SMAB-SR: A Sleeping Multi-Armed Bandit Framework for Secure Routing in Space-Air-Ground Integrated Networks

Yang Xu, *Senior Member, IEEE,* Xiaoyuan Liu, Jia Liu, *Senior Member, IEEE,*
Tarik Taleb, *Senior Member, IEEE,* Yusheng Ji, *Fellow, IEEE,* and Norio Shiratori, *Life Fellow, IEEE*

***Abstract*—Space-air-ground integrated network (SAGIN) represents a pivotal architecture for the future evolution of global mobile communications. However, its inherent high dynamics and stochastic nature pose significant challenges to conventional routing mechanisms. Moreover, the vast spatial-scale openness of SAGIN makes it particularly vulnerable to eavesdropping attacks. This paper presents a novel sleeping multi-armed bandit (SMAB) framework, designed to enable secure routing in SAGIN. Specifically, we first establish channel models for all types of links in SAGIN. Then, we theoretically analyze the statistical properties of secrecy capacity and end-to-end (E2E) delay for message transmission over arbitrary routes, and formulate the secure routing problem to maximize cumulative secure transmission throughput under the delay constraint. The uncertainty in the network state of SAGIN, along with the complexity of the optimization objective and constraint (non-convex, nonlinear, and coupled), renders the solution to the secure routing problem highly intractable. To this end, we leverage the MAB model to transform the secure routing problem into a budget-constrained arm-pulling problem and introduce the "sleeping" mode to capture route unavailability due to intermittent link failures. To effectively balance route exploration and exploitation, we further apply the upper confidence bound (UCB) method to design the SMAB-based secure routing algorithm (SMAB-SR), and derive its regret upper bound theoretically. Finally, extensive simulations verify that the SMAB-SR algorithm exhibits significant advantages in E2E secure transmission throughput compared to benchmarks and can maintain highly effective across various SAGIN configurations.**

***Index Terms*—Space-air-ground integrated network, secure routing, sleeping multi-armed bandit, regret analysis.**

I. Introduction

This work was supported in part by the National Natural Science Foundation of China under Grant No. 62372361; in part by JSPS KAKENHI Grant Numbers JP25K15087, JP24K02937; in part by the Project of Cyber Security Establishment with Inter-University Cooperation; in part by Zhejiang Provincial Natural Science Foundation of China under Grant No. LY24F020011; in part by the Federal Ministry of Research, Technology and Space (BMFTR), Germany, through the Project 6GEM+ under Grant No. 16KIS2411; and in part by the European Union's HE research and innovation program HORIZON-JUSNS-2023 through the 6G-Path project under Grant No. 101139172. *(Corresponding author: Jia Liu.)*

Yang Xu and Xiaoyuan Liu are with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China (e-mail: yxu@xidian.edu.cn; xiaoyuanliu@stu.xidian.edu.cn).

Jia Liu is with the Center for Strategic Cyber Resilience Research and Development, National Institute of Informatics, Tokyo 101-8430, Japan (e-mail: jliu@nii.ac.jp).

Tarik Taleb is with the Faculty of Electrical Engineering and Information Technology of the Ruhr University Bochum, Bochum, Germany (e-mail: tarik.taleb@rub.de).

Yusheng Ji is with the National Institute of Informatics, Tokyo 101-8430, Japan (e-mail: kei@nii.ac.jp).

Norio Shiratori with the Research and Development Initiative, Chuo University, Tokyo 112-8551, Japan (e-mail: norio.shiratori.e8@tohoku.ac.jp).

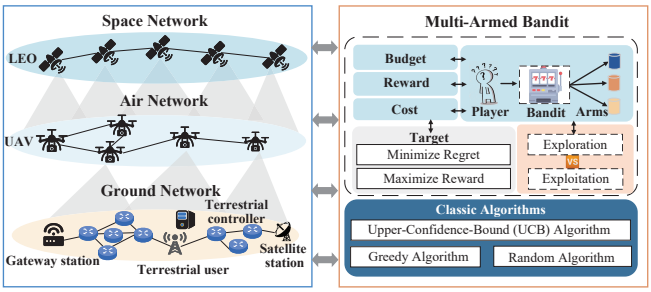


Fig. 1. SAGIN scenario and MAB principles.

With the rapid advancement of global informatization, conventional terrestrial communication networks are encountering critical bottlenecks in spatial coverage and quality-of-service (QoS) guarantees. Existing base stations, constrained by wired backhaul and fixed deployment paradigms, can support basic communication in urban regions but fail to adequately resolve difficulties such as signal voids in remote areas, transmission attenuation across complex terrains, and delayed emergency responses during disasters [1]. In this context, the space-air-ground integrated network (SAGIN) has emerged, which establishes a three-dimensional communication system through the deep integration of satellite constellations, high-altitude platforms (e.g., UAVs and airships), and terrestrial base stations (as illustrated in Fig. 1) [2]. SAGIN aims to achieve globally ubiquitous coverage while delivering low-latency and high-reliability services, positioning itself as the cornerstone architecture for mobile communication evolution in the 6G era [3].

However, the heterogeneity and dynamic nature of SAGIN pose significant challenges to its network operation, management, and optimization [4]. The SAGIN integrates heterogeneous communication nodes from different layers, with substantial disparities across layers in transmission bandwidth, latency characteristics, link stability, and so on. Moreover, the high-speed movement of low-Earth-orbit (LEO) satellites and the dynamic deployment of aerial platforms induce spatiotemporal uncertainties in network topology. Under these circumstances, conventional routing mechanisms based on fixed or quasi-static topology assumptions struggle to adapt to real-time variations in SAGIN's cross-domain link states and topological configurations [5]. Consequently, designing suitable routing strategies to ensure end-to-end (E2E) transmission efficiency and QoS in complex time-varying environments has become the critical challenge for achieving ubiquitous and reliable communication through SAGIN.

In recent years, preliminary studies have been conducted on routing mechanism design within the SAGIN architecture.

Tang *et al.* [6] proposed a reinforcement learning (RL)-based routing optimization strategy for SAGIN, which combines a double Q-learning algorithm with an improved delay-sensitive replay memory mechanism to facilitate routing decisions through historical routing information. Chen *et al.* [7] introduced software-defined networking into SAGIN and optimized the number of controllers and the positions of switch nodes in routing decisions using a k-means clustering algorithm and a multi-objective optimization model based on a genetic algorithm, effectively reducing routing delay and improving load balancing. Wei *et al.* [8] proposed an intelligent and reliable routing scheme based on deep RL to improve data transmission efficiency and reliability in LEO satellite networks. This scheme automatically selects routes and gradually increases the scenario complexity during training, showing strong adaptability and reliability in environments with frequent dynamic topology changes. Zhang *et al.* [9] developed a control domain partitioning-based adaptive routing mechanism for LEO satellite networks to mitigate topology dynamics and routing overhead challenges. The mechanism employs a non-dominated sorting framework for domain management to reduce control latency and implements localized link-status awareness across adjacent domains, enabling dynamic route optimization without global network updates.

A. Motivations and Contributions

The aforementioned studies represent significant progress in routing research for SAGIN, making notable contributions by enhancing network throughput, reliability, and operational efficiency. However, we note that existing routing mechanisms predominantly rely on predefined network state information, whereas the dual uncertainties inherent to SAGIN—stemming from dynamic topological variations and time-varying channel conditions—strictly constrain the real-time route quality assessment. This fundamental limitation inevitably impacts the optimality of routing decision-making. Consequently, how to dynamically adjust routing strategies in unknown, random, and time-varying environments constitutes a critical consideration for the routing design applicable to the SAGIN architecture. Furthermore, the open nature of wireless channels across SAGIN's vast spatial-scale coverage exposes message transmissions to potential unauthorized eavesdropping by multiple malicious users or platforms, resulting in exacerbated cumulative information leakage risks throughout E2E communication routes. Therefore, ensuring protection against malicious eavesdropping while maintaining reliable and efficient message transmission emerges as another pivotal concern in SAGIN routing mechanism design. Currently, secure routing design for SAGIN remains largely underexplored, and effective solutions have yet to be established.

Motivated by the preceding discussions, this paper proposes SMAB-SR, a sleeping multi-armed bandit framework, aiming to provide a secure routing solution for SAGIN. Our objective is to maximize the transmission throughput within a specified time while ensuring that messages are not intercepted by malicious eavesdroppers. Given that the network state is time-varying and uncertain, our approach follows an online decision-making principle, where routing decisions are continuously updated based on the statistical properties of route quality in terms of secure capacity and E2E delay, as well as the historical experiences, thus dynamically selecting the most suitable route for message transmission in real-time. We apply the multi-armed bandit (MAB) model to transform the routing problem into an arm-pulling problem. As a form of reinforcement learning, MAB, as illustrated in Fig. 1, involves a player facing a row of slot machines ("arms"). Pulling each arm incurs a cost and receives a reward, which correspond to the delay consumed over the route and the achievable secrecy throughput the route can provide, respectively. To solve the MAB problem, we design the SMAB-SR algorithm, which approximates the optimal arm selection in each round. Furthermore, we theoretically analyze the cumulative performance gap between the SMAB-SR algorithm and the ideal optimal strategy.

In summary, this paper makes the following contributions:

- *Novel Secure Routing Framework*: To the best of our knowledge, this is the first work that investigates the secure routing design in SAGIN to counteract eavesdropping attacks during E2E message transmission[1]. We propose a novel SMAB-based framework that dynamically updates route selection in real-time under network state uncertainties, aiming to optimize the cumulative secure transmission throughput while satisfying the delay constraint.
- *Route Quality Assessment*: We establish channel models for all types of links in SAGIN. Building on these models, we theoretically characterize the statistical properties of the secrecy capacity and E2E delay for message transmission over arbitrary routes. This analysis entails rigorous and intricate mathematical derivations encompassing integral operations and transcendental function manipulations.
- *SMAB-SR Algorithm and Regret Analysis*: We transform the secure routing problem into an MAB problem and introduce the "sleeping" mode to capture route unavailability caused by intermittent link failures. Leveraging the Upper Confidence Bound (UCB) strategy, we design the SMAB-SR algorithm to achieve an effective balance between route exploration and exploitation. Theoretically, we prove an upper bound on the SMAB-SR algorithm's regret, ensuring its long-term performance approaches that of the ideal optimal routing strategy.
- *Extensive Simulations and Comparisons*: We conduct comprehensive simulations across various SAGIN configurations. The results reveal that the SMAB-SR algorithm significantly outperforms benchmarks while approaching the performance of the ideal optimal algorithm, demonstrating its effectiveness and stability as a robust solution for enabling E2E secure transmission in SAGIN environments.

[1]Note that the security issue addressed in this paper specifically concerns routing security against eavesdropping on wireless message transmission, and does not cover other aspects such as authorization, authentication, encryption, and integrity.

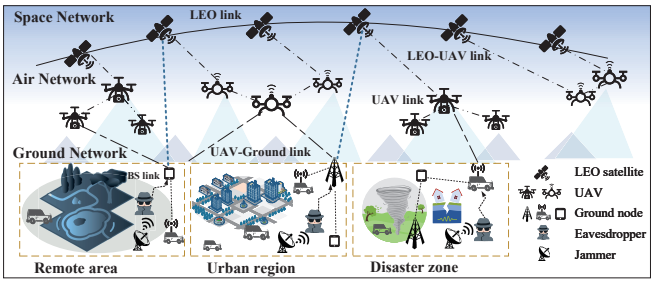


Fig. 2. Network model of SAGIN.

B. Paper Organization

The remainder of this paper is organized as follows. Section II presents the system model. Section III conducts theoretical modeling for route performance and formulates the secure routing problem. The sleeping MAB-based secure routing framework is elaborated in Section IV, while its regret bound and computational complexity are theoretically analyzed in Section V. The simulation results are provided in Section VI. We introduce the literature review in Section VII, followed by the conclusion in Section VIII.

II. System Model

In this section, we introduce the network model and channel model involved in this study.

A. Network Model

As illustrated in Fig. 2, we consider a generally configured SAGIN scenario, which integrates the space, air, and ground segments into a unified system. In the space segment, there is a group of low Earth orbit (LEO) satellites, denoted by set $\mathcal{S} = \{s_1, s_2, \ldots, s_{N_S}\}$, which are capable of covering the entire communication area of interest. In the air segment, there are multiple communication nodes such as unmanned aerial vehicles (UAVs), aerial platforms, etc. For simplicity and without loss of generality, we consider only UAVs, denoted by set $\mathcal{U} = \{u_1, u_2, \ldots, u_{N_U}\}$, which can be deployed in urban areas, remote regions, or disaster zones according to practical demands. In the ground segment, there are various types of communication devices, such as base stations, static or mobile access points, and user terminals, which we collectively refer to as ground nodes (GNs), represented by set $\mathcal{G} = \{g_1, g_2, \ldots, g_{N_G}\}$.

We investigate the end-to-end (E2E) transmission of messages in the aforementioned SAGIN scenario, that is, the message originates from a source GN, then is forwarded in a multi-hop manner across the SAGIN, and ultimately reaches another destination GN. We focus on a general pair of source and destination GNs, denoted by S and D, respectively, where $S, D \in \mathcal{G}$. We use the notation $\pi = \langle l_1, \ldots, l_k, \ldots, l_K \rangle$ to denote a K-hop route (path) from S to D, where an intermediate link l_k on route π represents the k-th hop from node S_k to the next-hop node D_k, i.e., $S_1 = S$, $S_k = D_{k-1}$, $D_K = D$, and $S_k, D_k \in \mathcal{S} \cup \mathcal{U} \cup \mathcal{G}$. The universal set of all candidate routes prescribed in the SAGIN is denoted as $\mathcal{R}^{\text{univ}}$.

Note that due to the dynamic nature of SAGIN, any link is intermittently connected (the connectivity usually depends on the signal-to-noise ratio requirement at the receiver), leading to the availability of all routes varying dynamically over time. We divide the timeline into time slots and ensure that an entire E2E routing process can be accomplished within a single time slot. The time slot is indexed by $\theta = 1, 2, \cdots$, and the network topology and channel states remain relatively static within a time slot but vary between successive slots. In a given time slot θ, multiple available candidate routes may exist between S and D, and their set is denoted as $\mathcal{R}(\theta) \subseteq \mathcal{R}^{\text{univ}}$. For clarity in subsequent presentations, slot-dependent symbols can be used interchangeably with their counterparts that omit the time slot index, provided that no ambiguity arises. For example, $\mathcal{R}(\theta)$ and $\mathcal{R}$ are used interchangeably. At the beginning of each time slot, S can choose one route $\pi \in \mathcal{R}$ according to some strategy, i.e., corresponding to a route selection round, and then implement an E2E message delivery over the selected route in this time slot.

In addition, we consider the presence of a group of malicious ground eavesdroppers who attempt to intercept communication on legitimate links. To counteract eavesdropping attacks, a group of friendly jammers is introduced to assist the transmission node by generating artificial interference to disrupt the eavesdropper's message reception. We assume that the locations of the eavesdroppers and the jammers follow an independent and homogeneous Poisson point process (PPP)[2], denoted by $\mathcal{E} = \{1, 2 \cdots\}$ and $\mathcal{J} = \{1, 2 \cdots\}$, with densities λ_E and λ_J, respectively.

B. Channel Model

We adopt the general assumption that the network is interference-limited [11]–[13]. When both the transmitter T_x and the receiver R_x are GNs, the large-scale path loss, along with the small-scale Rayleigh fading, are considered. h_{T_x,R_x} represents the channel gain between T_x and R_x at time slot θ, and $|h_{T_x,R_x}|^2$ is exponentially distributed with $\mathbb{E}\left\{|h_{T_x,R_x}|^2\right\} = 1$. d_{T_x,R_x} denotes the Euclidean distance between T_x and R_x, and α is the path loss exponent, typically ranging from 2 to 6. When either the transmitter T_x or the receiver R_x is a GN while the other is a UAV, $P^{loss}_{T_x,R_x}$ denotes the path loss between T_x and R_x at slot θ. According to [14], $P^{loss}_{T_x,R_x}$ can be formulated as

$$P^{loss}_{T_x,R_x} = 10\alpha \log_2(d_{T_x,R_x}) + C_{f,Tx} + \sigma_{T_x,R_x}, \quad (1)$$

where C_{f,T_x} is a constant depending on the operating frequency and antenna gain of the transmitter T_x, and σ_{T_x,R_x} is a Gaussian random variable. When either the transmitter or the receiver is a satellite, the Weibull-based channel model [15] is applied, and the channel gain G_{T_x,R_x} at slot θ is expressed as

$$G_{T_x,R_x} = \frac{G_{T_x} G_{R_x} \lambda^2}{4\pi d^2_{T_x,R_x}} 10^{-\frac{F^{rain}_{T_x,R_x}}{10}}, \quad (2)$$

where G_{T_x} and G_{R_x} denote the antenna gain of the transmitter and receiver, respectively, λ is the wavelength of the propagation signal, and $F^{rain}_{T_x,R_x}$ represents the rain attenuation, which follows the Weibull distribution.

[2]PPP is the most widely used spatial distribution model in stochastic geometry analysis. It can reasonably represent the uncertainty of network node locations while offering high analytical tractability [10].

Due to the numerous types of links involved in SAGIN, we categorize them based on the segment where the transmission node S_k is located and present the corresponding channel models as follows.

1) $S_k \in \mathcal{G}$: According to Shannon's Theorem, the achievable rate C_{S_k,D_k} from S_k to D_k at slot θ is given by

$$C_{S_k,D_k}=\begin{cases}\log_2\left(1+\frac{P_{S_k}\left|h_{S_k,D_k}\right|^2/d^{\alpha}_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\right), & D_k\in\mathcal{G}\\ \log_2\left(1+\frac{P_{S_k}10^{\frac{P^{loss}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J10^{\frac{P^{loss}_{J_j,D_k}}{10}}}\right), & D_k\in\mathcal{U}\\ \log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_JG_{J_j,D_k}}\right), & D_k\in\mathcal{S}\end{cases} \quad (3)$$

where P_{S_k} denotes the transmitting power of S_k and P_J denotes the jamming power of a jammer.

For an eavesdropper E_i, its achievable eavesdropping rate C_{S_k,E_i} is determined as

$$C_{S_k,E_i}=\log_2\left(1+\frac{P_{S_k}\left|h_{S_k,E_i}\right|^2/d^{\alpha}_{S_k,E_i}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,E_i}\right|^2/d^{\alpha}_{J_j,E_i}}\right). \quad (4)$$

2) $S_k \in \mathcal{U}$: The achievable rate C_{S_k,D_k} from S_k to D_k at slot θ is given by

$$C_{S_k,D_k}=\begin{cases}\log_2\left(1+\frac{P_{S_k}10^{\frac{P^{loss}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\right), & D_k\in\mathcal{G}\\ \log_2\left(1+\frac{P_{S_k}10^{\frac{P^{loss}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J10^{\frac{P^{loss}_{J_j,D_k}}{10}}}\right), & D_k\in\mathcal{U}\\ \log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_JG_{J_j,D_k}}\right). & D_k\in\mathcal{S}.\end{cases} \quad (5)$$

For an eavesdropper E_i, its achievable eavesdropping rate C_{S_k,E_i} is determined as

$$C_{S_k,E_i}=\log_2\left(1+\frac{P_{S_k}10^{\frac{P^{loss}_{S_k,E_i}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,E_i}\right|^2/d^{\alpha}_{J_j,E_i}}\right). \quad (6)$$

3) $S_k \in \mathcal{S}$: The achievable rate C_{S_k,D_k} from S_k to D_k at slot θ is given by

$$C_{S_k,D_k}=\begin{cases}\log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\right), & D_k\in\mathcal{G}\\ \log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J10^{\frac{P^{loss}_{S_k,D_k}}{10}}}\right), & D_k\in\mathcal{U}\\ \log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_JG_{J_j,D_k}}\right). & D_k\in\mathcal{S}\end{cases} \quad (7)$$

For an eavesdropper E_i, its achievable eavesdropping rate C_{S_k,E_i} is determined as

$$C_{S_k,E_i}=\log_2\left(1+\frac{P_{S_k}G_{S_k,E_i}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,E_i}\right|^2/d^{\alpha}_{J_j,E_i}}\right). \quad (8)$$

III. Route Performance Modeling and Secure Routing Problem Formulation

In this section, we provide the performance modeling for a general route in the considered SAGIN, in terms of secrecy capacity and E2E delay, and subsequently formulate the secure routing problem.

A. Secrecy Capacity

We adopt secrecy capacity (SC) as the key metric for evaluating message transmission security. SC represents the fundamental limit of secure communication in the presence of eavesdroppers, it is the maximum achievable data transmission rate at which messages can be reliably delivered to the intended receiver while ensuring that no meaningful information can be inferred by the eavesdropper. Let Z_{l_k} denote the SC of link l_k at time slot θ, then, from Wyner's result [16], Z_{l_k} can be determined as

$$Z_{l_k}=\left[C_{S_k,D_k}-\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}\right]^+\triangleq\left[C_{S_k,D_k}-C^{\max}_{S_k,E_i}\right]^+, \quad (9)$$

where $[x]^+=\max\{0,x\}$.

From Eq. (9), it can be observed that evaluating the SC of link l_k requires knowledge of C_{S_k,D_k} and $C^{\max}_{S_k,E_i}$. However, due to the dynamic nature of SAGIN, obtaining their instantaneous values is challenging. To this end, we provide the following lemmas to capture their statistical characteristics. Let $F_{C_{S_k,D_k}}(x)$ and $\mathbb{E}\{C_{S_k,D_k}\}$ denote the cumulative distribution function (CDF) and expectation of the legitimate link capacity, $F_{C^{\max}_{S_k,E_i}}(x)$ and $\mathbb{E}\{C^{\max}_{S_k,E_i}\}$ denote the CDF and expectation of the eavesdropping link capacity, respectively.

Lemma 1: When $S_k \in \mathcal{G}$, $F_{C_{S_k,D_k}}(x)$, $\mathbb{E}\{C_{S_k,D_k}\}$, $F_{C^{\max}_{S_k,E_i}}(x)$ and $\mathbb{E}\{C^{\max}_{S_k,E_i}\}$ are determined as Eqs. (10)~(13), respectively, shown at the top of the next page, where $\Gamma(\cdot)$ is the Gamma function.

Lemma 2: When $S_k \in \mathcal{U}$, $F_{C_{S_k,D_k}}(x)$, $\mathbb{E}\{C_{S_k,D_k}\}$, $F_{C^{\max}_{S_k,E_i}}(x)$ and $\mathbb{E}\{C^{\max}_{S_k,E_i}\}$ are determined as Eqs. (14)~(17), respectively, shown at the top of Page 6.

Lemma 3: When $S_k \in \mathcal{S}$, $F_{C_{S_k,D_k}}(x)$, $\mathbb{E}\{C_{S_k,D_k}\}$, $F_{C^{\max}_{S_k,E_i}}(x)$ and $\mathbb{E}\{C^{\max}_{S_k,E_i}\}$ are determined as Eqs. (18)~(21), respectively, shown at the top of Page 7.

Proof of Lemmas 1, 2, and 3: Please kindly refer to the supplementary material. ∎

Assume that in time slot θ, corresponding to the θ-th round of route selection, the E2E message transmission is conducted over route π. The SC Z_π of route π is determined by the minimum SC among all its links, i.e., there is

$$Z_\pi=\min_{l_k\in\pi}Z_{l_k}. \quad (22)$$

B. End-to-End Delay

Let W_{l_k} denote the delay experienced by the message on link l_k, which can be estimated as

$$W_{l_k}=\frac{d_{S_k,D_k}}{v}+\frac{\eta}{C_{S_k,D_k}}+W_{Q,l_k}. \quad (23)$$

The first term on the right-hand side (RHS) of Eq. (23) is the prorogation delay, where v denotes the signal propagation speed. The second term on the RHS of Eq. (23) is the transmission delay, where η indicates the length of the message. W_{Q,l_k} represents the queuing delay, which we estimate using the Kingman's formula [17]:

$$W_{Q,l_k}\approx\frac{\rho(c_a^2+c_s^2)}{2(1-\rho)\mu}, \quad (24)$$

$$F_{C_{S_k,D_k}}(x) \triangleq \mathbb{P}(C_{S_k,D_k} \le x) = \begin{cases} 1 - \exp\left[-2\pi\lambda_J \left(\frac{P_{S_k}/d_{S_k,D_k}^{\alpha}}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}} \Gamma(1-\frac{2}{\alpha})\Gamma(1+\frac{2}{\alpha})\right], & D_k \in \mathcal{G} \\ \frac{10}{\sqrt{\pi}}\left\{\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log_2(d_{S_k,D_k}) - \frac{C_{f,S_k}}{10} + \log_{10}\left[\frac{P_J}{P_{S_k}}(2^x-1)\right]\right\}, & D_k \in \mathcal{U} \\ \frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J d_{S_k,D_k}^2}{P_{S_k}} - 10 - 10\pi\lambda_J\Gamma(\frac{1}{2})\right]. & D_k \in \mathcal{S} \end{cases} \tag{10}$$

$$\mathbb{E}\{C_{S_k,D_k}\} = \begin{cases} \frac{1}{2\pi\lambda_J\left(\frac{P_{S_k}}{d_{S_k,D_k}^{\alpha} P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\ln 2}, & D_k \in \mathcal{G} \\ \frac{\sqrt{\pi}}{10\ln 2}\left[\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log_2(d_{S_k,D_k}) - \frac{C_{f,S_k}}{10} + \log_{10}\left(\frac{P_J}{P_{S_k}}\right)\right]^{-1}, & D_k \in \mathcal{U} \\ \frac{\ln 10}{10\ln 2}\cdot\frac{P_{S_k}}{P_J d_{S_k,D_k}^2}\cdot\exp\left(10 + 10\pi\lambda_J\Gamma\left(\frac{1}{2}\right)\right). & D_k \in \mathcal{S} \end{cases} \tag{11}$$

$$F_{C^{\max}_{S_k,E_i}}(x) \triangleq \mathbb{P}(C^{\max}_{S_k,E_i} \le x) = \exp\left[-\frac{\lambda_E}{\lambda_J(\frac{P_{S_k}}{(2^x-1)P_J})^{-\frac{2}{\alpha}}\Gamma(1-\frac{2}{\alpha})\Gamma(1+\frac{2}{\alpha})}\right] \tag{12}$$

$$\mathbb{E}\{C^{\max}_{S_k,E_i}\} = \frac{1}{\frac{\lambda_E}{\lambda_J}\left(\frac{P_{S_k}}{P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\ln 2} \tag{13}$$

where ρ is the utilization (ratio of arrival rate to service rate), μ is the service rate, and c_a and c_s are the coefficients of variation of the arrival time interval and the service time, respectively.

For the CDF of W_{l_k}, we have

$$\begin{aligned} F_{W_{l_k}}(x) &\triangleq \mathbb{P}(W_{l_k} \le x) \\ &= \mathbb{P}\left(\frac{d_{S_k,D_k}}{v} + \frac{\eta}{C_{S_k,D_k}} + W_{Q,l_k} \le x\right) \quad (25) \\ &= 1 - F_{C_{S_k,D_k}}\left(\frac{\eta}{x - \frac{d_{S_k,D_k}}{v} - W_{Q,l_k}}\right). \quad (26) \end{aligned}$$

Assuming that route π is selected in time slot θ, the E2E delay of the message W_π can be determined as

$$W_\pi = \sum_{l_k \in \pi} W_{l_k}. \tag{27}$$

C. Secure Routing Problem Formulation

For the SAGIN considered in this paper, our objective is to transmit as many messages as possible within a specified time while ensuring that the messages are not intercepted by malicious eavesdroppers. To achieve this goal, we need to design a routing strategy so that an appropriate route for E2E message transmission can be identified in each round. Therefore, we formulate the secure routing problem in SAGIN as the following optimization problem:

$$\max_{\pi\in\mathcal{R}(\theta)} \quad \sum_{\theta=1}^{\Theta(W_B)} \mathbb{E}\{Z_\pi(\theta)\} \tag{28a}$$

$$\text{s.t.} \quad \sum_{\theta=1}^{\Theta(W_B)} \mathbb{E}\{W_\pi(\theta)\} \le W_B. \tag{28b}$$

Eq. (28b) represents the constraint on the cumulative E2E delay, where W_B denotes the specified upper bound for the cumulative E2E delay, and $\Theta(W_B)$ denotes the maximum number of rounds that the E2E message transmission can be performed under given W_B. Eq. (28a) indicates the objective of maximizing the cumulative E2E secrecy capacity, by optimizing the selected route π in each round θ from the corresponding feasible route set $\mathcal{R}(\theta)$.

IV. Sleeping Multi-Armed Bandits-based Routing Solution

A. Conversion to MAB Problem

There are several key challenges in solving the secure routing problem (28): The inherent uncertainties in SAGIN (e.g., instantaneous channel state information, locations of eavesdroppers and jammers) and its dynamic, time-varying characteristics (e.g., network topology, channel conditions) result in the unavailability of complete information about the whole network state during each round of route selection, thereby making global optimization infeasible; From an optimization perspective, the expressions of both the objective function and the constraint in (28) are highly complex (non-convex, non-linear, etc.) and involve multiple cases, rendering classical optimization methods ineffective for solving the problem. The strong coupling between the constraint and the objective further complicates the optimization problem, that is, (28b) determines the maximum number of implemented routing rounds $\Theta(W_B)$, and $\Theta(W_B)$ is also the number of summation terms in (28a).

Conventional routing schemes are incapable of addressing the aforementioned challenges. To this end, we develop a multi-armed bandit (MAB)-based online reinforcement learning solution. The MAB problem is a foundational framework for decision-making under uncertainty. It is inspired by the analogy of a gambler at a casino facing a row of slot machines

$$F_{C_{S_k,D_k}}(x) \triangleq \mathbb{P}(C_{S_k,D_k} \le x) = \begin{cases} \frac{10}{\sqrt{\pi}}\left[e^{(-\pi\lambda_J)}-1\right] + \frac{20}{\sqrt{\pi}}\log_{10}\left(\frac{(2^x-1)P_J}{P_{S_k}10^{\alpha\log_2(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right), & D_k \in \mathcal{G} \\ \frac{10}{\sqrt{\pi}}\left\{\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log_2(d_{S_k,D_k}) - \frac{C_{f,S_k}}{10} + \log_{10}\left[\frac{P_J}{P_{S_k}}(2^x-1)\right]\right\}, & D_k \in \mathcal{U} \\ \frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J d^2_{S_k,D_k}}{P_{S_k}} - 10 - 10\pi\lambda_J\Gamma(\frac{1}{2})\right]. & D_k \in \mathcal{S} \end{cases} \tag{14}$$

$$\mathbb{E}[C_{S_k,D_k}] = \begin{cases} \frac{\sqrt{\pi}}{20\ln 2}\left(\frac{10}{\sqrt{\pi}}\left(e^{-\pi\lambda_J}-1\right) + \log_{10}\left(\frac{P_J}{P_{S_k}10^{\alpha\log_2(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)^{-1}, & D_k \in \mathcal{G} \\ \frac{\sqrt{\pi}}{10\ln 2}\left(\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log_2(d_{S_k,D_k}) - \frac{C_{f,S_k}}{10} + \log_{10}\left(\frac{P_J}{P_{S_k}}\right)\right)^{-1}, & D_k \in \mathcal{U} \\ \frac{\ln 10}{10\ln 2}\cdot\frac{P_{S_k}}{P_J d^2_{S_k,D_k}}\cdot\exp\left(10 + 10\pi\lambda_J\Gamma\left(\frac{1}{2}\right)\right). & D_k \in \mathcal{S} \end{cases} \tag{15}$$

$$F_{C^{\max}_{S_k,E_i}}(x) \triangleq \mathbb{P}(C^{\max}_{S_k,E_i} \le x) = \exp\left\{-\lambda_E + \frac{10}{\sqrt{\pi}}\lambda_E e^{-\pi\lambda_J} - \frac{10}{\sqrt{\pi}}\lambda_E - \frac{20}{\sqrt{\pi}}\lambda_E\log_{10}\left(\frac{(2^x-1)P_J}{P_{S_k}10^{\alpha\log_2(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right\} \tag{16}$$

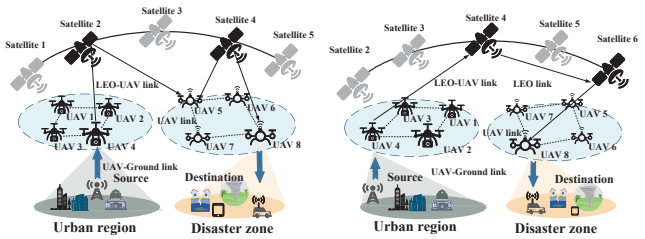
$$\mathbb{E}\{C^{\max}_{S_k,E_i}\} = \frac{\sqrt{\pi}}{20\lambda_E\ln 2}\left[-\lambda_E + \frac{10}{\sqrt{\pi}}\lambda_E e^{-\pi\lambda_J} - \frac{10}{\sqrt{\pi}}\lambda_E + \log_{10}\left(\frac{P_J}{P_{S_k}10^{\alpha\log_2(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right]^{-1} \tag{17}$$

("arms"), each with a potentially different and unknown probability of reward, and the gambler must decide which arm to pull over a series of rounds to maximize cumulative reward.

We convert SAGIN's secure routing problem (28) into an MAB problem as follows: We regard the source node as the player (i.e., gambler), and each feasible candidate route as an arm. Selecting a route corresponds to pulling an arm. The achievable secrecy capacity of the selected route and the incurred E2E delay serve as the associated reward and cost of pulling the arm, respectively. As long as the cumulative cost does not exceed the given budget W_B, additional rounds of arm pulling (route selection) can be performed. This process continues until the maximum allowable rounds $\Theta(W_B)$ are reached, at which the budget is exhausted and the game ends. In the standard formalization of the MAB problem, regret is a core concept, which measures the difference between the reward obtained by pulling the arm following a given strategy and the reward that could have been achieved by consistently pulling the best arm.

Definition 1 (Regret): Regret is defined as the difference between the total achieved reward of the optimal policy that has knowledge of the complete instantaneous network state in advance, and the reward achieved by the applied strategy in the case where such knowledge is unavailable. Let $\mathrm{Reg}(W_B)$ denote the regret under the budget W_B, R^*_{total} and R_{total} represent the total reward achieved by the optimal policy and the applied strategy, respectively. Then, $\mathrm{Reg}(W_B)$ is expressed as

$$\begin{aligned} \mathrm{Reg}(W_B) &= R^*_{total} - R_{total} \\ &= \sum_{\theta=1}^{\Theta^*(W_B)} Z_{\pi^*}(\theta) - \sum_{\theta=1}^{\Theta(W_B)} Z_{\pi}(\theta), \end{aligned} \tag{29}$$



(a) An available route in time slot θ. (b) An available route in time slot $\theta+1$.

Fig. 3. Dynamic variation of the available route over time.

where $*$ indicates the selection under the optimal policy.

Based on Definition 1, secure routing problem (28) can be converted into the following normal-form of an MAB problem:

$$\text{MAB:} \quad \min_{\pi\in\mathcal{R}(\theta)} \quad \mathrm{Reg}(W_B) \tag{30a}$$

$$\text{s.t.} \quad \sum_{\theta=1}^{\Theta(W_B)} \mathbb{E}\left\{W_\pi(\theta)\right\} \le W_B. \tag{30b}$$

B. SMAB-SR Algorithm

Note that there is a key difference between Problem (30) and a standard MAB problem. In a standard MAB problem, the set of arms that can be selected remains fixed in each round. However, due to the high dynamics of SAGIN, communication links may intermittently break, causing the set of available routes to change dynamically over time. Fig. 3 illustrates an example of dynamic variation of the available route between a source-destination pair. Due to changes in the positions of satellites and UAVs, some links used in time slot θ become interrupted in time slot $\theta+1$, leading to a change in the

$$F_{C_{S_k,D_k}}(x) \triangleq \mathbb{P}(C_{S_k,D_k} \le x) = \begin{cases} \exp\left[-10 - \frac{80\pi^2\lambda_J d_{S_k,D_k}^2 (2^x-1)P_J}{P_{S_k}G^2\lambda^2(1-\alpha)}\Gamma(1-\alpha)\right], & D_k \in \mathcal{G} \\ \exp\left[-2\pi\lambda_J \int_0^\infty \left(1-\exp\left(10^{\alpha \log_2 r}\right)\right) r dr\right] \cdot \exp\left[10\log_{10} \frac{(2^x-1)P_J 10^{\frac{C_{f,S_k}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d_{S_k,D_k}^2}}\right], & D_k \in \mathcal{U} \\ \mathbb{E}_{\mathcal{J}}\left\{\exp\left(10\log_{10}\sum_{J_j\in\mathcal{J}} \frac{1}{d_{J_j,D_k}^2}\right)\right\} \cdot \frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right]. & D_k \in \mathcal{S} \end{cases} \tag{18}$$

$$\mathbb{E}[C_{S_k,D_k}] = \begin{cases} \frac{P_{S_k}G^2\lambda^2(1-\alpha)\Gamma(1-\alpha)}{80\pi^2\lambda_J d_{S_k,D_k}^2 \ln 2 P_J}, & D_k \in \mathcal{G} \\ \frac{\ln 10}{10\ln 2} \cdot \frac{P_{S_k}G^2\lambda^2}{4\pi d_{S_k,D_k}^2 P_J 10^{\frac{C_{f,S_k}}{10}}}, & D_k \in \mathcal{U} \\ = \frac{\ln 10}{10\ln 2} \cdot \frac{P_{S_k}}{P_J d_{S_k,D_k}^2} \cdot \frac{2}{\mathbb{E}_{\mathcal{J}}\left\{\exp\left[10\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d_{J_j,D_k}^2}\right)\right]\right\}}. & D_k \in \mathcal{S} \end{cases} \tag{19}$$

$$F_{C^{\max}_{S_k,E_i}}(x) \triangleq \mathbb{P}(C^{\max}_{S_k,E_i} \le x) = \exp\left[-10 - \frac{80\pi^2\lambda_J d_{S_k,E_i}^2 (2^x-1)P_J}{P_{S_k}G^2\lambda^2(1-\alpha)}\Gamma(1-\alpha)\right] \tag{20}$$

$$\mathbb{E}\{C^{\max}_{S_k,E_i}\} = \frac{P_{S_k}G^2\lambda^2(1-\alpha)\Gamma(1-\alpha)}{80\pi^2\lambda_J d_{S_k,E_i}^2 \ln 2 P_J} \tag{21}$$

available E2E transmission route. To this end, we introduce a variant of MAB, i.e., the sleeping MAB (SMAB). In SMAB, only a subset of arms (corresponding to the current feasible routes) is available in each round, while the rest of the arms are sleeping. Within the SMAB framework, our objective is to establish a mapping from the set of available arms to the selected arm, thereby approaching the best arm and minimizing the regret.

We propose Algorithm 1 to determine the set of available E2E message transmission routes before each round of route selection. Algorithm 1 consists of two phases: connectivity calculation phase (Steps 1–11) and available route set determination phase (Steps 12–26). In the first phase, we compute the connectivity matrices $C^{G,S}$ and $C^{U,S}$, which represent the link connectivity status between satellites and GNs, as well as between satellites and UAVs, respectively. Step 1 initializes $C^{G,S}$ and $C^{U,S}$ as zero matrices. In Steps 2–6, we iterate over all GNs and satellites to compute the link connectivity value $C^{G,S}_{g,s,\theta}$ at round θ. In Steps 7–11, we iterate over all UAVs and satellites to compute the link connectivity value $C^{U,S}_{u,s,\theta}$ at round θ. In the second phase, we initialize the available route set as an empty set (Step 12), and define $f_\pi \in \{true, false\}$ as the flag indicating whether a candidate route π is available in the current round (Step 13). For each candidate route $\pi = \langle l_1, \ldots, l_k, \ldots, l_K \rangle$, which is composed of K-hop links, we sequentially check the connectivity status of every link it contains (Steps 16-17). If any link is found to be disconnected, the route is immediately marked as unavailable (Step 19), and the checking process is terminated (Step 20). When all K links are connected, we add π to $\mathcal{R}(\theta)$ (Step 24).

In the MAB problem (30), since the information about the reward associated with each arm is unknown, we need to explore pulling arms to gather knowledge, which can be used for subsequent arm selection. "Exploration" means trying different arms to collect information about their reward distributions, while "Exploitation" means pulling the arm believed to have the highest expected reward, based on current knowledge. Excessive exploration will consume significant cost on inferior choices, reducing the cumulative reward and slowing convergence to the optimal arm. On the other hand, excessive exploitation hinders the acquisition of sufficient knowledge about arm quality, potentially trapping the decision-maker in a suboptimal choice while failing to discover better alternatives. Moreover, it limits adaptability to changes in reward distributions in dynamic environments.

In order to develop a desirable strategy that can address the dilemma between exploration and exploitation, we utilize the concept of Upper Confidence Bound (UCB), whose core idea is to always maintain optimism in the face of uncertainty [18]. Specifically, for an arbitrary round θ, we estimate the quality of each available arm based on the knowledge learned from previous rounds. Let $\beta_\pi(\theta)$ and $\bar{Z}_\pi(\theta)$ denote the number of times arm π being selected and the average empirical reward of arm π until round θ, respectively. If arm π is pulled in round θ, the values of β_π and $\bar{Z}_\pi$ are updated as follows:

$$\begin{cases} \beta_\pi(\theta) = \beta_\pi(\theta-1) + 1, \\ \bar{Z}_\pi(\theta) = \dfrac{\bar{Z}_\pi(\theta-1)\cdot\beta_\pi(\theta-1) + Z_\pi(\theta)}{\beta_\pi(\theta-1)+1}. \end{cases} \tag{31}$$

Otherwise, these parameters remain unchanged from the previous round.

Algorithm 1 Available Route Set Determination.

Input: Satellite set $\mathcal{S}$, UAV set $\mathcal{U}$, ground node set $\mathcal{G}$, universal set of all candidate routes $\mathcal{R}^{\text{univ}}$, round θ

Output: Available route set $\mathcal{R}(\theta)$

/ Connectivity Calculation Phase */*

Initialize $C^{G,S}$ and $C^{U,S}$ as zero matrices;
for $g = 1$ to N_G **do**
  for $s = 1$ to N_S **do**
    Evaluate the connectivity status $C^{G,S}_{g,s,\theta}$ between GN g and satellite s;
  end for
end for
for $u = 1$ to N_U **do**
  for $s = 1$ to N_S **do**
    Evaluate the connectivity status $C^{U,S}_{u,s,\theta}$ between UAV u and satellite s.
  end for
end for

/ Available Route Set Determination Phase */*

Initialize $\mathcal{R}(\theta) = \emptyset$;
Define $f_\pi \in \{true, false\}$ as the flag indicating whether the candidate route π is available in the current round;
while $\forall \pi = \langle l_1, \ldots, l_k, \ldots, l_K \rangle \in \mathcal{R}^{\text{univ}}$ **do**
  $f_\pi = true$;
  for $k = 1$ to K **do**
    Check the connectivity of link l_k by referring to the corresponding entries in matrices $C^{G,S}$ and $C^{U,S}$;
    if l_k is disconnected **then**
      Set $f_\pi = false$;
      Break;
    end if
  end for
  if $f_\pi = true$ **then**
    Add π to the available route set $\mathcal{R}(\theta)$;
  end if
end while
return Available route set $\mathcal{R}(\theta)$;

According to the UCB policy of MAB, the UCB-based reward for arm π in round θ, denoted as $\widehat{Z}_\pi(\theta)$, is determined as

$$\widehat{Z}_\pi(\theta) = \bar{Z}_\pi(\theta) + U_\pi(\theta), \tag{32}$$

where $U_\pi(\theta) = \sqrt{\frac{2\ln\theta}{\beta_\pi(\theta)}}$ is the confidence interval of route π's reward. In each round, we prioritize selecting an arm from the set of available arms that has never been chosen before, to guarantee that all candidate arms can be explored. If no such unselected arms remain, we pull the arm with the maximal UCB-based reward. After each round of the routing process, we update the associated parameters of all arms for the selection in the next round.

Integrating the available route set determination and the UCB strategy, we propose the **SMAB**-based **s**ecure **r**outing (SMAB-SR) solution for problem (28) in SAGIN, as summarized in Algorithm 2. Specifically, Step 1 initializes the system parameters, where we use $B(\theta)$ to record the remaining budget after the θ-th round. In the main loop for pulling arms, Step 3 updates the index of the round, and Step 4 identifies the currently available route set $\mathcal{R}(\theta)$ by performing Algorithm 1. In Steps 5–10, if there exist arms in $\mathcal{R}(\theta)$ that have never been pulled before, we randomly pull such an arm and record it as π^θ; else, we sort the arms in $\mathcal{R}(\theta)$ in descending order according to their UCB-based reward and pull the arm π^θ with the maximal UCB reward. Once the routing process is complete, Step 11 calculates the associate reward $Z_{\pi^\theta}(\theta)$ and the cost $W_{\pi^\theta}(\theta)$. Steps 12–13 check whether the cost exceeds the remaining budget, if so, the arm-pulling process will terminate; else, we update the parameters in Steps 15–18, and the loop of the arm-pulling process will continue until the budget exhausts.

Algorithm 2 Sleeping MAB-Based Secure Routing Algorithm for SAGIN.

Input: Total budget W_B, universal set of all candidate routes $\mathcal{R}^{\text{univ}}$

Output: Total reward R_{total}, total cost W_{total}, total number of rounds $\Theta(W_B)$, the arm π^θ pulled in each round

/ Initialization */*

1: Initialize $\theta = 0$, $R_{total} = 0$, $W_{total} = 0$, $\Theta(W_B) = 0$, $B(\theta) = W_B$, and for all $\pi \in \mathcal{R}^{\text{univ}}$, initialize $\beta_\pi(0) = 0$;

/ Main loop for pulling arms */*

2: **while** true **do**
3: $\theta = \theta + 1$;
4: Perform **Algorithm 1** to obtain the currently available route set $\mathcal{R}(\theta)$;
5: **if** There exist $\pi \in \mathcal{R}(\theta)$ such that $\beta_\pi(\theta - 1) = 0$ **then**;
6: Randomly pull such an arm and denote it as π^θ;
7: **else**
8: Sort the arms according to the UCB-based reward $\widehat{Z}_\pi(\theta - 1)$ for $\pi \in \mathcal{R}(\theta)$;
9: Pull the arm π^θ with the maximal UCB-based reward;
10: **end if**
11: Calculate the reward $Z_{\pi^\theta}(\theta)$ and the cost $W_{\pi^\theta}(\theta)$;
12: **if** $W_{\pi^\theta}(\theta) > B(\theta - 1)$ **then**
13: **Break**;
14: **else**
15: $R_{total} = R_{total} + Z_{\pi^\theta}(\theta)$;
16: $W_{total} = W_{total} + W_{\pi^\theta}(\theta)$;
17: $B(\theta) = B(\theta - 1) - W_{\pi^\theta}(\theta)$;
18: Update $\beta_\pi(\theta)$, $\bar{Z}_\pi(\theta)$, and $\widehat{Z}_\pi(\theta)$ according to Eqs. (31) and (32) for all arms;
19: **end if**
20: **end while**
21: $\Theta(W_B) = \theta - 1$;
22: **return** R_{total}, W_{total}, $\Theta(W_B)$, π^θ for $\theta = 1, 2, \cdots, \Theta(W_B)$;

V. Theoretical Analysis for SMAB-SR Algorithm

A. Regret Bound Analysis

In this section, we analyze the regret upper bound of the proposed SMAB-SR algorithm. We index all candidate routes in the universal set $\mathcal{R}^{\text{univ}}$ in descending order based on

the route's actual average reward as $1, 2, \cdots, I$, and define $\Delta_{i,i+1} = Z_i - Z_{i+1}$ for $1 \le i \le I-1$. In addition, we define the minimum and maximum reward differences between all non-optimal arms and the optimal arm as follows:

$$\Delta_{\max} = Z_{\pi^*} - \min_{\pi' \neq \pi^*} Z_{\pi'}, \tag{33}$$

$$\Delta_{\min} = Z_{\pi^*} - \max_{\pi' \neq \pi^*} Z_{\pi'}. \tag{34}$$

Let $W_{\min}$ and $W_{\max}$ represent the smallest and largest cost for pulling an arm, respectively, and W^* represent the cost of the optimal arm. Then, we provide the following lemmas.

Lemma 4: With the definition of confidence interval $U_i(\theta) = \sqrt{\frac{2\ln\theta}{\beta_i(\theta)}}$, the following inequality holds for all $1 \le i \le I$ and $\theta \ge 1$:

$$\mathbb{P}\left\{\bar{Z}_i(\theta) \in [Z_i(\theta) - U_i(\theta), Z_i(\theta) + U_i(\theta)]\right\} \ge 1 - \frac{2}{\theta^4}. \tag{35}$$

Lemma 5: The following inequality holds:

$$\sum_{a=1}^{I-1}\sum_{i=a}^{I}\frac{\Delta_{a,a+1}}{\Delta_{a,i}^2} \le 2\sum_{a=1}^{I-1}\frac{1}{\Delta_{a,a+1}}. \tag{36}$$

Lemma 6: The total number of routing rounds $\Theta(W_B)$ under budget W_B is upper bounded by

$$\Theta(W_B) \le \frac{W_B\Delta_{\min}^2 W_{\min} - 4W_{\max}W^*}{W^*\left[\Delta_{\min}^2 W_{\min} - W_{\max}(4 + 2\Delta_{\min}^2)\right]}. \tag{37}$$

Proof of Lemmas 4, 5, and 6: Please kindly refer to the supplementary material. ∎

Built upon the above lemmas, we give the following theorem regarding the upper bound on the regret of the SMAB-SR algorithm.

Theorem 1: The regret of the SMAB-SR algorithm under budget W_B is upper bounded by

$$\begin{aligned}\mathrm{Reg}(W_B) \le &\left\{18\ln\frac{W_B\Delta_{\min}^2 W_{\min} - 4W_{\max}W^*}{W^*[\Delta_{\min}^2 W_{\min} - W_{\max}(4+2\Delta_{\min}^2)]}\right.\\ &\left. + O(1)\right\}\cdot\sum_{i=1}^{I-1}\frac{1}{\Delta_{i,i+1}}.\end{aligned} \tag{38}$$

Proof: For $1 \le a < i \le I$, let $N_{a,i}$ denote the total number of times (i.e., from the first round until the budget is exhausted) route i is selected under the condition that there exists at least an available route whose index is within $\{1, 2, \cdots, a\}$. Accordingly, under the same condition, let $R_{a,i}$ represent the remaining number of times route i will be selected after being chosen $Q_{a,i}$ times, i.e., $N_{a,i} = R_{a,i} + Q_{a,i}$. Moreover, we define the events E_1 and E_2 as follows:

$$\begin{cases}\mathrm{E}_1 \triangleq \bigcup\limits_{z=1}^{a}\left(\bar{Z}_i(\theta) + \sqrt{\dfrac{2\ln\theta}{r}} \ge \bar{Z}_z(\theta) + U_z(\theta)\right),\\ \mathrm{E}_2 \triangleq \bigcup\limits_{z\in\{1,\cdots,a\}\cup\{i\}}\bar{Z}_z(\theta) \in [Z_z(\theta) - U_z(\theta), Z_z(\theta) + U_z(\theta)].\end{cases} \tag{39}$$

When $Q_{a,i} = \lfloor\frac{9\ln\Theta(W_B)}{\Delta_{a,i}^2}\rfloor$, we have

$$\begin{aligned}R_{a,i} &= \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\mathbb{P}\big\{(\pi^\theta = i) \wedge (\text{route } i \text{ is chosen } r \text{ times})\\ &\quad \wedge (\mathcal{R}(\theta) \cap \{\text{the first } a \text{ routes}\} \neq \emptyset)\big\}\\ &\le \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\mathbb{P}\big\{(\pi^\theta = i) \wedge (\beta(\theta) = r) \wedge \mathrm{E}_1\big\}\\ &\le \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\mathbb{P}\big\{\mathrm{E}_1\big\}.\end{aligned} \tag{40}$$

Applying the law of total probability, we can obtain

$$\begin{aligned}\mathbb{P}\big\{\mathrm{E}_1\big\} &= \mathbb{P}\big\{\mathrm{E}_1|\mathrm{E}_2\big\}\mathbb{P}\big\{\mathrm{E}_2\big\} + \mathbb{P}\big\{\mathrm{E}_1|\mathrm{E}_2^c\big\}\mathbb{P}\big\{\mathrm{E}_2^c\big\}\\ &\le \mathbb{P}\big\{\mathrm{E}_1|\mathrm{E}_2\big\} + \mathbb{P}\big\{\mathrm{E}_2^c\big\},\end{aligned} \tag{41}$$

where E_2^c denotes the contrary event of E_2. According to Lemma 4, the probability that event E_2 happens is at least $1 - 2(a+1)\theta^{-4}$, so we have $\mathbb{P}\big\{\mathrm{E}_2^c\big\} \le 2(a+1)\theta^{-4}$.

Given that route i has been selected $Q_{a,i}$ times, the confidence interval $U_i(\theta)$ is at most $\sqrt{\frac{2\ln\Theta(W_B)\Delta_{a,i}^2}{9\ln\Theta(W_B)}} < \frac{\Delta_{a,i}}{2}$. Then, if event E_2 happens, there are $\bar{Z}_i(\theta) + U_i(\theta) \le (Z_i + U_i(\theta)) + U_i(\theta) < Z_i + \Delta_{a,i} = Z_a$, and $\bar{Z}_z(\theta) + U_z(\theta) \ge Z_z$, $\forall z \in \{i\} \cup \{1, 2, ..., a\}$. Substituting (41) into (40) yields

$$\begin{aligned}R_{a,i} &\le \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\left(\mathbb{P}\big\{\mathrm{E}_1|\mathrm{E}_2\big\} + \mathbb{P}\big\{\mathrm{E}_2^c\big\}\right)\\ &\le \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\left(\mathbb{P}\Big\{\bigcup_{z=1}^{a}(\bar{Z}_i(\theta) + U_i(\theta) \ge Z_z)\Big\} + \mathbb{P}\big\{\mathrm{E}_2^c\big\}\right)\\ &\le \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\mathbb{P}\big\{\bar{Z}_i(\theta) + U_i(\theta) \ge Z_a\big\}\\ &\quad + \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\frac{2(a+1)}{\theta^4}\\ &= \sum_{r=Q_{a,i}+1}^{\Theta(W_B)}\sum_{\theta=r}^{\Theta(W_B)}\frac{2(a+1)}{\theta^4} \le O\left(\frac{I}{\Theta(W_B)^2}\right) = O(1).\end{aligned} \tag{42}$$

Then, we have

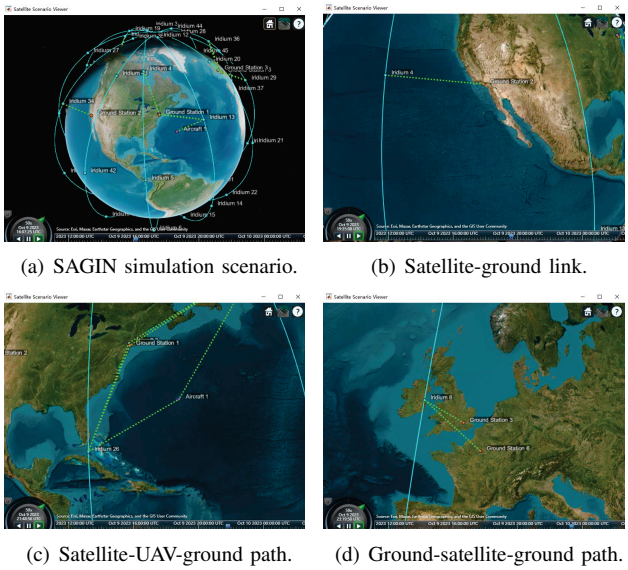
$$N_{a,i} = Q_{a,i} + R_{a,i} \le \frac{9\ln\Theta(W_B)}{{\Delta_{a,i}}^2} + O(1). \tag{43}$$

The regret can be re-expressed as

$$\begin{aligned}\mathrm{Reg}(W_B) &= \mathbb{E}\{\sum_{i=2}^{I}\sum_{a=1}^{i-1}(N_{a,i} - N_{a-1,i})\Delta_{a,i}\}\\ &= \sum_{i=2}^{I}\sum_{a=1}^{i-1}N_{a,i}(\Delta_{a,i} - \Delta_{a+1,i})\\ &\le [9\ln\Theta(W_B) + O(1)]\sum_{i=2}^{I}\sum_{a=1}^{i-1}\frac{\Delta_{a,i} - \Delta_{a+1,i}}{\Delta_{a,i}^2}.\end{aligned} \tag{44}$$

Substituting (36) and (37) into (44), we obtain (38).

This completes the proof. ∎



(a) SAGIN simulation scenario. (b) Satellite-ground link.

(c) Satellite-UAV-ground path. (d) Ground-satellite-ground path.

Fig. 4. Simulation Setup.

B. Computational Complexity Analysis

We begin by analyzing Algorithm 1, which consists of two phases: (1) Connectivity Calculation Phase. In this phase, the connectivity matrix $C^{G,S}$ is generated through a nested loop with the complexity of $O(N_G N_S)$, while the connectivity matrix $C^{U,S}$ requires another nested loop with the complexity of $O(N_U N_S)$. Hence, the complexity of this phase is $O(N_S(N_G + N_U))$; (2) Available Route Set Determination Phase. This phase iterates over all $|\mathcal{R}^{\text{univ}}|$ candidate routes and verifies the connectivity of K-hop links per route. So, the complexity of this phase is $O(|\mathcal{R}^{\text{univ}}|K)$. Therefore, the overall computational complexity T_{Alg1} of Algorithm 1 is $O(N_S(N_G + N_U) + |\mathcal{R}^{\text{univ}}|K)$.

For the SMAB-SR algorithm, the main loop iterates $\Theta(W_B)$ times, which is upper bounded by Eq. (37). Each iteration consists of three computational phases: (1) invoking Algorithm 1; (2) ranking the available routes based on the UCB criterion, with the complexity of $O(|\mathcal{R}(\theta)| \log_2 |\mathcal{R}(\theta)|)$; (3) updating the parameters of all available routes, with the complexity of $O(|\mathcal{R}(\theta)|)$. Based on the above analysis, the total computational complexity of the SMAB-SR algorithm is $O(\Theta(W_B) \cdot (T_{\text{Alg1}} + |\mathcal{R}(\theta)| \log_2 |\mathcal{R}(\theta)|)$.

VI. Simulation Results

A. Simulation Methodology

1) Simulation Setup: Our SAGIN simulation scenario and the entire simulation process–including SAGIN topology construction, 3D visualization, routing algorithm execution, and performance evaluation–are implemented using MATLAB and its associated toolboxes, primarily the *Satellite Communications Toolbox*[3]. Specifically, we employ the built-in *satelliteScenario* function from the *Satellite Communications Toolbox* to create an object that can define various network entities such as satellites, UAVs, and GNs. The default simulation scenario includes 48 LEO satellites from the Iridium constellation, 2 UAVs, and 6 GNs (base stations). We download the Two-Line Element (TLE) files of relevant LEO satellites from *CelesTrak*[4], which is a publicly accessible platform that provides up-to-date satellite orbital data in standardized formats[5]. Then, we assign the orbital parameters extracted from the TLE files to the satellites defined by the *satelliteScenario* object via built-in interfaces, enabling accurate replication of actual Iridium satellite trajectories within the MATLAB-based environment. In addition, the ground stations are distributed globally in cities such as New York, Tokyo, and Paris, while UAV flight routes are configured by specifying stopover points and flight durations between them. Fig. 4 illustrates the visualization interface of our simulation platform, including the overall SAGIN topology and detailed link information, which is primarily generated using the built-in *satelliteScenarioViewer* function of the toolbox.

In the simulation, we set the transmit power of legitimate nodes to 1W and the jamming power of jammers to 0.1W. The path loss exponent is set to 2, and the constant in the average path loss model is set to 10. The satellite antenna gain factor is configured as 5, while the densities of eavesdroppers and jammers are both set to 0.01. For each simulation, we run for a total of 15 hours and 30 minutes and we obtain the distance matrix every 10 seconds. The default settings of the main simulation parameters are summarized in Table I.

TABLE I
Default Settings of Main Simulation Parameters

Simulation Parameter	Default Setting
Number of LEO satellites	48
Satellite constellation	Iridium
Number of UAVs	2
Number of ground nodes	6
Transmit Power	1W
Jamming power	0.1W
Legitimate link bandwidth	1Gbps
Eavesdropping link bandwidth	1Gbps
Path loss exponent	2
Average path loss constant	10
Satellite antenna gain	5
the interference node density	0.01
the eavesdropping node density	0.01

2) Evaluation Metrics and Comparison Algorithms: In the simulation, we test the following routing performance evaluation metrics: i) Cumulative reward: the cumulative reward achieved by the algorithm as the number of rounds grows, which is equivalent to the cumulative E2E secure transmission throughput; ii) Average reward: the average reward (i.e., average secure transmission throughput) that can be received per round of routing, calculated as the total reward divided by the total number of rounds; iii) Cumulative regret: the cumulative reward gap between the adopted algorithm and the optimal strategy under ideal conditions; iv) Reward-cost-ratio: the reward gained per unit of cost consumed. Moreover,

[3]https://www.mathworks.com/products/satellite-communications.html.

[4]https://celestrak.org/.

[5]The authors gratefully acknowledge MathWorks for their development support in satellite communications research, and CelesTrak for providing satellite orbital data publicly.

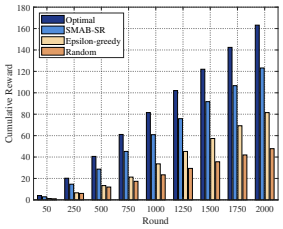


Fig. 5. Cumulative reward achieved by different algorithms.

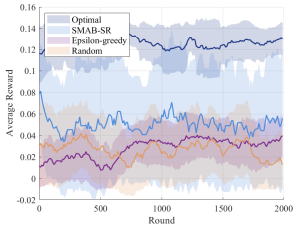


Fig. 6. Average reward achieved by different algorithms.

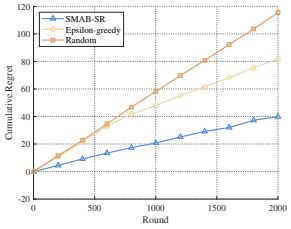


Fig. 7. Cumulative regret achieved by different algorithms.

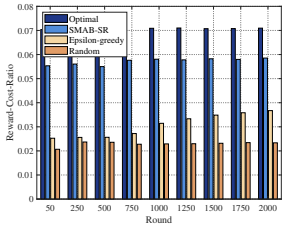


Fig. 8. Reward-cost-ratio of different algorithms.

we have developed three comparison algorithms in the simulation, called "Optimal", "Epsilon-greedy" and "Random". "Optimal" means that the algorithm knows the instantaneous information of all routes in advance and always selects the best route in each round. "Random" selects an available route randomly in each round. "Epsilon-greedy" uses $\epsilon \cdot W_B$ budget to randomly select an available route in the corresponding rounds (pure exploration phase), and then uses the remaining $(1-\epsilon) \cdot W_B$ budget to greedily select the available route with the highest average reward (pure exploitation phase). In the "Epsilon-greedy" algorithm, the default setting of ϵ is 0.2. Additionally, "Epsilon-greedy" updates the arm's metrics only in the exploration phase, while its exploitation phase relies solely on historical statistics. Note that SMAB-SR leverages the UCB-based reward to select a route, dynamically switching between exploration and exploitation according to the route's availability and the current round performance. Moreover, the arm's metrics are updated during both exploration and exploitation, and confidence bounds are incorporated to adjust decisions. These differences imply the advantage of SMAB-SR in terms of more reasonable budget usage and higher adaptability to dynamic environments in SAGIN.

B. Performance Comparison

We first compare the cumulative reward that can be achieved by different algorithms, as summarized in Fig. 5. We can observe that as the number of routing rounds increases, the cumulative reward achieved by all algorithms exhibits a monotonically increasing trend. Notably, the proposed SMAB-SR algorithm demonstrates a significantly faster growth in cumulative reward compared to the Epsilon-greedy algorithm and the Random algorithm, indicating that for a given number of routing rounds, SMAB-SR facilitates a greater amount of E2E secure data transmission. This result validates the effectiveness of the SMAB-SR solution to the challenging secure routing issue in SAGIN scenarios.

We then plot Fig. 6 to show the behaviors of the average reward of different algorithms. To reduce the fluctuations in the data, the *movmean* function is applied to smooth the data of reward for each algorithm, with a window size set to 10. The shaded region surrounding each curve represents the standard deviation range of the smoothed data, reflecting the fluctuation range of the reward data. The wider the shaded region, the greater the fluctuation of the reward data, and vice versa. It can be observed that the curve of the average reward of the SMAB-SR algorithm consistently remains above those of the Epsilon-greedy and Random algorithms, indicating that SMAB-SR enables a larger amount of secure E2E data transmission per round on average. Another interesting phenomenon is that the fluctuations in the average reward of SMAB-SR are greater than those of the other three algorithms. This is due to the rapid state variations in the SAGIN, to which SMAB-SR dynamically adapts by adjusting its route selection accordingly. It indicates that the SMAB-SR algorithm is well-suited to the dynamic nature of SAGIN and has greater potential to achieve a better balance between exploration and exploitation.

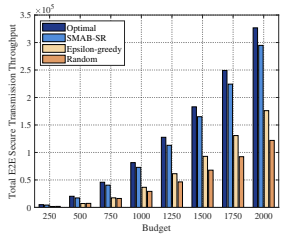
Fig. 7 compares the evolution of cumulative regret across different algorithms as the number of routing rounds increases. We can see that the SMAB-SR algorithm exhibits the slowest growth in cumulative regret among the three algorithms, indicating the smallest performance gap from the optimal strategy under ideal conditions. The Random algorithm shows an approximately linear increase in cumulative regret, as it does not optimize route selection based on historical knowledge. The Epsilon-greedy algorithm initially experiences a rapid increase in cumulative regret, which then slows down due to the transition from the pure exploration phase to the pure exploitation phase. For the SMAB-SR algorithm, the growth rate of cumulative regret gradually flattens as the number of rounds increases, demonstrating its ability to achieve a better exploration-exploitation balance and converge more quickly to an appropriate route selection.

In Fig. 8, we compare the reward-cost-ratio (RCR, i.e., the cumulative reward divided by the cumulative cost) across different algorithms. Since all information about the SAGIN is known in advance, the Optimal algorithm consistently achieves the highest RCR. The SMAB-SR algorithm ranks second, maintaining a relatively stable RCR as the number of rounds changes, indicating its ability to stably achieve a larger reward (i.e., greater E2E secure transmission throughput) at a certain cost (i.e., delay constraints). The Epsilon-greedy algorithm shows an increase in its RCR after transiting to the pure exploitation phase, but it still lags behind SMAB-SR. This is because Epsilon-greedy cannot adaptively explore and optimize route selection in response to the dynamics of SAGIN, as effectively as the SMAB-SR algorithm.

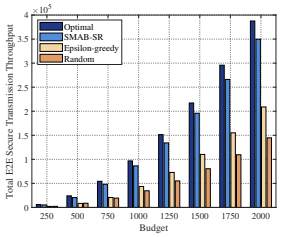
C. Impacts of System Parameters

We further conduct simulations under different settings of the number of UAVs, GNs, and satellites in the SAGIN, to demonstrate the impacts of system parameters on routing performance.

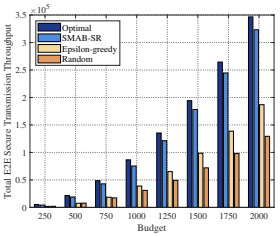
We summarize respectively in Fig. 9 and Fig. 10 how the total E2E secure transmission throughput and the total regret vary with the budget under different UAV quantity settings. The total throughput and total regret refer to the



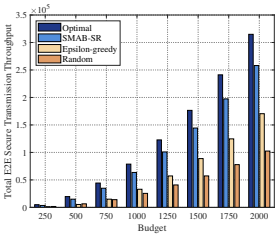
(a) 2 UAVs.



(b) 4 UAVs.

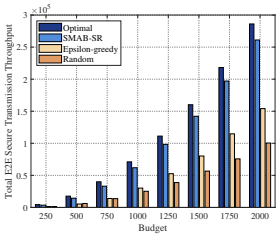


(c) 6 UAVs.

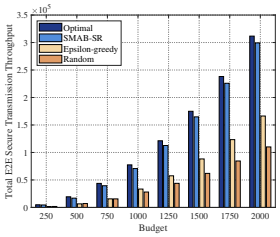


(d) 8 UAVs.

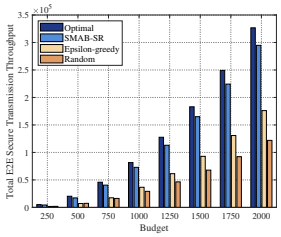
Fig. 9. Total E2E secure transmission throughput versus budget under different numbers of UAVs.



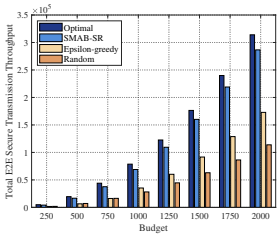
(a) 12 LEO satellites.



(b) 24 LEO satellites.

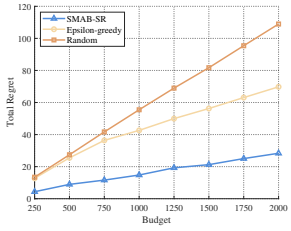


(c) 48 LEO satellites.

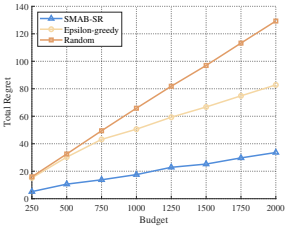


(d) 96 LEO satellites.

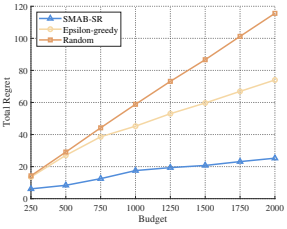
Fig. 11. Total E2E secure transmission throughput versus budget under different numbers of LEO satellites.



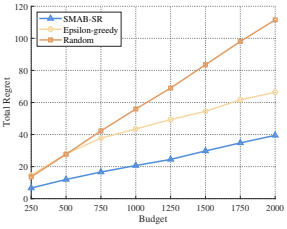
(a) 2 UAVs.



(b) 4 UAVs.

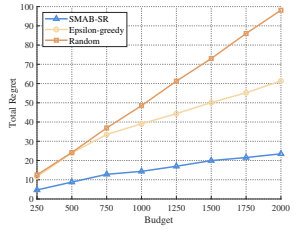


(c) 6 UAVs.

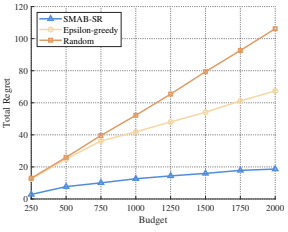


(d) 8 UAVs.

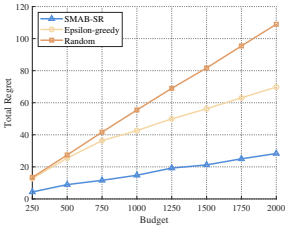
Fig. 10. Total regret versus budget under different numbers of UAVs.



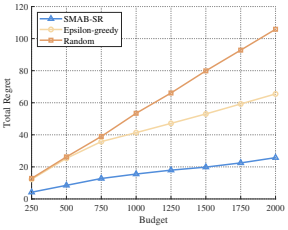
(a) 12 LEO satellites.



(b) 24 LEO satellites.



(c) 48 LEO satellites.

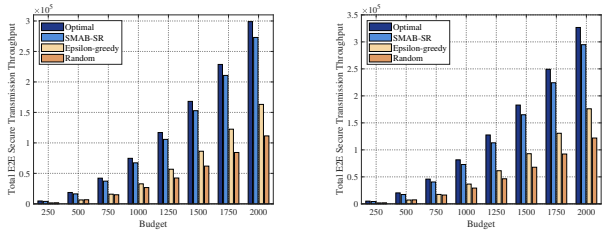


(d) 96 LEO satellites.

Fig. 12. Total regret versus budget under different numbers of LEO satellites.

cumulative throughput and cumulative regret achieved by the algorithm when the entire budget is exhausted, respectively. Each subfigure corresponds to a specific number of UAVs: (a) 2 UAVs, (b) 4 UAVs, (c) 6 UAVs, and (d) 8 UAVs. Other system parameters follow the default setting. Fig. 9 and Fig. 10 show that in all scenarios, as the total budget becomes larger, the total secure throughput and total regret of all algorithms increase (the same phenomenon can also be seen in the subsequent Figs. 11-14). This is because the increase in budget allows for more routing rounds to be executed. It can be observed from Fig. 9 that the total throughput of the SMAB-SR algorithm significantly outperforms the Epsilon-greedy and Random algorithms, and is very close to the Optimal algorithm. Moreover, an increase in the number of UAVs does not always improve the total throughput of the routing algorithm. For instance, the total throughput of all algorithms is generally higher when the number of UAVs is 4 compared to when it is 8. A possible explanation for this phenomenon is that the increase in UAVs leads to a larger number of suboptimal routes, which may cause the algorithm to explore more suboptimal routes, thereby decreasing the total throughput. From Fig. 10, it can be seen that the slope of the increase in total regret for the SMAB-SR algorithm is significantly smaller than that of the other algorithms as the budget increases, indicating that it can efficiently utilize the budget (i.e., delay constraint) to achieve near-optimal route selection.

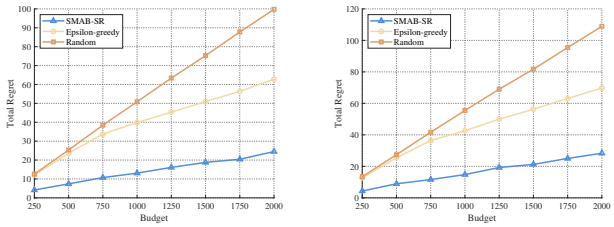
We then plot Fig. 11 and Fig. 12 to show the performance of total E2E secure transmission throughput and total regret of the algorithms under different LEO satellite quantity settings, respectively. Each subfigure corresponds to a specific number of LEO satellites: (a) 12 LEO satellites, (b) 24 LEO satellites, (c) 48 LEO satellites, and (d) 96 LEO satellites. We can observe from Fig. 11 and Fig. 12 that under all LEO satellite configurations, the performance of the SMAB-SR algorithm is significantly better than the comparison algorithms and can approach closely the Optimal algorithm. Comparing different LEO satellite configurations, it can be seen that increasing



(a) 4 ground nodes. (b) 6 ground nodes.

(c) 8 ground nodes. (d) 10 ground nodes.

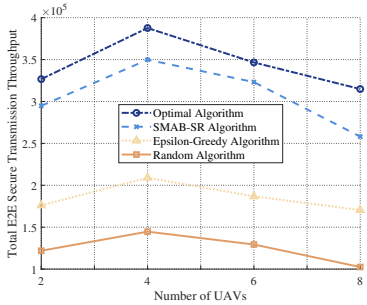
Fig. 13. Total E2E secure transmission throughput versus budget under different numbers of GNs.



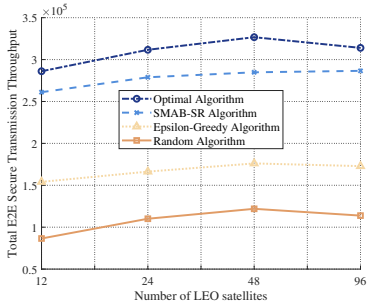
(a) 4 ground nodes. (b) 6 ground nodes.

(c) 8 ground nodes. (d) 10 ground nodes.

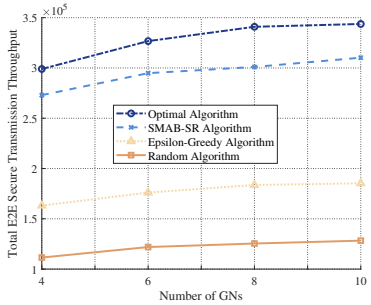
Fig. 14. Total regret versus budget under different numbers of GNs.



(a) Total secure transmission throughput versus number of UAVs.



(b) Total secure transmission throughput versus number of LEO satellites.



(c) Total secure transmission throughput versus number of GNs.

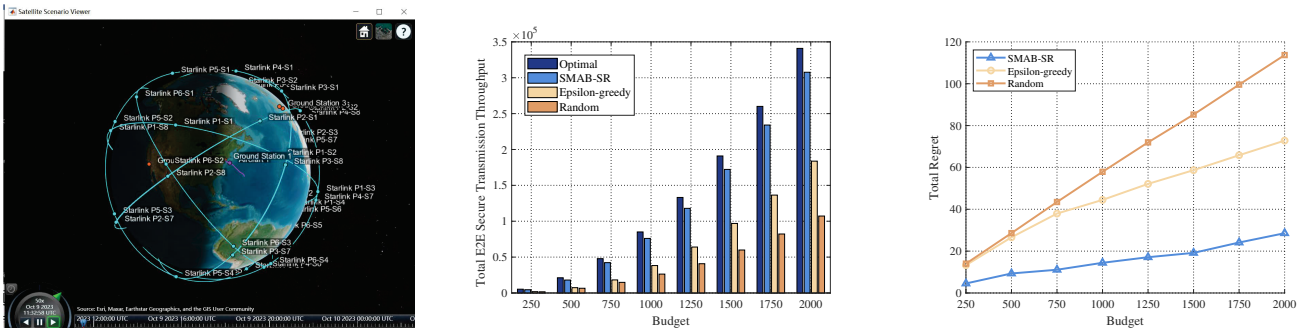
Fig. 15. Impacts of node quantity in SAGIN on total secure transmission throughput.

the number of LEO satellites can generally improve the total throughput of routing algorithms. However, when the number of LEO satellites becomes sufficiently large, the improvement in throughput will diminish, indicating a certain saturation effect in the deployment of LEO satellites in SAGIN.

Fig. 13 and Fig. 14 respectively summarize the variations of the total E2E secure transmission throughput and total regret of all algorithms as the budget increases under different GN quantity configurations. Each subfigure corresponds to a specific number of GNs: (a) 4 GNs, (b) 6 GNs, (c) 8 GNs, and (d) 10 GNs. Consistent with the previous results, under all GN quantity configurations, the SMAB-SR algorithm achieves performance close to the Optimal algorithm (e.g., with a total throughput gap of no more than 10%) and significantly outperforms the comparison algorithms. Similarly, as in the case of LEO satellites, increasing the number of GNs can improve throughput to some extent, but a saturation effect is also observed. Furthermore, we can see from Fig. 14 as well as Fig. 10 and Fig. 12 that, as the number of nodes in the SAGIN increases, the trend of total regret growth of the SMAB-SR algorithm with increasing budget remains almost the same. This indicates that when there are more available routes in the SAGIN, the SMAB-SR algorithm can still efficiently converge to the near-optimal route selection.

In Fig. 15, we summarize the impacts of network node quantity configurations in SAGIN on the total E2E secure transmission throughput achieved by algorithms, with the budget set to $W_B = 2000$. Fig. 15 shows that the throughput of all algorithms first increases and then decreases as the number of UAVs increases, while it gradually increases with the number of LEO satellites or GNs. Notably, regardless of variations in network node configurations, the proposed SMAB-SR algorithm can always approximate the Optimal algorithm very well, demonstrating its effectiveness and stability as a robust solution for enabling E2E secure transmission in SAGIN environments.

Finally, we replace the Iridium constellation in the simulation with the Starlink constellation at an orbital altitude of 550km, setting the number of satellite nodes to 48, and compare the performance of routing algorithms. The results are summarized in Fig. 16. Fig. 16(a) shows the topology of the Starlink constellation, which exhibits significant differences from that of the Iridium constellation. As observed from Figs. 16(b) and 16(c), under all budget settings, the SMAB-SR algorithm consistently outperforms the two benchmark algorithms, achieving higher secure throughput and lower regret, with performance close to that of the Optimal algorithm. This demonstrates the adaptability of SMAB-SR to diversified SAGIN scenarios. Furthermore, comparing the results across the two constellation configurations reveals that the secure



(a) Starlink constellation topology.

(b) Total E2E secure transmission throughput versus budget.

(c) Total regret versus budget.

Fig. 16. Performance comparison under the Starlink constellation.

throughput of SMAB-SR is slightly higher in the Starlink constellation scenario, indicating that constellation topology can indeed have a measurable impact on routing performance.

VII. Related Work

A. Space-Air-Ground Integrated Networks

With the rapid evolution of communication networks, SAGIN has been proposed as a promising 6G solution and become the forefront of future advanced network research [2]. SAGIN aims to integrate ground information infrastructure, air communication platforms, and satellite networks to enhance global coverage and ensure seamless connectivity [3]. However, this integrated architecture also brings several unique challenges, such as long propagation delay, high Doppler shifts, and frequent handovers, which will especially pose significant difficulties for the effective operation of traditional network routing protocols [19]. In addition, the literature [20] outlined the future trends of SAGIN in the context of 6G evolution and highlighted the associated security threats. To date, existing research has primarily focused on the underlying communication technologies, networking, and resource allocation in SAGIN, while the investigation on the design of upper-layer routing protocols, particularly those considering security aspects, remains relatively limited.

B. Routing Solutions in SAGIN

A survey of the existing literature indicates that there is some recent relevant work on routing in SAGIN or satellite networks. Geng *et al.* [21] developed an agent-based framework for satellite networks and introduced a clustering and game theory-based collaborative routing scheme to optimize network efficiency and reliability. Xu *et al.* [22] proposed a spatial location-aided fully distributed routing scheme for large-scale LEO satellite networks, where a multi-agent deep reinforcement learning mechanism was designed to optimize routing and minimize the average delay. Wang *et al.* [23] presented a multi-task routing strategy for integrated satellite-terrestrial networks, which combines satellite control and convolutional neural network-based path allocation with fuzzy logic to enhance load balance and quality of experience. Du *et al.* [24] explored location management for mega satellite networks, where a satellite-assisted routing scheme was proposed to reduce management overhead and avoid network congestion. Maity *et al.* [25] introduced a virtual network embedding-based routing protocol, which can transform the satellite network model into a virtual network and map real satellites to virtual nodes according to the physical locations and assigned geographic coordinates. Bhattacharjee *et al.* [26] developed a nonlinear optimization model for dynamic routing in LEO satellite mega-constellations, and designed heuristic algorithms to balance performance and computational complexity. Huang *et al.* [27] studied topology management and route planning for multi-layer satellite networks, and proposed a multi-objective inter-layer link allocation scheme to improve topology stability, transmission efficiency, and user experience. In [28], [29], the authors developed theoretical analysis frameworks based on stochastic geometry, for investigating the routing reliability in LEO satellite constellations. More recently, Duan *et al.* [30] introduced an adaptive snapshot mechanism for distributed LEO satellite networks to optimize route reconstruction, which can minimize routing overhead while maintaining performance under dynamic topologies and failures.

Note that the Markov Decision Process (MDP) is a robust theoretical framework for sequential decision-making under environmental uncertainty, which has been widely applied to routing design in various types of networks. For example, MDP-based routing methods have been proposed for delay-tolerant networks in [31], [32]. We believe that MDP also has strong potential for routing design in the uncertain environment of SAGIN. However, several challenges must be addressed. First, as the network scales, the state and action spaces expand rapidly, rendering the exact solution computationally intractable. Second, the pronounced spatiotemporal dynamics of SAGIN make accurate modeling of state transition probabilities particularly difficult. Third, the computational and storage requirements of solving MDPs may conflict with the resource constraints of SAGIN nodes. In contrast, the MAB framework developed in this work does not require explicit transition probability modeling and can be regarded as a lightweight decision-making mechanism under uncertainty. Overall, integrating MDP with lightweight online learning techniques represents a promising direction for addressing routing problems in SAGIN.

Table II summarizes recent significant research efforts on routing in satellite-related networks, presenting a comparative analysis between existing studies and our work in terms of network topology dynamics, link reliability, delay constraints,

TABLE II
SUMMARY OF RESEARCH WORKS ON ROUTING IN SATELLITE-RELATED NETWORKS

Research Work	Network Topology Dynamics	Link Reliability	Delay Constraint	Routing Security Against Eavesdropping Attacks
Geng *et al.* 2021 [21]		✓		
Xu *et al.* 2022 [22]	✓		✓	
Wang *et al.* 2022 [23]	✓	✓	✓	
Du *et al.* 2023 [24]		✓		
Maity *et al.* 2024 [25]	✓			
Bhattacharjee *et al.* 2024 [26]	✓			
Huang *et al.* 2024 [27]		✓		
Wang *et al.* 2024 [28], [29]		✓		
Duan *et al.* 2025 [30]	✓			
Our work	✓	✓	✓	✓

and routing security. As highlighted in the table, although notable progress has been made, substantial gaps remain, particularly in addressing routing security against eavesdropping attacks and coping with the dynamic and uncertain nature of SAGIN environments. Our work precisely aims to bridge these gaps and stimulate further research in this direction. For instance, beyond routing security against eavesdropping, other critical aspects, such as authorization, authentication, encryption, and integrity, are also indispensable to constructing a truly comprehensive security framework for SAGIN [33], [34]. Tackling these dimensions in a unified manner is therefore not only urgent but also pivotal to ensuring the robustness and trustworthiness of future SAGIN operations.

C. Applications of MAB in Routing Design

MAB has emerged as a prominent research topic in the field of reinforcement learning in recent years [35]–[37], garnering substantial attention across diverse applications such as adaptive resource allocation, recommendation systems, online learning, and dynamic decision-making, with notable research extensions into the network routing design.

Henri *et al.* [38] designed an MAB-based algorithm to explore and select optimal multipaths in hybrid networks while managing congestion, which adapts to dynamic conditions, ensuring optimal throughput in static scenarios and efficient path switching in response to changing environments. Zhou *et al.* [39] investigated online shortest path routing in networks with stochastically varying link states under potential denial-of-service attacks, where an MAB-based approach was developed to adaptively identify jammed and unjammed links through a novel exploration mechanism and a martingale-based adversarial learning framework. Bura *et al.* [40] focused on the regret analysis of caching algorithms from an MAB perspective, demonstrating that simple algorithms can achieve optimal regret by exploiting request distribution structure, thus avoiding high complexity. Fu *et al.* [41] formulated the optimal routing problem in a discrete-time system with parallel servers as an MAB with queues, aiming to maximize total utility under stochastic arrivals, service rates, and unknown utilities. The authors combined stochastic queue control and bandit techniques to design routing policies that minimize regret over a finite time horizon. Tabei *et al.* [42] proposed an MAB-based routing method for in-network cache-enabled networks, leveraging collaborative caching and reinforcement learning to improve cache hit ratios, reduce hop counts, and optimize content retrieval.

In the aforementioned MAB-based routing studies, there is no restriction on which route is selected (i.e., which arm is pulled) at a given time. However, due to the inherent dynamics and randomness in SAGIN environments, we introduce the concept of "sleeping", allowing temporarily disconnected routes in the arm set to become inactive. As a novel variant of MAB, SMAB has been partially explored in regret analysis [43] and applied in uplink grant allocation [44]. This work is the first to apply SMAB to routing design, enabling E2E secure transmission in SAGIN.

VIII. CONCLUSION

In this study, we have proposed SMAB-SR, a secure routing solution based on a sleeping multi-armed bandit (SMAB) framework, specifically tailored for SAGIN. This solution aims to maximize message transmission throughput within a specified time while ensuring that messages cannot be intercepted by malicious eavesdroppers. We first modeled the channels of all link types in SAGIN, evaluated the statistical characteristics of secrecy capacity and end-to-end delay for a general route, and formulated the secure routing problem under the delay constraint. Then, we introduced the SMAB framework, which transforms the route selection problem into an arm-pulling problem and utilizes the sleep mode to capture route unavailability caused by intermittent link failures. Furthermore, to address the exploration-exploitation dilemma in route selection, we designed the SMAB-SR algorithm based on the UCB approach and theoretically derived its regret upper bound. The proposed SMAB-SR algorithm has been thoroughly evaluated through extensive simulations and comparative analyses, demonstrating its effectiveness and stability under various SAGIN configurations.

Note that in this work, the analytical expressions for route quality evaluation are derived based on fundamental modeling assumptions, such as the spatial distribution of eavesdroppers following a Poisson point process (PPP). However, in practical SAGIN scenarios, attackers may adopt more sophisticated deployment strategies beyond random spatial distributions, for example, concentrating on critical link regions. Furthermore, unintentional interference from benign nodes introduces additional complexity to the spatial interference distribution in the network. Therefore, investigating route quality evaluation,

secure routing design, and the optimization of security defense strategies under more practical node distribution and interference models in SAGIN environments will be the next significant direction of our research.

REFERENCES

[1] W. Chen, X. Lin, J. Lee, A. Toskala, S. Sun, C. F. Chiasserini, and L. Liu, "5g-advanced toward 6g: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1592–1619, 2023.

[2] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 2018.

[3] Y. Xiao, Z. Ye, M. Wu, H. Li, M. Xiao, M.-S. Alouini, A. Al-Hourani, and S. Cioni, "Space-air-ground integrated wireless networks for 6g: Basics, key technologies and future trends," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 12, pp. 3327–3354, 2024.

[4] B. Shang, Y. Yi, and L. Liu, "Computing over space-air-ground integrated networks: Challenges and opportunities," *IEEE Netw.*, vol. 35, no. 4, pp. 302–309, 2021.

[5] D. Zhou, M. Sheng, J. Li, and Z. Han, "Aerospace integrated networks innovation for empowering 6g: A survey and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 975–1019, 2023.

[6] F. Tang, H. Hofner, N. Kato, K. Kaneko, Y. Yamashita, and M. Hangai, "A deep reinforcement learning-based dynamic traffic offloading in space-air-ground integrated networks (sagin)," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 276–289, 2022.

[7] C. Chen, Z. Liao, Y. Ju, C. He, K. Yu, and S. Wan, "Hierarchical domain-based multicontroller deployment strategy in sdn-enabled space–air–ground integrated network," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 6, pp. 4864–4879, 2022.

[8] W. Wei, L. Fu, H. Gu, X. Lu, L. Liu, S. Mumtaz, and M. Guizani, "Iris: Towards intelligent reliable routing for software defined satellite networks," *lEEE Trans. Commun.*, vol. 73, no. 1, pp. 454–468, 2025.

[9] P. Zhang, Z. Xian, M. Liao, H. Huang, and J. Yang, "Adaptive routing mechanism for leo satellite network based on control domain partition," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 70–82, 2025.

[10] R. Wang, M. A. Kishk, and M.-S. Alouini, "Modeling and analysis of non-terrestrial networks by spherical stochastic geometry: A survey," *IEEE Commun. Surveys Tuts.*, 2025, Early Access.

[11] A. Bansal, N. Agrawal, and K. Singh, "Rate-splitting multiple access for uav-based ris-enabled interference-limited vehicular communication system," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 936–948, 2023.

[12] Y. Liu, C. She, Y. Zhong, W. Hardjawana, F.-C. Zheng, and B. Vucetic, "Graph neural networks for quality of service improvement in interference-limited ultra-reliable and low-latency communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 3, pp. 3718–3732, 2024.

[13] C. Xu, P. Zhang, H. Yu, and Y. Li, "D3qn-based multi-priority computation offloading for time-sensitive and interference-limited industrial wireless networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 13 682–13 693, 2024.

[14] A. Al-Hourani and K. Gomez, "Modeling cellular-to-uav path-loss for suburban environments," *lEEE Wreless Commun. Lett.*, vol. 7, no. 1, pp. 82–85, 2018.

[15] S. A. Kanellopoulos, C. I. Kourogiorgas, A. D. Panagopoulos, S. N. Livieratos, and G. E. Chatzarakis, "Channel model for satellite communication links above 10ghz based on weibull distribution," *IEEE Commun. Lett.*, vol. 18, no. 4, pp. 568–571, 2014.

[16] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[17] J. F. Kingman, "The single server queue in heavy traffic," *Math. Proc. Cambridge Phil. Soc.*, vol. 57, no. 4, pp. 902–904, 1961.

[18] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, pp. 235–256, 2002.

[19] S. Mahboob and L. Liu, "Revolutionizing future connectivity: A contemporary survey on ai-empowered satellite-based non-terrestrial networks in 6g," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1279–1321, 2024.

[20] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A survey on space-air-ground-sea integrated network security in 6g," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 53–87, 2022.

[21] S. Geng, S. Liu, Z. Fang, and S. Gao, "An agent-based clustering framework for reliable satellite networks," *Reliab. Eng. Syst. Saf*, vol. 212, p. 107630, 2021.

[22] G. Xu, Y. Zhao, Y. Ran, R. Zhao, and J. Luo, "Spatial location aided fully-distributed dynamic routing for large-scale leo satellite networks," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 3034–3038, 2022.

[23] F. Wang, D. Jiang, Z. Wang, Z. Lv, and S. Mumtaz, "Fuzzy-cnn based multi-task routing for integrated satellite-terrestrial networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1913–1926, 2022.

[24] P. Du, J. Li, W. Bai, M. Sheng, and D. Zhou, "Dual location area based distributed location management for hybrid leo/meo mega satellite networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 2307–2321, 2023.

[25] I. Maity, T. X. Vu, and S. Chatzinotas, "Revine: Reinforcement learning-based virtual network embedding in satellite-terrestrial networks," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6316–6327, 2024.

[26] D. Bhattacharjee, P. G. Madoery, A. U. Chaudhry, H. Yanikomeroglu, G. K. Kurt, P. Hu, K. Ahmed, and S. Martel, "On-demand routing in leo mega-constellations with dynamic laser inter-satellite links," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 5, pp. 7089–7105, 2024.

[27] Y. Huang, B. Feng, A. Tian, P. Dong, S. Yu, and H. Zhang, "An efficient differentiated routing scheme for meo/leo-based multi-layer satellite networks," *IEEE Trans. Network Sci. Eng.*, vol. 11, no. 1, pp. 1026–1041, 2024.

[28] R. Wang, M. A. Kishk, and M.-S. Alouini, "Reliability analysis of multi-hop routing in multi-tier leo satellite networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 3, pp. 1959–1973, 2024.

[29] ——, "Ultra reliable low latency routing in leo satellite constellations: A stochastic geometry approach," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 5, pp. 1231–1245, 2024.

[30] R. Duan, H. Zhang, and X. Li, "Adaptive snapshot-based elastic route reconstruction for distributed leo satellite networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 2, pp. 3516–3530, 2025.

[31] F. D. Raverta, J. A. Fraire, P. G. Madoery, R. A. Demasi, J. M. Finochietto, and P. R. D'argenio, "Routing in delay-tolerant networks under uncertain contact plans," *Ad Hoc Netw.*, vol. 123, 2021, Art. no. 102663.

[32] P. R. D'argenio, J. A. Fraire, A. Hartmanns, and F. Raverta, "Comparing statistical and analytical routing approaches for delay-tolerant networks," in *Proc. Int. Conf. Quant. Eval. Syst. (QEST)*, 2022, pp. 337–355.

[33] R. Chothe, S. Ugale, D. Chandwadkar, and S. Shelke, "A combined cryptography and error correction system based on enhanced aes and ldpc," in *Proc. Int. Conf. Comput., Commun., Control, Autom. (ICCUBEA)*, 2023, pp. 1–6.

[34] Y. Cai, H. Yao, Y. Gong, F. Wang, N. Zhang, and M. Guizani, "Privacy-driven security-aware task scheduling mechanism for space-air-ground integrated networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 5, pp. 4704–4718, 2024.

[35] C. Shi and C. Shen, "Federated multi-armed bandits," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 11, 2021, pp. 9603–9611.

[36] V. Patil, G. Ghalme, V. Nair, and Y. Narahari, "Achieving fairness in the stochastic multi-armed bandit problem," *J. Mach. Learn. Res*, vol. 22, no. 174, pp. 1–31, 2021.

[37] G. Nie, Y. Y. Nadew, Y. Zhu, V. Aggarwal, and C. J. Quinn, "A framework for adapting offline algorithms to solve combinatorial multi-armed bandit problems with bandit feedback," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2023, pp. 26 166–26 198.

[38] S. Henri, C. Vlachou, and P. Thiran, "Multi-armed bandit in action: Optimizing performance in dynamic hybrid networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1879–1892, 2018.

[39] P. Zhou, J. Xu, W. Wang, Y. Hu, D. O. Wu, and S. Ji, "Toward optimal adaptive online shortest path routing with acceleration under jamming attack," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 1815–1829, 2019.

[40] A. Bura, D. Rengarajan, D. Kalathil, S. Shakkottai, and J.-F. Chamberland, "Learning to cache and caching to learn: Regret analysis of caching algorithms," *IEEE/ACM Trans. Netw.*, vol. 30, no. 1, pp. 18–31, 2022.

[41] X. Fu and E. Modiano, "Optimal routing to parallel servers with unknown utilities—multi-armed bandit with queues," *IEEE/ACM Trans. Netw.*, vol. 31, no. 5, pp. 1997–2012, 2023.

[42] G. Tabei, Y. Ito, T. Kimura, and K. Hirata, "Design of multi-armed bandit-based routing for in-network caching," *IEEE Access*, vol. 11, pp. 82 584–82 600, 2023.

[43] F. Li, J. Liu, and B. Ji, "Combinatorial sleeping bandits with fairness constraints," *IEEE Trans. Network Sci. Eng.*, vol. 7, no. 3, pp. 1799–1813, 2020.

[44] S. Ali, A. Ferdowsi, W. Saad, N. Rajatheva, and J. Haapola, "Sleeping multi-armed bandit learning for fast uplink grant allocation in machine type communications," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5072–5086, 2020.

Yang Xu (Senior Member, IEEE) received the B.E. and Ph.D. degrees in Telecommunications Engineering from Xidian University, Xi'an, China, in 2006 and 2014, respectively, where she is currently an associate professor with the School of Computer Science and Technology. She has published over 50 academic papers at premium international journals and conferences, like the IEEE TRANSACTIONS ON MOBILE COMPUTING (TMC), the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING (TDSC), the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (TWC), etc. Her research interests include space-air-ground integrated networks, wireless systems security, network economics, and emerging paradigms of intelligent networking, including AI-native network architecture design, semantic-aware communications, intelligent resource orchestration, and LLM-enabled autonomous network optimization for 6G systems.

Xiaoyuan Liu received the B.E. degree in Internet of Things Engineering from China University of Petroleum, Qingdao, China, in 2022, and received the M.E. degree in Computer Science and Technology from Xidian University, Xi'an, China, in 2025. Her research interests include space-air-ground integrated networks, multi-armed bandit, wireless network routing, etc.

Jia Liu (Senior Member, IEEE) received the B.E. degree in Telecommunications Engineering from Xidian University, Xi'an, China, in 2010, and received the Ph.D. degree in Systems Information Science from Future University Hakodate, Hakodate, Japan, in 2016. His research interests include resilient and intelligent networking, AI-native network architecture and protocol design, LLM-driven network automation and optimization, distributed intelligence and edge learning over wireless networks, etc. He received the IEEE Sapporo Section Encouragement Award in 2016 and 2020.

Tarik Taleb (Senior Member, IEEE) received the B.E. degree (with distinction) in Information Engineering and the M.Sc. and Ph.D. degrees in Information Sciences from Tohoku University in 2001, 2003, and 2005, respectively. He is currently a Full Professor at Ruhr University Bochum, where he heads the Institute of Networked Energy-Efficient Systems. Previously, he was a Professor with the Center for Wireless Communications at the University of Oulu. From October 2014 to December 2021, he served as an Associate Professor with the School of Electrical Engineering at Aalto University. Prior to joining Aalto University, he worked as a Senior Researcher and 3GPP Standards Expert with NEC Europe Ltd.. Until March 2009, he was an Assistant Professor with the Graduate School of Information Sciences at Tohoku University, where he worked in a laboratory fully funded by KDDI. From 2005 to 2006, he was a Research Fellow with the Intelligent Cosmos Research Institute. He was directly involved in the development and standardization of the Evolved Packet System as a member of the 3GPP System Architecture Working Group. His current research interests include AI-based network management, autonomous edge–cloud continuum management, architectural enhancements to mobile core networks, network softwarization and slicing, software-defined security, and mobile multimedia streaming.

Yusheng Ji (Fellow, IEEE) received the B.E., M.E., and Ph.D. degrees in Electrical Engineering from the University of Tokyo. She joined the National Center for Science Information Systems (NACSIS), Tokyo, Japan in 1990 as an assistant professor, then became an associate professor. In 2000, she became an associate professor, then a professor with National Institute of Informatics (NII), Tokyo, Japan, and also with the Graduate University for Advanced Studies, SOKENDAI, Japan from 2002. She is a Professor Emeritus at NII and SOKENDAI from April 2026. She has published more than 600 refereed papers in the areas of network resource management, traffic control and performance evaluation, and mobile computing. She has received numerous Paper Awards including IEEE Andrew P. Sage Best Transactions Paper Award, IEEE Communications Society Outstanding Paper Award, etc. She is a General Chair of IEEE INFOCOM 2026, and has served as TPC Co-chair, Symposium Co-chair and Track Co-chair for major conferences such as IEEE INFOCOM, ICC, GLOBECOM, VTC, etc. She is serving/has served editorship for transactions and magazine of IEEE, as well as IEICE and IPSJ. She is a fellow of IEEE, and a Distinguished Speeker of IEEE Vehicular Technology Society.

Norio Shiratori (Life Fellow, IEEE) received his Ph.D. degree from Tohoku University in 1977. He became an Assistant Professor and Associate Professor at the Research Institute of Electrical Communication, Tohoku University, in 1977 and 1984, respectively. In 1990, he was promoted to Full Professor at the School of Engineering, Tohoku University. In 1993, he commenced his role as Full Professor at the Research Institute of Electrical Communication, Tohoku University. In 1997, Prof. Shiratori became a Visiting Professor at UCLA (University of California, Los Angeles). In 1998, he was elevated to the status of IEEE Fellow. In 2004, Prof. Shiratori assumed the role of Japan representative for the International Federation for Information Processing (IFIP). In 2009, he served as the President of the Information Processing Society of Japan. In 2010, he became the Chair of the IEEE Sendai Section. In 2012, he was appointed as a Professor at the Global Information and Telecommunication Institute, Waseda University. In 2013, Prof. Shiratori assumed the role of Vice Chair of the IEEE Japan Council. Recognizing his significant contributions, he was honored with the title of IEEE Life Fellow in 2017. Since 2017, he has been serving as a Professor at Research and Development Initiative, Chuo University.

Prof. Shiratori has published over 15 books and over 600 refereed papers in computer science and related fields. He is a fellow of the Japan Foundation of Engineering Societies (JFES), the Information Processing Society of Japan (IPSJ), and the Institute of Electronics, Information and Communication Engineers (IEICE). He was a recipient of the Minister of MEXT Award from the Japanese Government in 2016, the Science and Technology Award from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) in 2009, the IEICE Achievement Award in 2001, the IEICE Contribution Award in 2011, the IPSJ Contribution Award in 2008, the IEICE Honorary Member in 2012, the IPSJ Honorary Member in 2013, IPSJ Memorial Prize Winning Paper Award in 1985, the IPSJ Best Paper Award in 1997, the IEICE Best Paper Award in 2001, the IEEE 5th SCE01 Best Paper Award in 2001, the IEEE ICPADS 2000 Best Paper Award in 2000, and the IEEE 12th ICOIN Best Paper Award.

Supplementary Material for the Paper "SMAB-SR: A Sleeping Multi-Armed Bandit Framework for Secure Routing in Space-Air-Ground Integrated Networks"

Yang Xu, *Senior Member, IEEE,* Xiaoyuan Liu, *Graduate Student Member, IEEE,* Jia Liu, *Senior Member, IEEE,* Tarik Taleb, *Senior Member, IEEE,* Yusheng Ji, *Fellow, IEEE,* and Norio Shiratori, *Life Fellow, IEEE*

APPENDIX A
PROOF OF LEMMA 1

Proof: When $S_k \in \mathcal{G}$, we determine the cumulative distribution function (CDF) of legitimate link capacity C_{S_k,D_k} and its expectation. Firstly, when $D_k \in \mathcal{G}$, there is

$$\begin{aligned}
\mathbb{P}\left(C_{S_k,D_k} \le x\right) &= \mathbb{P}\left\{\log_2\left(1+\frac{P_{S_k}|h_{S_k,D_k}|^2/d^{\alpha}_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}} P_J|h_{J_j,D_k}|^2/d^{\alpha}_{J_j,D_k}}\right) \le x\right\} \\
&= \mathbb{P}\left\{|h_{S_k,D_k}|^2 \le \frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J|h_{J_j,D_k}|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}/d^{\alpha}_{S_k,D_k}}\right\} \\
&= 1-\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left\{\exp\left[-\frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J|h_{J_j,D_k}|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}/d^{\alpha}_{S_k,D_k}}\right]\right\}\right\} \\
&= 1-\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}}\mathbb{E}_{h_{J_j,D_k}}\left[\exp\left(-\frac{(2^x-1)P_J|h_{J_j,D_k}|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}/d^{\alpha}_{S_k,D_k}}\right)\right]\right\} \\
&= 1-\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}}\int_0^{\infty}\exp\left[-\left(\frac{(2^x-1)P_J/d^{\alpha}_{J_j,D_k}}{P_{S_k}/d^{\alpha}_{S_k,D_k}}+1\right)y\right]dy\right\} \\
&= 1-\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}}\frac{1}{\frac{(2^x-1)P_J/d^{\alpha}_{J_j,D_k}}{P_{S_k}/d^{\alpha}_{S_k,D_k}}+1}\right\}. \qquad \text{(A-1)}
\end{aligned}$$

According to the probability generating functional (PGFL) of PPP of jammers, we can obtain

$$\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}} f(z_{J_j})\right\} = \exp\left[-\lambda_J \int_{R^2}\left(1-f(z_{J_j})\right)dz_{J_j}\right] = \exp\left[-2\pi\lambda_J\int_0^{\infty}\left(1-f(r)\right)rdr\right]. \tag{A-2}$$

Then we have

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\le x\right) &= 1-\exp\left[-2\pi\lambda_J\int_0^{\infty}\frac{1}{1+\frac{P_{S_k}/d^{\alpha}_{S_k,D_k}}{(2^x-1)P_J/r^{\alpha}}}rdr\right]\\ &= 1-\exp\left[-2\pi\lambda_J\left(\frac{P_{S_k}/d^{\alpha}_{S_k,D_k}}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\right].\end{aligned} \tag{A-3}$$

The expectation of C_{S_k,D_k} is determined as

$$\begin{aligned}\mathbb{E}\{C_{S_k,D_k}\} &= \int_0^{\infty}\exp\left[-2\pi\lambda_J\left(\frac{P_{S_k}/d^{\alpha}_{S_k,D_k}}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\right]dx\\ &= \frac{1}{2\pi\lambda_J\left(\frac{P_{S_k}}{d^{\alpha}_{S_k,D_k}P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\ln(2)}.\end{aligned} \tag{A-4}$$

Secondly, when $D_k\in\mathcal{U}$, there is

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\le x\right) &= \mathbb{P}\left(\log_2\left(1+\frac{P_{S_k}\cdot 10^{\frac{P^{loss}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\cdot 10^{\frac{P^{loss}_{J_j,D_k}}{10}}}\right)\le x\right)\\ &= \mathbb{P}\left(\frac{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}+\frac{\sigma_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}+\frac{\sigma_{J_j,D_k}}{10}}}\le 2^x-1\right)\\ &= \mathbb{P}\left(\sigma_{S_k,D_k}\le 10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}+\frac{\sigma_{J_j,D_k}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)\\ &= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{\sigma_{J_j,D_k}}\left[\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]\right\},\end{aligned} \tag{A-5}$$

where $y = 10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}+\frac{\sigma_{J_j,D_k}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)$. Then, we have

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k}\le x\right)\\
&=\frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{\sigma_{J_j,D_k}}\left\{\mathrm{erf}\left(10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}+\frac{\sigma_{J_j,D_k}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)\right\}\right\}\\
&=\frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{\sigma_{J_j,D_k}}\left\{\mathrm{erf}\left(10\log_{10}\left(\frac{10^{\frac{\sigma_{J_j,D_k}}{10}}(2^x-1)\sum_{J_j\in\mathcal{J}} P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)\right\}\right\}\\
&=\frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{\sigma_{J_j,D_k}}\left\{\mathrm{erf}\left(\sigma_{J_j,D_k}+A\right)\right\}\right\},
\end{aligned}\tag{A-6}$$

where $A = 10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)$. Then, we have

$$\mathbb{P}\left(C_{S_k,D_k}\le x\right)=\frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\int_{-\infty}^{\infty}\mathrm{erf}(z+A)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz\right\}=\frac{1}{2\sqrt{2\pi}}\mathbb{E}_{\mathcal{J}}\left\{\int_{-\infty}^{\infty}\mathrm{erf}(z+A)e^{-\frac{z^2}{2}}dz\right\}.\tag{A-7}$$

Using Taylor's formula to expand $\mathrm{erf}(z+A)$ to the first order, we obtain $\mathrm{erf}(z+A)\approx\frac{2}{\sqrt{\pi}}(z+A)$. Then, there is

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k}\le x\right)\approx\frac{1}{2\sqrt{2}\pi}\mathbb{E}_{\mathcal{J}}\left\{\int_{-\infty}^{\infty}\frac{2}{\sqrt{\pi}}(z+A)e^{-\frac{z^2}{2}}dz\right\}\\
&=\frac{1}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left(10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}} P_J\cdot 10^{\alpha\log(d_{J_j,D_k})+\frac{C_{f,J_j}}{10}}}{P_{S_k}\cdot 10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)\\
&=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\sum_{J_j\in\mathcal{J}}10^{\alpha\log(d_{J_j,D_k})}+\log_{10}\left(\frac{(2^x-1)P_J10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right\}\\
&=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\sum_{J_j\in\mathcal{J}}10^{\alpha\log(d_{J_j,D_k})}\right\}+\frac{10}{\sqrt{\pi}}\log_{10}\left(\frac{(2^x-1)P_J10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\\
&=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\sum_{J_j\in\mathcal{J}}10^{\log\left(d_{J_j,D_k}\right)}\right\}+\frac{10}{\sqrt{\pi}}\left[\alpha+\log_{10}\frac{(2^x-1)\,P_J10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right]\\
&=\frac{10}{\log 10\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log\left(\sum_{J_j\in\mathcal{J}}10^{\log\left(d_{J_j,D_k}\right)}\right)\right\}+\frac{10}{\sqrt{\pi}}\left[\alpha+\log_{10}\frac{(2^x-1)\,P_J10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right]
\end{aligned}\tag{A-8}$$

Expanding $\log\left(\sum_{J_j\in\mathcal{J}} 10^{\log\left(d_{J_j,D_k}\right)}\right)$ by the Taylor formula (first-order expansion) yields

$$\log\left(\sum_{J_j\in\mathcal{J}} 10^{\log\left(d_{J_j,D_k}\right)}\right) = \sum_{J_j\in\mathcal{J}} 10^{\log\left(d_{J_j,D_k}\right)} - 1. \tag{A-9}$$

Then, we obtain

$$\begin{aligned}&\mathbb{P}\left(C_{S_k,D_k}\le x\right)\\ &\approx \frac{10}{\log 10\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left[\sum_{J_j\in\mathcal{J}} 10^{\log\left(d_{J_j,D_k}\right)}\right] - \frac{10}{\log 10\sqrt{\pi}} + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right].\end{aligned} \tag{A-10}$$

Expanding $10^{\log\left(d_{J_j,D_k}\right)}$ by the Taylor formula (first-order expansion) yields

$$10^{\log\left(d_{J_j,D_k}\right)} \approx \ln 10\log\left(d_{J_j,D_k}\right) + 1. \tag{A-11}$$

Then, there is

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\le x\right) &\approx \frac{10}{\ln 10\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\sum_{J_j\in\mathcal{J}}\left\{\ln 10\log\left(d_{J_j,D_k}\right)+1\right)\right\}\\ &\quad - \frac{10}{\log 10\sqrt{\pi}} + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right]\\ &= \frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log\left(\prod_{J_j\in\mathcal{J}} d_{J_j,D_k}\right)\right\} + \frac{10}{\ln 10\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left(\sum_{J_j\in\mathcal{J}} 1\right)\\ &\quad - \frac{10}{\log 10\sqrt{\pi}} + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right].\end{aligned} \tag{A-12}$$

Expanding $\log\left(\prod_{J_j\in\mathcal{J}} d_{J_j,D_k}\right)$ using the Taylor formula (first-order expansion) yields

$$\log\left(\prod_{J_j\in\mathcal{J}} d_{J_j,D_k}\right) \approx \prod_{J_j\in\mathcal{J}} d_{J_j,D_k} - 1. \tag{A-13}$$

Then, we have

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\le x\right) &\approx \frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}} d_{J_j,D_k}\right\} - \frac{10}{\sqrt{\pi}} + \frac{10}{\ln 10\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}} 1\right\}\\ &\quad - \frac{10}{\log 10\sqrt{\pi}} + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right].\end{aligned} \tag{A-14}$$

Substituting (A-2) into (A-14) yields

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \le x\right) \\
&= \frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}} d_{J_j,D_k}\right\} + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}} - 1\right] \\
&= \frac{10}{\sqrt{\pi}}\exp\left[-2\pi\lambda_J\int_0^{\infty}(1-r)rdr\right] + \frac{10}{\sqrt{\pi}}\left[\alpha + \log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}} - 1\right] \\
&= \frac{10}{\sqrt{\pi}}\left\{\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log\left(d_{S_k,D_k}\right) - \frac{C_{f,S_k}}{10} + \log_{10}\left[\frac{P_J}{P_{S_k}}\left(2^x-1\right)\right]\right\}. \qquad \text{(A-15)}
\end{aligned}$$

The expectation can be determined as

$$\begin{aligned}
&\mathbb{E}\{C_{S_k,D_k}\} \\
&= \int_{-\infty}^{+\infty} x\cdot\left(\frac{d}{dx}\left(\frac{10}{\sqrt{\pi}}\left\{\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log\left(d_{S_k,D_k}\right) - \frac{C_{f,S_k}}{10} + \log_{10}\left[\frac{P_J}{P_{S_k}}\left(2^x-1\right)\right]\right\}\right)\right)dx \\
&= \frac{\sqrt{\pi}}{10\ln 2}\int_{-\infty}^{+\infty} x\cdot\frac{2^x}{2^x-1}dx\cdot\left(\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log\left(d_{S_k,D_k}\right) - \frac{C_{f,S_k}}{10} + \log_{10}\left(\frac{P_J}{P_{S_k}}\right)\right)^{-1} \\
&= \frac{\sqrt{\pi}}{10\ln 2}\left(\alpha - 1 + \frac{C_{f,J_j}}{10} - \alpha\log\left(d_{S_k,D_k}\right) - \frac{C_{f,S_k}}{10} + \log_{10}\left(\frac{P_J}{P_{S_k}}\right)\right)^{-1} \qquad \text{(A-16)}
\end{aligned}$$

Thirdly, when $D_k \in \mathcal{S}$, there is

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= \mathbb{P}\left(\log_2\left(1 + \frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J G_{J_j,D_k}}\right) \leqslant x\right) = \mathbb{P}\left(\frac{P_{S_k}\frac{G_{S_k}G_{D_k}\lambda^2}{4\pi d_{S_k,D_k}^2}10^{-\frac{F_{S_k,D_k}^{rain}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\frac{G_{J_j}G_{D_k}\lambda^2}{4\pi d_{J_j,D_k}^2}10^{-\frac{F_{J_j,D_k}^{rain}}{10}}} \leqslant 2^x - 1\right) \\
&= \mathbb{P}\left(F_{S_k,D_k}^{rain} \geqslant -10\log_{10}\frac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}\frac{P_J}{d_{J_j,D_k}^2}10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right) \\
&= 1 - \mathbb{P}\left(F_{S_k,D_k}^{rain} < -10\log_{10}\frac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}\frac{P_J}{d_{J_j,D_k}^2}10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right). \qquad \text{(A-17)}
\end{aligned}$$

Note that F_{S_k,D_k}^{rain} and F_{J_j,D_k}^{rain} follow a Weibull distribution with parameters $\lambda = 1$ and $k = 1$, so we have

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= 1 - \left\{ 1 - \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{F_{J_j,D_k}^{rain}} \left\{ \exp\left[10\log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} \frac{P_J}{d_{J_j,D_k}^2} 10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}} \right] \right\} \right\} \right\} \\
&= \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{F_{J_j,D_k}^{rain}} \left\{ \exp\left[10\log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} \frac{P_J}{d_{J_j,D_k}^2} 10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}} \right] \right\} \right\} \\
&= \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{F_{J_j,D_k}^{rain}} \left\{ \exp\left[10\log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} \frac{P_J}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}} - F_{J_j,D_k}^{rain} \right] \right\} \right\}. \qquad \text{(A-18)}
\end{aligned}$$

Substituting $B = 10\log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} \frac{P_J}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}$ and $y = F_{J_j,D_k}^{rain}$ into (A-18) yields

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) = \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_y \{ \exp[B - y] \} \right\} = \frac{1}{2} \mathbb{E}_{\mathcal{J}} \left\{ \exp\left[10\log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} \frac{P_J}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}} \right] \right\} \\
&= \frac{1}{2} \mathbb{E}_{\mathcal{J}} \left\{ \exp\left[10\log_{10} \sum_{J \in \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2} + 10\log_{10} \frac{\left(2^x - 1\right) P_J d_{S_k,D_k}^2}{P_{S_k}} \right] \right\} \\
&= \frac{1}{2} \exp\left[10\log_{10} \frac{\left(2^x - 1\right) P_J d_{S_k,D_k}^2}{P_{S_k}} \right] \cdot \mathbb{E}_{\mathcal{J}} \left\{ \exp\left[10\log_{10} \sum_{J_j \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2} \right] \right\}. \qquad \text{(A-19)}
\end{aligned}$$

Expanding $\log_{10} \sum_{J_j \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2}$ by the Taylor formula (first-order expansion) yields

$$\log_{10} \sum_{J_j \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2} = \sum_{J_j \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2} - 1. \qquad \text{(A-20)}$$

Substituting (A-20) into (A-19), there is

$$\begin{aligned}
\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) &= \frac{1}{2} \exp\left[10\log_{10} \frac{\left(2^x - 1\right) P_J d_{S_k,D_k}^2}{P_{S_k}} \right] \cdot \mathbb{E}_{\mathcal{J}} \left\{ \exp\left[10 \sum_{J_j \in \mathcal{J}} \frac{1}{d_{J_j,D_k}^2} - 10 \right] \right\} \\
&= \frac{1}{2} \exp\left[10\log_{10} \frac{\left(2^x - 1\right) P_J d_{S_k,D_k}^2}{P_{S_k}} - 10 \right] \cdot \mathbb{E}_{\mathcal{J}} \left\{ \prod_{J_j \in \mathcal{J}} \exp\left(\frac{10}{d_{J_j,D_k}^2} \right) \right\}. \qquad \text{(A-21)}
\end{aligned}$$

Substituting (A-2) into (A-21), we have

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= \frac{1}{2}\exp\left[10\log_{10}\frac{\left(2^x-1\right)P_J d_{S_k,D_k}^2}{P_{S_k}} - 10\right] \cdot \exp\left[-2\pi\lambda_J \int_0^\infty \left(1-\exp\left(\frac{10}{r^2}\right)\right) r dr\right] \\
&= \frac{1}{2}\exp\left[10\log_{10}\frac{\left(2^x-1\right)P_J d_{S_k,D_k}^2}{P_{S_k}} - 10 - 10\pi\lambda_J\Gamma\left(\frac{1}{2}\right)\right] \qquad \text{(A-22)}
\end{aligned}$$

The expectation can be determined as

$$\begin{aligned}
\mathbb{E}(C_{S_k,D_k}) &= \int_{-\infty}^{+\infty} x \cdot \frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J d_{S_k,D_k}^2}{P_{S_k}} - 10 - 10\pi\lambda_J\Gamma(\frac{1}{2})\right] \times \frac{10\times 2^x \ln 2}{\ln 10(2^x-1)} dx \\
&= \frac{5\ln 2}{\ln 10}\int_{-\infty}^{+\infty} \frac{x\cdot 2^x}{2^x-1}\exp\left[10\log_{10}\frac{(2^x-1)P_J d_{S_k,D_k}^2}{P_{S_k}} - 10 - 10\pi\lambda_J\Gamma(\frac{1}{2})\right] dx \\
&= \frac{\ln 10}{10\ln 2}\cdot\frac{P_{S_k}}{P_J d_{S_k,D_k}^2}\cdot\exp\left(10 + 10\pi\lambda_J\Gamma\left(\frac{1}{2}\right)\right) \qquad \text{(A-23)}
\end{aligned}$$

Next, we derive the CDF and expectation of the eavesdropping link capacity $\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}$.

$$\begin{aligned}
&\mathbb{P}\left(\max_{E_i\in\mathcal{E}} C_{S_k,E_i} \leq x\right) \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\mathbb{P}\left[\log_2\left(1+\frac{P_{S_k}|h_{S_k,E_i}|^2/d_{S_k,E_i}^\alpha}{\sum_{J_j\in\mathcal{J}}P_J|h_{J_j,E_i}|^2/d_{J_j,E_i}^\alpha}\right)\leq x\right]\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\mathbb{P}\left[|h_{S_k,E_i}|^2 \leq \frac{(2^x-1)\sum_{J_j\in\mathcal{J}}P_J|h_{J_j,E_i}|^2/d_{J_j,E_i}^\alpha}{P_{S_k}/d_{S_k,E_i}^\alpha}\right]\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\left\{1-\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,E_i}}\left[\exp\left(-\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}P_J|h_{J_j,E_i}|^2/d_{J_j,E_i}^\alpha}{P_{S_k}/d_{S_k,E_i}^\alpha}\right)\right]\right\}\right\}\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\left\{1-\mathbb{E}_{\mathcal{J}}\prod_{J_j\in\mathcal{J}}\int_0^\infty \exp\left[-\left(\frac{(2^x-1)P_J/d_{J_j,E_i}^\alpha}{P_{S_k}/d_{S_k,E_i}^\alpha}+1\right)y\right]dy\right\}\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\left\{1-\mathbb{E}_{\mathcal{J}}\prod_{J_j\in\mathcal{J}}\frac{1}{\frac{(2^x-1)P_J/d_{J_j,E_i}^\alpha}{P_{S_k}/d_{S_k,E_i}^\alpha}+1}\right\}\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\left\{1-\exp\left[-\pi\lambda_J\int_0^\infty\frac{1}{1+\frac{P_{S_k}/d_{S_k,E_i}^\alpha}{(2^x-1)P_J/r^\alpha}}r dr\right]\right\}\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\left[1-\exp\left(-\pi\lambda_J\frac{(2^x-1)P_J}{P_{S_k}/d_{S_k,E_i}^\alpha}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\right)\right]\right\}. \qquad \text{(A-24)}
\end{aligned}$$

According to the PGFL of PPP of eavesdroppers, we have

$$\begin{aligned}\mathbb{P}\left(\max_{E_i\in\mathcal{E}} C_{S_k,E_i}\le x\right) &= \exp\left[-2\pi\lambda_E\int_0^\infty \exp\left(-\pi\lambda_J\frac{(2^x-1)P_J}{P_{S_k}/r^\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)r\,dr\right]\\ &= \exp\left[-2\pi\lambda_E\cdot\frac{1}{2\pi\lambda_J\left(\frac{P_{S_k}}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)}\right]\\ &= \exp\left[-\frac{\lambda_E}{\lambda_J\left(\frac{P_{S_k}}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)}\right]. \end{aligned} \tag{A-25}$$

The expectation can be determined as

$$\begin{aligned}&\mathbb{E}\left\{\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}\right\}\\ &= \int_0^\infty \mathbb{P}\left(\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}>x\right)dx = \int_0^\infty\left[1-\mathbb{P}\left(\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}\le x\right)\right]dx\\ &= \int_0^\infty \exp\left(\frac{\lambda_E}{\lambda_J}\left(\frac{P_S}{(2^x-1)P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\right)dx\\ &= \frac{\lambda_J}{\lambda_E}\left(\frac{P_J}{P_S}\right)^{\frac{2}{\alpha}}\frac{1}{\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)}\int_0^\infty\frac{1}{2^x-1}dx\cdot\ln(2)\\ &= \frac{1}{\frac{\lambda_E}{\lambda_J}\left(\frac{P_S}{P_J}\right)^{-\frac{2}{\alpha}}\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)\ln(2)}. \end{aligned} \tag{A-26}$$

This completes the proof. ■

APPENDIX B
PROOF OF LEMMA 2

Proof: When $S_k \in \mathcal{U}$, we determine the CDF of legitimate link capacity C_{S_k,D_k} and its expectation. Firstly, when $D_k \in \mathcal{G}$, there is

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) &= \mathbb{P}\left(\log_2\left(1+\frac{P_{S_k}10^{\frac{P^{loss}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\right)\leqslant x\right)\\ &= \mathbb{P}\left(\frac{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}+\frac{\delta_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\leqslant 2^x-1\right)\\ &= \mathbb{P}\left(\delta_{S_k,D_k}\leqslant 10\log_{10}\frac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right)\\ &= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}e^{-\frac{t^2}{2}}dt\right\}\right\},\end{aligned} \tag{B-1}$$

where $y = 10\log_{10}\dfrac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}$.

Then, we have

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right) &= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left[\frac{1}{2}\operatorname{erf}\left(10\log_{10}\frac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right)\right]\right\}\\ &= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left[\frac{1}{2}\operatorname{erf}\left(10\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{\left|h_{J_j,D_k}\right|^2}{d^{\alpha}_{J_j,D_k}}\right)+10\log_{10}\left(\frac{\left(2^x-1\right)P_J}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right)\right)\right]\right\}.\end{aligned} \tag{B-2}$$

Let $A = 10\log_{10}\left(\dfrac{\left(2^x-1\right)P_J}{P_{S_k}10^{\alpha\log\left(d_{S_k,D_k}\right)+\frac{C_{f,S_k}}{10}}}\right)$, there is

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right) = \frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left\{\operatorname{erf}\left[10\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{\left|h_{J_j,D_k}\right|^2}{d^{\alpha}_{J_j,D_k}}\right)+A\right]\right\}\right\}. \tag{B-3}$$

Expanding $\operatorname{erf}(z+A)$ by the Taylor formula (first-order expansion) yields

$$\operatorname{erf}(z+A) = \frac{2}{\sqrt{\pi}}(z+A) = \frac{2}{\sqrt{\pi}}\left[10\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{\left|h_{J_j,D_k}\right|^2}{d^{\alpha}_{J_j,D_k}}\right)+A\right]. \tag{B-4}$$

Substituting (B-4) into (B-3) yields

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right)\approx\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{h_{J_j,D_k}}\left\{\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{\left|h_{J_j,D_k}\right|^2}{d^{\alpha}_{J_j,D_k}}\right)\right\}\right\}+\frac{2A}{\sqrt{\pi}}. \tag{B-5}$$

According to Jensen's inequality, we have $\mathbb{E}_{X_i}\left(\log_{10}\sum_{i=1}^{n}X_i\right) \geqslant \log_{10}\mathbb{E}_{X_i}\left(\sum_{i=1}^{n}X_i\right) = \log_{10}\left(\sum_{i=1}^{n}\mathbb{E}_{X_i}\left(X_i\right)\right)$. Then, there is

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right) \geqslant \frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\left[\sum_{J_j\in\mathcal{J}}\left(\mathbb{E}_{h_{J_j,D_k}}\left(\frac{\left|h_{J_j,D_k}\right|^2}{d^{\alpha}_{J_j,D_k}}\right)\right)\right]\right\}+\frac{2A}{\sqrt{\pi}}$$
$$=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}\right)\right\}+\frac{2A}{\sqrt{\pi}}=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}\right)\right\}+\frac{2A}{\sqrt{\pi}}. \quad \text{(B-6)}$$

Expanding $\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}\right)$ by the Taylor formula (first-order expansion) yields

$$\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}\right)\approx\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}-1. \quad \text{(B-7)}$$

Then, we have

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right)\approx\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\sum_{J_j\in\mathcal{J}}\frac{1}{d^{\alpha}_{J_j,D_k}}-1\right\}+\frac{2A}{\sqrt{\pi}}$$
$$=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\sum_{J_j\in\mathcal{J}}\ln\left(e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}\right)-1\right\}+\frac{2A}{\sqrt{\pi}}=\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\ln\left(\prod_{J_j\in\mathcal{J}}e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}\right)-1\right\}+\frac{2A}{\sqrt{\pi}}. \quad \text{(B-8)}$$

Expanding $\ln\left(\prod_{J_j\in\mathcal{J}}e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}\right)$ by the Taylor formula (first-order expansion) yields

$$\ln\left(\prod_{J_j\in\mathcal{J}}e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}\right)-1\approx\prod_{J_j\in\mathcal{J}}e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}-1. \quad \text{(B-9)}$$

Then, we have

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right)\approx\frac{10}{\sqrt{\pi}}\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}}e^{\frac{1}{d^{\alpha}_{J_j,D_k}}}\right\}-\frac{10}{\sqrt{\pi}}+\frac{2A}{\sqrt{\pi}}. \quad \text{(B-10)}$$

Substituting (A-2) into (B-10), there is

$$\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right)=\frac{10}{\sqrt{\pi}}\exp\left[-2\pi\lambda_J\int_0^{\infty}\left(1-e^{\frac{1}{r^{\alpha}}}\right)rdr\right]-\frac{10}{\sqrt{\pi}}+\frac{2A}{\sqrt{\pi}}$$
$$=\frac{10}{\sqrt{\pi}}[\exp(-\pi\lambda_J)-1]+\frac{2A}{\sqrt{\pi}}. \quad \text{(B-11)}$$

The expectation can be determined as

$$\begin{aligned}
&\mathbb{E}(C_{S_k,D_k}) \\
&= \int_{-\infty}^{+\infty} x \cdot \frac{d}{dx}\left(\frac{10}{\sqrt{\pi}}\left[\exp(-\pi\lambda_J)-1\right] + \frac{2}{\sqrt{\pi}} \cdot 10\log_{10}\left(\frac{(2^x-1)P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right) dx \\
&= \int_{-\infty}^{+\infty} x \cdot \left(\frac{2}{\sqrt{\pi}} \cdot \frac{10}{\ln 10} \cdot \frac{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}{(2^x-1)P_J} \cdot \frac{P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}} \cdot 2^x \ln 2\right) dx \\
&= \frac{\sqrt{\pi}}{20\ln 2}\int_{-\infty}^{+\infty} x \cdot \frac{2^x}{2^x-1}dx \cdot \left(\frac{10}{\sqrt{\pi}}(e^{-\pi\lambda_J}-1) + \log_{10}\left(\frac{P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)^{-1} \\
&= \frac{\sqrt{\pi}}{20\ln 2}\left(\frac{10}{\sqrt{\pi}}(e^{-\pi\lambda_J}-1) + \log_{10}\left(\frac{P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)^{-1}. \qquad \text{(B-12)}
\end{aligned}$$

The derivation processes when $D_k \in \mathcal{U}$ and when $D_k \in \mathcal{S}$ are similar to those of equation (A-16) and equation (A-22), respectively. So, we omit them here.

Next, we determine the CDF and expectation of eavesdropping link capacity $\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}$. We have

$$\begin{aligned}
&\mathbb{P}\left(\max_{E_i\in\mathcal{E}} C_{S_k,E_i} \le x\right) = \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}} \mathbb{P}\left[C_{S_k,E_i} \le x\right]\right\} \\
&= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}} \frac{10}{\sqrt{\pi}}[\exp(-\pi\lambda_J)-1] + \frac{2\cdot 10\log_{10}\left(\frac{(2^x-1)P_J}{P_{S_k}10^{\alpha\log\left(d_{S_k,E_i}\right)+\frac{C_{f,S_k}}{10}}}\right)}{\sqrt{\pi}}\right\}. \qquad \text{(B-13)}
\end{aligned}$$

According to the PGFL of PPP of eavesdroppers, there is

$$\mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}} f\left(z_{E_i}\right)\right\} = \exp\left[-\lambda_E\int_{R^2}\left(1-f\left(z_{E_i}\right)\right)dz_{E_i}\right] = \exp\left[-2\pi\lambda_E\int_0^{\infty}(1-f(r))rdr\right]. \qquad \text{(B-14)}$$

Then, we can obtain

$$\mathbb{P}\left(\max_{E_i\in\mathcal{E}} C_{S_k,E_i} \le x\right) = \exp\left(-\lambda_E + \frac{10}{\sqrt{\pi}}\lambda_E\exp(-\pi\lambda_J) - \frac{10}{\sqrt{\pi}}\lambda_E - \frac{2A}{\sqrt{\pi}}\lambda_E\right). \qquad \text{(B-15)}$$

The expectation can be determined as

$$\begin{aligned}
&\mathbb{E}\left[\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}\right] \\
&=\int_{-\infty}^{+\infty} x\cdot\frac{d}{dx}\left(\exp\left\{-\lambda_E+\frac{10}{\sqrt{\pi}}\lambda_E e^{-\pi\lambda_J}-\frac{10}{\sqrt{\pi}}\lambda_E-\frac{20}{\sqrt{\pi}}\lambda_E\log_{10}\left(\frac{(2^x-1)P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right\}\right)dx \\
&=\frac{\sqrt{\pi}}{20\lambda_E\ln(2)}\left(-\lambda_E+\frac{10}{\sqrt{\pi}}\lambda_E e^{-\pi\lambda_J}-\frac{10}{\sqrt{\pi}}\lambda_E+\log_{10}\left(\frac{P_J}{P_{S_k}10^{\alpha\log(d_{S_k,D_k})+\frac{C_{f,S_k}}{10}}}\right)\right)^{-1}.
\end{aligned} \tag{B-16}$$

This completes the proof. ■

APPENDIX C
PROOF OF LEMMA 3

Proof: When $S_k\in\mathcal{S}$, we determine the CDF and expectation of legitimate link capacity C_{S_k,D_k}. Firstly, when $D_k\in\mathcal{G}$, there is

$$\begin{aligned}
\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right)&=\mathbb{P}\left[\log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\right)\leqslant x\right] \\
&=\mathbb{P}\left[\frac{P_{S_k}\frac{G^2\lambda^2}{4\pi d^2_{S_k,D_k}}10^{-\frac{F^{rain}_{S_k,D_k}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}\leqslant 2^x-1\right] \\
&=\mathbb{P}\left[F^{rain}_{S_k,D_k}\geqslant -10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right)\right] \\
&=1-\mathbb{P}\left[F^{rain}_{S_k,D_k}<-10\log_{10}\left(\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}P_J\left|h_{J_j,D_k}\right|^2/d^{\alpha}_{J_j,D_k}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right)\right].
\end{aligned} \tag{C-1}$$

Given that F_{S_k,D_k}^{rain} follows a Weibull distribution with $\lambda = 1$ and $k = 1$, we have

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= 1 - \left\{ 1 - \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{h_{J_j,D_k}} \left\{ \exp\left(10 \log_{10} \left(\frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} P_J \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{\frac{P_{S_k} G^2 \lambda^2}{4\pi d_{S_k,D_k}^2}} \right) \right) \right\} \right\} \right\} \\
&= \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{h_{J_j,D_k}} \left\{ \exp\left(10 \log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} P_J \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{\frac{P_{S_k} G^2 \lambda^2}{4\pi d_{S_k,D_k}^2}} \right) \right\} \right\}. \qquad \text{(C-2)}
\end{aligned}$$

Using first order taylor approximation, there is

$$10 \log_{10} \frac{\left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} P_J \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{\frac{P_{S_k} G^2 \lambda^2}{4\pi d_{S_k,D_k}^2}} \approx \frac{10 \left(2^x - 1\right) \sum_{J_j \in \mathcal{J}} P_J \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{\frac{P_{S_k} G^2 \lambda^2}{4\pi d_{S_k,D_k}^2}} - 10. \qquad \text{(C-3)}$$

Then, we can obtain

$$\begin{aligned}
\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) &\approx \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{h_{J_j,D_k}} \left\{ \exp\left(\frac{10 \left(2^x - 1\right) P_J \sum_{J_j \in \mathcal{J}} \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2} - 10 \right) \right\} \right\} \\
&= \frac{1}{e^{10}} \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{h_{J_j,D_k}} \left\{ \exp\left(\frac{10 \left(2^x - 1\right) P_J \sum_{J_j \in \mathcal{J}} \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2} \right) \right\} \right\} \\
&= \frac{1}{e^{10}} \mathbb{E}_{\mathcal{J}} \left\{ \prod_{J_j \in \mathcal{J}} \mathbb{E}_{h_{J_j,D_k}} \left\{ \exp\left(\frac{10 \left(2^x - 1\right) P_J \left| h_{J_j,D_k} \right|^2 / d_{J_j,D_k}^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2} \right) \right\} \right\} \\
&= \frac{1}{e^{10}} \mathbb{E}_{\mathcal{J}} \left\{ \prod_{J_j \in \mathcal{J}} \left\{ \int_0^{\infty} \exp\left[\left(\frac{10 \left(2^x - 1\right) P_J / d_{J_j,D_k}^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2} - 1 \right) y \right] dy \right\} \right\} \\
&= \frac{1}{e^{10}} \mathbb{E}_{\mathcal{J}} \left\{ \prod_{J_j \in \mathcal{J}} \frac{1}{1 - \frac{10(2^x - 1) P_J / d_{J_j,D_k}^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2}} \right\}. \qquad \text{(C-4)}
\end{aligned}$$

Substituting (A-2) into (C-4) yields

$$\begin{aligned}
\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) &= \frac{1}{e^{10}} \exp\left[-2\pi\lambda_J \int_0^{\infty} \left(1 - \frac{1}{1 - \frac{10(2^x - 1) P_J / r^{\alpha}}{P_{S_k} G^2 \lambda^2 / 4\pi d_{S_k,D_k}^2}} \right) r dr \right] \\
&= \frac{1}{e^{10}} \exp\left[-2\pi\lambda_J \frac{4\pi d_{S_k,D_k}^2 \cdot 10 \left(2^x - 1\right) P_J}{P_{S_k} G^2 \lambda^2} \cdot \frac{1}{1 - \alpha} \Gamma(1 - \alpha) \right] \\
&= \exp\left[-10 - \frac{80\pi^2 \lambda_J d_{S_k,D_k}^2 \left(2^x - 1\right) P_J}{P_{S_k} G^2 \lambda^2 (1 - \alpha)} \Gamma(1 - \alpha) \right]. \qquad \text{(C-5)}
\end{aligned}$$

The expectation is determined as

$$\begin{aligned}\mathbb{E}(C_{S_k,D_k}) &= \int_{-\infty}^{+\infty} x \cdot \left(\frac{d}{dx} \exp\left[-10\right] \cdot \exp\left[-\frac{80\pi^2 \lambda_J d_{S_k,D_k}^2 (2^x-1) P_J}{P_{S_k} G^2 \lambda^2 (1-\alpha)} \Gamma(1-\alpha) \right] \right) dx \\ &= \int_{-\infty}^{+\infty} x \cdot \exp\left[-10\right] \cdot \exp\left[-\frac{80\pi^2 \lambda_J d_{S_k,D_k}^2 (2^x-1) P_J}{P_{S_k} G^2 \lambda^2 (1-\alpha)} \Gamma(1-\alpha) \right] \cdot \left(-\frac{80\pi^2 \lambda_J d_{S_k,D_k}^2 P_J}{P_{S_k} G^2 \lambda^2 (1-\alpha)} \Gamma(1-\alpha) \cdot 2^x \ln 2 \right) dx \\ &= \frac{P_{S_k} G^2 \lambda^2 (1-\alpha) \Gamma(1-\alpha)}{80\pi^2 \lambda_J d_{S_k,D_k}^2 \ln 2 P_J} \end{aligned} \tag{C-6}$$

Secondly, when $D_k \in \mathcal{U}$, we have

$$\begin{aligned}&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) = \mathbb{P}\left[\log_2 \left(1 + \frac{P_{S_k} G_{S_k,D_k}}{\sum_{J_j \in \mathcal{J}} P_J 10^{\frac{P_{J_j,D_k}^{loss}}{10}}} \right) \leqslant x \right] \\ &= \mathbb{P}\left[\frac{P_{S_k} \frac{G^2\lambda^2}{4\pi d_{S_k,D_k}^2} 10^{-\frac{F_{S_k,D_k}^{rain}}{10}}}{\sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10} + \frac{\delta_{J_j,D_k}}{10}}} \leqslant 2^x - 1 \right] \\ &= \mathbb{P}\left[10^{-\frac{F_{S_k,D_k}^{rain}}{10}} \leqslant \frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10} + \frac{\delta_{J_j,D_k}}{10}}}{P_{S_k} \frac{G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} \right] \\ &= \mathbb{P}\left[F_{S_k,D_k}^{rain} \geqslant -10 \log_{10} \left(\frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10} + \frac{\delta_{J_j,D_k}}{10}}}{P_{S_k} \frac{G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} \right) \right] \\ &= 1 - \mathbb{P}\left[F_{S_k,D_k}^{rain} < -10 \log_{10} \left(\frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10} + \frac{\delta_{J_j,D_k}}{10}}}{P_{S_k} \frac{G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} \right) \right]. \end{aligned} \tag{C-7}$$

Given that F_{S_k,D_k}^{rain} follow a Weibull distribution with parameters $\lambda = 1$ and $k = 1$, there is

$$\begin{aligned}&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\ &= 1 - \left\{ 1 - \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{\delta_{J_j,D_k}} \left\{ \exp\left[10 \log_{10} \frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10} + \frac{\delta_{J_j,D_k}}{10}}}{\frac{P_{S_k} G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} \right] \right\} \right\} \right\} \\ &= 1 - \left\{ 1 - \mathbb{E}_{\mathcal{J}} \left\{ \mathbb{E}_{\delta_{J_j,D_k}} \left\{ \exp\left[10 \log_{10} \frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10}}}{\frac{P_{S_k} G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} + 10 \log_{10} 10^{\frac{\delta_{J_j,D_k}}{10}} \right] \right\} \right\} \right\} \\ &= \mathbb{E}_{\mathcal{J}} \left\{ \exp\left[10 \log_{10} \frac{(2^x-1) \sum_{J_j \in \mathcal{J}} P_J 10^{\alpha \log\left(d_{J_j,D_k}\right) + \frac{C_{f,J_j}}{10}}}{\frac{P_{S_k} G^2\lambda^2}{4\pi d_{S_k,D_k}^2}} \right] \cdot \mathbb{E}_{\delta_{J_j,D_k}} \left\{ e^{\delta_{J_j,D_k}} \right\} \right\}. \end{aligned} \tag{C-8}$$

Since δ_{J_j,D_k} follows a standard normal distribution, then we have

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) &= \mathbb{E}_{\mathcal{J}}\left\{\exp\left[10\log_{10}\frac{\left(2^x-1\right)\sum_{J_j\in\mathcal{J}}P_J 10^{\alpha\log\left(d_{J_j,D_k}\right)+\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\right\}\\ &= \mathbb{E}_{\mathcal{J}}\left\{\exp\left[10\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}+10\log_{10}\sum_{J_j\in\mathcal{J}}10^{\alpha\log\left(d_{J_j,D_k}\right)}\right]\right\}\\ &= \exp\left[10\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\mathbb{E}_{\mathcal{J}}\left\{\exp\left[\sum_{J_j\in\mathcal{J}}10^{\alpha\log\left(d_{J_j,D_k}\right)}\right]\right\}\\ &= \exp\left[10\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\mathbb{E}_{\mathcal{J}}\left\{\prod_{J_j\in\mathcal{J}}\exp\left[10^{\alpha\log\left(d_{J_j,D_k}\right)}\right]\right\}. \quad \text{(C-9)}\end{aligned}$$

Substituting (A-2) into (C-9) yields

$$\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) = \exp\left[5\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\exp\left[-2\pi\lambda_J\int_0^\infty\left(1-\exp\left(10^{\alpha\log(r)}\right)\right)rdr\right]. \quad \text{(C-10)}$$

The expectation is given by

$$\begin{aligned}\mathbb{E}(C_{S_k,D_k}) &= \int_{-\infty}^{+\infty}x\cdot\frac{d}{dx}\left(\exp\left[5\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\exp\left[-2\pi\lambda_J\int_0^\infty(1-\exp(10^{\alpha\log(r)}))rdr\right]\right)dx\\ &= \int_{-\infty}^{+\infty}x\cdot\left(\exp\left[-2\pi\lambda_J\int_0^\infty(1-\exp(10^{\alpha\log(r)}))rdr\right]\cdot\frac{d}{dx}\exp\left[5\log_{10}\frac{\left(2^x-1\right)P_J 10^{\frac{C_{f,J_j}}{10}}}{\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}}}\right]\right)dx\\ &= \frac{\ln 10}{10\ln 2}\cdot\frac{P_{S_k}G^2\lambda^2}{4\pi d^2_{S_k,D_k}P_J 10^{\frac{C_{f,J_j}}{10}}}. \quad \text{(C-11)}\end{aligned}$$

Thirdly, when $D_k \in \mathcal{S}$, there is

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= \mathbb{P}\left[\log_2\left(1+\frac{P_{S_k}G_{S_k,D_k}}{\sum_{J_j\in\mathcal{J}}P_J G_{J_j,D_k}}\right) \leqslant x\right] \\
&= \mathbb{P}\left[\frac{P_{S_k}\frac{G^2\lambda^2}{4\pi d_{S_k,D_k}^2}10^{-\frac{F_{S_k,D_k}^{rain}}{10}}}{\sum_{J_j\in\mathcal{J}}P_J\frac{G^2\lambda^2}{4\pi d_{J_j,D_k}^2}10^{-\frac{F_{J_j,D_k}^{rain}}{10}}} \leqslant 2^x-1\right] \\
&= \mathbb{P}\left\{F_{S_k,D_k}^{rain} \geqslant -10\log_{10}\left[\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J 10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right]\right\} \\
&= 1-\mathbb{P}\left\{F_{S_k,D_k}^{rain} < -10\log_{10}\left[\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J 10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right]\right\}.
\end{aligned} \tag{C-12}$$

Given that F_{S_k,D_k}^{rain} and F_{J_j,D_k}^{rain} follow a Weibull distribution with $\lambda=1$ and $k=1$, we have

$$\begin{aligned}
&\mathbb{P}\left(C_{S_k,D_k} \leqslant x\right) \\
&= 1-\left\{1-\mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{F_{J_j,D_k}^{rain}}\left\{\exp\left[10\log_{10}\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J 10^{-\frac{F_{J_j,D_k}^{rain}}{10}}}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}\right]\right\}\right\}\right\} \\
&= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_{F_{J_j,D_k}^{rain}}\left\{\exp\left[10\log_{10}\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J}{d_{J_j,D_k}^2}}{\frac{P_{S_k}}{d_{S_k,D_k}^2}}-F_{J_j,D_k}^{rain}\right]\right\}\right\}.
\end{aligned} \tag{C-13}$$

Let $A = 10\log_{10}\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J}{d^2_{J_j,D_k}}}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}$ and $y = F^{rain}_{J_j,D_k}$, then we can obtain

$$\begin{aligned}\mathbb{P}\left(C_{S_k,D_k}\leqslant x\right) &= \mathbb{E}_{\mathcal{J}}\left\{\mathbb{E}_y\{\exp[A-y]\}\right\} = \frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\exp\left(10\log_{10}\frac{(2^x-1)\sum_{J_j\in\mathcal{J}}\frac{P_J}{d^2_{J_j,D_k}}}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}\right)\right\}\\ &= \frac{1}{2}\mathbb{E}_{\mathcal{J}}\left\{\exp\left[10\log_{10}\frac{(2^x-1)P_J}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}+10\log_{10}\left(\sum_{J_j\in\mathcal{J}}\frac{1}{d^2_{J_j,D_k}}\right)\right]\right\}\\ &= \frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}\right]\cdot\mathbb{E}_{\mathcal{J}}\left[\exp\left(10\log_{10}\sum_{J_j\in\mathcal{J}}\frac{1}{d^2_{J_j,D_k}}\right)\right]. \quad \text{(C-14)}\end{aligned}$$

The expectation is given by

$$\begin{aligned}&\mathbb{E}(C_{S_k,D_k})\\ &= \int_{-\infty}^{+\infty} x\cdot\frac{d}{dx}\left(\frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}\right]\cdot\mathbb{E}_{\mathcal{J}}\left[\exp\left(10\log_{10}\sum_{J_j\in\mathcal{J}}\frac{1}{d^2_{J_j,D_k}}\right)\right]\right)dx\\ &= \int_{-\infty}^{+\infty} x\cdot\mathbb{E}_{\mathcal{J}}\left[\exp\left(10\log_{10}\sum_{J_j\in\mathcal{J}}\frac{1}{d^2_{J_j,D_k}}\right)\right]\cdot\frac{1}{2}\exp\left[10\log_{10}\frac{(2^x-1)P_J}{\frac{P_{S_k}}{d^2_{S_k,D_k}}}\right]\cdot\frac{10\cdot 2^x\ln 2}{(2^x-1)\ln 10}dx\\ &= \frac{\ln(10)}{10\ln(2)}\cdot\frac{P_{S_k}}{P_J d^2_{S_k,D_k}}\cdot\frac{2}{\mathbb{E}_J\left[\exp\left(10\log_{10}\sum_{J_j\in J}\frac{1}{d^2_{J_j,D_k}}\right)\right]}. \quad \text{(C-15)}\end{aligned}$$

Similarly, the CDF and expectation of eavesdropping link capacity $\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}$ can be determined as

$$\begin{aligned}\mathbb{P}\left(\max_{E_i\in\mathcal{E}}C_{S_k,E_i}\leq x\right) &= \mathbb{E}_{\mathcal{E}}\left\{\prod_{E_i\in\mathcal{E}}\mathbb{P}\left[C_{S_k,E_i}\leq x\right]\right\}\\ &= \exp\left[-10-\frac{80\pi^2\lambda_J d^2_{S_k,E_i}(2^x-1)P_J}{P_{S_k}G^2\lambda^2(1-\alpha)}\Gamma(1-\alpha)\right]. \quad \text{(C-16)}\end{aligned}$$

$$\mathbb{E}\left\{\max_{E_i\in\mathcal{E}}\{C_{S_k,E_i}\}\right\} = \frac{P_{S_k}G^2\lambda^2(1-\alpha)\Gamma(1-\alpha)}{80\pi^2\lambda_J d^2_{S_k,E_i}\ln(2)P_J}. \quad \text{(C-17)}$$

This completes the proof. ■

APPENDIX D
PROOF OF LEMMA 4

Chernoff-Hoeffding Bound: Let $X_1, \ldots, X_n$ be random variables with common range $[0, 1]$ such that $\mathbb{E}\{X_t | X_1, \ldots, X_{t-1}\} = \mu$. Let $S_n = X_1 + \cdots + X_n$. Then, for all $a \geq 0$, the following inequalities hold:

$$\mathbb{P}\{S_n \geq n\mu + a\} \leq e^{-2a^2/n} \quad \text{and} \quad \mathbb{P}\{S_n \leq n\mu - a\} \leq e^{-2a^2/n}.$$

We apply the Chernoff-Hoeffding bound for the probabilities $\mathbb{P}\left\{\bar{Z}_i(\theta) \leq Z_i(\theta) - U_i(\theta)\right\}$ and $\mathbb{P}\left\{\bar{Z}_i(\theta) \geq Z_i(\theta) + U_i(\theta)\right\}$, respectively. When confidence interval $U_i(\theta) = \sqrt{\frac{2\ln\theta}{\beta_i(\theta)}}$, there are

$$\mathbb{P}\left\{\bar{Z}_i(\theta) \leq Z_i(\theta) - U_i(\theta)\right\} \leq e^{-4\ln\theta} = \frac{1}{\theta^4}, \tag{D-1}$$

$$\mathbb{P}\left\{\bar{Z}_i(\theta) \geq Z_i(\theta) + U_i(\theta)\right\} \leq e^{-4\ln\theta} = \frac{1}{\theta^4}. \tag{D-2}$$

Then, we can obtain

$$\begin{aligned}
&\mathbb{P}\left\{\bar{Z}_i(\theta) \in [Z_i(\theta) - U_i(\theta), Z_i(\theta) + U_i(\theta)]\right\} \\
&= \mathbb{P}\left\{[\bar{Z}_i(\theta) \geq Z_i(\theta) - U_i(\theta)] \wedge [\bar{Z}_i(\theta) \leq Z_i(\theta) + U_i(\theta)]\right\} \\
&= 1 - \mathbb{P}\left\{\bar{Z}_i(\theta) \leq Z_i(\theta) - U_i(\theta)] - \mathbb{P}[\bar{Z}_i(\theta) \geq Z_i(\theta) + U_i(\theta)\right\} \\
&= 1 - \frac{2}{\theta^4}
\end{aligned} \tag{D-3}$$

This completes the proof. ■

APPENDIX E
PROOF OF LEMMA 5

Proof: The expression (36) can be rewritten as (by replacing each a with i on the right-hand side)

$$\sum_{1 \leq a < i \leq N} \frac{\Delta_{i-1,i}}{\Delta_{a,i}^2} \leq 2 \sum_{i=2}^{N} \frac{1}{\Delta_{i-1,i}}.$$

For a fixed $a \in \{1, 2, \cdots, N\}$, we write $\sum_{i:i>a} \frac{1}{\Delta_{a,i}^2}$ as follows:

$$\begin{aligned}\sum_{i:i>a} \frac{1}{\Delta_{a,i}^2} &= \sum_{i=2}^{N} \mathbf{1}\{i > a\} \frac{1}{\Delta_{a,i}^2} = \int_{x=0}^{\infty} \#\{i : i > a, \Delta_{a,i}^{-2} \geq x\} \mathrm{d}x \\ &= \int_{x=0}^{\infty} \#\{i > a, \Delta_{a,i} \leq x^{-1/2}\} \mathrm{d}x = -2 \int_{y=\infty}^{0} \#\{i > a, \Delta_{a,i} \leq y\} y^{-3} \mathrm{d}y \\ &\quad \text{(Changing the variable of integration } x^{-1/2} = y\text{)} \\ &= 2 \int_{y=0}^{\infty} \#\{i > a, \Delta_{a,i} \leq y\} y^{-3} \mathrm{d}y. \end{aligned} \tag{E-1}$$

For $i \geq 0$ and $y \geq 0$, let $a_y(i)$ be the minimum $a \leq i$ such that $\Delta_{a,i}$ is no more than y. That is,

$$a_y(i) := \arg\min\{a : a \leq i, \Delta_{a,i} \leq y\}. \tag{E-2}$$

Then, we have

$$\begin{aligned}\sum_{1 \leq a < i \leq N} \frac{\Delta_{a,a+1}}{\Delta_{a,i}^2} &= \sum_{i=2}^{N} \sum_{a=1}^{i-1} \frac{\Delta_{a,a+1}}{\Delta_{a,i}^2} = \sum_{a=1}^{N-1} \Delta_{a,a+1} \sum_{i:i>a} \frac{1}{\Delta_{a,i}^2} \\ &= 2 \sum_{a=1}^{N-1} \Delta_{a,a+1} \left(\int_{y=0}^{\infty} \#\{i : i > a, \Delta_{a,i} \leq y\} y^{-3} \mathrm{d}y \right) \\ &\quad \text{(From Eq. (E-1))} \\ &= 2 \int_{y=0}^{\infty} y^{-3} \left(\sum_{a=1}^{N-1} \Delta_{a,a+1} \cdot \#\{i > a, \Delta_{a,i} \leq y\} \right) dy \\ &\quad \text{(Changing the order of integration and summation)} \\ &= 2 \int_{y=0}^{\infty} y^{-3} \left(\sum_{a=1}^{N-1} \Delta_{a,a+1} \sum_{i=a+1}^{N} \mathbf{1}\{i > a, \Delta_{a,i} \leq y\} \right) dy \\ &\quad \text{(Expanding} \#\{\cdot\} \text{into sum of } \mathbf{1}\{\cdot\}\text{)} \\ &= 2 \int_{y=0}^{\infty} y^{-3} \left(\sum_{i=2}^{N} \sum_{a=1}^{i-1} \Delta_{a,a+1} \mathbf{1}\{i > a, \Delta_{a,i} \leq y\} \right) dy. \end{aligned} \tag{E-3}$$

According to (E-2), there is

$$\begin{aligned}\sum_{1\le a<i\le N}\frac{\Delta_{a,a+1}}{\Delta_{a,i}^2}&=2\int_{y=0}^{\infty}y^{-3}\left(\sum_{i=2}^{N}\sum_{a=a_y(i)}^{i-1}\Delta_{a,a+1}\right)dy=2\int_{y=0}^{\infty}y^{-3}\left(\sum_{i=2}^{N}(\mu_{a_y(i)}-\mu_i)\right)dy\\&=2\sum_{i=2}^{N}\int_{y=0}^{\infty}y^{-3}(\mu_{a_y(i)}-\mu_i)dy\\&\text{(Changing the order of summation and integration)}\\&=2\sum_{i=2}^{N}\int_{y=\Delta_{i-1,i}}^{\infty}y^{-3}(\mu_{a_y(i)}-\mu_i)dy\\&\text{(For } y<\Delta_{i-1,i}, a_y(i)=i, \text{ so the integrand is equal to } 0)\\&\le 2\sum_{i=2}^{N}\int_{y=\Delta_{i-1,i}}^{\infty}y^{-3}\cdot y dy\\&\text{(Since}\quad \mu_{a_y(i)}-\mu_i\le y)\\&=2\sum_{i=2}^{N}\int_{y=\Delta_{i-1,i}}^{\infty}y^{-2}dy=2\sum_{i=2}^{N}\frac{1}{\Delta_{i-1,i}}.\end{aligned}\tag{E-4}$$

This completes the proof. ■

APPENDIX F
PROOF OF LEMMA 6

Proof: We introduce $\kappa(\theta)$ as a counter to record the number of times that non-optimal arms have been selected until the θ-th round. That is, in each round, if the optimal route is selected, $\kappa(\theta)$ will not change; else, $\kappa(\theta)=\kappa(\theta-1)+1$. We use the notation $I(\theta)\in\{0,1\}$ to denote the indicator, in which $I(\theta)=1$ means the corresponding counter $\kappa(\theta)$ is incremented in round θ, and $I(\theta)=0$ otherwise. Based on this, we have the following result:

$$\begin{aligned}\kappa(\Theta(W_B))&=\sum_{\theta=2}^{\Theta(W_B)}\mathbb{I}\{I(\theta)=1\}\le\lambda+\sum_{\theta=1}^{\Theta(W_B)}\mathbb{I}\{\widehat{Z}_{\pi^\theta}(\theta)\ge\widehat{Z}_{\pi^*}(\theta),\kappa(\theta)\ge\lambda\}\\&\le\lambda+\sum_{\theta=1}^{\Theta(W_B)}\sum_{\beta_{\pi^\theta}(\theta)=\lambda}^{\theta}\sum_{\beta_{\pi^*}(\theta)=1}^{\theta}\mathbb{I}\{\widehat{Z}_{\pi^\theta}(\theta)\ge\widehat{Z}_{\pi^*}(\theta)\}.\end{aligned}\tag{F-1}$$

If the event $\bar{Z}_{\pi^\theta}(\theta)+U_{\pi^\theta}(\theta)\ge\bar{Z}_{\pi^*}(\theta)+U_{\pi^*}(\theta)$ happens, we can know that at least one of the following three cases must be true (based on the proof by contradiction): (a) $\bar{Z}_{\pi^*}(\theta)\le\widehat{Z}_{\pi^*}(\theta)-U_{\pi^*}(\theta)$; (b) $\bar{Z}_{\pi^\theta}(\theta)\ge\widehat{Z}_{\pi^\theta}(\theta)+U_{\pi^\theta}(\theta)$; (c) $\widehat{Z}_{\pi^*}(\theta)<\widehat{Z}_{\pi^\theta}(\theta)+2U_{\pi^\theta}(\theta)$. According to

the Chernoff-Hoeffding bound, the upper bound of the probability that case (a) happens is given by

$$\mathbb{P}\{\bar{Z}_{\pi^*}(\theta) \leq \widehat{Z}_{\pi^*}(\theta) - U_{\pi^*}(\theta)\} \leq e^{-2\beta_{\pi^*}(\theta)U_{\pi^*}(\theta)^2} = \theta^{-2}. \tag{F-2}$$

Similarly, we can prove the probability that case (b) happens has an identical upper bound. Regarding case (c), we have

$$\widehat{Z}_{\pi^*}(\theta) - \widehat{Z}_{\pi^\theta}(\theta) - 2U_{\pi^\theta}(\theta) < \Delta_{\min} - 2\sqrt{\frac{\ln\theta}{\beta_{\pi^\theta}(\theta)}} < \Delta_{\min} - 2\sqrt{\frac{\ln\Theta(W_B)}{\lambda}} < 0. \tag{F-3}$$

We can see that if $\lambda \geq \frac{4\ln\Theta(W_B)}{\Delta_{\min}^2}$, case (c) will not hold.

Substituting the above analysis into Eq. (F-1) yields

$$\kappa(\Theta(W_B)) \leq \frac{4\ln\Theta(W_B)}{\Delta_{\min}^2} + \sum_{\theta=1}^{\Theta(W_B)} \theta^2 2\theta^{-2} = \frac{4\ln\Theta(W_B)}{\Delta_{\min}^2} + 2\Theta(W_B). \tag{F-4}$$

Comparing $\Theta(W_B)$ and $\Theta^*(W_B)$, we have

$$\begin{aligned}\Theta(W_B) &\leq \Theta^*(W_B) + W_{\max}T[\kappa(\Theta(W_B))] \\ &\leq \frac{B}{W^*} + \frac{W_{\max}}{W_{\min}}\left\{\frac{4\ln\Theta(W_B)}{\Delta_{\min}^2} + 2\Theta(W_B)\right\}.\end{aligned} \tag{F-5}$$

Since $\ln\Theta(W_B) \leq \Theta(W_B) - 1$ for $\forall\Theta(W_B) > 0$, there is

$$\begin{aligned}\Theta(W_B) &\leq \frac{W_B}{W^*} + \frac{W_{\max}}{W_{\min}}\left\{\frac{4(\Theta(W_B)-1)}{\Delta_{\min}^2} + 2\Theta(W_B)\right\} \\ &\Rightarrow \Theta(W_B) \leq \frac{W_B\Delta_{\min}^2 W_{\min} - 4W_{\max}W^*}{W^*[\Delta_{\min}^2 W_{\min} - W_{\max}(4+2\Delta_{\min}^2)]}.\end{aligned} \tag{F-6}$$

This completes the proof. ■