Enhancing Physical Layer Security in Low Earth Orbit Satellite Networks via Multi-Agent Reinforcement Learning

Yongjae Lee[a], Taehoon Kim[b,*], Inkyu Bang[a,*]
[a]Department of Intelligence Media Engineering, Hanbat National University
[b]Department of Computer Engineering, Hanbat National University

yjlee@edu.hanbat.ac.kr, thkim@hanbat.ac.kr, ikbang@hanbat.ac.kr

Abstract

In this paper, we propose a multi-agent deep reinforcement learning (MADRL)-based adaptive beamforming and artificial noise (AN) strategy for enhanced physical layer security in low Earth orbit (LEO) satellite networks. Multiple satellites utilize joint scheduling to cooperatively transmit data with AN against eavesdropping threats from adversarial nodes (e.g., enemy UAVs). Each satellite independently decides its transmission mode (idle, data transmission, or AN generation) to maximize secrecy rate, leveraging a centralized training for decentralized execution (CTDE) paradigm with the soft actor-critic (SAC) algorithm; adaptive beamforming and AN are generated post-action. Training exploits only the adversary's statistical channel information. The proposed scheme's performance is verified against conventional methods in terms of secrecy rate.

Ⅰ. Introduction

Low Earth orbit (LEO) satellite communications offer global coverage but are vulnerable to eavesdropping due to the broadcast nature of wireless links [1]. Physical layer security (PLS) aims to secure transmissions by exploiting wireless channel characteristics [2]. Key PLS techniques include beamforming towards legitimate users and transmitting artificial noise (AN) to confuse eavesdroppers [3]. However, jointly optimizing these in multi-satellite systems is challenging.

Deep reinforcement learning (DRL) has emerged as a promising tool for dynamic PLS design [4]. We investigate a multi-agent DRL (MADRL) approach for cooperative beamforming and AN generation in multi-satellite systems. The main contributions of our paper are summarized as follows:

1) We formulate MADRL framework where LEO satellites cooperatively select transmission modes (idle, data, AN) and design beamformers using soft actor-critic (SAC) within a centralized training for decentralized execution (CTDE) paradigm;
2) We train the proposed scheme only exploiting statistical channel information of the adversary, and further verify its performance compared with conventional schemes and baseline methods in terms of secrecy rate.

Ⅱ. System Model

We consider a system model N LEO satellites, each with M antennas, serving a single ground base station (GBS) against a single eavesdropper (i.e., adversary), as described in Fig. 1. We consider the Rician fading channel model where the channel vector $\boldsymbol{h}_{(k,i)} \in C^{M\times 1}$ from satellite i to the GBS (or eavesdropper) is denoted by $h_{k,i} = \sqrt{\frac{K}{K+1}} h_{k,i}^{LOS} + \sqrt{\frac{1}{K+1}} h_{k,i}^{NLOS}$, where K is the Rician K-factor.

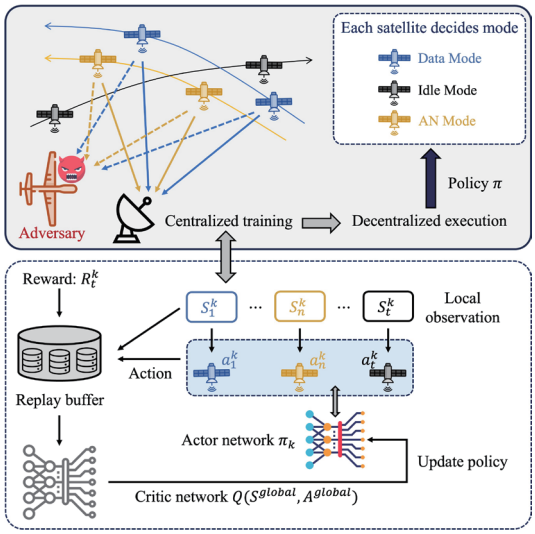

Fig 1. System Model

Each satellite independently employs one of the three operation modes (i.e., m_i): idle ($m_i = 0$, no transmission), data ($m_i = 1$ transmission of data symbol $s_{d,i}$ with beamforming vector $\boldsymbol{w}_{d,i}$), or AN ($m_i = 2$, transmission of AN signal $x_{a,i}$ using beamforming vector $\boldsymbol{w}_{a,i}$). We assume unit power for $s_{d,i}$, $\boldsymbol{w}_{d,i}$, $x_{a,i}$, and $\boldsymbol{w}_{a,i}$, and maximum transmit power P_m for each satellite. Further, $\mathcal{N}_d$ and $\mathcal{N}_a$ denote the index sets of data and the AN transmitting satellite, respectively.

For scheduled satellite index i^*, the signal-to-interference-plus-noise Ratio (SINR) at GBS is given by

$$SINR_g = \frac{\left|h_{g,i^*}^H \boldsymbol{w}_{d,i^*}\right|^2}{\sum_{j\in\mathcal{N}_d, j\neq i^*}\left|h_{g,j^*}^H \boldsymbol{w}_{d,j}\right|^2 + \sum_{k\in\mathcal{N}_a}\left|h_{g,k^*}^H \boldsymbol{w}_{a,k}\right|^2 + 1/\rho}$$

*: corresponding authors

where $\rho = P_m/\sigma_n^2$ denotes transmit signal-to-noise ratio (SNR) of the satellite for given noise variance σ_n^2. $SINR_E$ is defined similarly. Thus, the secrecy rate are $R_k = \log_2(1 + SINR_k)$. The system secrecy rate is $R_{sec} = \max\{R_g - R_e, 0\}$. The objective is to maximize the average secrecy rate $\mathbb{E}[R_{sec}]$ by optimizing satellite modes and beamformers via MADRL.

Ⅲ. Proposed MADRL-based Secure Transmission

We employ a CTDE framework with SAC. Each satellite acts as an agent.

1) Observation: Agent $i's$ local observation $s_{i,t}$ includes its channel to GBS, $\boldsymbol{h}_{g,i}$. For training with partial eavesdropper CSI.
2) Action: Each agent i selects a hybrid action $a_{i,t}$: a discrete mode $m_{i,t} \in \{0,1,2\}$ and a continuous beamforming control vector $\boldsymbol{z}_{i,t}$. If $m_{i,t} = 1$ (Data mode), $w_{d,i,t} = \boldsymbol{h}_{g,i}/\|\boldsymbol{h}_{g,i}\|$ (maximum ratio transmission, MRT). If $m_{i,t} = 2$ (AN mode), $w_{a,i,t}$ is derived from $\boldsymbol{z}_{i,t}$ and projected into the null-space of $\boldsymbol{h}_{g,i}$ to minimize interference to GBS.
3) Reward: The shared reward r_t is the instantaneous $R_{sec,t}$.

Each agent i employs an actor network $\pi_i(\cdot\,|s_{i,t}; \theta_i^\pi)$ and a temperature parameter α_i for entropy regularization. A centralized critic network $Q(s_t, a_t; \theta^Q)$ estimate Q-values using the global state $s_t = \{s_{1,t}, \cdots, s_{N,t}\}$ and the joint action $a_t = \{a_{1,t}, \cdots, a_{N,t}\}$.

The actor network for agent i is updated to maximize its objective function $J(\theta_i^\pi)$, which encourages achieving higher cumulative rewards while maintaining policy randomness (entropy):

$$J(\theta_i^\pi) = \mathbb{E}_{(s^{(t)}, a^{(t)})\sim\mathfrak{D}}\Big[Q\left(s^{(t)}, a_i^{(t)}, a_{-i}^{(t)}; \theta^Q\right) - \alpha \log \pi_i(a_i^{(t)}|s_i^t; \theta_i^\pi)\Big]$$

where $a_{-i}^{(t)}$ denotes the joint action of other agents and $\mathfrak{D}$ is the experience replay buffer.

The critic network is trained by minimizing the squared temporal difference (TD) error. This involves using a target value $y^{(t)}$ that incorporates the entropy of the next state's policy:

$$y^{(t)} = r^{(t)} + \gamma \mathbb{E}_{a^{(t+1)}\sim\pi(\cdot|s^{(t+1)})}\Big[Q'\left(s^{(t+1)}, a^{(t+1)}; \theta^Q\right) - \alpha \log \pi(a^{(t+1)}|s^{(t+1)})\Big]$$

Here, Q' is the target critic network and γ is the discount factor.

Ⅳ. Numerical Results and Conclusion

Simulations consider $N = 10$ satellites, $M = 4$ antennas/satellite, and $K = 5$. We compare our proposed MADRL schemes: "Proposed MADRL" against heuristic baselines: 1) Random MRT: one random satellite uses MRT; 2) Max MRT: best satellite for GBS uses MRT; 3) Max MRT with Random AN: Max MRT with another random satellite sending null-space AN.

Fig. 2 shows that the proposed MADRL scheme consistently outperforms the baselines. This superiority arises from the agents adaptively selecting transmission modes (Idle, Data, AN) and optimizing beamformers via learned policies using a CTDE with the SAC algorithm. Notably, the proposed MADRL scheme demonstrates significant gains even with only statistical channel information of the adversary, rather than full instantaneous CSI, under practical constraints, thereby proving its efficacy. This underscores the potential of MADRL for dynamic and secure resource allocation in complex multi-satellite environments. Future research directions may include advanced channel and hardware modeling, multi-user and multiple eavesdropper scenarios, and energy efficiency.

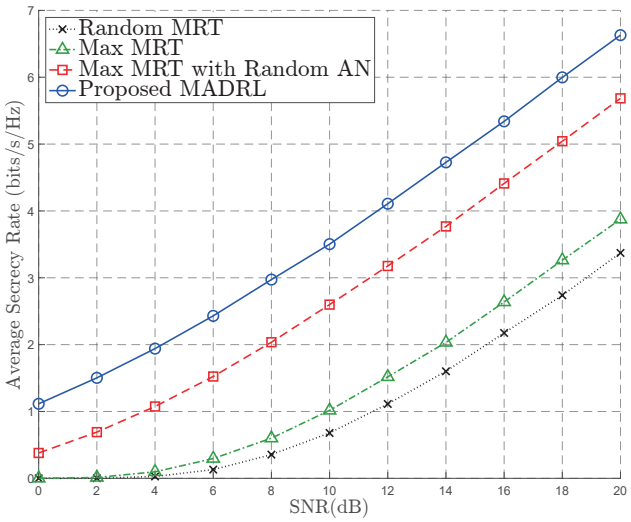

Fig 2. Average secrecy rate versus SNR

ACKNOWLEDGMENT

This work was partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2024-00444170, Research and international collaboration on trust model-based intelligent incident response technologies in 6G open network environment, 50%) and IITP-ITRC(Information Technology Research Center) grant funded by the Korea government(MSIT)(IITP-2025-RS-2024-00437886, 50%).

References

[1] S. salim, N. Moustafa, M. Reisslein, Cybersecurity of satellite communications systems: A comprehensive survey of the space, ground, and links segments, IEEE Communications Surveys & Tutorials (2024).

[2] B. Li, Z. Fei, C. Zhou, Y. Zhang, Physical-layer security in space information networks: A survey, IEEE Internet of things journal 7 (1) (2019) 33-52.

[3] S. Goel, R. Negi, Guaranteeing secrecy using artificial noise, IEEE transactions on wireless communications 7 (6) (2008) 2180-2189.

[4] M. M. Razaq, Y. Jiao, L. Peng, P.-H. Ho, Deep reinforcement learning-based physical layer security framework for internet of medical things, IEEE Transactions on Consumer Electronics (2024).