Online Joint Power Allocation and Task Scheduling for LEO Satellite Networks

Zheyuan Li*, Lijun He†, Juncheng Wang‡, Ziye Jia§, Yanting Wang*, and Chau Yuen¶
* School of Software, Northwestern Polytechnical University, China
† School of Information and Control Engineering, China University of Mining and Technology, China
‡ Department of Computer Science, Hong Kong Baptist University, China
§ Ministry of Industry and Information Technology, Nanjing University of Aeronautics and Astronautics, China
¶ School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore
Email: schwarz@mail.nwpu.edu.cn, lijunhe.xd@gmail.com, jcwang@comp.hkbu.edu.hk, jiaziye@nuaa.edu.cn, yantingwang@nwpu.edu.cn, chau.yuen@ntu.edu.sg

2025 IEEE Wireless Communications and Networking Conference (WCNC) | 979-8-3503-6836-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/WCNC61545.2025.10978643

***Abstract*—The excessive proliferation of Low Earth Orbit (LEO) satellites inescapably bring the explosive growth of space data in LEO Satellite Networks (LSNs). Meanwhile, the stochastic arrivals of space data together with the time-varying satellite-ground links in LSNs pose significant challenges for offloading a large volume of space data from LSNs to ground stations. To circumvent these challenges, we systematically study the energy-constrained online data offloading problem to jointly optimize power allocation and task scheduling for LSNs. First, we leverage Lyapunov optimization to decouple our formulated long-term stochastic joint optimization problem into a set of per-time-slot subproblems. Then, each subproblem is decoupled into a task scheduling problem and a power allocation problem. Next, we derive the optimal solution to the power allocation problem and propose a multi-armed bandit based quasi-optimal solution to the task scheduling problem. Finally, extensive simulation results show that our proposed algorithm has superior performance over the state-of-the-art solutions.**

***Index Terms*—LEO Satellite Networks, online data offloading, power allocation, task scheduling.**

I. Introduction

The rapid growth of wireless networking services constantly put forward the urgent demands for seamless global coverage and higher data transmission rate in the future 6G communication networks [1]. It is unfeasible to deploy the expensive base stations and other communication facilities to achieve seamless global coverage with high-speed data transmission. Due to the wider coverage and higher communication rate, Low Earth Orbit Satellite Networks (LSNs) have been considered as a promising solution to compensate conventional terrestrial networks [2]–[4]. Currently, LSNs provide wide coverage services in a broad range of sectors, such as military and environmental monitoring [5]. In this context, the excessive proliferation of Low Earth Orbit (LEO) satellites brings the explosive growth of space data in LSNs [6]. Typically, the majority of these collected space data needs to be offloaded from the LEO satellites to a set of Ground Stations (GSs) for further processing and analysis. Nevertheless, the high-speed motion of LEO satellites makes LSNs line-of-sight communication networks in nature. This restricts the data offloading capability of LSNs. In brief, how to design an efficient data offloading scheme to alleviate the contradiction between the explosive growth of space data and the limited intermittent communication resources is the key to improve the data offloading efficiency of LSNs.

Early works on data offloading for satellite networks focused on *offline* schemes [7]–[9], which cannot handle stochastic arrivals of space data as well as time-varying network resources. More recent works studied *online* data offloading problems to capture the stochastic arrivals of space data [10] or to circumvent the unknown environment dynamics of LSNs [11]–[13]. Specifically, the data offloading scheme in [10] ignored energy consumption and focused on scheduling the stable inter-satellite links in data relay satellite networks instead of the time-varying Satellite-Ground Links (SGLs) in LSNs. The works of [11] and [12] studied the total task latency minimization under the energy consumption constraint, without considering the metric of the total offloaded data amount. Our prior work [13] online scheduled time-invariant SGLs for space-air-ground integrated networks. These motivate us to pose the following key question: How to *online* schedule time-varying SGLs in LSNs to accommodate the stochastic arrivals of space data under the energy consumption constraint?

In this work, we jointly optimize power allocation and task scheduling for LSNs in an online manner to maximize the total amount of offloaded data, subject to the time-averaged energy consumption constraints. The main contributions of this paper are summarized as follows: First, we formulate an online joint power allocation and task scheduling optimization problem for energy constrained online data offloading in LSNs. Different from prior online data offloading problems for LSNs, we accommodate both stochastic space data arrivals

This work was supported in part by the National Natural Science Foundation of China under Grant 62201463 and 62301251 and in part by the Natural Science Foundation of Jiangsu Province of China under Project BK20220883. (Corresponding author: Lijun He)

2025 IEEE Wireless Communications and Networking Conference (WCNC) | 979-8-3503-6836-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/WCNC61545.2025.10978643

and time-varying SGLs. Due to the tight coupling between task scheduling and power allocation over time slots, we leverage Lyapunov optimization to convert the joint online optimization problem into a set of per-time-slot subproblems. Each subproblem is then equivalently decoupled into a task scheduling problem embedded with a power allocation problem. We further explore the unique structures of power allocation problem and task scheduling problem to derive an optimal power allocation solution and a multi-armed bandit based quasi-optimal task scheduling solution, respectively. Finally, simulation results demonstrate that our proposed Energy-constrained Online Data Offloading (EODO) algorithm outperforms the alternatives.

II. System Model and Problem Formulation

Consider a time-slotted LSN consisting of one LEO satellite and a set of GSs. The scheduling horizon of the LSN is divided into a set $\mathcal{T} = \{1, 2, \ldots, T\}$ of time slots. Each time slot $t \in \mathcal{T}$ indexes the time segment of $[t, t+1]$ with the same duration of τ. The LEO satellite can offload its carried task data to a set $\mathcal{G} = \{1, 2, \ldots, G\}$ of GSs in each time slot $t \in \mathcal{T}$. The high-speed motion of the LEO satellite leads to intermittent contacts to the GSs over time. Therefore, the LEO satellite can only connect to a subset $\mathcal{G}(t)$ of GSs for each time slot t.

A. Task and Channel Model

Consider a task model with sequential dependencies, i.e., a new task is triggered and generated at the beginning of each time slot, and the LEO satellite processes all tasks in order without any overlapping in time slots. We assume that a new task with the data amount of $a(t)$ is generated in time slot $t \in \mathcal{T}$. We let $\boldsymbol{a} = \{a(1), a(2), \ldots, a(T)\}$ be the data amount of all data offloading tasks. We further introduce a variable $x_{g,t}$ to represent the data offloading strategy, such that $x_{g,t} = 1$ indicates that the LEO satellite selects the GS $g \in \mathcal{G}(t)$ to offload task data in time slot $t \in \mathcal{T}$; and $x_{g,t} = 0$ otherwise. We let $R_g(t)$ denote the achievable data rate of SGLs between the LEO satellite and GS g in time slot t, given by

$$R_g(t) = B_c \log_2 \left(1 + \text{SNR}_g(t)\right), \ \forall g \in \mathcal{G}(t), t \in \mathcal{T}, \quad (1)$$

where B_c is the dedicated bandwidth allocated to the LEO satellite and $\text{SNR}_g(t)$ is the signal-to-noise ratio of the SGL between the LEO satellite and GS g in time slot t. The value of $\text{SNR}_g(t)$ is evaluated by

$$\text{SNR}_g(t) = \Psi_g(t) P_g(t), \ \forall g \in \mathcal{G}(t), t \in \mathcal{T}, \quad (2)$$

where $\Psi_g(t)$ and $P_g(t)$ represent the channel gain and the allocated transmit power (in Watts) of the SGL between the LEO satellite and GS g in time slot t, respectively. The value of $\Psi_g(t)$ is given by

$$\Psi_g(t) = \frac{h^{tr}(t) h_g^{re}(t) L_f(t) L_l(t)}{\sigma_g^2(t)}, \quad (3)$$

where $h^{tr}(t)$ is the transmit antenna gain of the LEO satellite in time slot t, $h_g^{re}(t)$ is the gain of data receiver antenna on GS g in time slot t, $L_f(t)$ is the free space loss in time slot t, $L_l(t)$ is the total line loss in time slot t, and $\sigma_g^2(t)$ is the noise power in time slot t [14].

B. Queue Update and Energy Consumption Model

The arrival and departure procedure of task data in the LEO satellite for each time slot t is modeled as a queue, denoted by $Q(t)$. This queue is dynamically updated over time slots according to the following rule:

$$Q(t+1) = \max\left[Q(t) - b(t), 0\right] + a(t), \ \forall t \in \mathcal{T}, \quad (4)$$

where $b(t)$ is the data volume offloaded from the LEO satellite to one GS in time slot t. The definition of the offloaded data volume $b(t)$ is given by

$$b(t) = R(t) d(t), \ \forall t \in \mathcal{T}. \quad (5)$$

In (5), $R(t)$ is the offload transmission rate of the LEO satellite in time slot t, expressed as

$$R(t) = \sum_{g \in \mathcal{G}(t)} x_{g,t} R_g(t), \ \forall t \in \mathcal{T}, \quad (6)$$

and $d(t)$ indicates the transmission time for the LEO satellite to offload the task data in time slot t, evaluated by

$$d(t) = \sum_{g \in \mathcal{G}(t)} x_{g,t} d_g(t), \ \forall t \in \mathcal{T}. \quad (7)$$

In (7), $d_g(t)$ represents the transmission time between the LEO satellite and GS g within time slot t, given by

$$d_g(t) = \begin{cases} \frac{Q(t)}{R_g(t)} & Q(t) < b(t), \\ \tau & \text{otherwise}. \end{cases} \quad (8)$$

Furthermore, we denote $e(t)$ as the energy consumption for the LEO satellite to offload its task data in time slot t, which can be calculated as

$$e(t) = \sum_{g \in \mathcal{G}(t)} x_{g,t} P_g(t) d(t), \ \forall t \in \mathcal{T}. \quad (9)$$

C. Joint Online Optimization Problem Formulation

By jointly optimizing power allocation and task scheduling, we formulate an online data offloading problem for LSNs as the following stochastic optimization problem:

$$\begin{aligned} \textbf{P0}: \ & \max_{\boldsymbol{P}(t), \boldsymbol{x}(t)} \ \lim_{|\mathcal{T}| \to \infty} \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \mathbb{E}\left\{b(t)\right\} \\ \text{s.t.} \ & \text{C1}: \ \lim_{|\mathcal{T}| \to \infty} \sup \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \mathbb{E}\left\{e(t)\right\} \le \bar{e}, \\ & \text{C2}: \ P_{\min} \le P_g(t) \le P_{\max}, \ \forall t \in \mathcal{T}, g \in \mathcal{G}(t), \\ & \text{C3}: \ \sum_{g \in \mathcal{G}(t)} x_{g,t} = 1, \ \forall t \in \mathcal{T}, \\ & \text{C4}: \ x_{g,t} \in \{0, 1\}, \ \forall t \in \mathcal{T}, g \in \mathcal{G}(t), \end{aligned}$$

where $\boldsymbol{P}(t) = \{P_g(t)\}$, $\boldsymbol{x}(t) = \{x_{g,t}\}$, $P_{\min}$ is the minimum transmit power, $P_{\max}$ is the maximum transmit power, and $\bar{e}$

indicates the average energy level allowably consumed in the LEO satellite. C1 indicates that the energy consumption of the LEO satellite can not exceed its average level of allowable energy consumption from the long-term average perspective. C2 indicates the range of power control of the LEO satellite. C3 reflects that the LEO satellite selects at most one GS in each time slot. C4 is the binary constraint.

III. Proposed Solution

In this section, we introduce a virtual queue and leverage Lyapunov optimization to transform **P0** into a set of per-slot optimization problems, followed by an efficient solution for each per-slot optimization problem.

A. Virtual Queue and Per-Slot Optimization Problem

To handle the long-term energy constraint C1, we introduce a virtual queue $Z(t)$ with the following updating rule:

$$Z(t+1) = \max\{Z(t) + e(t) - \bar{e}, 0\}. \tag{10}$$

We further introduce a Lyapunov function of $L(Z(t)) = \frac{1}{2}Z^2(t)$ to convert C1 into the stability of virtual queue $Z(t)$. On basis of this, we provide the definition of conditional Lyapunov drift for each time slot t to show the change in the Lyapunov function from one slot to the next as follows:

$$\Delta\left(Z(t)\right) = \mathbb{E}\{L(Z(t+1)) - L(Z(t))|Z(t)\},\ \forall t \in \mathcal{T}.$$

As such, the stability of virtual queue $Z(t)$ can be equivalently transformed into minimizing $\Delta\left(Z(t)\right)$. Furthermore, we introduce a Drift-Plus-Penalty (DPP) metric of $\Delta\left(Z(t)\right) - V\mathbb{E}\left\{b(t)\right\}$ to transform **P0** into the following problem:

$$\begin{aligned}\textbf{P1}:\ &\min_{\boldsymbol{P}(t),\boldsymbol{x}(t)}\ \lim_{|\mathcal{T}|\to\infty}\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\left(\Delta\left(Z(t)\right) - V\mathbb{E}\left\{b(t)\right\}\right)\\ &\text{s.t. C2-C4,}\end{aligned}$$

where V is a positive parameter to balance the stability of virtual queue $Z(t)$ and the offloaded data amount. Directly solving **P1** is still difficult, since its objective function is non-convex. As such, we turn to calculating an upper bound of the optimal objective value of **P1**. Using the definitions of $L\left(Z(t)\right)$ and $\Delta(Z(t))$, we have $\Delta\left(Z(t)\right) - V\mathbb{E}\left\{b(t)\right\} \le \frac{1}{2}(P_{\max}\tau - \bar{e})^2 - V\mathbb{E}\left\{b(t)\right\} + \mathbb{E}\{Z(t)(e(t)-\bar{e})|Z(t)\}$ and obtain an upper bound problem of **P1** as follows:

$$\begin{aligned}\textbf{P2}:\ &\min_{\boldsymbol{P}(t),\boldsymbol{x}(t)}\ \lim_{|\mathcal{T}|\to\infty}\frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\Big(\mathbb{E}\{Z(t)(e(t)-\bar{e})|Z(t)\}\\ &- V\mathbb{E}\left\{b(t)\right\}\Big)\\ &\text{s.t. C3-C5.}\end{aligned}$$

We transform **P2** into the following per-time-slot subproblem:

$$\begin{aligned}\textbf{P3}:\ &\min_{\boldsymbol{P}(t),\boldsymbol{x}(t)}\ \sum_{g\in\mathcal{G}}x_{g,t}U_g(t)\\ &\text{s.t. C6}:\ P_{\min}\le P_g(t)\le P_{\max}, \forall g\in\mathcal{G}(t),\\ &\text{C7}:\ \sum_{g\in\mathcal{G}}x_{g,t}=1,\\ &\text{C8}:\ x_{g,t}\in\{0,1\},\ \forall g\in\mathcal{G}(t),\end{aligned}$$

where $U_g(t) = \left(Z(t)P_g(t) - VR_g(t)\right)d_g(t)$. Although **P3** is a mixed integer programming problem, we next explore its special structure to obtain a high-quality solution.

B. Solution for Per-Slot Optimization Problem

We separate **P3** into two levels of optimization. At the lower optimization level optimization, we optimize power allocation $\boldsymbol{P}(t)$ with fixed task scheduling $\boldsymbol{x}(t)$:

$$\begin{aligned}\textbf{P4}:\ &\min_{P_g(t)}\ U_g(t)\\ &\text{s.t. C6.}\end{aligned}$$

At the higher optimization level optimization, we optimize task scheduling $\boldsymbol{x}(t)$ when fixed power allocation $\boldsymbol{P}(t)$:

$$\begin{aligned}\textbf{P5}:\ &\min_{\boldsymbol{x}(t)}\ \sum_{g\in\mathcal{G}}x_{g,t}U_g^*(t)\\ &\text{s.t. C7-C8,}\end{aligned}$$

where $U_g^*(t)$ is an optimal value to **P4**.

*1) **P4** Solution:* The definition of $d_g(t)$ in (8) indicates that it is a piecewise function. As such, we discuss two cases of this piecewise function. We note that the partition point of $d_g(t)$ can be calculated as $P_0^g(t) = \frac{1}{\Psi_g(t)}(2^{\frac{Q(t)}{\tau B_c}} - 1)$. Since $\frac{Q(t)}{\tau B_c} \ge 0$, we have three cases for the range of $P_0^g(t)$:

Case 1: Since $P_0^g(t) \in [P_{\min}, P_{\max}]$, we recast (8) as the following form:

$$d_g(t) = \begin{cases}\tau & P_g(t)\in[P_{\min}, P_0^g(t)],\\ \frac{Q(t)}{R_g(t)} & P_g(t)\in(P_0^g(t), P_{\max}].\end{cases}$$

When $P_g(t) \in [P_{\min}, P_0^g(t)]$, we have $d_g(t) = \tau$ and rewrite **P4** as the following problem:

$$\begin{aligned}\textbf{P6}:\ &\min_{P_g(t)}\ \left(Z(t)P_g(t) - VR_g(t)\right)\tau\\ &\text{s.t. } P_{\min}\le P_g(t)\le P_0^g(t).\end{aligned}$$

We further substitute (1) and (2) into the objective function of **P6** to yield: $\Phi\left(P_g(t)\right) = \left(Z(t)P_g(t) - VB_c\log_2\left(1+\Psi_g(t)P_g(t)\right)\right)\tau$. Then, we deduce its first-order and second-order derivation as follows:

$$\nabla\left(\Phi\left(P_g(t)\right)\right) = Z(t) - VB_c\frac{\Psi_g(t)}{\left(1+\Psi_g(t)P_g(t)\right)\ln 2}, \tag{11}$$

$$\nabla^2\left(\Phi\left(P_g(t)\right)\right) = \frac{VB_c\left(\Psi_g(t)\right)^2}{\left(1+\Psi_g(t)P_g(t)\right)^2\ln 2}. \tag{12}$$

From (12), we see that $\nabla^2\left(\Phi\left(P_g(t)\right)\right) \ge 0$. This indicates that **P6** is convex. As such, we use the fact of $\nabla\left(\Phi\left(P_g(t)\right)\right) = 0$ to compute an extremal point to **P6**, given by

$$P_g^0(t) = \frac{VB_c}{Z(t)\ln 2} - \frac{1}{\Psi_g(t)}. \tag{13}$$

If $P_{\min} \leq P_g^0(t) \leq P_0^g(t)$, we use (13) to yield: $\Psi_g(t)VB_c/\left(1+\Psi_g(t)P_0^g(t)\ln 2\right) \leq Z(t) \leq \Psi_g(t)VB_c/\left(1+\Psi_g(t)P_{\min}\ln 2\right)$. Then, we have the optimal solution to **P6** satisfying $P_1^g(t) = P_g^0(t)$. If $P_g^0(t) > P_0^g(t)$, we have $Z(t) < \Psi_g(t)VB_c/(1+\Psi_g(t)P_0^g(t)\ln 2)$. Then, when $P_{\min} \leq P_g(t) \leq P_g^0(t)$, we can obtain: $\triangledown\left(\Phi\left(P(t)\right)\right) < 0$. As such, we have the optimal solution to **P6** satisfying $P_1^g(t) = P_0^g(t)$. If $P_g^0(t) < P_{\min}$, we have $Z(t) > \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\min})\ln 2}$. Then, when $P_{\min} \leq P_g(t) \leq P_0^g(t)$, we have $\triangledown\left(\Phi\left(P_g(t)\right)\right) > 0$. As such, we have the optimal solution to **P6** satisfying $P_1^g(t) = P_{\min}$. Thus, we have

$$P_1^g(t) = \begin{cases} P_0^g(t) & Z(t) < \frac{\Psi_g(t)VB_c}{\left(1+\Psi_g(t)P_0^g(t)\right)\ln 2}, \\ P_{\min} & Z(t) > \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\min})\ln 2}, \\ P_g^0(t) & \text{otherwise.} \end{cases}$$

When $P_g(t) \in (P_0^g(t), P_{\max}]$, we have $d_g(t) = \frac{Q(t)}{R_g(t)}$ and rewrite **P4** as the following problem:

$$\textbf{P7}: \min_{P_g(t)} \left(Z(t)\frac{P_g(t)}{R_g(t)} - V\right)Q(t)$$
$$\text{s.t. } P_0^g(t) \leq P_g(t) \leq P_{\max}.$$

We further introduce a new variable $\eta_g(t) = \frac{1}{R_g(t)}$ to replace $R_g(t)$ and $P_g(t)$. Combining this with (1) and (2), we have

$$\eta_g(t) = \frac{1}{B_c \log_2\left(1+\Psi_g(t)P_g(t)\right)}. \tag{14}$$

From (14), we can easily derive:

$$P_g(t) = \frac{1}{\Psi_g(t)}\left(2^{\frac{1}{B_c\eta_g(t)}} - 1\right). \tag{15}$$

We use (14) and (15) to recast **P7** as the following form:

$$\textbf{P8}: \min_{\eta_g(t)} \ \varphi\left(\eta_g(t)\right)Q(t)$$
$$\text{s.t. C9}: \ \chi\left(P_{\max}\right) \leq \eta_g(t) \leq \chi\left(P_0^g(t)\right),$$

where $\varphi\left(\eta_g(t)\right) = \frac{Z(t)}{\Psi_g(t)}\eta_g(t)\left(2^{\frac{1}{B_c\eta_g(t)}} - 1\right) - V$ and $\chi\left(P_g(t)\right) = \frac{1}{B_c\log_2(1+\Psi_g(t)P_g(t))}$. We further derive the first-order and second-order derivation of $\varphi\left(\eta_g(t)\right)$:

$$\triangledown\varphi\left(\eta_g(t)\right) = \frac{Z(t)}{\Psi_g(t)}\left(2^{\frac{1}{\eta_g(t)B_c}}\left(1 - \frac{\ln 2}{\eta_g(t)B_c}\right) - 1\right), \tag{16}$$

$$\triangledown^2\varphi\left(\eta_g(t)\right) = \frac{Z(t)(\ln 2)^2}{\Psi_g(t)\left(\eta_g(t)\right)^3\left(B_c\right)^2}2^{\frac{1}{\eta_g(t)B_c}}. \tag{17}$$

We see from (17) that the second derivative of $\varphi\left(\eta_g(t)\right)$ is positive, i.e., $\triangledown^2\varphi\left(\eta_g(t)\right) > 0$, which indicates that $\triangledown\varphi\left(\eta_g(t)\right)$ monotonously increases as $\eta_g(t)$ increasing. Note that the upper bound of $\triangledown\varphi\left(\eta_g(t)\right)$ is zero, i.e., $\lim_{\eta_g(t)\to+\infty}\triangledown\varphi\left(\eta_g(t)\right) = 0$. Thus, we can easily obtain $\triangledown\varphi\left(\eta_g(t)\right) < 0$, when $\chi\left(P_{\max}\right) \leq \eta_g(t) \leq \chi\left(P_0^g(t)\right)$. This reveals that $\varphi\left(\eta_g(t)\right)$ monotonously decreases as $\eta_g(t)$ increasing. As such, it is easy to obtain the optimal solution to **P8**, obtaining that the optimal power solution to **P7** is $P_2^g(t) = P_0^g(t)$. In brief, the optimal power to **P6** is $P_g^*(t) = \arg\ \min\{\Phi(P_1^g(t))\tau, \varphi\left(\chi\left(P_2^g(t)\right)\right)Q(t)\}$.

Case 2: Since $P_g^0(t) \in (P_{\max}, \infty)$, we have $d_g(t) = \tau, P_g(t) \in [P_{\min}, P_{\max}]$ and recast **P4** as

$$\textbf{P9}: \ \min \ \left(Z(t)P_g(t) - VR_g(t)\right)\tau$$
$$\text{s.t. C6.}$$

We see that $P_g^0(t)$ is also an extremal point of **P9**, since it has the same form as **P6**. As such, we have the following three cases: If $P_{\min} \leq P_g^0(t) \leq P_{\max}$, we use (13) to yield: $\frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\max})\ln 2} \leq Z(t) \leq \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\min})\ln 2}$. Then, the optimal solution to **P9** is $P_g^0(t)$. If $P_g^0(t) > P_{\max}$, we have $Z(t) < \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\max})\ln 2}$. Then, the optimal solution to **P9** is $P_{\max}$. If $P_g^0(t) < P_{\min}$, we have $Z(t) > \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\min})\ln 2}$. Then, the optimal solution to **P9** is $P_{\min}$. As such, we have the optimal solution to **P9** as follows:

$$P_g^*(t) = \begin{cases} P_{\max} & Z(t) < \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\max})\ln 2}, \\ P_{\min} & Z(t) > \frac{\Psi_g(t)VB_c}{(1+\Psi_g(t)P_{\min})\ln 2}, \\ P_g^0(t) & \text{otherwise.} \end{cases}$$

Case 3: Since $P_0^g(t) \in (0, P_{\min})$, we have $d_g(t) = \frac{Q(t)}{R_g(t)}, P_g(t) \in [P_{\min}, P_{\max}]$ and recast **P4** as follows:

$$\textbf{P10}: \min_{P_g(t)} \left(Z(t)\frac{P_g(t)}{R_g(t)} - V\right)Q(t)$$
$$\text{s.t. C6.}$$

Next, we further transform **P10** into the following problem:

$$\textbf{P11}: \min_{\eta_g(t)} \ \varphi\left(\eta_g(t)\right)Q(t)$$
$$\text{s.t. C10}: \ \chi\left(P_{\max}\right) \leq \eta_g(t) \leq \chi\left(P_{\min}\right).$$

Since **P11** has the same form of **P8**, we can obtain that $\triangledown^2\varphi\left(\eta_g(t)\right) > 0$ and the upper bound of $\triangledown\varphi\left(\eta_g(t)\right)$ is zero. Therefore, we have $\triangledown\varphi\left(\eta_g(t)\right) < 0$ which indicates that $\varphi\left(\eta_g(t)\right)$ monotonously decreases as $\eta_g(t)$ increasing. Hence, the optimal solution to **P11** is $\chi(P_{\min})$, and the optimal power to **P10** is obtained by $P_g^*(t) = P_{\min}$.

*2) **P5** Solution:* The challenge to solve **P5** is how to acquire the channel gains of the time-varying yet time-delayed SGLs in real time. To this end, we resort to the Multi-Armed Bandit (MAB) framework [15] to tackle **P5** via learning the unknown yet delayed $\Psi_g(t), \forall g, t$. In particular, we treat each GS as an arm, each associated with a cost unknown beforehand, i.e., the objective value of **P5**. As such, **P5** can be equivalently reformulated as a MAB problem to minimize the cost in each time slot through the arm selection. Nevertheless, the challenge for the MAB problem is to strike a balance of choosing the best arm currently available and exploring potentially better arms. To address this challenge, we leverage an Upper Confidence Bound (UCB) policy [15] to achieve a good balance between exploration and exploitation by constantly calculating the

confidence radius. The main steps of the proposed UCB policy are below. First, we need to obtain an estimation of the channel gain $\hat{\Psi}_g(t)$, given by

$$\hat{\Psi}_g(t) = \bar{\Psi}_g(t) + \sqrt{\frac{3\ln t}{2n_g(t-1)}}, \tag{18}$$

where $\bar{\Psi}_g(t)$ represents the sample mean of $\Psi_g(t)$, given by

$$\bar{\Psi}_g(t) = \begin{cases} \frac{1}{n_g(t)} \sum_{k=1}^{t} x_{g,t}\Psi_g(t) & n_g(t) > 0, \\ 0 & n_g(t) = 0. \end{cases} \tag{19}$$

Wherein, $n_g(t)$ represents the number of selected times of GS g until the end of time slot t, i.e.,

$$n_g(t) = \sum_{k=1}^{t} x_{g,t}, \ \forall t \in \mathcal{T}. \tag{20}$$

Particularly, we set $\hat{\Psi}_g(t) = 0$ when $n_g(t-1) = 0$. We note that the second term of (18) is the confidence radius of $\hat{\Psi}_g(t)$, indicating the accuracy of the estimated channel gain $\hat{\Psi}_g(t)$. The smaller the confidence radius, the closer $\hat{\Psi}_g(t)$ is to the mean of $\Psi_g(t)$. The definition of $\hat{\Psi}_g(t)$ reveals the method of adaptive exploration in the UCB policy. That is, when the confidence radius is large, the predicted value $\hat{\Psi}_g(t)$ is relatively large, indicating that it may be a potentially better choice for future exploration. Second, we use $\hat{\Psi}_g(t)$ to acquire the estimated transmission rate for GS g in time slot t, denoted by $\hat{R}_g(t)$. Specifically, we combine (1) with (2) to calculate the value of $\hat{R}_g(t)$, i.e., $\hat{R}_g(t) = B_c \log_2\left(1 + \hat{\Psi}_g(t)P_g^*(t)\right)$. Third, we use $\hat{R}_g(t)$ to calculate the estimated objective function value of **P5** equal to $\sum_{g\in\mathcal{G}} x_{g,t}\left(Z(t)P_g^*(t) - V\hat{R}_g(t)\right)\hat{d}_g(t)$, where $\hat{d}_g(t) = \min(\tau, \frac{Q(t)}{\hat{R}_g(t)})$. Finally, we use the UCB policy to yield the task scheduling strategy, such that we have $x_{g,t}^* = 1$ when

$$g^*(t) = \arg\min_{g\in\mathcal{G}(t)} \left(Z(t)P_g^*(t) - V\hat{R}_g(t)\right)\hat{d}_g(t), \ \forall t. \tag{21}$$

Algorithm 1 Proposed EODO Algorithm

Input: Parameter $\mathcal{G}(t)$, $Q(t)$, $Z(t)$, $P_{\min}$, $P_{\max}$, τ, B_c, $\bar{e}$.

```
for t = 1, ..., T do
  for g = 1, ..., G do
    if g ∈ G(t) then
      Use (18) to compute Ψ̂_g(t).
      Use P4 solution to obtain P_g^*(t).
    end if
  end for
  Use (21) to obtain x_{g,t}^*.
  Use (4) and (10) to update Q(t) and Z(t).
  Use (20) and (19) to update n_g(t) and Ψ̄_g(t).
end for
```

Output: $\boldsymbol{P}^*(t) = \{P_g^*(t)\}$, $\boldsymbol{x}_{g,t}^* = \{x_{g,t}^*\}$.

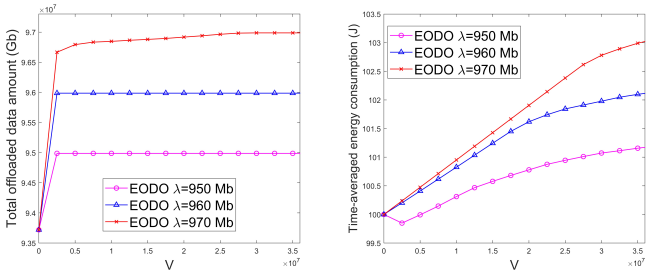
Otherwise, we have $x_{g,t}^* = 0$. We summarize the main steps above in Algorithm 1.

IV. Numerical Results

In this section, we evaluate the proposed EODO algorithm through the following extensive simulations. The simulation scenario of LSN is consist of one LEO satellite and 100 GSs. The total number of time slots is set to 10^5. The size of each time slot is set to 1 second, i.e., $\tau = 1$. In each time slot, the LEO satellite can access only 20 GSs due to the dynamic characteristic of the LSN. We denote λ by the average amount of data arriving. The data amount of the LEO satellite is generated by a Poisson distribution with $\lambda = 960$ Mb. We further set $L_f = 10^{-23}$, $L_l = 1$, $B_c = 40$ GHz, and $\sigma_g^2(t) = -97.579$ dB. We randomly assign five values within $[32.42, 37.42]$ to $h^{tr}(t)$ and $h_g^{re}(t)$ for each time slot. We set $\bar{e} = 100$ J, $V = 3.0 \times 10^7$, $P_{\min} = 0$ W and $P_{\max} = 200$ W.

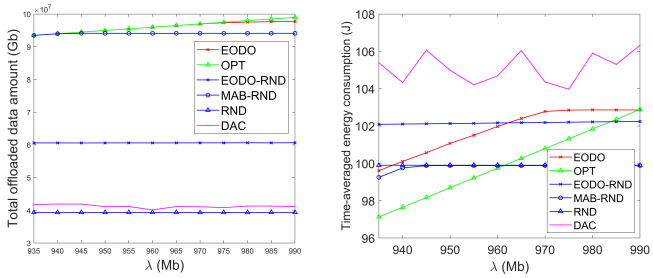
To show the convergence evolution of EODO with different λ, Fig. 1a and Fig. 1b plot the total offloaded data amount and the time-averaged energy consumption versus the value of parameter V, respectively. We observe that the total offloaded data amount under different values of V first increases and then inevitably converges to a maximum value as the value of V grows. This indicates that setting V large enough can guarantee the convergence of EODO. From Fig. 1b, we observe that the time-averaged energy consumption under different values of V increases with the gradual decreasing rate. This illustrates that a larger value of V would increase energy consumption. In brief, setting a suitable value of V in EODO enables the larger total offloaded data amount with less energy consumption.

There are five baseline algorithms involving the Deep Actor-Critic (DAC) [11] algorithm, the Optimal Power allocation and Task scheduling (OPT) algorithm, the EODO with Random GS selection (EODO-RND) algorithm, the MAB with Random power allocation (MAB-RND) algorithm, and the Random (RND) algorithm. DAC is an efficient deep reinforcement learning based algorithm consisting of the two neural networks of the actor network and the critic network. The actor network is responsible for making the joint decisions of channel selection and power allocation. Then, the critic network evaluates the generated decisions in terms of the offloaded data amount and energy consumption. The interaction of the two networks as the time slot evolving promotes the decision efficiency of the actor network while enhancing the evaluation accuracy of the critic network. The OPT algorithm leverages the **P4** solution in Section III-B1 to obtain the optimal power allocation and selects one GS associated with the best objective value of **P5** in each time slot with the known $\Psi_g(t)$. The EODO-RND algorithm uses EODO to optimize power control but randomly selects the available GS in each time slot. The MAB-RND algorithm utilizes the UCB policy to solve **P5** but optimizes the power allocation in random manner within $[P_{\min}, P_{\max}]$. Under RND, each GS is randomly selected and the power allocation of each time slot is randomly generated.



(a) (b)

Fig. 1. Verification of convergence. (a) Total offloaded data amount versus V. (b) Time-averaged energy consumption versus V.



(a) (b)

Fig. 2. Comparisons with alternatives. (a) Total offloaded data amount versus λ. (b) Time-averaged energy consumption versus λ.

The performance comparisons of EODO against the baselines above are shown in Fig. 2a and Fig. 2b. In Fig. 2a, we compare EODO with the five baseline algorithms in terms of the total offloaded data amount versus the value of λ. From Fig. 2a, we observe that EODO offloads more task data than the schemes of EODO-RND, MAB-RND, and RND as the value of λ increases. This is because EODO jointly optimizes the variables of the power allocation and task scheduling to improve the total offloaded data amount of LSNs, while the alternatives only optimize at most one optimization variable. Furthermore, we observe that EODO is very close to the scheme of OPT in terms of the total offloaded data amount with the increasing λ. This indicates that the proposed EODO algorithm is quite efficient to improve the data offloaded efficiency of LSNs. Fig. 2b shows that the energy consumption of EODO first increases and then converges to a constant. It is because that EODO needs to follow the constraint C1 in **P0**. From Fig. 2a and Fig. 2b, we observe that compared with DAC, EODO schedules more offloaded data in the less cost of energy consumption. The hidden reason is that EODO achieves the optimal power allocation for each time slot to significantly reduce the solution space of **P0**.

V. Conclusion

In this paper, we have studied energy-constrained online data offloading for LSNs to accommodate both the stochastic space data arrivals and the time-varying yet time-delayed SGLs. We convert the joint online power allocation and task scheduling problem into a set of per-time-slot subproblems, and propose an efficient two-layer optimization scheme to solve them. Simulation results demonstrate that the proposed EODO algorithm significantly outperforms the state-of-the-art algorithms and is close to the optimal scheme.

References

[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys & Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart. Feb. 2023.

[2] H. Al-Hraishawi, H. Chougrani, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "A survey on nongeostationary satellite systems: The communication perspective," *IEEE Commun. Surveys & Tuts.*, vol. 25, no. 1, pp. 101–132, 1nd Quart. Feb. 2023.

[3] X. Cao, B. Yang, Y. Shen, C. Yuen, Y. Zhang, Z. Han, H. V. Poor, and L. Hanzo, "Edge-assisted multi-layer offloading optimization of LEO satellite-terrestrial integrated networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 2, pp. 381–398, Dec. 2023.

[4] M. Ouyang, R. Zhang, B. Wang, J. Liu, T. Huang, L. Liu, J. Tong, N. Xin, and F. R. Yu, "Network coding-based multipath transmission for LEO satellite networks with domain cluster," *IEEE IoT J.*, vol. 11, no. 12, pp. 21 659–21 673, Mar. 2024.

[5] L. He, J. Li, Y. Wang, J. Zheng, and L. He, "Balancing total energy consumption and mean makespan in data offloading for space-air-ground integrated networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 1, pp. 209–222, Jan. 2024.

[6] L. He, K. Guo, H. Gan, and L. Wang, "Collaborative data offloading for earth observation satellite networks," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1116–1120, May 2022.

[7] S. Rojanasoonthon, "Parallel machine scheduling with time windows." Ph.D. dissertation, Univ. Texas, Austin, TX, 2004.

[8] F. He, Q. Liu, T. Lv, C. Liu, H. Huang, and X. Jia, "Delay-bounded and minimal transmission broadcast in LEO satellite networks," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May 2016.

[9] X. Jia, T. Lv, F. He, and H. Huang, "Collaborative data downloading by using inter-satellite links in LEO satellite networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1523–1532, Mar. 2017.

[10] C.-Q. Dai, C. Li, S. Fu, J. Zhao, and Q. Chen, "Dynamic scheduling for emergency tasks in space data relay network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 795–807, Jan. 2020.

[11] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[12] X. Gao, J. Wang, X. Huang, Q. Leng, Z. Shao, and Y. Yang, "Energy-constrained online scheduling for satellite-terrestrial integrated networks," *IEEE Trans. Mobile Comput.*, vol. 22, no. 4, pp. 2163–2176, Apr. 2023.

[13] L. He, Z. Jia, K. Guo, H. Gan, Z. Han, and C. Yuen, "Online joint data offloading and power control for space-air-ground integrated networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 18 126–18 141, Dec. 2024.

[14] D. Zhou, M. Sheng, Y. Wang, J. Li, and Z. Han, "Machine learning-based resource allocation in satellite networks supporting Internet of remote things," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6606–6621, Oct. 2021.

[15] X. Gao, J. Wang, X. Huang, Q. Leng, Z. Shao, and Y. Yang, "Energy-constrained online scheduling for satellite-terrestrial integrated networks," *IEEE Trans. Mobile Comput.*, vol. 22, no. 4, pp. 2163–2176, Sept. 2023.