**Enhancing NOMA Network Security with RIS-UAV Integration: Exploring PPO Technique**

Syed Zain Ul Abideen1\*,Abdul Wahid1, Mian Muhammad Kamal2,

Touseef Hussain1

1. College of Computer Science and Technology, Qingdao University, Qingdao, Shandong, China. (e-mail: zain208shah@gmail.com, wahidjan999@gmail.com, touseefhussain098@gmail.com ).
2. Joint International Research Laboratory of Information Display and Visualization, Southeast University, Nanjing, People’s Republic of China. (e-mail: mmkamal@seu.edu.cn)

\***CorrespondingAuthor**: Syed Zain Ul Abideen (email: zain208shah@gmail.com)

***Abstract***—In order to improve the security aspects of a nonorthogonal multiple access (NOMA) downlink network, we investigate the use of the proximal policy optimization (PPO) technique in this study. This network is uniquely augmented with reconfi- gurable intelligent surfaces (RIS) mounted on unmanned aerial vehicles (UAVs). The main objective of this article is to prevent possible eavesdroppers from accessing the network while still assuring continuous and secure connectivity for legitimate users. The network configuration includes a UAV outfitted with a RIS, which plays an important role in signal propagation. This work aims to increase the secrecy rates of all user communications in eavesdropping prone locations. By adaptively modifying the phase shifts and power distributions through the RIS the proposed PPO algorithm not only shows adaptability and effectiveness in securing wireless communications, but also highlights the significant advances made in communication system security through this technology.

***Index Terms*—Reconfigurable intelligent surface, Secrecy, UAV, DRL**

**Introduction:**

Next-level wireless communication systems, particularly those designed for the sixth generation (6G) networks, are expected to set new standards in connectivity characterized by high reliability, extremely fast data rates, and minimal latency. RIS and UAV assisted communication systems are identified as crucial components contributing to this significant advancement [1]. In such a scenario, NOMA is highlighted as an impactful access method intended to enhance spectral efficiency and cater to the increasing demand for wireless connectivity [2].

RIS have emerged as a strategic approach to alter the propagation environment. Through the use of programmable meta surfaces containing numerous passive reflecting elements RIS can manipulate the phase and amplitude of incoming signals in order to tailor propagation paths according to specific communication needs [3]. The potential of incorporating this technology into wireless networks supported by unmanned aerial vehicles (UAVs) is significant as it offers a flexible and dynamic

method for achieving extensive network coverage [4]. Unmanned Aerial Vehicles (UAVs) enhance the communications sector's versatility by means of their capacity to be adjusted and placed carefully in the atmosphere enabling rapid and precise deployments. This approach demonstrates outstanding effectiveness in increasing network coverage, capacity, and overcoming service deficiencies in physical networks [5]. When utilized in conjunction with RIS and with NOMA protocol [6] support, UAVs substantially enhance the efficacy of resource allocation and the network-wide experience for users [7]. The development of RIS as an important technology signifies a substantial progression in PLS establishing a wireless communications environment that is both dynamic and manageable [8]. Signal phases are managed by RIS via complicated control systems. It consists of a variety of inexpensive passive components with integrated PIN diodes that alter the electromagnetic field. By means of modulation, the intelligibility of the signal is improved for authorized users, while the signals intended for unauthorized eavesdropping are interrupted. RIS differentiates itself from conventional PLS methods that rely on artificial noise [9] or complicated multi-antenna beamforming [10] by means of its inactive operational features, thereby avoiding the significant costs associated with RF chains.

Moreover, the flexible nature of RIS facilitates its straightforward incorporation into prevailing network systems and makes it ideal for attachment to a variety of structures within urban landscapes, as well as to wearable technology [11].

Nevertheless, the growing interconnections and complexity of systems also increase the risk of security breaches. The physical layer of wireless communication systems is particularly vulnerable to eavesdropping and other cyber threats, highlighting the need for new protective strategies [12]. In this respect RIS offer a valuable means to enhance the security of the physical layer. Through strategic manipulation of signal reflections [13], RISs can reduce signal reception at potential eavesdropper locations while strengthening it for authorized recipients, establishing a secure communication environment.

The field of industrial automation is experiencing a dynamic evolution due to the rise of innovative AI algorithm frameworks like RL, DL, and particularly DRL. These technologies are streamlining the path toward real-time automation and sustained advancement. DRL is at the forefront, distinguished by its capability to effectively process the intricate flow of data within communication systems and to navigate the complex management of system and resource control, often presented in non-linear and challenging non-convex scenarios. This level of adeptness is achieved even in the computationally rigorous task of deciphering and network formation for understanding wireless channels, conducted without reliance on established channel models or documented patterns of user movements [14]. Moreover, DRL’s strength lies in its strategic identification of optimal solutions for complex optimization issues, a skill honed by analyzing the patterns of rewards obtained from interactions in a wireless context, thus contributing vitally to the advancement of cutting-edge algorithmic designs [11].

## A. Related Work

RIS technologies have quickly become a groundbreak- ing element in wireless networks, finding applications in areas like UAV communications [19], [20], NOMA [15], [16] and CoMP [17], [18]. Their rise is largely due to their significant contributions to network performance improvements.

The development discussed was facilitated by pion- eering studies in [15], which proposed an effective strategy for implementing RIS in NOMA systems managing the trade-off between power efficiency and the sum rate proficiently. This was achieved by carefully applying the SCA technique, which made it possible to improve beamforming and phase shifts on a periodic schedule.

Expanding on this research [16] extended the study to evaluate how RIS affects the performance of NOMA systems, proposing algorithms that are adapted for various RIS setups. The investigation went on [17], who examined the role of RIS in CoMP commun-ications, considering a spectrum of scenarios from ideal to less than ideal, and applying the dual Lagran- gian approach to optimize reflection coefficients. Simultaneously [18], focused on enhancing the long-term energy efficiency in STAR-RIS facilitated CoMP networks by combining methods for active and passive beamforming optimization with a combination of partial programming and DRL approaches to achieve close to optimal results.

The discussion expands much more to include UAV communications, as [19], addressing the challenge of enhancing sum rates within RIS-supported multi-UAV NOMA frameworks. This entailed a holistic optimi- zation strategy that encompassed UAV positionning, power management, RIS reflection matrix configurations, and the sequencing of NOMA deco- ding, all resolved through a BCD iterative methodology. Complementary to this [20], pioneered a novel DRRL algorithm designed to optimize both UAV flight patterns and beamforming processes active at the UAV and passive at the STAR-RIS simultaneously thus showcasing the expansive utility and significant advantages of RIS in advancing the capabilities of contemporary wireless communication systems.

Recent research incorporating RIS has significantly advanced the topic of Physical Layer Security (PLS). Important advancements include the [21], novel RIS configuration that protects downlink NOMA systems from attackers and the [22], method that enhances beamforming in secure wireless systems that use RIS. Robust optimization strategies were proposed by [23], to address the problem of imperfect eavesdropper CSI. By combining STAR-RIS with NOMA in a novel way. [24], improved security by creating artificial noise [25], explored the use of RIS to optimize secure energy efficiency in the context of UAVs and [26], used aerial RIS to adjust signal distributions for NOMA systems in order to improve secrecy rates, highlighting the revolutionary effect of RIS on improving wireless communication security.

## B. Motivation and Contribution

This work is motivated by the critical role that RIS play in the development of 6G wireless communi- cation technology. We are at the beginning of a new age in wireless communications, and with it comes a growing need for improved energy efficiency, cost reduction and spectrum utilization. RIS provides an innovative resolution to these issues, thanks to its capability to reshape electromagnetic wave propa- gation, especially when direct lines of commun- ication are compromised. Furthermore, the increasing worries about security breaches in wireless networks emphasize the critical need to incorporate effective physical layer security (PLS) strategies. RIS techno- logy presents an innovative way to secure data transmission against illegal access and eavesdropping while simultaneously enhancing service quality through innovative non-line-of-sight connections. By making precise adjustments to RIS elements, it is possible to enhance Quality of Service (QoS) and improve security. Additionally, the spatial selectivity of RIS can be utilized to promote signal secrecy for authorized users and limit it for prospective eaves- droppers. The combination of RIS and NOMA together with the potential incorporation of UAV showcases the groundbreaking potential of these technologies in creating a wireless communication environment that is more flexible, efficient and secure. The potential to leverage these state-of-the-art techn- ologies to tackle the challenges of dynamic multiple user environments where traditional optimization techniques fall short because of their nonlinear nature is what drives this research.

The DDPG algorithm is employed in our research to simultaneously optimize power distribution phase shifting and UAV positioning. This enables us to dynamically adapt to changing channel conditions and user requirements. By implementing this compreh- ensive optimization strategy, the system is capable of efficiently allocating wireless resources in real time, therefore guaranteeing peak performance across a wide range of operational scenarios and optimizing spectral efficiency. Moreover the system may constantly improve its approaches responding to changing network dynamics and improving overall utilization of resources in wireless communication environments through using AI-driven learning algorithms.

Our work presents distinctive PLS approaches that attempt to enhance the system’s security measures against potential eavesdropping threats in addition to optimizing system performance. Our solution streng- thens the security and integrity of wireless trans- missions by incorporating PLS mechanisms like secure beamforming and transmission techniques into the optimization framework. This reduces the possibility of unauthorized access and information interception. The practical relevance and significance of our research in protecting wireless communication systems from malicious actors is highlighted by the need of this proactive security approach which is especially important in wireless communication netw- orks handling confidential information or operating in unsafe environments. With the assistance of extensive computer simulations, we intend to establish that the DDPG-based strategy is effective in fortifying the physical layer against vulnerabilities and enhancing the overall performance of the network. To influence the future development of secure and efficient wireless communication networks the objective of this study is to maximize the utilization of RIS, NOMA and UAV technologies.

**SYSTEM MODEL**

The NOMA communication system which is facilitated by UAVs is an advanced framework designed to bolster security at the physical layer. The system comprises a base station (BS) which transmits data to a collection of K end users denoted as K an integer set spanning from 1 to K. A UAV is outfitted with a RIS comprising N passive modulators which aid in the redirection and reflection of the signals. This UAV functions independently at an established alti- tude over Area A and commences its mission from a designated base station. The system is designed for flat fading channels, assuming that there is accurate chan- nel state information available at both the BS and the UAV-RIS. By omitting considerations for energy use and operation time simplifications are made. This ela- borate configuration risks potential covert surveillance by eavesdroppers marked as E = {1,2,...,E}, who atte- mpt to secretly intercept the transmissions. Notably, by employing the DDPG method, the proposed algorithm adjusts to shifting channel conditions across various time slots while maintaining uniformity in each slot ensuring robustness and reliability in fluctuating communication settings. The path that signals take from the base station (BTS) to the interconnect system (RIS) is shown by G ∈ CN²1. The paths that signals take from the RIS to each user (k) and eavesdropper (e) are shown by hr,k ∈ CN×1 and hr,e ∈ CN×1, which show how signal reflections and RIS modulations can change signal strength. The reception at each *k*-th end-user is mathematically modeled as:

where encapsulates the RIS's phase shift capabilities.
 represents the power allocated to the -th user's signal , and symbolizes the additive white Gaussian noise, inherent in wireless communication, at the -th end-user's receiver.
 denotes the BS's power allocation coefficient for the -th user, constrained within the interval [0,1]. The sum of these coefficients across all users equal 1. The transmitted signal for the -th user is represented by , designed such that its expected power equals 1, denoted as . The noise affecting the -th user's signal, denoted as , follows a complex normal distribution with zero mean and variance . The RIS, positioned on a UAV, is located at with a height , while the BS is at the origin (0,0) with height . The horizontal position of each -th user is given by . The distance between the BS and RIS is calculated as , and the distance between the RIS and the -th user is .

Surveillance by the -th eavesdropper yields the intercepted signal as:

For each legitimate user , the channel gain considering the path loss can be expressed as:

where is the path loss coefficient, and and represent the distances from the base station (BS) to the RIS and from the RIS to the user , respectively.

To incorporate Eve into this model, we can extend the scenario to include the channel vector from the RIS to Eve, , and the distance from the RIS to Eve, . The channel gain for Eve would then be similar to that of the legitimate users, adjusted for Eve's position:

In the SINR calculations for implementing Successive Interference Cancellation (SIC) among NOMA users, it's important to account for the potential interception by Eve. The SINR for a user decoding the signal intended for a weaker user is given by:

For Eve attempting to decode the signal intended for user , the SINR would be:



**Fig. 1: UAV-Assisted Reconfigurable intelligent surface NOMA Downlink System**

This formulation allows the system to evaluate the risk posed by Eve's interception attempts and adjust the power allocation and phase shifts accordingly to ensure secure communication. To ensure effective SIC and maintain the system's security, the data rates for the legitimate users must be optimized to maximize the difference between their SINRs and Eve's SINR, effectively increasing the secrecy rate and making the system robust against eavesdropping. To ensure physical layer confidentiality, the covert communi- cation rate or secrecy rate for each -th user, which quantifies the secure information transmission rate, is defined as:

where is the legitimate communication rate to the -th user, and is the potential information rate accessible to the -th eavesdropper. These rates are articulated by:

effectively capturing the dynamics of secure and potentially compromised communication paths.

The principal objective in this advanced communi- cation paradigm is to maximize the sum of all users' secrecy rates, which is pivotal for ensuring robust secure communication against eavesdropping threats.

**Problem Formulation**

In our comprehensive RIS-UAV-NOMA downlink network model, focused on enhancing PLS alongside communication efficiency, our objective extends to not only maximizing the sum rate but also incur- porating an aspect of security by optimizing several key parameters. These parameters include the power allocation at the Base Station (BS), the phase-shifting of the RIS, and the horizontal positioning of the UAV. The integration of these aspects leads to a complex optimization problem that can be articulated as follows:

In this formulation, represents the secrecy rate for user , enhancing the PLS by considering the potential eavesdropping threats. Constraint (5b) ensures the Quality of Service (QoS) for all users by guaranteeing that the secrecy rate for each user is above a predefined minimum . Constraint (5c) is pivotal for the implementation of Successive Interference Cancellation (SIC), ensuring that the decoding process for NOMA can be executed effectively. The constraint (5d) encapsulates the total transmission power limitation at the BS, ensuring efficient power usage. Constraint (5e) specifies the operational area for the UAV, ensuring it remains within a designated feasible region . Lastly, constraint (5f) governs the phase shifts applied by the RIS, with each element confined within a to range.

Given the non-convex nature of this optimization problem, primarily due to the intricate interplay between the variables , finding a global optimal solution presents significant challenges. To navigate these complexities, we propose sophisticated and efficient solution framework based on Deep Reinforcement Learning (DRL), specifically utilizing PPO. This approach is designed to tackle the non-convexity and high dimensionality inherent in the problem, offering a robust and low-complexity method to achieve near-optimal solutions, thereby ensuring an enhanced and secure communication network.

**DRL-Based Approach for Securing NOMA Communications**

In this section we detail the Deep Reinforcement Learning (DRL) strategy utilized to optimize the UAV-assisted NOMA communication system which is aimed at enhancing confidentiality through intelligent and dynamic control of the RIS-equipped UAV. Notably, DRL excels in environments requiring complex decision-making by learning optimal state-action mappings directly from interactions with the environment without the need for predefined training [27].

DRL algorithms are distinguished into two primary types: On-policy and Off-policy methods. On-policy algorithms such as PPO Asynchronous Advantage Actor-Critic (A3C) and Phasic Policy Gradient (PPG) [28] evaluate and improve the same policy used to make decisions in the environment. These methods maintain the exploration and exploitation balance by updating the policy incrementally using a function known as the clipped surrogate objective which ensures that the new policy does not deviate significantly from the existing policy.

Conversely, Off-policy algorithms like Deep Determ- inistic Policy Gradient (DDPG) operate independently of the agent policy, allowing the exploration of new strategies by learning from actions that are outside the current policy. This characteristic can be advantageous in environments where acquiring new data is costly or risky.

Given the UAV-assisted NOMA system's setup, where a base station communicates with multiple users via a UAV-mounted RIS, both On-policy and Off-policy methods are applicable for optimizing signal reflections and directing transmissions securely away from potential eavesdroppers. However, for our spec- ific implementation challenges—characterized by the need for rapid adaptation to dynamic communication channels and the imperative of minimizing compu- tational overhead On-policy methods are preferable.

Thus, we adopt PPO, an On-policy algorithm known for its effectiveness in continuous action spaces and its computational efficiency. PPO will be used to dynamically adjust the RIS's phase shifts () to maximize the secure transmission rates while minim- izing potential interceptions by eavesdroppers. The algorithm will interact with the environment to maxi- mize the expected sum of secrecy rates across all communications, which is mathematically defined for each user in the system as:

Where and represent the transmission rates to the legitimate users and the potential eavesdroppers, respectively. The aim of our study is to optimize the configuration of the RIS to maximize and minimize across all channels thereby enhancing the NOMA communication system's overall security. The subse- quent subsections will provide an overview of the DRL formulation and an in-depth analysis of the PPO algorithm that has been customized for our particular application. In the following subsections, we will outline the DRL formulation and delve into the specifics of the PPO algorithm tailored for our application.

**Proposed PPO**

The implementation of the PPO algorithm is essential within the integrated system that combines RIS technology with UAV-assisted NOMA downlink com- munications, aiming to bolster Physical Layer Security (PLS). Initially, this section provides a concise intro- duction to PPO before examining the modifications made to the PPO framework to address the particular optimization challenges presented by this system.

PPO as opposed to conventional algorithms such as Deep Q-Networks (DQN), which are designed for dis- crete action spaces is specifically optimized for envir- onments characterized by continuous action spaces. PPO being a model-free on-policy algorithm utilizes a stochastic policy gradient approach that is improved by cropped probability ratios in the objective function. By establishing a robust equilibrium between explo- ration and exploitation this design empowers PPO to efficiently traverse environments characterized by a multitude of continuous actions.

**The PPO Framework**: PPO employs an actor-critic structure where policy and value functions are distinct. Its efficacy stems from the equilibrium it maintains between exploration of new strategies and exploitation of known rewards. This balance is achieved by constraining the magnitude of updates to the policy, a process regulated by a clipping mechanism in the objective function that curtails excessively large and potentially disruptive updates. The equation you've provided can be condensed into two lines as follows:

 (12)

where is a small number like 0.1 or 0.2, which defines the clipping range to moderate updates, thus stabilizing the training process.

 **The PPO Processing**: The state space delineates the complete set of observable variables at each discrete time step undefined, providing a full repre- sentation of the system's current condition. It includes the previous secrecy metrics (undefined) for each user undefined, the most recent RIS phase modulations and power distribution along with the current coordinates of the UAV. These elements collectively form the state space, offering the necessary information to inform subsequent decisions.

**PPO Algorithm for RIS-UAV-NOMA Communication System**

The state space () encompasses critical elements such as the UAV's position, represented by coordinates , the phase shifts of the RIS denoted as , and the channel state information (CSI) pertinent to all users and eavesdroppers. This collection of data points forms the foundation for decision-making within the system.

The action space () includes several key adjustments crucial for system optimization: the power allocation coefficients () for each user, the RIS phase shift matrix (), and the future positioning of the UAV, given by . These elements allow for dynamic control over the system's operational parameters to meet performance and security objectives effectively.

The reward () is calculated by evaluating the secrecy rates of users, with the objective of optimizing both the network performance and security. This ensures that the system not only functions efficiently but also maintain stringent security standards, reflecting a dual-focus approach in the optimization process.

**Proposed PPO-based Algorithm
Initialization:**

Randomly initialize the policy network π(a|s; θ^π) and the value network V(s; θ^V) with their respective weights θπ and θV.

Prepare the experience replay buffer 𝓓 with capacity C.

Set learning parameters: learning rate β, discount factor γ, clipping parameter ε, and minibatch size N\_B.
Procedure:
For each iteration i = 1 to I:
For each episode j = 1 to J:
Initialize the UAV's position v(x, y), RIS phase shifts Φ, and obtain the channel state information G^(j) and h\_rk^(j) for all users and eavesdroppers.
Calculate the initial state s\_1.
For each timestep t = 1 to T:
Select an action a\_t based on the current policy π(a|s\_t; θ^π) that adjusts the RIS phase shifts, power allocations, and the UAV's positioning.
Calculate the new secrecy rates R\_{s,k}^{(t)} for each user to determine the total secrecy rate, which is used to compute the reward r\_t.
Capture the updated state s\_{t+1}, including changes in the UAV’s position, RIS settings, and channel state information.
Store the transition data (current state, action, reward, new state) in the replay buffer 𝓓, which supports the learning process by providing historical data for future analysis.

Sample a minibatch of N\_B transitions from 𝓓 for training.

For each sampled transition, compute advantages using the difference between returns and value estimates.

Update the policy by maximizing the clipped surrogate objective function:
L(θ^π) = 𝔼[min(ratio × advantage, clip(ratio, 1-ε, 1+ε) × advantage)],
where ratio = π(a|s; θ^π)/π\_old(a|s).

Update the value network by minimizing the loss function:
L(θ^V) = 𝔼[(V(s; θ^V) - Returns)^2].

**Simulation Results**

In our study, we utilize a Proximal Policy Optimization (PPO) approach designed specifically for an RIS-enabled UAV-NOMA communication syst- em, aiming to assess its potential in enhancing both the performance and security of the system. The simul- ation strategically places the Base Station (BS) at the origin and positions the RIS-equipped UAV at an initial coordinate of (40,0). The designated area for users, marked by the vertices (35,35), (45,35), (45,45), and (35,45), accommodates users whose locations are fixed for each simulation run.

The system maintains Line-of-Sight (LoS) connectivity between the BS and the RIS, as well as from the RIS to the users, utilizing a Rician fading model expressed as:

 (13)

In this model, denotes the line-of-sight component, the non-line-of-sight component influenced by Rayleigh fading, and the Rician K-factor set at 8. A path loss exponent () of 2 is used, with channel conditions being predetermined at the beginning of each episode.

The UAV is positioned at an altitude of 20 meters, matching the elevation of the Base Station (BS). The system's noise power is configured to dB, and it requires a minimum user rate of bps/Hz. The policy (actor) and value (critic) networks within the PPO framework are structured using dual-layer fully connected neural networks, incorporating ReLU activation functions in the first layer and tanh in the output layer to ensure efficient gradient propagation.

Key hyperparameters guiding the simulation include a learning rate () of 0.0001, a discount factor () of 0.95, a soft update rate () of 0.004, and a replay buffer with a capacity of 40,000. The simulation unfolds over 400 episodes, with each episode consi- sting of 200 steps, and utilizes a minibatch size of 16. Exploration within the system is driven by noise characterized by a complex Gaussian distribution with zero mean and a variance of 0.1, promoting varied policy exploration.

To account for potential security threats, the simula- tion incorporates the RIS-to-Eve channel, , and modifies Eve's channel gain relative to that of legitimate users, specifically adjusted for her distinct location. This inclusion is essential for realistically assessing and mitigating the risk of eavesdropping in the communication system.

This configuration allows the PPO algorithm to dynamically tune power allocations () and phase adjustments () to continuously optimize the security and efficiency of communications in response to evolving threats and communication dynamics.



**Fig 2: Number of RIS elements**

In Fig 2, the line graph illustrates how the secrecy rate, measured in bps/Hz, escalates with an increase in the number of RIS elements within a communication system. Three distinct curves represent various transmit power levels—specifically 15 dBm, 20 dBm, and 25 dBm. Each curve trends upwards, clearly showing that the addition of RIS elements is positively correlated with the improvement of the system's secrecy rate.
Notably, at the point where the system employs 40 RIS elements, there are marked percentage increases in secrecy rates: a 16.6 percent rise at a transmit power of 15 dBm, and an even more significant 42.8 percent surge at 25 dBm. These figures suggest substantial gains in secrecy performance due to the increased number of RIS elements, with higher transmit powers yielding more pronounced benefits. While the graph does not specify the baseline from which these percentages are derived, it is evident that incorporating a greater number of RIS elements can greatly enhance the secure transmission capabilities of the system, particularly when operated at higher power levels.

Fig 3 illustrates a comparison of secrecy rates, measured in bps/Hz, as a function of varying transmit powers at the BS, denoted in dBm. The configurations under examination include systems with 40 RIS elements at a user-to-eavesdropper ratio of 6, as well as systems with 30 RIS elements at user-to-eavesdropper ratios of 6 and 4. A consistent upward trend across these scenarios reveals that higher transmit powers correlate with increased secrecy rates. The highest secrecy rate is observed in the system boasting the greatest number of RIS elements coupled with the largest user-to-eavesdropper ratio, highlighting the significant roles both factors play in the enhancement of secure communications.



**Fig 3 : Transmit power**

****

**Fig 4: Number of episodes**

In Fig 4, the plotted data illustrates the correlation between accumulated rewards and the number of completed episodes for three different configurations of the parameter . These configurations are represented by the blue, red, and black lines, corresponding to values of 0.001, 0.01, and 0.1, respectively. Each line shows a reward increase as the episode count climbs to 400. The blue line exhibits a notable ascent and eventually levels off above a reward value of 12. The red line presents a steady, moderate incline, reaching stability around a reward value of 8. The black line, on the other hand, after a promising start, plateaus at the lowest final reward value, slightly above 4. This visualization underscores the influence of the parameter on the learning trajectory, where a setting of appears to lead to the highest reward attainment after 400 episodes.



**Fig 5: Impact of various transmit power at BS on Secrecy rate**

Fig 5 charts the relationship between secrecy rates in bps/Hz and the base station's transmit power, displayed in dBm, across a range of communication techniques. The graph reveals that the PPO-FD method, featuring Full Duplex with Proximal Policy Optimization and illustrated by pink stars, significa- ntly leads in performance, climbing sharply with increased power. Full Duplex without the addition of RIS, depicted with grey circles, shows a consistent but less dramatic rise. Half Duplex setups, both with RIS (blue crosses) and without (orange triangles), demons- trate only slight gains, with the latter being the least improved. The SR-FD strategy, represented by cyan squares, exhibits moderate progress, better than Half Duplex without RIS but not as effective as Full Duplex approaches. This visualization underscores the efficacy of combining PPO and RIS in Full Duplex systems for superior secrecy rate enhancements as transmit power is amped up.

**Conclusion**

This study investigates the potential of using an Intelligent RIS together with a UAV to protect data transmission in a downlink NOMA network. The approach utilizes the PPO algorithm to precisely adjust the phase shifts of the IRS allocate power to the BS, and position the UAV spatially. To ensure SIC confid- entiality while increasing overall data rate, a decoding sequence adaptable to fluctuating channel conditions is necessary. Simulation results demonstrate that this PPO-centric approach enhances network performance as a whole and effectively handles variations in IRS configurations and user counts. Importantly, it imple- ments heightened preventive measures against unauth- orized entry, strengthening network security against surveillance while guaranteeing unrestricted access for authorized users; thus representing significant improv- ement in wireless network security.

**References**

[1] J. Li, S. Xu, J. Liu, Y. Cao, and W. Gao, “Reconfigurable Intelligent Surface Enhanced Secure Aerial-Ground Communication,” IEEE Transactions on Communications, vol. 69, no. 9, pp. 6185–6197, Jun, 2021.

[2] V. N. Vo, C. So-In, H. Tran, D.-D. Tran, and T. P.Huu, “Performance Analysis of an Energy-Harvesting IoT System Using a UAV Friendly Jammer and NOMA Under Cooperative Attack,” IEEE Access, vol. 8, pp. 221 986–222 000, Dec, 2020.

[3] A. S. Abdalla, T. F. Rahman, and V. Marojevic, “UAVs with reconfigurable intelligent surfaces: Appli-cations, challenges, and opportunities,” arXiv preprintarXiv:2012.04775, 2020.

[4] W. U. Khan, A. Mahmood, C. K. Sheemar, E. Lagunas,S. Chatzinotas, and B. Ottersten, “Reconfigurable Intelligent Surfaces for 6G Non-Terrestrial Networks: AssistingConnectivity from the Sky,” IEEE Internet of Things Magazine, vol. 7, no. 1, pp. 34–39, Jan, 2024.

[5] T. Alladi, Naren, G. Bansal, V. Chamola, and M. Guizani, “Secauthuav: A novel authentication scheme for uavground station and uav-uav communication,” IEEE Transactions on Vehicular Technology, vol. 69, no. 12, pp.15 068–15 077, Oct, 2020.

[6] S. Jiao, X. Xie, and Z. Ding, “Deep reinforcement learning based optimization for irs based uav-noma downlink networks,” arXiv preprint arXiv:2106.09616, 2021.

[7] J. Yao and J. Xu, “Joint 3D Maneuver and Power Adaptation for Secure UAV Communication With CoMP Reception,” IEEE Transactions on Wireless Communications,vol. 19, no. 10, pp. 6992–7006, July, 2020.

[8] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, “Physics-Based Modeling and Scalable Optimization of Large Intelligent Reflecting Surfaces,” IEEE Transactions on Communications, vol. 69, no. 4, pp. 2673–2691, Dec, 2021.

[9] F. Zhou, Z. Li, J. Cheng, Q. Li, and J. Si, “Robust AN-Aided Beamforming and Power Splitting Design for Secure MISO Cognitive Radio With SWIPT,” IEEE Transactions on Wireless Communications, vol. 16, no. 4, pp. 2450–2464, Mar, 2017.

[10] X. Zhang, X. Zhou, and M. R. McKay, “Enhancing Secrecy With Multi-Antenna Transmission in Wireless Ad Hoc Networks,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 11, pp. 1802–1814, Aug, 2013.

[11] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, “Reconfigurable Intelligent Surfaces: Principles and Opportunities,” IEEE Communications Surveys Tutorials, vol. 23, no. 3, pp. 1546–1577, May, 2021.

[12] X. Sun, D. W. K. Ng, Z. Ding, Y. Xu, and Z. Zhong, “Physical layer security in uav systems: Challenges and opportunities,” IEEE Wireless Communications, vol. 26, no. 5, pp. 40–47, Oct, 2019.

[13] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, “Improving physical layer security for reconfigurable intelligent surface aided noma 6g networks,” IEEE Transactions on Vehicular Technology, vol. 70, no. 5, pp. 4451–4463, Mar, 2021.

[14] C. Huang, G. C. Alexandropoulos, A. Zappone, C. Yuen, and M. Debbah, “Deep Learning for UL/DL Channel Calibration in Generic Massive MIMO Systems,” in ICC 2019- 2019 IEEE International Conference on Communications

(ICC), Shanghai, China, July, 2019, pp. 1–6.

[15] T. Wang, F. Fang, and Z. Ding, “An SCA and Relaxation Based Energy Efficiency Optimization for Multi-User RIS-Assisted NOMA Networks,” IEEE Transactions on Vehicular Technology, vol. 71, no. 6, pp. 6843–6847, Mar, 2022.

[16] J. Chen, L. Guo, J. Jia, J. Shang, and X. Wang, “Resource Allocation for IRS Assisted SGF NOMA Transmission: A MADRL Approach,” IEEE Journal on Selected Areas in Communications, vol. 40, no. 4, pp. 1302–1316, 2022.

[17] J. Chen, Y. Xie, X. Mu, J. Jia, Y. Liu, and X. Wang, “Energy Efficient Resource Allocation for IRS Assisted CoMP Systems,” IEEE Transactions on Wireless Communications, vol. 21, no. 7, pp. 5688–5702, Jan, 2022.

[18] J. Chen, Z. Ma, Y. Zou, J. Jia, and X. Wang, “DRL-based Energy Efficient Resource Allocation for STAR-RIS Assisted Coordinated Multi-cell Networks,” in GLOBECOM 2022 - 2022 IEEE Global Communications Conference,Jan, 2022, pp. 4232–4237.11

[19] X. Mu, Y. Liu, L. Guo, J. Lin, and H. V. Poor, “Intelligent Reflecting Surface Enhanced Multi-UAV NOMA Networks,” IEEE Journal on Selected Areas in Communications, vol. 39, no. 10, pp. 3051–3066, Jun, 2021.

[20] J. Zhao, Y. Zhu, X. Mu, K. Cai, Y. Liu, and L. Hanzo, “Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface (STAR-RIS) Assisted UAV Communications,” IEEE Journal on Selected Areas in Communications, vol. 40, no. 10, pp. 3041–3056, Aug, 2022.

[21] Z. Tang, T. Hou, Y. Liu, J. Zhang, and C. Zhong, “A Novel Design of RIS for Enhancing the Physical Layer Security for RIS-Aided NOMA Networks,” IEEE Wireless Communications Letters, vol. 10, no. 11, pp. 2398–2401, Aug, 2021.

[22] M. Cui, G. Zhang, and R. Zhang, “Secure Wireless Communication via Intelligent Reflecting Surface,” IEEE Wireless Communications Letters, vol. 8, no. 5, pp. 1410–1414,May, 2019.

[23] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, “Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces,” IEEE Journal on Selected Areas in Communications, vol. 38, no. 11, pp. 2637–2652, July,2020.

[24] Y. Han, N. Li, Y. Liu, T. Zhang, and X. Tao, “Artificial Noise Aided Secure NOMA Communications in STAR-RIS Networks,” IEEE Wireless Communications Letters, vol. 11, no. 6, pp. 1191–1195, Mar, 2022.

[25] H. Long, M. Chen, Z. Yang, Z. Li, B. Wang, X. Yun, and M. Shikh-Bahaei, “Joint Trajectory and Passive Beamforming Design for Secure UAV Networks with RIS,” in 2020 IEEE Globecom Workshops (GC Wkshps, Mar, 2020, pp.1–6.

[26] D. Wang, Y. Zhao, Y. Lou, L. Pang, Y. He, and D. Zhang, “Secure NOMA Based RIS-UAV Networks: Passive Beam-forming and Location Optimization,” in GLOBECOM 2022- 2022 IEEE Global Communications Conference, Rio deJaneiro, Brazil, Jan, 2022, pp. 3168–3173

[27] Y. Li, W. Han, and Y. Wang, “Deep reinforcement learning with application to air confrontation intelligent decision making of manned/unmanned aerial vehicle cooperativesystem,” IEEE Access, vol. 8, pp. 67 887–67 898, 2020.

[28] S. S. Gu, T. Lillicrap, R. E. Turner, Z. Ghahramani,

B. Sch ̈olkopf, and S. Levine, “Interpolated policy gradient:Merging on-policy and off-policy gradient estimation for deep reinforcement learning,” Advances in neural information processing systems, vol. 30, 2017.