Secure Data Transmission Based on Reinforcement Learning and Position Confusion for Internet of UAVs

Xiuzhen Zhu, Limei Lin, Yanze Huang, Xiaoding Wang, Youxiong Que, Behrouz Jedari, and Md. Jalil Piran, *Senior Member, IEEE*

***Abstract*—Ensuring the stability and security of unmanned aerial vehicle (UAV) communication, especially during long-distance missions, is essential for safeguarding against potential attacks. Large-scale UAV communication faces challenges, including eavesdropping threat, data tampering, replay threat, and man-in-the-middle threat. We propose a security information transmission solution based on reinforcement learning and location confusion algorithm (RLPC-SIT) to achieve a secure data transmission between UAVs. First, we leverage the principles of reinforcement learning to identify the most stable transmission routes. Second, we employ location confusion techniques to blur each location of the transmitting UAV with respect to other UAVs. Furthermore, we utilize the concept of message authentication to encrypt the transmitted data, thus making it inaccessible to malicious nodes and preventing forgery. The results of our theoretical analysis and simulation-based experiments indicate that our approach outperforms other security schemes.**

***Index Terms*—Message authentication, position confusion, reinforcement learning, unmanned aerial vehicles (UAVs).**

Manuscript received 27 November 2023; revised 26 January 2024; accepted 16 February 2024. Date of publication 21 February 2024; date of current version 7 June 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62171132, Grant 62102088, and Grant U1905211; in part by the Fok Ying Tung Education Foundation under Grant 171061; in part by the National Key Research and Development Program of China under Grant 2022YFD2301100; in part by the Agriculture Research System of China under Grant CARS-17; and in part by the Natural Science Foundation of Fujian Province under Grant 2021J05228. *(Corresponding authors: Limei Lin; Md. Jalil Piran.)*

Xiuzhen Zhu and Limei Lin are with the College of Computer and Cyber Security, Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350117, Fujian, China (e-mail: fjnuzxz0715@163.com; linlimei@fjnu.edu.cn).

Yanze Huang is with the Fujian Provincial Key Laboratory of Big Data Mining and Applications, School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China (e-mail: yzhuang@fjut.edu.cn).

Xiaoding Wang is with the College of Computer and Cyber Security, Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350117, Fujian, China, and also with the National Key Laboratory for Tropical Crop Breeding, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Sanya 572024, China (e-mail: wangdin1982@fjnu.edu.cn).

Youxiong Que is with the National Key Laboratory for Tropical Crop Breeding, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Sanya 572024, China, and also with the Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture and Rural Affairs, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China (e-mail: queyouxiong@126.com).

Behrouz Jedari is with RAN L1, Nokia Corporation, 02610 Espoo, Finland (e-mail: behrouz.jedari@nokia.com).

Md. Jalil Piran is with the Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea (e-mail: piran@sejong.ac.kr).

Digital Object Identifier 10.1109/JIOT.2024.3368200

I. Introduction

THE Internet of Unmanned Aerial Vehicle (IoUAV) is an emerging field that amalgamates unmanned aerial vehicle (UAV) technology with Internet capabilities to facilitate remote operation, data sharing, intelligent decision making, and collaborative efforts. IoUAV concept encompasses multiple facets, including communication, data processing, security, and application domains. Meanwhile, IoUAV enables UAVs to seamlessly connect with the global Internet ecosystem. With the continuous development of UAV technology, UAVs have got wide-ranging applications across various domains, with distinctive roles [1], [2]. UAVs possess multiple characteristics, including flexibility, portability, and high customizability, which make them indispensable tools in numerous fields. In the realm of reconnaissance, UAVs can execute surveillance missions, gather intelligence, and monitor specific areas without exposing pilots to hazardous situations [3]. In the context of emergency response, UAVs can rapidly reach disaster sites, providing urgent medical services [4]. In the field of communication, UAVs can serve as mobile communication base stations, offering emergency communication support, restoring communication networks, or expanding communication coverage [5].

Fig. 1 illustrates a scenario of UAV information exchange. The UAVs with/without "claws" indicate malicious/normal UAVs, and the links with/without a "link" depict information transmission through malicious/normal UAVs. In this scenario, UAVs perform tasks in various regions and utilize other UAVs as relay nodes to transmit data. Additionally, malicious UAVs can manipulate or forge information, leading to the reception of fake data by destination nodes. Simultaneously, malicious UAVs can result in communication interruptions and an increase in information transmission delays. The selection of communication routes for UAVs is of paramount importance for ensuring the security of information transmission [6], [7], [8]. When UAVs are deployed for missions at remote distances, they often need to relay feedback to their operators. In such scenarios, UAVs may rely on other UAVs as intermediate nodes to relay information back

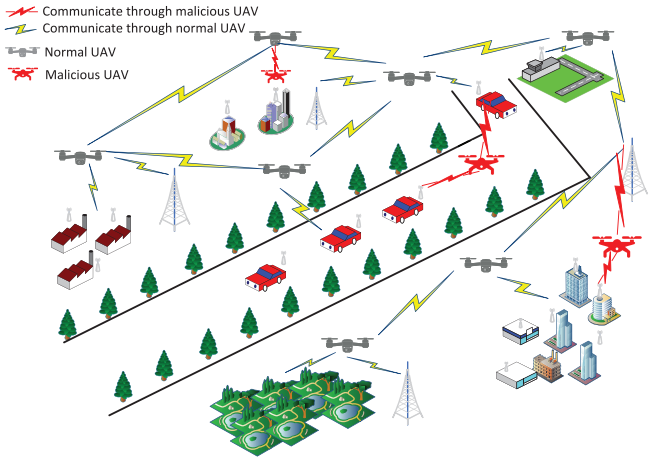

Fig. 1. Data transmission scenarios in the Internet of UAV.

to the operators. However, the position of UAVs changes constantly due to their mobility which results in a dynamic network topology. This dynamic topology can lead to unstable communication links which can create opportunities for malicious activities, such as distributed denial-of-service attacks or unauthorized access, making devices more susceptible to threats like network sniffing or eavesdropping. Therefore, establishing a stable communication link is a prerequisite for ensuring secure communication among UAVs [9], [10], [11].

When a UAV transmits information to the next UAV, they often reveal their current location coordinates [12], [13]. The action which reveal their current location can compromise the UAV's location. Once malicious UAVs obtain location coordinates, they can use this information to infer the tasks being carried out by the UAV, which can expose crucial location details and leading to mission failure [14], [15]. Additionally, if the transmitted information is intercepted by adversaries, the exposure of its content could enable them to forge erroneous feedback, resulting in serious repercussions [16]. For example, in military operations, a large number of UAVs are often deployed for reconnaissance. UAVs can clandestinely observe enemy military deployments and personnel maneuvers, and subsequently transmit the acquired intelligence. This provides the other party with the advantage of taking the initiative in planning. If the enemy discovers the coordinates of the deployed UAVs, they often attempt to destroy the UAVs or intercept the transmitted information, manipulate the data, and provide false information in return.

In summary, large-scale UAV communication faces challenges, including 1) eavesdropping threat; 2) data tampering threat; 3) replay threat; and 4) man-in-the-middle threat. To solve the above problems, this article introduces a novel approach with the following objectives: 1) rapidly and stably transmitting information; 2) safeguarding the location privacy of UAVs; and 3) ensuring the integrity and confidentiality of the transmitted data. The main contributions of this article can be summarized as follows.

1) We propose a reinforcement learning (RL) and location obfuscation solution (RLPC-SIT) to achieve a stable data transmission between UAVs. In RLPC-SIT, we first define a current reward by using sensitivity, stability, sustainability, and distance, which are quantified in (5) and (8)–(10). The current reward could be used to compute the Q-value of surrounding UAVs, which could be applied to determine the next forwarded node. The path obtained by the above choosing node process is optimal due to each next node with the highest Q-value.
2) We employ a location confusion method which enables a UAV to broadcasts messages to its surroundings, receives feedback from neighboring UAVs, and records their position information. The information to be sent is combined with the surrounding location information to create multiple data packets. Based on the Q-values of the UAVs surrounding the UAV, the generated data packets are sent to the next UAV. The next UAV receives multiple data packets, each containing different position information. The drone cannot determine which position information is the real one.
3) We employ an encryption method for the authentication of data transmission. The data to be transmitted is encrypted using the public key of the destination node, and then further encrypted using the private key of the current UAV SK_i. When the next UAV receives the data, it decrypts using the public key of the previous one PK_i; if decryption is unsuccessful, the data is discarded. The UAV selects the next one based on the Q-value, encrypts the data using its private key SK_{i+1}, and transmits it to the next UAV. The same process continues until the message reaches the destination node.
4) Simulation experiments show that RLPC-SIT effectively achieves fast and stable information transmission while protecting the position information of unmanned aircraft. RLPC-SIT demonstrates significant advantages in terms of information delivery rate compared to existing schemes, such as PBQR, GPSR, and AODV [9]. Furthermore, compared to the MOP and EPPS schemes [17], our solution exhibits notable advantages in terms of traceability and entropy.

The remainder of this article is organized as follows. Section II presents related works on routing and privacy protection of UAV. Section III defines the system model. Section IV elaborates the implementation details of the proposed scheme. Section V conducts a security analysis of our proposed solution Section VI evaluates our solution through experimental simulation and compares it with existing solutions. Section VII concludes this article.

II. Related Work

Numerous research has studied data routing in the Internet of UAVs. This section introduces state-of-the-art technologies in UAVs from two aspects: 1) optimal route by machine learning and 2) privacy protection of UAV.

A. UAV Routing

Baek et al. [18] utilized UAVs in wireless sensor networks. In their method, UAVs gathered data from sensors strategically positioned within the network. The solution proposed to prioritize selecting UAV hover points at Voronoi vertices,

optimizing data collection efficiency by maximizing the coverage of neighboring sensors. Yao et al. [10] focused on the offline route planning of UAV for the coverage search mission in a river region. They used Gaussian mixture model to approximate the prior likelihood distribution and they proposed positive/negative greedy method to expand or contract waypoints. Baek et al. [19] made the assumption that the UAV's position for each sensor remains constant. They then focused on enhancing the UAV's hover time by employing the Lagrange multiplier technique. They introduced a geometry-based update algorithm that can be employed to establish preliminary, workable UAV routes for the given task. Ultimately, they derived a near-optimal UAV route by iteratively refining the initial feasible route. Throughout this iterative process, adjustments were made to both the UAV's hover positions and the duration of their stays.

Coelho et al. [20] introduced a novel real-time routing challenge where various types of UAVs are tasked with the collection and delivery of packages. These airborne vehicles have the ability to pick up multiple parcels simultaneously as long as they fit within their maximum capacity. Drawing inspiration from a multifaceted perspective of real-world systems, the study incorporates seven distinct objective functions. The goal is to minimize these objectives using a mixed-integer linear programming model, which is then solved through the application of a matheuristic algorithm. Zhang et al. [21] introduced a UAV swarm network architecture structured in layers, along with an analysis of the ideal quantity of UAVs required. Additionally, they devised a low-latency routing algorithm that relies on partial location data and the network's connectivity. Sajid et al. [22] introduced a hybrid approach known as the hybrid genetic and simulated annealing algorithm, which focuses on reducing travel time in solving the UAV-routing problem. In addition, they put forward the UAV-Oriented MinMin algorithm, aimed at minimizing the makespan when addressing the UAV-route scheduling problem. Hong et al. [23] proposed a proactive topology-aware scheme based on investigating the relationship between the swarm formation control and the network topology to track the network topology change.

Above-mentioned works introduced in this section encompass various applications of UAVs in wireless sensor networks and package delivery systems. Specifically, they explore optimization methods and algorithms for path planning, data collection efficiency, task scheduling, and network topology. However, to achieve a stable routing for intermittent connectivity topologies is overlooked in the majority of existing research.

B. Privacy Protection of UAV

Deebak and Al-Turjman [24] introduced an S-IoD framework for UAV environments, designed for autonomous data collection. To enhance the efficiency of the authentication protocol while minimizing computational overhead, they introduced a lightweight privacy-preserving scheme (L-PPS). L-PPS incorporates elements like secret tokens and dynamic user authentication to expedite the authentication process between communication entities. Ch et al. [25] introduced a solution leveraging blockchain technology (BCT) to enhance the security and privacy of data generated by UAV. Their proposed design was assessed through the implementation of an IoT application within a virtual vehicle monitoring system. The technical details regarding device instructions, authentication, and data integrity are securely stored in a cloud platform. BCT's storage employs Pentatope-based Elliptic curve cryptography and SHA for data privacy assurance. Subsequently, the data is archived on a public blockchain rooted in Ethereum to facilitate seamless BCT transactions. The system relies on the Ganache platform for BCT, ensuring robust data protection and privacy.

Lv et al. [26] adopted BCT to address the privacy preservation challenge associated with UAV big data. Their proposed privacy protection approach employs a cryptographic system based on number theory research units for encrypting blockchain data. They have conducted a privacy analysis to substantiate the fulfillment of security prerequisites. Li et al. [27] introduced a lightweight symmetric encryption algorithm that relies on SM4, coupled with a relevant key negotiation and updating mechanism. This was implemented to safeguard the confidentiality of communication contents. Additionally, they introduced an adapted BLS signature scheme that combines with the Merkle Hash tree to ensure the integrity and authenticity of transmitted data packets. Furthermore, they put forth an online/offline revocable identity-based group signature scheme, seamlessly integrating it into their framework to achieve UAV anonymity, traceability, and revocability. Importantly, this integration comes with the advantage of minimal key management overhead and high operational efficiency.

Wang et al. [28] presented an architectural framework for collaborative learning among UAVs, harnessing BCT. This framework facilitates the secure exchange of local model updates and validation of contributions without the need for a central curator. In addition, they developed a privacy-preserving algorithm, leveraging local differential privacy techniques, to safeguard the privacy of updated local models while maintaining desirable learning accuracy. Furthermore, they utilized a two-tier RL-based incentive system to encourage UAVs to share high-quality models, even when precise knowledge of network parameters is unavailable in practical scenarios. Wang et al. [29] introduced SEAL, an all-encompassing framework designed to tackle the challenges of strategy-proof, equitable, and privacy-preserving computation offloading for UAVs. SEAL employs a strategy-proof reverse combinatorial auction mechanism to optimize the offloading of tasks for UAVs while adhering to practical constraints, guaranteeing economic resilience, and maintaining polynomial-time efficiency. Using smart contracts and hash chain micro payments, SEAL implements a fair on-chain exchange protocol to achieve the seamless completion of batch payments and computing results in multiround auctions. Moreover, they developed a privacy-preserving off-chain auction protocol with the aid of a trusted processor to effectively safeguard the bid privacy of the vehicles involved.

The aforementioned research primarily focuses on data collection, security, privacy protection, and computation offloading in UAV environments. However, there is a lack of research on UAV message authentication and location privacy protection in the context of five-a-side football.

In summary, despite the extensive research in areas such as UAV routing and privacy protection, there is still a lack of relevant studies on location privacy protection and secure information transmission for UAVs operating in intermittent networks.

III. System Model

With global temperatures rising and climate change intensifying, agriculture is facing unprecedented challenges. Traditional crop monitoring methods fall short in providing real-time and comprehensive information. Crop monitoring is crucial for agricultural management as it directly affects crop growth, health, and ultimately, yield. In this regard, UAVs equipped with real-time imaging systems have emerged as a revolutionary technology. They fly over farmlands, carrying high-resolution cameras and sensors capable of capturing precise images and data.

In order to gain a more comprehensive understanding of the application of UAVs in agriculture, and to better comprehend the communication relationships and behavioral patterns among drones, as well as to provide farmers with more accurate, efficient, and intelligent agricultural management solutions, we can consider the following points. First, UAVs offer real-time monitoring, enabling farmers to immediately understand the condition of their fields. This timeliness allows them to quickly respond to issues, such as pest outbreaks, disease spread, or drought, reducing losses and increasing productivity. Second, agricultural UAVs can provide personalized monitoring for each individual plot of land, helping farmers to use resources more efficiently and minimize waste. Additionally, agricultural UAVs improve cost-effectiveness in agricultural production. Compared to purchasing expensive satellite imagery services, using agricultural UAVs typically costs less. Farmers can rent or buy UAVs as needed, without the need for costly service subscriptions. Finally, these UAVs not only provide detailed data analysis but also integrate advanced artificial intelligence technology, enabling intelligent analysis and prediction of farmland, providing farmers with precise agricultural production advice, helping them make wiser decisions, and improving the efficiency and quality of agricultural production.

However, large-scale UAV communication faces challenges including unstable communication links, insufficient protection of drone location privacy, and insecurity in the transmitted information. To deal these problems, we introduce the concept of contact graph, which reflects the connectivity relationships among individuals or entities. We abstract the UAV data transmission system as a contact graph, where UAVs are represented as nodes. Nodes are connected by edges to indicate direct communication between UAVs. Since UAVs are in constant flight, edges between nodes may not always be present. Each drone in the network is considered to be a state of the agent; thus, the set of all nodes in the network forms the state set. The communication in the network is abstracted as the behavior of the nodes. The nodes can only communicate with neighboring nodes, so the neighboring nodes of a node constitute the action set. Some of the notations and their definitions used in this article are listed in Table I.

TABLE I
List of Notations and Their Description

Symbols	Definition
S	The set of environmental states
m	Message to be send from drone to destination
S_t	The action taken by the agent at time t
R_t	the reward obtained by the agent at time t
γ	discount factor
$V_\pi(s_t)$	Value function of strategy π in state s_t
SL	The set of location information
(i,j,k)	Location coordinates of UAV
F_i	Representative of a specific drone i
$D(F_1,F_2)$	Distance between UAVs F_1 and F_2
$D(t)$	The distance between UAVs at time t
$SFR(t)$	Speed factor at time t
$RPD(t)$	Relative position distance between UAVs at time t
$STR(t)$	Stability factor at time t
$COF(t)$	Continuity factor at time t
$DDN(t)$	Distance from destination node at time t
α	Learning factor
PK_E	Public key of operator to encrypt m
SK_{AU}	Private key of operator to authentication m
Res	Request tamp
Loc	Position information
ID	Identity information
RA	Communication radius of UAV
$NE(t)$	Degree of UAV at time t
$N(t)$	The number of surrounding UAV at t
AS_{A_E}	Expect size of anonymous set
AS_A^t	Anonymous set at time t
P_{tsr}	Probability of traceable success ratio

A. Threat Model

To ensure communication security in UAVs, following potential threats are modeled.

1) *Eavesdropping Threat:* This threat involves attackers monitoring and intercepting communication content during the transmission process of the unmanned aircraft system. Attackers can employ various techniques, such as wireless signal interception or network monitoring, to steal sensitive information being transmitted, such as location data, instructions, or other critical data.
2) *Data Tampering Threat:* This threat involves attackers maliciously altering data during the transmission process of the unmanned aircraft communication. Attackers may modify transmitted instructions, sensor data, or other crucial information to deceive the unmanned aircraft system or disrupt its normal operation.
3) *Replay Threat:* This threat involves attackers capturing and recording encrypted data from unmanned aircraft communication and later replaying that data. This can result in the unmanned aircraft system receiving duplicate instructions or data, leading to errors or security vulnerabilities.

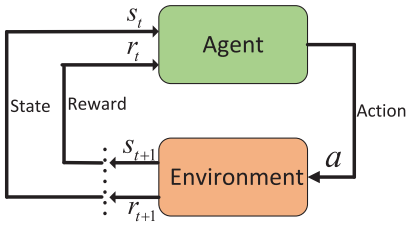

Fig. 2. RL system model.

4) *Man-in-the-Middle Threat:* This threat involves attackers inserting themselves as intermediaries in the communication path of the unmanned aircraft. They intercept, modify, or deceive the communication content. Attackers can carry out such attacks on the communication link between the unmanned aircraft and the ground control station or other intermediate points.

B. Reinforcement Learning

RL is a process in which an agent continually interacts with its environment to improve its decision making and actions. Initially, the agent perceives the current environmental state and then takes a specific action based on the current state, thus altering the current environment and transitioning to the next environmental state. The agent receives a reward from the current environment, as shown in Fig. 2.

In the process of an agent selecting actions, each step's action is determined based on the previous actions. Assuming the probability of taking a certain action in the decision-making process is P which is described as

$$P\big[S_{t+1}|S_t\big] = P\big[S_{t+1}|S_t, S_{t-1}, \ldots\big]. \tag{1}$$

The equation above primarily signifies that each step's decision action is determined by the preceding actions.

When devising strategies in the current environment, the uncertainty of the environment and the long-term objectives should be considered. Therefore, we construct a value function to evaluate the rewards obtained when implementing a strategy. Since each step taken by the agent offers various strategy choices, the reward value function returned at each iteration is distinct. Therefore, we represent the expected returns of adopting a particular strategy in the current state using the following equation:

$$V_\pi(s) = E_\pi\left[\sum_{k=0}^{+\infty} \gamma^k R_{t+k+1}|S_t = s\right] \tag{2}$$

where, γ represents the discount factor. It means that the farther away from the current state, the lower the expected reward.

For strategy π, the value function of the system at time t in state s_t is represented as

$$V_\pi(s_t) = R_t + \gamma \sum_{s_{t+1}\in S} P(s_{t+1}|s_t, a_t) V_\pi(s_{t+1}). \tag{3}$$

The objective is to find a strategy that maximizes the rewards obtained by the agent, that is

$$V_\pi^*(s_t) = \max\left\{R_t + \gamma \sum_{s_{t+1}\in S} P(s_{t+1}|s_t, a_t) V_\pi(s_{t+1})\right\}. \tag{4}$$

C. Position Confusion

In the UAV system, UAVs use other UAVs as intermediate nodes to transmit information to these intermediaries, who then forward the information to the operator. representative of a specific drone. The information sent by UAVs includes sensitive details like their current location and identity. As UAVs continuously move, they periodically provide feedback to the operator, resulting in their positions being continually revealed to other UAVs. If malicious nodes are present among the forwarding nodes, they can obtain the flight trajectories and the locations of UAVs, leading the leakage of drone location privacy.

To address the issue of location privacy leakage in UAV systems, we employ a location obfuscation-based method. UAVs utilize GPS which is a global positioning system to determine their location coordinates, then broadcast the request to nearby UAVs, collecting responses and recording the location information of surrounding UAVs. UAVs combine the data they intend to send with the location information to form multiple data packets, each containing a set of location coordinates. Consequently, when the next hop node receives these data packets, it cannot ascertain the true sender of the information

D. Message Encryption and Authentication

When forwarding messages, ensuring the confidentiality of transmitted data is crucial, as interception by malicious nodes may expose the message to adversaries. We employ an encryption mechanism to safeguard the integrity and trustworthiness of data transmission. Initially, we encrypt a message using the source node's public key, ensuring that even if intercepted, malicious nodes cannot access the intended content. Subsequently, the message is authenticated using the UAV's private key. For example, as the information is transmitted from UAV F_A to UAV F_B. F_A encrypts the messages which are included different UAV's location by his private key, and then F_B decrypts them by using F_A's public key. It is worth noting that any UAVs can decrypt the messages using F_A's public key. However, in the event of interception by a malicious UAV, the inability to re-encrypt the message without F_A's private key ensures that legitimate UAVs cannot decrypt the information. So, if F_B can decrypts the message by F_A's public key successfully, it will transmit the message to the next UAV. And if F_B decrypts the message by F_A's public key unsuccessfully, F_B will discard the message. This process guarantees the trustworthiness of the information (see Fig. 3).

IV. Proposed Scheme

A. RL for Optimal Path Selection

When the data arrives a node F_i at time t_i, the node F_i will periodically broadcast request to its neighboring nodes and receive feedback from neighboring nodes at time t_i' within the communication range. The node F_i will select

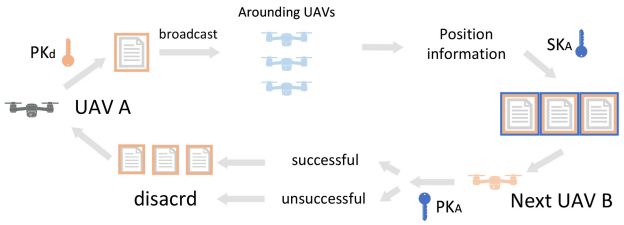

Fig. 3. Message encryption and authentication: F_A transmitted data to UAV F_B. F_A encrypts the messages which are included different UAV's location by his private key, and then F_B decrypts them by using F_A's public key. And if F_B decrypts the message by F_A's public key unsuccessfully, F_B will discard the message.

next forwarded node based on four key factors: 1) sensitivity; 2) stability; 3) sustainability; and 4) distance between node F_i and destination node. The node F_i uses the above four factors to calculate the reward for each neighbor node, and then selects the node with the greatest reward as the next hop node. We will quantify these four factors through mathematical expressions and use them to calculate the reward.

1) Sensitivity: Due to the long-distance nature of tasks typically performed by UAVs and the need for timely feedback as tasks unfold, messages are rarely directly relayed to the source node. Instead, they often pass through multiple intermediate nodes, which increase the message delivery time. To expedite the delivery of messages to the source node, it is essential to convey information to intermediate nodes with strong signal strength and rapid forwarding capabilities. Therefore, we define a sensitivity factor $\mathrm{SFR}(t_i)$ to quantify the influence of neighbor nodes on node F_i at time t_i as follows:

$$\mathrm{SFR}(t_i) = e^{1-\frac{t_i'-t_i}{t_i}}. \tag{5}$$

2) Stability: Due to the continuous mobility of UAVs, the connections between nodes are highly unstable, which results in a constantly changing state of the UAV network topology. If the contact time between nodes is too short, it can lead to the connection failures, causing information forwarding to be unsuccessful. Therefore, we have defined the stability factor $\mathrm{STR}(t_i)$ to quantify the stability of the connection between the node F_i and its neighbor F_j. Assume that the coordinates of the node F_i are (x_i, y_i, z_i) and the coordinates of the node F_j are (x_j, y_j, z_j). According to the distance coordinate formula, the actual distance $D_{t_i}(F_i, F_j)$ between two nodes F_i and F_j at time t is calculated by

$$D_{t_i}(F_i, F_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}. \tag{6}$$

We also assume that the communication range of a node is RA. If the relative distance between two nodes F_i and F_j changes very little during the time period, it indicates that there is a relatively stable connection between them. The relative positional distance between two UAVs F_i and F_j during the time period $t_i \sim t_i'$, denoted as $\mathrm{RPD}(t_i)$, is described as

$$\mathrm{RPD}(t_i) = \begin{cases} e^{-|D_{t_i}(F_i,F_j) - D_{t_i'}(F_i,F_j)|}, & D_{t_i}(F_i, F_j) \le RA \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

To reduce errors, we record m measurements between time windows t_i and t_i', denoted as $t_i^1, t_i^2, \ldots, t_i^m$, and calculate their average to define the stability factor $\mathrm{STR}(t_i)$ as follows:

$$\mathrm{STR}(t_i) = \frac{1}{m}\sum_{j=1}^{m} \mathrm{RPD}\left(t_i^j\right). \tag{8}$$

3) Sustainability: When a node has a large number of neighbors, it has a higher probability of selecting different neighbors for message forwarding, leading to a better chance of creating sustainable paths. As time progresses, the position of a node changes which leads to variations in the number of neighboring nodes. Therefore, relying on the number of neighboring nodes at a single time instant can result in significant errors. Let $\mathrm{NE}(t_i)$ be the node degree at time t_i. Similar to the stability factor, to reduce errors, we take we record m measurements between time windows t_i and t_i', denoted as $t_i^1, t_i^2, \ldots, t_i^m$, and calculate their average to define the sustainability factor $\mathrm{COF}(t_i)$ as follows:

$$\mathrm{COF}(t_i) = \frac{1}{m}\sum_{j=1}^{m}\left(1 - \frac{1}{\log_2\left(\mathrm{NE}\left(t_i^j\right) + 1\right)}\right). \tag{9}$$

4) Distance From Destination Node: In addition to the aforementioned three factors, we need to account for the direction of the information transmission, in order to find the optimal relay node in the direction of the destination node. Therefore, we introduce a distance factor, $\mathrm{DDN}(t_i)$, to quantify the direction of information transmission. After receiving responses, UAVs calculate distance $D_{t_i}(F_i, F_d)$ between their current location coordinates (x_i, y_i, z_i) and the coordinates of the destination node (x_d, y_d, z_d). The distance $D_{t_i}(F_i, F_d) = \sqrt{(x_i - x_d)^2 + (y_i - y_d)^2 + (z_i - z_d)^2}$. The distance factor can be described as follows:

$$\mathrm{DDN}_{t_i}(F_i, F_d) = e^{\frac{1}{1+D_{t_i}(F_i,F_d)}}. \tag{10}$$

5) Update Q-Value: We use $R(t_i)$ to represent the reward obtained at time t_i, and the calculation formula for $R(t_i)$ can be expressed as follows:

$$R(t_i) = \mathrm{SFR}(t_i) \times \mathrm{STR}(t_i) \times \mathrm{COF}(t_i) \times \mathrm{DDN}(t_i). \tag{11}$$

We can conclude the formula of Q-value as follows:

$$Q\left(s_{t_{i+1}}, a_{t_{i+1}}\right) = (1-\alpha)Q\left(s_{t_i}, a_{t_i}\right) + \alpha\left[R(t_i) + \gamma \times \max Q\left(s_{t_i}, a_{t_i}\right)\right]. \tag{12}$$

From the above equation, it is evident that when the connections between nodes are more stable, their responses are more sensitive, sustainability is stronger, distance from destination node is closer, the immediate reward $R(t_i)$ and Q-value are higher. Each node operates according to its current action with the maximum Q-value.

In order to better explain the selection process of the next forward node, we will take our experimental data as an example to illustrate the solution. In our experiment, there are 100 UAVs. The model assumes that when data is transmitted to the destination node, all UAVs make the maximum effort to participate in the information transfer. According to the

principles of RL, the maximum reward is obtained when the task is completed. That is, when the data is transferred to the destination node, the reward obtained at this moment is the greatest. Let $R_{\max}$ be the maximum reward. When the data reaches the destination node, the UAV ceases further broadcasting, thus we can obtain that $t_i = t_i'$. By (5), we have

$$\mathrm{SFR}(t_i) = e^{1-\frac{t_i'-t_i}{t_i}} = e^{1-\frac{0}{t_i}} = e^{1-0} = e. \tag{13}$$

Second, the relative position information of the UAV is not recorded when the data reaches the destination node. Hence, we can obtain that $D_{t_i}(F_i, F_d) = D_{t_i'}(F_i, F_d)$. According to (7), we have

$$\mathrm{RPD}(t_i) = e^{-|D_{t_i}(F_i,F_d) - D_{t_i'}(F_i,F_d)|} = e^{-|0|} = 1. \tag{14}$$

Therefore, combining with (7) and (8), we have

$$\mathrm{STR}(t_i) = \frac{1}{m}\sum_{j=1}^{m} \mathrm{RPD}\left(t_i^j\right) = \frac{1}{m}\sum_{j=1}^{m} 1 = 1. \tag{15}$$

Third, according to model assumption, the maximum degree of the destination node is 100 at time t_i^j, i.e., $\mathrm{NE}(t_i^j) = 100$. According to (9), we have

$$\begin{aligned}\mathrm{COF}(t_i) &= \frac{1}{m}\sum_{j=1}^{m}\left(1 - \frac{1}{\log_2\left(\mathrm{NE}\left(t_i^j\right)+1\right)}\right)\\ &= \frac{1}{m}\sum_{j=1}^{m}\left(1 - \frac{1}{\log_2(100+1)}\right)\\ &\approx 0.85.\end{aligned} \tag{16}$$

Fourth, when the data reaches the destination node at time t_i, the distance between the current location and destination location at time t_i is 0, i.e.,

$$D_{t_i}(F_i, F_d) = \sqrt{(x_i - x_d)^2 + (y_i - y_d)^2 + (z_i - z_d)^2} = 0 \tag{17}$$

By (10), we have

$$\mathrm{DDN}(t_i) = e^{\frac{1}{1+D_{t_i}(F_i,F_d)}} = e^{\frac{1}{1+0}} = e. \tag{18}$$

According to (11), we can obtain that the reward of the arrival destination node at time t_i as follows:

$$\begin{aligned}R(t_i) &= \mathrm{SFR}(t_i) \times \mathrm{STR}(t_i) \times \mathrm{COF}(t_i) \times \mathrm{DDN}(t_i)\\ &= e \times 1 \times 0.85 \times e\\ &\approx 6.\end{aligned} \tag{19}$$

Therefore, when data is transmitted to the destination node, the reward at this point is maximal $R_{\max} = 6$. Otherwise, it means that the information has not been transmitted to the destination node.

The network nodes in the network constitute a set of environmental states. Each node employs message broadcasting and response reception for sampling. Given that both states and actions are discrete, and the reward function has an upper bound, our algorithm can converge within a finite time frame. Consequently, our RL approach is effective. The next section presents our proposed reinforcement learning algorithm to find the optimal path is as follows (see Algorithm 1).

Algorithm 1 Optimal Route Selection

Input:
State set $S = \{s_{t_1}, s_{t_2}, \ldots, s_{t_n}\}$ which is the location of the UAVs; action set $A = \{a_{t_1}, a_{t_2}, \ldots, a_{t_n}\}$ which is the next UAV that the data packet is transmitted.
Output:
The optimal route.

for each episode **do**
  for each step of episode **do**
    F_i chooses a_t from A derived from Q-value;
    F_i broadcasts the message, records the current time as t_i;
    F_i broadcasts message during from t_i to t_i';
    F_i receives responses and records each time as t_i';
    for j from $i+1$ to n **do**
      F_i computes $SFR(t_i) = e^{1-\frac{t_i'-t_i}{t_i}}$;
      F_i computes $STR(t_i) = \frac{1}{m}\sum_{j=1}^{m} RPD(t_i^j)$;
      F_i computes $COF(t_i) = \frac{1}{m}\sum_{j=1}^{m}\left(1 - \frac{1}{log_2(NE(t_i^j)+1)}\right)$;
      F_i computes $DDN(t_i) = e^{\frac{1}{D(t_i)}}$;
      F_i computes $R(t_i) = SFR(t_i) \times STR(t_i) \times COF(t_i) \times DDN(t_i)$;
      F_i update Q-value $Q(s_{t_{i+1}}, a_{t_{i+1}} = (1-\alpha)Q(s_{t_i}, a_{t_i}) + \alpha\left[R(t_i) + \gamma \times maxQ(s_{t_i}, a_{t_i})\right]$;
      F_i selects next forwarded UAV which the Q-value is the maximum;
    end for
  end for
end for
return The optimal route.

B. Position Confusion and Message Encrypt

When UAVs are providing feedback to the operator during task execution, they require the assistance of intermediate UAVs as relay nodes to facilitate information transmission. The UAV sends data packets to the next UAV, and the format of the data packet *Res* can be described as $Res := \{Loc, T, \{ID, \{m\}_{PK_d}\}_{SK_i}\}$. *Loc* represents the geographical location, T denotes the feedback time. $\{m\}_{PK_d}$ represents the message m which is encrypted with the destination's public key PK_d, while $\{ID, \{m\}_{PK_d}\}_{SK_i}$ is encrypted with the i_{th} UAV's private key SK_i. The benefit of this approach is to provide identity authentication for the feedback messages, preventing adversaries from accessing the current data packets and impersonating an identity to feed incorrect information to the operator.

The UAVs broadcast messages to their surroundings, receiving responses from neighboring UAVs and obtaining their location information. At the same time, the location information of the UAVs are recorded as loc_1, loc_2, loc_n, ... The location information are used to form a series of data packets, denoted as $Res_1 : \{\mathrm{loc}_1, T, \{ID, \{m\}_{PK_d}\}_{SK_i}\}$, $Res_2 : \{\mathrm{loc}_2, T, \{ID, \{m\}_{PK_d}\}_{SK_i}\}$, and so on. The UAVs subsequently relay these data packets to the base station. It is important to note that PK_d and SK_i are counterfeit keys, and therefore, the packets containing PK_i and SK_i cannot be tampered by malicious UAVs. Because the private key of a UAV is kept exclusively by itself and not publicly disclosed, other UAVs can use the UAV's public key for decryption. However, since the core information to be transmitted is

Algorithm 2 Position Confusion and Message Authentication

UAV F_A broadcasts message to the surrounding area;
Surrounding UAVs response message;
UAV F_A records the position coordinates of the surrounding UAVs, such as loc_1, loc_2, loc_3, …, loc_n;
for i from 1 to n **do**
  F_A calculates the $D_{t_A}(F_A, F_i)$;
  if $D_{t_A}(F_A, F_i) > R_{\min}$ **then**
    A deletes loc_i;
  else
    F_A generates request stamp $Res_i = \{Loc, T, \{ID, \{m\}_{PK_d}\}_{SK_A}\}$;
  end if
end for
F_A sends Res to next UAV by Q-value;
for each UAV **do**
  F_i use F_{i-1}'s public key to decrypt the data packet.
  if F_i decrypt unsuccessfully **then**
    F_i discards the packet
  end if
end for
for Each Res **do**
  if Destination node uses SK_d and the advanced's private key to decrypt the Res successfully **then**
    Destination node receives the m successfully;
  else
    Destination node discards the Res;
  end if
end for

encrypted, malicious UAVs cannot view the content of the information. For example, when F_A selects the next UAV based on the Q-value, if the data packet is intercepted by a malicious node during transmission and decrypted using F_A's public key, it cannot decrypt the core information because it does not know the private key of a destination node. Additionally, without F_A's private key, it is impossible to re-encrypt after decryption. When the next UAV receives the decrypted data packet and finds that it cannot be decrypted using F_A's public key, it discards the packet. Therefore, the message cannot be forged by other malicious nodes, and only data packets that can be decrypted successfully are allowed to remain. This approach not only ensures that the location of UAVs is not revealed but also guarantees protection against receiving forged feedback, thereby enhancing the credibility of the data. Meanwhile, when a UAV uses the location information of nearby UAVs, it should satisfy the condition that both UAVs are within each other's communication range, denoted as $D(F_A, F_B) \leq R_{\min}$, where $D(F_A, F_B)$ refers to the distance between UAVs F_A and F_B, $R_{\min}$ refers to the minimum communication radius between F_A and F_B. Algorithm 2 is the pseudocode for position fuzzy scheme.

V. Security Analysis

We analyze the effectiveness of the location privacy by considering the size of the anonymous set and traceability. The anonymity set represent the uncertainty between the UAV's real location and other possibility locations. Traceable rate indicates the possibility that the attacker can track the UAV. The detailed descriptions are listed as below.

A. Anonymity Set Size

An anonymity set in an anonymous system refers to a group of users at a specific time point who share similar characteristics or attributes with a particular user. The anonymity set plays a crucial role in protecting user privacy, enhancing anonymity, strengthening security, and countering statistical analysis in anonymous systems. We use loc_A to represent the real route of UAV F_A when it send the message. The anonymity set of F_A denoted by AS_A is the set of all possible location Loc'_A that can be confused with loc_A. The size of UAV F_A's anonymous set $|AS_A|$ is the number of elements in the set. Assuming that $p(\text{loc}_A, \text{loc}_B)$ indicates the probability of the attacker regards loc_B as the real route of UAV F_A, then the anonymity set can be expressed as follows:

$$AS_A = \{SL_B \mid p(\text{loc}_A, \text{loc}_B) \neq 0\}. \tag{20}$$

Since UAVs record the location information of surrounding UAVs when broadcasting messages, leading to the transmission of false data packets to the next node. Let us assume that at time t_0, a total of N_{t_0} surrounding UAV location information has been recorded. The information of the location will result in N_{t_0} false messages being sent to the next UAV, and at this point, the anonymous set size is $N_{t_0}+1$. When the UAV arrives at the next location at time t_1, it will once again broadcast to determine the number of surrounding UAVs at the current time, denoted as N_{t_1}. Thus, the UAV's anonymous set at this time becomes $N_{t_0} + N_{t_1} + 1$. This process continues until the UAV finish the task. Assuming it has to pass through η location in total, we can derive the expected size of the anonymous set AS_{A_E} as

$$AS_{A_E} = \frac{1}{\eta} \sum_{i=0}^{\eta} N_{t_i}. \tag{21}$$

B. Entropy of Anonymity Set

Entropy represents the degree of uncertainty in the relationship between the real location of the UAV F_A and all other possible location. Entropy is often used as a precise measure of location privacy in UAV

$$H_A = -\sum_{\text{loc}_B \in AS_{A_E}} p(\text{loc}_A, \text{loc}_B) \times \log_2(A, B). \tag{22}$$

When the entropy is high, the confused degree of the UAV's location in the anonymity set get larger and the location privacy of the UAV is more secure.

C. Traceable Rate

When a UAV sends data packets to next UAV at time t_0, the probability that the base station can correctly identify the true location is $(1)/(N_{t_0} + 1)$. At time t_1, when the UAV arrives at the next location, the probability that this base station can correctly identify the true location becomes $(1)/(N_{t_0} + N_{t_1} + 1)$. We assume that the anonymity set of time t is AS_A^t. Therefore, the overall traceability rate P_{tsr} for the entire path is denoted as follows:

$$P_{tsr} = \prod_{i=0} \frac{1}{AS_A^{t_i}}. \tag{23}$$

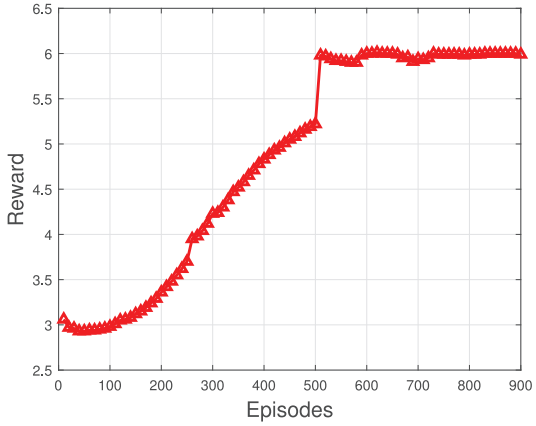

Fig. 4. Trend of reward value over episodes.

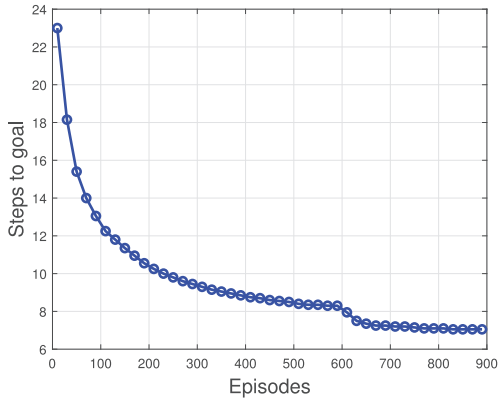

Fig. 5. Trend of steps to goal over episodes.

D. Delivery Rate

Delivery rate refers to the probability of successfully delivering a message to the destination node. We conducted a transmission of 1000 messages using UAVs and observed the number of messages that arrived at different cutoff times. Assuming that n messages reached the destination node, the delivery rate P_{delivery} is calculated as

$$P_{\text{delivery}} = \frac{n}{1000}. \tag{24}$$

VI. Simulation

We use Network Simulator-2 to simulate our system. We configure a rectangular area with dimensions of 5 km × 5 km, within which obstacles are randomly placed. The number of UAVs is 100. The obstacles are distributed randomly with areas ranging from 10 to 1000 m^2. We set the speed range of the UAVs to be between 10 and 15 m/s. In terms of simulating data transfer rate and latency, each simulation experiment has a duration of 250 s and involves transmitting 1000 pieces of information. When UAVs send request stamps to the destination node, the UAVs generate false coordinate information before sending a request. We assume a time interval of 3 s between sending requests. In terms of data transfer rate and latency, we compare our scheme against three other schemes: 1) PBOR; 2) GPSR; and 3) AODV [9]. In the context of path anonymity, we compare our scheme with MOP and EPPS [17].

A. RL Training Effect

We conducted a total of 900 training rounds, during which we recorded data every ten rounds. Figs. 4 and 5 depict the training outcomes. It is evident that when the training reached 500 rounds, the nodes identified the optimal pathway for information transmission, at which point the reward was maximized. We also simulated the random selection approach and the local optimal strategy, and the rewards for both the random selection and local optimal strategies were lower than those for the RL approach. Further details are available in Table II.

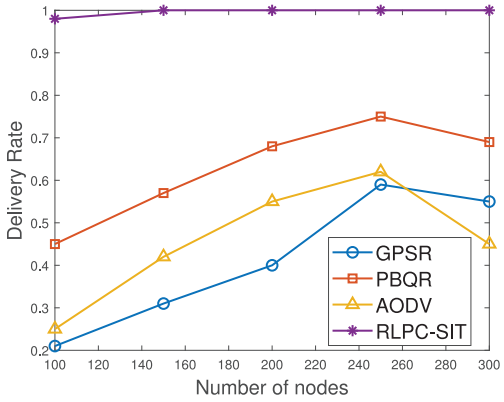

Fig. 6. Comparison of delivery rate under the different numbers of nodes.

TABLE III
Comparison of Delivery Rate Under the Different Numbers of Nodes

Nodes	GPSR	PBQR	AODV	RLPC-SIT
100	0.21	0.45	0.25	0.98
150	0.31	0.57	0.42	1
200	0.4	0.68	0.55	1
250	0.59	0.75	0.62	1
300	0.55	0.69	0.45	1

B. Packet Delivery Ratio and Delay

Fig. 6 and Table III show that as the number of nodes increases, the data transmission rate also increases. This is because with more nodes, there are more choices when selecting the next-hop node, increasing the probability of selecting more responsive and stable nodes. When the number of nodes reaches 250, the data packet transmission rates for the PBQR, GPSR, and AODV schemes are 74%, 59%, and 62%, respectively. In contrast, RLPC-SIT achieves a data transmission rate of 100%.

Fig. 7 and Table IV illustrate the relationship between packet delivery rate and time. We configured the nodes to be 200 in number. It can be observed that the packet delivery success rates for each approach increase as time progresses. When the time reaches 250 ms, GPSR, PBQR, and AODV achieve data transmission success rates of 21%, 45%, and

TABLE II
COMPARISON OF REWARD FOR THREE SCHEMES

	Reward in episodes									
	10	50	100	150	200	250	300	350	400	450
reinforcement learning	3.06	2.93	2.98	3.12	3.36	3.7	4.23	4.52	4.83	5.05
local optimum	2.35	2.54	2.58	2.70	2.91	3.21	3.67	3.92	4.19	4.38
random selection	3.35	3.70	2.53	2.77	2.79	3.62	3.29	3.95	3.51	2.56
	460	500	550	600	650	700	750	800	850	900
reinforcement learning	5.08	5.22	5.92	6	6	5.94	5.99	5.99	6	5.99
local optimum	4.40	4.52	5.13	5.20	5.20	5.15	5.19	5.19	5.20	5.19
random selection	3.58	3.08	3.14	3.88	2.75	3.82	3.58	3.41	2.73	3.73

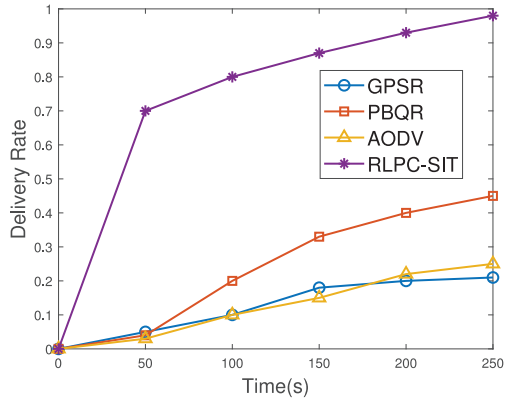

Fig. 7. Comparison of delivery rate under the different times.

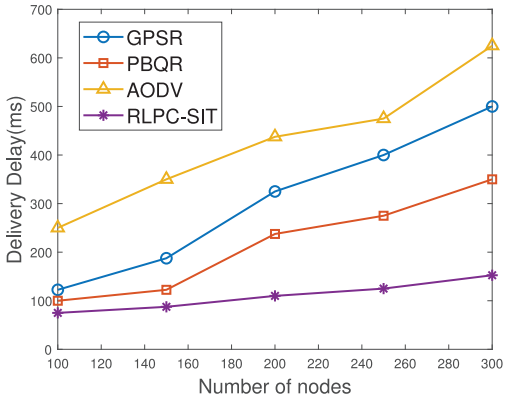

Fig. 8. Comparison of delivery delay under the different numbers of nodes.

TABLE IV
COMPARISON OF DELIVERY RATE UNDER THE DIFFERENT TIMES

Time(s)	GPSR	PBQR	AODV	RLPC-SIT
50	0.05	0.04	0.03	0.7
100	0.1	0.2	0.1	0.8
150	0.18	0.33	0.15	0.87
200	0.2	0.4	0.22	0.93
250	0.21	0.45	0.25	0.98

25%, respectively. In contrast, our approach attains a success rate of 98%.

Fig. 8 illustrates that our proposed scheme outperforms PBQR, GPSR, and AODV in terms of end-to-end latency. The improved performance is attributed to the comprehensive consideration of the destination node's location, the quality of neighboring nodes, and the use of RL methods to avoid getting trapped in local optima during node selection. Consequently, the latency is significantly reduced compared to the other three schemes. From the experimental results, it can be observed that as the number of nodes increases, there are more choices available when nodes select their next-hop nodes, which requires additional computation time. And the increase in the number of hops from the source to the destination node also leads to an increase in latency For instance, with 100 nodes, the latency for PBQR, GPSR, and AODV schemes is 40, 50, and 100 ms, respectively, while our scheme achieves a latency of 30 ms.

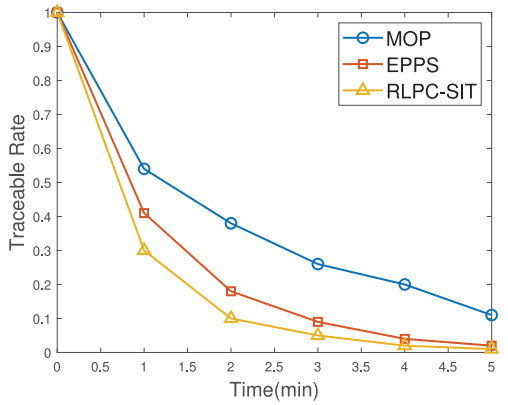

Fig. 9. Trend of traceable rate under the different times.

C. Traceable Rate and Entropy

Fig. 9 shows the relationship between traceable rate and time. It can be observed that as time progresses, the traceability decreases. This is because, with the passage of time, UAVs send virtual locations to different base stations, significantly expanding the anonymity set of UAVs, which, in turn, leads to a decrease in UAV traceability. From the experimental results, it is evident that at the 5-min mark, the traceability for the MOP scheme is 0.1, for the EPPS scheme is 0.02, and for our scheme is 0.01. Although our scheme and the EPPS scheme exhibit similar performance after 5 min, it is worth noting that from the outset, our scheme's anonymity level is lower

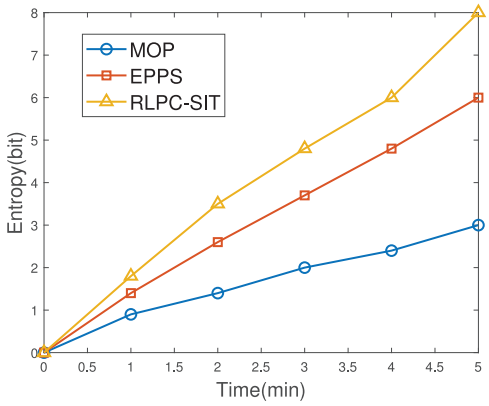

Fig. 10. Trend of entropy under the different times.

than that of the EPPS scheme, which gives our scheme a competitive advantage.

Fig. 10 illustrates the entropy of the anonymity sets for three strategies. Initially, the entropy is zero because the anonymity set size is one, indicating that the UAVs do not broadcast any information. At the 5-min mark, the entropy of the anonymity set for RLPC-SIT is 8 bits, while for MOP and EPPS, the anonymity set sizes are 3 and 6 bits, respectively. Clearly, the effectiveness of our proposed approach surpasses that of the other two methods.

VII. Conclusion

This article presented a dedicated solution for ensuring the security of UAV information transmission. Building upon the use of RL to identify the optimal information transmission routes, we employed location obfuscation techniques to safeguard the privacy of the transmitting UAV's location. To enhance the integrity and confidentiality of the transmitted information, we encrypted the data, preventing malicious nodes from accessing the content of the transmitted information and forging it. Theoretical and simulation results demonstrated that our solution offers superior security compared to other approaches. In future work, we will further investigate the application of more complex encryption technologies in large-scale drone networks.

Acknowledgment

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] M. Petrlík, T. Báča, D. Heřt, M. Vrba, T. Krajník, and M. Saska, "A robust UAV system for operations in a constrained environment," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2169–2176, Apr. 2020.

[2] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2nd Quart., 2015.

[3] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: State-of-the-art in applications and challenges," *J. Netw. Comput. Appl.*, vol. 166, Sep. 2020, Art. no. 102706.

[4] Z. Huang, C. Chen, and M. Pan, "Multiobjective UAV path planning for emergency information collection and transmission," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6993–7009, Aug. 2020.

[5] S. Hayat, E. Yanmaz, T. X. Brown, and C. Bettstetter, "Multi-objective UAV path planning for search and rescue," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2017, pp. 5569–5574.

[6] M. M. Azari, G. Geraci, A. Garcia-Rodriguez, and S. Pollin, "UAV-to-UAV communications in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6130–6144, Sep. 2020.

[7] Y. Zeng, J. Lyu, and R. Zhang, "Cellular-connected UAV: Potential, challenges, and promising technologies," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 120–127, Feb. 2019.

[8] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, Apr. 2019.

[9] Y. Sun, Y. Lin, and Y. Tang, "A reinforcement learning-based routing protocol in VANETs," in *Proc. Int. Conf. Commun., Signal Process., Syst.*, 2017, pp. 2493–2500.

[10] P. Yao, Z. Xie, and P. Ren, "Optimal UAV route planning for coverage search of stationary target in river," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 822–829, Mar. 2019.

[11] C. Qu, W. Gai, J. Zhang, and M. Zhong, "A novel hybrid grey wolf optimizer algorithm for unmanned aerial vehicle (UAV) path planning," *Knowl.-Based Syst.*, vol. 194, Apr. 2020,Art. no. 105530.

[12] C. T. Cicek, H. Gultekin, B. Tavli, and H. Yanikomeroglu, "UAV base station location optimization for next generation wireless networks: Overview and future research directions," in *Proc. 1st Int. Conf. Unmanned Veh. Syst.-Oman (UVS)*, 2019, pp. 1–6.

[13] C. Liu, W. Yuan, Z. Wei, X. Liu, and D. W. K. Ng, "Location-aware predictive beamforming for UAV communications: A deep learning approach," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 668–672, Mar. 2021.

[14] A. S. Abdalla, K. Powell, V. Marojevic, and G. Geraci, "UAV-assisted attack prevention, detection, and recovery of 5G networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 40–47, Aug. 2020.

[15] Y. Zhou et al., "Improving physical layer security via a UAV friendly jammer for unknown eavesdropper location," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11280–11284, Nov. 2018.

[16] S. Enayati, D. Goeckel, A. Houmansadr, and H. Pishro-Nik, "Location privacy protection for UAVs in package delivery and IoT data collection," *IEEE Internet Things J.*, vol. 10, no. 23, pp. 20586–20601, Dec. 2023.

[17] J. Cui, J. Wen, S. Han, and H. Zhong, "Efficient privacy-preserving scheme for real-time location data in vehicular Ad-Hoc network," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3491–3498, Oct. 2018.

[18] J. Baek, S. I. Han, and Y. Han, "Energy-efficient UAV routing for wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1741–1750, Feb. 2020.

[19] J. Baek, S. I. Han, and Y. Han, "Optimal UAV route in wireless charging sensor networks," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1327–1335, Feb. 2020.

[20] B. N. Coelho et al., "A multi-objective green UAV routing problem," *Comput. Oper. Res.*, vol. 88, pp. 306–315, Dec. 2017.

[21] Q. Zhang, M. Jiang, Z. Feng, W. Li, W. Zhang, and M. Pan, "IoT enabled UAV: Network architecture and routing algorithm," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3727–3742, Apr. 2019.

[22] M. Sajid, H. Mittal, S. Pare, and M. Prasad, "Routing and scheduling optimization for UAV assisted delivery system: A hybrid approach," *Appl. Soft Comput.*, vol. 126, Sep. 2022, Art. no. 109225.

[23] L. Hong, H. Guo, J. Liu, and Y. Zhang, "Toward swarm coordination: Topology-aware inter-UAV routing optimization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10177–10187, Sep. 2020.

[24] B. D. Deebak and F. Al-Turjman, "A smart lightweight privacy preservation scheme for IoT-based UAV communication systems," *Comput. Commun.*, vol. 162, pp. 102–117, Oct. 2020.

[25] R. Ch, G. Srivastava, T. R. Gadekallu, P. K. R. Maddikunta, and S. Bhattacharya, "Security and privacy of UAV data using blockchain technology," *J. Inf. Secur. Appl.*, vol. 55, Dec. 2020, Art. no. 102670.

[26] Z. Lv, L. Qiao, M. S. Hossain, and B. J. Choi, "Analysis of using blockchain to protect the privacy of drone big data," *IEEE Netw.*, vol. 35, no. 1, pp. 44–49, Jan./Feb. 2021.

[27] T. Li et al., "Energy-efficient and secure communication toward UAV networks," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10061–10076, Jun. 2021.

[28] Y. Wang, Z. Su, N. Zhang, and A. Benslimane, "Learning in the air: Secure federated learning for UAV-assisted crowdsensing," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1055–1069, Apr.–Jun. 2020.

[29] Y. Wang, Z. Su, T. H. Luan, J. Li, Q. Xu, and R. Li, "SEAL: A strategy-proof and privacy-preserving UAV computation offloading framework," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 5213–5228, 2023.