Collaborative Multi-Resource Allocation in Terrestrial-Satellite Network Towards 6G

Shu Fu, Jie Gao, *Senior Member, IEEE*, and Lian Zhao, *Senior Member, IEEE*

***Abstract*— Terrestrial-satellite networks (TSNs) are envisioned to play a significant role in the sixth-generation (6G) wireless networks. In such networks, hot air balloons are useful as they can relay the signals between satellites and ground stations. Most existing works assume that the hot air balloons are deployed at the same height with the same minimum elevation angle to the satellites, which may not be practical due to possible route conflict with airplanes and other flight equipment. In this paper, we consider a TSN containing hot air balloons at different heights and with different minimum elevation angles, which creates the challenge of non-uniform available serving time for the communication between the hot air balloons and the satellites. Jointly considering the caching, computing, and communication (3C) resource management for both the ground-balloon-satellite links and inter-satellite laser links, our objective is to maximize the network energy efficiency. Firstly, by proposing a tapped water-filling algorithm, we schedule the traffic to relay among satellites according to the available serving time of satellites. Then, we generate a series of configuration matrices, based on which we formulate the relation between relay time and the power consumption involved in the relay among satellites. Finally, the collaborative resource allocation problem for TSN is modeled and solved by geometric programming with Taylor series approximation. Simulation results demonstrate the effectiveness of our proposed scheme.**

***Index Terms*— Terrestrial-satellite network, 6G, hot air balloon, satellite relay, energy efficiency.**

I. Introduction

In 6G, terrestrial-satellite networks (TSN) will become the key leverage to explore airspace resources for communications [1]–[4]. In such networks, laser based inter-satellite links working at Terahertz frequency can support high-capacity transmissions among low earth orbit (LEO) satellites [5]. Meanwhile, the large distance between satellites and ground stations necessitates relays, such as hot air balloons, between the terrestrial and satellite segments [6]. A TSN integrates caching, computing, and communication (3C), each of which may consume considerable energy. Therefore, maximizing the energy efficiency performance will be one of the main design objectives of TSN in 6G [7], [8].

Communications in a TSN involve steps of data traffic collection, transmission, and relay. In a previous work [1], we studied such a TSN with the above steps but without considering hot air balloons as relays or laser link among satellites. In practice, a TSN can be very complex due to the following reasons.

Firstly, there can be relays such as hot air balloons between satellites and ground stations. In real-world environment, the relays may be deployed at different heights due to a variety of reasons, such as avoiding airplane routes. Additionally, the minimum elevation angle from the hot air balloons to the satellites may not be identical either. Define the time during which a relay and its serving satellite can communicate as *time windows*. Then, the time windows for each relay to communicate to a satellite can be unique. Modeling and studying the above real-world TSN scenario is challenging and important. However, relays with different heights and minimum elevation angles are not yet considered in the existing works, to the best of our knowledge.

Secondly, regarding the laser links among the satellites, it is necessary to determine the number of lasers per satellite as well as develop a proper scheme for scheduling inter-satellite traffic. If the number of lasers per satellite is large, delay is low, power consumption can be high, and scheduling can be complex. By contrast, if the number of lasers per satellite is small, power consumption is low, but the delay is large.

Finally, in additional to the energy efficiency of communications between the relays and satellites and among the satellites, as mentioned above, the communications between ground stations and hot air balloons should be involved while considering the network energy efficiency. Moreover, the energy consumption of caching and computing should also be included, in additional to that of communication. This requires a joint allocation of the 3C resources.

Existing works contributed to TSNs from different aspects. In [9], a framework to efficiently deploy customized service function chains was proposed in terrestrial-satellite hybrid cloud networks, which enables computation and data traffic off-loading in a TSN. In [10], a genetic algorithm based

Manuscript received December 26, 2020; revised April 2, 2021; accepted May 6, 2021. Date of publication May 21, 2021; date of current version November 11, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61701054, Grant 62071398, and Grant U20A20157, in part by the Fundamental Research Funds for the Central University under Grant 2020CDJQY-A001, and in part by the Huawei Collaboration Project under Grant HF2019115002. The associate editor coordinating the review of this article and approving it for publication was H. Zhu. *(Corresponding authors: Shu Fu; Lian Zhao.)*

Shu Fu is with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China, and also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China (e-mail: shufu@cqu.edu.cn).

Jie Gao is with the Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI 53233 USA (e-mail: j.gao@marquette.edu).

Lian Zhao is with the Department of Electrical, Computer, and Biomedical Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: l5zhao@ryerson.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3080578.

Digital Object Identifier 10.1109/TWC.2021.3080578

method was proposed to increase the coverage of a TSN. In [11], a novel network architecture was proposed to provide seamless and high data-rate wireless service in a TSN, where the terrestrial and satellite segments are jointly designed. In [12], a dynamic space-air-ground resource pool was proposed by a software-defined space-air-ground integrated framework to effectively manage the network in a seamless, efficient, and cost-effective manner. In [13], an unmanned aerial vehicle (UAV) aided space-air-ground network was proposed to provide seamless coverage and high system throughput. However, the system models used in most of the existing works considered the case that all hot air balloons were deployed at an equal height and with the same minimum elevation angle. In [14], a UAV based relay system was proposed in a non-terrestrial network. In [15], UAV was leveraged to optimize computation offloading with minimum energy consumption in mobile edge computing.

In terms of the inter-satellite relay links, a routing and wavelength assignment algorithm was proposed to reduce the system cost of inter-satellite laser communication in [16]. This work focused on the allocation of optical wavelengths, without considering the optimal number of lasers. In [17], joint wavelength allocation and routing to decrease the system energy consumption was investigated. In [18], a hybrid satellite-terrestrial relay framework was proposed for a TSN with multi-antenna satellites, and a user-relay selection method was developed to minimize system outage probability. However, the relationship between the number of lasers for inter-satellite relaying and the resulting relay time needs further investigation.

Motivated by the above observations, we aim to maximize the system energy efficiency in a TSN while considering: 1) hot air balloons as relays between satellites and ground stations at different heights and minimum elevation angles; 2) the relation between the delay of inter-satellite relay and the number of configured lasers per satellite; 3) overall energy efficiency considering the 3C in the integrated TSN.

We first model the system, in which hot air balloons hover at different heights with different elevation angles. The time window between a hot air balloon and a satellite is determined by the minimum elevation angle and their distance. Thus, the time windows are heterogeneous among satellites due to the different heights of hot air balloons. For inter-satellite links, as there are multiple source, relay, and target satellites, the amount of traffic to be relayed among satellites can be represented by a traffic matrix. We propose an algorithm to determine the traffic matrix from finding multiple sub-traffic-matrices (STMs) in a water-filling manner [19], [20]. The algorithm guarantees that the relaying happens during the time windows of the involved source and target satellites. The next step is to relay the data traffic represented by the STMs to the corresponding target satellites. This step is completed by a well-designed configuration matrices generation algorithm. We derive the relationship between the relay time and the number of lasers, based on which we can sufficiently utilize the relay time and reduce the number of lasers used for inter-satellite communications. Finally, we maximize the system energy efficiency considering constraints including the number of lasers and the traffic delay, *etc*. By geometric programming with Taylor series approximation, we derive the optimal parameters to maximize the system energy efficiency.

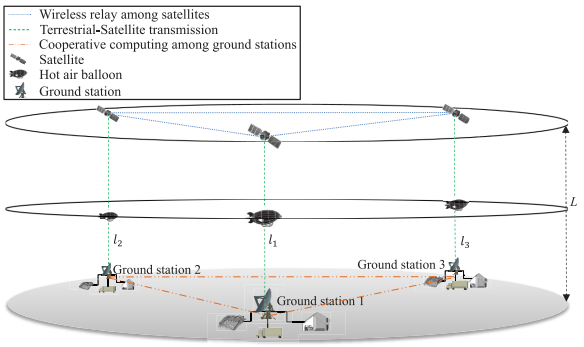


Fig. 1. System model.

Our main contributions can be summarized as follows:

1) we propose an algorithm in a water-filling manner [19], [20] to relay data among satellites. The proposed algorithm can guarantee that the data is relayed during the time windows of the involved source and target satellites, thereby avoiding overflow.
2) we propose a configuration matrices generation method to relay data traffic in multiple schedules. We also derive the optimal number of lasers subject to the available time for relay among satellites, and achieve a tradeoff between the delay and the required number of lasers for inter-satellite links.
3) we maximize the system energy efficiency considering various real-world constraints, including the period of a satellite serving a ground station, the transmission power of ground stations, hot air balloons, and satellites, and the number of lasers at each satellite, *etc*. The novelty of our system model is that the hot air balloons hover at different height with different minimum elevation angle. To the best of our knowledge, this is the first work to address the difference in hovering height and minimum elevation angle problem of hot air balloons.

The remainder of this paper is organized as follows. Section II presents the network model. Section III presents the problem formulation. Section IV presents the STMs determination. Section V presents the configuration matrices generation method. Section VI solves the collaborative multi-resource allocation optimization problem. Numerical results demonstrate the performance gain of the proposed algorithms in Section VII. We conclude the paper in Section VIII.

Notation: Standard notations are used in this paper. $|\boldsymbol{x}|$ is the number of entries in vector $\boldsymbol{x}$, and $\boldsymbol{x}(y)$ is the y-th element in $\boldsymbol{x}$, where $1 \leq y \leq |\boldsymbol{x}|$. $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denote rounding up and rounding down operations on a real number, respectively.

II. Network Model

In this section, we introduce the network model of the considered TSN by segments, starting from the hot air balloons.

A. Hot Air Balloons

The considered system model is illustrated in Fig. 1, where S satellites constitute a low earth orbit (LEO) satellite network.

The satellites provide service for S ground stations and each satellite has a dedicated ground station. Denote the number of satellites by S. We assume that the distance from all the satellites to the ground is the same, which is denoted by L. All satellites have the same orbital period given by

$$T^{\mathrm{LEO}} = 2\pi\sqrt{\frac{(L+r_E)^3}{\mu}}, \tag{1}$$

where $\mu = 398,601.58\mathrm{km}^3/\mathrm{s}^2$ is the Kepler constant, and r_E is the radius of the Earth. One hot air balloon is hovering above and serving its ground station to play the role of a dedicated relay between its ground station and the corresponding satellite. Considering the practical wireless and geographical environment, the distance from the hot air balloons to the ground stations generally differs with each other. We denote the distance from hot air balloon i ($1 \le i \le S$) to its ground station by l_i.

Due to the orbiting motion of LEO satellites, a hot air balloon can connect with a satellite periodically. The continuous serving time between the hot air balloon and the satellite in the T^{LEO} is called a time window, which is determined by the minimal elevation angle from the hot air balloon to the satellite. An air balloon can exchange data with its satellite only within the corresponding time window.

Specifically, denote the minimal elevation angle from the hot air balloon i to its satellite by β_i and the geocentric angle of the satellite i by α_i. Generally, a hot air balloon at a lower height has a larger minimum elevation to the satellite due to the thicker atmosphere the larger distance for the wireless transmission between the hot air balloon and the satellite. Then, it holds that [1]:

$$\alpha_i = \arccos\left(\frac{r_E + l_i}{L + r_E}\cos\beta_i\right) - \beta_i. \tag{2}$$

The time window of the satellite i denoted by $\widetilde{T}_i$, where $\widetilde{T}_i \le T^{\mathrm{LEO}}$, is given by

$$\begin{aligned}\widetilde{T}_i &= \frac{2\alpha_i}{2\pi}T^{\mathrm{LEO}} \\ &= 2\left(\arccos\left(\frac{r_E + l_i}{L + r_E}\cos\beta_i\right) - \beta_i\right)\sqrt{\frac{(L+r_E)^3}{\mu}}.\end{aligned} \tag{3}$$

The ground stations are connected through a wired network to form a cooperative computing resource pool as shown in Fig. 1. This network has the computing ability to generate and execute the scheme of collaborative multi-resource allocation to be developed.

B. LEO Satellites

The LEO satellite network considered in this work is shown in Fig. 2. We assume that, when the satellites circle the earth, they simultaneously arrive at the positions right above their corresponding hot air balloons. This is due to the identical orbital period T^{LEO} for all satellites. As shown in Fig. 2, the time window for each satellite is equally divided into two parts by the straight line from the ground station to the hot air balloon.

Assume that the time windows of the S satellites in a LEO satellite network have $\widetilde{T}_1 \ge \widetilde{T}_2 \ge \ldots \ge \widetilde{T}_S$. As shown in the lower part of Fig. 2, the time windows for satellites are symmetrical with respect to the vertical dashed line in the center, which denotes the instant when the satellites are right above their ground stations and hot air balloons. Within the time window of satellite i, the following tasks should be completed: uplink transmission from the ith hot air balloon to the ith satellite, which takes a time duration of T_i^B; relay for the ith satellite, which takes a time duration of T_i^Z; and downlink transmission from the ith satellite to the ith hot air balloon, which takes a time duration of T_i^D. As shown in the lower part of Fig. 2, with $\widetilde{T}_1 \ge \widetilde{T}_2 \ge \ldots \ge \widetilde{T}_S$, we assume that $T_1^Z \ge T_2^Z \ge \ldots T_S^Z$.

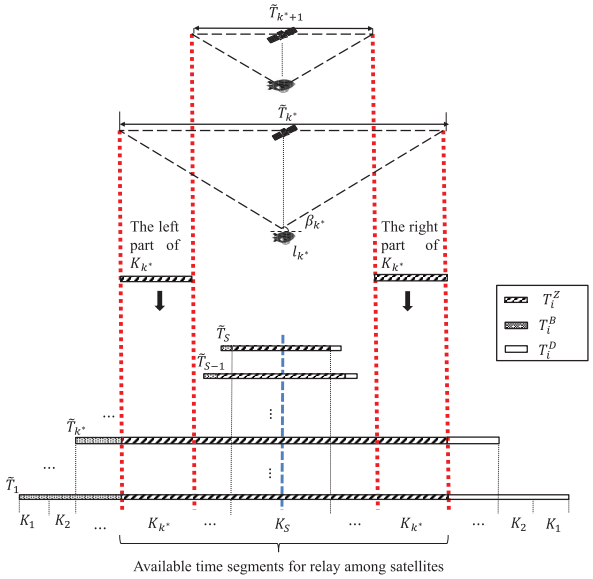


Fig. 2. LEO satellite model.

As shown in the bottom part of Fig. 2, as there are S satellites and S time windows, the largest time window can be divided into S time segments. The time segments with sequence numbers $\{k^*, k^*+1, \ldots, S\}$ are employed for traffic relay among satellites, where the value of k^* can be determined by an iterative manner according to the system performance. Denote the vth time segment by K_v. Time segment K_v, $v \in \{k^*, k^*+1, \ldots, S-1\}$, is equally divided into two parts, which are symmetrical with respect to the central dashed straight line. In the sequel, the time length of time segment K_v refers to the total length of these two parts, denoted as w_v.

For the wireless channel between a satellite and a hot air balloon, as well as that between a hot air balloon and a ground station, we use the omnidirectional path loss model as in existing works [1], [21]–[23]. The path-loss $\mathcal{C}$ (dB) for a link with distance l is denoted as:

$$\mathcal{C} = 92.44 + 20 \times \log_{10} l + 20 \times \log_{10} f, \tag{4}$$

where f is the system operating frequency and equals to 3 GHz in the S-band. The unit of the distance l is kilometer (km), and the unit of the frequency f is GHz in the above path loss model. Due to the line of sight (LOS) channels between the ground stations and the hot air balloons, and that between the

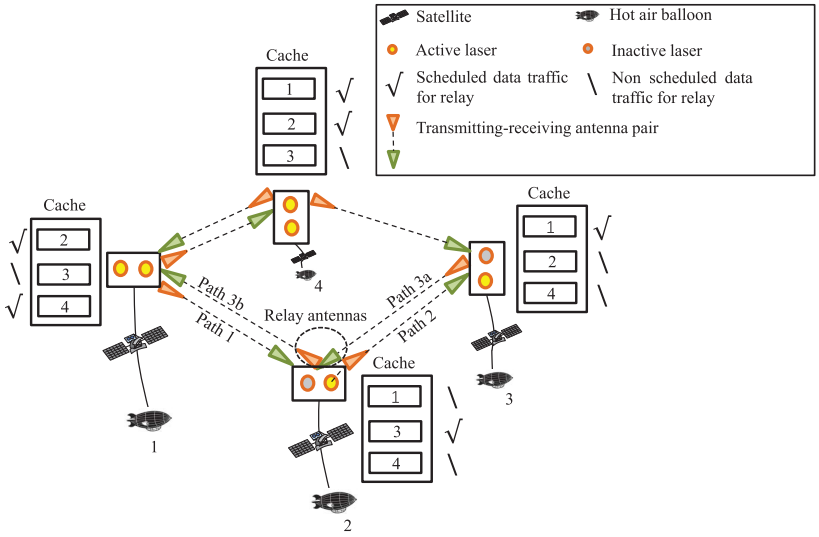


Fig. 3. Relay among satellites.

hot air balloons and the satellites, we consider only path loss in the wireless channel power gain, $|H|^2$, as follows.

$$|H|^2 = 10^{-\frac{c}{10}}. \tag{5}$$

For the LEO satellite network, we use the Starlink project launched by SpaceX as a reference [24]. LEO satellites use inter-satellite-links (ISLs) to form a satellite backbone network in a lattice grid manner [25]. For the inter-satellite links, we assume that Terahertz signal [26]–[28] or optical signal is employed. Each satellite is configured with multiple lasers to generate the signals at high frequency, and multiple antennas to transmit and receive the signals.

Signals working at high frequency are highly directional due to the narrow wave beam. As an instance, in Fig. 3, the transmitting-receiving antenna pair must be aligned. This feature can lead to the requirement of multiple schedules to relay data, because of the limited number of lasers and antenna pairs at each satellite. In Fig. 3, each satellite has a cache for storing the data to be transmitted to other satellites. We call a laser currently generating a high frequency signal as an active laser, and a laser that is currently not in use as an inactive laser. In the example shown in Fig. 3, let the maximum number of lasers be 2. Then, at least two schedules are needed to complete a full traffic relay cycle for this four satellites network. For example at satellite 1, the two lasers can generate the signals to be transmitted to satellite 2 and satellite 4, respectively. Thus, the data traffic heading to the satellite 3 will be transmitted in the next schedule.

C. Traffic Relaying Among Satellites

Traffic relaying among satellites is illustrated in Fig. 3. Each data traffic stream is generated by a laser and sent by a transmitting antenna. The data traffic is relayed among satellites and received by the receiving antenna of the target satellite. In practice, direct one-hop communication is possible between two satellites with the distance less than or equal to a pre-determined threshold. If the distance from the source satellite to the target satellite is larger than the threshold, a relay satellite is required. For instance, in Fig. 3, satellite 3 sends its data to satellite 1 via path 3 consisting of link 3a and link 3b with satellite 2 as the relay. The inter-satellite data traffic is represented by a matrix $\boldsymbol{A} = \{a_{ij}\}$, where a_{ij} is the amount of data from the source satellite i to the target satellite j. We have $a_{ii} = 0$ because satellite i does not need to send data to itself.

We assume that the number of schedules of the data is ϑ. Denote the average active number of lasers per satellite by $\overline{m}$, which is bounded by the maximum value of $m_{\max}$. Denote the power to launch a laser by P^Z. Denote the delay of aligning antennas before a scheduled transmission by δ_Y. The basic energy consumption including circuit energy consumption, *etc.*, for lasers of the satellites is

$$\varepsilon^Z = P^Z S \overline{m} \delta_Y \vartheta. \tag{6}$$

The velocity of electromagnetic signal propagation in space is $V = 3 \times 10^8$ m/s. Define the maximal distance among all paths of source-target satellites by Ω. Then, the overhead involved in data transmission among satellites is $\delta_R = \Omega / V$. Prior to a scheduled transmission, overhead δ exists, which includes the alignment of antennas among satellites δ_Y, and the data transmission time on paths of source-target satellites δ_R. The total overhead time, denoted by δ, is formulated as:

$$\delta = \delta_Y + \delta_R. \tag{7}$$

In eq. (7), for simplification, we assume that δ_Y is the same for each satellite. Since we consider the upper bound of δ_R, δ can be treated as a constant in each schedule.

In Fig. 3, a pair of transmitting-receiving antennas can act as the relay for data transmission. The average traffic relay rate can be calculated by $\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij} / T_{k^*}^Z$, where $T_{k^*}^Z$ is the maximum relay time as shown in Fig. 2. Denote the static power consumption of one laser by P_0^L W/bps, then, the static

energy consumption of lasers is

$$\varepsilon_0^L = \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij} / T_{k^*}^Z \right) \times \overline{m} P_0^L S \times \widetilde{T}_{k^*}. \tag{8}$$

Denote the dynamic power consumption of one laser by P_1^L W/bps, and the capacity of the inter-satellite channels by C_0 bps. Then, the dynamic energy consumption of lasers is

$$\varepsilon_1^L = C_0 \times \overline{m} P_1^L S \times T_{k^*}^Z. \tag{9}$$

In practice, P_1^L may depend on the distance between satellites and the signal frequency. In this work, P_1^L is assumed to be a constant for the considered TSN.

III. Problem Formulation

In this section, we present the steps of communications involved in the considered TSN and formulate a network energy efficiency maximization problem considering constraints on 3C resources.

A. Data Transmission and Relay

A service period in the system consists of 6 steps and is described as follows.

Step 1) Accumulation of data at ground stations. Data is gathered and accumulated at the ground stations for time T^A before each transmission, with T^A given by

$$T^A = n_0 \times T^{\text{LEO}}, \tag{10}$$

where $n_0 \geq 1$ represents the number of circles the satellite orbiting over the Earth in between two transmissions, and n_0 T^{LEO} is referred to as the serving period of satellites. We denote the maximum value of n_0 by $n_{\max}$ so that $n_{\max} T^{\text{LEO}}$ is the maximum tolerable system delay at the ground stations. Denote the average data arrival rate by λ_i and the power consumption for caching data during the accumulation duration by P^A W/bit. Then, the energy consumption for caching the data at the *i*th ground station is $P^A \lambda_i T^A$.

Step 2) Calculating the traffic scheduling and routing scheme. After the traffic accumulation, the computing servers of the ground stations cooperatively calculate the scheme of traffic scheduling and routing with a pre-determined algorithm. Denote the computing capacity of the cooperative computing pool at the ground stations by C^{HAB} cycles/second (cps) and the computing demand of the data scheduling and routing in a network with *S* satellite by ηS cycles, where η is the average computing load for adding one satellite in the satellite relay network. The computing delay T^C is given by

$$T^C = \frac{\eta S}{C^{\text{HAB}}}. \tag{11}$$

In terms of the energy consumption of the computing resource, we denote the power consumption of the computing by P^C W/cps. The corresponding energy consumption of computing ε^C is

$$\varepsilon^C = P^C \eta S. \tag{12}$$

Step 3) Transmission of data traffic from ground stations to hot air balloons. Let T_i^G denote the data transmission delay from the ith ground station to its serving hot air balloon with transmission power P_i^G. Assume that the transmission bandwidth is B_0, which is identical for all ground station to hot air balloon transmissions, and the power spectral density of the Gaussian white noise is σ^2 W/Hz. By eq. (5), the wireless channel power gain is denoted by $|H_i^G|^2 = 10^{-\frac{106.3+20\times\log_{10} l_i}{10}}$. Denote the antenna gain by G_T, with $|H_i^G|^2$, we can prove that

$$P_i^G = \frac{B_0 \sigma^2 l_i^2 \times 10^{11.44} \left(2^{\frac{n_0 T^{\text{LEO}} \lambda_i}{B_0 T_i^G}} - 1 \right)}{G_T}. \tag{13}$$

The delay caused by the transmission is $t_1 = l_i / V$. In order to efficiently utilize the time window of $\widetilde{T}_i$ for the inter-satellite relay and the rely between hot air balloons and satellites, we have the following constraint,

$$\begin{aligned} T_i^G + T^C + \frac{l_i}{V} &= T_i^G + \frac{\eta S}{C^{\text{HAB}}} + \frac{l_i}{V} \\ &\leq T^{\text{LEO}} - \widetilde{T}_i, \forall i. \end{aligned} \tag{14}$$

Constraint (14) means that the delay in step 2) plus step 3) should be within $T^{\text{LEO}} - \widetilde{T}_i$. Accordingly, the following three steps should be completed within the corresponding time windows.

Step 4) Data transmission from hot air balloons to satellites. The transmission time of the *i*th hot air balloon to the corresponding serving satellite is denoted by T_i^B, which should be a portion of the time window of satellite i. Denote the transmission power of the *i*th hot air balloon by P_i^B. Denote the transmission bandwidth of a hot air balloon by B_1, which is identical for all balloon to satellite transmissions. Denote the path loss based wireless channel power gain by $|H_i^B|^2 = 10^{-\frac{106.3+20\times\log_{10}(L-l_i)}{10}}$. Then, with G_T and $|H_i^B|^2$, we can prove that

$$P_i^B = \frac{B_1 \sigma^2 (L - l_i)^2 \times 10^{11.44} \left(2^{\frac{n_0 T^{\text{LEO}} \lambda_i}{B_1 T_i^B}} - 1 \right)}{G_T}. \tag{15}$$

The corresponding transmission delay is $t_2 = (L - l_i)/V$.

Step 5) Data relaying among satellites. After the data transmission from hot air balloons to satellites, the inter-satellite relaying is planned in multiple schedules for forwarding data to target satellites.

Step 6) Data transmission from satellites to hot air balloons. After the inter-satellite relaying, the data will be

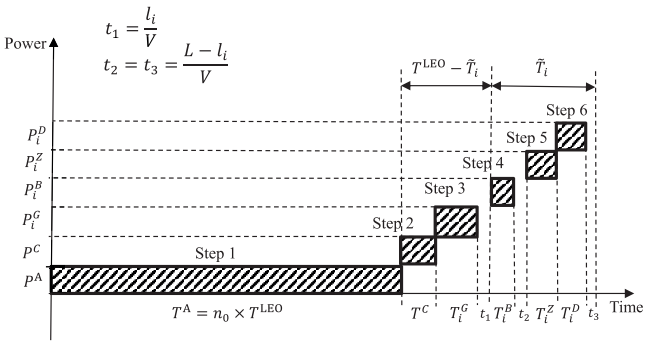


Fig. 4. Illustration of the six steps of a communication cycle.

transmitted from satellites to hot air balloons, which takes a duration of length T_i^D, and then forwarded to the target ground stations to complete a ground station to ground station data communication cycle. Denote the average data traffic arrival rate at satellite *i*, as a result of inter-satellite relaying, by μ_i. It holds that $\sum_{i=1}^{S} \lambda_i = \sum_{i=1}^{S} \mu_i$. The amount of the data received at satellite *i* in duration $\widetilde{T}_i$ is then $n_0 T^{\text{LEO}} \mu_i$. Denote the transmission power of sending data from satellite *i* to its hot air balloon by P_i^D. Denote the bandwidth by B_2, which is identical for all satellite to balloon transmissions. The corresponding channel gain is $|H_i^D|^2 = 10^{-\frac{106.3+20\times\log_{10}(L-l_i)}{10}}$. Then, with G_T and $|H_i^B|^2$, we have

$$P_i^D = \frac{B_2 \sigma^2 \left(L - l_i\right)^2 \times 10^{11.44} \left(2^{\frac{n_0 T^{\text{LEO}} \mu_i}{B_2 T_i^D}} - 1\right)}{G_T}. \tag{16}$$

The propagation delay in this step is $t_3 = (L - l_i)/V$.

Next, the data collected at hot air balloons from satellites will be used for specific objectives such as store-and-forward for traffic to the corresponding ground stations.

B. Optimization Model

In this sub-section, we focus on formulating the system energy efficiency maximization problem. First, we illustrate the constraints of data transmission in the aforementioned 6 steps as in Fig. 4. The ordinate in Fig. 4 is the average power consumption for each step, and the abscissa is the time for each step. The propagation delay is denoted by t_1 for step 3, t_2 for step 4, and t_3 for step 6, respectively, as shown in Fig. 4. To avoid data overflow, T^A in step 1 is the upper bound for each of the next steps. Since $T^A \geq T^{\text{LEO}}$ as shown in eq. (10), we have

$$T^C + T_i^G + \frac{l_i}{V} \leq T^{\text{LEO}} - \widetilde{T}_i < T^A, \tag{17}$$

and

$$T_i^B + T_i^D + 2 \times \frac{L - l_i}{V} + T_i^Z \leq \widetilde{T}_i < T^A. \tag{18}$$

There exists a trade-off between the delay and power consumption in each step, which will be discussed later.

Denote the total energy efficiency from Step 1 to Step 6 by E^{Total}. Considering the analysis above, $1/E^{\text{Total}}$ can be formulated by eq. (19), as shown at the bottom of the next page. In eq. (19), the denominator is the total amount of the data transmitted in one serving period. The overall energy consumption, used for caching, computing, wireless transmission (including the lasers), is formulated in the numerator of eq. (19). As there are many variables, we define a set of all variables in eq. (20), as shown at the bottom of the next page.

With eqs. (19) and (20), the target optimization problem can be formulated as

$$\begin{aligned} &\underset{\boldsymbol{Var}}{\text{minimize}} \ \frac{1}{E^{\text{Total}}}, \\ &\textit{s.t.} \quad (17) \ \text{and} \ \ (18). \end{aligned} \tag{P1}$$

Generally speaking, the energy consumption of antenna pairs depends on routing paths. The energy consumption of traffic scheduling for relay among satellites depends on the number of used lasers per satellite. In this work, we do not include the energy consumption for routing among the satellites in the free space and the antenna pairs, since this part of energy consumption is relatively small.

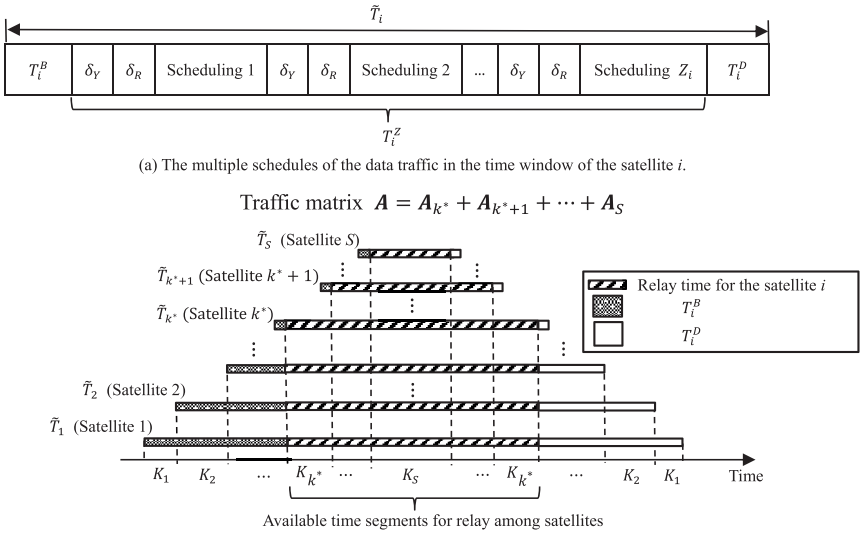
IV. STMs Determination

As aforementioned, data relayed via inter-satellite communications can be represented by a traffic matrix $\boldsymbol{A}$. To solve the optimization problem (P1), we first study STMs determination based on different time window lengths of the satellites. The objective of the STMs determination is allocating the data traffic to different time segments. The relay scheduled for each time segment must happen within the time windows of the corresponding source and target satellites. Using the case shown in Fig. 3 for instance, each satellite caches the data to be relayed to the other three satellites. The different time window lengths of satellite determine the different available times for inter-satellite relaying, and the data for relaying indicated by the traffic matrix should be decomposed into multiple STMs, each for one specific time segment.

As shown in Fig. 5 (a), the time window $\widetilde{T}_i$ contains T_i^B, T_i^D, and relay time T_i^Z. Without loss of generality, we assume that the length of the time window satisfies $\widetilde{T}_1 \geq \widetilde{T}_2 \geq \ldots \geq \widetilde{T}_S$. We also assume that the maximum available time for inter-satellite relaying in Step 5 is within the relay time of the k^*th satellite, *i.e.*, $T_{k^*}^Z$, which is formulated as

$$T_i^Z = \begin{cases} \alpha \widetilde{T}_i, & i \geq k^*, \\ \alpha \widetilde{T}_{k^*}, & i < k^*, \end{cases} \tag{21}$$

where $0 < \alpha < 1$. By eq. (21), we have $T_1^Z = T_2^Z = \ldots = T_{k^*}^Z$ due to $T_i^Z = \alpha \widetilde{T}_{k^*}$ for $i < k^*$. When T_i^B and T_i^D are determined, the value of α can be found. As we have assumed, the length of the vth time segment is w_v. Denote the length of the vth time segment by τ_v when $\alpha = 1$. Then, we have $w_v = \alpha \tau_v$, where τ_v is a constant determined by the length of time windows, and α will be optimized later.



(a) The multiple schedules of the data traffic in the time window of the satellite *i*.

(b) The segmentation of the data traffic transmission.

Fig. 5. Traffic segmentation for traffic scheduling.

The data traffic matrix $\boldsymbol{A}$ can be decomposed into S STMs $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_S\}$, each representing the data traffic to be relayed in the corresponding time segment. Since only the time segments with the sequence numbers $\{k^*, k^*+1, \ldots, S\}$ are employed for traffic relay, $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{k^*-1}\}$ are all-zero matrices, *i.e.*, $\boldsymbol{A} = \sum_{i=1}^{S} \boldsymbol{A}_i = \sum_{i=k^*}^{S} \boldsymbol{A}_i$. The element of the *i*th row and the *j*th column of $\boldsymbol{A}_v$ is denoted by a_{ij}^v, which represents the amount of data traffic transmitted from the *i*th satellite to the *j*th satellite during time segment K_v. In the rest of this Section, we will solve the STMs generation problem.

A. Geometric-Water-Filling Algorithm

The objective of the STMs determination to different time segments with various length is to guarantee that the relay time is within the time windows of the corresponding source and destination satellites and to achieve data load balancing. Water-filling is well-known for deriving the optimal solutions for power allocation problems [19] as well as load balancing problems [20]. In the following, we introduce the geometric-water-filling (GWF) algorithm. More details can be found in [19], [20].

As shown in Fig. 6, we assume a water tank with k uneven bottom steps. The width of the steps is denoted by vector $\boldsymbol{W}$ and height by vector $\boldsymbol{H}$, representing the time segments width and already allocated data traffic density (*i.e.*, data traffic per unit time), respectively. The total volume of water, D, represents the total amount of data for distribution. The dashed horizontal line represents the formed water surface. The optimal distribution of these data amount is depicted by

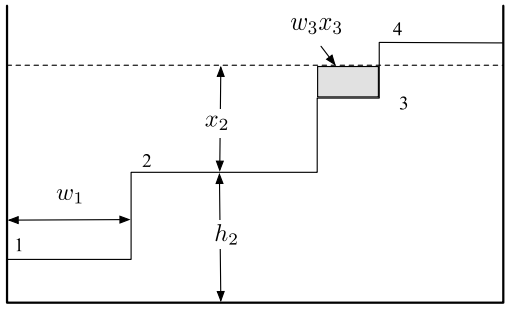


Fig. 6. Illustration of GWF algorithm.

the water-filling algorithm. The solution returns a vector, $\boldsymbol{X}$, the height of the distributed water above the steps, denoting the added traffic density for the corresponding time segment. In Fig. 6, the 4th step height h_4 is higher than the water surface, the solution of x_4 is then zero.

We can represent the geometric water-filling function as

$$\boldsymbol{X} = \mathrm{GWF}(k, \boldsymbol{W}, \boldsymbol{H}, D). \tag{22}$$

The data amount scheduled for the ith time segment is $w_i x_i$. The total data amount satisfies $D = \sum_{i=1}^{k} w_i x_i$. The detailed solution approach can be referred in [19], [20]. Without loss of focus, we shall directly apply eq. (22) as a functional block to solve the STMs for satellites traffic relay.

B. Tapped Geometric-Water-Filling Based STMs Determination

In this subsection, we present the proposed algorithm for STMs determination. During the time window for the k^*th

$$\frac{1}{E^{\mathrm{Total}}} = \frac{\sum_{i=1}^{S} \left(P^A \lambda_i n_0 T^{\mathrm{LEO}} + P_i^G T_i^G + P_i^B T_i^B + P_i^D T_i^D\right) + \varepsilon^C + \varepsilon^Z + \varepsilon_0^L + \varepsilon_1^L}{n_0 T^{\mathrm{LEO}} \times \sum_{i=1}^{S} \lambda_i}. \tag{19}$$

$$\boldsymbol{Var} = \{n_0, \overline{m}, P_i^G, P_i^B, P_i^D, T_i^G, T_i^B, T_i^D, T_i^Z, \vartheta\}. \tag{20}$$

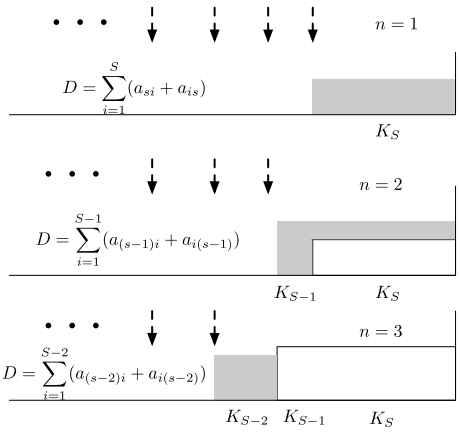


Fig. 7. Illustration of T-GWF algorithm.

Algorithm 1 T-GWF Based STMs Determination Algorithm

Initialization and Input:

Initialization: k^*: the index of the widest time window being used for relay; $\boldsymbol{A}$: traffic matrix $(S \times S)$; $\boldsymbol{A}_v$: zero matrices $(S \times S)$, $v = k^*, \cdots, S$ (STMs); $\boldsymbol{W} = [0, \cdots, w_{k^*}, \cdots, w_S]$: width of the time segment; $\boldsymbol{H}$ = zeros$(1:S)$; $D_{\text{total}} = \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij}$; $\boldsymbol{D}_{\text{allocated}} = 0$; $n = 1$.

1: **for** $(m = S : -1 : k^*)$ **do**
2: **if** $m > k^*$ **then**
3: $\boldsymbol{D}_{\text{row}} = \boldsymbol{A}(m, 1:m)$;
4: $\boldsymbol{D}_{\text{col}} = \boldsymbol{A}(1:m, m)$;
5: $D = \sum_{i=1}^{m}(\boldsymbol{D}_{\text{row}}(i) + \boldsymbol{D}_{\text{col}}(i))$;
6: $\boldsymbol{D}_{\text{allocated}} \leftarrow D_{\text{allocated}} + D$;
7: **else**
8: $D = D_{\text{total}} - D_{\text{allocated}}$;
9: $\boldsymbol{D}_{\text{matrix}} = \boldsymbol{A}(1:k^*, 1:k^*)$;
10: **end if**
11: $\boldsymbol{X} = \text{GWF}(n, \boldsymbol{W}(m:S), \boldsymbol{H}(m:S), D)$;
12: $\boldsymbol{H}(m:S) \leftarrow \boldsymbol{H}(m:S) + \boldsymbol{X}$;
13: $n \leftarrow n + 1$;
14: **for** $(v = m : S)$ **do**
15: $\eta = W(v) \cdot X(v)/D$;
16: **if** $m > k^*$ **then**
17: $\boldsymbol{A}_v(m, 1:m) \leftarrow \boldsymbol{A}_v(m, 1:m) + \widetilde{\eta} \cdot \boldsymbol{D}_{\text{row}}$;
18: $\boldsymbol{A}_v(1:m, m) \leftarrow \boldsymbol{A}_v(1:m, m) + \widetilde{\eta} \cdot \boldsymbol{D}_{\text{col}}$;
19: **else**
20: $\boldsymbol{A}_v(1:k^*, 1:k^*) \leftarrow \boldsymbol{A}_v(1:k^*, 1:k^*) + \widetilde{\eta} \cdot \boldsymbol{D}_{\text{matrix}}$;
21: **end if**
22: **end for**
23: **end for**
24: **Return**: $\boldsymbol{A}_v$, $k^* \le v \le S$.

satellite, we have a series of time segments, denoted as $K_{k^*}, \cdots, K_S$ with width $w_{k^*}, \cdots, w_S$ respectively. Data of the total traffic matrix $\boldsymbol{A}$ is allocated to STMs, $\boldsymbol{A}_{k^*}, \cdots, \boldsymbol{A}_S$, one for each time segment. We propose tapped geometric-water-filling (T-GWF) algorithm as described in Algorithm 1 to generate the STMs.

In the initialization stage of Algorithm 1, we set all STMs as all-zero matrices. The elements of the time segment width vector $\boldsymbol{W}$ are set to 0 for time segments K_1 to K_{k^*-1} and the corresponding true width for time segments K_{k^*} to K_S. D_{total} and $D_{\text{allocated}}$ represent the total data amount and the allocated amount, respectively.

In Line 1, a "For" loop is used, running from the K_Sth time segment to the K_{k^*}th time segment. Before the K_{k^*}th time segment, we start from time segment K_S, where the Sth row and the Sth column of the traffic matrix $\boldsymbol{A}$ is stored in $\boldsymbol{D}_{\text{row}}$ and $\boldsymbol{D}_{\text{col}}$ respectively in Lines 3 and 4. The total amount of data for the allocation in this round is denoted by D as shown in Line 5.

In Line 11, the GWF algorithm is applied to conduct data traffic allocation. The notation $\boldsymbol{W}(m : S)$ denotes a vector consisting of the mth element to the Sth element of vector $\boldsymbol{W}$. We design a mechanism recursively updating the STMs through GWF operation. We model that the time segments are separated by a series of taps, shown as the vertical arrowed dashed-lines in Fig.7. Starting from $n = 1$, when GWF works for the K_Sth segment, and all the taps are in their closed positions. Then when $n = 2$, the right most tap is removed and the data traffic can be distributed in time segments K_{S-1} to K_S. For all the time segments involved in the traffic allocation, the newly allocated traffic is frozen and added on the top of the current step height to update the step height for the next round GWF calculation. The taps are lifted one by one from right to left, until the last tap for the time segment K_{k^*}. We refer to the proposed Algorithm 1 as T-GWF algorithm for this nature.

In Line 12 of Algorithm 1, the step height is updated. In the following "For" loop, Before the K_{k^*}th time segment, the STMs are updated with the scaled mth row and mth column of the traffic matrix $\boldsymbol{A}$. The scale factor $\widetilde{\eta}$ is the ratio of the allocated traffic for a specific participating time segment to the total allocated traffic.

The last step of the outer "For" loop is to allocate data when $m = k^*$, *i.e.*, for the time segment K_{k^*}. The data traffic for allocation is a sub-matrix of $\boldsymbol{A}$, consisting of its first row to the k^*th row, and from the first column to the k^*th column, as listed in Line 9. In the inner "For" loop, the STMs are updated in Line 20. Different with Lines 17 and 18, where each update is carried for a row and a column respectively, each update is adding a scaled sub-matrix.

The proposed T-GWF have two great advantages. The first one is to distribute the data traffic among participating time segments to guarantee the relay being in the time windows of the involved source and destination satellites and to achieve load balance. GWF can efficiently solve this problem. The second advantage is the structure of the water-filling solution facilitating the STMs update.

Water-filling structured solution by T-GWF also reveals that the obtained water level for each time segment represents the corresponding data rate, or equivalently, data traffic density. For time segment K_{k^*} to K_S, the water level is non-decreasing. This reflects that satellites relaying data traffic is not uniform across the time segments, but tends to be more intense towards the central point of the time windows. This is

reasonable since the closer to the central point, the more data traffic uploaded from the air balloons is available.

V. Configuration Matrices Generation

After the STMs determination, the data traffic represented by the STM $\boldsymbol{A}_v$, with the dimension of $S \times S$, is relayed in the time segment of $\alpha\tau_v$. Define a set of *configuration matrices* with the dimension of $S \times S$ so that each row and each column for a configuration matrix has at most one '1', respectively. The rest of the elements for the configuration matrix are all zeroes. The time for transmitting one bit of data, *i.e.*, the inverse of the data transmission rate, is ϕ seconds, where $\phi = 1/C_0$. Assume that when $n_0 = 1$, each configuration matrix has a coefficient φ_v, which denotes the number of ϕ to relay the data traffic in the configuration matrix. Then, when $n_0 \geq 1$, each configuration matrix generated from the STM $\boldsymbol{A}_v$ has a coefficient $n_0\varphi_v$. For simplicity, we assume that φ_v is the same for each configuration matrix.

To determine the required number of active lasers per satellite, *i.e.*, m_v, we assume that Φ_v configuration matrices are firstly generated for $\boldsymbol{A}_v$. If the Φ_v configuration matrices are implemented by satellites with one laser per satellite, the delay is by $\Phi_v n_0 \varphi_v \phi + \Phi_v \delta$. Then, when m_v lasers are configured for each satellite, the satellites can simultaneously relay data traffic according to m_v configuration matrices, which can be formulated as

$$\frac{1}{m_v}\left(\Phi_v n_0 \varphi_v \phi + \Phi_v \delta\right) \leq \alpha\tau_v. \tag{23}$$

In practice, m_v should be small to reduce the power consumption and the hardware cost of the satellites. We can have

$$\frac{\Phi_v}{m_v}\left(n_0 \varphi_v \phi + \delta\right) = \alpha\tau_v. \tag{24}$$

From eq. (24), a smaller Φ_v leads to a larger φ_v. This is because that a smaller Φ_v increases ineffective occupation time of the transmitting-receiving antenna pairs after the data traffic has been relayed in a configuration matrix. On the other hand, Φ_v should be no less than S to cover all the elements in $\boldsymbol{A}_v$.

When $n_0 = 1$, we denote the maximum summation of the elements in either a row or a column of $\boldsymbol{A}_v$ by $\widetilde{A}_v$. Then, when $n_0 \geq 1$, the maximum summation of the elements in either a row or a column of $\boldsymbol{A}_v$ is $n_0 \times \widetilde{A}_v$. We propose Theorem 1 as follows.

Theorem 1 Define φ_v as follows, with a unit of bits:

$$\varphi_v = \frac{n_0 \widetilde{A}_v}{\Phi_v - S}, \Phi_v > S. \tag{25}$$

Then, $\boldsymbol{A}_v$ can be covered by at most Φ_v configuration matrices. The corresponding time delay for transmitting φ_v bits of data is $\varphi_v\phi$.

Proof: See Appendix A. ∎

Considering Theorem 1, in eq. (24), we have

$$\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \leq \alpha\tau_v, \Phi_v > S. \tag{26}$$

This guarantees the feasibility of the configuration matrices generation.

By eqs. (24) and (25), we can further derive that

$$m_v = \frac{\Phi_v}{\alpha\tau_v} \times \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta\right), \Phi_v > S. \tag{27}$$

In practice, m_v can be rounded up to an integer, *i.e.*, $m = \lceil m_v \rceil$.

To maximize the energy efficiency of the active lasers, we define the average number of active lasers $\overline{m}$ as

$$\begin{aligned}\overline{m} &= \frac{\alpha \sum_{v=k^*}^{S} m_v \tau_v}{\alpha \sum_{\widetilde{v}=k^*}^{S} \tau_{\widetilde{v}}} = \frac{\alpha \sum_{v=k^*}^{S} m_v \tau_v}{T_{k^*}^Z} \\ &= \frac{\sum_{v=k^*}^{S} \Phi_v \times \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta\right)}{T_{k^*}^Z}, \Phi_v > S.\end{aligned} \tag{28}$$

By eq. (6), ε^Z can be re-formulated by

$$\varepsilon^Z = P^Z \frac{\left(\Phi_v\right)^2}{T_{k^*}^Z}\left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta\right) S\delta_Y. \tag{29}$$

By eqs. (8) and (9), the static energy consumption ε_0^L and dynamic energy consumption ε_1^L can be formulated in eqs. (30) and (31), respectively, as follows:

$$\varepsilon_0^L = \frac{n_0 \left(\sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij}\right) P_0^L S \widetilde{T}_k \sum_{v=k}^{S} \Phi_v \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta\right)}{\left(T_k^Z\right)^2}. \tag{30}$$

$$\varepsilon_1^L = C_0 P_1^L S \sum_{v=k}^{S} \Phi_v \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta\right). \tag{31}$$

In Fig. 8, using a LEO satellite network with 4 satellites, we demonstrate an example of the configuration matrices generation for a specific time segment *v* with the sub-traffic-matrix $\boldsymbol{A}_v$. We assume that $C_0 = 4 \times 10^9$ bps, $\delta = 2$ s, $\alpha\tau_v = 1000$ s, and $\Phi_v = 7$. Then, $\phi = 1/C_0 = 0.25 \times 10^{-9}$ s/bit, *i.e.*, it takes 0.25×10^{-9} seconds for relaying 1 bit of data among satellites. By Theorem 1, we can find that $\varphi_v = 6 \times 10^9$ bits. This suggests that the transmission delay is $\varphi_v\phi = 1.5$ s for a schedule. Furthermore, m_v can be calculated by eq. (27) and rounded up to an integer, which gives $m_v = 1$.

$\boldsymbol{A}_v$ is decomposed into Φ_v configuration matrices, each of which has a coefficient φ_v. The specific elements in a configuration matrix are determined by the traffic scheduling, which in turn depends on the data traffic routing among satellites and the number of transmitting-receiving antenna pairs, *etc*. We show an example of traffic scheduling for the Φ_v configuration matrices in Fig. 8. Since $m_v = 1$, each satellite can transmit at most one data stream at any given time. Thus, it takes 7 schedules for completing the relay among satellites. For the third configuration matrix, satellite 1 transmits to satellite 4, while satellite 2 transmits to satellite 3, using satellite 1 and satellite 4 as relays. In our future work, we will discuss the data traffic scheduling in each configuration matrix and the corresponding routing schemes for relay among satellites.

VI. Collaborative Multi-Resource Allocation

In this section, we present the approximation and transformation to solve the original optimization problem of (P1).

A. Optimization With Taylor Approximation for Collaborative Multi-Resource Allocation

By the expansion of the problem (P1) with eqs. (10) to (13), (15), (16), (21) and (29) to (31) as well as the constraints of $1 \leq n_0 \leq n_{\max}$ and $\overline{m} \leq m_{\max}$, optimization problem (P1) can be transformed into problem (P2), as shown at the bottom of the page.

Let $\Phi'_v = \Phi_v - S$. For a variable x, by Taylor series, we have $e^x = \sum_{t=1}^{+\infty} x^t/t!$, which leads to

$$2^x = e^{x \ln 2} = 1 + \sum_{t=1}^{+\infty} (x \ln 2)^t / t!. \tag{32}$$

Then, optimization problem (P2) can be further transformed into problem (P3), which is a standard geometric programming problem solvable by CVX [29].

In (P3), as shown at the bottom of the page, $t_{\max}$ is a positive integer that denotes the maximum number of terms used in the Taylor series. We propose Algorithm 2 to solve problem (P3). In Algorithm 2, $t_{\max}$ is determined in an iterative manner until the values of all optimization variables converge. It is difficult to determine the optimal value of k^*, especially when the number of time segments is large. In Algorithm 2, the value of k^* is determined heuristically, and its value stops updating when the system energy efficiency does not increase further as k^* increases.

VII. Numerical Results

A. Simulations

Table I gives the general simulation parameters. The noise temperature is assumed to be 260 K. We consider satellites

$$\begin{aligned}
\text{minimize}\quad & \sum_{i=1}^{S} \frac{B_0 l_i^2 T_i^G}{n_0 G_T} \left(e^{\frac{\ln 2 \times n_0 T^{\text{LEO}} \lambda_i}{B_0 T_i^G}} - 1 \right) + \sum_{i=1}^{S} \frac{B_1 (L-l_i)^2 T_i^B}{n_0 G_T} \left(e^{\frac{\ln 2 \times n_0 T^{\text{LEO}} \lambda_i}{B_1 T_i^B}} - 1 \right) \\
& + \sum_{i=1}^{S} \frac{B_2 (L-l_i)^2 T_i^D}{n_0 G_T} \left(e^{\frac{\ln 2 \times n_0 T^{\text{LEO}} \mu_i}{B_2 T_i^D}} - 1 \right) + \frac{P^C \eta S}{10^{11.44} \times \sigma^2 n_0} \\
& + \frac{\sum_{v=k^*}^{S} \left(\frac{\Phi_v n_0 P_0^L S \widetilde{T}_{k^*} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij} \right) \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \right)}{\left(T_{k^*}^Z \right)^2} + C_0 P_1^L S \Phi_v \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \right) + P^Z \frac{(\Phi_v)^2}{T_{k^*}^Z} \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \right) S \delta_Y \right)}{10^{11.44} \times \sigma^2 n_0}, \\
\textit{s.t.}\quad & T_i^G + \frac{\eta n_0 T^{\text{LEO}} \times \sum_{i=1}^{S} \lambda_i}{C^{\text{HAB}}} + \frac{l_i}{V} + \widetilde{T}_i \leq T^{\text{LEO}}, 1 \leq i \leq S, \\
& T_i^B + T_i^D + 2 \times \frac{L - l_i}{V} + \alpha \widetilde{T}_i \leq \widetilde{T}_i, k^* \leq i \leq S, T_i^B + T_i^D + 2 \times \frac{L - l_i}{V} + \alpha \widetilde{T}_{k^*} \leq \widetilde{T}_i, 1 \leq i < k^*, \\
& \overline{m} = \sum_{v=k^*}^{S} \frac{\Phi_v}{T_{k^*}^Z} \times \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \right) \leq m_{\max}, \quad \text{and} \quad \frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \leq \alpha \tau_v, k^* \leq v \leq S, \\
& 1 \leq n_0 \leq n_{\max} \quad \text{and} \quad 0 < \alpha < 1.
\end{aligned} \tag{P2}$$

$$\begin{aligned}
\text{minimize}\quad & F\left(\alpha, n_0, T_i^G, T_i^B, T_i^D, \{\Phi'_v\}\right) \\
\textit{s.t.}\quad & \sum_{i=1}^{S} \sum_{t=1}^{t_{\max}} \frac{B_0 l_i^2 T_i^G \left(\frac{\ln 2 \times n_0 T^{\text{LEO}} \lambda_i}{B_0 T_i^G} \right)^t}{n_0 G_T t!} + \sum_{i=1}^{S} \sum_{t=1}^{t_{\max}} \frac{B_1 (L-l_i)^2 T_i^B \left(\frac{\ln 2 \times n_0 T^{\text{LEO}} \lambda_i}{B_1 T_i^B} \right)^t}{n_0 G_T t!} \\
& + \sum_{i=1}^{S} \sum_{t=1}^{t_{\max}} \frac{B_2 (L-l_i)^2 T_i^D \left(\frac{\ln 2 \times n_0 T^{\text{LEO}} \mu_i}{B_2 T_i^D} \right)^t}{n_0 G_T t!} + \frac{P^C \eta S}{10^{11.44} \times \sigma^2 n_0} \\
& + \frac{\sum_{v=k^*}^{S} \left(\frac{S(\Phi'_v + S) n_0 P_0^L \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij} \right) \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi'_v} + \delta \right)}{\alpha^2 \widetilde{T}_{k^*}} + C_0 P_1^L S (\Phi'_v + S) \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi'_v} + \delta \right) \right)}{10^{11.44} \times \sigma^2 n_0} \\
& + \frac{\sum_{v=k^*}^{S} \left(P^Z \frac{\left(\Phi'_v + S \right)^2}{\alpha \widetilde{T}_{k^*}} \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi'_v} + \delta \right) S \delta_Y \right)}{10^{11.44} \times \sigma^2 n_0} \leq F, \\
& T_i^G + \frac{\eta n_0 T^{\text{LEO}} \times \sum_{i=1}^{S} \lambda_i}{C^{\text{HAB}}} + \frac{l_i}{V} + \widetilde{T}_i \leq T^{\text{LEO}}, 1 \leq i \leq S, \\
& T_i^B + T_i^D + 2 \times \frac{L - l_i}{V} + \alpha \widetilde{T}_i \leq \widetilde{T}_i, k^* \leq i \leq S, \quad \text{and} \quad T_i^B + T_i^D + 2 \times \frac{L - l_i}{V} + \alpha \widetilde{T}_{k^*} \leq \widetilde{T}_i, 1 \leq i < k^*, \\
& \overline{m} = \sum_{v=k^*}^{S} \frac{\Phi_v}{\alpha \widetilde{T}_{k^*}} \times \left(\frac{n_0 \phi \widetilde{A}_v}{\Phi_v - S} + \delta \right) \leq m_{\max}, \quad \text{and} \quad \frac{n_0 \phi \widetilde{A}_v}{\Phi'_v} + \delta \leq \alpha \tau_v, k^* \leq v \leq S, \\
& 1 \leq n_0 \leq n_{\max}, \quad \text{and} \quad 0 < \alpha < 1.
\end{aligned} \tag{P3}$$

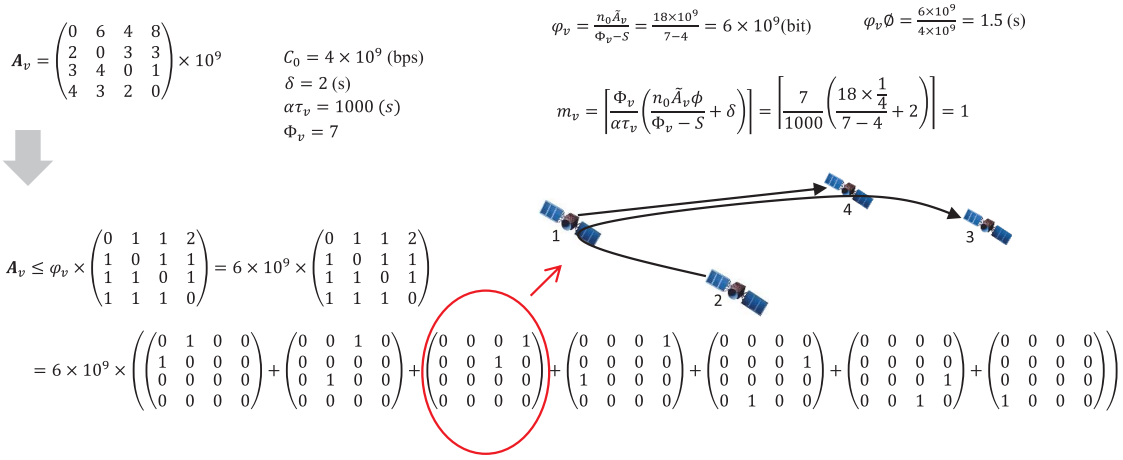


Fig. 8. An example showing the configuration matrices generation.

Algorithm 2 Algorithm for Solving Optimization Problem (P3)

Initialization and Input:
$\{B_0, B_1, B_2\}$; $\{l_i\}$; σ^2; V; $\{\lambda_i\}$; $\{\mu_i\}$; S; δ_Y; P_0^L; P^Z; T^{LEO}; Initialize $t_{\max} = 0$ and $E^{\text{Total}} = 0$; Initialize $k^* = 1$; Initialize $\widetilde{E} = 0$.

for ($k^* = 1$, $k^* \le S$, k^*++) **do**
  Implement Algorithm 1 to determine $\{\boldsymbol{A}_v\}$;
  while (E^{Total} has not converged) or ($t_{\max} = 0$) **do**
    $t_{\max} + 1 \to t_{\max}$;
    Let $\Phi'_v = \Phi_v - S$;
    Solve the geometric programming problem (P3) with $t_{\max}$ and k^*.
    Calculate $\{m_v\}$ by eq. (27);
    Let $\Phi_v = \Phi'_v + S$, $\forall v$;
    Calculate $\overline{m}$ by eq. (28);
    Calculate E^{Total} by eq. (19).
  end while
  if ($E^{\text{Total}} > \widetilde{E}$) **then**
    $E^{\text{Total}} \to \widetilde{E}$;
    $k^* + 1 \to k^*$;
  else
    break;
  end if
end for
Return: E^{Total}.

deployed at the height of 550 km above the Earth, same as the setting used in the Starlink LEO network. This leads to a maximum routing distance Ω being equal to half of the orbit circumference, *i.e.*, $\Omega = \pi(L + r_E) = \pi(550 + 6371) \approx 2.17 \times 10^4$ km. Then, the maximum delay of data routing among satellites δ_R is $\delta_R = \Omega/V = 7.25 \times 10^{-2}$ s.

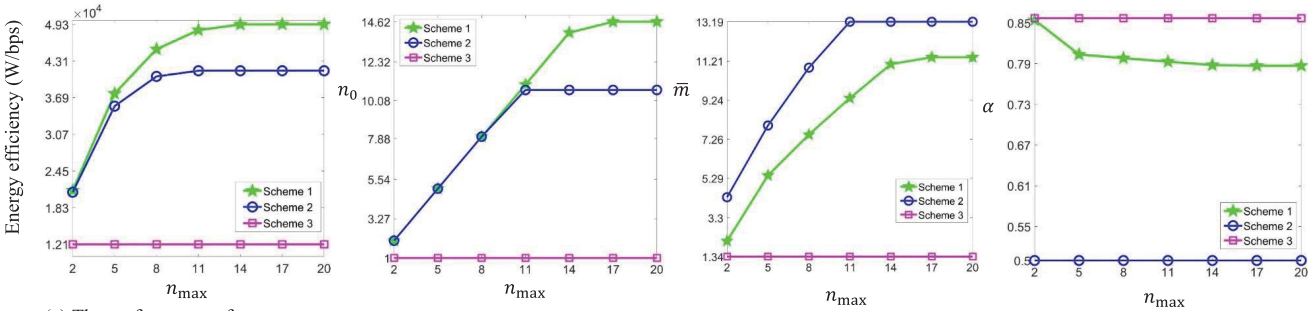
The hot air balloon with the minimum height is deployed at $l_{\min} = 20$ km, and the one with the maximum height at $l_{\max} = 75$ km above the earth. The height of the remaining $S-2$ hot air balloons is evenly distributed between $l_{\min}$ and $l_{\max}$. The hot air balloon at lower height generally has a larger minimum elevation to the satellite. We set the minimum elevation, for the hot air balloon at the maximum and the minimum height to the Earth, to $\beta^{\min} = 5°$ and $\beta^{\max} = 45°$, respectively. The minimum elevation of the other $S-2$ hot air balloons is evenly distributed between $\beta^{\min}$ and $\beta^{\max}$. The traffic matrix $\boldsymbol{A}$ is randomly generated, where the diagonal elements are 0, and the other entries of $\boldsymbol{A}$ are random and uniformly distributed over [0, θ] bits. Unless otherwise specified, we assume that $\theta = 10^4$ bits.

Since the system model that we consider is new, there is no existing algorithm that considers the same network model and scenario for comparison. To evaluate the performance of our proposed scheme, we name the scheme using both Algorithm 1 and Algorithm 2 as scheme 1. Let scheme 2 be the semi-fixed method using Algorithm 1 and Algorithm 2 with $\alpha = 0.5$ and $k^* = 1$. Let scheme 3 be the semi-fixed method using Algorithm 1 and Algorithm 2 with $n_0 = 1$ and $k^* = 1$. Then, the performance of scheme 1 is compared with those of scheme 2 and scheme 3.

In Fig. 9, the performance of the three schemes versus the maximum value of n_0, *i.e.*, $n_{\max}$, is illustrated. In Fig. 9(a), as $n_{\max}$ increases, the energy efficiency of the schemes 1 and 2 firstly increases due to the increased throughput and relatively small energy consumption. As $n_{\max}$ further increases, the energy efficiency gradually saturates because of the ever-increasing energy consumption. Scheme 3 has a fixed energy efficiency performance because its n_0 remains 1, which means the resulting energy efficiency is independent on $n_{\max}$. As shown in the figure, our proposed scheme 1 outperforms the other schemes in terms of energy efficiency. The gap is much smaller between the performance of schemes 1 and 2 than between the performance of schemes 1 and 3. This suggests that more frequent transmission in a terrestrial-satellite network cannot guarantee a larger energy efficiency, because the large power consumption only supports a limited amount of data traffic corresponding to a relatively small $n_{\max}$. In Fig. 9(b), the required n_0 increases for both schemes 1 and 2. This is because when the amount of data is relatively small, an increasing n_0 leads to a larger amount of data traffic, which enables higher energy efficiency. From Fig. 9(c), we can verify that the required $\overline{m}$ increases for schemes 1 and 2 to meet the relay requirement among satellites

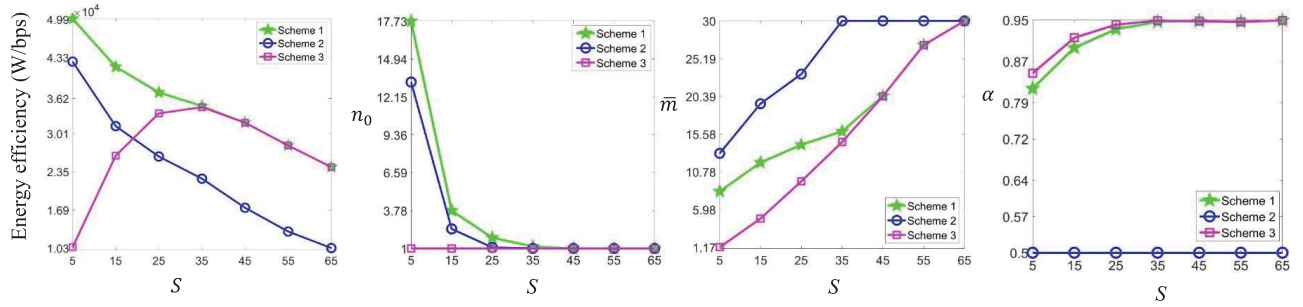
TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
$t^{\max}$	10	S	5	L	550 km	G_T	$10^{\frac{15}{10}}$ (15 dB)
V	3×10^8 m/s	η	10^{10} cycles/bit	C^{HAB}	10^{12} cycles/second	C_0	10^9 bps
B_0	10^8 Hz	B_1	10^8 Hz	B_2	10^8 Hz	P^A	10^{-10} W/bit
P^C	10^{-6} W/cps	P_0^L	10^{-15} W/bps	P_1^L	10^{-15} W/bps	P^Z	10^{-3} W/laser
Ω	2.17×10^4 km	δ_Y	1 second	$n_{\max}$	20	$m_{\max}$	50



(a) The performance of system energy efficiency with $n_{\max}$ (b) The performance of n_0 with $n_{\max}$ (c) The performance of $\overline{m}$ with $n_{\max}$ (d) The performance of α with $n_{\max}$

Fig. 9. The system performance with $n_{\max}$.

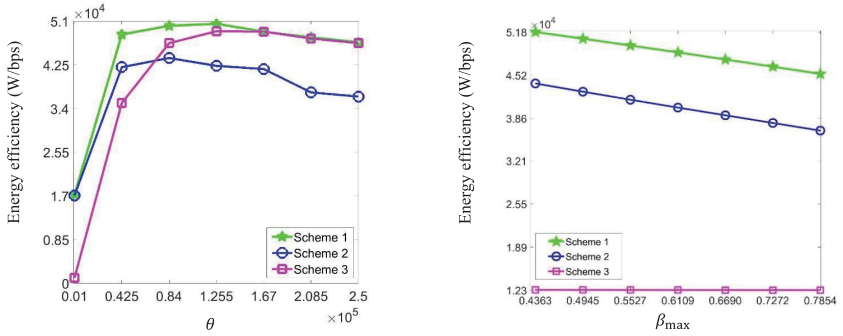


(a) The performance of system energy efficiency with θ (b) The performance of n_0 with S (c) The performance of $\overline{m}$ with S (d) The performance of α with S

Fig. 10. The system performance with S.

due to the increasing amount of data traffic. As $n_{\max}$ further increases, $\overline{m}$ remains constant for schemes 1 and 2, because the amount of data traffic does not change. In Fig. 9(d), α for scheme 1 gradually decreases, because the increasing amount of data traffic requires a longer time interval to transmit between satellites and hot air balloons.

In Fig. 10, the performance of the three schemes versus the number of satellites S is illustrated. In Fig. 10(a), as S increases, the energy efficiency performance of schemes 1 and 2 monotonously decreases. This is because a larger S results in a larger power consumption dominating the performance of energy efficiency in wireless transmission and the lasers for relay. For scheme 3, as S increases, the energy efficiency firstly increases because the increasing amount of data traffic brought by the increased satellites. As S further increases, the energy efficiency of the schemes 1 and 3 gradually merges. This is because, as S increases, the optimal n_0 is 1 to decrease the energy consumption of lasers and wireless transmission. This can be verified in Fig. 10(b), in which the required n_0 approaches 1 for all the schemes as S increases. In Fig. 10(b), the n_0 for scheme 1 is larger than that of scheme 2 when $n_{\max} < 25$. This suggests that scheme 1 can support a larger amount of data traffic due to the optimal resource allocation. In Fig. 10(c), as S further increases, $\overline{m}$ also increases due to the increasing amount of data traffic. The optimal $\overline{m}$ for scheme 2 is larger than those of schemes 1 and 3. This is caused by the smaller α for scheme 2, which can be verified in Fig. 10(d). In Fig. 10(d), as S increases, the required α for both the scheme 1 and scheme 3 gradually increases, because the increased power consumption by lasers for relay among satellites. As S further increases, the required α of schemes 1 and 3 gradually approaches to each other and does not change. This is because that n_0 approaches 1 for both schemes. Then, the performance of scheme 1 becomes the same with that of scheme 3.



(a) The performance of system energy efficiency with θ

(b) The performance of system energy efficiency with $\beta_{\max}$

Fig. 11. The performance of the system energy efficiency with θ and $\beta_{\max}$, respectively.

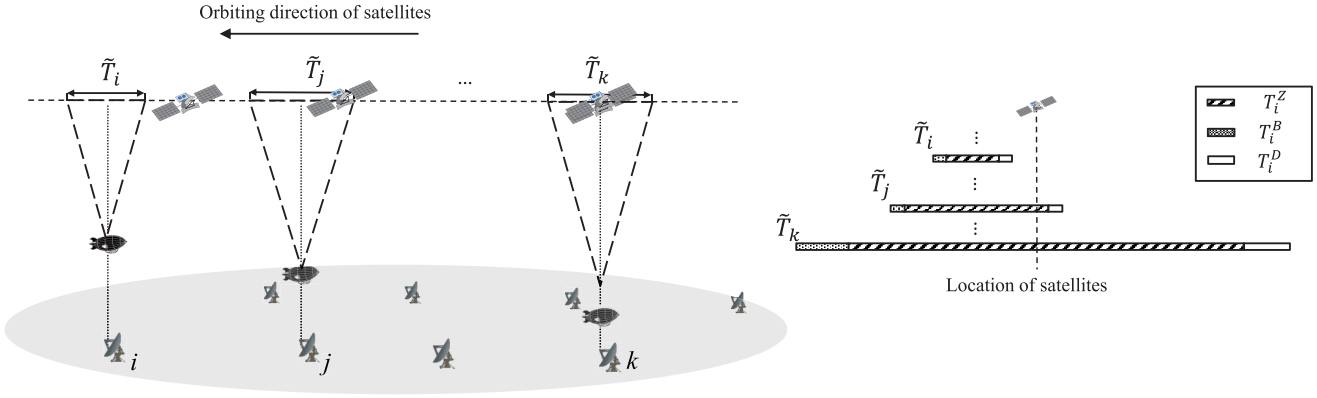


Fig. 12. A more complex relationship among time windows.

In Fig. 11, we illustrate the energy efficiency versus θ, the maximum amount of data to relay between arbitrary two satellites, and $\beta_{\max}$, the maximum minimal elevation angle, respectively. In Fig. 11(a), as θ increases, the energy efficiency firstly increases for all the schemes. This is because a larger θ can increase the data traffic amount to improve the energy efficiency. As θ further increases, the energy efficiency arrives the peak and then gradually falls for all three schemes, because the high power consumption used for communications. In Fig. 11(b), as $\beta_{\max}$ increases, the energy efficiency of schemes 1 and 2 monotonously decreases. This is because a larger $\beta_{\max}$ leads to a smaller time windows between hot air balloons and satellites. Accordingly, a shorter relay time among satellites causes higher power consumption of relay by lasers and reduced energy efficiency. For the scheme 3, the energy efficiency remains constant, because a smaller n_0 will not cause higher power consumption by lasers when $\beta_{\max}$ is not very large.

B. Discussions

We have studied a relatively complex system model in this work, including: 1) different hovering heights of the hot air balloons with different minimum elevation angle from hot air balloons to the satellites; 2) the caching, computing, and communication power consumption in this integrated system; 3) different lengths of time windows and the corresponding STMs determination for relay among satellites; 4) the configuration matrix generation to form multiple configuration matrices corresponding to the schedules; 5) an integrated optimization of the variables to maximize the overall system energy efficiency.

Two directions can be considered for future works, which can advance the research in terrestrial-satellite networks.

Firstly, joint traffic scheduling and routing scheme is of interest. As illustrated in Fig. 8, the joint traffic scheduling and routing can be considered to determine the elements in a specific configuration matrix. The elements in configuration matrices will determine the number of the active lasers and affect the number of transmitting-receiving antenna pairs. Additionally, the routing scheme of traffic relay among satellites also affects the performance of the total routing distance and the number of transmitting-receiving antenna pairs in use at each satellite.

Secondly, in this work, we assume that the time windows are overlapping and symmetrical with respective to a central line, as shown in the lower part of Fig. 2. The central line represents the assumption that the satellites will simultaneously arrive right above their corresponding hot air balloons. For a large satellite network, such assumption might be strong. Instead, as shown on the left-hand side of Fig. 12, the starting point of the time windows for the satellites may differ with each other, leading to asymmetric time windows as shown on the right-hand side of Fig. 12. In the figure, when the satellite k

is right above its hot air balloon, the time window of satellite j just starts, and the time window of satellite i is yet to start. In this scenario, the STM determination needs further investigation.

Lastly, applications, such as content provision [30], can be studied based on our proposed framework for TSN.

VIII. Conclusion

In this work, we studied the energy efficiency performance of TSN while jointly considering the caching, computing, and communication resources. We considered a practical scenario, in which different hot air balloons can be deployed at different heights with different minimum elevation angles. To effectively utilize the time windows between satellites and hot air balloons, we investigated STM determination according to the length of the time windows, through which we achieved balanced load for relaying and guaranteed that relaying happens within the time windows of the corresponding source and target satellites. We also proposed a configuration matrix generation algorithm to obtain the optimal number of lasers per satellite under the constraint of the available relay time. Then, we solved the collaborative multi-resource allocation problem to optimize the transmission power, the serving period of satellites, and the required number of lasers per satellites at the relays for the maximum system energy efficiency. Simulation results verified the effectiveness of our proposed scheme, and potential future directions were discussed.

Appendix
Proof of Theorem 1

Proof: since the maximum line summation of $\boldsymbol{A}_v$ is $n_0\widetilde{A}_v$, *i.e.*, $n_0 \times \left(\sum_{i=1}^{S} a_{ij}^v\right) \le n_0\widetilde{A}_v$ and $n_0 \times \left(\sum_{j=1}^{S} a_{ij}^v\right) \le n_0\widetilde{A}_v$. Define $\boldsymbol{Q} = \{q_{ij}\}$ and $\boldsymbol{R} = \{r_{ij}\}$, Let

$$\boldsymbol{A}_v = \frac{n_0\widetilde{A}_v}{\Phi_v - S} \times \boldsymbol{Q} + \boldsymbol{R}, \tag{33}$$

where

$$\begin{aligned}\sum_{i=1}^{S} q_{ij} &= \sum_{i=1}^{S} \left\lfloor \frac{n_0 a_{ij}^v}{n_0\widetilde{A}_v/\left(\Phi_v - S\right)} \right\rfloor \le \left\lfloor \frac{n_0 \times \left(\sum_{i=1}^{S} a_{ij}^v\right)}{n_0\widetilde{A}_v/\left(\Phi_v - S\right)} \right\rfloor \\ &\le \left\lfloor \frac{n_0\widetilde{A}_v}{n_0\widetilde{A}_v/\left(\Phi_v - S\right)} \right\rfloor \le \Phi_v - S, \end{aligned} \tag{34}$$

and $\sum_{j=1}^{S} q_{ij} \le \Phi_v - S$.

Hence, the maximum line summation of $\boldsymbol{A}_v$ is $\Phi_v - S$. According to graph theory [31], the corresponding bipartite graph of $\boldsymbol{Q}$ has the maximum endpoint degree $\Phi_v - S$. Then, $\boldsymbol{Q}$ can be decomposed into $\Phi_v - S$ configuration matrices with the coefficient of 1 for each. Moreover, by eq. (33), we have $r_{ij} < n_0\widetilde{A}_v/(\Phi_v - S)$. This suggests that $\boldsymbol{R}$ can be covered by at most S configuration matrices with the coefficient of $n_0\widetilde{A}_v/\left(\Phi_v - S\right)$ for each. Then, $\boldsymbol{A}_v$ can be covered by at most Φ_v configuration matrices, each of which the coefficient of $\varphi_v = n_0\widetilde{A}_v/\left(\Phi_v - S\right)$. ■

Acknowledgment

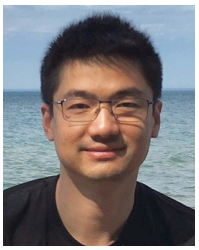
The authors thank the support and effort from the members of the Network Technologies Laboratory.

References

[1] S. Fu, J. Gao, and L. Zhao, "Integrated resource management for terrestrial-satellite systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3256–3266, Mar. 2020.
[2] C. Qiu, H. Yao, F. R. Yu, F. Xu, and C. Zhao, "Deep Q-learning aided networking, caching, and computing resources allocation in software-defined satellite-terrestrial networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5871–5883, Jun. 2019.
[3] P. V. R. Ferreira *et al.*, "Reinforcement learning for satellite communications: From LEO to deep space operations," *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 70–75, May 2019.
[4] Y. Liang, J. Tan, H. Jia, J. Zhang, and L. Zhao, "Realizing intelligent spectrum management for integrated satellite and terrestrial networks," *J. Commun. Inf. Netw.*, vol. 6, no. 1, pp. 32–43, 2021.
[5] Z. Chen *et al.*, "A survey on terahertz communications," *China Commun.*, vol. 16, no. 2, pp. 1–35, Feb. 2019.
[6] L. Zhang, Y.-C. Liang, and D. Niyato, "6G visions: Mobile ultra-broadband, super Internet-of-Things, and artificial intelligence," *China Commun.*, vol. 16, no. 8, pp. 1–14, Aug. 2019.
[7] J. Gao, L. Zhao, and X. She, "The study of dynamic caching via state transition field—The case of time-invariant popularity," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5924–5937, Dec. 2019.
[8] S. Fu, J. Wu, H. Wen, Y. Cai, and B. Wu, "Software defined wireline-wireless cross-networks: Framework, challenges, and prospects," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 145–151, Aug. 2018.
[9] B. Feng, G. Li, G. Li, Y. Zhang, H. Zhou, and S. Yu, "Enabling efficient service function chains at terrestrial-satellite hybrid cloud networks," *IEEE Netw.*, vol. 33, no. 6, pp. 94–99, Nov. 2019.
[10] C. Dai, G. Zheng, and Q. Chen, "Satellite constellation design with multi-objective genetic algorithm for regional terrestrial satellite network," *China Commun.*, vol. 15, no. 8, pp. 1–10, Aug. 2018.
[11] B. Di, L. Song, Y. Li, and H. V. Poor, "Ultra-dense LEO: Integration of satellite access networks into 5G and beyond," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 62–69, Apr. 2019.
[12] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 101–109, Jul. 2017.
[13] J. Wang, C. Jiang, Z. Wei, C. Pan, H. Zhang, and Y. Ren, "Joint UAV hovering altitude and power control for space-air-ground IoT networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1741–1753, Apr. 2019.
[14] S. Fu, Y. Tang, N. Zhang, L. Zhao, S. Wu, and X. Jian, "Joint unmanned aerial vehicle (UAV) deployment and power control for Internet of Things networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4367–4378, Apr. 2020.
[15] M. Li, N. Cheng, J. Gao, Y. Wang, L. Zhao, and X. Shen, "Energy-efficient UAV-assisted mobile edge computing: Resource allocation and trajectory optimization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3424–3438, Mar. 2020.
[16] X. Sun and S. Cao, "A routing and wavelength assignment algorithm based on two types of LEO constellations in optical satellite networks," *J. Lightw. Technol.*, vol. 38, no. 8, pp. 2106–2113, Apr. 15, 2020.
[17] B. Wu, S. Fu, and H. Wen, "Joint scheduling and routing for QoS guaranteed packet transmission in energy efficient reconfigurable WDM mesh networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 8, pp. 1533–1541, Aug. 2014.
[18] P. K. Upadhyay and P. K. Sharma, "Max-max user-relay selection scheme in multiuser and multirelay hybrid satellite-terrestrial relay systems," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 268–271, Feb. 2016.
[19] P. He, L. Zhao, S. Zhou, and Z. Niu, "Water-filling: A geometric approach and its application to solve generalized radio resource allocation problems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3637–3647, Jul. 2013.
[20] P. He, M. Li, L. Zhao, B. Venkatesh, and H. Li, "Water-filling exact solutions for load balancing of smart power grid systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 140–1397, Jul. 2018.
[21] X. Zhu, C. Jiang, L. Kuang, N. Ge, and J. Lu, "Non-orthogonal multiple access based integrated terrestrial-satellite networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2253–2267, Oct. 2017.

[22] B. Deng, C. Jiang, J. Yan, N. Ge, S. Guo, and S. Zhao, "Joint multi-group precoding and resource allocation in integrated terrestrial-satellite networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8075–8090, Aug. 2019.
[23] Y. Zhang, L. Yin, C. Jiang, and Y. Qian, "Joint beamforming design and resource allocation for terrestrial-satellite cooperation system," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 778–791, Feb. 2020.
[24] C. Henry. *SpaceX Gets OK to Re-Space Starlink Orbits, SpaceNews*. Accessed: Dec. 20, 2019. [Online]. Available: https://spacenews.com/spacex-gets-ok-to-re-space-starlink-orbits/
[25] J. Hu, L. Cai, C. Zhao, and J. Pan, "Directed percolation routing for ultra-reliable and low-latency services in low earth orbit (LEO) satellite networks," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Nov. 2020, pp. 1–6.
[26] X. Wang *et al.*, "Performance analysis of TeraHertz unmanned aerial vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16330–16335, Dec. 2020.
[27] K. Tekbiyik, A. R. Ekti, G. K. Kurt, A. Gorcin, and H. Yanikomeroglu, "A holistic investigation of terahertz propagation and channel modeling toward vertical heterogeneous networks," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 14–20, Nov. 2020.
[28] H. Yuan, N. Yang, K. Yang, C. Han, and J. An, "Enabling massive connections using hybrid beamforming in terahertz micro-scale networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–7.
[29] M. Chiang, *Geometric Programming for Communication Systems*. Hanover, MA, USA: Now, 2005.
[30] S. Zhang, J. Li, H. Luo, J. Gao, L. Zhao, and X. S. Shen, "Low-latency and fresh content provision in information-centric vehicular networks," *IEEE Trans. Mobile Comput.*, early access, Sep. 8, 2020, doi: 10.1109/TMC.2020.3025201.
[31] R. Diestel, *Graph Theory*, 2nd ed. New York, NY, USA: Springer-Verlag, 2000.

Shu Fu received the Ph.D. degree in communication and information system from the University of Electronic Science and Technology of China, Chengdu, China, in June 2016, focusing on resource allocation in wireless networks and wavelength division multiplexing (WDM) based optical networks. He is currently an Associate Professor with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China. His research interests include next generation of wireless networks, integrated networks, and artificial intelligence.



Jie Gao (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical and computer engineer from the University of Alberta, Edmonton, AB, Canada, in 2009 and 2014, respectively. From 2017 to 2019, he was a Post-Doctoral Fellow with Ryerson University, Toronto, ON, Canada. From 2019 to 2020, he was a Research Associate with the University of Waterloo, Waterloo, ON, Canada. He joined the Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI, USA, as an Assistant Professor, in August 2020. His research interests include machine learning for communications and networking, the Internet of Things (IoT) and the industrial IoT solutions, and next-generation networks in general. He was a recipient of the IEEE Transactions on Vehicular Technology Top Reviewer Award in 2018, the Ontario Centres of Excellence TalentEdge Fellowship in 2016, and the Natural Science and Engineering Research Council of Canada Postdoctoral Fellowship in 2016. He is an Editor of IEEE ACCESS and *Peer-to-Peer Networking and Applications* (Springer).

Lian Zhao (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Canada. She is currently a Professor with the Department of Electrical, Computer, and Biomedical Engineering, Ryerson University, Toronto, ON, Canada. Her research interests include wireless communications, resource management, mobile edge computing, caching and communications, and vehicular ad-hoc networks. She is a Senior Member of the IEEE Communication Society and the Vehicular Technology Society. From 2015 to 2018, she served as a Committee Member for the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grants Evaluation Group for Electrical and Computer Engineering. She received the Best Land Transportation Paper Award from IEEE Vehicular Technology Society in 2016, the Top 15 Editor Award in 2016 for IEEE Transaction on Vehicular Technology, the Best Paper Award from the 2013 International Conference on Wireless Communications and Signal Processing (WCSP), and the Canada Foundation for Innovation (CFI) New Opportunity Research Award in 2005. She has been serving as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. She served as the Co-Chair of Wireless Communication Symposium for IEEE Globecom 2020 and IEEE ICC 2018, the Local Arrangement Co-Chair for IEEE VTC Fall 2017 and IEEE Infocom 2014, and the Co-Chair of Communication Theory Symposium for IEEE Globecom 2013. She is a licensed Professional Engineer in the Province of Ontario. She has been an IEEE Communication Society (ComSoc) Distinguished Lecturer (DL).