THE ASIAN BULLETIN OF BIG DATA MANAGMENT
Vol. 4. Issue 2 (2024)
https://doi.org/ 10.62019/abbdm.v4i02.13

ASIAN BULLETIN OF BIG DATA MANAGEMENT
ISSN (Print): 2959-0795
http://abbdm.com/
ISSN (online): 2959-0809

IRS Integrated UAV Communication to Enhance Physical Layer Security: A Deep Reinforcement Learning approach

Touseef Hussain, Muhammad Shahwar, Syed Zain Ul Abdin *, Zakria Zaheen, Naveed Jan, Muhammad Shoaib

Chronicle

Article history

Received: April 20, 2024
Received in the revised format: May 3, 2024
Accepted: May 14, 2024
Available online: May 17, 2024

Touseef Hussain, Muhammad Shahwar, Syed Zain Ul Abdin, Zakria Zaheen are currently affiliated with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China.
Email:touseefhussain098@gmail.com
Email: shahwarmughal10@gmail.com
Email: zainshah208@gmail.com
Email: zakriyazaheen10@gmail.com

Naveed Jan is currently affiliated with the Department of Information Engineering Technology, University of Technology Nowshera, 24100, KPK, Pakistan.
Email: naveed.jan@uotnowshera.edu.pk

Muhammad Shoaib is currently affiliated with the Department of Information and Communication Engineering, Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan.
Email: malik.engrshoaib@gmail.com

Corresponding Author*

Abstract

The paper investigates the security of an intelligent reflecting surface (IRS) assisted unmanned aerial vehicle (UAV) network, where a base station (BS) transmits confidential information to the ground user (GU) via IRS-assisted UAV. The study explores using UAVs with IRS to enhance the security and reliability of wireless communication systems, particularly in the presence of eavesdroppers and friendly jammers. The beamforming at IRS-assisted UAV and UAV trajectory are jointly formulated as a non-convex optimization problem, which is solved by the deep reinforcement learning (DRL) algorithm to maximize the sum secrecy rate of GU with the aid of jammer. We proposed a dual-DDPG (D3PG) algorithm that utilizes the deep deterministic policy gradient (DDPG) structure to effectively address these dual non-convex problems of the UAV-trajectory optimization and the UAV-IRS beamforming optimization. The proposed algorithm's effectiveness and robustness are demonstrated through simulation results, with the IRS significantly enhancing the sum secrecy rate. Extensive simulations show that the proposed DRL-based D3PG scheme outperforms the traditional optimization schemes and ordinary DQN schemes.

Keywords: Secure wireless communication, UAV communication, Deep Deterministic Policy Gradient.

INTRODUCTION

In the past, military operations relied on human pilots for missions with high levels of risk. However, there has been a recent discovery of more applications for Unmanned Aerial Vehicles (UAVs) in civilian domains (Zhang et al., 2022). They encompass various tasks such as rescue and search operations, inspection, and policing. The setup includes various components and multiple links to communication. Each link has the responsibility of transmitting specific types of information and data. In these networks, there should be three different types of links based on the transmitted information: radio communication, Satellite link, and User 2 User(U2U) (Lu et al., 2022). The radio communication links transmit telemetry data, control audio, and video information. In addition, satellite links are responsible for transmitting GPS, meteorological, and weather information, as well as the data transferred through radio communication links. Due to the UAVs' ability (Yang et al., 2020) to move quickly, easy deployment, floating capability, and low maintenance cost, they can be used in various civil applications. Effective communication is vital for the seamless operation of UAVs in the cutting-edge next-generation (6G) network, which is driven by the power of AI. Unmanned Aerial Vehicles (UAVs) possess the capacity to deliver efficient and secure wireless communication. This is because of their versatility and user-friendly interface. critical emergency areas where Ground BS are unavailable (Lacovelli et al., 2023). It can work in the In certain high-traffic locations, such as a bustling sports or music event, the utilisation of flying UAVs can be implemented to provide immediate service and relieve the burden on cellular networks. One more application of UAVs (Zhang et al., 2023) involves serving as a flying relay to establish connections between nodes that are far apart (Wang et al., 2022). The rapid development of UAVs has sparked a wide range of applications in various sectors, such as military, civilian, and commercial (Samir et al., 2021). These applications include aerial inspection, cargo transport, search and rescue operations, as well as video streaming, among others. UAVs are becoming increasingly popular as aerial communication platforms.

They are being used to improve the overall performance and secrecy rate of the system. Furthermore, the incorporation of UAVs mobile communication or with BS to greatly improve network reliability and throughput. UAVs have a greater tendency to establish LoS wireless links, which leads to improved air-ground(A2G) communication quality. Typically, high-frequency waves have a tendency to be blocked by physical barriers because of their strong directivity and limited diffraction. Wireless propagation environments can be reconfigured by the IRS through software-controlled reflection. IRS is a planar array that houses numerous passive reflective elements. These elements have the ability to individually adjust the phase of the incoming signal and reflect the waves accordingly. IRS has been identified as a promising contender for the upcoming wireless network. Nevertheless, when it comes to wireless communication systems, the UAV-assisted systems have a unique challenge. Their air-to-ground channels, known as LoS, make them more vulnerable to eavesdropping. This means that potential eavesdroppers can easily intercept these systems, which poses a greater risk to the confidentiality (Guo et al., 2021) of the data. Based on the specific applications of UAVs, such as communication networks, it is clear that there is a growing demand to improve the security of data transmission on UAV-assisted systems. These systems often rely on transmitting sensitive information over wireless links. Therefore, they can be vulnerable to interception by individuals with malicious intentions, such as eavesdroppers and jammers. In the past, conventional cryptographic methods have been employed to safeguard communication channels.

However, with the rapid advancements in computer technologies, these methods may be vulnerable to sophisticated attacks. In addition to encryption, PLS techniques also rely on the unique properties of the wireless channel to enhance security. Integrating IRS into UAV communication infrastructure is an increasingly popular solution for addressing PLS issues. These surfaces are constructed using affordable passive mirrors that can manipulate wave propagation to boost signal strength and direct signals towards the intended receivers while reducing interference. Furthermore, the study focuses on IRS-enabled secure UAV communication. Utilising the principles of trajectory optimisation, power transmission, and the reflective properties of IRS, we aim to enhance the security of data transmission by minimising the impact of eavesdropping along with friendly jammer. The research work considers previous studies on UAV communication systems, physical layer security, and optimisation techniques. Its aim is to provide a fresh perspective on the development of secure communication protocols for UAV networks, with a special focus on integrating IRS and DRL algorithms to solve complex optimisation problems. The research findings will greatly contribute to the advancement of UAV-assisted communication systems, leading to the development of more secure and reliable wireless networks in the future.

The focus of the research is on the security of the physical layer of a UAV network. In this scenario, the base station acts as a transmitter, sending sensitive data to the ground station or Ground User(GU) with the assistance of a UAV-IRS. The secure transmission of data is compromised by the presence of an eavesdropper who can also be near the jamming device. In these networks, there are ongoing efforts to mitigate any interference and ensure the secure transmission of data to its intended destination. An effective approach to ensuring security at the physical layer involves utilising carefully planned UAV paths and controlling transmit power. This strategy considers both mobility and power limitations to achieve a cooperative solution. In particular, solving the optimisation problem for trajectory and the backscatter coefficient matrix can be quite challenging due to its non-convex nature. Therefore, the DRL-based algorithm is utilised due to the presence of two intricate convex problems or the utilisation of two DRL algorithms to achieve optimal optimisation for UAV and IRS.

LITERATURE REVIEW

The previous work primarily focuses on security of data by using the baseline and AO schemes. Although, up until now, there have been limited studies on incorporating IRS into UAV communication, specifically in terms of physical layer security. The latest research on the IRS has mostly concentrated on optimizing power and spectral efficiency, as well as improving channel estimation. Additionally, these studies have also examined the study of capacity and data rates. The area of deep learning-based design for communications with IRS and analysis of the reliability of IRS-aided communication has also been explored. Regarding the studies on UAV-enabled communications, there are two primary areas of focus. One involves the use of UAVs to assist with communications, while the other involves utilising cellular networks for UAV communications.

In in this work (Tariq et al., 2023), the authors aim at discussing an IRS-assisted UAV network, where a UAV with an IRS acts as a passive relay. The primary challenge involved maximising the system's secrecy rate by jointly optimising the location of the UAV and the phase shifts of the IRS. An iterative algorithm is proposed to address this issue, where the location of the UAV is optimised using fixed phase shifts, and the phase shifts are optimised based on the updated UAV location. In this paper (Naeem et al., 2022), the authors have explored the use of an UAV with an air-to-ground friendly jammer to enhance the security of communications between a legitimate transmitter-receiver pair, even in the presence of an unknown eavesdropper. Authors analyse the effect of UAV jamming power and its spatial deployment on the reliability and security of the system. They investigate how these factors impact the outage probability of the legitimate receiver and the intercept probability of the eavesdropper. They have developed a new security measure called the intercept probability security region (IPSR) based on the IP. The IPSR defines the specific region within a target area where the IP falls below a certain threshold. Afterwards, a low-complexity iterative algorithm was developed to maximise the IPSR. This was achieved by optimising the 3-D deployment and jamming power of the UAV jammer in a joint manner. Considering the rapid advancements in technology, it is crucial to reconsider the design of wireless networks to effectively incorporate emerging technologies like federated learning (FL) and software defined networks deployed over the air. IRS is a cutting-edge technology that has shown immense promise in improving communication, particularly in difficult environments. By enhancing the control of the channel between the transmitter and receiver, IRS has the ability to revolutionise the way we communicate. For this research (Dong et al., 2024), the authors are examining the implementation of IRS using unmanned aerial vehicles to improve wireless communications in battlefield situations.

They also explore the security considerations in this type of deployment by utilising a combination of DRL and defensive deception techniques. Utilising data-driven power allocation in communication channels with the help of RL, we can effectively obscure the attack surface, entice jammers to specific channels, and ultimately reduce the impact of denial-of-service attacks. UAVs, when used in conjunction with IRSs, are leading the way in technology to improve the capacity of wireless communications channels. This is accomplished by utilizing the three-dimensional mobility of UAVs in conjunction with the intelligent radio capabilities of IRSs. In this article (Wu et al., 2024), the authors studied the design of secure transmission for a UAV network with the assistance of an IRS, while considering the presence of an eavesdropper. Optimising the trajectory of the UAV, phase and the amplitude with the IRS matrix is vital to maximize the secrecy rate. In order to tackle this complex issue, the authors break it down upto 3 smaller problems, then utilised an iteration-based algorithm to solve them in an alternating manner.

Initially, a closed-form solution is obtained for the active beamforming. By employing optimal transmit beamforming, the issue of optimizing passive beamforming in fractional programming is transformed into associated parametric sub-problems. Furthermore, the method of successive convex approximation is employed to tackle the non-convex issue of optimizing the trajectory of the UAV. This requires transforming the problem into a convex form, which establishes a minimum value for the original problem.

This study (Yang et al., 2020) presents friendly jammer-based approach where jammer-UAV assist the transmit-UAV in defending from GEs. More precisely, the UAV-based transmitter transmits secret information to the Users, while the jammer-UAV use 3D beamforming to send AN signal to the GEs. The authors proposed a DRL-based MADRL approach, specifically the multi-agent deep deterministic policy gradient (MADDPG) algorithm. This research's main objective is to enhance the overall secrecy rate by solving the convex optimization problem of UAV's trajectory and IRS variables. To obtain this objective MADDPG algorithm utilises a centralised training strategy and distributed-execution. In order to enhance learning efficiency and convergence, a novel approach called continuous action attention MADDPG (CAA-MADDPG) has been proposed.

This paper (Song et al., 2022) examines a communication system that IRS and UAVs. The UAV is deployed to provide service to user equipment (UE), in the presence of many walls mounted IRSs. The authors strive to enhance the energy efficiency of the system by optimising the trajectory of the UAV and the phase shifts of reflecting elements of IRS. This optimisation takes into account the movement of the UE and the selection of IRSs, all with the goal of reducing energy consumption and maximising the data rate of the UE. Given the complexity of the system and the ever-changing environment, it can be quite difficult to come up with simple algorithms using traditional optimisation methods. To tackle this problem, they initially suggest an algorithm based on deep Q-network (DQN) that discretizes the trajectory, offering the benefit of reduced training time. In addition, they suggest an algorithm based on DDPG to address the situation where there is a continuous trajectory, in order to improve performance.

Over the past few years, there has been a significant amount of research conducted on IRS-assisted communication in (Omar et al., 2023). These studies mainly cover the advantages, applications, hardware architecture, and signal model of IRS. The design of the transmitter's beamforming and the passive beamforming of the IRS are done together. The work (Xu et al., 2022) involves the development of an IRS assisted anti-jamming UAV-communication system. In this system, a UAV collects data from a GU even in the presence of multiple jammers with imperfect CSI. To enhance the desired signal and suppress the interference, an IRS is deployed. The IRS achieves this by adjusting the phase shifts of its reflecting elements. In (Lin et al., 2023) the authors solved the formulated problem of IRS-based UAV communication. These works didn't include either the jammer in UAV-based secure communication or they ignored they tried to achieve secrecy rate and UAV trajectory optimization by traditional optimization methods. We proposed the complex relationship of the UAV, an IRS mounted on it, a BS and the GU in the presence of an eavesdropper along with a friendly jammer. We employed the full potential of DRL based dual-DDPG (D3PG) to solve the trajectory and IRS backscatter optimization problem of UAV-based IRS. previous studies focused on situations where UAVs were used exclusively as mobile BS/jammers, making it difficult to achieve a high-level physical layer security. We employed the mobility and adaptability of UAV to enhance the use-case of IRS. The main contribution of our work are as follows:

The Motivated by previous research and the challenges it presents, we are examining the security of a UAV network's physical layer. In this scenario, a base station transmits sensitive information to a ground receiver with the assistance of a UAV-IRS, all while an eavesdropper attempts to intercept the confidential data. There is a jammer in the vicinity of the Eves that is causing confusion and interference, and efforts are being made to counteract the jammer and ensure that the confidential information can reach its intended destination. We collaboratively enhance the physical layer security by optimising the trajectories and transmit power of the UAV, taking into account mobility and power constraints. The resulting optimisation problem for the trajectory optimization of UAV, and the IRS backscatter coefficient matrix optimisation is non-convex can't be solved with traditional schemes. Thus, we employed the DRL-based algorithm to solve the said non-convex optimisation problems. It is important to note that, here two different and complex types of convex problems are encountered, therefore we have to use dual neural networks or we have to apply the DRL based algorithm twice, to achieve the desired optimization for both UAV and the IRS.

SYSTEM MODEL

Consider an IRS-assisted secure communication system including a ground BS, in the air IRS-assisted UAV, a GU, one malicious eavesdropper, and one friendly jammer as shown in Figure. The IRS has Lr number of x-y oriented elements, while GU and eve are equipped with a single antenna. The received signal at IRS can be expressed as:

$y_L(t) = \ G\,x(t) + n_L$, **(1)**

x(t) implies for the transmit vector for BS n_L ϵ $C^{N_L x 1}$, stands for the Additive White Gaussian Noise(AWGN) with zero mean, , G ϵ $C^{N_b x\ N_L}$, represents channel from BS to UAV-IRs , N_b shows the number of BS antennas and N_L represents the reflecting elements of IRS. For this channel model we have the G for the BS to IRS , IRS-GU and IRS-Eve channels h_{Lu} and h_{Le} respectively. And the phase shift matrix ɣ is represented as:

ɣ=*diag*$(\alpha_1 e^{j\theta_1},\ \alpha_2 e^{j\theta_2}, \alpha_3 e^{j\theta_3} \ldots \alpha_L e^{j\theta_L})^{H}$,**(2)**

j = $\sqrt{-1}$, θ ∈ [0, 2π] , a ∈ [0,1] and θ or a represents the amplitude and phase shifts of the IRS matrix.

CHANNEL MODEL

Suppose a system that uses a three-dimensional Cartesian coordinate. In this system, the ground base station is located at the origin (0,0,0). The UAV's position is represented by (x, y, H), while the GU, jammer and Eve are pointed at (xu, 0,0), (x_j, y_j,0) , (x_e, y_e, 0) respectively. The distances from the BS to the UAV, the UAV to the GU, the UAV to the eavesdropper, jammer to the GU and jammer to the eavesdropper are represented by d_{bL} , d_{Lu}, d_{Le}, d_{ju} and d_{je} respectively. The connection between the UAV and the GU or eavesdropper can be described as an A-2G channel, which is influenced by the propagation characteristics of the surrounding environment, as well as the altitude and elevation angle of the UAV.

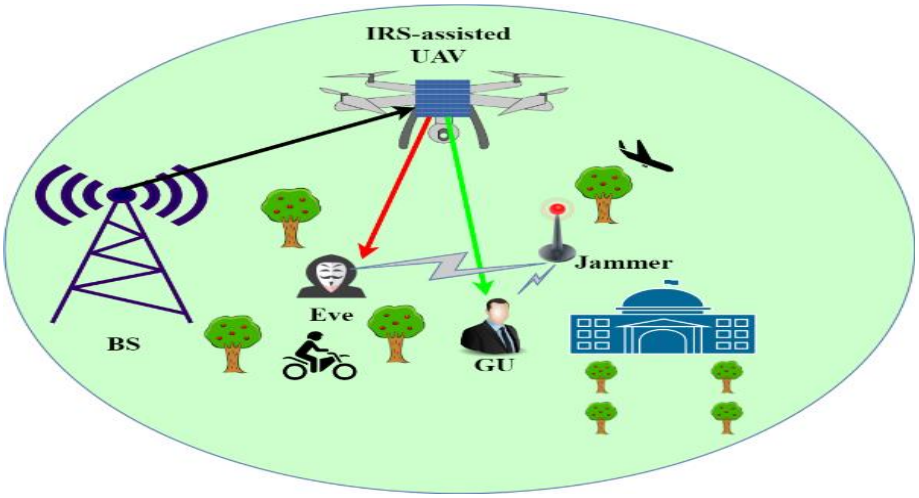

Figure 1.
System model and channels

In the absence of obstacles, the A2G channel encounters free space path loss. However, in urban environments, the presence of buildings or trees might cause extra path loss in addition to the current free space path loss. Thus, the A2G channel model can be depicted as

$$PL_{\zeta} = P_{FS}(t) + \chi_{\zeta}, \qquad \textbf{(3)}$$

Where $P_{FS}(t)$, is the notation used to describe the free space path loss and is is further described as:

$$P_{FS}(t) = 20 \log(d(t)) + 20 \log(f) + 20 \log(4\pi/c), \qquad \textbf{(4)}$$

In the above equation *d(t)* is the instantaneous distance between the UAV-IRS and GU, represented in 2-D cartesian plane as $d(t) = \sqrt{(\mathrm{x(t)} - x_0)2 + (\mathrm{y(t)} - y_0)2 + \mathrm{H(t)}2}$.

In the above equations the 'f' and 'c' represents the carrier frequency and the light speed, χ_{ζ} is the average path-loss affected by LoS and NLoS factors and ζ expressed as LoS and NLoS ζ ϵ {LoS, NLoS}. With the free space path-loss, the channel gain between IRS-UAV to GU, IRS-UAV to BS, IRS-UAV to Eve for the LoS are respectively calculated as:

$$\mathbf{h}_{\mathbf{Lu}}^{\mathbf{H}^2} = d_{Lu}{}^{-2}\Omega_0 = \frac{\Omega_0}{(x - x_u)^2 + y^2 + H^2} \qquad (\mathbf{5})$$

$$\mathbf{G}^2 = d_{bL}{}^{-2}\Omega_0 = \frac{\Omega_0}{x^2 + y^2 + H^2} \qquad (\mathbf{6})$$

$$\mathbf{h}_{\mathbf{Le}}^{\mathbf{H}^2} = d_{Le}{}^{-2}\Omega_0 = \frac{\Omega_0}{(x - x_e)^2 + (y - y_e)^2 + H^2} \qquad (\mathbf{7})$$

$$\mathbf{h}_{\mathbf{ju}} = d_{ju}{}^{-2}\Omega_0 = \frac{\Omega_0}{\left(x_j - x_u\right)^2 + y^2} \qquad (\mathbf{8})$$

$$\mathbf{h}_{\mathbf{je}} = d_{je}{}^{-2}\Omega_0 = \frac{\Omega_0}{\left(x_j - x_e\right)^2 + \left(y_j - y_e\right)^2} \qquad (\mathbf{9})$$

Ω_0 represents the power of the channel at the reference distance of d_0 = 1 m. LoS and NLoS may encounter through different set of probability, therefore the probability LoS link can be expressed as:

$$P_{LoS}(t) = \frac{1}{1 + a\exp\left[-b(\sin^{-1}(H(t)/d(t) - a)\right]} \textbf{(10)}$$

Where, a and b are environment related variables. For improving the accuracy of jamming with detection avoidance by receivers, we employ beamforming to send jamming signals via UAV jammers.

Beamforming is a powerful technique that allows for precise signal transmission to a specific direction. Precise targeting of signals can be achieved by manipulating the elevation and angle of azimuth of the beam. Therefore, the jamming signals can be accurately directed towards the GEs, without causing any interference to the GU. The received data rate at time slot t can be expressed as

$$R_{GU}(t) = \log_2\left(1 + \frac{P_T}{PL_{UV,}\sigma_{GU}^2}\right), \qquad \textbf{(11)}$$

Here P_T is the maximum power transmitted from the BS, PL_{UV} is the path loss from the UAV, to the GU and σ_{GU}^2 is the noise power for the GU. It is important to note there is no signal effect at the GU, because the signal reaching at the GU will be cancelled out by GU, because we have friendly jammer that will only affect or disturbs the eavesdropper in the scenario. Similarly, the signal reaching at the Eavesdropper expressed as:

$$R_e = \log_2 \left[1 + \frac{P_T}{\overline{PL}_{Ve}\left(\frac{P_J}{PL_e}|\mathbf{a}(N_h,N_v,\theta,\phi)|^2+\sigma_e^2\right)}\right] \textbf{(12)}$$

In the eq (12) $|\mathbf{a}(N_h, N_v, \theta, \phi)|$ is the steering vector function of UAV given in (Kaur et al., 2024) and σ_e^2 is the noise power of the eavesdropper.

Problem Formulation

When it comes to find the secrecy rate, we follow the Shannon's formula of the secrecy rate which is represented as:

$$R_s((\mathrm{x},\mathrm{y}),\gamma) = \log\left[1 + \frac{P\left|(h_L^H \gamma \mathbf{G})\right|^2}{\mu^2}\right] - \log\left[1 + \frac{P|(g_r^H \gamma\, \mathbf{M})|^2}{\mu^2}\right] \qquad (\mathbf{13})$$

Where μ^2 is represents the noise power, now

Let we assume that $\mathbf{v}$ = $(\alpha_1 e^{j\theta_1},\ \alpha_2 e^{j\theta_2}, \alpha_3 e^{j\theta_3}\ldots\ldots\ldots\ldots\alpha_L e^{j\theta_L})^H$, then $\gamma = \mathrm{diag}\,(\mathbf{v}^*)$. With these approximations, hence the problem can be formulated as:

$$\begin{aligned}
&\max_{\gamma,(x,y)} R_s,\\
&\text{s.t. } \gamma = \mathrm{diag}(\alpha_1 e^{j\theta_1},\ \alpha_2 e^{j\theta_2},\ \alpha_3 e^{j\theta_3}\ldots\alpha_L e^{j\theta_L})^H,\\
&where, \theta_L \in [0,2\pi],\quad \alpha_L \in [0,1]\\
&|\mathbf{v}(i)| = 1, i = 1, \ldots, \mathrm{L} \qquad (\mathbf{14}),
\end{aligned}$$

DRL-BASED SOLUTION

With the goal to address problem (14), we present a Dual DDPG (D3PG) algorithm that enables the agent to acquire knowledge of the beamforming and trajectory policies without any prior understanding of the system. Given the complex relationship between the UAV trajectory P and the highly dynamic CSI, it becomes challenging to handle such complicated problem. Therefore, in order to address this problem, Instead of utilizing a single agent in the standard network based on DRL, two DDPG networks are formed to segregate these factors. Specifically, the initial network utilizes the CSI, or H_S, in as a state to calculate G and ɣ. When the Agent starts interacting with the environment the coordinates of location and direction of UAV,GU including eavesdropper, are used as the state to determine the movement of the UAV. This movement includes the flight in x-axis ϕ[n], UAV's direction Ψ[n] at each time slot. In the D3PG scheme both of the neural networks have same reward function and optimizer.

ACTIVE AND PASSIVE BEAMFORMING OPTIMIZATION

Fostering on previous research, a pioneering DDPG network is utilized to acquire the most effective strategy for the reflecting beamforming matrix ɣ for the IRS through active interaction on surrounding system model. The working parameters of proposed DRL model are described as follows:

- ***State*** $s_{n,1}$***:*** It represents the agent's initial interaction with taking input as CSI and other UAV parameters in the n-th time slot. In addition, h doesn't have the value in small scale component beforehand by the UAV. The collection of small-scale information could be in real-time to monitor the network's status. This algorithm can adapt to changes in the environment as they occur.

- ***Action*** $a_{n,1}$ ***:*** First, IRS's passive beamforming matrix ɣ and the active beamforming matrix 'G' as part of the action. It's important to mention that G and ɣ split into M = Re{G} + Im{G} , or ɣ = Re{ɣ} + Im{ɣ} in order to address the issue with real input.

- ***Reward*** $r_{n,1}$: The agent interacts with the environment and then receives some reward as action on the basis of interaction. The reward function in this case expressed as: $r_{n,1}$ = tanh(R^{SEC} – K_1Þm – K_2Þs – K_3Þg), where Þm , Þs , Þg are the penalties for the environment. If the required conditions haven't met in the following constraints, then the by policy it will give it a penalty. In the reward function – K_1, K_2, K_3 are the are the weights variables used to keep balance between the penalty and sum rate. This can be approximated as 1 - Pr{R_{sec} ≥ R_{sec}^{min},} ≈ N_{outage}/N_{sample}, where N_{outage} represents the total samples where the R_{sec} is less than required R_{sec},, and N_{sample} represents the total number of generated samples during the process.

UAV-TRAJECTORY OPTIMIZATION

The second DDPG agent is employed in the neural network to acquire the UAV's motion ϕ[n], and direction Ψ[n] for G or ɣ. The reward taking function from the state is described as:

- The state $s_{n,2}$: DDPG algorithm takes continuous state space, which consist of big data collection. Therefore, to find best convergence rate and optimal policy we only consider the UAV location into the neural network.

- The Action $a_{n,2}$: Agent's action comes with the flight distance horizontally ϕ[n], and the direction ψ[n]. After that, The UAV's trajectory can be optimized by utilizing D3PG. The trajectory of the UAV at the n-th time slot can be represented as: $q[n+1]-q[n]= \phi[n](\cos\psi[n]e_x + \sin\psi[n]e_y)$, where e_x, e_y are the unit vector on the X-axis and the Y-axis.

- The Reward $r_{n,2}$: Both of neural networks are set to work on same reward policy, that's why reward function of DDPG's UAV trajectory optimization will also be same.

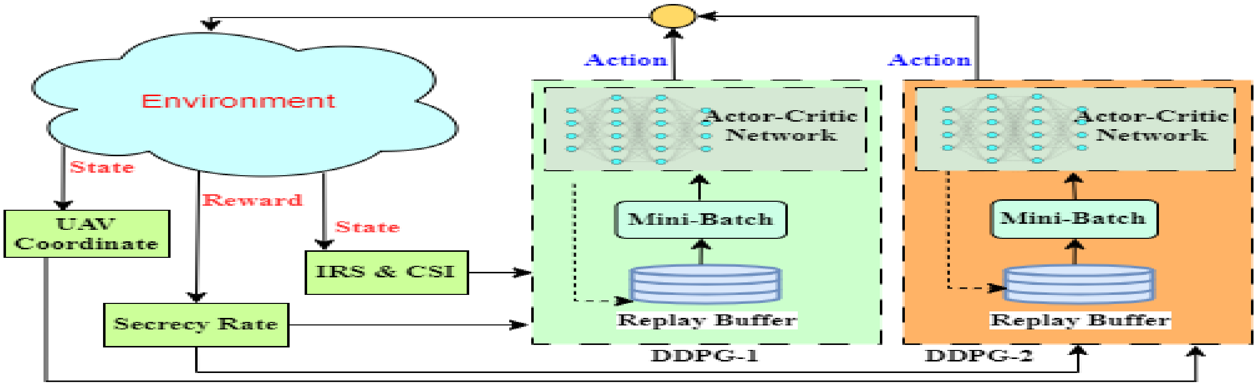

Figure 2.
DRL based Model for UAV based secrecy rate maximization

When the training process nears convergence, the initial network determines the optimal method for both active and passive beamforming, while the second network computes the ideal trajectory. The utilization of a mutually agreed upon reward function and the exchange of information regarding the surroundings augment the synchronization between these two networks, allowing them to learn and develop a beneficial strategy. Consequently, the beamforming matrix (G, ɣ), and the UAV trajectory Q are acquired using the proposed D3PG technique. The D3PG system we propose operates in real-time and is capable of capturing the instantaneous Channel State Information (CSI) at every time slot, including the rapidly changing elements. In contrast, the offline mode involves preloading beamforming policy and UAV's trajectory beforehand and remains unaffected by alterations in the dynamic environment.

Computational Complexity Analysis of proposed scheme

This section specifically addresses the computational difficulty in the proposed D3PG algorithm. More precisely, let L represent the layer numbers of the DNN utilized in agent's networks, and let n_i represent the number of neurons in the i-th layer. Concerning the training mode, the computational cost for a single DNN to assess and update in a solitary step can be mathematically represented as ©(N_b($\sum_{i=1}^{L-1} n_{i,}\ \ n_i+1$)) , while N_b represents the mini-batch size used in the DDPG. The D3PG algorithm consists of a finite number of DNNs and requires N_{ep} * N_{step} steps to complete training. As a result, the overall computational complexity of the D3PG has been reduced, which ultimately effects on required secrecy rate. The Overall reduced computational complexity of algorithm for training is ©(N_{ep} N_{step} N_b($\sum_{i=1}^{L-1} n_{i,}\ \ n_i+1$)). When working online, we can greatly decrease the computational complexity at each step by terminating the training process whenever the network's performance reaches a stable state. This aids in maintaining the computational complexity at a desirable level and results in a satisfactory convergence of the method.

RESULTS AND DISCUSSIONS

Here, we present numerical results to assess the efficiency of the offered D3PG algorithm. Our initial DDPG network, consists of two fully-connected hidden layers are utilized, with [128,256] neurons in both the actor and critic networks. We've used Adam optimizer for training the actor network, employing a learning rate of 0.0001, while the critic network is trained with a learning rate of 0.001. The 2nd network also has a similar structure to the first network, but with the number of layers as [128,256]. The D3PG model undergoes training over 1400 episodes, with each episode consisting of 300 time slots. The starting positions of the UAV and the fixed RIS are defined as (0m, 20m, 40m) and (0m, 40m, 15m) respectively. The eavesdropper is located at coordinates (42m, −7m, 0m). We've used the tanh activation function at the output and Relu activation function at the output and input layers respectively. Similarly, the power transmitted for the BS and Jammer varies between the 15 to 40 and 18 to 36 dBm.

The fig 3 illustrates the relationship between the secrecy rate (bps/Hz) and the transmit power (dBm) in a communication channel. The guaranteed secrecy rate refers to a transmission rate that ensures the confidentiality of information by effectively protecting it from unauthorized eavesdroppers. The x-axis Transmit Power (dBm) - In this context, the transmitter emits signals with varying power levels, higher dBm values guarantees a stronger signal. The Y-axis represents the Secrecy Rate, which measures the rate at which secret data can be transmitted over the channel while maintaining data confidentiality in a secure manner.

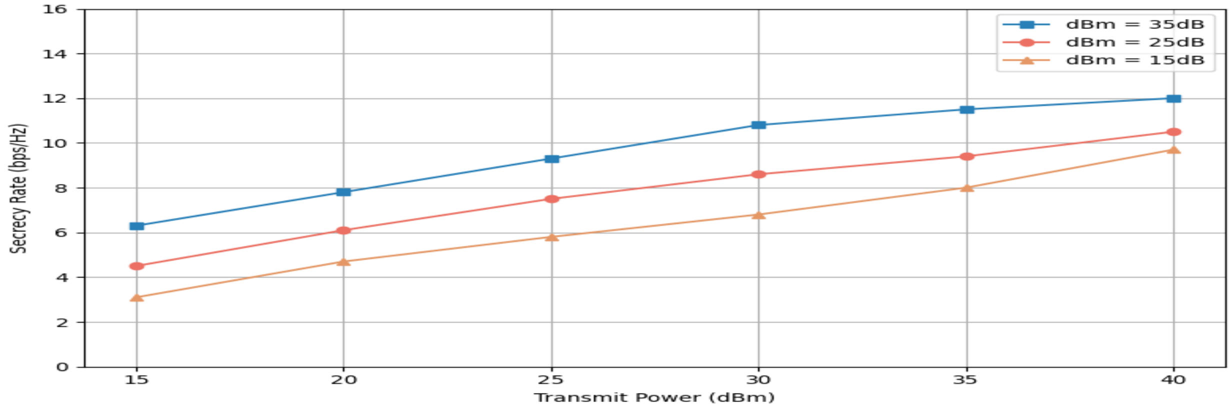

Figure 3.
Transmit power P_t Vs Secrecy rate

Reduced power lessens the strength of the signal and signal is prone to being disguised by the surrounding noise and can also be captured by any eavesdroppers, hence lowering the secrecy level. Increasing the transmit power results in a faster successful transmission of a strong signal. A strong signal will mitigate the impact of noise and the Instances of communication errors may not occur as frequently in comparison to transmission faults. The enhanced quality of the signal raises the likelihood of preventing others from intercepting and deciphering it.

The figure 4 illustrates three lines representing the results of three algorithms (DQN, SCA, and Proposed DDPG) utilized for controlling the IRS elements. Among all the configurations of reflective elements examined, the D3PG exhibits the highest secrecy rate. This suggests that the DDPG technique possesses sufficient capacity to effectively utilize IRS aspects in order to achieve a high level of secrecy rate. In the DQN line, the secrecy rate is lower than that of the DDPG line, but higher than that of the SCA line when the number of IRS elements is changed. The statement suggests that the DQN algorithm incorporates IRS features to some extent in order to achieve a level of secrecy rate.

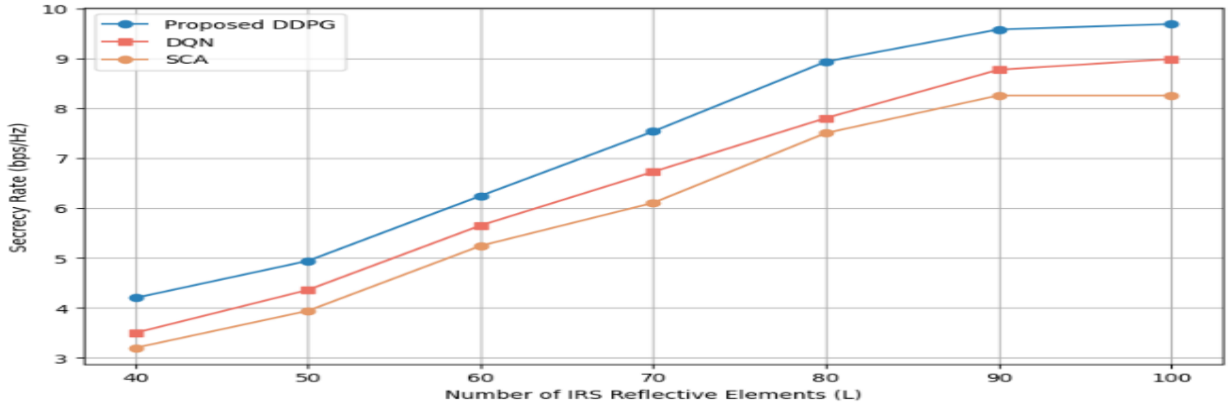

Figure 4.
IRS elements Vs Secrecy rate

However, the secrecy rates achieved by DQN are not as impressive as those produced by the DDPG algorithm. The successive convex approximation (SCA) graph represents the maximum level of confidentiality allowed for the IRS parameters. Consequently, the SCA algorithm is not preferred when it comes to using an IRS to achieve a higher level of secrecy rate. The secrecy rate of the three algorithms increases as the number of reflecting parts of the IRS increases. Therefore, it can be inferred that incorporating additional IRS components will improve the level of security in communication channels. By employing narrower beam steering, the IRS has the potential to accurately guide the signal towards the intended user rather than unintended listeners. Simultaneously, this can enhance the signal-to-noise ratio at the receiver, so impeding an adversary's capacity to intercept the communication.

Increasing the number of IRS element results in a more advanced manipulation of radio waves. In our case by effectively optimizing the G and ɣ we can effectively use the IRS to increase the secrecy rate and the Data rate of the GU. By selectively focusing on specific signals or manipulating their diffraction, the IRS can create more intricate interference, thus obstructing the capacity of eavesdroppers to reassemble the data. The graph indicates that DDPG performs better than DQN, most likely because it is specifically designed for continuous control (such as UAV movement) and has less overestimation. Optimizing the trajectory of the UAV is of utmost importance. Moreover, employing two distinct DDPG agents, one dedicated to trajectory and the other to secrecy (through IRS manipulation), enables them to concentrate on their respective objectives without any disruption. The specialization and potential synergy among the agents are likely factors that contribute to the overall improved performance reported in the dual DDPG strategy.

Figure 5 illustrates the performance of the DDPG algorithm when applied to the friendly jammer system, which effectively prevents eavesdroppers from accessing confidential data. The figure demonstrates that the DDPG algorithm effectively controls jammers to disrupt the eavesdropper without compromising the conversation. DDPG has the capacity to identify the eavesdropper's location, their level of capability, and their vulnerability to jamming. It can then generate jamming signals based on this information.

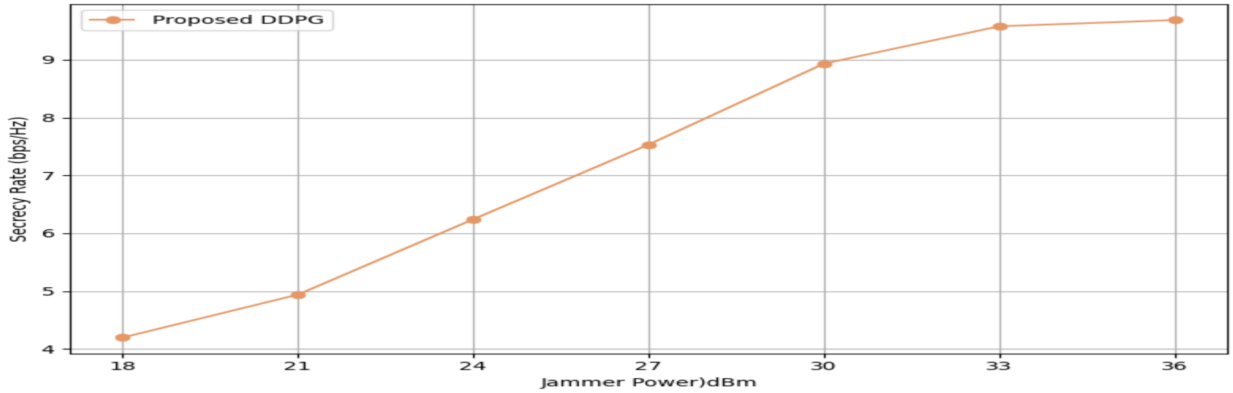

Figure 5.
Jamming power Vs secrecy rate

DDPG accurately calculates the jamming power based on environmental parameters. Adjusting the level of jamming can either enhance the obstruction area surrounding authorized communication or reduce interference on your own channel. The algorithm selects a jammer that employs techniques such as frequency hopping and noise injection to prevent the eavesdropper from accurately determining the true signal. The DDPG algorithm demonstrates its advantage in protecting communication channels from friendly jammers by maintaining a stable or increasing secrecy rate as the jamming force grows. This example demonstrates the proficiency of reinforcement learning in producing intelligent jamming tactics that confuse eavesdroppers.

This figure 6, illustrates the number of training episodes, where each episode represents a complete encounter cycle with the environment. During each episode, rewards are obtained and the action policy is updated based on these rewards. The Y-axis represents the average reward obtained by the DDPG algorithm in each episode. Indeed, in this context, the term "reward" refers to the extent of information revealed throughout the episode. An important indicator is the algorithm's ability to accurately determine the optimal rate of achieving the secret as the number of episodes increases.

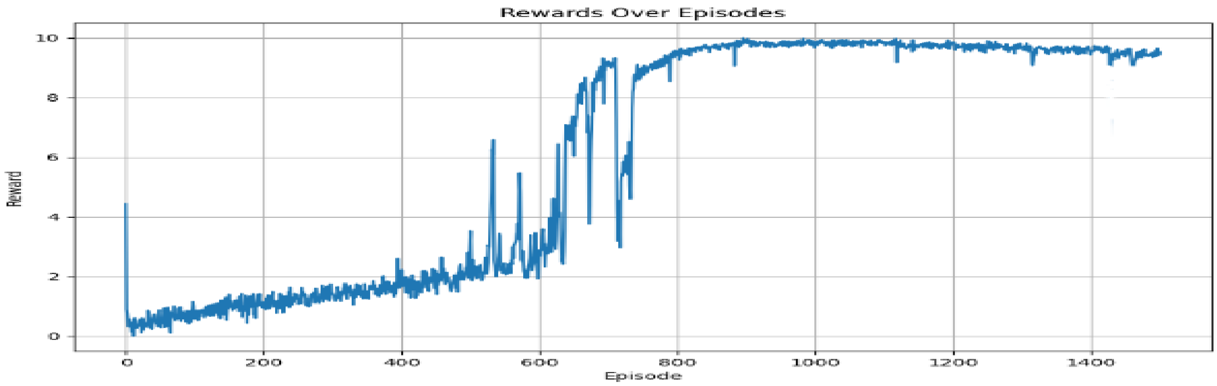

Figure 6.
No of Episodes Vs Reward

The figure represents curve that shows the relation to the number of episodes (X-axis) vs the reward (Y-axis). Based on this figure, we can assert that the DDPG algorithm is approaching convergence. Convergence, in this context, refers to the algorithm's gradual improvement in performance (accuracy grade) over time. As the number of episodes increases, the average reward consistently remains close to a certain value, indicating that the learning process effectively generates a practical method for achieving a high secrecy rate.

DECLARATIONS

Acknowledgement: We appreciate the generous support from all the supervisors and their different affiliations.

Funding: No funding body in the public, private, or nonprofit sectors provided a particular grant for this research.

Availability of data and material: In the approach, the data sources for the variables are stated.

Authors' contributions: Each author participated equally to the creation of this work.

Conflicts of Interests: The authors declare no conflict of interest.

Consent to Participate: Yes

Consent for publication and Ethical approval: Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent**.**

REFERENCES

Dong, R., Wang, B., Cao, K., Tian, J., & Cheng, T. (2024). Secure Transmission Design of RIS Enabled UAV Communication Networks Exploiting Deep Reinforcement Learning. *IEEE Transactions on Vehicular Technology*.

Guo, X., Chen, Y., & Wang, Y. (2021). Learning-based robust and secure transmission for reconfigurable intelligent surface aided millimeter wave UAV communications. *IEEE Wireless Communications Letters*, *10*(8), 1795-1799.

Iacovelli, G., Coluccia, A., & Grieco, L. A. (2023). Multi-UAV IRS-assisted Communications: Multi-Node Channel Modeling and Fair Sum-Rate Optimization via Deep Reinforcement Learning. *IEEE Internet of Things Journal*.

Kaur, R., Bansal, B., Majhi, S., Jain, S., Huang, C., & Yuen, C. (2024). A survey on reconfigurable intelligent surface for physical layer security of next-generation wireless communications. *IEEE Open Journal of Vehicular Technology*.

Lin, R., Qiu, H., Jiang, W., Jiang, Z., Li, Z., & Wang, J. (2023). Deep Reinforcement Learning for Physical Layer Security Enhancement in Energy Harvesting Based Cognitive Radio Networks. *Sensors*, *23*(2), 807.

Lu, X., Xiao, L., Li, P., Ji, X., Xu, C., Yu, S., & Zhuang, W. (2022). Reinforcement learning-based physical cross-layer security and privacy in 6G. *IEEE Communications Surveys & Tutorials*, *25*(1), 425-466.

Naeem, F., Kaddoum, G., Khan, S., Khan, K. S., & Adam, N. (2022). IRS-empowered 6G networks: deployment strategies, performance optimization, and future research directions. *IEEE Access*, *10*, 118676-118696.

Omar, S. S., Abd El-Haleem, A. M., Ibrahim, I. I., & Saleh, A. M. (2023). Capacity Enhancement of Flying-IRS Assisted 6G THz Network using Deep Reinforcement Learning. *IEEE Access*.

Samir, M., Elhattab, M., Assi, C., Sharafeddine, S., & Ghrayeb, A. (2021). Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach. *IEEE Transactions on Vehicular Technology*, *70*(4), 3978-3983.

Song, W., Rajak, S., Dang, S., Liu, R., Li, J., & Chinnadurai, S. (2022). Deep learning enabled IRS for 6G intelligent transportation systems: A comprehensive study. *IEEE Transactions on Intelligent Transportation Systems*.

Tariq, Z. U. A., Baccour, E., Erbad, A., & Hamdi, M. (2023). Reinforcement Learning for Resilient Aerial-IRS Assisted Wireless Communications Networks in the Presence of Multiple Jammers. *IEEE Open Journal of the Communications Society*.

Wang, L., Wang, K., Pan, C., & Aslam, N. (2022). Joint trajectory and passive beamforming design for intelligent reflecting surface-aided UAV communications: A deep reinforcement learning approach. *IEEE Transactions on Mobile Computing*.

Wu, M., Guo, K., Li, X., Nauman, A., An, K., & Wang, J. (2024). Optimization Design in RIS-Assisted Integrated Satellite-UAV-Served 6G IoT: A Deep Reinforcement Learning Approach. *IEEE Internet of Things Magazine*, *7*(1), 12-18.

Xu, J., Kang, X., Zhang, R., Liang, Y. C., & Sun, S. (2022). Optimization for master-UAV-powered auxiliary-aerial-IRS-assisted IoT networks: An option-based multi-agent hierarchical deep reinforcement learning approach. *IEEE Internet of Things Journal*, 9(22), 22887-22902.

Yang, H., Xiong, Z., Zhao, J., Niyato, D., Wu, Q., Poor, H. V., & Tornatore, M. (2020). Intelligent reflecting surface assisted anti-jamming communications: A fast reinforcement learning approach. *IEEE transactions on wireless communications*, *20*(3), 1963-1974.

Yang, H., Xiong, Z., Zhao, J., Niyato, D., Xiao, L., & Wu, Q. (2020). Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications. *IEEE Transactions on Wireless Communications*, *20*(1), 375-388.

Zhang, L., Lai, S., Xia, J., Gao, C., Fan, D., & Ou, J. (2022). Deep reinforcement learning based IRS-assisted mobile edge computing under physical-layer security. *Physical Communication*, *55*, 101896.

Zhang, Yingzheng, Jufang Li, Guangchen Mu, and Xiaoyu Chen. "Deep reinforcement learning enabled UAV-IRS-assisted secure mobile edge computing network." *Physical Communication* 61 (2023): 102173.