 continuously approached UAV3. After UAV4 completed the certification, it successfully joined the existing data collection mission and started to collaborate with other UAVs. This shows that the way LS-QMIX rewards different parts of the task and the ability of LCAP-SIoT to grow can help new UAVs join easily and make sure that adding new UAVs does not harm the performance of the multi-UAV system during missions, highlighting how well LS-QMIX works when more UAVs are added.

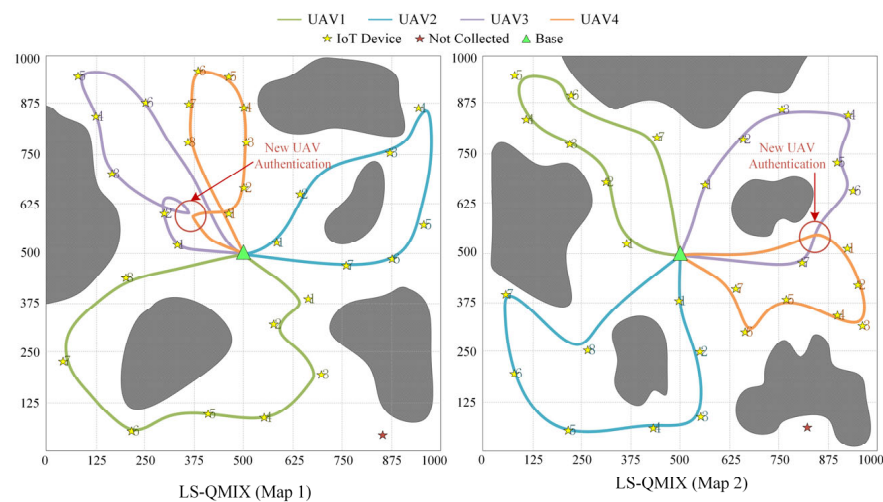


Figure 7. Flight path.

5.6. Authentication Success Rate

In LS-QMIX, a newly joined UAV (UAV4) needs to complete authentication through distributed authentication with a trusted UAV (UAV3). The authentication success rate is affected by (1) the communication range between UAVs, (2) the density of obstacles in the mission area, and (3) the initial distance between the new UAV and the trusted UAV. To evaluate the impact of these factors, we designed the following experiments: the obstacle density is set to be low (10%), medium (30%), and high (60%), and the initial position of UAV4 is randomly set to be at a distance from UAV3 (ranging from 200 m to 800 m). Each scenario was run for 100 test rounds, and the number of successful authentications was recorded.

Figure 8 demonstrates the trend of the authentication success rate with initial distance for different obstacle densities. Increasing obstacle density decreases the authentication success rate, e.g., at an initial distance of 400 m, the success rate decreases from 96% in the low-density scenario to 91% in the high-density scenario, which is due to the obstacles causing obstruction in the communication paths between the UAVs, and the communication becomes NLOS due to obstacle interference, which increases the path loss. In addition, an increase in the initial distance at the same obstacle density also leads to a decrease in the success rate, e.g., in the low-density scenario, the success rate decreases from 98% to 93%

when the distance is increased from 200 m to 600 m. This is due to the fact that the LS-QMIX algorithm reward function encourages the UAV to approach to complete the authentication, and the increase in distance makes the cost of approaching higher, as well as the difficulty of path planning due to the limitation of communication range. Despite these challenges, the authentication success rate of the LS-QMIX algorithm remains between 86% and 98%. It shows that the LCAP-SIoT protocol has high reliability.

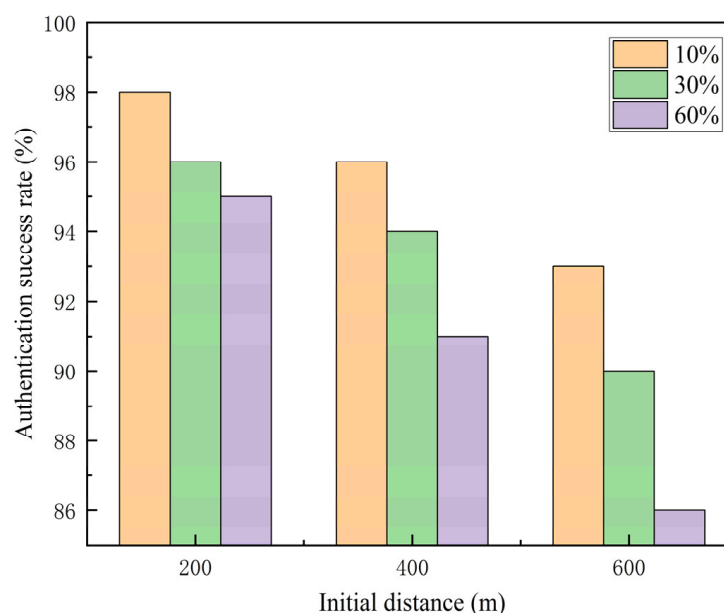


Figure 8. Comparison of authentication success rates.

5.7. Analysis of Task Completion Time and Energy Consumption

This subsection analyzes and compares the performance of LS-QMIX (4 UAVs) and QMIX (3 UAVs) with the number of IoT devices at 20, 40, and 60, with the new UAV (UAV4) joining at the 50th time slot and running 100 tests for each scenario.

LS-QMIX increases the initial overhead due to the authentication phase but is expected to shorten the total time through the collaboration of the new UAVs. The comparison of task completion times is shown in Figure 9, where LS-QMIX reduces about 3.9%, 6.4%, and 9.9% compared to QMIX when the number of devices is 20, 40, and 60, respectively. LS-QMIX needs to complete the authentication to UAV4 in the early stage of the task, and this overhead mainly comes from the path adjustment and communication interaction between UAV4 and UAV3. Despite the increased initial time, the ability of the four UAVs to collaborate, after authentication is completed, outperforms that of the three UAVs in QMIX. As the number of IoT devices increases from 20 to 60, the relative time savings of LS-QMIX improves from 3.9% to 9.9%, this indicates that LS-QMIX scales more efficiently with increasing task size, validating its suitability for large-scale UAV-IoT deployments. Although LS-QMIX performs well in task completion time, the percentage of time savings (3.9%) is relatively limited when the number of IoT devices is small (20). This may be due to the fact that the overhead of the authentication phase is higher in smaller tasks.

As shown in Figure 10, the total energy consumption of LS-QMIX is slightly higher than that of QMIX at all device sizes. This is because, even though LS-QMIX shortens the flight distance for one UAV by assigning tasks at the same time, the overall flight distance goes up because UAV4 and UAV3 have to complete the authentication process first, along with the scattered placement of IoT devices, which results in more total flight distance and higher energy use. However, as the number of IoT devices increased from 20 to 60, the gap decreased from 22.8% to 4.2%. This trend stems from the fact that UAV4 has a higher share

of the authentication phase in small-scale tasks, with limited time to engage in collaborative data collection. In contrast, the data collection phase accounts for a higher percentage of the time spent on larger-scale tasks. Overall, LS-QMIX successfully manages the rise in energy use during high-density tasks and shows it works well when increasing the number of UAVs.

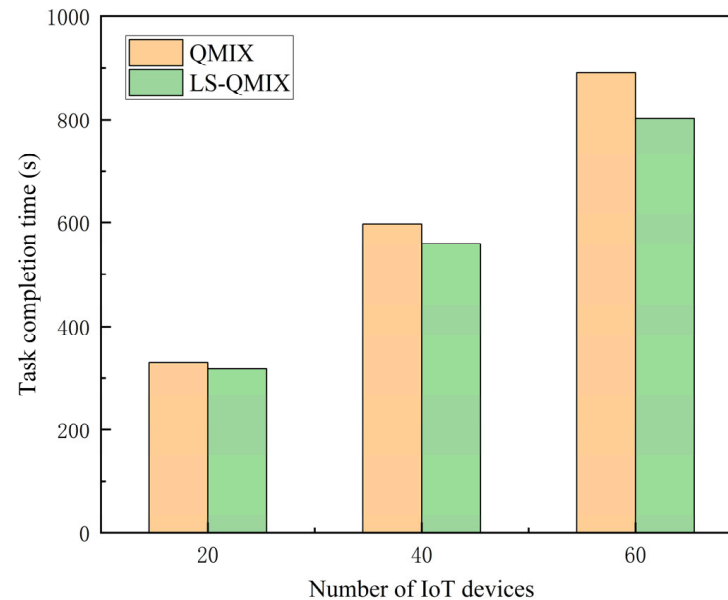


Figure 9. Comparison of task completion time.

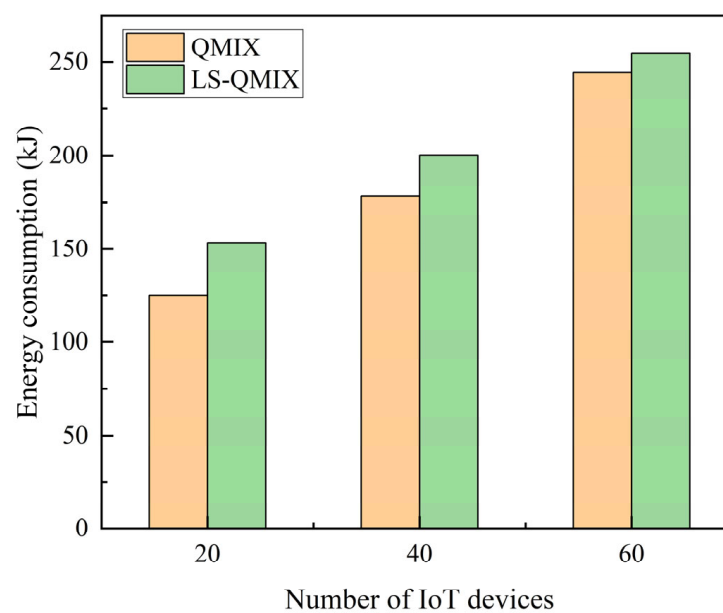


Figure 10. Comparison of energy consumption.

6. Conclusions

In this study, we present a cooperative approach to address the security risks and dynamic scaling requirements related to multi-UAV collection systems by integrating multi-agent reinforcement learning with a lightweight authentication protocol. First, employing distributed token authentication and key negotiation mechanisms based on the chained PUF and Merkle tree authentication mechanism, LCAP-SIoT is designed to efficiently resist security threats like cloning and masquerade attacks, while also facilitating rapid access of new UAVs. Then, using a phase-aware reward function that guarantees security

and adaptability in dynamic environments, we integrate LCAP-SIoT with the QMIX algorithm to handle the integration of newly joined UAVs and propose a scalable and secure multi-UAV data collection method. This method guides the new UAV to complete mutual authentication and then engages in collaborative data collection. The simulation results show that LCAP-SIoT has a low security cost, and LS-QMIX maintains the success rate of authentication between 86% and 98% for the added UAVs in multiple scenarios and maintains the data collection efficiency. Future work will further optimize the model construction of LS-QMIX to explore dynamic obstacle avoidance mechanisms, heterogeneous UAV collaboration mechanisms, and performance validation in real-world environments to support secure data collection tasks in more complex scenarios.

Author Contributions: Conceptualization, J.Z.; formal analysis, J.Z.; investigation, Y.W.; methodology, J.Z., Y.W. and G.H.; project administration, G.H.; resources, G.H.; software, Y.W.; validation, D.C.; writing—original draft, J.Z. and Y.W.; writing—review and editing, D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Open Fund Project of Key Laboratory of Ocean Observation Technology, MNR, 2023klootA04.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Du, J. Application Analysis of IoT Technology in Smart Cities. In Proceedings of the 2021 2nd International Conference on E-Commerce and Internet Technology (ECIT), Hangzhou, China, 5–7 March 2021; pp. 264–269.
2. Ketu, S.; Mishra, P.K. Internet of Healthcare Things: A Contemporary Survey. *J. Netw. Comput. Appl.* **2021**, *192*, 103179. [\[CrossRef\]](#)
3. Ang, L.-M.; Seng, K.P.; Wachowicz, M. Embedded Intelligence and the Data-Driven Future of Application-Specific Internet of Things for Smart Environments. *Int. J. Distrib. Sens. Netw.* **2022**, *18*, 15501329221102371. [\[CrossRef\]](#)
4. Akan, O.B.; Dinc, E.; Kuscü, M.; Cetinkaya, O.; Bilgin, B.A. Internet of Everything (IoE)—From Molecules to the Universe. *IEEE Commun. Mag.* **2023**, *61*, 122–128. [\[CrossRef\]](#)
5. Krejčí, J.; Babiuch, M.; Suder, J.; Krys, V.; Bobovský, Z. Internet of Robotic Things: Current Technologies, Challenges, Applications, and Future Research Topics. *Sensors* **2025**, *25*, 765. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Messaoudi, K.; Oubbati, O.S.; Rachedi, A.; Lakas, A.; Bendouma, T.; Chaib, N. A Survey of UAV-Based Data Collection: Challenges, Solutions and Future Perspectives. *J. Netw. Comput. Appl.* **2023**, *216*, 103670. [\[CrossRef\]](#)
7. Cao, Y.; Cheng, X.; Mu, J. Concentrated Coverage Path Planning Algorithm of UAV Formation for Aerial Photography. *IEEE Sens. J.* **2022**, *22*, 11098–11111. [\[CrossRef\]](#)
8. Wang, W.; Liu, Y.; Srikant, R.; Ying, L. 3M-RL: Multi-Resolution, Multi-Agent, Mean-Field Reinforcement Learning for Autonomous UAV Routing. *IEEE Trans. Intell. Transport. Syst.* **2022**, *23*, 8985–8996. [\[CrossRef\]](#)
9. Liu, Z.; Cao, Y.; Chen, J.; Li, J. A Hierarchical Reinforcement Learning Algorithm Based on Attention Mechanism for UAV Autonomous Navigation. *IEEE Trans. Intell. Transport. Syst.* **2023**, *24*, 13309–13320. [\[CrossRef\]](#)
10. Wu, J.; Sun, Y.; Li, D.; Shi, J.; Li, X.; Gao, L.; Yu, L.; Han, G.; Wu, J. An Adaptive Conversion Speed Q-Learning Algorithm for Search and Rescue UAV Path Planning in Unknown Environments. *IEEE Trans. Veh. Technol.* **2023**, *72*, 15391–15404. [\[CrossRef\]](#)
11. Zhang, J.; Guo, Y.; Zheng, L.; Yang, Q.; Shi, G.; Wu, Y. Real-Time UAV Path Planning Based on LSTM Network. *J. Syst. Eng. Electron.* **2024**, *35*, 374–385. [\[CrossRef\]](#)
12. Lv, H.; Chen, Y.; Li, S.; Zhu, B.; Li, M. Improve Exploration in Deep Reinforcement Learning for UAV Path Planning Using State and Action Entropy. *Meas. Sci. Technol.* **2024**, *35*, 056206. [\[CrossRef\]](#)
13. Fu, X.; Huang, X.; Pan, Q.; Pace, P.; Aloï, G.; Fortino, G. Cooperative Data Collection for UAV-Assisted Maritime IoT Based on Deep Reinforcement Learning. *IEEE Trans. Veh. Technol.* **2024**, *73*, 10333–10349. [\[CrossRef\]](#)
14. Wei, Z.; Zhu, M.; Zhang, N.; Wang, L.; Zou, Y.; Meng, Z.; Wu, H.; Feng, Z. UAV-Assisted Data Collection for Internet of Things: A Survey. *IEEE Internet Things J.* **2022**, *9*, 15460–15483. [\[CrossRef\]](#)
15. Shobiri, B.; Pourali, S.; Migault, D.; Boureau, I.; Preda, S.; Mannan, M.; Youssef, A. LURK-T: Limited Use of Remote Keys with Added Trust in TLS 1.3. *IEEE Trans. Netw. Sci. Eng.* **2024**, *11*, 6313–6327. [\[CrossRef\]](#)

16. Bian, Y.; Zheng, F.; Wang, Y.; Lei, L.; Ma, Y.; Zhou, T.; Dong, J.; Fan, G.; Jing, J. AsyncGBP+: Bridging SSL/TLS and Heterogeneous Computing Power With GPU-Based Providers. *IEEE Trans. Comput.* **2024**, *74*, 356–370. [\[CrossRef\]](#)
17. Shan, T.; Wang, Y.; Zhao, C.; Li, Y.; Zhang, G.; Zhu, Q. Multi-UAV WRSN Charging Path Planning Based on Improved Heed and IA-DRL. *Comput. Commun.* **2023**, *203*, 77–88. [\[CrossRef\]](#)
18. Ejaz, M.; Gui, J.; Asim, M.; El-Affendi, M.A.; Fung, C.; Abd El-Latif, A.A. RL-Planner: Reinforcement Learning-Enabled Efficient Path Planning in Multi-UAV MEC Systems. *IEEE Trans. Netw. Serv. Manag.* **2024**, *21*, 3317–3329. [\[CrossRef\]](#)
19. Silviriati; Narottama, B.; Shin, S.Y. UAV Coverage Path Planning With Quantum-Based Recurrent Deep Deterministic Policy Gradient. *IEEE Trans. Veh. Technol.* **2023**, *73*, 7424–7429. [\[CrossRef\]](#)
20. Liang, J.; Li, Y.; Yin, G.; Xu, L.; Lu, Y.; Feng, J.; Shen, T.; Cai, G. A MAS-Based Hierarchical Architecture for the Cooperation Control of Connected and Automated Vehicles. *IEEE Trans. Veh. Technol.* **2023**, *72*, 1559–1573. [\[CrossRef\]](#)
21. Liang, J.; Yang, K.; Tan, C.; Wang, J.; Yin, G. Enhancing High-Speed Cruising Performance of Autonomous Vehicles through Integrated Deep Reinforcement Learning Framework. *IEEE Trans. Intell. Transp. Syst.* **2025**, *26*, 835–848. [\[CrossRef\]](#)
22. Wang, L.; Wang, K.; Pan, C.; Xu, W.; Aslam, N.; Hanzo, L. Multi-Agent Deep Reinforcement Learning-Based Trajectory Planning for Multi-UAV Assisted Mobile Edge Computing. *IEEE Trans. Cogn. Commun. Netw.* **2021**, *7*, 73–84. [\[CrossRef\]](#)
23. Wu, R.; Gu, F.; Liu, H.; Shi, H. UAV Path Planning Based on Multicritic-Delayed Deep Deterministic Policy Gradient. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 9017079. [\[CrossRef\]](#)
24. Asghar Khan, M.; Ullah, I.; Kumar, N.; Afghah, F.; Barb, G.; Noor, F.; Alqahtany, S. A Certificate-Based Ring Signcryption Scheme for Securing UAV-Enabled Private Edge Computing Systems. *IEEE Access* **2024**, *12*, 83466–83479. [\[CrossRef\]](#)
25. Nyangaresi, V.O.; Jasim, H.M.; Mutlaq, K.A.-A.; Abduljabbar, Z.A.; Ma, J.; Abduljaleel, I.Q.; Honi, D.G. A Symmetric Key and Elliptic Curve Cryptography-Based Protocol for Message Encryption in Unmanned Aerial Vehicles. *Electronics* **2023**, *12*, 3688. [\[CrossRef\]](#)
26. Dong, C.; Jiang, F.; Li, X.; Yao, A.; Li, G.; Liu, X. A Blockchain-Aided Self-Sovereign Identity Framework for Edge-Based UAV Delivery System. In Proceedings of the 2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid), Melbourne, Australia, 10–13 May 2021; pp. 622–624.
27. Li, G.; He, B.; Wang, Z.; Cheng, X.; Chen, J. Blockchain-Enhanced Spatiotemporal Data Aggregation for UAV-Assisted Wireless Sensor Networks. *IEEE Trans. Ind. Inf.* **2022**, *18*, 4520–4530. [\[CrossRef\]](#)
28. Zhang, L.; Xu, J.; Obaidat, M.S.; Li, X.; Vijayakumar, P. A PUF-based Lightweight Authentication and Key Agreement Protocol for Smart UAV Networks. *IET Commun.* **2022**, *16*, 1142–1159. [\[CrossRef\]](#)
29. Tian, C.; Ma, J.; Li, T.; Zhang, J.; Ma, C.; Xi, N. Provably and Physically Secure UAV-Assisted Authentication Protocol for IoT Devices in Unattended Settings. *IEEE Trans. Inf. Forensic Secur.* **2024**, *19*, 4448–4463. [\[CrossRef\]](#)
30. Wang, H.-M.; Zhang, Y.; Zhang, X.; Li, Z. Secrecy and Covert Communications against UAV Surveillance via Multi-Hop Networks. *IEEE Trans. Commun.* **2020**, *68*, 389–401. [\[CrossRef\]](#)
31. Kuzulugil, K.; Hasirci, Z.; Cavdar, I.H. Optimum Reference Distance Based Path Loss Exponent Determination for Vehicle-to-Vehicle Communication. *Turk. J. Electr. Eng. Comput. Sci.* **2020**, *28*, 2956–2967. [\[CrossRef\]](#)
32. Li, Z.; Tong, P.; Liu, J.; Wang, X.; Xie, L.; Dai, H. Learning-Based Data Gathering for Information Freshness in UAV-Assisted IoT Networks. *IEEE Internet Things J.* **2023**, *10*, 2557–2573. [\[CrossRef\]](#)
33. Lei, Y.; Zeng, L.; Li, Y.-X.; Wang, M.-X.; Qin, H. A Lightweight Authentication Protocol for UAV Networks Based on Security and Computational Resource Optimization. *IEEE Access* **2021**, *9*, 53769–53785. [\[CrossRef\]](#)
34. Bansal, G.; Sikdar, B. Location Aware Clustering: Scalable Authentication Protocol for UAV Swarms. *IEEE Netw. Lett.* **2021**, *3*, 177–180. [\[CrossRef\]](#)
35. Pu, C.; Li, Y. Lightweight Authentication Protocol for Unmanned Aerial Vehicles Using Physical Unclonable Function and Chaotic System. In Proceedings of the 2020 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN), Orlando, FL, USA, 13–15 July 2020; pp. 1–6.
36. Wazid, M.; Das, A.K.; Kumar, N.; Vasilakos, A.V.; Rodrigues, J.J.P.C. Design and Analysis of Secure Lightweight Remote User Authentication and Key Agreement Scheme in Internet of Drones Deployment. *IEEE Internet Things J.* **2019**, *6*, 3572–3584. [\[CrossRef\]](#)
37. Srinivas, J.; Das, A.K.; Kumar, N.; Rodrigues, J.J.P.C. TCALAS: Temporal Credential-Based Anonymous Lightweight Authentication Scheme for Internet of Drones Environment. *IEEE Trans. Veh. Technol.* **2019**, *68*, 6903–6916. [\[CrossRef\]](#)
38. Ali, Z.; Chaudhry, S.A.; Ramzan, M.S.; Al-Turjman, F. Securing Smart City Surveillance: A Lightweight Authentication Mechanism for Unmanned Vehicles. *IEEE Access* **2020**, *8*, 43711–43724. [\[CrossRef\]](#)
39. He, D.; Kumar, N.; Khan, M.; Lee, J. Anonymous Two-Factor Authentication for Consumer Roaming Service in Global Mobility Networks. *IEEE Trans. Consum. Electron.* **2013**, *59*, 811–817. [\[CrossRef\]](#)

40. Banerjee, S.; Odelu, V.; Das, A.K.; Chattopadhyay, S.; Rodrigues, J.J.P.C.; Park, Y. Physically Secure Lightweight Anonymous User Authentication Protocol for Internet of Things Using Physically Unclonable Functions. *IEEE Access* **2019**, *7*, 85627–85644. [[CrossRef](#)]
41. Biham, E.; Chen, R.; Joux, A.; Carribault, P.; Lemuet, C.; Jalby, W. Collisions of SHA-0 and Reduced SHA-1. In *Advances in Cryptology—EUROCRYPT 2005*; Cramer, R., Ed.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3494, pp. 36–57, ISBN 978-3-540-25910-7.

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.