IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 4, 2019, accepted April 23, 2019, date of publication April 30, 2019, date of current version May 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914167

Load Balancing Based on Cache Resource Allocation in Satellite Networks

ERBAO WANG, (Student Member, IEEE), HONGYAN LI, (Member, IEEE), AND SHUN ZHANG, (Member, IEEE)
State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

Corresponding author: Hongyan Li (hyli@xidian.edu.cn)

The work of E. Wang, H. Li, and S. Zhang was supported in part by the National Natural Science Foundation of China under Grant 91638202, Grant 61871456, and Grant 61871455, and in part by the National Key Research and Development Program of China under Grant 2016YFB0501004 and Grant 2017YFB1010002.

ABSTRACT During a high-speed movement, the satellites are connected intermittently, so the queue length becomes larger and a cache overflow appears. In this paper, the abundant storage resources of the multilayered satellite network (MLSN) are used to avoid the packet loss caused by a cache overflow of the Low Earth Orbit (LEO) satellites. However, due to the limited storage space of the Geostationary Earth Orbit (GEO) satellites, an effective load balance scheme which addresses two problems: LEO satellites competition in a non-cooperative fashion and content popularity utilization, is needed. Therefore, we propose a load balancing scheme based on the Stackelberg game, containing Members of a Game Algorithm and Distributed Cache Price Bargaining Algorithm. In addition, a storage technology based on content popularity (Popularity Matching Algorithm) is introduced. The numerical results show that the proposed methods are effective in pricing, cache resource allocation of GEO satellites, and load balancing of LEO satellites.

INDEX TERMS Cache resource allocation, load balancing, Martingale theory, popularity of contents, Stackelberg game, satellite networks.

I. INTRODUCTION

Satellite network has been becoming a major information infrastructure with global coverage. Therefore, much attention has been paid to this field recently, especially, to the multilayered satellite architectures [1]–[8]. Namely, it is expected that the multilayered satellite architectures will dominate in the future satellite networks because the combination of different satellites layers can obtain much better performance than these layers separately. In the multilayered satellite networks [9], load balancing strategy is designed to detour traffic from LEO satellites to the GEO satellite [1], [2], [10], [11]. Nishiyama *et al*. [1] proposed a load balancing scheme by adopting a traffic distribution model based on the network capacity estimation and theoretical analysis of the congestion rate of each layer, where the model was assumed as a square lattice. However, this model is not general, and it does not take into account the channel influence. Kawamoto *et al*. [10] introduced a model based on the Multilayered Satellite Networks (MLSNs) by distributing the packets between the two layers to minimize the packet delivery delay, where the effect of the method on the packet delivery delay was analyzed by considering propagation and queuing latencies. However, in the scheme, the access and physical channel were not considered. Song *et al*. [2] proposed a traffic-light-based intelligent routing strategy for the Non-Geostationary (NGEO) satellite networks, where a set of traffic lights was used to indicate the congestion status. Especially, the packet route could be adjusted dynamically according to the real-time traffic light color. In this scheme, a public waiting queue could temporarily store packets which were not forwarded. However, the cache size was limited, which means the public waiting queue had a size limitation. In [11], an optimal portion of traffic flows of the low layer network is decided based upon the newly arrived traffic estimate, where the model is considered as square lattice. However, the analysis does not take into account the effects of the channel.

In addition, in the multilayered satellite networks, due to the limited storage resources of the upper-layer satellite, when many nodes need to upload data, they compete. Therefore, we assume that the game theory can be used to solve the resource allocation problem (e.g. [12]–[16]). Accordingly, we

The associate editor coordinating the review of this manuscript and approving it for publication was Bora Onat.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

propose a cache resource allocation method based on the Stackelberg game consisting of a leader (an GEO satellite) and multiple followers (LEO satellites), where the storage resource of the leading GEO satellite must first fulfill its demand, and then the remaining storage space can be rented to the LEO satellites. Considering the remaining (surplus) storage space of the leading GEO satellite as a resource, the system is optimized so that the nodes congestion can be alleviated.

The condition for finding the solution to the proposed game is that the traffic backlog of LEO and GEO satellites has to be known. However, the information about the backlog change is difficult to obtain timely, so it is estimated by the Martingale theory, providing the advantage of transferring the scheduling problem to the time-shifting martingales. F.Ciuiu [17] proposed a novel representation of a queueing system based on the martingale-envelope. Particularly, in [18], the backlog bound was given by abstracting the random access and physical channel as a service process. In this work, we utilize the martingale theory to estimate the backlog of LEO and GEO satellites. With the aim to use storage resources effectively, the data attribute is analyzed, namely the content popularity. The main contributions of this paper are summarized as follows:

1) To obtain an effective load balancing in a satellite network, the backlog of LEO and GEO satellites is estimated, and the characteristics of a terrestrial-satellite link are fully considered. To describe the interference uncertainty, the Gilbert-Elliott model is adopted. In addition, the fading channel characteristics are considered. A load balancing scheme based on the Stackelberg game containing the game members and the Distributed Cache Price Bargaining Algorithm, is proposed. The storage technology based on content popularity (Popularity Matching Algorithm) is also proposed.
2) Since it is difficult to obtain the actual real-time cache occupancy, it is estimated by the martingale process, where the terrestrial-satellite link is abstracted as a service supermartingale, and the arrival process of satellite data is considered to be the departure process of a satellite user after being served by the uplink wireless channel according to the network calculus theory. In our uplink wireless channel model, the Gilbert-Elliott based channel model and the Slot-ALOHA random access protocol are fully used.
3) To minimize the congestion of LEO satellites and maximize the profit of the GEO satellite, we propose a Stackelberg game model based on the congestion level of LEO satellites, where the non-convex problem is transformed into a convex optimization problem using a threshold function which is a strictly monotonically-increasing function of the number of game participants. In the proposed game, a new profit model is developed for LEO satellites by assuming the GEO cache space is a resource. In this way, both the congestion level of the cache space and the non-cooperative nature of LEO satellites are taken into consideration.

The remainder of the paper is organized as follows. Section II introduces the system and channel model and defines the data arrival and service processes. Section III proposes a load balancing scheme based on the Stackelberg game. Section IV presents the simulation results and analysis. Section V concludes the paper.

II. RELATED WORK

In this section, we mainly summarize some important related studies, which is are grouped into two categories: terrestrial wireless caching networks and satellite caching networks.

A. TERRESTRIAL WIRELESS CACHING NETWORKS

Caching the most popular content on a network edge can reduce the content download delay and consumption of backhaul link resources. Specifically, the wireless network edge-caching includes device-to-device (D2D) [19]–[21], small base station [22], [23], base station [24], and virtual network [25].

In [19], a method for maximizing the hit rate based on a truncated Zipf distribution is proposed. In [20], the performance of large-scale cache-enabled D2D networks is studied, and the optimal probabilistic caching strategy of mobile helpers was investigated based on the content delivery probability. In [21], a joint transmission and caching strategy was proposed based on the pre-awareness of user needs. In [26], to solve the congestion of the backhaul channel, a joint cache and downlink resource optimization framework was proposed, where the D2D communication technology was fully utilized. In [22], based on content popularity, two collaborative transport mechanisms were proposed to improve cache efficiency. In [25], a collaborative optimization scheme for wireless spectrum resources and content placement was proposed to minimize the rejection rate of user requirements in a virtual network environment.

In [23], a commercialized small-cell caching was considered to provide faster local video transmissions to users. To simultaneously maximize the profit of network service providers and video retailers, the Stackelberg game was used as a model, and the Stackelberg equilibrium was obtained by solving a non-convex optimization problem. The mechanism was characterized by the base station as a resource for pricing and trading. At the same time, In [27], a pricing-based content placement mechanism was also proposed in named data networking to maximize the profit of both service providers and Internet service providers. Essentially, the cache was traded as a resource.

B. SATELLITE CACHING NETWORK

The initial Cache-Satellite Distribution Service (CSDS) was mainly based on proxy services [28], [29], in other words, communication was performed through ground stations. In [28], a model for CSDS was proposed to capture the intricate interaction between the proxy caches, where the

authors analyzed the business flow properties from the HTTP log, which was similar to the popular distribution. In [29], the satellite ground station collected the selected content information and broadcasted it periodically to the participating proxies via satellite. At present, content placement is primarily considered to have an on-board cache [30]–[32]. S. Liu [30] proposed a caching algorithm by optimizing the content placement based on a many-to-many matching game in the LEO satellite constellation network. A two-layer caching model for content delivery was introduced in [31], where a nonlinear integer programming problem was solved by the genetic algorithm.

TABLE 1. Summary of notation.

Notation	Description
K	Number of LEO satellites
n_0	Zero-mean additive white Gaussian noise
p_s	Satellite transmitting power
H_{su}	Channel coefficient
w_{su}	Bandwidth of satellite-terrestrial link
p_{gb}	Transition probabilities from g to b
$I(T)$	Indicator function
p_m	Probability of user's transmission
P^{us}	State transition matrix
λ_f	Arrival rate content f
$A^h(T)$	Accumulation of high-priority arrival
$A^f(T)$	Accumulation of low-priority arrival
$M_{A^t}(T)$	Martingale-envelope of arrival process
$M_S(T)$	Martingale-envelope of service process
$\sigma_{SP,L}$	SP based Backlog
$\sigma_{FIFO,L}$	FIFO based Backlog
κ_h	High-priority arrival rates
κ_f	Low-priority arrival rates
$S^q(T)$	Queue service process
κ_s	Queue service rate
$S(T)$	Equivalent service process
κ_{ss}	Equivalent service rate
R_a	Arrival rate
R_{su}	Downlink channel service rate
R_{us}	Uplink channel service rate
f	Content index
F	Number of Content
s	Coefficient of popularity distribution
$\boldsymbol{\eta}$	Pricing vector
$s_{L_k}(q_{L_k})$	Cache occupation function
$\boldsymbol{q}$	Portion vector
$I(q_{L_k})$	Congestion Index (CI)

III. SYSTEM MODEL

In this section, we present the system model and introduce the channel model, arrival process, service process, and Zipf distribution. The notations used in the following are listed in Table 1.

The multilayered satellite network is presented in Fig.1, where it can be seen that in such a network, one GEO satellite covers K LEO satellites. The GEO and LEO satellites communicate with their ground stations via a terrestrial-satellite link. The GEO satellite communicates with its associated LEO satellites along the inter-satellite links, which are assumed to be ideal because of the good electromagnetic scattering environment. Furthermore, it is assumed that coverage areas of the GEO satellites do not overlap.[1] Therefore, in the following, the GEO satellite and its associated (served) K LEO satellites are considered.

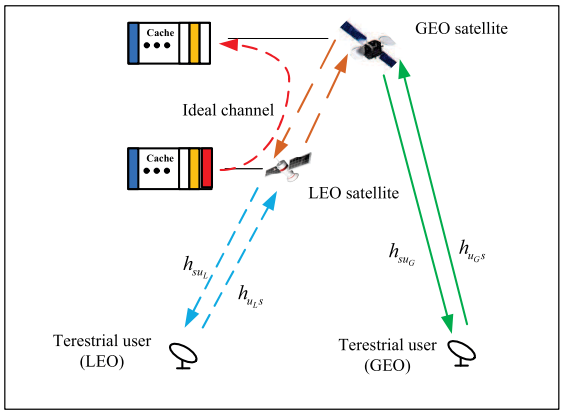

FIGURE 1. An example of caching system for satellite networks.

In this work, we assume that each satellite node is equipped with several storage chips to construct the cache space. However, the volume of the GEO satellite is much bigger than that of LEO satellites. Correspondingly, the size of the cache space of the GEO satellite is much larger than of the LEO satellites. Thus, the probability of congestion of the LEO satellites may be much bigger than that of the GEO satellite, which means that the GEO satellite may have some idle cache space to place the received content from the LEO satellites. However, uploading content of the LEO satellites to the GEO satellite is closely related to the content popularity, space size of the GEO satellite, and queuing method. The storage system of LEO/GEO is presented in Fig.1, where the right rectangle represents the cache.

A. CHANNEL MODEL

In order to analyze the content backlog in different layers of satellites, we introduce the channel models of the uplink and downlink channels in the following. In the channel models, the term "satellite" refers to LEO or GEO satellite. Similarly, the term "satellite users" corresponds to the LEO satellite users or GEO satellite users, respectively.

1) DOWNLINK WIRELESS CHANNEL MODEL

The signal received through the downlink at the satellite user side can be expressed as

$$Y_{su} = \sqrt{p_s} H_{su} x_{su} + n_0, \tag{1}$$

[1]Since this work mainly analyzes the load balancing of satellite data streams, we do not consider the overlap of coverage, which is beyond the scope of our analysis.

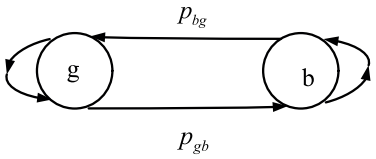

FIGURE 2. Gilbert–Elliott model can be abstracted as a two-state Markov chain.

where n_0 represents the zero-mean additive white Gaussian noise at the satellite user, p_s denotes the satellite transmitting power, and x_{su} is the corresponding signal. In addition, H_{su} denotes the channel from the satellite to the ground user side, and it can be characterized by the Gilbert-Elliott model to include the interference uncertainty. Explicitly, the evolution states of the terrestrial-satellite link can be depicted by a two-state Markov chain model which is shown in Fig.2, where g denotes the *good* state; b denotes the *bad* state, and p_{gb} and p_{bg} separately represent the transition probabilities from g to b and from b to g state. The corresponding state transition matrix can be expressed as

$$P = \begin{bmatrix} 1 - p_{gb} & p_{gb} \\ p_{bg} & 1 - p_{bg} \end{bmatrix}. \tag{2}$$

In 'good' state, the transmission rate can be expressed as

$$R_{su} = w_{su} \log \left(1 + \frac{\sqrt{p_s}||h_{su}||^2}{n_0} \right), \tag{3}$$

where w_{su} is the bandwidth of a satellite-terrestrial link, and $||h_{su}||^2$ is the channel coefficient in the *good* state. Moreover, the probability density function (pdf) of h_{su}'s [33], [34] can be defined by

$$f_{||h_{su}||^2} = \frac{1}{2b_{su}} \left(\frac{2b_{su}m_{su}}{2b_{su}m_{su} + \Omega} \right)^{m_{su}} \exp\left(-\frac{x}{2b_{su}} \right) \cdot {}_1F_1 \left(m_{su}; 1; \frac{\Omega x}{2b_{su}(2b_{su}m_{su} + \Omega)} \right), \tag{4}$$

where Ω is the average power of the line-of-sight (LOS) component, $2b_{su}$ is the average power of the multipath component, and m_{su} is the Nakagami-m parameter.

2) UPLINK WIRELESS CHANNEL MODEL

The signal received through the satellite uplink can be written expressed as

$$Y_{us} = \sqrt{p_u} H_{us} x_{us} + n_0, \tag{5}$$

where n_0 represents the zero-mean additive white Gaussian noise at the satellite user side, p_u denotes the transmitting power of a satellite user, and x_{us} is the corresponding signal. Generally, the downlink channel works in the orthogonal mode, while the uplink adopt S-ALOHA network protocol because of the large propagation delay. To meet the challenge, by using the polarization transmission structure in the satellite uplink [35]–[39],[2] the access is presented by the value of the indicator function $I(T)$, where 1 denotes the successful access, and 0 denotes unsuccessful access.

$$I(T) = \begin{cases} 1, & p_m(1-p_m)^N \\ 0, & 1 - p_m(1-p_m)^N \end{cases}, \tag{6}$$

In (6), p_m denotes the probability of each satellite user accessing the channel, and N is the number of users who to access the channel at the same time. Further, H_{us} denotes the channel from the ground user to the satellite and can be characterized by cascading the random access channel and Gilbert-Elliott model. Similar to the downlink, the corresponding state transition matrix P^{us} can be written as

$$\begin{bmatrix} (1 - p_{gb}^{us})p_m(1-p_m)^N & p_{gb}^{us}p_m(1-p_m)^N \\ p_{bg}^{us}1 - p_m(1-p_m)^N & (1 - p_{bg}^{us})(1 - p_m(1-p_m)^N) \end{bmatrix} \tag{7}$$

In the *good* state, the transmission rate can be expressed as

$$R_{us} = w_{us} \log \left(1 + \frac{\sqrt{p_u}||h_{us}||^2}{n_0} \right), \tag{8}$$

where w_{us} is the bandwidth of a satellite-terrestrial uplink, and h_{us} is the channel coefficient in the *good* state.

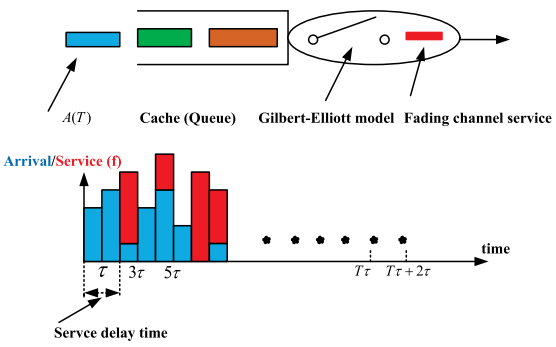

FIGURE 3. An example of arrival and service of content.

B. ARRIVAL PROCESS AND SERVICE PROCESS

The arrival service model considered in this work is presented in Fig.3. In this model, λ_f content arrives in a time slot τ, and the state of the Gilbert-Elliott channel keeps constant during a time slot but changes from one slot to another. The amount of the service content depends on both the fading channel characteristics in the good state and queuing service rate in the cache. For notation simplicity, we assume that a time slot lasts for τ second ($\tau = 1$ means that $A^t(T\tau) = A^t(T)$, $T = 1, 2, \ldots,$).

[2]In the references, the early S-Aloha and the recent satellite network access protocol are provided. The latest protocols mainly consider how to improve the access success rate, so our analysis structure can be directly used without a need for many changes. Therefore, we adopt the S-Aloha protocol.

In our system, the content arrival process at a satellite node can be divided into three sub-processes, i.e., data generation process at a satellite user side, channel access process of a satellite user, and the Gilbert-Elliott channel service process. Next, in order to introduce a static prioritized (SP) scheduling algorithm, we divide the arriving traffic into two types high-priority arrival $A^h(T)$ and the low-priority arrives $A^f(T)$ and $A^t(T) = A^h(T) + A^f(T)$. Therefore, the following can be derived using the Martingale-Envelope theory.[3]

Lemma 1: The arrival process $A^t(T)$ is a supermartingale, where the martingale-envelope [17] is defined by

$$M_{A^t}(T) = h_{A^h}(a_T)h_{A^f}(d_T^f)h_{s^q}(s_T) \times e^{\theta(A^h(\kappa,T)-(T-\kappa)\kappa_f+A^f(T)-Tk_h+T\kappa_s-S^q(T))}. \quad (9)$$

In (9), the part $h_{A^h}(a_T), h_{A^f}(d_T^f), h_{s^q}(s_T)$ represents a monotonically increasing function, and κ_h and κ_f are the high-priority and low-priority arrival rates of an LEO/GEO satellite, respectively, $S^q(T)$ is queue service process, $\theta > 0$ and κ_s depends on the queue service rate in the cache.

Proof: Please refer to A section in the appendix.

In the above proof, the arrival process involves the generation of user data and the satellite channel service process. According to the idea of system cascading,[4] we regard this arrival process as the arrival process of the satellite cache.

Corollary 1: The equivalent service process $S(T)$ is a supermartingale, where the martingale-envelope is defined as follows

$$M_S(T) = h_S(s_T)e^{\theta^*(\{S(T)-T\kappa_{ss}\})} \quad (10)$$

where $h_S(s_T)$ is a monotonically increasing function, $\theta^* > 0$ and κ_{ss} depends on R_{su}. Note that κ_{ss} is the rate at which the content arrive at the user.

C. ZIPF DISTRIBUTION

A cache of an LEO satellite is capable of storing F different contents, whose popularity follows the Zipf's law as [29], [31], [40] which is given by

$$t_f = \frac{1/f^s}{\sum_{i=1}^{F}(1/i^s)}, \quad \forall f \in \mathcal{F}, \quad (11)$$

where f denotes the ranked content according to the descending order of popularity, s is the coefficient which controls the popularity distribution of the contents, and $0 \leq s \leq 1$. At $s = 0$, the distribution is uniform, which means that all the content will be identical regarding the popularity.

Fig.4 shows the popularity varied with different types of content and different values of the Zipfs factor s. As the Zipfs factor s increased, the popularity distribution of content got skewed according to (11). Therefore, the effectiveness of the load balancing could be improved by setting a different popularity parameter s to obtain an effective content selection.

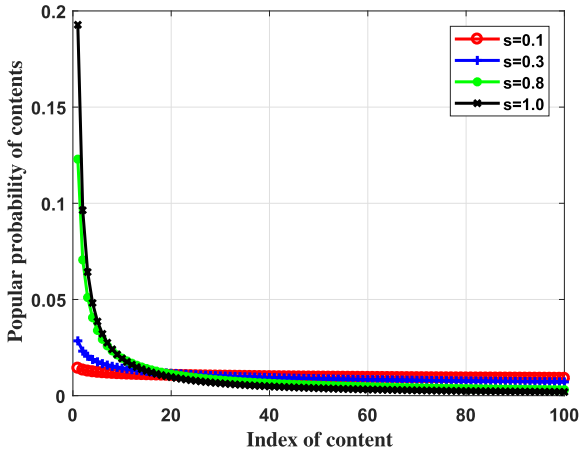

FIGURE 4. The popular probability of contents vs. the index of content under different popular parameter *s*.

IV. LOADING BALANCE SCHEME BASED ON STACKELBERG GAME

With the development of the content delivery networks (CDNs), the cache has stopped being only a small memory, and it has become feasible to store the time tolerance content in the cache. In this paper, we aim to make full use of the cache resources of satellites to balance the content placement. Specifically, in this section, we estimate the backlog by the martingale method, which is described in detail in the following. In addition, we present also a competition mechanism of the multiple nodes.

A. BACKLOG ESTIMATION

Since the information about the backlog change is difficult to obtain timely, we estimate the backlog by using the martingale method. As it is well known, the content backlog depends on the data generation rate, transmission rate, and queuing method. Moreover, the transmission rate is time-varying due to the performance uncertainty of a wireless channel. Besides, the general queuing theory can achieve only an average performance for a given queue length, such as an average delay. However, given a certain probability, we need to focus on the data backlog. Fortunately, the martingale process can provide a theoretical basis for the calculation of the probability. To determine the satellite backlog (GEO or LEO), we firstly set the delay upper bound based on the data generation rate, transmission rate, and queuing method.

Theorem 1: The delay upper bound can be expressed as follows

$$p(\mathrm{W}(T) > \kappa) \leq \frac{E(M_{A^t}(0))E(M_S(0))}{H} e^{-\theta_1^*\kappa\kappa_{ss}}, \quad (12)$$

where $H := \min\{h_{A^t}(a_T)h_S(s_T) : a_T - s_T > 0\}$, $R_a = \min\{\kappa_s, R_{us}\}$, $\kappa_{ss} > R_{su}$, $R_a < \kappa_{ss}$ and θ_1^* is any

[3] refer to F section of the appendix

[4] Similar to signal and system theory in the circuit analysis, service curves can be viewed as the impulse response of a linear system. In fact, the output of a system is the input of the second system in the tandem of the many system. Therefore, the tandem of the many systems can be substituted by a single equivalent system $S(\tau, t)$ that is composed by min-plus convolution of the individual service processes.

value satisfying:

$$\theta_1^* \in \left\{ \theta > 0 \middle| \frac{R_a \ln[sp(\mathbf{M}^\theta)]}{\theta} \leq \frac{\ln[e^{-\theta\kappa_{ss}}]}{-\theta} \right\} \tag{13}$$

Proof: Please refer to B section of the appendix.

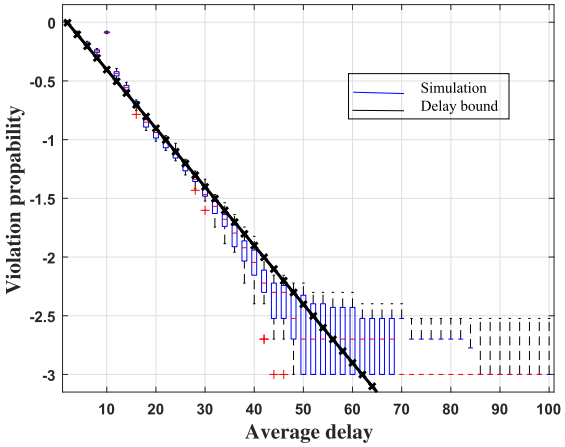

FIGURE 5. Violation probability vs Average delay Based on FIFO ($\lambda_f = 0.6$, $p_m = 0.1$ and $R_{su} = 1$).

In this section, we compare the Martingale method by simulations under asymptotic regimes. The simulation parameters were $\lambda_f = 0.6, p_m = 0.1$ and $R_{su} = 1$. In Fig.5, the y-axis is presented in logarithm format, and blue box plots show the simulation results, while black line denotes the delay bound, where the irregular tail behavior of the box plots is caused by the restriction of the simulation runs to 10^6 contents. As Fig.5 shows that the delay bound $p(\mathrm{W}(T) > \kappa)$ given by the Martingale bounds is quite tight. Since the method of Martingale-envelope can achieve a tight upper bound, we use this delay bound directly in the backlog estimation.

1) BACKLOG ESTIMATION BASED ON FIFO

The cache analysis is mainly focused on the scheduling algorithms, such as First-In First-Out (FIFO) and Static Priority (SP). In this subsection, according to the delay analysis, the cache queue violation probability based on the FIFO scheduling algorithm can be written as follows: For any $\sigma_{FIFO,L} \geq 0$, the violation probability bound $p_{FIFO,L}$ is as follows

$$\frac{E(M_{A^t}(0))E(M_S(0))}{H} e^{-\theta_1^* \sigma_{FIFO,L}}, \tag{14}$$

where H refers to section B in the appendix and θ_1^* can be any value satisfying the following condition:

$$\theta_1^* \in \left\{ \theta > 0 \middle| R_a \leq \frac{\kappa_{ss} \ln[sp(\mathrm{M}^\theta)]}{\theta} \right\} \tag{15}$$

This proof is similar to that of Theorem 1, and so it will not be explained in detail here. Then, we use the backlog as an estimation of the cache space.

$$\sigma_{FIFO,L} = \frac{1}{\theta_1^*} \ln \frac{E(M_{A^t}(0))E(M_S(0))}{H p_{FIFO,L}}, \tag{16}$$

2) BACKLOG ESTIMATE BASED ON SP

The SP scheduling policy can assign a specific priority to every service flow, where $A^h(T)$ is assumed to have a higher priority. Similar to the delay analysis, the cache queue $Q_{SP,L}$ violation probability can be defined as follows: for any $\sigma_{SP,L} \geq 0$, where $\sigma_{SP,L}$ is the backlog corresponding to $A^t(T)$ in the cache, the following violation probability bound $p^h_{SP,L}$ is given by

$$\frac{E(M_{Af}(0))E(M_{A^h}(0))E(M_S(0))}{H} e^{-\theta_2^*(\sigma_{SP,L} - \sigma^{a,h}_{SP,L})}, \tag{17}$$

where $\sigma^{a,h}_{SP,L}$ is the backlog corresponding to $A^h(T)$ in the cache and θ^* is any value satisfying the following condition:

$$\theta_2^* \in \left\{ \theta > 0 \middle| R_a \leq \frac{\kappa_{ss} \ln[sp(\mathrm{M}^\theta)]}{\theta} \right\} \tag{18}$$

This proof is similar to that of Theorem 1, and so it will not be elaborated in detail here. The backlog based on the SP can be expressed as

$$\sigma_{SP,L} = \frac{1}{\theta_2^*} \ln \frac{E(M_{Af}(0))E(M_{A^h}(0))E(M_S(0))}{H p_{SP,L}} - \sigma^{a,h}_{SP,L} \tag{19}$$

The backlog simulation was conducted using different shadowing scenarios for the satellite links, including heavy shadowing and light shadowing, which are given in Table 2 [34]. For the sake of simplicity, we let $\kappa_s > R_{us}$, and $\kappa_{ss} > R_{su}$. The other simulation parameters are given in Table 2.

TABLE 2. Shadowing values.

Shadowing	b_0	m	Ω
Heavy shadowing	0.063	0.739	$8.97 \times 10e4$
Light shadowing	0.158	19.4	1.29

Using the FIFO scheduling algorithm, we simulated the violation probability of the backlog in the LEO satellite cache. As shown in Fig.6, as the backlog increased, the probability of violation decreased. Based on the backlog given by (16), and under the assumption that the upstream channel scenario was light shadowing, and the same backlog, when the downlink channel scenario was the heavy shadowing, the probability of backlog violation was higher compared with the downlink channel scenario of light shadowing. In other words, under the heavy shadowing channel conditions, the cache was more likely to accumulate the content. In addition, using the SP scheduling algorithm, we also simulated the violation probability of the backlog in the LEO satellite cache. As presented in 7, as the backlog increased, the probability of violation decreased, as given by (16). Here, we give a comparison of two high-priority services with the arrival rate of $A^h(T) = 1$ and $A^h(T) = 0.5T$ respectively. In the comparison, the high priority service rate $A^h(T) = 1$ was fixed, the probability of violation was paralleled with that of the service without priority, and when the priority service

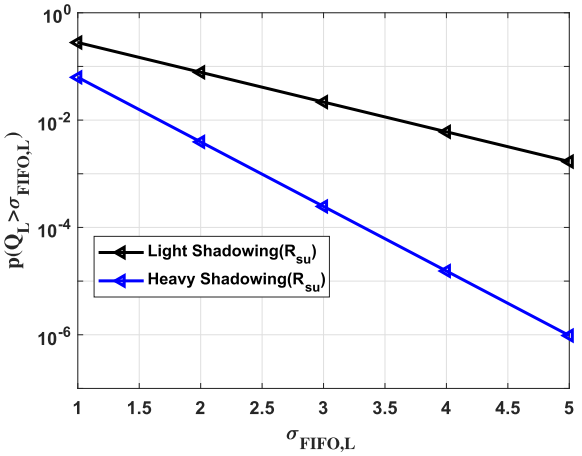

FIGURE 6. Violation probability vs Backlog Estimation Based on FIFO (The uplink and downlink channel fading models are the light and heavy shadowing respectively).

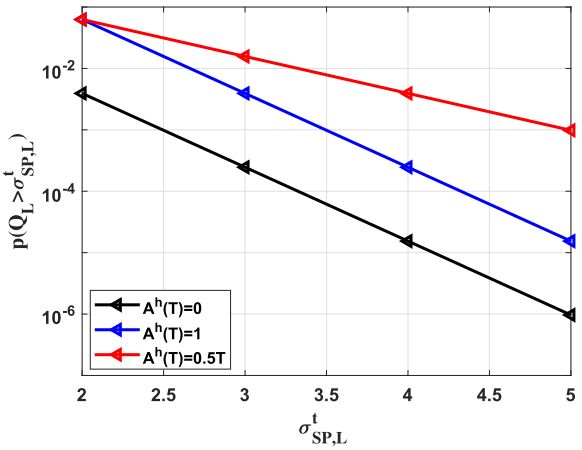

FIGURE 7. Violation probability vs Backlog Estimation Based on SP (The uplink and downlink channel fading models are the light and heavy shadowing respectively).

rate $A^h(T) = 0.5T$ was an affine function, the probability of violation deviated from that of the service without priority.

In Fig.7, the black line is the backlog in the case of a single type of service arrival, which is the same as the blue line of Fig.6, that is, the backlog of the SP and the FIFO of the single type service is the same. Given the same probability of violation, compared to the FIFO based on two different types of traffic, the SP backlog is greater. The reason for the large backlog in the case of SP is that high-priority services are often served, resulting in a large backlog of low-priority services.

B. DISTRIBUTED CACHE PRICE BARGAINING

In this section, we consider only the multi-node case, where the objective of a low layer satellite (LEO satellites) has a lower congestion probability, as low as possible. In other words, LEO satellites are expected to upload as much data as possible to the GEO satellite. However, due to the limited storage space of an GEO satellite, the LEO satellites compete for the resources in a non-cooperative fashion. At the GEO satellite side, the objective is to maximize the revenue by selling the storage space to the LEO satellites. In the following, the game process is analyzed, and the Stackelberg equilibrium of the game is investigated.

1) STACKELBERG GAME FORMULATION

The Stackelberg game is a strategy game in the economic field, where a leader and several followers compete with each other for given resources. Specifically, the leader moves first, and then followers move sequentially. In this work, the GEO satellite is treated as a leader because its storage must satisfy its requirements first, and only the surplus space, if there is any, can be rented to the LEO satellites. On the other side, the LEO satellites want to use the caching resources of the GEO satellite. Therefore, the LEO satellites are seen as followers. The GEO satellite imposes a set of prices per unit of the received data to each LEO satellite. Then, the LEO satellites (followers) update their cache allocation strategies to maximize their individual utilities. The revenue of the GEO satellite (leader) can be calculated by

$$R_{\mathrm{GEO}}(\boldsymbol{\eta}, q) = \sum_{k=1}^{K} \eta_{L_k} s_{L_k}(q_{L_k}), \tag{20}$$

where $\boldsymbol{\eta} = [\eta_{L_1}, \eta_{L_2}, \ldots, \eta_{L_K}]$ is a pricing vector, $q = [q_{L_1}, q_{L_2}, \ldots, q_{L_K}]$ is a portion vector of the accumulated storage for LEO satellites, and $s_{L_k}(q_{L_k}) = q_{L_k}\sigma_L$ is the cache occupation function of data received from a low layer node k, where σ_L is a backlog of LEO satellite. Due to the limited surplus space of the GEO satellite, the price vector $\boldsymbol{\eta}$ can be optimized by solving the following problem:

$$\begin{aligned} \textit{Problem 1:}\quad & \max\ R_{\mathrm{GEO}}(\boldsymbol{\eta}, q) \\ & \text{s.t.} \sum_{k=1}^{K} s(q_{L_k}) \le h(\sigma_G(p_G)). \end{aligned} \tag{21}$$

Where $h(\sigma_G)$ is surplus storage of GEO satellite. Next, we design a function that reflects the congestion level of the cache space and the non cooperative nature of the nodes.

Definition 3.1: Congestion Index (CI) is defined by

$$I(q_{L_k}) = \ln\left(1 + \frac{q_{L_k}\sigma_{L_k}}{\sum_{j\neq k}^{K} q_{L_j}\sigma_{L_j}}\right), \quad \sum_{j\neq k}^{K} q_{L_j}\sigma_{L_j} \neq 0$$
$$\forall k,\ j,\ \ q_{L_k} > 0, q_{L_j} > 0 k,\ j \in \{1, 2, \ldots, K\} \tag{22}$$

The above-given definition is rational. In the Shannon formula, the power of other users to the user interference is used as a denominator. Here, each LEO satellite is expected to upload the maximized portion of its traffic flow to the GEO satellite. However, due to limited surplus space of the GEO satellite, the LEO satellites compete in a non-cooperative fashion. Therefore, we provide a similar definition.

Besides, the ratio q_{L_k} can reflect the nature of the congestion degree. Also, a logarithmic function will not change the relation of a variable, so we will use the logarithm as a congestion function.

Then, the profit of the low layer nodes can be defined by

$$\mathrm{R_{LEO}} = cf(\sigma_{L_k})I(q_{L_k}) - \eta_{L_k} s(q_{L_k}), \tag{23}$$

where c is the utility gain per unit congestion index, and $f(\sigma_{L_k})$ represents the probability that the GEO satellite prefers to accept data from L_k, where $f(\sigma_{L_k}) \propto \sigma_{L_k}$. The assumption is rational because the GEO satellite prefers to accept data that will change its storage amount the most. Thus, it wants to receive as much data as possible from the nodes participating in the congestion.

The profit of a low layer node (an LEO satellite) consists of two parts, revenue and cost. If a low layer node increases the amount of data redirected to the GEO satellite, the congestion index $I(q_{L_k})$ increases, and also the revenue. Meanwhile, with the increase in data amount, the data will consume more space on the GEO satellite. Thus, LEO satellites have to buy larger cache from the GEO satellite, which will enhance the cost. Therefore, to achieve the maximal profit of an LEO satellite, the optimal cache allocation scheme of the GEO satellite is given by solving the following problem:

$$\textit{Problem 2:} \quad \max \ R_{\mathrm{LEO}}(q_{L_k}, -q_{L_k}, \eta_{L_k}), \tag{24}$$

Then, Problem 1 and 2 formulate a Stackelberg game.

2) STACKELBERG EQUILIBRIUM (SE)

For the proposed Stackelberg game, the SE is defined as follows.

Definition 3.1: Let $\boldsymbol{\eta}^*$ and $q^*_{L_k}$ be the solutions for Problem 1 and Problem 2, respectively. Then, the point $(\boldsymbol{\eta}^*, \boldsymbol{q}^*)$ is the SE for the proposed Stackelberg game if the following conditions are satisfied:

$$R_{\mathrm{GEO}}(\boldsymbol{\eta}^*, \boldsymbol{q}^*) \geq (\boldsymbol{\eta}, \boldsymbol{q}^*), \tag{25}$$

$$R_{\mathrm{LEO}}(q^*_{L_k}, -\boldsymbol{q}^*, \boldsymbol{\eta}^*) \geq (q_{L_k}, -\boldsymbol{q}^*, \boldsymbol{\eta}^*). \tag{26}$$

where $\boldsymbol{\eta} \succeq 0$ and $\boldsymbol{q} \succeq 0$. Generally, the SE for the Stackelberg game can be obtained through finding a perfect Nash Equilibrium (NE) for its subgame. In our proposed game, LEO satellites compete in a strictly non-cooperative fashion. Thus, a non-cooperative subgame for the cache scheduling can be formulated at the low layer node (LEO satellites) side. For a non-cooperative game, the NE is defined as an operating point, at which any player in the game can not improve utility by changing its strategy unilaterally.

Proposition 1: The SE for the Stackelberg game formulated by Problems 1 and 2 is expressed as $(\boldsymbol{\eta}^*, \boldsymbol{q}^*)$, where $\boldsymbol{\eta}^*$is given by (25), and $\boldsymbol{q}^*$ is given by (26). In a centralized manner, the proposed game can be implemented as follows.

3) GAME PROBLEM OPTIMIZATION

In the previous section, we analyze the backlog bound, and give a backlog violation probability p_L or p_G. In this section, it is proven that the objective function of Problem 2 is a concave function of q_{L_k} and the constraint is affine. Therefore, Problem 2 is a convex problem. Since the duality gap between the problem and its dual optimization problem is zero, the problem is solvable.

Non-uniform Pricing Scheme: In this part, we solve the optimization Problem 2 in our game. The Lagrangian can be written as

$$F(\eta_{L_k}, \alpha_{L_k}) = cf(\sigma_{L_k}) \ln \left(1 + \frac{q_{L_k}\sigma_{L_k}}{\sum\limits_{j\neq k}^{K} q_{L_j}\sigma_{L_j}} \right) - \sigma_{L_k}\eta_{L_k}q_{L_k} - \alpha_k q_{L_k}. \tag{27}$$

Lemma 2: For a given η_k, the optimal solution of Problem 2 is

$$q^*_{L_k} = \left(\frac{cf(\sigma_{L_k})}{\sigma_{L_k}\eta_{L_k}} - y_{L_k} \right)^+, \quad \forall k, \tag{28}$$

where $y_{L_k} = \sum\limits_{j\neq k}^{K} q_{L_j}\sigma_{L_j}$.

Proof: Please refer to section C in the appendix.

Since $q^*_{L_k}$ is an optimal portion in a low layer satellite, $q^*_{L_k}$ is larger than zero. When $\eta_{L_k} \geq \frac{cf(\sigma_{L_k})}{\sigma_{L_k} y_{L_k}}$, the data will not be uploaded. This indicates that node L_k will be removed from the game.

Substituting the optimal $q^*_{L_k}$ into (21), we can rewrite Problem 1 as

$$\textit{Problem 3:} \quad \max \sum_{k=1}^{K} \left(cf(\sigma_{L_k}) - \sigma_{L_k}\eta_{L_k}y_{L_k} \right)^+ \tag{29}$$

$$\text{s.t.} \sum_{k=1}^{K} \left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}} - \sigma_{L_k}y_{L_k} \right)^+ \leq h(\sigma_G). \tag{30}$$

Note that the above problem is a non-convex problem. Then, we assume that all nodes participate in the game so that Problem 3 will be transferred into a convex problem without the limit of the indicator. Then we can simplify Problem 3

$$\textit{Problem 4:} \quad \max \sum_{k=1}^{K} \left(cf(\sigma_{L_k}) - \sigma_{L_k}\eta_{L_k}y_{L_k} \right) \tag{31}$$

$$\text{s.t.} \sum_{k=1}^{K} \left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}} - \sigma_{L_k}y_{L_k} \right) \leq h(\sigma_G). \tag{32}$$

Theorem 2: The optimal solution for Problem 4 can be expressed as $\boldsymbol{\eta}^* = [\eta^*_{L_1}, \eta^*_{L_2}, ..., \eta^*_{L_K}]$, where

$$\eta^*_{L_k} = \sqrt{\frac{f(\sigma_{L_k})}{y_{L_k}}} \frac{c\sum\limits_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}}{\left(h(\sigma_G) + \sum\limits_{k=1}^{K}\sigma_{L_k}y_{L_k} \right)}, \quad \forall k. \tag{33}$$

Proof: Please refer to section D in the appendix.

In the following, we discuss the constraints on $h(\sigma_G)$. For $\eta_{L_k} \geq \frac{cf(\sigma_{L_k})}{\sigma_{L_k} y_{L_k}}$, node L_k will not upload the data.

Theorem 3: We can obtain the minimum value of $h(\sigma_G)$ with K nodes participating the game. Mathematically, it is expressed by

$$h(\sigma_G)^{min} = \sqrt{\frac{y_{L_k}}{\sigma_{L_k} f(\sigma_{L_k})}} \sum_{k=1}^{K} \sqrt{f(\sigma_{L_k}) y_{L_k}} - \sum_{k=1}^{K} \sigma_{L_k} y_{L_k}, \tag{34}$$

Proof: Please refer to section E in the appendix.

Theorem 4: For $v_1 > v_2$ and $x_{v_1} > x_{v_2}$,

$$x_{L_K} = \sqrt{\frac{y_{L_K}}{\sigma_{L_K} f(\sigma_{L_K})}} \sum_{k=1}^{K} \sqrt{f(\sigma_{L_k}) y_{L_k}} - \sum_{k=1}^{K} \sigma_{L_k} y_{L_k}. \tag{35}$$

The above function is a strictly monotonically increasing function with respect to K.

Proof: Please refer to section F in the appendix.

According to above-presented results, the optimal solution of Problem 2 is given by the following theorem.

Theorem 5: Assuming that all preferences of LEO satellites are sorted in the following order $\sqrt{\frac{f(\sigma_{L_1})}{y_{L_1}}} > \sqrt{\frac{f(\sigma_{L_2})}{y_{L_2}}} > \ldots > \sqrt{\frac{f(\sigma_{L_K})}{y_{L_K}}}$, the optimal solution of Problem 3 is as follows: where $Z_{L_K} = \frac{c \sum_{k=1}^{K} \sqrt{f(\sigma_{L_K}) y_{L_k}}}{(h(\sigma_G) + \sum_{k=1}^{K} \sigma_{L_k} y_{L_k})}$.

Proof: If $h(\sigma_G) > x_{L_K}$, we can obtain the optimal η^* with Theorem 2. For $\forall k$, if $x_{L_{K-1}} < h(\sigma_G) < x_{L_K}$, we can obtain the corresponding η^*. □

Actually, the results given by (36), as shown at the bottom of this page, are reasonable because the less the surplus cache resources of the GEO satellite is, the fewer nodes can participate in the game. Then, the Stackelberg game for non-uniform pricing is completely solved. The SE of the Stackelberg game is then given as follows.

First, the GEO satellite initializes the cache price $\boldsymbol{\eta}$. Next, the GEO satellite collects information such as $f_{L_k}(\sigma_L)$ and each LEO satellite computes its optimal fractional $q^*_{L_k}$ of space occupation $q^*_{L_k}$ of the LEO satellite based on the received $\boldsymbol{\eta}$ which is defined by (36). Then, this information is uploaded through the backhaul link. Further, the GEO satellite computes the threshold vector $\boldsymbol{x} = [x_{L_k}, x_{L_{k-1}}, \ldots, x_{L_1}]$ by Theorem 4. Then, for a given surplus space $h(\sigma_G)$ of the GEO satellite, the number of LEO satellites participating in the game is computed by the Members of the Game algorithm described in Algorithm 1. After that, according to the obtained threshold vector $\boldsymbol{x}$, the GEO satellite decides the fractional capacity price for each LEO satellite based on its surplus space $h(\sigma_G)$. Then, the prices are fed back to the LEO satellites through the backhaul links between the GEO and LEO satellites. Finally, the LEO satellites decide the amount of data which will be uploaded to the GEO satellite according to (51). This process is described in detail in Algorithm 2.

Algorithm 1 Members of a Game

1: **input**:
surplus space $h(\sigma_G)$ based on violation probability p_G;
2: The GEO satellite initializes the cache price $\boldsymbol{\eta}$, and broadcasts surplus space $h(\sigma_G)$ based on violation probability p_G to all LEO satellites.
3: Count $\boldsymbol{f}(\boldsymbol{\sigma_L})$ based violation probability p_L.
$(\boldsymbol{f}(\boldsymbol{\sigma_L}) = [f_1(\sigma_L), f_2(\sigma_L), \ldots, f_K(\sigma_L)])$, collect $\boldsymbol{q}$ $(\boldsymbol{q} = [q_{L_1}, q_{L_2}, \ldots, q_{L_U}])$.
4: Set $U = K$, sort the U users according to $\sqrt{\frac{f_{L_1}(\sigma_L)}{y_{L_1}}} > \sqrt{\frac{f_{L_2}(\sigma_L)}{y_{L_2}}} > \ldots > \sqrt{\frac{f_{L_U}(\sigma_L)}{y_{L_U}}}$.
5: Calculate x_{L_U}, where

$$x_{L_U} = \left(\sqrt{\frac{y_{L_U}}{\sigma_L f_{L_U}(\sigma_L)}} \sum_{k=1}^{U} \sqrt{f_{L_U}(\sigma_L) y_{L_k}} - \sum_{k=1}^{U} y_{L_k} \right)$$

6: Comparing the x_{L_U} with $\sqrt{\frac{f_{L_U}(\sigma_L)}{y_{L_U}}}$, if $x_{L_U} < h(\sigma_G)$ remove user N from the game, set $U = U - 1$, and go to step 3. Otherwise, go to step 6.
7: With Z_{L_U} and L, the cache price $f_{L_k}(\sigma_L)$ for the LEO satellite node k is given by $\eta^* = \begin{cases} Z_{L_k} \sqrt{\frac{y_{L_k}}{f_{L_k}(\sigma_L)}}, \textit{if}, k \leq K \\ \infty, \textit{otherwise} \end{cases}$
8: **output**: U users (Members of a Game);

$$\eta^* = \begin{cases} Z_{L_K} \left[\sqrt{\frac{f(\sigma_{L_1})}{y_{L_1}}}, \sqrt{\frac{f(\sigma_{L_2})}{y_{L_2}}}, \ldots, \sqrt{\frac{f(\sigma_{L_K})}{y_{L_K}}} \right], & h(\sigma_G) > x_{L_K} \\ Z_{L_{K-1}} \left[\sqrt{\frac{f(\sigma_{L_1})}{y_{L_1}}}, \sqrt{\frac{f(\sigma_{L_2})}{y_{L_2}}}, \ldots, \sqrt{\frac{f(\sigma_{L_K})}{y_{L_{K-1}}}} \right], & x_{L_{K-1}} < h(\sigma_G) < x_{L_K} \\ \vdots & \\ Z_{L_2} \left[\sqrt{\frac{f(\sigma_{L_1})}{y_{L_1}}}, \sqrt{\frac{f(\sigma_{L_2})}{y_{L_2}}} \right], & x_{L_2} < h(\sigma_G) < x_{L_3}, \end{cases} \tag{36}$$

Algorithm 2 Distributed Cache Price Bargaining

1: **input**: U users (Members of a Game);
2: The GEO satellite initializes the cache price $\boldsymbol{\eta}$, and broadcasts $\boldsymbol{\eta}$ to all LEO satellites.
3: Each LEO satellite computes its optimal amount of data $q^*_{L_k}$ based on the received $\boldsymbol{\eta}$ by (36), and uploads with $q^*_{L_k}$.
4: The GEO satellite calculates the total received data $\sum_{k=1}^{K} q_{L_k}$, and updates the price $\boldsymbol{\eta}$ based on $\sum_{k=1}^{K} q_{L_k}$. Assume that ζ is a small positive constant that controls the algorithm accuracy. Then,
5: **if** $\sum_{k=1}^{K} q_{L_k} > h(\sigma_G) + \zeta$ **then**
6: The GEO satellite increases the cache price by $\Delta\eta$;
7: **else**
8: The GEO decreases the cache price by $\Delta\eta$, where $\Delta\eta > 0$ is a small step size.
9: **end if**
10: After that, the GEO satellite broadcasts the new cache price to all LEO satellites.
11: Step 2 and Step 3 are repeated until $|\sum_{k=1}^{K} q_{L_k} - h(\sigma_G)| < \zeta$.
12: **output**: $(\boldsymbol{\eta}^*, \boldsymbol{q}^*)$;

Uniform Pricing Scheme: In this part, we consider a uniform scheme. The GEO satellite sets a uniform price for all the LEO satellites. In this case, similar to Lemma 2, the optimal solution of the problem is given by

$$q^*_{L_k} = \left(\frac{cf(\sigma_{L_k})}{\sigma_{L_k}\eta} - y_{L_k}\right)^+, \quad \forall k, \tag{37}$$

Then, the optimization problem can be simplified to

$$\textit{Problem 5:} \quad \max \sum_{k=1}^{K} \left(cf(\sigma_{L_k}) - \sigma_{L_k}\eta y_{L_k}\right)^+ \tag{38}$$

$$\text{s.t.} \sum_{k=1}^{K} \left(\frac{cf(\sigma_{L_k})}{\eta} - \sigma_{L_k} y_{L_k}\right)^+ \leq h(\sigma_G). \tag{39}$$

Similar to problem 3, we obtain the price as follows:

Lemma 3: The optimal solution of Problem 5 can be expressed as $\boldsymbol{\eta}^* = [\eta^*_{L_1}, \eta^*_{L_2}, ..., \eta^*_{L_K}]$, where

$$\eta^*_{L_k} = \frac{c\sum_{j=1}^{k} f(\sigma_{L_j})}{\sum_{j=1}^{k} \sigma_{L_j} y_{L_j} + h(\sigma_G)}, \quad \forall j. \tag{40}$$

Proof: The proof is similar to theorem Theorem 2.

Theorem 6: The optimal solution to Problem 5 is given as follows:

$$\eta^* = \begin{cases} \tilde{\eta}_{L_K}, & if,\ h(\sigma_G) > \tilde{x}_{L_K} \\ \tilde{\eta}_{L_{K-1}}, & if,\ \tilde{x}_{L_{K-1}} < h(\sigma_G) < \tilde{x}_{L_K} \\ \vdots & \\ \tilde{\eta}_{L_1}, & if,\ \tilde{x}_{L_1} < h(\sigma_G) < \tilde{x}_{L_2} \end{cases} \tag{41}$$

where, $\tilde{\eta}_{L_k} = \frac{c\sum_{k=1}^{K} f(\sigma_{L_K})}{\sum_{k=1}^{K} \sigma_{L_k} y_{L_k} + h(\sigma_G)}$ and $\tilde{x}_{L_K} = \left(\frac{\sum_{k=1}^{K} y_{L_k}\sigma_{L_K} - \sum_{k=1}^{K} y_{L_k}}{\sum_{k=1}^{K} y_{L_k}\sigma_{L_k}}\right)$.

A detailed proof is omitted because it is similar to that of Theorem 5.

For a given $h(\sigma_G)$, it is easy to note that the optimal price is unique. Thus, we can obtain SE for the Stackelberg game. It is clear that the uniform pricing scheme is inferior to the non-uniform pricing scheme in terms of maximizing the revenue of the GEO satellite.

C. CONTENT UPLOADING BASED ON POPULARITY

To utilize the storage space of the GEO satellite effectively, we sort the popularity of the LEO satellites contents using the Popularity Matching Algorithm and upload unpopular content to the GEO satellite. Note that in this algorithm the caching cost is inversely proportional to the content popularity, so unpopular content will be uploaded to the GEO satellite. Specifically, in the first stage, the GEO satellite broadcasts the information on its surplus storage space $h(\sigma_G)$ and the standard caching size C_b to all the LEO satellites. In the second stage, the LEO satellites strictly compete in a non-cooperative fashion. Besides, Algorithm 2 introduces the Distributed Cache Price Bargaining Algorithm, and Algorithm 1 introduces the Members of the Game Algorithm. Based on the optimal result, content can be sorted regarding t_f based on the popularity ($(t_1 < t_2 < ... < t_F)$), where t_f is the fth content of the LEO satellite, and these contents are matched to C_b according to the popularity in ascending order. In the third stage, the LEO satellite is connected to the GEO satellite so that its contents can be upload to the GEO satellite. In order to show the relationship between the three algorithms presented in this paper more clearly, we present the flow chart in Fig. 8. In the initialization phase, first an error calculation accuracy value ζ is given, and then Algorithm 1 is implemented. If $|\sum_{k=1}^{K} q_{L_k} - h(\sigma_G)| < \zeta$ is satisfied, then Algorithm 3 is implemented; otherwise, Algorithm 2 is implemented.

Algorithm 1 gives the number of members participating in the game. The algorithm is correct and the effectiveness depends on the accuracy of the threshold x_{L_K} which is analyzed in the second subsection. The other is Distributed Cache Price Bargaining, where Algorithm 2 is correct because of the existence of the optimization results, and the effectiveness of which depends on the accuracy of the pricing error ζ.

V. NUMERICAL RESULTS AND THE ANALYSIS

In this section, several numerical simulations are provided to validate the performance of the load balancing strategies based on content popularity. In the simulations, we used the STK (Satellite Tool Kit) tool to generate an Iridium-like constellation, as shown in Fig.10. This constellation contained 66 LEO satellites evenly distributed over six orbital planes. At the same time, we added three GEO satellites to this scene. In the STK coverage analysis, six LEO satellites were

Algorithm 3 Popularity Matching Algorithm

1: **input**: q^*;
2: The GEO satellite broadcasts the standard caching size C_b and surplus storage space $h(\sigma_G)$ based on violation probability $p_G$to LEO satellites.
3: A congestion LEO satellite sort in t_f with popularity $(t_1 < t_2 < ... < t_F)$.
4: Transforming contents into equal-sized C_b. That is, let contents with popularity $(t_1 < t_2 < ... < t_F)$ match caching size C_b.
5: The LEO satellite is connected to the GEO satellite for uploading these matching content.
6: **output**: $(t_1 < t_2 < ... < t_f)$;

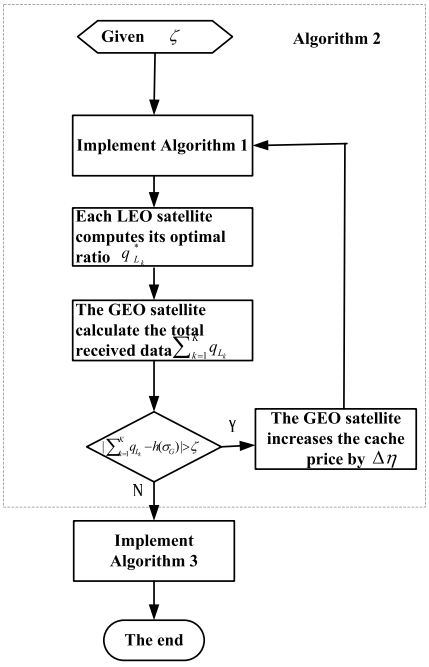

FIGURE 8. Relationship flow chart between the three algorithms.

covered by a GEO satellite (satellite111) for a long time. Next, we analyzed the simulation results based on these seven satellites, as shown in Fig.10.

Fig.11 shows that the price difference decreased with the increase in $h(\sigma_G)$ which was consistent with (36); namely, as given by (36), the factor η_k decreases with the increase in $h(\sigma_G)$. Fig.11, it can be observed that, at a given $h(\sigma_G)$, the price of node L_3 was the lowest, while that of node L_1 was the highest. Therefore, we assumed that $f_{L_1}(\sigma_L) > f_{L_2}(\sigma_L) > f_{L_3}(\sigma_L)$, where $f_{L_1}(\sigma_L)$ was a larger preference for L_1.

We set the parameter s of Zipf to 0.1, 0.8 just to better reflect the influence of parameters on the number of members participating in the game. As we can see from Fig.5, the larger the s parameter, the steeper the distribution curve, and the better the discrimination of the popularity of the content. Therefore, we set the parameter s to two numbers

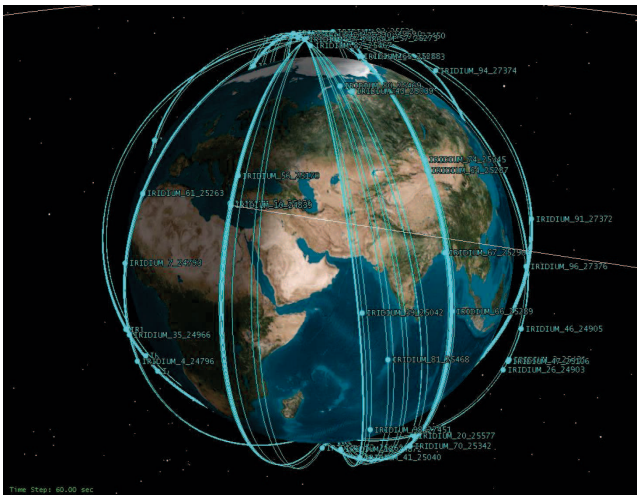

FIGURE 9. 3D graphic based on STK (An Iridium-like LEO satellite constellation and three GEO satellites).

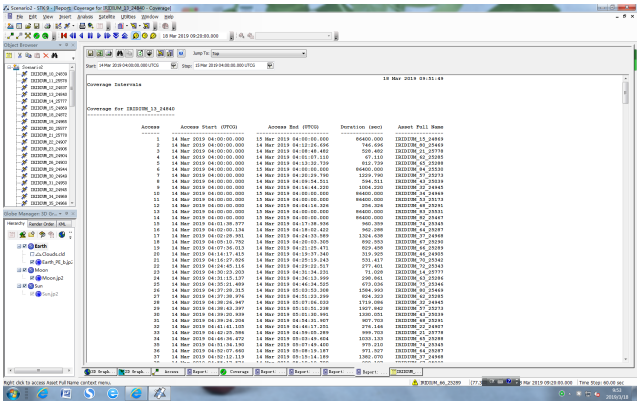

FIGURE 10. Coverage for satellite111 report.

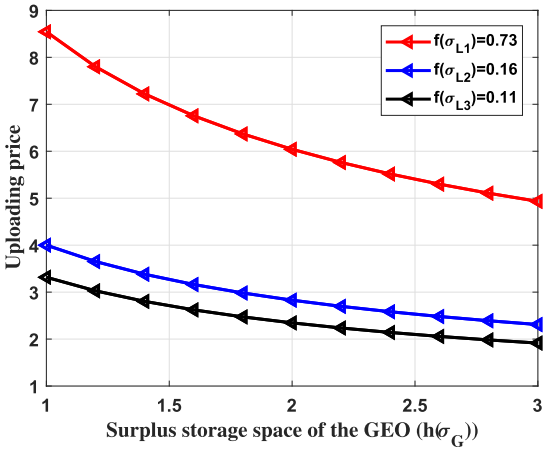

FIGURE 11. Uploading price vs. the surplus storage space $h(\sigma_G)$ of the GEO satellite.

with a large span. Fig.12 shows the number of participants in the game versus $h(\sigma_G)$ at two value for s, $s = 0.1$ and $s = 0.8$. In Fig.12, it is shown that more LEO satellites were able to

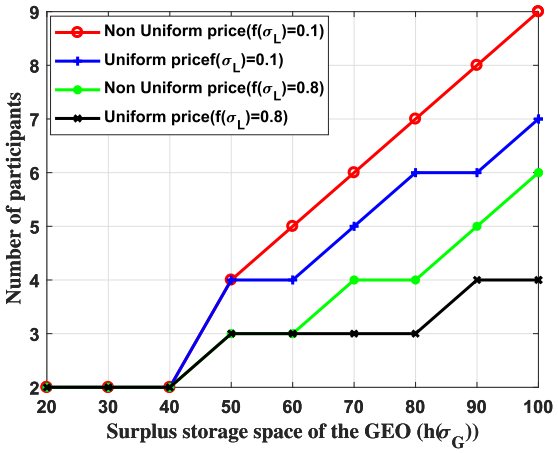

FIGURE 12. Number of participants vs. the surplus storage space $h(\sigma_G)$ of the GEO satellite ($s = 0.1, 0.8$).

participate in the game at a larger $h(\sigma_G)$. Again, the non-uniform price outperformed the uniform price because of the capability to accommodate more LEO satellites for a given $h(\sigma_G)$. By comparing the scenarios at $f(\sigma_L) = 0.1$ and $f(\sigma_L) = 0.8$, we found that for a given $h(\sigma_G)$, at $f(\sigma_L) = 0.1$ more LEO satellites were accommodated in the game than at $f(\sigma_L) = 0.8$. Therefore, we assumed that distribution of the backlog uniformity was the Zipf distribution. At $f(\sigma_L) = 0.1$, the distribution was nearly uniform, while at $f(\sigma_L) = 0.8$, the nodes making the congestion more easily were dominant in the network.

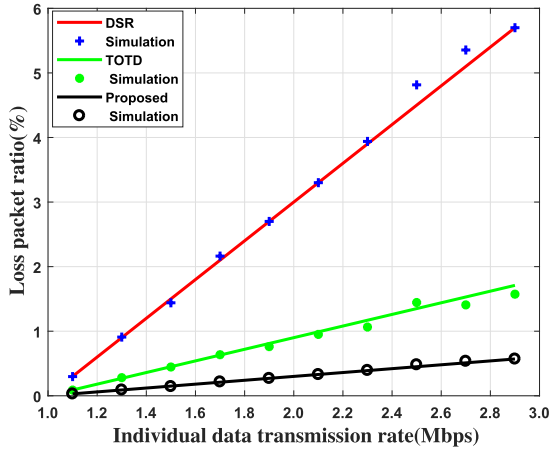

FIGURE 13. Loss packet ratio vs Individual data transmission rate.

With the aim to profoundly analyze the load balancing performance of the proposed method, we compared it with the shortest path routing protocol (DSR) and the TOTD (we referred to the method proposed in [1] as TOTD) balance mechanism regarding the packet loss rate. As shown in Fig.13, the DSR had the highest packet loss rate, followed by the TOTD, and the proposed method had the lowest packet loss rate. The reason for such results was that DSR caused the node traffic to be unbalanced. This phenomenon is particularly prominent in the lattice topology [41] because congestion node cannot upload traffic to another node under the DSR mode. In Fig.13, the results of the TOTD method which considered uploading traffic to the upper nodes are presented. However, the proportion of the upload depended mainly on the optimal threshold which largely depended on channel conditions. However, the TOTD considered only the propagation delay influence. As presented in Fig.13, the proposed method significantly reduced the packet loss rate, and the method based on the Stackelberg game promoted the load balance of low layer satellite (LEO satellite) services under the premise of ensuring the use of high layer satellite (GEO satellite) cache.

VI. CONCLUSION

In this paper, we focus on the problem of load balancing in multi-layers satellite network. In addition, to resolve the competition of LEO satellites in a non-cooperative fashion, we first estimate the backlog of the LEO and GEO satellites to identify the congestion state of the LEO satellites and the surplus storage space of the GEO satellite. Based on the backlog estimation, a non-cooperative cache resource allocation problem is formed because the low layer node (the LEO satellite) is subject to the maximum tolerable surplus storage space constraint at the GEO satellite. We formulate the Stackelberg game model, which contains a new profit model of LEO satellites, to maximize the profit of both GEO and LEO satellites jointly. A non-convex problem is converted to a convex optimization problem by using the threshold function, and the Stackelberg equilibrium is achieved. The threshold function is strictly monotonically increasing with the number of participants in the game. Finally, several numerical simulations are conducted to evaluate the performance of the proposed method regarding the number of participant in the game depending on the storage space of the GEO satellite.

APPENDIX A
PROOF OF LEMMA1

The arrival process $A^t(T)$ is a supermartingale, where $g(Y_{T\tau})$ can be written as

$$g(Y_T(\tau)) = \begin{cases} I(T)R_{us}, & good \\ 0, & bad \end{cases} \tag{42}$$

Proof: Before the proceeding, the transition matrix $\mathbf{M}$ for x_N is defined by

$$\mathbf{M}_{i,j} = P(Y_{T+1} = j | Y_T = i). \tag{43}$$

Correspondingly, the exponential transformation along the columns of the matrix $\mathbf{M}_{i,j}$ [18] can be expressed as

$$\mathbf{M}^{\theta}_{i,j} = \mathbf{M}_{i,j} e^{\theta g(Y_j)}. \tag{44}$$

The channel service process can be expressed as

$$\begin{aligned} & E\left[e^{\theta}\left(A^t(T+1) - g(Y(T+1))\right)\middle| Y_1, Y_2, \cdots, Y_T\right] \\ & \leq e^{\theta\left(A^t(T) - R_{us}T\right)} E\left[e^{\theta Y_{T+1}}\middle| Y_T\right] \\ & \leq e^{\theta\left(A^t(T) - R_{us}T\right)} e^{-\theta p_m (1-p_m)^N R_{us}} sp\left(\mathbf{M}^{\theta}\right) \\ & \leq e^{\theta\left(A^t(T) - R_{us}T\right)}, \end{aligned} \tag{45}$$

where $sp(\mathbf{M}^{\theta})$ is $\mathbf{M}^{\theta}_{i,j}$ which is the spectral radius, and the inequality is workable in $e^{-\theta R_{us} p_m (1-p_m)^N} sp\left(\mathbf{M}^{\theta}\right) \geq 1$ case.

In this work, we assume the high-priority arrival $A^h(T)$, and low-priority arrives at $A^f(T)$. Then, the martingale envelope can be defined by

$$M_{A^t}(T) = h_{A^h}(a_T) h_{A^f}(a^f_T) h_{s^q}(s_T) \times e^{\theta(A^h(\kappa,T)-(T-\kappa)\kappa_f + A^f(T) - Tk_h + T\kappa_s - S^q(T))},$$

where $h_{A^h}(a_T)$, $h_{A^f}(a^f_T)$, $h_{s^q}(s_T)$ are monotonically increasing functions, and κ_h and κ_f are the high-priority arrival and low-priority arrival rates of the LEO/GEO satellite, respectively, $S^q(T)$ is the queue service process, $\theta^* > 0$ and κ_s depends on the queue service rate. □

APPENDIX B
PROOF THEOREM1

According to the definition of the martingale envelope given by (63), we consider the arrival supermartingale envelope $M_{A^t}(T)$ and service supermartingale envelope $M_s(T)$ are respectively expressed by

$$M_{A^t}(T) = h_{A^t}(a^f_T) e^{\theta^*(\{(A^t(T,\kappa)-(T-\kappa)R_a)\})}$$
$$M_S(T) = h_S(s_T) e^{\theta^*(\{(S(T)-n\kappa_{ss})\})}$$

In this work, we assume that random processes of arrivals and services are independent. Thus, a process $M(T)$ is also a supermartingale in the time domain $V := \kappa, \kappa+1, \ldots$, which is given by

$$M(T) = h_{A^t}(a^t_T) h_S(s_T) e^{\theta^*(\{(A^f(T,\kappa)-(T-\kappa)R_a)\})} . e^{\theta^*(\{(S(T)-T\kappa_{ss})\})} \quad (46)$$

Therefore, by the optional stopping theorem applied to the stopping time, the process can be written as

$$\begin{aligned} E[M(0)] &= E[M(\kappa \wedge T)] \\ &\geq E[M(\kappa \wedge T) | \{\kappa \leq T\}] \\ &\geq E[h_{A^t}(a^t_T) h_S(s_T)] \\ &\quad E[e^{\theta^*((A^t(T,\kappa)-(T-\kappa)R_a - S(T)+T\kappa_{ss}))}] p(\kappa \leq T) \\ &\geq H e^{\theta^* \kappa \kappa_{ss}} p(\kappa \leq T), \end{aligned} \quad (47)$$

where, $H := \min\{h_{A^t}(a_T) h_S(s_T) : a_T - s_T > 0\}$

$$\begin{aligned} p(\mathrm{W}(T) > \kappa) &= p(A^t(0, T-k) \geq D(T)) \\ &\leq p(A^t(T,\kappa) \geq \min_{0 \leq m \leq T} \{A^t(m) + S(m,T)\}) \\ &\leq p(A^t(T,\kappa) - \min_{0 \leq m \leq T} \{A^t(m) + S(m,T)\} \geq 0) \\ &\leq p(\max_{T \geq \kappa} \{A^t(T,\kappa) - (T-\kappa)\kappa_{ss} \\ &\quad + TR_a - S(T)\} \geq \kappa \kappa_{ss}) \\ &\leq \frac{E(M_{A^t}(0)) E(M_S(0))}{H} e^{-\theta^* \kappa \kappa_{ss}} \end{aligned}$$

APPENDIX C
PROOF LEMMA2

The optimal portion $q^*_{L_k}$ for node k can be achieved by deriving $F(\boldsymbol{\eta}, \boldsymbol{\alpha})$ with respect to q_{L_k} and solving the Karush-Kuhn-Tucher (KKT) conditions given by: $\frac{\partial(F(\boldsymbol{\eta},\boldsymbol{\alpha}))}{\partial q_{L_k}} = 0$, $\alpha_{L_k} \geq 0, q_{L_k} > 0$.

By applying the Lagrangian multipliers to the objective function, we get

$$F(\eta_{L_k}, \alpha_{L_k}) = cf(\sigma_{L_k}) \ln \left(1 + \frac{q_{L_k}\sigma_{L_k}}{\sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}}\right) - \sigma_{L_k}\eta_{L_k} q_{L_k} - \alpha_k q_{L_k}. \quad (48)$$

To solve problem3.2, let $\frac{\partial(F(\boldsymbol{\eta},\boldsymbol{\alpha}))}{\partial q_{L_k}} = 0$,

$$\frac{\frac{c\sigma_{L_k} f(\sigma_{L_k})}{\sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}}}{1 + \frac{q_{L_k}\sigma_{L_k}}{\sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}}} - \sigma_{L_k}\eta_{L_k} - \alpha_k = 0 \quad (49)$$

Since $\alpha_k q_{L_k} = 0, q_{L_k} > 0$, therefore, $\alpha_k = 0$. Setting α_k equal to be 0 yields to

$$\frac{\frac{c\sigma_{L_k} f(\sigma_{L_k})}{\sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}}}{1 + \frac{q_{L_k}\sigma_{L_k}}{\sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}}} - \sigma_{L_k}\eta_{L_k} = 0 \quad (50)$$

Then it follows that

$$q^*_{L_k} = \left(\frac{cf(\sigma_{L_k})}{\sigma_{L_k}\eta_{L_k}} - y_{L_k}\right)^+, \quad \forall k,$$

where $y_{L_k} = \sum_{j \neq k}^{K} q_{L_j}\sigma_{L_j}$.

APPENDIX D
PROOF THEOREM2

By applying the Lagrangian multipliers to the objective function, we get

$$\begin{aligned} L(\boldsymbol{\eta}, \alpha, \boldsymbol{\gamma}) &= \sum_{k=1}^{K} \left(\eta_{L_k} y_{L_k} \sigma_{L_k} - cf(\sigma_{L_k})\right) \\ &\quad + \alpha \left[\sum_{k=1}^{K} \left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}} - \sigma_{L_k} y_{L_k}\right) - h(\sigma_G)\right] \\ &\quad - \sum_{k=1}^{K} \gamma_{L_k} \eta_{L_k}. \end{aligned} \quad (51)$$

Then, the KKT conditions can be written as follows:

$$\frac{\partial L(\boldsymbol{\eta}, \alpha, \boldsymbol{\gamma})}{\partial \eta_{L_k}} = 0 \quad \forall k, \quad (52)$$

$$\alpha\left[\sum_{k=1}^{K}\left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}}-\sigma_{L_k}y_{L_k}\right)-h(\sigma_G)\right]=0, \tag{53}$$

$$\alpha\geq 0,\gamma_{L_k}\geq 0,\eta_{L_k}>0,\gamma_{L_k}\eta_{L_k}=0,\quad \forall k, \tag{54}$$

$$\sum_{k=1}^{K}\left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}}-\sigma_{L_k}y_{L_k}\right)-h(\sigma_G)\leq 0. \tag{55}$$

Following (52), we have $\frac{\partial L(\boldsymbol{\eta},\alpha,\boldsymbol{\gamma})}{\partial\eta_k}=\sigma_{L_k}y_{L_k}-\frac{c\alpha f(\sigma_{L_k})}{\eta_{L_k}^2}-\gamma_{L_k}$. Since $\gamma_{L_k}\eta_{L_k}=0,\eta_{L_k}>0$, therefore, $\gamma_{L_k}=0$. Setting (52) equal to be 0 yields to

$$\eta_{L_k}^{*}=\sqrt{\frac{c\alpha f(\sigma_{L_k})}{\sigma_{L_k}y_{L_k}}}. \tag{56}$$

Supposing that $\sum_{k=1}^{K}\left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}}-\sigma_{L_k}y_{L_k}\right)-h(\sigma_G)\neq 0$, based on (53), we have $\alpha=0$. Then, $\eta_{L_k}=0,\forall k$ according to (56), which contradicts to $\eta_{L_k}>0,\forall k$. Hence the aforementioned preassumption does not hold. Thus, we have $\sum_{k=1}^{K}\left(\frac{cf(\sigma_{L_k})}{\eta_{L_k}}-\sigma_{L_k}y_{L_k}\right)-h(\sigma_G)=0$. By substituting (56) into (53), we get

$$\sqrt{\alpha}=\frac{\sqrt{c}\sum_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}}{h(\sigma_G)+\sum_{k=1}^{K}\sigma_{L_k}y_{L_k}},\quad \forall k. \tag{57}$$

Then it follows that

$$\eta_{L_k}^{*}=\sqrt{\frac{f(\sigma_{L_k})}{y_{L_k}}}\frac{c\sum_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}}{\left(h(\sigma_G)+\sum_{k=1}^{K}\sigma_{L_k}y_{L_k}\right)},\quad \forall k.$$

This completes the proof.

APPENDIX E
PROOF THEOREM3

According to $\eta_{L_k}\leq\frac{cf(\sigma_{L_k})}{\sigma_{L_k}y_{L_k}}$, a low layer satellite k can upload its traffic to the GEO satellite. The condition can be written as the follows:

$$\sqrt{\frac{f(\sigma_{L_k})}{y_{L_k}}}\frac{c\sum_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}}{\left(h(\sigma_G)+\sum_{k=1}^{K}\sigma_{L_k}y_{L_k}\right)}\leq\frac{cf(\sigma_{L_k})}{\sigma_{L_k}y_{L_k}}. \tag{58}$$

In addition, (58) can be expressed as follows:

$$h(\sigma_G)\geq\left(\sqrt{\frac{y_{L_k}}{\sigma_{L_k}f(\sigma_{L_k})}}\sum_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}-\sum_{k=1}^{K}\sigma_{L_k}y_{L_k}\right),\quad \forall k. \tag{59}$$

Then we can determine the minimum value of $h(\sigma_G)$ when K' nodes participate in the game. Mathematically, this can be written as follows:

$$h(\sigma_G)^{min}=\sqrt{\frac{y_{L_K}}{\sigma_{L_K}f(\sigma_{L_K})}}\sum_{k=1}^{K}\sqrt{f(\sigma_{L_k})y_{L_k}}-\sum_{k=1}^{K}\sigma_{L_k}y_{L_k},$$

where $\frac{f(\sigma_{L_k})}{\sigma_{L_k}y_{L_k}}$ can be sorted in $\frac{f(\sigma_{L_1})}{\sigma_{L_1}y_{L_1}}>\frac{f(\sigma_{L_2})}{\sigma_{L_2}y_{L_2}}>\ldots>\frac{f(\sigma_{L_k})}{\sigma_{L_K}y_{L_K}}$.

This completes the proof.

APPENDIX F
PROOF THEOREM4

Consider $v_1,v_2=1,2,\ldots,K$ and $v_1=v_2+1$. Then, we can prove $x_{L_{v_1}}>x_{L_{v_2}}$. We have

$$\begin{aligned}
x_{L_{v_1}}&=\sqrt{\frac{y_{L_{v_1}}}{\sigma_{L_{v_1}}f(\sigma_{L_{v_1}})}}\sum_{k=1}^{v_1}\sqrt{f(\sigma_{L_{v_1}})y_{L_k}}-\sum_{k=1}^{v_1}\sigma_{L_k}y_{L_k}\\
&=\sqrt{\frac{y_{L_{v_2}}}{\sigma_{L_{v_2}}f(\sigma_{L_{v_2}})}}\sum_{k=1}^{v_2}\sqrt{f(\sigma_{L_k})y_{L_k}}-\sum_{k=1}^{v_2}\sigma_{L_k}y_{L_k}\\
&\quad+\sqrt{\frac{y_{v_1}}{\sigma_{L_{v_1}}f(\sigma_{L_{v_1}})}}\sum_{j=v_2+1}^{v_1}\sqrt{f(\sigma_{L_j})y_j}\\
&\quad-\sum_{k=1}^{v_1}\sigma_{L_{v_1}}y_k+\sum_{k=1}^{v_2}\sigma_{L_{v_2}}y_k\\
&>\sqrt{\frac{y_{L_{v_2}}}{\sigma_{L_{v_2}}f(\sigma_{L_{v_2}})}}\sum_{k=1}^{v_2}\sqrt{f(\sigma_{L_k})y_{L_k}}-\sum_{k=1}^{v_2}\sigma_{L_k}y_{L_k}\\
&>\sqrt{\frac{y_{L_{v_2}}}{\sigma_{L_{v_2}}f(\sigma_{L_{v_2}})}}\sum_{k=1}^{v_2}\sqrt{f(\sigma_{L_k})y_{L_k}}-\sum_{k=1}^{v_2}\sigma_{L_k}y_{L_k}\\
&>x_{L_{v2}}.
\end{aligned}$$

This completes the proof.

APPENDIX G
MARTINGALE

The Martingales form an extremely useful class of random processes which appear in many fields (e.g., finance, machine learning, information theory, etc.). Consider a sequence of random variables $X_0,X_1,X_2,\cdots,X_n$. The sequence $\{X_i\}$ is called a discrete-time martingale if it satisfies the following condition:

$$E(X_{i+1}|X_1,X_2,\cdots,X_i)=X_i. \tag{60}$$

For any i, it holds that

$$\begin{aligned}
E(X_i)&=E[E[X_i|X_1,X_2,\cdots,X_{i-1}]]\\
&=E[X_{i-1}]=\cdots=E[X_1].
\end{aligned} \tag{61}$$

The sequence $\{X_i\}$ is called a discrete-time supermartingale if it satisfies the following condition:

$$E[X_i|X_1,X_2,\cdots,X_{i-1}]\leq X_{i-1}. \tag{62}$$

Stopping time of concept: Define a non-negative, integer-valued random variable T as a stopping time of a sequence $\{Z_n,n>0\}$, if the event $T=n$ such that it depends only

on the value of random variables $Z_1, Z_2, \cdots, \cdots Z_n$. For instance, consider a set of the system with n nodes sharing the cache, the traffic arrival at the system between times 0 and t is denoted by $A(t) = \sum_{i=1}^{i=n} A_i(t)$ and the departure from the system is denoted by $D(t) = \sum_{j=1}^{j=m} D_j(t)$.

We define the stopping time N as the first time when the queue length Q in the cache exceeds q. The Martingale-Envelope [17] is defined as follows: For $p \geq 0$ and monotonically increasing functions $h_1, h_2, \cdots, h_p : \mathbb{R}_+ \to \mathbb{R}_+$ and $\theta > 0$, it is said that the flow A admits a $(\Pi h, \theta, c)$-Martingale-Envelope for every $m \geq 0$. The process is given by

$$\Pi\left(\overrightarrow{h(a_n)}e^{\theta(A^r(m,n)-(n-m)c)}\right) \leq M_m(n), \tag{63}$$

Also, it is almost surely bounded by a supermartingale $M_m(n)$.

REFERENCES

[1] H. Nishiyama, Y. Tada, N. Kato, N. Yoshimura, M. Toyoshima, and N. Kadowaki, "Toward optimized traffic distribution for efficient network capacity utilization in two-layered satellite networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1303–1313, Mar. 2013.

[2] G. Song, M. Chao, B. Yang, and Y. Zheng, "TLR: A traffic-light-based intelligent routing strategy for NGEO satellite IP networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3380–3393, Jun. 2014.

[3] L. Jiang, G. Cui, S. Liu, W. Wang, D. Liu, and Y. Chen, "Cooperative relay assisted load balancing scheme based on stackelberg game for hybrid GEO-LEO satellite network," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[4] H. Nishiyama, D. Kudoh, N. Kato, and N. Kadowaki, "Load balancing and QoS provisioning based on congestion prediction for GEO/LEO hybrid satellite networks," *Proc. IEEE*, vol. 99, no. 11, pp. 1998–2007, Nov. 2011.

[5] L. Jun, Z. Ji-Wei, and X. Nan, "Research and simulation on an autonomous routing algorithm for GEO_LEO satellite networks," in *Proc. 4th Int. Conf. Intell. Comput. Technol. Automat.*, vol. 2, Mar. 2011, pp. 657–660.

[6] X. Feng, M. Yang, and Q. Guo, "A novel distributed routing algorithm based on data-driven in GEO/LEO hybrid satellite network," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[7] T. C. House, "An analysis format for client-server performance using GEO & LEO satellite networks (inmarsat vs. globalstar)," in *Proc. Adv. Comput., Inf., Syst. Sci., Eng.*, Dordrecht, The Netherlands: Springer, 2006, pp. 165–170.

[8] M. Jiang, Y. Liu, W. Xu, F. Tang, Y. Yang, and L. Kuang, "An optimized layered routing algorithm for GEO/LEO hybrid satellite networks," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Aug. 2016, pp. 1153–1158.

[9] I. F. Akyildiz, E. Ekici, and M. D. Bender, "MLSR: A novel routing algorithm for multilayered satellite IP networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 3, pp. 411–424, Jun. 2002.

[10] Y. Kawamoto, H. Nishiyama, N. Kato, and N. Kadowaki, "A traffic distribution technique to minimize packet delivery delay in multilayered satellite networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3315–3324, Sep. 2013.

[11] Y. Wang *et al.*, "Tailored load-aware routing for load balance in multilayered satellite networks," in *Proc. 82nd IEEE Veh. Technol. Conf.*, Sep. 2015, pp. 1–5.

[12] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: An overview," *Comput. Netw.*, vol. 54, no. 14, pp. 2537–2561, 2010.

[13] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A stackelberg game approach," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 538–549, Apr. 2012.

[14] G. Vazquez-Vilar, C. Mosquera, and S. K. Jayaweera, "Primary user enters the game: Performance of dynamic spectrum leasing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 1–5, Dec. 2010.

[15] P. Xu, X. Fang, M. Chen, and Y. Xu, "A stackelberg game-based spectrum allocation scheme in macro/femtocell hierarchical networks," *Comput. Commun.*, vol. 36, no. 14, pp. 1552–1558, Aug. 2013.

[16] J. Li, H. Chen, Y. Chen, Z. Lin, B. Vucetic, and L. Hanzo, "Pricing and resource allocation via game theory for a small-cell video caching system," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 8, pp. 2115–2129, Aug. 2016.

[17] F. Poloczek and F. Ciucu, "Scheduling analysis with martingales," *Perform. Eval.*, vol. 79, pp. 56–72, Sep. 2014.

[18] F. Poloczek and F. Ciucu, "Service-martingales: Theory and applications to the delay analysis of random access protocols," in *Proc. IEEE INFOCOM*, Apr. 2015, pp. 945–953.

[19] Q. Li, Y. Zhang, A. Pandharipande, X. Ge, and J. Zhang, "D2D-assisted caching on truncated zipf distribution," *IEEE Access*, vol. 7, pp. 13411–13421, 2019.

[20] X. Song, Y. Geng, X. Meng, J. Liu, W. Lei, and Y. Wen, "Cache-enabled device to device networks with contention-based multimedia delivery," *IEEE Access*, vol. 5, pp. 3228–3239, 2017.

[21] M. Gregori, J. Gómez-Vilardebó, J. Matamoros, and D. Gündüz, "Wireless content caching for small cell and D2D networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1222–1234, May 2016.

[22] Z. Chen, J. Lee, T. Q. S. Quek, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3401–3415, May 2017.

[23] M. Hajimirsadeghi, N. B. Mandayam, and A. Reznik, "Joint caching and pricing strategies for popular content in information centric networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 3, pp. 654–667, Mar. 2017.

[24] F. Zhou, L. Fan, N. Wang, G. Luo, J. Tang, and W. Chen, "A cache-aided communication scheme for downlink coordinated multipoint transmission," *IEEE Access*, vol. 6, pp. 1416–1427, 2018.

[25] T. D. Tran and L. B. Le, "Joint resource allocation and content caching in virtualized content-centric wireless networks," *IEEE Access*, vol. 6, pp. 11329–11341, 2018.

[26] N.-S. Vo, T. Q. Duong, M. Guizani, and A. Kortun, "5G optimized caching and downlink resource sharing for smart cities," *IEEE Access*, vol. 6, pp. 31457–31468, 2018.

[27] A. Ndikumana, N. H. Tran, T. M. Ho, D. Niyato, Z. Han, and C. S. Hong, "Joint incentive mechanism for paid content caching and price based cache replacement policy in named data networking," *IEEE Access*, vol. 6, pp. 33702–33717, 2018.

[28] A. Armon and H. Levy, "Cache satellite distribution systems: Modeling, analysis, and efficient operation," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 2, pp. 218–228, Feb. 2004.

[29] S.-G. Chang, "Caching strategy and service policy optimization in a cache-satellite distribution service," *Telecommun. Syst.*, vol. 21, no. 2, pp. 261–281, Dec. 2002.

[30] S. Liu, X. Hu, Y. Wang, G. Cui, and W. Wang, "Distributed caching based on matching game in leo satellite constellation networks," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 300–303, Feb. 2018.

[31] H. Wu, J. Li, H. Lu, and P. Hong, "A two-layer caching model for content delivery services in satellite-terrestrial networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[32] G. Gür and S. Kafiloğlu, "Layered content delivery over satellite integrated cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 6, no. 3, pp. 390–393, Jun. 2017.

[33] A. Abdi, W. C. Lau, M.-S. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: First- and second-order statistics," *IEEE Trans. Wireless Commun*, vol. 2, no. 3, pp. 519–528, May 2003.

[34] K. An, M. Lin, W.-P. Zhu, Y. Huang, and G. Zheng, "Outage performance of cognitive hybrid satellite–Terrestrial networks with interference constraint," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 9397–9404, Nov. 2016.

[35] G. Choudhury and S. Rappaport, "Diversity ALOHA—A random access scheme for satellite communications," *IEEE Trans. Commun.*, vol. 31, no. 3, pp. 450–457, Mar. 1983.

[36] J. Bai and G. Ren, "Polarized MIMO slotted ALOHA random access scheme in satellite network," *IEEE Access*, vol. 5, pp. 26354–26363, 2017.

[37] H. Baek, J. Lim, and S. Oh, "Beacon-based slotted ALOHA for wireless networks with large propagation delay," *IEEE Commun. Lett.*, vol. 17, no. 11, pp. 2196–2199, Nov. 2013.

[38] M. Lee, J.-K. Lee, H. Noh, and J. Lim, "Stability of reservation-contention resolution diversity slotted ALOHA for satellite networks," in *Proc. IEEE Mil. Commun. Conf.*, Oct./Nov. 2012, pp. 1–6.

[39] M. Lee, J.-K. Lee, J.-J. Lee, and J. Lim, "R-CRDSA: Reservation-contention resolution diversity slotted ALOHA for satellite networks," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1576–1579, Oct. 2012.

[40] M. Furqan, C. Zhang, W. Yan, A. Shahid, M. Wasim, and Y. Huang, "A collaborative hotspot caching design for 5G cellular network," *IEEE Access*, vol. 6, pp. 38161–38170, 2018.
[41] G. Barrenetxea, B. Berefull-Lozano, and M. Vetterli, "Lattice networks: Capacity limits, optimal routing, and queueing behavior," *IEEE/ACM Trans. Netw.*, vol. 14, no. 3, pp. 492–505, Jun. 2006.

ERBAO WANG received the M.S. degree in control engineering from Xidian University, Xi'an, China, in 2007, where he is currently pursuing the Ph.D. degree in military communications with the State Key Lab of ISN. His current research interests include satellite networks, game theory, and martingale theory analysis.

HONGYAN LI (M'08) received the M.S. degree in control engineering from Xi'an Jiaotong University, Xi'an, China, in 1991, and the Ph.D. degree in signal and information processing from Xidian University, Xi'an, in 2000, where she is currently a Professor with the State Key Laboratory of Integrated Service Networks. Her research interests include spatial information networks, cognitive networks, integration of heterogeneous networks, and mobile ad hoc networks.

SHUN ZHANG received the B.S. degree in communication engineering from Shandong University, Jinan, China, in 2007, and the Ph.D. degree in communications and signal processing from Xidian University, Xi'an, China, in 2013. He is currently with the State Key Laboratory of Integrated Services Networks, Xidian University. His research interests include MIMO-OFDM systems, relay networks, and detection and parameter estimation theory.