The Korean Institute of Communications and Information Sciences

Available online at www.sciencedirect.com

ScienceDirect

ICT Express 10 (2024) 1199–1205

www.elsevier.com/locate/icte

Dynamic content-cached satellite selection and routing for power minimization in LEO satellite networks

Jeongmin Seo[a], Dongho Ham[b], Jeongho Kwak[a,*]

[a] Department of Artificial Intelligence, Daegu Gyeongbuk Institute of Science and Technology, Daegu, Republic of Korea
[b] Department of Electrical Engineering and Computer Science, Daegu Gyeongbuk Institute of Science and Technology, Daegu, Republic of Korea

Received 13 February 2024; received in revised form 9 June 2024; accepted 8 September 2024
Available online 12 September 2024

Abstract

Efficient delivery of content to areas where terrestrial Internet service is unavailable can be possible via content caching at low earth orbit (LEO) satellites. Cached content in several LEO satellites must be delivered via inter-satellite links (ISLs) with appropriate routing techniques. Until now, content caching and routing techniques have been optimized independently. To tackle this issue, the optimization of selecting a content-cached satellite and routing is jointly performed, using the example of Earth observation data cached across multiple satellites. In this paper, we first formulate a dynamic power minimization problem constrained by the queue stability of all LEO satellites, where the control variables are the selection of content-cached satellite and routing in every satellite. To solve this long-term time-averaged problem, we leverage Lyapunov optimization framework to transform the original problem into a series of slot-by-slot problems. Moreover, we prove that the average power consumption and the average queue backlog by the proposed algorithm can be upper-bounded via theoretical analysis. Finally, through extensive simulations, we demonstrate that our proposed algorithm surpasses existing independent content-retrieval algorithms in terms of power consumption, queue backlog, and fairness.

Keywords: Content delivery; Earth observation data; LEO satellite networks; Load balancing; Lyapunov optimization

1. Introduction

Recently, companies like SpaceX, Telesat, Amazon, and OneWeb have been making significant efforts to provide global Internet access using low earth orbit (LEO) satellites. In addition, LEO satellite networks not only provide typical Internet services but also handle the transmission of content delivery, such as video clips on earth observation. Recently, the satellite traffic for this content delivery has been rapidly increasing, leading to severe congestion of inter-satellite links (ISLs) between source and destination [1]. Fortunately, however, LEO satellite networks are structured in a mesh network topology, which makes it easy to route data through alternative paths. Therefore, we can alleviate link congestion and enhance transmission rate by distributing traffic load through alternative routes.

Content caching is a well-known technique that significantly reduces backhaul loads and content retrieval time in terrestrial networks. Similarly, caching content on LEO satellites can significantly reduce the retrieval time of the content thanks to the short distance between the destination and the cached satellites. For example, in the case of earth observation, observed video clips can be transmitted efficiently by pre-caching them on nearby satellites with the destination area. It would be the best to pull this content from the nearest satellite with low-congested links. However, since the wireless channels and the content request pattern change over time, the content must be dynamically retrieved from the satellite among multiple cached satellites with consideration of the time-varying parameters.

The transmission power of satellites in ISLs is crucial, as they rely on solar panels to store energy in their batteries, unlike routers in terrestrial networks [2]. High power consumption for data transmission shortens the lifespans of the satellites, especially for small LEO satellites with approximately a five-year span. Moreover, recent LEO satellites have provided more various functions, which escalate the usage of

* Corresponding author.
E-mail addresses: jeongminseo@dgist.ac.kr (J. Seo), dhham97@dgist.ac.kr (D. Ham), jeongho.kwak@dgist.ac.kr (J. Kwak).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2024.09.004

power. Therefore, it is essential to develop the power-efficient operations of the data transmission in LEO satellites.

Hitherto, a few studies have addressed content caching and routing algorithm together in satellite networks, *e.g.*, [3] where they not only independently optimize the content caching and routing but also propose the routing algorithm without a consideration of link congestion. In this paper, we are first to propose a joint cached node selection and congestion-aware routing, namely JCCR algorithm, that stabilizes transmission queues and minimizes the power consumption of all satellites in the ISL-based LEO satellite networks by leveraging Lyapunov drift-plus-penalty framework [4]. The proposed algorithm captures dynamic aspects of the wireless channel gains, network loads of ISLs, and content request rates. Our contributions are as follows.

- To the best of our knowledge, this paper is the first to optimize cached satellite selection and routing jointly. We propose the distributed JCCR algorithm that determines the cached satellite node, delivery links, and transmission power every time slot to minimize average transmission power constrained by queue stability.
- We theoretically prove the upper bounds of average transmission power and average queue backlog forr the proposed algorithm.
- Through extensive simulations, we demonstrate that JCCR consumes less power for the same delay performance and shows superior performance in terms of queue stability and fairness compared to the existing independent control algorithms.

2. Related work

Recently, advancements in satellite onboard computers enable us to efficiently deliver content by storing it directly on the satellites. For example, some papers [5,6] addressed content placement in satellites aiming to minimize the latency of the content delivery. However, they relied on one-shot optimization and did not account for dynamic variations of satellite topology and wireless channel gains. Meanwhile, there were a few studies on predicting cached contents in the satellites and routing control algorithm [3]. However, they focused solely on determining the content-cached satellites and used the shortest path routing leveraging the conventional Dijkstra algorithm.

There were various routing schemes aimed at minimizing data transmission delay [1] or transmission failure [7] for LEO satellite networks. Originally, the back-pressure routing algorithm is a widely used technique in terrestrial multi-hop networks, which dynamically distributes traffic loads over the entire networks [8,9]. In satellite networks, Deng et al. [10] redesigned the backlog of backpressure routing technique using the inter-satellite distance to suit the satellite environment while reducing delivery time and ensuring network stability. More practically, Chen et al. [11] proposed a theoretical model to estimate the ISL hop count between any users by considering constellation in mega-constellation networks based on LEO satellites. Afterward, a similar group [12] proposed an explicit or analytical algorithm to solve the Shortest distance path problem based on LEO constellation topology and ISL distance. They contribute practical insights that can be leveraged to approach the joint optimization of cached satellite selection and routing more realistically.

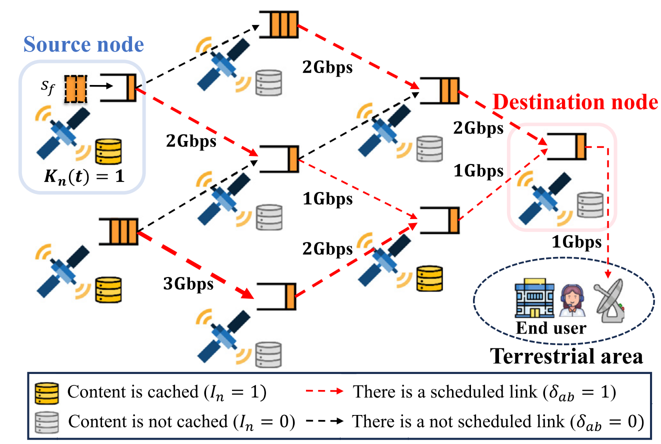

Fig. 1. A content-cached satellite selection and routing framework. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In summary, most prior works did not consider the transmission power consumption of satellites nor model the time-varying nature of wireless channels and the content requests. Moreover, there were no effective joint optimization algorithms of dynamic routing and cached satellite selection. Hence, we first propose an online-fashioned optimal cached satellite selection, delivery links, and transmission power control algorithm.

3. System model

Fig. 1 illustrates a content-cached satellite selection and routing framework. There are N satellites in networks. We consider earth observation data (*e.g.*, video clips) as the content. In this framework, we account for a series of time slots indexed by $t \in \mathcal{T} = \{0, 1, \dots, T-1\}$ where each time slot has a duration of Δt (in seconds). Each satellite node $n \in \mathcal{N}$ is connected by optical ISLs. The content is only requested from an area covered by the destination satellite. We assume that a few satellites are equipped with a cache to store content, and the same content is cached in the satellites. When a user requests the content, one of content-cached nodes[1] is selected as the source node. Then, each node determines the delivery link and transmission power considering the queue backlog and wireless channel states of the link between neighbor satellites. Finally, at the destination node, the content is downloaded to the terrestrial networks with a fixed transmission power.

[1] Hereinafter, we name satellite as node.

3.1. Satellite content delivery model

The transmission power consumed by satellite node n at time slot t is defined as $p_n(t) \in \{0, p_{n,min}, \ldots, p_{n,max}\}$. By applying Shannon's capacity formula [13], the transmission rate $\mu_{nb}(t)$ of the link is modeled as follows.

$$\mu_{nb}(t) = \delta_{nb}(t) B_n \log_2\left(1 + \frac{p_n(t) h_{nb}(t)}{\sigma_{th}^2}\right), \tag{1}$$

where $\delta_{nb}(t) \in \{0, 1\}$ represents the delivery link selection variable among candidate nodes connected to the current node n, σ_{th}^2 is the thermal noise power, and B_n is the ISL bandwidth. Here, $\delta_{nb}(t) = 1$ indicates that the link (n, b) is determined as the delivery link at time slot t. The thermal noise power is expressed by $\sigma_{th}^2 = \langle i_{th}^2 \rangle \cdot B_n$ where $\langle i_{th}^2 \rangle$ is the average value of the power spectrum density of thermal noise. Here, we have $\langle i_{th}^2 \rangle = kT_0$, where k denotes the Boltzmann's constant and T_0 represents the temperature of the system. Moreover, the channel gain $h_{nb}(t)$ between node n and b at time slot t is expressed as follows [2,14].

$$h_{nb}(t) = \eta^t \eta^r G^t G^r L^t L^r \left(\frac{w}{4\pi d_{nb}(t)}\right)^2, \tag{2}$$

where $\eta^{t/r}$, $G^{t/r}$, $L^{t/r}$ represent the optical efficiency, transmitter/receiver gain, and pointing loss factor of the transmitter/receiver, respectively, w is the operating wavelength, and $d_{nb}(t)$ represents the distance between node n and b at time slot t. The transmitter gain G^t is defined as $G^t = 16/(\Theta^t)^2$, where Θ^t is the full transmitting divergence angle. The receiver gain G^r is defined as $G^r = (D^r \pi / w)^2$, where D^r is the receiver telescope diameter. The transmitter/receiver pointing loss $L^{t/r}$ is defined as $L^{t/r} = \exp(-G^{t/r}(\theta^{t/r})^2)$, where $\theta^{t/r}$ is the transmitter/receiver pointing error.

A node n has a queue backlog $Q_n(t)$, which represents the congestion level for delivering content to the destination node. The transmitted content from node n to other nodes is denoted as $X_n(t) = \sum_{b=1}^{N} \mu_{nb}(t)$, and the content transmitted from other nodes to node n is denoted as $Y_n(t) = \sum_{a=1}^{N} \mu_{an}(t)$, where $X_n(t) \le X_{max}$, $Y_n(t) \le Y_{max}$. The caching indicator of the content in node n is denoted by $I_n \in \{0, 1\}$, *e.g.*, $I_n = 1$ when node n caches the content. Let $U(t) \in \{0, 1\}$ be an indicator of whether a user requests the content, assumed to be independent and identically distributed in every time slot. When the user requests the content, *i.e.*, $U(t) = 1$, one of content-cached nodes is determined as a source node. The source node indicator for node n can be denoted by $K_n(t) \in \{0, 1\}$. Note that $K_n(t)$ changes as time goes by. If a source node is selected, *i.e.*, $K_n(t) = 1$, the requested content with a size of s_f arrives at the transmission queue $Q_n(t)$ of source node n. Transmission queue $Q_n(t)$ for node n evolves as

$$Q_n(t+1) = \Big[Q_n(t) - X_n(t)\Big]^+ + Y_n(t) \underbrace{+ s_f K_n(t)}_{\text{arrival by content request}}, \tag{3}$$

where $[x]^+ = \max(x, 0)$. After the content arrives at the destination node, similar to (1), the arrived packets are transmitted to the terrestrial area with a constant transmit power so that the end user can receive the content.

4. Problem formulation

We consider a long-term optimization problem aiming to minimize the average transmission power consumption constrained by network queue stability of all nodes as follows.

$$\begin{aligned}
(\mathbf{P}): \ & \min_{\boldsymbol{K},\boldsymbol{\delta},\boldsymbol{p}} \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} p_n(t), \\
& s.t. \quad (\text{C1}): \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} Q_n(t) < \infty, \forall n \in \mathcal{N}, \\
& \qquad\ \ (\text{C2}): K_n(t) \in \{0, 1\}, \quad \forall I_n = 1,
\end{aligned}$$

where $(\boldsymbol{K}, \boldsymbol{\delta}, \boldsymbol{p}) \triangleq \{(K_n(t), \delta_{nb}(t), p_n(t))_{t=0}^{\infty}, \quad \forall n \in \mathcal{N}\}$. The constraint (C1) implies that the queue length of every node must be maintained as finite, *i.e.*, all content requests should be serviced within a finite time. The constraint (C2) means that cached nodes can be selected as a source node or not.

4.1. Lyapunov optimization approach

We design a joint cache selection and congestion-aware routing (JCCR) algorithm by employing the Lyapunov optimization approach [4]. To achieve this, we first define the Lyapunov function $L(t)$ and Lyapunov drift $\Delta\big(L(t)\big)$ as follows.

$$\begin{aligned}
& L(t) \triangleq \frac{1}{2} \sum_{n=1}^{N} Q_n(t)^2, \\
& \Delta\big(L(t)\big) \triangleq \mathbb{E}\Big[L(t+1) - L(t) \,\big|\, \boldsymbol{Q}(t)\Big],
\end{aligned} \tag{4}$$

where $\boldsymbol{Q}(t) = \{Q_n(t), \forall n \in \mathcal{N}\}$. Next, we define the Lyapunov drift-plus-penalty function as follows.

$$\Delta\big(L(t)\big) + V\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t) \,\big|\, \boldsymbol{Q}(t)\Big], \tag{5}$$

where V is a trade-off parameter between power and queueing delay. Here, the penalty function is the sum of the transmission power consumed during time slot t.[2] Now, the minimization problem of (5) every time slot becomes equivalent to the original problem, **(P)** [4].

Then, using the following Lemma 1, we derive an upper bound of the Lyapunov drift-plus-penalty function.

[2] Note that the minimization problem of average transmission power can be directly decomposed into a series of power minimization problem every time slot.

Lemma 1. *For all feasible sets of control variables* ($K_n(t)$, $\delta_{nb}(t)$, $p_n(t)$)*, we have:*

$$\begin{aligned} &\Delta\big(L(t)\big) + V\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t) \mid \boldsymbol{Q}(t)\Big] \\ &\le B + V\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t) \mid \boldsymbol{Q}(t)\Big] \\ &\quad + \mathbb{E}\Big[\sum_{n=1}^{N}\Big(Q_n(t)\Big(Y_n(t) + s_f K_n(t) - X_n(t)\Big)\Big) \mid \boldsymbol{Q}(t)\Big], \end{aligned} \tag{6}$$

where $B = \frac{1}{2}N\Big(\Big(Y_{max} + s_f\Big)^2 + \Big(X_{max}\Big)^2\Big)$.

Proof. From the queueing dynamics (3), utilizing the value of X_{max}, Y_{max} and the fact that $[x]^2 \le x^2$, we have:

$$\begin{aligned} Q_n(t+1)^2 \quad &\le Q_n(t)^2 + 2Q_n(t)\Big(Y_n(t) + s_f K_n(t) - X_n(t)\Big) \\ &+ \Big(Y_{max} + s_f\Big)^2 + \Big(X_{max}\Big)^2. \end{aligned} \tag{7}$$

For all nodes, by summing up (7) and the expected transmission power consumption, we obtain the upper bound of the Lyapunov drift-plus-penalty function (6). □

JCCR can be developed by minimizing the upper bound of the Lyapunov drift-plus-penalty function for each time slot.

By expressing for node n, we can reorganize the upper bound as the following problem.

$$\begin{aligned} \min_{K_n(t), \delta_{nb}(t), p_n(t)} \quad & V\sum_{n=1}^{N} p_n(t) + \sum_{n=1}^{N} s_f K_n(t) Q_n(t) \\ & - \sum_{n,b=1}^{N} \mu_{nb}(t)\Big[Q_n(t) - Q_b(t)\Big]. \end{aligned} \tag{8}$$

5. Joint cache selection and congestion-aware routing (JCCR) algorithm

The JCCR algorithm jointly determines the source node, delivery link, and transmission power ($K_n(t)$, $\delta_{nb}(t)$, $p_n(t)$) in a distributed manner every time slot. Since this algorithm is operated in an online manner, it can achieve long-term optimization without requirements of any statistics on future wireless channel states and content arrivals. Then, the JCCR algorithm is operated according to the following procedures.

Determining source node phase: For every time slot, the JCCR algorithm first checks whether there is an arrival for the content request or not. If there is an arrival for the content request, the algorithm makes a decision of a source node among content-cached nodes so as to satisfy the following problem.

$$\min_{K_n(t)} \quad \sum_{n=1}^{N} s_f K_n(t) Q_n(t), \quad \forall I_n = 1. \tag{10}$$

Here, we assume that the queue backlog information (*i.e.*, $Q_n(t)$) of cached nodes can be received via a centralized satellite such as medium earth orbit (MEO) or geostationary earth orbit (GEO) satellites. The data size of this information is approximately a few bytes, so the load of exchanging information among satellites is negligible, and thereby, there is no issue with operating the JCCR algorithm in a distributed manner. Despite its simplicity being its strength, choosing a node with a short queue backlog as the source node also significantly reduces transmission power consumption.

Algorithm 1 JCCR Algorithm: At each time slot t.

initialize: σ_{th}^2, $\eta^{t/r}$, $L^{t/r}$, $G^{t/r}$, w, and I_n.
input: $Q_n(t)$, $d_{nb}(t)$, $U(t)$. **output:** $K_n(t)$, $\delta_{nb}(t)$, $p_n(t)$.
Determine source node phase:
if a user requests a content **then**
  Find $K_n^*(t) = \text{argmin}_{K_n(t)} \sum_{n=1}^{N} s_f K_n(t) Q_n(t), \quad \forall I_n = 1$.
end if
Buffering phase:
if node n is the selected source node **then**
  Update the queue by $Q_{n^*}(t+1) \leftarrow Q_{n^*}(t) + Y_{n^*}(t) + s_f$.
else
  Update the queue by $Q_n(t+1) \leftarrow Q_n(t) + Y_n(t)$.
end if
Routing phase:
if node n is not the destination node **then**
  for all transmittable candidate nodes of node b **do**
    Calculate channel gain $h_{nb}(t)$ by (2).

  Find ($\delta_{nb}^*(t)$, $p_n^*(t)$)

$$= \underset{\delta_{nb}(t), p_n(t)}{\text{argmin}} \; V p_n(t) - \sum_{b=1}^{N} \mu_{nb}(t)\Big[Q_n(t) - Q_b(t)\Big]. \tag{9}$$

  end for
  Get the next hop b, transmit packets p on link (n, b).
else
  Download to the terrestrial with constant power.
end if
Update the queue by $Q_n(t+1) \leftarrow \Big[Q_n(t) - X_n(t)\Big]^+$.

Buffering phase: At each node, packets for the requested content are arrived at the transmission queue with respect to the determined source node. Here, the input of the transmission queue for the source node is the sum of all incoming packets from neighbor nodes and the amount of requested single content. Then, the transmission queue of the source node evolves as follows.

$$Q_{n^*}(t+1) \leftarrow Q_{n^*}(t) + Y_{n^*}(t) + s_f. \tag{11}$$

On the other hand, if the node is not a source node, the input of the transmission queue for the node is the sum of all incoming packets from neighbor nodes as follows.

$$Q_n(t+1) \leftarrow Q_n(t) + Y_n(t). \tag{12}$$

Routing phase: If node n is not a destination node, the delivery link and the transmission power are determined by solving the following problem.

$$\min_{\delta_{nb}(t),\, p_n(t)} V p_n(t) - \sum_{b=1}^{N} \mu_{nb}(t)\Big[Q_n(t) - Q_b(t)\Big]. \tag{13}$$

Here, the first term with V strives to minimize the transmission power, and the second term without V operates so as to minimize the difference in the queue backlog between node n and b. For the same V value, as the transmission power increases, the JCCR algorithm selects a delivery link that is able to stabilize queues with a small transmission power. On the other hand, as the queue difference between n and b becomes larger, the JCCR algorithm increases the transmission power to reduce the queue difference regardless of wireless channel states. The power-delay trade-off is controlled by the parameter V. As V increases, the transmission power can be reduced by trading a longer delay. By applying the determined delivery link and transmission power from (13), the packets are transmitted to the next hop, *i.e.*, node b. Then, the transmission queue evolves as follows.

$$Q_n(t+1) \leftarrow Q_n(t) - X_n(t). \tag{14}$$

On the other hand, if node n is a destination node, the content is downloaded to the terrestrial networks with constant transmission power. It is worth noting that because the JCCR algorithm stabilizes the queues of all nodes, it can compensate for the weakness of distributed load balancing in satellite networks, which is the inability to consider the entire traffic load [15].

5.1. Theoretical analysis

The long-term optimization problem **(P)** always has an optimal policy as the following Theorem 1.

Theorem 1. *For any average arrival $\mathbb{E}[s_f U(t)] = \lambda$ within the capacity region Λ,*[3] *$\lambda \in \Lambda$, there exists a stationary randomized policy π^* that determines $(K_n(t), \delta_{nb}(t), p_n(t))$ every time slot while satisfying the following condition.*

$$\mathbb{E}\Big[X_n(\delta_{nb}(t)^{\pi^*}, p_n(t)^{\pi^*})\Big] \geq \mathbb{E}\Big[Y_n(t) + s_f K_n(t)^{\pi^*}\Big],$$
$$\mathbb{E}\Big[p_n(t)^{\pi^*}\Big] = \bar{p}(\lambda), \tag{15}$$

where $\bar{p}(\lambda)$ is the minimum average transmission power that can be achieved through control to serve a given arrival rate λ. Then, we can derive the upper bounds of average transmission power and queue backlog as follows.

Theorem 2. *Assume that there exists $\epsilon > 0$ such that $\lambda + \epsilon \in \Lambda$. Then, the JCCR algorithm achieves the following performance bounds.*

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\Big[\sum_{n=1}^{N} Q_n(t)\Big] \leq \frac{B + V\bar{p}(\lambda+\epsilon)}{\epsilon},$$
$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t)\Big] \leq \bar{p}(\lambda+\epsilon) + \frac{B}{V}. \tag{16}$$

[3] It is the set of arrivals of a requested content that the system can serve within a finite time

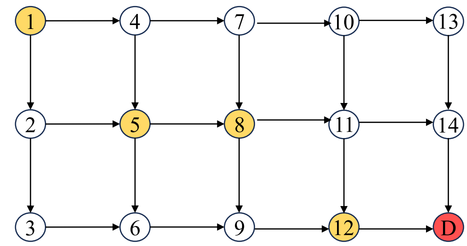

Fig. 2. Network topology. Content-cached nodes are displayed in yellow-colored, and the destination node is displayed in red-colored. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Simulation settings.

Parameter	Value
User request	$\mathcal{B}(\frac{1}{2})$
ISLs bandwidth [GHz]	10
Downlink bandwidth [GHz]	2
ISL transmission power [W]	[0, 0.1, ..., 1] [2]
Downlink transmission power [W]	20
Content size [Gbits]	290
Distance between satellites [km]	$\mathcal{N}(3000, 1000^2)$ [2]
Laser wavelength [nm]	1550 [2]
Transmitter/Receiver optical efficiency	0.8 [2]
Full transmitting divergence angle [μrad]	15 [2]
Receiver telescope diameter [mm]	80 [2]
Transmitter/Receiver pointing error [μrad]	1 [2]
Noise power density of the destination node [dBm/Hz]	−174 [16]
Downlink wireless channel gain	$\mathcal{N}(0.5, 0.1^2)$
Temperature [Kelvin]	289.85 [14]

Proof. Given the arrival rate of $\lambda + \epsilon \in \Lambda$, applying the stationary randomized control policy according to Theorem 1 to Lemma 1 is as follows.

$$\Delta\big(L(t)\big) + V\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t)\,\Big|\, \boldsymbol{Q}(t)\Big]$$
$$\leq B + V\bar{p}(\lambda+\epsilon) - \epsilon\mathbb{E}\Big[\sum_{n=1}^{N} Q_n(t)\,\Big|\,\boldsymbol{Q}(t)\Big]. \tag{17}$$

By taking expectation over the random $\boldsymbol{Q}(t)$ and summing over $t = \{0, 1, \ldots, T-1\}$, we have:

$$\mathbb{E}\Big[L(T) - L(0)\Big] + V\sum_{t=0}^{T-1}\mathbb{E}\Big[\sum_{n=1}^{N} p_n(t)\Big]$$
$$\leq TB + TV\bar{p}(\lambda+\epsilon) - \epsilon\sum_{t=0}^{T-1}\mathbb{E}\Big[\sum_{n=1}^{N} Q_n(t)\Big]. \tag{18}$$

By dividing both sides by T, disregarding the negative terms, and taking as $T \to \infty$ of each in terms of power and queue, this completes the proof of Theorem 2. □

6. Performance evaluation

Fig. 2 illustrates the network topology, following a Manhattan network [17]. Table 1 represents major parameters for our MATLAB simulations. We evaluate the performance in terms

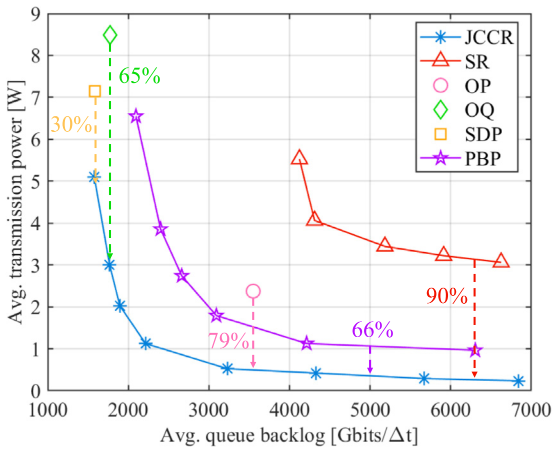

Fig. 3. Performance comparison of power-delay trade-off. The trade-off parameters V of JCCR, SR and PBP are $0 \sim 2.2 \cdot 10^{14}$, $0 \sim 8 \cdot 10^{12}$, and $0 \sim 7 \cdot 10^{13}$, respectively.

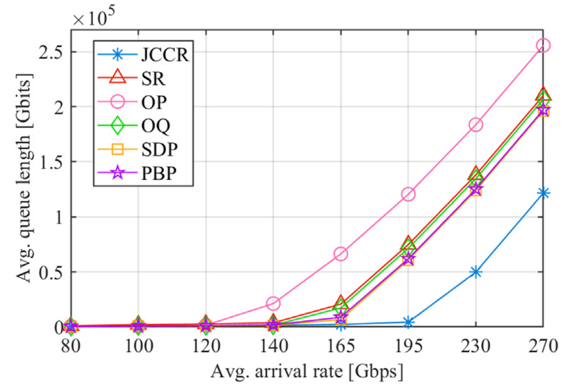

Fig. 4. Avg. queue backlog vs. avg. arrival rate. The trade-off parameter V of JCCR, SR, and PBP is 0.

of average power consumption, queue backlog, and fairness for the queueing backlog among nodes, employing Jain's fairness index [18], defined as follows.

$$\mathcal{J}(a_1, a_2, \ldots, a_n) = \frac{\left(\sum_{i=1}^{n} a_i\right)^2}{n \sum_{i=1}^{n} a_i^2}. \quad (19)$$

We compare the performance with the following existing algorithms.

- SR: This is a static routing algorithm in which the content is retrieved from node 1, and the routing path is predetermined. According to (13), each node adjusts only the transmission power in each time slot [19].
- OP: This algorithm dynamically transmits packets from node 1 with a minimum power of 0.1 W to the node with the best wireless channel [20].
- OQ: This algorithm dynamically transmits packets from node 1 with a maximum power of 1 W to the node with the shortest queue backlog in the queue [21].
- SDP: This is the shortest distance path, a widely used routing technique in existing content caching research in satellite networks. This dynamically transmits packets from node 1 with a maximum power of 1 W to the node with the best wireless channel [3,6].
- PBP: This algorithm considers transmission power based on back-pressure routing. From node 1, the delivery link and transmission power are dynamically controlled to reduce the largest queue backlog difference and power according to (13). This allows us to obtain the impact of dynamic cache node selection [22].

6.1. Simulation results

(1) Power-delay trade-off. Fig. 3 represents the trade-off between average transmission power consumption and queueing backlog. We can find that JCCR can save approximately 84% of transmission power by trading the average queue backlog by only 1000 Gbits (66 Gbits per node) every time slot. When wireless channel states are good, the transmission power increases to send more packets to the next node. On the other hand, when wireless channel states are bad, the transmission power decreases to save average satellite power, which increases the length of the transmission queue. Furthermore, we can observe that JCCR saves considerably on average transmission power compared to other algorithms with the same queue backlog. Interestingly, even when JCCR does not consider transmission power saving ($V = 0$), it achieves lower power and queue backlog simultaneously compared to both OQ and SDP, meaning that dynamically selecting the source node helps to save the average power consumption and average queue backlog. Additionally, the 66% power savings of JCCR than PBP is also thanks to the dynamic source node selection. It is worth noting that JCCR consumes less power than OP by adapting to the wireless channel and queue congestion, reducing average power consumption. Although SR considers transmission power saving, the performance of JCCR is better due to the dynamic routing path selection.

(2) Network stability. Fig. 4 depicts network stability according to the average arrival rate. Here, the arrival rate can vary depending on the content size and/or request rate of the content. We can find that JCCR outperforms other comparing algorithms in terms of achievable stability region, *i.e.*, JCCR can accommodate more content-requested arrivals to serve the content requests within a finite time since it adaptively selects the source node with respect to the dynamic queue backlog and considers queue backlog differences between nodes. In addition, the transmission queue of OP diverges at the earliest stage since it does not consider queue backlog at all.

(3) Delay fairness. Fig. 5 illustrates delay fairness among nodes. For a fair comparison, SR considers fairness only for six nodes corresponding to the path. We can observe that JCCR achieves the highest performance, with Jain's fairness index of 0.91. In addition, OP, OQ, and SDP show lower fairness performance since they do not consider fairness between nodes. While PBP and SR select the source node in a static manner, JCCR selects the source node in a dynamic manner, which improves delay fairness. Therefore, we can prove that dynamically selecting the source node and routing depending on the time-varying queue statistics are crucial in terms of traffic distribution over entire networks.

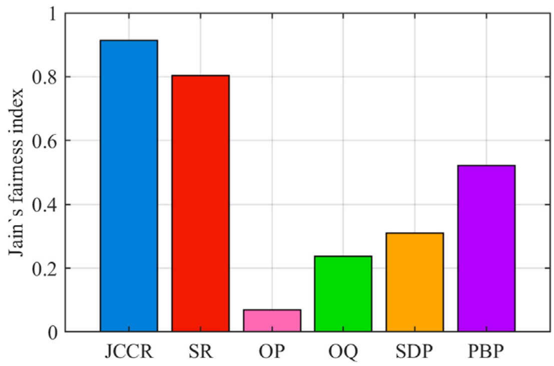

Fig. 5. Jain's fairness index. The trade-off parameter V of JCCR, SR, and PBP is $3 \cdot 10^{13}$.

7. Conclusion

As a way to prolong the lifespan of LEO satellites and avoid congestion, we proposed a content-cached satellite selection and routing framework that delivers content to users through ISLs when content is stored on multiple satellites. On top on that, we proposed the distributed JCCR algorithm that minimizes average transmission power constrained by queue stability, ensuring performance without additional approximations. Additionally, the performance bounds of the proposed algorithm for average power and average queue backlog were shown via theoretical analysis. Through extensive simulations, we observed that JCCR demonstrates superior performance in terms of queue backlog, transmission power consumption, and traffic distribution fairness compared to existing algorithms.

CRediT authorship contribution statement

Jeongmin Seo: Conceptualization, Formal analysis, Methodology. **Dongho Ham:** Resources, Writing – original draft. **Jeongho Kwak:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that there is no conflict of interest in this paper.

Acknowledgements

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2022-0-00704).

References

[1] P. Wang, S. Sourav, H. Li, B. Chen, One pass is sufficient: A solver for minimizing data delivery time over time-varying networks, in: Proc. of IEEE INFOCOM, New York City, NY, USA, 2023, pp. 1–10.
[2] J. Liang, A.U. Chaudhry, E. Erdogan, H. Yanikomeroglu, Link budget analysis for free-space optical satellite networks, in: Proc. of IEEE WoWMoM, Belfast, United Kingdom, 2022, pp. 471–476.
[3] J. Tang, J. Li, L. Zhang, K. Xue, Q. Sun, J. Lu, Content-aware routing based on cached content prediction in satellite networks, in: Proc. of IEEE Globecom, Rio de Janeiro, Brazil, 2022, pp. 6541–6546.
[4] M. Neely, Stochastic network optimization with application to communication and queueing systems, Synth. Lect. Commun. Netw. (2010) 1–211.
[5] S. Liu, X. Hu, Y. Wang, G. Cui, W. Wang, Distributed caching based on matching game in LEO satellite constellation networks, IEEE Commun. Lett. 22 (2) (2017) 300–303.
[6] X. Zhu, C. Jiang, L. Kuang, Z. Zhao, Cooperative multilayer edge caching in integrated satellite-terrestrial networks, IEEE Trans. Wireless Commun. 21 (5) (2021) 2924–2937.
[7] Z. Lai, H. Li, Y. Wang, Q. Wu, Y. Deng, J. Liu, Y. Li, J. Wu, Achieving resilient and performance-guaranteed routing in space-terrestrial integrated networks, in: Proc. of IEEE INFOCOM, New York City, NY, USA, 2023, pp. 1–10.
[8] Z. Jiao, B. Zhang, C. Li, H.T. Mouftah, Backpressure-based routing and scheduling protocols for wireless multihop networks: A survey, IEEE Wirel. Commun. 23 (1) (2016) 102–110.
[9] L. Hai, Q. Gao, J. Wang, H. Zhuang, P. Wang, Delay-optimal backpressure routing algorithm for multihop wireless networks, IEEE Trans. Veh. Technol. 67 (3) (2017) 2617–2630.
[10] X. Deng, L. Chang, S. Zeng, L. Cai, J. Pan, Distance-based backpressure routing for load-balancing LEO satellite networks, IEEE Trans. Veh. Technol. 72 (1) (2022) 1240–1253.
[11] Q. Chen, G. Giambene, L. Yang, C. Fan, X. Chen, Analysis of inter-satellite link paths for LEO mega-constellation networks, IEEE Trans. Veh. Technol. 70 (3) (2021) 2743–2755.
[12] Q. Chen, L. Yang, Y. Zhao, Y. Wang, H. Zhou, X. Chen, Shortest path in LEO satellite constellation networks: An explicit analytic approach, IEEE J. Sel. Areas Commun. 42 (5) (2024) 1175–1187.
[13] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.
[14] J. Hu, L. Cai, C. Zhao, J. Pan, Directed percolation routing for ultra-reliable and low-latency services in low earth orbit (LEO) satellite networks, in: Proc. of IEEE Vehicular Technology Conference, Victoria, BC, Canada, 2020, pp. 1–6.
[15] F. Alagoz, O. Korcak, A. Jamalipour, Exploring the routing strategies in next-generation satellite networks, IEEE Wirel. Commun. 14 (3) (2007) 79–88.
[16] J. Li, K. Xue, D.S. Wei, J. Liu, Y. Zhang, Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks, IEEE Trans. Wireless Commun. 19 (4) (2020) 2367–2381.
[17] L. Wood, A. Clerget, I. Andrikopoulos, G. Pavlou, W. Dabbous, IP routing issues in satellite constellation networks, Int. J. Satell. Commun. 19 (1) (2001) 69–92.
[18] R. Jain, The Art of Computer Systems Performance Analysis, vol. 1, Wiley, Hoboken, NJ, USA, 1991.
[19] C. Li, W. He, H. Yao, T. Mai, J. Wang, S. Guo, Knowledge graph aided network representation and routing algorithm for LEO satellite networks, IEEE Trans. Veh. Technol. 72 (4) (2023) 5195–5207.
[20] E. Ekici, I.F. Akyildiz, M.D. Bender, A distributed routing algorithm for datagram traffic in LEO satellite networks, IEEE/ACM Trans. netw. 9 (2) (2001) 137–147.
[21] T. Taleb, D. Mashimo, A. Jamalipour, K. Hashimoto, Y. Nemoto, N. Kato, Sat04-3: Elb: An explicit load balancing routing protocol for multi-hop ngeo satellite constellations, in: Proc. of IEEE Globecom, San Francisco, USA, 2006, pp. 1–5.
[22] M.J. Neely, R. Urgaonkar, Optimal backpressure routing for wireless networks with multi-receiver diversity, Ad Hoc Netw. 7 (5) (2008) 862–881.