Secure Access Selection for Space-Air-Ground Integrated Network Based on Deep Learning

Zhaowei Wang* Zhisheng Yin*, Xiucheng Wang†, Cheng Nan†, and Tom H. Luan*
* School of Cyber Engineering, Xidian University, Xi'an, 710071, China
†School of Telecommunications Engineering, Xidian University, Xi'an, 710071, China
Email: 1916040218@s.upc.edu.cn, {zsyin, tom.luan}@xidian.edu.cn, xcwang_1@stu.xidian.edu.cn, dr.nan.cheng@ieee.org

***Abstract*—Due to the openness and wide broadcast coverage of the Space-air-ground integrated network (SAGIN), SAGIN faces serious eavesdropping threats. Physical layer security can effectively address the security transmission issues in SAGIN, but traditional algorithms suffer from high complexity in multi-user, multi-antenna scenarios. Deep learning can be called multiple times after training, with simple computation that saves computational resources. In this paper, we investigate the scenario where multi-mode users have the ability to access satellites, drones, or ground base stations in the presence of eavesdroppers. We propose a deep Q-network(DQN) based deep reinforcement learning approach to select the best access strategy with the highest security condition, and optimize the power allocation under this access strategy through unsupervised learning to maximize the secrecy rate. The numerical results demonstrate that our proposed power optimization method and access strategy outperform traditional power allocation schemes in achieving better security transmission.**

***Index Terms*—Space-air-ground integrated network, physical layer security, power allocation, deep learning**

I. Introduction

SAGIN, as a novel network design that integrates satellites, aerial platforms, and ground communication systems, integrates heterogeneous networks to enhance network resource utilization and wireless access transmission capacity, attracting wide attention in academia and industry [1]. SAGIN's heterogeneity, self-organization, and time-varying characteristics offer significant advantages for various services and applications, but also pose challenges such as resource allocation and management, power control, and secure transmission [2]–[4]. In particular, due to the open nature and broad coverage of wireless channels, as well as the dynamic topology of satellite networks, communication in SAGIN is vulnerable to eavesdropping threats. Moreover, the complex and extensive geographical environment provides ample hiding space for potential attackers and eavesdroppers (Eve), resulting in serious security issues [5]. Ensuring secure transmission in SAGIN networks has become an urgent problem to be addressed.

In the context of SAGIN, multi-mode terminals are devices that can seamlessly connect to satellite, unmanned aerial vehicles (UAVs), base stations (BSs), and other heterogeneous networks. They can autonomously select the most secure mode of access for communication based on the perceived channel characteristics. The selection strategy of terminals is influenced by the randomness of wireless channels, and the access strategies of different multi-mode terminal devices can mutually affect each other. Therefore, determining the most secure communication access method has become a key consideration for multi-mode terminals.

Physical layer security, which utilizes the characteristics of wireless channels and employs physical layer techniques for secure communication, can serve as an effective complement to traditional encryption techniques based on key systems, providing comprehensive information security protection [6]. Physical layer security leverages the differences in randomness between channels to achieve lightweight secure transmission. When the quality of the main channel is better than that of the eavesdropping channel, legitimate users can achieve secure transmission through lossless encoding. However, in scenarios with a large number of antennas or users, traditional physical layer security optimization algorithms can become very complex [7].

Artificial intelligence has brought new opportunities to the research of physical layer security with its vigorous development. Deep learning, as a significant branch of artificial intelligence, demonstrates excellent performance in handling large-scale data. Its main objective is to construct models and extract features from sample data by learning the underlying patterns, which is suitable for predicting the future and making decisions based on current data. And after training, the inference computation of deep learning is much simpler than traditional algorithms, as it can be called multiple times after training, which can save computational resources. Currently, numerous experts and scholars have proposed solutions to physical layer security issues based on deep learning, and achieved promising results [8]–[11]. However, there is limited literature that addresses the security access selection problem of multi-mode terminals in heterogeneous networks using deep learning, which motivates the focus of this work.

In this paper, we consider the same-frequency interference caused by spectrum sharing among satellites, UAVs, and base stations in the heterogeneous network scenario of SAGIN. We aim to maximize the signal-to-interference-plus-noise ratio (SINR) at legitimate users and deteriorate the SINR at eavesdroppers. To achieve this, we propose a cascaded framework that combines DQN based reinforcement learning and unsupervised learning to obtain the optimal access selection strategy and power allocation scheme under the current access strategy. Furthermore, we compare the performance of our proposed approach with other common power allocation methods.

II. SYSTEM MODEL

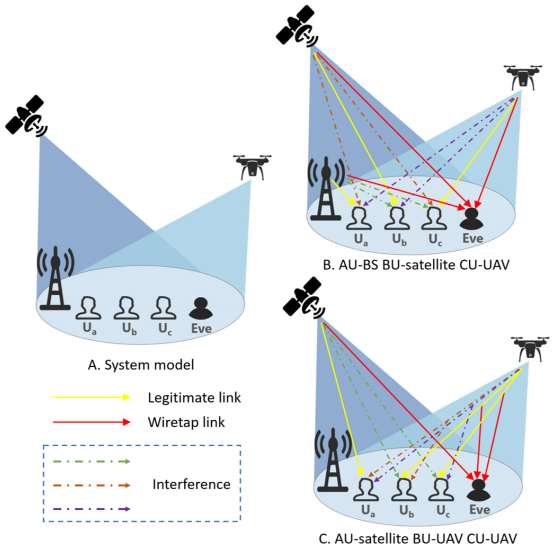

Fig. 1. Downlink transmission for multi-modal users in SAGIN

As shown in Figure 1, we consider a scenario where there are three multi-mode users and one eavesdropper in the overlapping coverage area of a satellite, a UAV , and a BS. The satellite, UAV, and BS share the spectrum to improve spectrum efficiency. User A (U_a), user B (U_b), and user C (U_c) can access any one of the satellite, UAV, or BS, depending on the current channel conditions. In the eavesdropping scenario of SAGIN, Eve exists within the overlapping coverage area of the satellite, UAV, and BS, and shares the same spectrum with legitimate users. Therefore, confidential information may leak to Eve. Users select the most secure access mode by perceiving the current channel environment, and the selected access node transmits signals to the target user with optimized transmission power to achieve the most secure communication transmission. Scenarios B and C in Figure 1 show two possible access modes. In scenario B, U_a accesses the BS, U_b accesses the satellite, and U_c accesses the UAV, with the eavesdropper eavesdropping on the information transmitted by the satellite, BS, and UAV, respectively. In scenario C, U_a accesses the satellite, U_b accesses the UAV, and U_c accesses the UAV, with the eavesdropper eavesdropping on the information transmitted by the satellite to U_a, and information from UAV to U_b and U_c, respectively. For any user, if they are not currently accessing a particular node but other users are accessing it, it may cause interference to the current user. For example, in Scenario B, the satellite and UAV may cause interference to U_a. Similarly, when multiple users are accessing the same node, the information transmitted by that node to one user may also cause interference to other users. For example, in Scenario C, the information sent by the UAV to U_b may cause interference to U_c.

A. Physical Layer Channel Models

In the access scenario of multi-modal users in SAGIN, there are three types of communication links, including satellite-to-ground link, air-to-ground link, and ground-to-ground link

The formula for the satellite-to-ground link is given by [12]:

$$h = \sqrt{C_L b \beta} \exp\left(-j\theta\right), \tag{1}$$

Where C_L represents free space loss, and its formula is given as $C_L = \left(\lambda/4\pi\right)^2 / \left(d^2 + h^2\right)$, Where λ represents the wavelength of the signal, d represents the horizontal distance from the satellite beam center to the ground user, and h represents the height of the satellite. β represents the channel gain caused by rain attenuation, which follows a log-normal random variable, i.e., $\ln\left(\beta_{dB}\right) \sim \mathcal{N}\left(\mu, \delta^2\right)$ with β_{dB} is the dB form of β. θ is a phase vector uniformly distributed in the range $[0, 2\pi)$. b represents the satellite beam gain, which is defined as:

$$b = G\left(\frac{J_1\left(u_0\right)}{2u_0} - 36\frac{J_3\left(u_0\right)}{u_0^2}\right)^2, \tag{2}$$

In the equation, G represents the maximum gain of the satellite antenna, $u_0 = 2.07123\sin\left(\alpha\right)/\sin\left(\alpha_{3dB}\right)$, Where α is the elevation angle between the beam center and the user, α_{3dB} is the $3dB$ angle of the satellite beam, $J_1\left(\cdot\right)$ and $J_3\left(\cdot\right)$ are the first and third order Bessel functions of the first kind. Therefore, assuming $h_{S,a}$, $h_{S,b}$, $h_{S,c}$, $h_{S,e}$ are the channel power gain from the satellite to U_a, U_b, U_c, and Eve, respectively.

The air-to-ground link can be defined as

$$a = \sqrt{G_L}\left(\sqrt{\frac{K}{K+1}}a_{LoS} + \sqrt{\frac{1}{K+1}}a_{Ray}\right), \tag{3}$$

Where G_L represents the path loss, given by: $G_L = g_0/\left(U_d^2 + U_h^2\right)$, g_0 represents the channel power gain at a reference distance of 1 m, U_d is the horizontal distance from UAV to the target user, and U_h is the height of the UAV. Small-scale fading follows the Rician channel model, where K is the Rician factor, a_{LoS} represents the line-of-sight Rician fading component, and a_{Ray} represents the non-line-of-sight Rayleigh fading component. Therefore, assuming $h_{U,a}$, $h_{U,b}$, $h_{U,c}$, $h_{U,e}$ represent the channel power gain from the UAV to U_a, U_b, U_c, and Eve, respectively.

The ground link can be defined as

$$g = \sqrt{\alpha} g_0, \tag{4}$$

The parameter α represents large-scale fading, with the formula $\alpha = C_0 r^{-4}$, C_0 denotes the channel power gain at a reference distance of 1 m, while r represents the distance between the base station and the user. g_0 represents small-scale fading, following a $Nakagami - m$ distribution. Therefore, assuming $h_{B,a}$, $h_{B,b}$, $h_{B,c}$, $h_{U,e}$ represent the channel power gain from the base station to U_a, U_b, U_c, and Eve, respectively.

B. The SAGIN downlink channel model.

In SAGIN downlink communication, legitimate users request satellite, UAV, or BS to transmit confidential information to them. However, the choice of which transmitting device to request depends on the security performance of different access methods under the current channel state, while also

meeting the requirements of basic communication transmission rates.

By defining the set of access points and the user index set, i.e., $\mathcal{A} = \{Sat, UAV, BS\}$ and $\mathcal{U} = \{a, b, c\}$, respectively, we calculate the achievable rate at such multi-mode users, which can be expressed as

$$R_{au} = \sum_{i \in \mathcal{A}} x_{i,a} \log_2 \left(1 + \frac{p_{i,a} h_{i,a}}{\sum\limits_{u \in \mathcal{U}, u \neq a, j \in \mathcal{A}} p_{j,u} h_{j,a} + \delta_{i,a}^2} \right), \tag{5}$$

where $p_{i,a}$ denotes the power allocation to U_a from its access point $i \in \mathcal{A}$, and similarly,

$$R_{bu} = \sum_{i \in \mathcal{A}} x_{i,b} \log_2 \left(1 + \frac{p_{i,b} h_{i,b}}{\sum\limits_{u \in \mathcal{U}, u \neq b, j \in \mathcal{A}} p_{j,u} h_{j,b} + \delta_{i,b}^2} \right), \tag{6}$$

$$R_{cu} = \sum_{i \in \mathcal{A}} x_{i,c} \log_2 \left(1 + \frac{p_{i,c} h_{i,c}}{\sum\limits_{u \in \mathcal{U}, u \neq c, j \in \mathcal{A}} p_{j,u} h_{j,c} + \delta_{i,c}^2} \right). \tag{7}$$

R_{au}, R_{bu}, R_{cu} represent the communication rate for receiving confidential information by U_a, U_b, U_c, respectively. $p_{i,u}$ represents the transmission power allocated to user u by access point i. For any $u \in \mathcal{U}$, it is true that $\sum_{i \in \mathcal{A}} x_{i,u} = 1$, because for any user, at a certain moment, they can only choose one access point to access. Similarly, for any $i \in \mathcal{A}$, $u \in \mathcal{U}$, it is true that $x_{i,u} \in \{0, 1\}$, this is because there are only two possible scenarios for the relationship between any user and any access point, either a user is connected or not connected. $\delta_{i,a}^2$, $\delta_{i,b}^2$, $\delta_{i,c}^2$ represents the noise power of the channels. In practical scenarios, the noise power of different channels may vary, but for the ease of calculation, let $\delta^2 = \delta_{i,a}^2 = \delta_{i,b}^2 = \delta_{i,c}^2 = 1$.

The eavesdropper's communication rate is:

$$R_{ae} = \sum_{i \in \mathcal{A}} x_{i,a} \log_2 \left(1 + \frac{p_{i,a} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq a, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta_{ae}^2} \right), \tag{8}$$

$$R_{be} = \sum_{i \in \mathcal{A}} x_{i,b} \log_2 \left(1 + \frac{p_{i,b} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq b, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta_{be}^2} \right), \tag{9}$$

$$R_{ce} = \sum_{i \in \mathcal{A}} x_{i,c} \log_2 \left(1 + \frac{p_{i,c} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq c, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta_{ce}^2} \right). \tag{10}$$

R_{ae}, R_{be}, and R_{ce} represent the communication rates when the eavesdropper listens to the signals transmitted from the access points of U_a, U_b, and U_c, respectively. δ_{ae}^2, δ_{be}^2, δ_{ce}^2 represents the noise power of the eavesdropping channel. In practice, the noise power of different channels may vary, but for convenience in calculations, let $\delta^2 = \delta_{ae}^2 = \delta_{be}^2 = \delta_{ce}^2 = 1$.

According to the Wiretap Channel Model [13], the secrecy rate is defined as

$$R = \left[R_u - R_e \right]^+, \tag{11}$$

Therefore, the corresponding secrecy rates when eavesdropping on different users are defined as

$$R_a = \sum_{i \in \mathcal{A}} x_{i,a} \left(\log_2 \left(1 + \frac{p_{i,a} h_{i,a}}{\sum\limits_{u \in \mathcal{U}, u \neq a, j \in \mathcal{A}} p_{j,u} h_{j,a} + \delta^2} \right) - \log_2 \left(1 + \frac{p_{i,a} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq a, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta^2} \right) \right), \tag{12}$$

$$R_b = \sum_{i \in \mathcal{A}} x_{i,b} \left(\log_2 \left(1 + \frac{p_{i,b} h_{i,b}}{\sum\limits_{u \in \mathcal{U}, u \neq b, j \in \mathcal{A}} p_{j,u} h_{j,b} + \delta^2} \right) - \log_2 \left(1 + \frac{p_{i,b} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq b, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta^2} \right) \right), \tag{13}$$

$$R_c = \sum_{i \in \mathcal{A}} x_{i,c} \left(\log_2 \left(1 + \frac{p_{i,c} h_{i,c}}{\sum\limits_{u \in \mathcal{U}, u \neq c, j \in \mathcal{A}} p_{j,u} h_{j,c} + \delta^2} \right) - \log_2 \left(1 + \frac{p_{i,c} h_{i,e}}{\sum\limits_{u \in \mathcal{U}, u \neq c, j \in \mathcal{A}} p_{j,u} h_{j,e} + \delta^2} \right) \right). \tag{14}$$

R_a, R_b, R_c represent the secrecy rates of U_a, U_b, and U_c respectively.

C. Problem Formulation

In order to improve the secrecy performance of multiple multimode users in heterogeneous networks, the power allocation can be optimized for different access choices to maximize the average secrecy rate, and then the access scheme with the highest average secrecy rate can be selected for access.

Based on the system model, this paper studies the access selection strategy for three access points and three access users, with a total of 27 access selection schemes.The constraints for the objective function in this paper are as follows:

$$\mathcal{P}\ 1\ : \underset{p_{i,u}, x_{i,u}}{Max} \left(R_a + R_b + R_c \right), \tag{15}$$

$$s.t.: \quad R_{au} \geq Q_{\min}, \tag{15a}$$

$$R_{bu} \geq Q_{\min}, \tag{15b}$$

$$R_{cu} \geq Q_{\min}, \tag{15c}$$

$$\sum_{u \in \mathcal{U}} x_{i,u} p_{i,u} \leq P_S, i \to Sat, \tag{15d}$$

$$\sum_{u \in \mathcal{U}} x_{i,u} p_{i,u} \leq P_U, i \to UAV, \tag{15e}$$

$$\sum_{u \in \mathcal{U}} x_{i,u} p_{i,u} \leq P_B, i \to BS, \tag{15f}$$

$$x_{i,u} \in \{0,1\}. \tag{15g}$$

Where (15) is the objective function, which is related to the allocation of transmission power and access selection strategy. (15a)-(15c) represent the QoS constraints of legitimate users; (15d)-(15f) represent the transmission power constraints of access points; (15e) represents the access selection between the current access point and the user. It is worth noting that at a fixed time, a user can only access one access point, but an access point can be accessed by multiple users simultaneously. For example, in scenario C of Figure 1, when both U_b and U_c access the UAV simultaneously, the corresponding constraint (15e) is $p_{i,b} + p_{i,c} \leq P_U$

III. Power Optimization and Access Selection Based on Deep Learning

In the scenario studied in this paper, there are three access points and three multi-mode users, resulting in 27 access selection strategies. For each of the 27 access selection strategies, power allocation strategies are optimized separately with the objective of maximizing the sum of $R_a + R_b + R_c$ for each access strategy. The access mode that maximizes $R_a + R_b + R_c$ is chosen as the optimal access strategy.

A. Power optimization

The power optimization network is shown in Figure 2 as follows: We adopt unsupervised deep learning for power

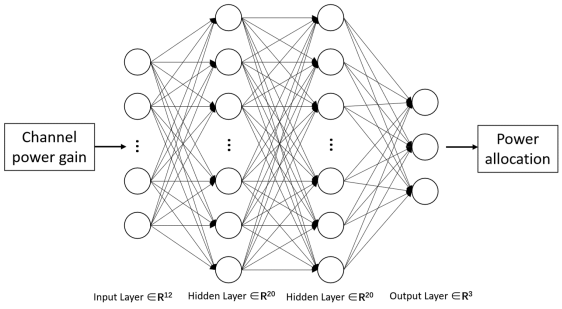

Fig. 2. Power allocation network

optimization.A four-layer Multilayer Perceptron (MLP) is used for training in each network. The first hidden layer has 20 neurons, and the second hidden layer also has 20 neurons. Dropout layers are added after the hidden layers to prevent overfitting, with a dropout rate set to 0.2. ReLU is used as the activation function after each layer.During training, the parameters and weights of the neural network are continuously adjusted using the stochastic gradient descent (SGD) algorithm. $R_a + R_b + R_c$ will converge continuously until the optimal secrecy performance is achieved. Additionally, the power allocation for each user corresponding to the output of the neural network can be obtained. The neural network is capable of effectively learning the statistical characteristics of the channels, and jointly optimizing the channel gains of legitimate users and eavesdroppers to maximize system security and rate.

- State: The input of the network is a vector whose size depends on the number of current access points, access users, and eavesdroppers. The values in the vector represent the channel power gains between the access points and access users, as well as between the access points and eavesdroppers. Specifically, in this paper, the input vector of neural network consists of 12 channel power gain values: $h_{S,a}$, $h_{S,b}$, $h_{S,c}$, $h_{S,e}$, $h_{U,a}$, $h_{U,b}$, $h_{U,c}$, $h_{U,e}$, $h_{B,a}$, $h_{B,b}$, $h_{B,c}$, and $h_{B,e}$, respectively.
- Action: The output is a vector representing the optimized power allocation values $p_{i,a}$, $p_{i,b}$, and $p_{i,c}$. The optimization objective is to maximize the security performance under the current access strategy.
- Loss: The constraint conditions are incorporated into the loss function through penalty terms.

$$\begin{aligned} loss = &- (R_a + R_b + R_c) + \lambda_1 (-R_{au} + Q_{\min}) \\ &+ \lambda_2 (-R_{bu} + Q_{\min}) + \lambda_3 (-R_{cu} + Q_{\min}) \\ &+ \lambda_4 \left(\sum_{u \in \mathcal{U}, i \to Sat} x_{i,u} p_{i,u} - P_S \right) \\ &+ \lambda_5 \left(\sum_{u \in \mathcal{U}, i \to UAV} x_{i,u} p_{i,u} - P_U \right) \\ &+ \lambda_6 \left(\sum_{u \in \mathcal{U}, i \to BS} x_{i,u} p_{i,u} - P_B \right) \end{aligned} \tag{16}$$

B. Access Selection

The purpose of conducting access selection training is to determine which access selection, after optimization, can achieve better security performance when given channel vectors. When the number of access points and access users increases significantly, making access selection decisions can reduce system running time, enhance system scalability, save computational resources, and reduce system latency.

We utilized DQN for access selection training, where the objective of the neural network is to approximate the Q-network. Training was conducted using a three-layer MLP with dropout layers added to prevent overfitting, The hidden layers of the neural network had 40 neurons, and dropout layers were used to prevent overfitting, with a dropout parameter set to 0.2.

- State: The input of the network is a vector whose size depends on the current number of access points, access users, and eavesdroppers. The vector consists of channel power gains between access points, access users, and eavesdroppers. In this paper, the input vector includes 12 channel power gains: $h_{S,a}$, $h_{S,b}$, $h_{S,c}$, $h_{S,e}$, $h_{U,a}$, $h_{U,b}$, $h_{U,c}$, $h_{U,e}$, $h_{B,a}$, $h_{B,b}$, $h_{B,c}$, and $h_{B,e}$.

- Action: After passing through the network, the output is the optimized average secrecy performance for 27 possible access selection strategies. The access strategy with the highest secrecy performance is chosen as the action.
- Reward: The reward is the product of the difference between the output of the network and the maximization of $R_a + R_b + R_c$ under various access strategies obtained by power optimization, and the penalty factor.

After obtaining the two training models mentioned above, when the input channel vector is given, the network model for access selection is first used to determine the current access selection. Then, the power optimization network model corresponding to the current access selection is used to obtain the power allocation that maximizes the security performance under the current access selection.

IV. Simulation and Discussion

TABLE I
Channel Parameter Settings

Space-to-ground channel parameters				
Space-to-ground link	U_a	U_b	U_c	Eve
Horizontal distance	2200 m	2000 m	2250 m	2250 m
Satellite altitude	600 km			
Carrier frequency	2 GHz			
Rain attenuation parameters	$\mu_{dB} = -3.125$, $\delta^2 = 1.6$			
3dB Beamwidth	0.4^o			
Maximum beam gain	52 dB			
Terrestrial link parameters				
Terrestrial link	U_a	U_b	U_c	Eve
Horizontal distance	250 m	120 m	250 m	200 m
Channel power gain	-38.46 dB			
Nakagami-m channel parameters	$m = 2$, $\Omega = 1$			
Air-to-ground channel parameters				
Air-to-ground link	U_a	U_b	U_c	Eve
Horizontal distance	300 m	260 m	100 m	120 m
UAV altitude	120 m			
Channel power gain	-40 dB			
Rician factor	$m = 2$, $\Omega = 1$			
Rician channel factor	$\sigma = 2$			
Rayleigh channel factor	$s = 2$, $\sigma = 1$			

In this section, we first establish an SAGIN simulation platform and then conduct simulations to evaluate the secrecy rate performance of multi-mode users. Specifically, the system parameters are set as shown in Table I. In the satellite-to-ground link, the satellite orbit height is set to 600 km, the maximum beam gain is 52 dB, and the 3 dB beamwidth of the satellite beam is set to 0.4^o. The rain attenuation follows a log-normal random variable, i.e.: $\mathcal{N}(-3.125, 1.6)$. The horizontal distances between the satellite and U_a, U_b, U_c, Eve are set to 2200 m, 2000 m, 2250 m, and 2250 m respectively. In the terrestrial link, the channel power gain at a reference distance of 1 m is -38.46 dB, and the distances between the BS and U_a, U_b, U_c, Eve are set to 250 m, 120 m, 250 m, and 200 m respectively. The Nakagami-m fading severity is set to $m = 2$ with an average power of $\Omega = 1$. In the air-to-ground link, the UAV altitude is set to 120 m, the channel power gain at a reference distance of 1 m is -40 dB, and the Rice factor K is set to 10 dB. The Rice channel factor $\sigma = 2$, and the Rayleigh channel factor $s = 2$, $\sigma = 1$. The carrier frequency for the satellite, UAV, and BS downlink is set to 2 GHz, and the distances between the UAV and U_a, U_b, U_c, Eve are set to 300 m, 260 m, 100 m, and 120 m respectively.

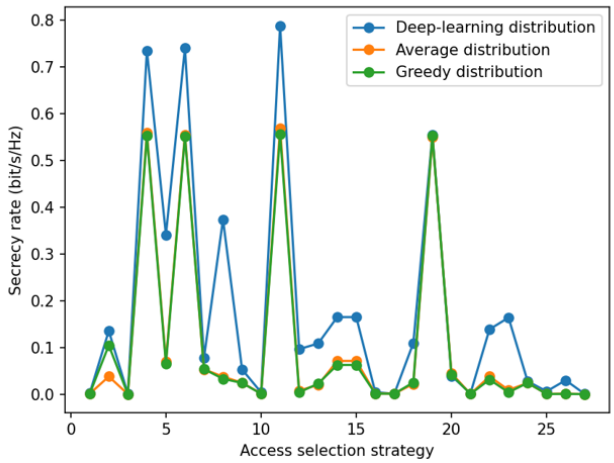

Fig. 3. Comparison of secrecy rates under different access selection strategies

In Figure 3, the horizontal axis corresponds to the 27 access selection strategies, and the vertical axis corresponds to the average secrecy rate. From the figure 3, it can be observed that the power optimization scheme proposed in this paper based on deep learning outperforms the average power allocation scheme and greedy power allocation scheme in terms of secrecy performance under all access selection strategies.

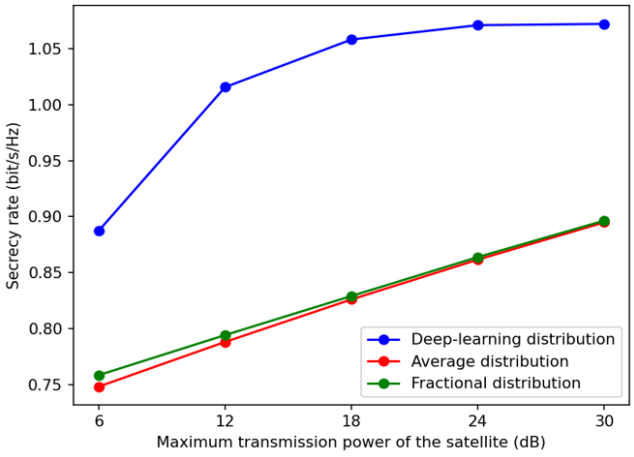

Fig. 4. Comparison of secrecy rates under satllite power upper limit

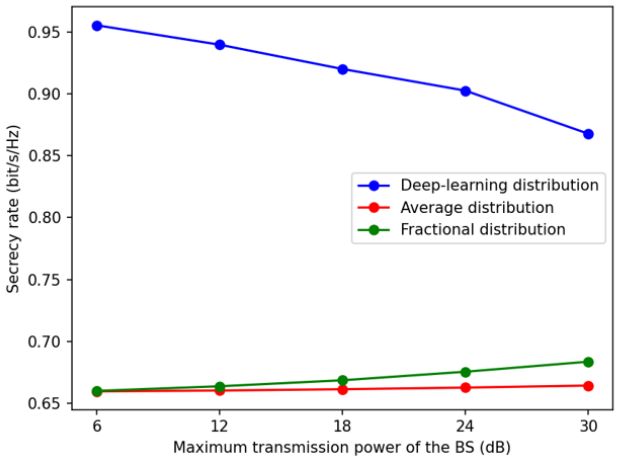

Fig. 5. Comparison of secrecy rates under BS power upper limit

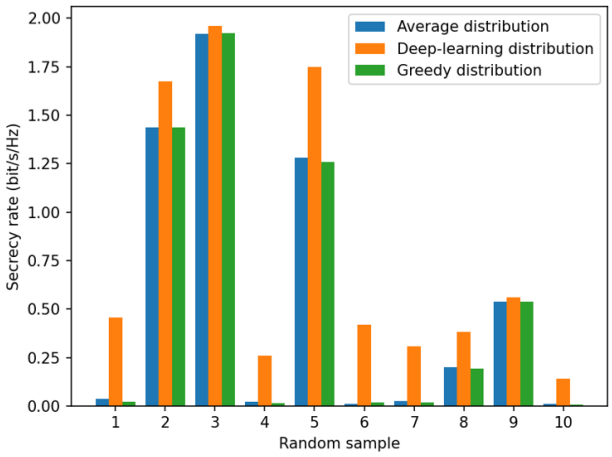

Fig. 6. Performance comparison

Figure 4 is a comparison of the security performance between the proposed deep reinforcement learning-based access selection strategy and traditional greedy power allocation, average power allocation, brute-force search, and the proposed deep learning-based power allocation. We randomly selected ten sets of data from the dataset for testing. It can be seen from the figure 4 that our proposed deep learning-based power optimization scheme outperforms average power allocation and greedy power allocation, and approaches the optimal solution obtained by exhaustive search.

V. Conclusion

This paper investigates the secure access problem of three multi-mode users in a SAGIN scenario, considering constraints on transmit power and basic communication rate. A novel approach is proposed that utilizes a cascaded DQN-based deep reinforcement learning network and an unsupervised learning network to obtain the most secure access selection strategy, and optimizes the power allocation under the current access strategy to improve the link transmission quality for legitimate users and degrade it for non-cooperative entities, thereby maximizing the system's average secrecy rate. The proposed algorithm outperforms average power allocation and greedy power allocation schemes in terms of secrecy performance. Future work will consider multi-antenna and pre-coding techniques to further improve the secrecy performance.

References

[1] N. CHENG, J. HE, Z. YIN, C. ZHOU, H. WU, F. LYU, H. ZHOU, and X. SHEN, "6g service-oriented space-air-ground integrated network: A survey," *Chinese Journal of Aeronautics*, vol. 35, no. 9, pp. 1–18, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1000936121004738

[2] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang, M. Zhu, B. Sheng, D. Wang, Z. Pan, P. Zhu, Y. Yang, Z. Liu, P. Zhang, X. Tao, S. Li, Z. Chen, X. Ma, C.-L. I, S. Han, K. Li, C. Pan, Z. Zheng, L. Hanzo, X. S. Shen, Y. J. Guo, Z. Ding, H. Haas, W. Tong, P. Zhu, G. Yang, J. Wang, E. G. Larsson, H. Q. Ngo, W. Hong, H. Wang, D. Hou, J. Chen, Z. Chen, Z. Hao, G. Y. Li, R. Tafazolli, Y. Gao, H. V. Poor, G. P. Fettweis, and Y.-C. Liang, "Towards 6g wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, no. 1, November 2020. [Online]. Available: https://europepmc.org/articles/PMC7714900

[3] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6g wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[4] L. Liu and Q. Liu, *Research on Intelligent Access of Space-Air-Ground Integrated Network*, 10 2022, pp. 66–79.

[5] Z. Yin, N. Cheng, T. H. Luan, Y. Hui, and W. Wang, "Green interference based symbiotic security in integrated satellite-terrestrial communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 9962–9973, 2022.

[6] S. A. Hoseini, F. Bouhafs, and F. den Hartog, "A practical implementation of physical layer security in wireless networks," in *2022 IEEE 19th Annual Consumer Communications Networking Conference (CCNC)*, 2022, pp. 1–4.

[7] J. Xing, T. Lv, and X. Zhang, "Cooperative relay based on machine learning for enhancing physical layer security," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1–6.

[8] L. Li, Y. Hu, H. Zhang, W. Liang, and A. Gao, "Deep learning based physical layer security of d2d underlay cellular network," *China Communications*, vol. 17, no. 2, pp. 93–106, 2020.

[9] R. Dong, B. Wang, and K. Cao, "Deep learning driven 3d robust beamforming for secure communication of uav systems," *IEEE Wireless Communications Letters*, vol. PP, pp. 1–1, 04 2021.

[10] M. Chu, A. Liu, C. Jiang, V. K. N. Lau, and T. Yang, "Wireless channel prediction for multi-user physical layer with deep reinforcement learning," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[11] T. Bao, J. Zhu, H.-C. Yang, and M. O. Hasna, "Secrecy outage performance of ground-to-air communications with multiple aerial eavesdroppers and its deep learning evaluation," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1351–1355, 2020.

[12] Z. Yin, N. Cheng, T. H. Luan, and P. Wang, "Physical layer security in cybertwin-enabled integrated satellite-terrestrial vehicle networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 4561–4572, 2022.

[13] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.