IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 25, 2019, accepted January 18, 2020, date of publication January 22, 2020, date of current version January 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968594

A Hierarchical Approach to Resource Allocation in Extensible Multi-Layer LEO-MSS

YITAO LI[1], NA DENG[2,3], (Member, IEEE), AND WUYANG ZHOU[1], (Member, IEEE)
[1]Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230026, China
[2]School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China
[3]National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

Corresponding author: Wuyang Zhou (wyzhou@ustc.edu.cn)

This work was supported in part by the Project under Grant JHZK2018-203-TS-YZ, in part by the National Natural Science Foundation of China under Grant 61701071, and in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2019D03.

ABSTRACT Low earth orbit mobile satellite system (LEO-MSS) is the major system to provide communication support for mobile terminals beyond the coverage of terrestrial communication systems. However, the quick movement of LEO satellites and current single-layer system architecture impose restrictions on the capability to provide satisfactory service quality, especially for the remote and non-land regions with high traffic requirement. To tackle this problem, high-altitude platforms (HAPs) and terrestrial relays (TRs) are introduced to cover hot-spot regions, and the current single-layer system becomes an LEO-HAP multi-layer access network. Under this setup, we propose a hierarchical resource allocation approach to circumvent the complex management caused by the intricate relationships among different layers. Specifically, to maximize the throughputs, we propose a dynamic multi-beam joint resource optimization method in LEO-ground downlinks based on the predicted movement of LEO satellites. Afterwards, we propose the dynamic resource optimization method of HAP-ground downlinks when LEO satellites and HAPs share the same spectrum. To solve these problems, we use the Lagrange dual method and Karush-Kuhn-Tucker (KKT) conditions to find the optimal solutions. Numerical results show that the proposed architecture outperforms current LEO-MSS in terms of average capacity. In addition, the proposed optimization methods increase the throughputs of LEO-ground downlinks and HAP-ground downlinks with an acceptable complexity.

INDEX TERMS Radio resource allocation, multi-beam satellite, multi-layer satellite network, LEO mobile satellite system, space-air-ground integrated network.

I. INTRODUCTION

A. MOTIVATION

Compared with terrestrial cellular communication systems, low earth orbit mobile satellite systems (LEO-MSSs) have a prominent superiority in global coverage, which can not only provide sufficient capacity in remote and non-land regions [1], but also play a critical role in emergency communication [2]. However, due to the large coverage radius and quick movement of LEO satellites, there are still some imperfections in LEO-MSS. LEO-MSS cannot satisfy the demands of huge capacity in ultra-dense regions because the coverage radius of most LEO satellite beams is more than 100 km. Besides, low earth orbit (LEO) satellites move at an extremely high velocity and bring great difficulties to guarantee the quality-of-service (QoS) of mobile terminals (MTs) via conventional raido resource management [3]. To overcome these imperfections, a promising option is to devise a novel architecture by introducing extra access nodes for ultra-dense regions, thus forming an integrated network.

How to manage the resource of this multi-layer integrated network and maximize the system throughtput becomes another problem. Since LEO satellites are equipped with multi-beams and the scenarios of LEO satellites systems are different from existing terrestrial systems and geostationary (GEO) satellite systems, especially after introducing novel architectures, it is hard for the existing methods to provide great performance and implement smoothly. Specifically, the coverage radius of each LEO satellite beam is up to hundreds of kilometers, making the interference to

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

be a serious problem. Due to the frequently passive group handover caused by the quick movement of LEO satellites, quality of service (QoS) can hardly be guaranteed by using traditional methods. In addition, the long propagation delay and its periodical change reduce the accuracy and real-time property of channel state information.

Fortunately, the orbits of LEO satellites are always regular and predictable [4], making it possible to predict the channel and propagation delay. This brings promising benefits to propose novel radio resource managment strategies. In summary, to overcome the inherent defects of LEO-MSSs, an extensible LEO-MSSs architecture needs to be proposed and the corresponding resource allocation methods need to be introduced to adapt to the satellite scenarios and maximize the system throughtput.

B. RELATED WORK

The network architectures that contain satellites have got attention of researchers [5]–[8]. In [5], the authors proposed a flexible architecture for heterogeneous satellite-terrestrial networks and presented its functions and protocols with integration of network convergence, routing scalability alleviation, mobility support, traffic engineering, efficient content delivery, etc. Nevertheless, they mainly studied the network topology structure of integrating satellite and terrestrial networks but did not give practical examples, reducing the practicability of their work. Besides, they neither analyze the corresponding performance or quantify the performance gain, nor propose efficient radio resource management strategy to increase the capacity. In [6], a hybrid satellite-aerial-terrestrial network architecture is introduced for the emergency scenarios. They compared the advantages and disadvantages of satellites, high-altitude platforms (HAP), low-altitude platforms (LAP) and terrestrial networks. They also presented some promising technologies of this network. However, they did not present link selection and spectrum planning scheme to improve the architecture. And the architecture cannot be generalized to mobile communication scenarios for the remote or non-land regions beyond the coverage of terrestrial networks because some facilities are temporary (e.g. headquarters) or cannot be deployed at remote or non-land regions (e.g. emergency vehicles). The authors in [7] proposed a software defined space-air-ground integrated vehicular (SSAGV) network architecture, including satellites, HAPs and cellular networks, to serve vehicles. The SSAGV network is divided into application layer, control layer, infrastructure layer and physical systems. They also introduced hierarchical network operations and big data-assisted operations to improve the network. However, they did not consider the node deployment, mobility management and spectrum planning scheme of their architectures. Besides, the SSAGV network architecture uses all kinds of entities in space, air and ground segment, making it too complicated. In fact, to be generalized to mobile communication scenarios in practical systems, parts of the entities need to be merged together. The authors in [8] proposed a more flexible and adaptive architecture based on hybrid satellite terrestrial 5G networks with software defined features. They also investigated access strategy, coverage probability and radio resource management aiming at two types of services. However, this architecture merely focused on the convergence of GEO satellites and terrestrial 5G network, without mentioning LEO satellites and aerial network, which can not improve the network capacity of the regions beyond the coverage of terrestrial communication systems, such as disaster regions.

Another line of the existing research work has focused on the resource management of satellite systems [9]–[12] in recent years. In [11], the authors presented various optimal dynamic capacity allocation schemes in the multi-beam GEO system based on the theory of Monge arrays. They focused on reducing the complexity of their algorithms. Nevertheless, in order to simplify the problem, the system model and parameters are too utopian, making the algorithm lack of practical significance in actual systems. In [12], the authors proposed a novel algorithm of power allocation and interference pricing for a multi-beam GEO satellite by using dynamic game model. The problem was modeled by fixing the Nash equilibrium. However, they did not take the traffics of different users into account. The proposed algorithm was used for uplink but the required parameters are difficult to obtain at the side of user terminals. In [13] and [14], the authors focused on the spectrum sharing problem between satellite network and terrestrial network. They proposed efficient methods to solve the problem respectively. But the scenario in this paper is more specific and complicated than that of them, so we have to consider more factors and approaches. In [15], the authors presented a general interference analysis model of a cognitive network with GEO and LEO satellites. To enhance the spectral efficiency and protect the primary system, they proposed an optimization algorithm with beamhopping and adaptive power control techniques. However, they neglected the issue of different propagation delay from satellites to users, which would cause inaccuracy and reduce the throughput. Besides, due to the different scenarios, we consider an integrated network instead of a cognitive network, where the methods in [15] can not achieve an excellent performance.

To maximize the capacity of the regions beyond the coverage of terrestrial communications, both network architectures and resource allocation methods should be improved. Different from the current architectures, we provide more fine-grained works in the following three important aspects that all of the four papers ignored.

1) We provide the deployment scheme of different kinds of facility nodes.

2) We present different kinds of links and the scheme of link selection.

3) We give suggestions of spectrum planning of different space-air-ground links and inter-satellite links with multi-beams.

The channel of satellites-ground links and HAP-ground links is different from cellular networks. In this paper,

the larger propagation delay and predictability of channels caused by the movement of LEO satellites are considered, which are the salient properties of the extensible LEO-MSS networks. In our proposed strategy, we make full use of the predictability of channels by designing corresponding utility function and increase the throughput by considering the tradeoff between delay and the predicted channel gains.

C. CONTRIBUTIONS

In this paper, we focus on improving the network capacity of the regions beyond the coverage of terrestrial communications by proposing a multi-layer architecture based on LEO-MSS and optimizing its resource allocation. Generally, we summarize the contributions of this paper as follows.

- Via introducing HAPs and TRs into the traditional single-layer network architecture of LEO mobile satellite system, a multi-layer architecture is formed where LEO satellites provide wide-area coverage and the introduced nodes cover the hot-spot regions to increase the capacity.
- To implement the radio resource management efficiently in this system, we design a novel hierarchical resource manager (HRM) framework based on the multi-beam structure and solve the delay management problem.
- We propose a multi-beam joint dynamic radio resource optimization method in LEO-ground downlinks to maximize the throughputs based on the predicted information according to regular orbits of LEO satellites. After introducing HAPs, the system becomes a LEO-HAP multi-layer access network. According to the proposed resource management framework, we propose a dynamic resource optimization method of HAP-ground downlinks when LEO satellites and HAPs share the same spectrum.
- To solve these problems, we use Lagrange dual method and Karush-Kuhn-Tucker (KKT) conditions to find the optimal solutions. Afterwards, we propose corresponding gradient descent algorithms for each optimization problems to get the results with an acceptable complexity.
- Simulations are implemented to analyze the performance improvement of the proposed architecture and optimization methods. Numerical results show that the presented architecture has a much better average capacity than LEO-MSS. In addition, the proposed optimization methods further increase the throughputs of the LEO-ground and HAP-ground downlinks.

The remainder of this paper is structured as follows. In Section II, the extended multi-layer architecture based on multi-beam LEO-MSS and its performance analysis are presented. In section III a novel radio resource optimization method is proposed and modeled. Solution techniques and algorithms are proposed in Section IV. Numerical results are presented and discussed in Section V. Finally, the conclusions and future directions are drawn in Section VI.

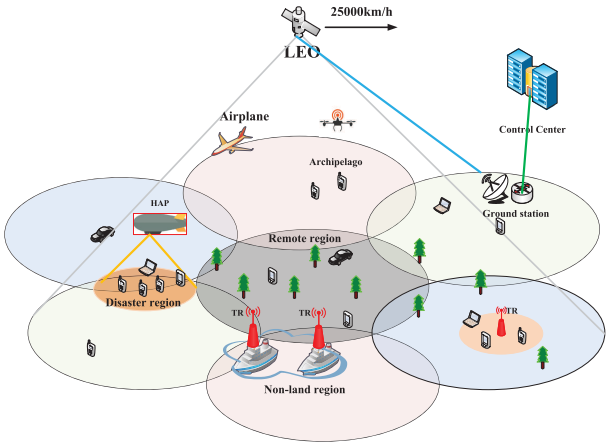


FIGURE 1. Multi-beam access scenarios.

II. EXTENSIBLE MULTI-LAYER LEO-MSS ARCHITECTURE

In this section, we will present the multi-layer architecture and analyze the capacity performance of the presented architecture. The architecture is introduced in terms of node types and functions, node deployment scheme, link selection and frequency spectrum planning. Capacity is calculated according to user distribution model and channel model.

A. EXTENSIBLE MULTI-LAYER LEO-MSS ARCHITECTURE

In general, typical traditional LEO-MSS of Walker constellation (e.g. the Iridium and Globalstar) consists of 48-80 LEO satellites with multi-beam coverages on the ground, several earth stations (ES) and one data control center. On this basis, the proposed architecture further introduces HAPs and TRs as the access nodes in hot-spot regions and geostationary earth orbit (GEO) satellites as routing nodes. HAPs could be airships or helium balloons with the altitude of about 20 kilometers to 100 kilometers high in the air [16], [17]. In addition, HAPs can move at a low velocity if necessary. TRs are full duplex relays and have an extra directional antenna pointing to the space to communicate with satellites. Since there is no need to equip TRs with wired network facilities, the deployment of TRs is unlimited, low-cost and flexible. Fig. 1 shows the access scenarios of this system, and the proposed extensible multi-layer architecture is shown in Fig. 2, where the orange circles are hot spot regions covered by HAPs or TRs, and the green circle is the remote region covered by satellites.

In the proposed architecture, mobile terminals (MTs) are not permitted to connect to GEO satellites directly because the high energy consumption is beyond the scope of the MTs. MTs in hot-spot regions preferentially connect to HAPs and TRs, while other MTs can only connect to LEO satellites. Different from the sparse regions with low user density, the hot-spot regions with higher user density are divided into two types according to the area size. The small hot-spot regions, including big ships, airplanes and small islands, are preferentially equipped with TRs. The broad hot-spot regions,

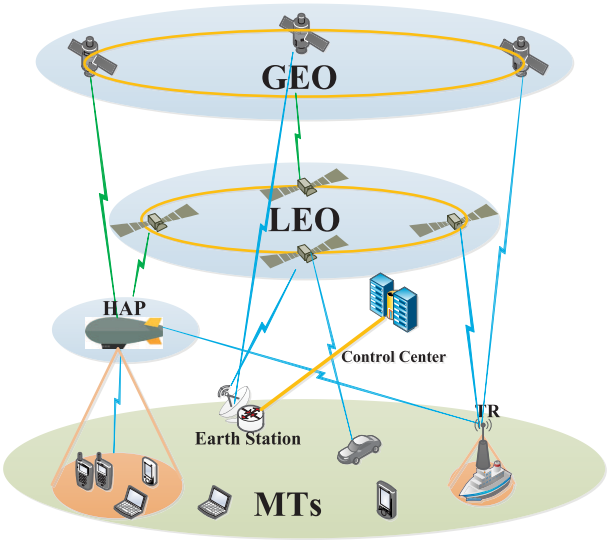


FIGURE 2. The proposed extensible multilayer architecture.

including archipelagoes, disaster regions and military bases, are preferentially served by HAPs.

Compared with current LEO-MSS system, the advantages of the proposed extensible multi-layer architecture are summarized as follows.

- Via HAPs and TRs, the proposed architecture can provide communication services to small mobile terminals that are not able to connect to LEO satellites directly.
- Since HAPs are able to move purposefully, the flexibility of system is significantly increased.
- The proposed architecture can increase the system capacity greatly, especially in the hot-spot regions.
- The proposed architecture can reduce the handover rate of the whole system through HAPs and TRs, thus significantly reduce the system signaling cost and the difficulty of radio resource management.
- As routing nodes, GEO satellites can reduce the complexity of routing, handover and rerouting rate significantly.

In the proposed architecture, HAPs and TRs are deployed according to the spatial distributions of MTs. Firstly, MTs get their own geographical location information by using satellite positioning technology [18], [19], and report their own geographical location information periodically to the system according to mobility management strategy [20]. Then, the control center performs the fitting of the reported data and obtain the spatial distribution law of the MTs. Finally, HAPs and TRs are deployed accurately according to the fitting results [21], [22].

There are two types of links in this system, i.e., access link and routing link. Access link consists of MT-TR link, MT-HAP link and MT-LEO link. MTs in hot-spot regions covered by HAPs and TRs prioritize MT-HAP link and MT-TR link, while MTs in sparse regions select MT-LEO link directly. Routing link consists of GEO-GEO link, LEO-GEO link, HAP-GEO link, TR/ES-GEO link, LEO-LEO link, HAP-LEO link, TR/ES-LEO link, HAP-HAP link, etc. Routing links connect all the nodes together and form an integrated network. Specifically, GEO satellites are used as routing nodes preferentially to reduce the hop counts if the transmitter and the receiver are far from each other, whereas LEO satellites are used as routing nodes preferentially to reduce delay if the transmitter and the receiver are not far from each other.

TABLE 1. Node connection and spectral selection.

Node Connection	Spectral Selection
GEO-GEO Inter-Satellite Link	*Laser*
GEO-LEO Inter-Satellite Link	*EHF*
GEO-HAP Space-Air Link	*EHF*
GEO-TR/ES Space-Ground Link	*Ka Band*
LEO Intra-Orbit Inter-Satellite Link	*Laser*
LEO Inter-Orbit Inter-Satellite Link	*EHF*
LEO-HAP Space-Air Link	*EHF*
LEO-TR/ES Space-Ground Link	*Ka Band*
LEO-MT Space-Ground Link	*Ka Band*
HAP-HAP Inter-HAP Link	*EHF*
HAP-MT Air-Ground Link	*Ka Band*
TR-MT relay Link	*C Band*

As Ka (typically $17.7GHz - 20.2GHz$ and $26.5GHz - 40GHz$) [23], [24] band has a higher frequency and wider spectrum than C ($4GHz - 8GHz$) band, it is used in satellite-ground links and HAP-ground links to avoid co-channel interference (CCI) with terrestrial cellular networks and provide global coverage. It is worth noting that C band can be used in TR-ground links since TRs are usually deployed in the regions where terrestrial cellular communication systems are underdeveloped. Thus we do not need to consider CCI between TR-ground links and satellite-ground (or HAP-ground) links. Extremely high frequency (EHF, $40GHz - 300GHz$) [25], [26] and laser [27] beam can be used in inter-satellite links and satellite-HAP links because they belong to the free-space links without atmosphere fading. Specifically, laser beam can be used in GEO inter-satellite links and LEO intra-orbit inter-satellite links. EHF can be used in GEO-LEO inter-satellite links, HAP-GEO links, HAP-LEO links, LEO inter-orbit inter-satellite links, and links between HAPs. Detailed node connection and spectral selection are shown in Table 1.

HAP-ground links share the same Ka spectrum with LEO-ground links. The QoS of MTs is affected by CCI. Therefore, frequency reuse schemes are used among HAP-ground links and LEO-ground links. Let B_{total} be the total available Ka spectrum bandwidth of LEO-ground links and HAP-ground links, $1/N$ be the frequency reuse factor of LEO beams. When the coverage region of a HAP is adjacent to that of n LEO satellite beams, the average bandwidth available for the HAP can be calculated approximatively as

$$B_{\text{H}} = \frac{N-n}{N} B_{\text{total}}. \tag{1}$$

B. CAPACITY ANALYSIS

1) USER DISTRIBUTION

LEO-MSS plays a major role in remote regions, non-land regions and emergency regions, where different types of service regions have their own user distributions. Specifically, it is justified that the user distribution in remote and non-land regions follows a sparse uniform distribution, denoted by

$$f_0(x,y) = \begin{cases} \frac{1}{S}, & (x,y) \in D \\ 0, & (x,y) \notin D \end{cases} \tag{2}$$

where S is the area of current regions D. In hot-spot regions such as disaster regions, the user distribution is assumed to follow normal distribution,[1] denoted by

$$\begin{aligned} &f_i(x,y) \\ &= \frac{1}{2\pi\sigma_{i,1}\sigma_{i,2}\sqrt{1-\rho_i^2}} \exp\left\{-\frac{1}{2\left(1-\rho_i^2\right)}\left[\frac{\left(x-\mu_{i,1}\right)^2}{\sigma_{i,1}^2}\right.\right. \\ &\left.\left.-2\rho_i\frac{\left(x-\mu_{i,1}\right)\left(y-\mu_{i,2}\right)}{\sigma_{i,1}\sigma_{i,2}}+\frac{\left(y-\mu_{i,2}\right)^2}{\sigma_{i,2}^2}\right]\right\} \end{aligned} \tag{3}$$

for hot-spot region i, where $(\mu_{i,1}, \mu_{i,2})$ is the center of the ith hot-spot region, $\sigma_{i,1}^2$, $\sigma_{i,2}^2$ are the variance of x and y, and ρ is the covariance between x and y. Therefore, the compound user distribution in the proposed system consists of one whole sparse region and several hot-spot regions, given by

$$f(x,y) = \omega_0 f_0(x,y) + \sum_{i=1}^{I} \omega_i f_i(x,y), \tag{4}$$

where

$$\sum_{i=0}^{I} \omega_i = 1. \tag{5}$$

I is the amount of hot-spot regions, ω_0 is the percentage of MTs in sparse ordinary regions, and ω_i is the percentage of MTs in the ith hot-spot region.

2) CHANNEL MODEL

The channel model mainly consists of the path loss, pitch angle fading, atmospheric fading and Rician small-scale fading. The specific expression is given by [28]

$$G = \left(\frac{C_{\text{light}}}{4\pi d f_c}\right)^2 \cdot G_H(\psi) \cdot A(d) \cdot \varphi, \tag{6}$$

where f_c is the carrier frequency, C_{light} is the velocity of light, d is the propagation distance between the transmitter and the receiver. The propagation distance between HAPs or satellites and the MT is given by

$$d = \sqrt{h^2 + \left(x_j - o_{i,1}\right)^2 + \left(y_j - o_{i,2}\right)^2}. \tag{7}$$

[1]The actual user distribution should be obtained according to the empirical data. The analysis in this paper can be easily extended to arbitrary user distributions.

h is the altitude of HAPs or satellites, $(o_{i,1}, o_{i,2})$ is the position of sub-satellite point or sub-HAP point, and (x_j, y_j) is the position of jth MT.

In (6), $G_H(\psi)$ is the pitch angle fading, expressed as

$$G_H(\psi) = Ap_{\text{eff}} \cdot \cos(\psi)^{\eta} \frac{32\log 2}{2\left(2\arccos\left(\sqrt[\eta]{0.5}\right)\right)^2}, \tag{8}$$

where Ap_{eff} is the antenna aperture efficiency (assumed to be unity) and η is the antenna factor determining the coverage radius.[2] ψ is the pitch angle, given by

$$\psi = \arctan\frac{\sqrt{\left(x_j - o_{i,1}\right)^2 + \left(y_j - o_{i,2}\right)^2}}{h}. \tag{9}$$

$A(d)$ in (6) is the atmospheric fading, expressed as

$$A(d) = 10^{\left(\frac{3d\chi}{10h}\right)}, \tag{10}$$

where χ is the attenuation through the clouds and rain in dB/km. And φ in (6) is the Rician small-scale fading.

The received power can be expressed as

$$P_{\text{re}} = P_{\text{tr}} \cdot G \cdot G_{\text{re}}, \tag{11}$$

where P_{tr} is the transmission power and G_{re} is the antenna receiving gain.

3) CAPACITY

In this subsection, we analyze the improvement of system capacity brought by HAPs and TRs. Since EHF band used in satellite-HAP links can provide a much wider bandwidth than Ka band used in space-ground links, and the TR-MT links are with much smaller path loss than satellite-TR links, so the capacity of the whole system is mainly limited by space-ground links, which include LEO-MT links, HAP-MT links and satellite-TR links.

The spectrum of LEO beams are N color reused (Fig. 1 is the example of $N = 4$). The wide remote region is served by LEO beams and each hot-spot region is served by an HAP or a TR. Since the antenna array of the HAP is highly directed to the hot-spot region, the interference from an HAP is assumed to be negligible to the users served by LEO beams or other HAPs. The interfering sources of LEO beams and HAPs are the LEO beams that use the same color spectrum band.

Due to the spectrum reuse scheme, HAPs share the same spectrum with LEO beams. Therefore, the available spectrum for HAPs can be reused to provide additional capacity for hot-spot regions. Also, the smaller path loss of HAP-MT link provides larger capacity than LEO-MT link. For TRs, compared with the non-directional antennas of MTs, the equipped directional antennas can provide a much larger receiving gain from the associated beam while reduce the receiving gain from the interfering beams.

[2]The antennas in this paper are assumed to be non-steerable, which means the antennas cannot change their pointing direction according to the demands.

Let $i \in \{1, 2, \ldots, M_L, \ldots, NM_L\}$, $H \in \{1, 2, \ldots, M_H\}$ and $T \in \{1, 2, \ldots, M_T\}$ be the LEO beams, HAPs and TRs, where M_L is the total number of LEO beams per color, M_H and M_T are the total number of HAPs and TRs, respectively. $\phi_n = \{(n-1)M_L + 1, (n-1)M_L + 2, \ldots, nM_L\}$ is the set of LEO beams using the nth color spectrum band, where $n \in \{1, 2, \ldots, N\}$. For the MT located in (x, y) served by the LEO beam i, $i \in \phi_n$, the capacity of LEO-MT link is expressed as

$$C_i(x,y) = B_i \log_2\left(1 + \frac{P_{\text{tr}}(i) \cdot G_i(x,y) \cdot G_{\text{re}}}{\sum_{i' \in \phi_n, i' \neq i} P_{\text{tr}}(i') \cdot G_{i'}(x,y) \cdot G_{\text{re}} + \sigma^2}\right), \tag{12}$$

where i' is the interfering beams, B_i is the available spectrum for the MT located in (x, y) covered by LEO beam i, $P_{\text{tr}}(i)$ and $P_{\text{tr}}(i')$ are the transmission power of LEO beam i and interfering beam i', respectively. $G_i(x, y)$ and $G_{i'}(x, y)$ are the channel gains from LEO beam i and interfering beam i' to MT located in (x, y). G_{re} is the antenna receiving gain of MT. σ^2 is the noise power.

For the HAP-MT link in hot-spot region within the coverage of beam i, $i \in \phi_n$, the average bandwidth available for the HAP-MT link to share is $B_{\text{H}} = \frac{N-1}{N} \cdot B_{\text{total}}$, which is larger than $\frac{B_{\text{total}}}{N}$ for LEO-MT link. Therefore, after introducing HAPs, part of the bandwidth can be reused to increase the capacity of the hot-spot regions covered by HAPs. The additional capacity of HAP-MT link within the coverage of LEO beam i is expressed as

$$C_{\text{H}}(x,y) = B_{\text{H}} \log_2\left(1 + \frac{P_{\text{tr}}(\text{H}) \cdot G_{\text{H}}(x,y) \cdot G_{\text{re}}}{\sum_{i' \in \phi_{n'}, n' \neq n} P_{\text{tr}}(i') \cdot G_{i'}(x,y) \cdot G_{\text{re}} + \sigma^2}\right), \tag{13}$$

where $P_{\text{tr}}(\text{H})$ is the power of HAP and $G_{\text{H}}(x, y)$ is the channel gain from HAP to MT located in (x, y). i' is the interfering beam using the spectrum band with HAP. $\phi_{n'}$ is the set of interfering beams for the HAP.

The capacity of satellite-TR-MT link is limited by the capacity of satellite-TR link. For the TR located in (x, y) served by LEO beam i, $i \in \phi_n$, the directional antenna of TR provides a much larger antenna gain from beam i and smaller antenna gain from the interfering beam i'. The capacity of LEO-TR link is expressed as

$$C_{i,\text{T}}(x,y) = B_{i,\text{T}} \log_2\left(1 + \frac{P_{\text{tr}}(i) \cdot G_i(x,y) \cdot G^i_{\text{re}}(\text{TR})}{\sum_{i' \in \phi_n} P_{\text{tr}}(i') \cdot G_{i'}(x,y) \cdot G^{i'}_{\text{re}}(\text{TR}) + \sigma^2}\right), \tag{14}$$

where $G^i_{\text{re}}(\text{TR})$ and $G^{i'}_{\text{re}}(\text{TR})$ are the antenna receiving gain from beam i and the interfering beam i', respectively. $G^i_{\text{re}}(\text{TR}) > G^{i'}_{\text{re}}(\text{TR})$.

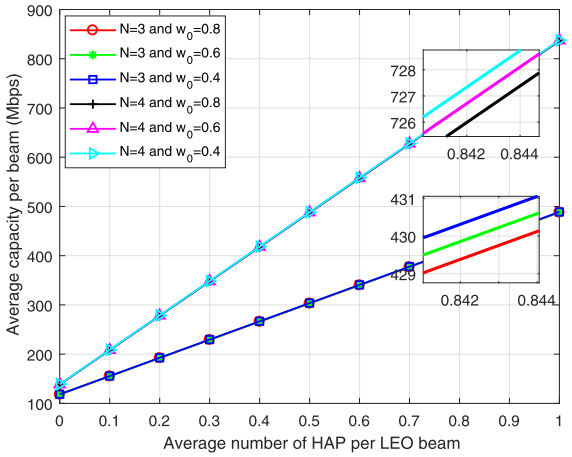


FIGURE 3. Average capacity with HAP in space-ground link.

For all MTs, the average total capacity of LEO-MT links, LEO-TR links and HAP-MT links is given by

$$C_{\text{total}} = \sum_{i=1}^{NM_L} C_i + \sum_{T=1}^{M_T} C_T + \sum_{H=1}^{M_H} C_H, \tag{15}$$

where C_i, C_T and C_H are the average total capacity of LEO-MT links, LEO-TR-MT links and HAP-MT links, respectively.

Fig. 3 and Fig. 4 show the numerical results of average capacity improvement brought by HAPs and TRs respectively. Important simulation parameters are shown in Table 3 in Section V.

The capacity improvement brought by introducing HAPs is shown in Fig. 3. It demonstrates that the average space-ground link capacity increases greatly with the help of HAPs, which supports the effectiveness in the system capacity of the proposed architecture. It indicates that N and the number of HAPs are the main factors of average capacity improvement. The bigger N is, the larger average space-ground link capacity per beam is, and the more dramatically HAPs improve average space-ground link capacity. The first reason is that CCI decreases when N increases. The second reason is that the bandwidth of HAPs is larger when N is larger. After N is determined, the average capacity of space-ground link increases linearly with the number of HAPs, because the equivalent bandwidth increases linearly as the number of HAPs increases. The percentage of MTs in hot-spot regions slightly affects the average channel gain of all MTs. So the average capacity changes slightly with different ω_0, which can be ignored compared with the impact of HAP number.

Fig. 4 presents the improvement of average space-ground link capacity brought by the introduction of TRs. Powerful antennas of TRs increase SINR equivalently. Different from HAPs, the number of TRs does not contribute to the increase of average space-ground link capacity actually. The main factor that influences the average space-ground link capacity is the percentage of MTs served by TRs. Average space-ground link capacity increases linearly with the percentage of

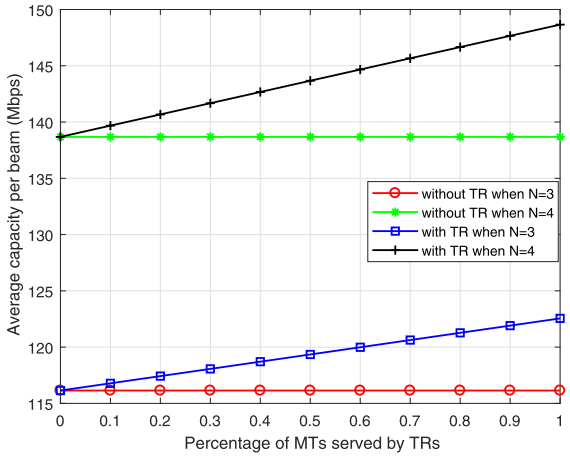


FIGURE 4. Average capacity with TR in space-ground link.

MTs that are served by TRs. Moreover, average space-ground link capacity increases more greatly when $N = 4$ than when $N = 3$.

System capacity improvement of HAPs is much better than that of TRs. Therefore, to maximize system capacity in hot-spot regions, we recommend to deploy HAPs preferentially. If deployment cost and sustainable utilization are taken into consideration comprehensively, TR is also an option. Deployment should be made according to specific indicators.

III. RADIO RESOURCE MANAGEMENT OPTIMIZATION METHOD

Due to the benefits of HAPs and TRs explained in II-B.3, the presented extensible multilayer LEO-MSS architecture provides better performance on the average capacity of space-air-ground links. To further increase the throughputs, a better radio resource management method needs to be proposed. The radio resource management of TRs is the same with that of base stations in terrestrial networks. Therefore, we do not consider the radio resource management problem of TRs in this paper. In this section, we will present a radio resource management framework according to the characteristics of multi-beam satellites and the multi-layer system. Afterwards, we will deal with the delay problem in the proposed multi-layer LEO-MSS which was ignored by researchers when they studied the radio resource management problem in satellites systems. Finally, we propose the radio resource optimization methods in the proposed multi-layer LEO-MSS based on the framework.

According to the implementation framework, the radio resource management has three steps. The first step is time-slot resource allocation, implemented in the queue management module of the HRM in LEO satellites and HAPs, and formulated as OP1. The second step is spectrum resource allocation which is also implemented in the HRM. The third step is the power optimization and it is implemented in the power control module of HRM in LEO satellites, and formulated as OP2. HAPs allocate the spectrum and power

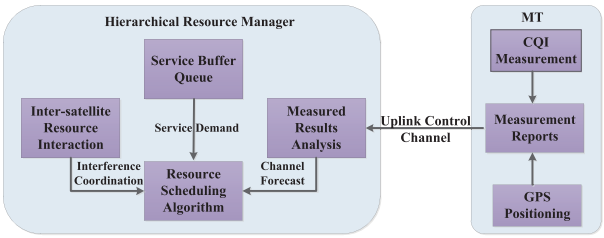


FIGURE 5. HRM framework.

according to the optimization results of LEO beams. The spectrum resource allocation of HAP is formulated as OP3 and implemented in the spectrum allocation module of HRM in HAP. The power allocation in HAP is formulated as OP4 and implemented in the power control module of HRM in HAP.

A. RADIO RESOURCE MANAGEMENT FRAMEWORK

According to the proposed extensible multi-layer LEO-MSS architecture, the access scenarios of multi-beam LEO satellites and HAPs are shown in Fig. 1. Each LEO satellite is equipped with 7 or 19 beams in general [30].

The function of radio resource management is mainly implemented in the HRM module. HRM contains resource scheduling algorithm module, measurement results analysis module, inter-satellite resource interaction module and service buffer queue, as shown in Fig. 5. And the HRM of all beams is integrated into one module in the LEO satellite in order to control jointly during each transmit time interval (TTI), where TTI is the minimum time interval unit of resource scheduling and equal to one time-slot in this system.

The measurement result analysis module is responsible for analyzing the reported channel quality information and positioning results, and providing the resource scheduling algorithm module with the analysis results. The GPS positioning results are used to estimate the path loss and pitch angle fading, while the channel quality indication (CQI) measurement results are used to calculate the sum of atmospheric fading and Rician small-scale fading. Thus, the channel gains from satellites to MTs are estimated and used as important parameters in resource allocation. The inter-satellite resource interaction module is responsible for sharing spectrum and power allocated information with other satellites and HAPs to coordinate the interference. Since HAPs share the same spectrum with LEO satellites, the resources of HAPs are allocated according to the allocation results of LEO satellites sent by LEO satellites through LEO-HAP links. The service buffer queue is responsible for storing and sorting packets according to the allocated results of the resource scheduling algorithm module.

The resource scheduling algorithm module is the core module of HRM and divided into three submodules in order, e.g. time-slot allocation, spectrum allocation and power optimization, as shown in Fig. 6. Firstly, the time-slot allocation submodule is responsible for allocating time-slot resource to MTs according to the designed utility function and the

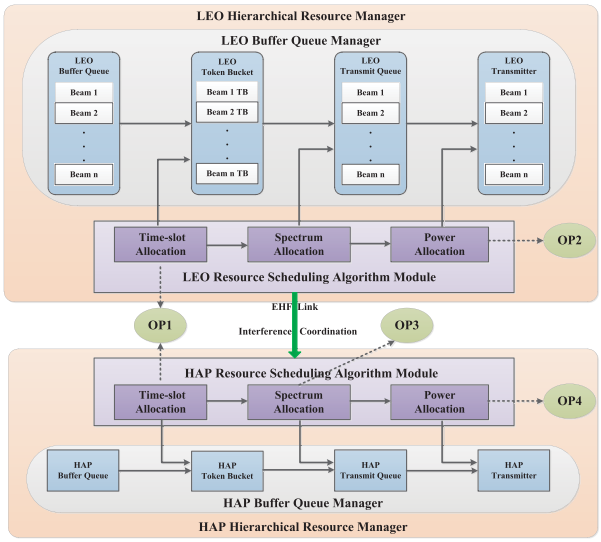


FIGURE 6. Resource scheduling algorithm module.

TABLE 2. Delay management.

Time	Meaning
t_0	timeslot of starting calculation in LEO
$(t_1 - t_0)$	calculation and processing delay of LEO
t_1	timeslot of LEO sending results to HAP
t_2	timeslot of LEO sending packets to MTs
$(t_3 - t_1)$	propagation delay from LEO to HAP
t_3	timeslot of HAP receiving results from LEO
$(t_4 - t_3)$	calculation and processing delay of HAP
t_4	timeslot of HAP sending packets to MTs
$(t_5 - t_4)$	propagation delay from HAP to MTs
$(t_5 - t_2)$	propagation delay from LEO to MTs
t_5	timeslot of MTs receiving packets from LEO and HAPs

formulated optimization problem OP1. Secondly, the spectrum allocation submodule allocates the spectrum resource for the selected MTs of OP1. Finally, the power optimization submodule allocates the transmission power for all the selected MTs. Since HAPs share the same spectrum with LEO satellites, the available resource for HAPs depends on the allocation results of LEO satellites. Thus, we design the hierarchical step by step resource allocation scheme.

B. DELAY MANAGEMENT

Compared with the radio resource management in terrestrial communication systems, there are more challenges in multi-layer LEO-MSS brought by the propagation delay of LEO satellites to ground. The propagation delay of MTs in the center of the satellites coverage is smaller than that of MTs in the edge of the satellites coverage. And the LEO-ground link propagation delay of the same MT changes periodically. Therefore, an extra delay management strategy should be presented to achieve accurate resource optimization, as shown in Fig. 7 and Table 2.

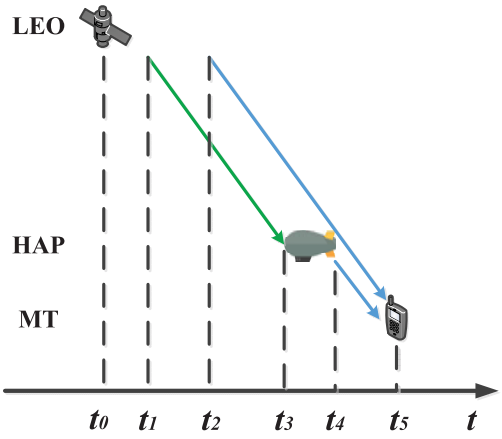


FIGURE 7. Delay management.

According to the framework, all the resources of satellites and HAPs are allocated hierarchically, as shown in Fig. 6. Since the location of LEO satellites changes periodically, $(t_3 - t_1)$ and $(t_5 - t_2)$ also change periodically. The calculation and processing delay $(t_1 - t_0)$ and $(t_4 - t_3)$ is fixed. $(t_2 - t_1) = (t_4 - t_3)$. Thus, $(t_2 - t_0)$ is fixed. t_3 depends on the distance between LEO satellite and HAP. t_5 depends on the distance between LEO satellite and MTs. Thus, for the same HAP and MT, t_3, t_4 and t_5 change periodically. For different HAPs and MTs, t_3, t_4 and t_5 are different at the same time-slot. Therefore, t_0 and t_2 are used as the reference time-slots of calculating and transmitting. And LEO satellites send control signallings to inform HAPs and MTs the time of receiving data. With this scheme, the interference of different beams in the actual system can be calculated accurately.

C. RADIO RESOURCE ALLOCATION IN LEO SATELLITES

According to the multi-layer architecture, HAPs share the same spectrum with LEO satellites. The resource allocations of HAPs are affected by the resource allocation results of LEO satellites. Therefore, we first study the resource allocation in LEO satellites. As shown in Fig. 6, the resource allocation is divided into time-slot resource allocation, spectrum allocation and power allocation in the HRM.

1) UTILITY DESIGN AND DYNAMIC TIME-SLOT ALLOCATION STRATEGY IN LEO

In LEO satellites, the first step of resource allocation is time-slot resource allocation. The channel gains of LEO satellites are regular, periodic and predictable, which is helpful for increasing the throughput of LEO-ground links. To make full use of this point, we design a utility function and propose a dynamic optimization strategy to allocate the time-slot resource to different MTs according to the designed utility function.

The velocity of LEO satellites at the altitude of 1500km is about 25000km/h, which is more than 100 times higher than velocity of MTs. Thus, it is assumed to ignore the change of MTs' geographical locations until the next report.

Besides, the atmospheric fading and Rician small-scale fading of each links are assumed to be unchanged until the next CQI report after dealing with the predicted information. Therefore, the channel gains between satellites and MTs can be forecasted accurately.

Beams are assigned N colors. There are no interference between beams of different color. Thus, we take one color beams as an example. Let I be the total number of this color of beams and K be the total number of resource block (RB). The total MT number of beam $i \in \{1, 2, \ldots, I\}$ is J_i.

We consider the priority of MTs, the delay of services, MTs' channel gains and their variation trend in the utility function. The utility function is designed according to the following observations.

(1) Some important MTs have a higher priority when allocated resource.

(2) We assume the packet whose delay is larger than its maximum time delay will be dropped. The delay utility should decrease rapidly when delay increases and approaches the maximum time delay [31].

(3) The channel gain utility should increase when the channel gain increases. If the channel gain is the same, the channel gain utility of the decline trend should be a little larger than that of the rise trend because we consider the predicted channel gain of the following period of time too.

Thus, the proposed utility function in this paper contains the delay utility function of the target services and the channel gain utility function of the target MTs. The utility function of MT $j_i \in \{1, 2, \ldots, J_i\}$ at timeslot t is defined as

$$U_{mt}^{i,j_i,t} = \omega_{j_i}(t) \frac{\omega_g(t) \cdot U_g^{i,j_i}(t)}{\omega_d(t) \cdot \max\left\{U_d^{j_i}(t), \varepsilon_d\right\}}, \tag{16}$$

where ω_d and ω_g are the weights of the delay utility function and channel gain utility function, respectively. ω_{j_i} is the weight of MT j_i. ε_d is the threshold of the delay utility function, to balance the tradeoff between the fairness and throughputs.

The delay utility function $U_d^{j_i}$ is defined as

$$U_d^{j_i}(t) = 1 - \exp\left(\frac{t - d_{j_i}^{\max}}{p_d}\right) \tag{17}$$

and shown in Fig. 8, where $d_{j_i}^{max}$ is the maximum time delay of the current service packet of the jth MT and p_d is a parameter depending on the service types.

The channel gain utility function U_g^{i,j_i} is defined as

$$U_g^{i,j_i}(t) = G_t^{i,j_i} \cdot \exp\left(p_g\left(G_t^{i,j_i} - G_{t+\Delta t}^{i,j_i}\right)\right) \tag{18}$$

and shown in Fig. 9, where G_t^{i,j_i} is the channel gain at t between MT j_i and beam i. p_g is the weight of the channel gain variation trend.

With the designed utility function, we propose a dynamic time-slot allocation strategy based on the predicted information to allocate the optimal time-slot resource for all MTs.

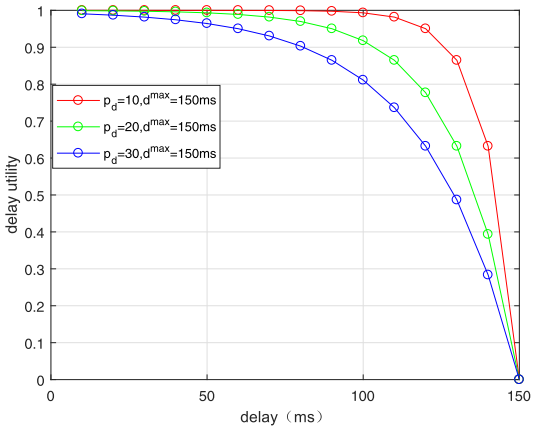


FIGURE 8. Delay utility.

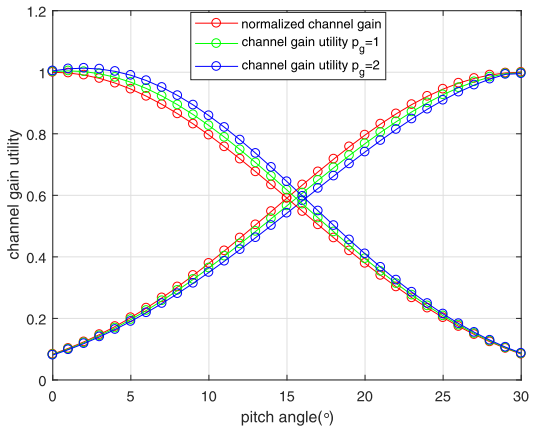


FIGURE 9. Channel gain utility.

This strategy is implemented in the queue management module of HRM.

The matrix $\boldsymbol{A}_t = [a_{i,j_i,t}]$ is defined as the allocation result. $a_{i,j_i,t} = 1$ refers to allocate timeslot t to MT j in beam i and $a_{i,j_i,t} = 0$ refers to not allocate timeslot t to MT j in beam i. The proposed dynamic resource allocation strategy is formulated as OP1.

$$\begin{aligned} OP1 \quad & \max \sum_{i=1}^{I} \sum_{j_i=1}^{J_i} a_{i,j_i,t} U_{mt}^{i,j_i,t} \\ & s.t.\ C1: \sum_{j_i=1}^{J_i} a_{i,j_i,t} \le K, \quad \forall i, t \\ & C2: a_{i,j_i,t} \in \{0,\ 1\}, \quad \forall j_i, i, t. \end{aligned} \tag{19}$$

OP1 is a linear integer programming problem. C1 is the constraint of total bandwith. We can get the optimal results of $\boldsymbol{A}_t$ by sorting $U_{mt}^{i,j_i,t}$.

2) SPECTRUM ALLOCATION AND POWER OPTIMIZATION IN LEO

After dynamic time-slot allocation, spectrum and power should be allocated in step 2 and step 3 to maximize the signal

to interference plus noise ratios (SINR) and the throughput of multi-beam satellites.

The spectrums are reused by multi-beams. MTs are allocated to the spectrums according to the current usage rate of spectrums one by one in each beam. Spectrum with the minimum current usage rate is allocated to the MT with the largest channel gain. The usage rates of all the spectrums are updated via the spectrum usage rate matrixes in the HRM of LEO satellites.

After the frequency allocation, all MTs allocated with RB is expressed as matrix $\boldsymbol{J} = [j_{i,k}]$. $j_{i,k}$ is the MT covered by beam i allocated with RB k. The rate of $j_{i,k}$ is

$$R_{j_{i,k},t} = B\log_2\left(1 + \frac{P_{i,k,t} \cdot G_{i,j_{i,k},t}}{\sum\limits_{i' \neq i}^{I} P_{i',k,t} \cdot G_{i',j_{i,k},t} + \sigma^2}\right), \quad (20)$$

where $P_{i,k,t}$ is the transmission power of beam i with RB k at t, $G_{i,j_{i,k},t}$ is the channel gain between beam i and MT $j_{i,k}$ at t, and B is the bandwith of each RB. The throughput of the whole system at timeslot t is defined as

$$Throughput_t = \sum_{i=1}^{I}\sum_{k=1}^{K} R_{j_{i,k},t}. \quad (21)$$

The multi-beam joint power optimization problem is formulated as OP2.

$$\begin{aligned} OP2 \quad & \max \ Throughput_t \\ & s.t \ C1: R_{\min}^{j_{i,k},t} \leq R_{j_{i,k},t} \leq R_{\max}^{j_{i,k},t}, \quad \forall j_{i,k}, t \\ & C2: \sum_{k=1}^{K} P_{i,k,t} \leq P_{\max}^{i}, \quad \forall i, t \\ & C3: 0 \leq P_{i,k,t} \leq P_{\max}, \quad \forall i, k, t, \end{aligned} \quad (22)$$

C1 is the con tes of each MTs. C2 is the constraint of the total power of each beam. C3 is the constraint of the transmission power of each RB.

D. RADIO RESOURCE ALLOCATION IN HAP

Existing work [32] has discovered that there would be great challenges in handover management and resource allocation for HAP with multi-beams because of the quasi-static characteristic of HAP. Therefore, we equip HAP with single beam in this paper.

LEO satellite sends the allocation results to HAP immediately. Then HAP works out the available spectrums and allocates the time-slot, spectrum and power to MTs. As HAPs cover the hot-spot regions in this system, the throughput is the most important metric. Thus, we propose a throughput-maximization method based on the above HRM framework.

1) RATE MAXIMIZATION TIME-SLOT AND SPECTRUM ALLOCATION IN HAP

As HAPs and LEO satellites share the same frequency band in this system, the available spectrum resource and co-channel interference in HAP downlink change periodically because of the periodical movement of LEO satellites. Thus, HAPs calculate the interference of all the downlink spectrum according to the transmission power of LEO satellites and compare them with the the interference threshold value ξ. The spectrum whose interference is less than ξ is considered available to the HAP downlink.

Let J^h be the total number of MTs covered by HAP h at t and $q_{h,k,t} = \sum\limits_{i=1}^{I} P_{i,k,t} G_{i,h,t}$ be the total interference of spectrum RB k ($k \in \{1, 2, \cdots, N \times K\}$). The spectrum RB k is available to HAP h at t when $q_{h,k,t} \leq \xi$. Let K^h be the total available spectrum RB at t calculated by HAP h. The maximum transmission power of each RB is $P'_{\max}$. The channel gain between HAP h and MT j^h ($j^h \in \{1, 2, \cdots, J^h\}$) at t is $G_{h,j^h,t}$. The rate of J^h with RB k^h ($k^h \in \{1, 2, \cdots, K^h\}$) is

$$R_{j^h,k^h,t}^{\max} = B\log_2\left(1 + \frac{P'_{\max} G_{h,j^h,t}}{\sum\limits_{i=1}^{I} P_{i,k^h,t} G_{i,j^h,t} + \sigma^2}\right). \quad (23)$$

Let

$$\boldsymbol{R}_t^{\max} = \begin{bmatrix} R_{1,1,t}^{\max} & R_{1,2,t}^{\max} & \cdots & R_{1,K^h,t}^{\max} \\ R_{2,1,t}^{\max} & R_{2,2,t}^{\max} & \cdots & R_{2,K^h,t}^{\max} \\ \vdots & \vdots & & \vdots \\ R_{J^h,1,t}^{\max} & R_{J^h,2,t}^{\max} & \cdots & R_{J^h,K^h,t}^{\max} \end{bmatrix} \quad (24)$$

and the resource allocation matrix

$$\boldsymbol{B}_t = \begin{bmatrix} b_{1,1,t} & b_{1,2,t} & \cdots & b_{1,K^h,t} \\ b_{2,1,t} & b_{2,2,t} & \cdots & b_{2,K^h,t} \\ \vdots & \vdots & & \vdots \\ b_{J^h,1,t} & b_{J^h,2,t} & \cdots & b_{J^h,K^h,t} \end{bmatrix}, \quad (25)$$

where $b_{j^h,k^h,t} = 1$ means RB k^h is allocated to MT j^h at t.

The proposed maximized total rate allocation problem is formulated as OP3.

$$\begin{aligned} OP3 \quad & \max \sum_{k^h=1}^{K^h}\sum_{j^h=1}^{J^h} b_{j^h,k^h,t} \cdot R_{j^h,k^h,t}^{\max} \\ & s.t. \ C1: \sum_{j^h=1}^{J^h} b_{j^h,k^h,t} \leq 1, \quad \forall k^h, t \\ & C2: \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \leq 1, \quad \forall j^h, t \\ & C3: \ b_{j^h,k^h,t} \in \{0, 1\}, \quad \forall j^h, k^h, t. \end{aligned} \quad (26)$$

OP3 is a linear integer programming problem and we will propose the corresponding algorithm to find the optimal allocation results.

2) POWER OPTIMIZATION IN HAP

Aiming at maximizing the throughputs of HAP-ground downlink, we formulate the power optimization

problem as OP4.

$$\begin{aligned} OP4 \quad & \max \ Throughput_t^h \\ & s.t\ C1: R_{\min}^{j^h,t} \le R_{j^h,t} \le R_{\max}^{j^h,t}, \quad \forall j^h, t \\ & C2: \sum_{k^h=1}^{K^h} P_{k^h,t} \le P_{\max}^h, \quad \forall t \\ & C3: 0 \le P_{k^h,t} \le P'_{\max}, \quad \forall k^h, t, \end{aligned} \tag{27}$$

where $P_{\max}^h$ is the maximum total transmission power in HAP downlink. The throughputs of HAP-ground downlink is defined as

$$Throughput_t^h = \sum_{j^h=1}^{J^h} R_{j^h,t}, \tag{28}$$

where

$$R_{j^h,t} = \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} B \log_2\left(1+\frac{P_{k^h,t}G_{h,j^h,t}}{\sum_{i=1}^{I} P_{i,k^h,t}G_{i,j^h,t}+\sigma^2}\right). \tag{29}$$

OP4 is a convex optimization problem. C1 is the constraint of the maximum and minimum rates of each MTs. C2 is the constraint of the total power. C3 is the constraint of the transmission power of each RB.

IV. PROPOSED SOLUTION TECHNIQUES AND ALGORITHMS

As mentioned above, OP1 and OP3 are linear integer programming problems. Thus, in this section, we will focus on how to obtain the solutions of OP2 and OP4.

A. PROPOSED SOLUTION TECHNIQUES OF OP2

OP2 is a non-convex optimization problem. There is a strong duality when K and I are big enough [33], which means the optimal results of the dual problem is the same with the primal problem. Thus, we use the Lagrangian dual method to find the optimal results. The Lagrangian function of OP2 is

$$\begin{aligned} & L(P_{i,k,t}, \lambda^{\min}, \lambda^{\max}, \mu) \\ & = \sum_{i=1}^{I}\sum_{k=1}^{K} R_{j_{i,k},t} \\ & \quad + \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{i,k,t}^{\min}\left(R_{j_{i,k},t} - R_{\min}^{j_{i,k},t}\right) \\ & \quad + \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{i,k,t}^{\max}\left(R_{\max}^{j_{i,k},t} - R_{j_{i,k},t}\right) \\ & \quad + \sum_{i=1}^{I} \mu_{i,t}\left(P_{\max}^{i} - \sum_{k=1}^{K} P_{i,k,t}\right), \end{aligned} \tag{30}$$

where $(\lambda^{\min}, \lambda^{\max}) = \{(\lambda_{i,k,t}^{\min}, \lambda_{i,k,t}^{\max}), k \in K, i \in I\} \ge 0$ and $\mu = \{\mu_{i,t}, i \in I\} \ge 0$ are Lagrangian dual variables. The lagrangian equation of OP2 is

$$h(\lambda^{\min}, \lambda^{\max}, \mu) = \sup_{P_{i,k,t}\ge 0} L(P_{i,k,t}, \lambda^{\min}, \lambda^{\max}, \mu). \tag{31}$$

The dual problem of OP2 is

$$\min_{(\lambda^{min}, \lambda^{max}, \mu)\ge 0} h(\lambda^{\min}, \lambda^{\max}, \mu). \tag{32}$$

$L(P_{i,k,t}, \lambda^{\min}, \lambda^{\max}, \mu)$ is a convex function when $P_{i,k,t} \ge 0$. According to the KKT conditions, we have

$$\sum_{i'=1}^{I} P_{i',k,t}G_{i',j_{i,k},t} = \frac{\left(1+\lambda_{i,k,t}^{\min}-\lambda_{i,k,t}^{\max}\right)BG_{i,j_{i,k},t}\log_2 e}{\mu_{i,t}} - \sigma^2. \tag{33}$$

Letting

$$c_{i,k,t} = \frac{\left(1+\lambda_{i,k,t}^{\min}-\lambda_{i,k,t}^{\max}\right)BG_{i,j_{i,k},t}\log_2 e}{\mu_{i,t}} - \sigma^2, \tag{34}$$

and

$$\begin{aligned} \boldsymbol{C}_{k,t} &= \begin{bmatrix} c_{1,k,t} & c_{2,k,t} & \cdots & c_{I,k,t} \end{bmatrix} \\ \boldsymbol{P}_{k,t} &= \begin{bmatrix} P_{1,k,t} & P_{2,k,t} & \cdots & P_{I,k,t} \end{bmatrix} \\ \boldsymbol{G}_{k,t} &= \begin{bmatrix} G_{1,j_{1,k},t} & G_{1,j_{2,k},t} & \cdots & G_{1,j_{I,k},t} \\ G_{2,j_{1,k},t} & G_{2,j_{2,k},t} & \cdots & G_{2,j_{I,k},t} \\ \vdots & \vdots & & \vdots \\ G_{I,j_{1,k},t} & G_{I,j_{2,k},t} & \cdots & G_{I,j_{I,k},t} \end{bmatrix}, \end{aligned} \tag{35}$$

we have

$$\boldsymbol{P}_{k,t} = \boldsymbol{C}_{k,t}\boldsymbol{G}_{k,t}^{-1} \forall k, t, \tag{36}$$

where $\boldsymbol{G}_{k,t}^{-1}$ is the inverse matrix of $\boldsymbol{G}_{k,t}$, expressed as

$$\boldsymbol{G}_{k,t}^{-1} = \begin{bmatrix} g_{1,1}^{k,t} & g_{1,2}^{k,t} & \cdots & g_{1,I}^{k,t} \\ g_{2,1}^{k,t} & g_{2,2}^{k,t} & \cdots & g_{2,I}^{k,t} \\ \vdots & \vdots & & \vdots \\ g_{I,1}^{k,t} & g_{I,2}^{k,t} & \cdots & g_{I,I}^{k,t} \end{bmatrix}. \tag{37}$$

Thus, the optimal transmission power is

$$P_{i,k,t}^* = \sum_{i'=1}^{I} c_{i',k,t} g_{i',i}^{k,t} \forall i, k, t. \tag{38}$$

And the Lagrangian equation is expressed as

$$\begin{aligned} & h(\lambda^{\min}, \lambda^{\max}, \mu) \\ & = \sup_{P_{i,k,t}\ge 0} L(P_{i,k,t}, \lambda^{\min}, \lambda^{\max}, \mu) \\ & = \sup_{P_{i,k,t}\ge 0} \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{i,k,t}^{\min}\left(R_{j_{i,k},t}^* - R_{\min}^{j_{i,k},t}\right) \\ & \quad + \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{i,k,t}^{\max}\left(R_{\max}^{j_{i,k},t} - R_{j_{i,k},t}^*\right) \\ & \quad + \sum_{i=1}^{I}\sum_{k=1}^{K} R_{j_{i,k},t}^* \\ & \quad + \sum_{i=1}^{I} \mu_{i,t}\left(P_{\max}^{i} - \sum_{k=1}^{K} P_{i,k,t}^*\right). \end{aligned} \tag{39}$$

$h(\lambda^{\min}, \lambda^{\max}, \mu)$ is a convex function, but its gradient may not exist at some points. We use subgradient method [34] to get the optimal results. Let

$$\begin{aligned}
\lambda^{\min}_{i,k,t}(l+1) &= \left[\lambda^{\min}_{i,k,t}(l) - \beta_1^{i,k,t}(l)\left(R_{j_{i,k},t} - R^{j_{i,k},t}_{\min}\right)\right]^+ \\
\lambda^{\max}_{i,k,t}(l+1) &= \left[\lambda^{\max}_{i,k,t}(l) - \beta_2^{i,k,t}(l)\left(R^{j_{i,k},t}_{\max} - R_{j_{i,k},t}\right)\right]^+ \\
\mu_{i,t}(l+1) &= \left[\mu_{i,t}(l) - \beta_3^{i,t}(l)\left(P^i_{\max} - \sum_{k=1}^{K} P_{i,k,t}\right)\right]^+, \qquad (40)
\end{aligned}$$

where $(x)^+ \equiv \max\{0, x\}$, $l \in \{1, 2, \ldots, l_{\max}\}$ is the iteration. $\beta^{i,k,t}_{m=1,2,3}(l) = \beta^{i,k,t}_{m=1,2,3}(1)/l$ is the step length.

B. PROPOSED SOLUTION TECHNIQUES OF OP4

OP4 is a convex optimization problem and we use the KKT conditions to solve it. The Lagrangian function of OP4 is

$$\begin{aligned}
&L(P_{k^h,t}, \lambda^{\min}, \lambda^{\max}, \lambda) \\
&= \sum_{j^h=1}^{J^h} R_{j^h,t} + \sum_{j^h=1}^{J^h} \lambda^{\min}_{j^h,t}\left(R_{j^h,t} - R^{j^h,t}_{\min}\right) \\
&\quad + \sum_{j^h=1}^{J^h} \lambda^{\max}_{j^h,t}\left(R^{j^h,t}_{\max} - R_{j^h,t}\right) \\
&\quad + \lambda_t\left(P^h_{\max} - \sum_{k^h=1}^{K^h} P_{k^h,t}\right), \qquad (41)
\end{aligned}$$

where $(\lambda^{\min}, \lambda^{\max}) = \{(\lambda^{\min}_{j^h,t}, \lambda^{\max}_{j^h,t}), j^h \in J^h\} \geq 0$ and $\lambda_t \geq 0$ are Lagrangian variables.

According to the KKT conditions, we have

$$\frac{\sum_{j^h=1}^{J^h} b_{j^h,k^h,t} B c_{j^h,k^h,t} \log_2 e\left(1 + \lambda^{\min}_{j^h,t} - \lambda^{\max}_{j^h,t}\right)}{1 + c_{j^h,k^h,t} P_{k^h,t}} - \lambda_t = 0, \qquad (42)$$

where

$$c_{j^h,k^h,t} = \frac{G_{h,j^h,t}}{\sum_{i=1}^{I} P_{i,k^h,t} \cdot G_{i,j^h,t} + \sigma^2}. \qquad (43)$$

Letting

$$\tau_{j^h,k^h,t} = b_{j^h,k^h,t} c_{j^h,k^h,t} \log_2 e, \qquad (44)$$

the optimal transmission power is expressed as

$$P^*_{k^h,t} = \frac{\left(1 + \lambda^{\min}_{j^h,t} - \lambda^{\max}_{j^h,t}\right) B \sum_{j^h=1}^{J^h} \tau_{j^h,k^h,t} - \lambda_t}{\lambda_t c_{j^h,k^h,t}}. \qquad (45)$$

The lagrangian equation is expressed as

$$\begin{aligned}
&h(\lambda^{\min}, \lambda^{\max}, \lambda) \\
&= \sup_{P_{k^h,t} \geq 0} L(P_{k^h,t}, \lambda^{\min}, \lambda^{\max}, \lambda) \\
&= \sup_{P_{k^h,t} \geq 0} \sum_{j^h=1}^{J^h} R^*_{j^h,t} + \sum_{j^h=1}^{J^h} \lambda^{\min}_{j^h,t}\left(R^*_{j^h,t} - R^{j^h,t}_{\min}\right) \\
&\quad + \sum_{j^h=1}^{J^h} \lambda^{\max}_{j^h,t}\left(R^{j^h,t}_{\max} - R^*_{j^h,t}\right) \\
&\quad + \lambda_t\left(P^h_{\max} - \sum_{k^h=1}^{K^h} P^*_{k^h,t}\right). \qquad (46)
\end{aligned}$$

The dual problem of OP4 is

$$\min_{(\lambda^{\min}, \lambda^{\max}, \lambda) \geq 0} h(\lambda^{\min}, \lambda^{\max}, \lambda). \qquad (47)$$

The dual problem is a convex optimization problem and we use gradient descent method to solve it. The partial derivatives are

$$\begin{aligned}
\frac{\partial h}{\partial \lambda^{\min}_{j^h,t}} &= \frac{B^2 \sum_{j^h=1}^{J^h} \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \tau_{j^h,k^h,t}}{\lambda_t \lambda^{\min}_{j^h,t} \ln 2}\left(1 - \lambda^{\max}_{j^h,t}\right) \\
&\quad + \frac{B^2 \sum_{j^h=1}^{J^h} \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \tau_{j^h,k^h,t}}{\lambda_t \ln 2} + R^*_{j^h,t} \\
&\quad - \frac{\lambda^{\min}_{j^h,t} B \sum_{j^h=1}^{J^h} \tau_{j^h,k^h,t}}{c_{j^h,k^h,t}} - R^{j^h,t}_{\min}, \qquad (48) \\
\frac{\partial h}{\partial \lambda^{\max}_{j^h,t}} &= \frac{B^2 \sum_{j^h=1}^{J^h} \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \tau_{j^h,k^h,t}}{\lambda_t \lambda^{\max}_{j^h,t} \ln 2}\left(\lambda^{\min}_{j^h,t} - 1\right) \\
&\quad + \frac{B^2 \sum_{j^h=1}^{J^h} \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \tau_{j^h,k^h,t}}{\lambda_t \ln 2} - R^*_{j^h,t} \\
&\quad + \frac{\lambda^{\max}_{j^h,t} B \sum_{j^h=1}^{J^h} \tau_{j^h,k^h,t}}{c_{j^h,k^h,t}} + R^{j^h,t}_{\max} \qquad (49)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial h}{\partial \lambda_t} &= P^h_{\max} - \frac{K^h}{c_{j^h,k^h,t}} \\
&\quad - \frac{\left(1 + \lambda^{\min}_{j^h,t} - \lambda^{\max}_{j^h,t}\right)^2 B^2 \sum_{j^h=1}^{J^h} \sum_{k^h=1}^{K^h} b_{j^h,k^h,t} \tau_{j^h,k^h,t}}{\lambda_t \ln 2}. \qquad (50)
\end{aligned}$$

Thus, we get

$$\begin{aligned}
\lambda^{\min}_{j^h,t}(l) &= \left[\lambda^{\min}_{j^h,t}(l-1) - \beta_1^{j^h,t}(l) \frac{\partial h}{\partial \lambda^{\min}_{j^h,t}}(l-1)\right]^+ \\
\lambda^{\max}_{j^h,t}(l) &= \left[\lambda^{\max}_{j^h,t}(l-1) - \beta_2^{j^h,t}(l) \frac{\partial h}{\partial \lambda^{\max}_{j^h,t}}(l-1)\right]^+ \\
\lambda_t(l) &= \left[\lambda_t(l-1) - \beta_3^{j^h,t}(l) \frac{\partial h}{\partial \lambda_t}(l-1)\right]^+, \qquad (51)
\end{aligned}$$

where l ($l \in \{1, 2, \ldots, l_{\max}\}$) is the iteration. $\beta^{j^h,t}_{m=1,2,3}(l) = \beta^{j^h,t}_{m=1,2,3}(1)/l$ is the the step length.

C. PROPOSED ALGORITHMS

To get the results of the above optimization problems, we design corresponding algorithms as follows. A subgradient algorithm is proposed for OP2, shown in Algorithm 1 and run in the HRM of LEO satellites. The complexity of the proposed algorithm is $O(LIK)$, where L is the maximum iterations in the proposed algorithms. The proposed algorithm for OP3 is shown in Algorithm 2 and run in the HRM of HAP. The complexity of the proposed algorithm is $O(\min\{J^h J^h K^h, J^h K^h K^h\})$. A gradient algorithm is proposed for OP4, shown in Algorithm 3 and run in the HRM of HAP. The complexity of the proposed algorithm is $O(LJ^h)$. The complexities of all the implement algorithms are linear, making it realistic to use the proposed methods in this paper at the future LEO-HAP systems.

Algorithm 1 Algorithm for OP 2

Initialize l_{max}, threshold ε and step length $\beta^{i,k,t}_{m=1,2,3}(1)$.
Initialize Lagrangian $\lambda^{\min}_{i,k,t}(1)$, $\lambda^{\max}_{i,k,t}(1)$ and $\mu_{i,t}(1)$.
Initialize $l = 1$, $n = 0$ and $h(0)$.
while $n = 0$ **do**
  for $i = 1 : I$ **do**
    for $i = 1 : K$ **do**
      Update $P^*_{i,k,t}(l)$
    end for
  end for
  Calculate $h(l)$.
  if $|h(l) - h(l-1)| \geq \varepsilon$ and $l \leq l_{\max}$ **then**
    for $i = 1 : I$ **do**
      for $i = 1 : K$ **do**
        Update $\beta^{i,k,t}_{m=1,2,3}(l+1)$, Lagrangian $\lambda^{\min}_{i,k,t}(1+1)$, $\lambda^{\max}_{i,k,t}(1+1)$ and $\mu_{i,t}(1+1)$.
      end for
    end for
    $l = l + 1$.
  else
    $n = 1$
  end if
end while
Output optimal power $P^*_{i,k,t}$.

Algorithm 2 Algorithm for OP 3

Initialize matrix $\boldsymbol{B}_t$, calculate matrix $\boldsymbol{R}^{\max}_t$.
Step 1: Find out the maximum element $R^{\max}_{j^h_1,k^h_1,t}$ of matrix $\boldsymbol{R}^{\max}_t$. Let $b_{j^h_1,k^h_1,t} = 1$. Delete the rows and columns that includes $R^{\max}_{j^h_1,k^h_1,t}$ and get the new matrix $\boldsymbol{R}'^{\max}_t$.
Step 2: Repeat step 1 until matrix $\boldsymbol{R}^{\max}_t$ is deleted completely.
Output optimal matrix $\boldsymbol{B}_t$.

Algorithm 3 Algorithm for OP 4

Initialize l_{max}, threshold ε and step length $\beta(1)$.
Initialize Lagrangian $\lambda^{\min}_{j^h,t}(1)$, $\lambda^{\max}_{j^h,t}(1)$ and $\lambda_t(1)$.
Initialize $l = 1$, $n = 0$ and $h(0)$.
while $n = 0$ **do**
  for $j^h = 1 : J^h$ **do**
    Update $P^*_{k^h,t}(l)$
  end for
  Calculate $h(l)$.
  if $|h(l) - h(l-1)| \geq \varepsilon$ and $l \leq l_{\max}$ **then**
    for $j^h = 1 : J^h$ **do**
      Update $\beta(l+1)$, Lagrangian $\lambda^{\min}_{i,k,t}(1+1)$, $\lambda^{\max}_{i,k,t}(1+1)$ and $\mu_{i,t}(1+1)$.
    end for
    $l = l + 1$.
  else
    $n = 1$
  end if
end while
Output optimal power $P^*_{k^h,t}$.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we implement simulations to calculate the capacity improvement of the presented multi-layer architecture compared with traditional LEO-MSS. The instantaneous capacity of this system cyclically changes due to quick cyclic movement of LEO satellites. In our simulations, we calculate the average capacity per cycle as an overall performance of the proposed architecture.

Then we implement simulations to compare the average throughputs of the proposed forecast-based multi-beam joint dynamic radio resource optimization method with different parameters and services in LEO satellites and HAPs. The simulations are implemented by MATLAB R2017a and the simulation results will be presented and analyzed in the following.

A. SYSTEM PARAMETERS

The default system parameters are shown in Table 3.

To validate the capacity improvement of the presented multi-layer architecture, we set the normalized antenna receiving gain of the TR $G_{\text{reTR}} = 3dB$ and that of MTs $G_{\text{reMT}} = 0dB$. The center of hot-spot regions follows a uniform distribution. The average radius of small hot-spot regions and large hot-spot regions are respectively 2km and 20km. We compare downlink capacity when N is 3 and 4.

To validate the proposed optimization method, we consider a rectangular region covered by 24 LEO beams and one HAP beam for simplicity. $N = 4$. The bandwith of each RB $B = 6MHz$ and $K = 10$. The length of one timeslot $t_s = 10\ ms$. $P_{\max} = 3dBW$ and $P^i_{\max} = 10dBW$. $P'_{\max} = -20dBW$ and $P^h_{\max} = -10dBW$. $\omega_{j_i} = \omega_d = \omega_g = 1$ and $p_g = 1$. $\omega_0 = 0.9$ and $\omega_1 = 0.1$. $\Delta t = 100ms$. Packets arrive every $100ms$ on average and the packet sizes are normally distributed with

TABLE 3. Simulation parameters.

Parameters	Value
Altitude of LEO Satellite	*1500 km*
$V_{\mathbf{LEO}}$	*25000 km/h*
LEO Satellite Antenna η	*20*
LEO Beam Coverage Radius of 3 dB	*200 km*
Altitude of HAP	*20-100 km*
HAP Antenna η	*1*
TR Antenna η	*20*
Receiving Gain of TR $G_{\mathbf{reTR}}$	*3dB*
Receiving Gain of MT $G_{\mathbf{reMT}}$	*0dB*
Noise Power Spectral Density	*-173dBm/Hz*
Average Radius of Small Hot-Spot Regions	*2 km*
Average Radius of Large Hot-Spot Regions	*20 km*
Carrier Frequency of Ka Band	*30 GHz*
$B_{\mathbf{total}}$	*240 MHz*
Bandwith of Each RB B	*6 MHz*
Timeslot t_s	*10 ms*

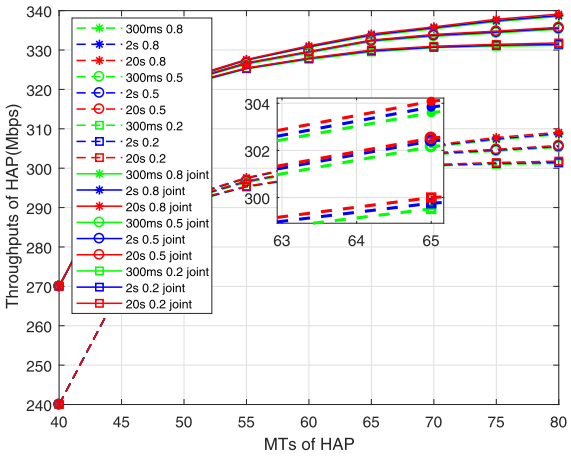


FIGURE 10. Average throughput of HAP.

the means of 500k bits. The maximum delay tolerances are 300ms, 2s, and 20s respectively. We set ε_d =0.8, 0.5 and 0.2 to find the effect of the threshold of delay utility function.

B. SIMULATION RESULTS AND DISCUSSION

We do 100 Monte-carlo simulations and each lasts 2 minutes and contains 12000 timeslots. Finally we get the following average results in Fig. 10 and Fig. 11.

Fig. 10 and Fig. 11 demonstrate that the proposed multi-beam joint power optimization method improves the average throughputs of both HAP and LEO beams greatly at more than 10%. The proposed dynamic strategy has a better improvement on throughputs in LEO satellites when the maximum delay tolerances of the target services is longer because the packets can be transmitted with larger channel gains. Besides, the average throughputs of LEO beams have a larger upper limit and rising trend when ε_d increases. However, the maximum delay tolerances of the target services have no

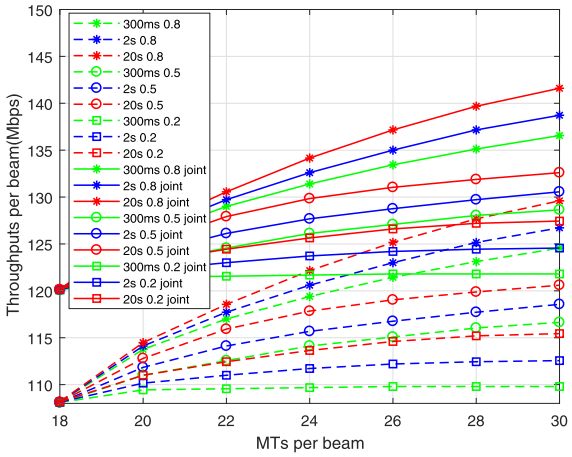


FIGURE 11. Average throughput per beam.

impact on the average throughputs of HAP. And the impact of different ε_d is much less in HAP compared with in LEO satellites. This is because the channel gain of HAP-ground links does not change periodically.

Summarily, HAP provides more than two times average throughputs than LEO satellite beam. The proposed multi-beam power optimization method has a significant improvement on the average throughputs and can be implemented easily based on the proposed framework. The proposed dynamic time-slot allocation strategy also has an improvement on the average throughputs because it makes full use of the predicted information and the maximum delay tolerances of different service types. To maximize the average throughputs, ε_d should be set close to 1 in the proposed dynamic strategy.

VI. CONCLUSION AND FUTURE DIRECTION

In this paper, we focused on improving the network capacity of the regions beyond the coverage of terrestrial communication systems. As LEO-MSS can not provide enough capacity especially in hot-spot regions, we proposed an extended multi-layer network architecture based on multi-beam LEO-MSS by introducing HAPs and TRs to cover hot-spot regions, and analyzed the capacity improvement with different parameters. To further increase the throughputs, we proposed a multi-beam joint dynamic radio resource optimization method in LEO-ground downlinks using the predicted information according to the movement of LEO satellites based on the proposed efficient HRM framework and delay management scheme. Also, we proposed the dynamic resource optimization method of HAP-ground downlinks in hot-spot regions when LEO satellites and HAPs share the same spectrum. To solve these problems, the Lagrange dual method and Karush-Kuhn-Tucker (KKT) conditions is used to find the optimal solutions. Afterwards, corresponding gradient descent algorithms are proposed for each optimization problems to get the results. Numerical results show that the proposed architecture yields a significant improvement on

capacity, especially in hot-spot regions. Moreover, the proposed optimization methods have increased the throughputs of LEO-ground downlinks and HAP-ground downlinks with an acceptable algorithm complexity.

The proposed extended multi-layer architecture and the dynamic resource optimization methods in this paper can be used in future space-air-ground communication network systems to provide better communication services for the regions beyond the coverage of terrestrial communication systems. The quick movement of LEO satellites and the propagation delay of LEO-ground links bring not only great challenges to radio resource management, but also difficulties in handover management. Thus, improved handover strategies should be introduced in future works.

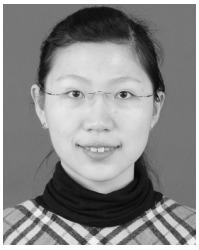
ACKNOWLEDGMENT

This article was presented in part at the 2019 IEEE 90th Vehicular Technology Conference (VTC2019-FALL).

REFERENCES

[1] C. Sacchi, K. Bhasin, N. Kadowaki, and F. Vong, "Technologies and applications of future satellite networking [guest editorial]," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 154–155, May 2015.
[2] T. Pecorella, L. S. Ronga, F. Chiti, S. Jayousi, and L. Franck, "Emergency satellite communications: Research and standardization activities," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 170–177, May 2015.
[3] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, May 2018.
[4] Z. Qu, G. Zhang, H. Cao, and J. Xie, "LEO satellite constellation for Internet of Things," *IEEE Access*, vol. 5, pp. 18391–18401, 2017.
[5] B. Feng, H. Zhou, H. Zhang, G. Li, H. Li, S. Yu, and H.-C. Chao, "HetNet: A flexible architecture for heterogeneous satellite-terrestrial networks," *IEEE Netw.*, vol. 31, no. 6, pp. 86–92, Nov. 2017.
[6] Y. Wang, Y. Xu, Y. Zhang, and P. Zhang, "Hybrid satellite-aerial-terrestrial networks in emergency scenarios: A survey," *China Commun.*, vol. 14, no. 7, pp. 1–13, Jul. 2017.
[7] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 101–109, Jul. 2017.
[8] J. Zhang, X. Zhang, M. A. Imran, B. Evans, Y. Zhang, and W. Wang, "Energy efficient hybrid satellite terrestrial 5G networks with software defined features," *J. Commun. Netw.*, vol. 19, no. 2, pp. 147–161, Apr. 2017.
[9] Z. Ji, Y. Wang, W. Feng, and J. Lu, "Delay-aware power and bandwidth allocation for multiuser satellite downlinks," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 1951–1954, Nov. 2014.
[10] M. Jia, X. Zhang, X. Gu, Q. Guo, Y. Li, and P. Lin, "Interbeam interference constrained resource allocation for shared spectrum multibeam satellite communication systems," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6052–6059, Aug. 2019.
[11] A. J. Roumeliotis, C. I. Kourogiorgas, and A. D. Panagopoulos, "Optimal capacity allocation strategies in smart gateway satellite systems," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 56–59, Jan. 2019.
[12] F. Li, K.-Y. Lam, X. Liu, J. Wang, K. Zhao, and L. Wang, "Joint pricing and power allocation for multibeam satellite systems with dynamic game model," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2398–2408, Mar. 2018.
[13] W. Feng, N. Ge, and J. Lu, "Coordinated satellite-terrestrial networks: A robust spectrum sharing perspective," in *Proc. 26th Wireless Opt. Commun. Conf. (WOCC)*, Apr. 2017, pp. 1–5.
[14] C. Liu, W. Feng, Y. Chen, C.-X. Wang, and N. Ge, "Optimal beamforming for hybrid satellite terrestrial networks with nonlinear PA and imperfect CSIT," *IEEE Wireless Commun. Lett.*, to be published.
[15] C. Wang, D. Bian, S. Shi, J. Xu, and G. Zhang, "A novel cognitive satellite network with GEO and LEO broadband systems in the downlink case," *IEEE Access*, vol. 6, pp. 25987–26000, 2018.
[16] A. Mohammed, A. Mehmood, F.-N. Pavlidou, and M. Mohorcic, "The role of high-altitude platforms (HAPs) in the global wireless connectivity," *Proc. IEEE*, vol. 99, no. 11, pp. 1939–1953, Nov. 2011.
[17] N. Vaiopoulos, H. G. Sandalidis, and D. Varoutas, "Using a HAP network to transfer WiMAX OFDM signals: Outage probability analysis," *J. Opt. Commun. Netw.*, vol. 5, no. 7, p. 711, Jul. 2013.
[18] X. Luo, S. Li, and H. Xu, "Results of real-time kinematic positioning based on real GPS L5 data," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 8, pp. 1193–1197, Aug. 2016.
[19] M. Kirkko-Jaakkola, J. Parviainen, J. Collin, and J. Takala, "Improving TTFF by two-satellite GNSS positioning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3660–3670, Oct. 2012.
[20] A. Shahriar, M. Atiquzzaman, and S. Rahman, "Mobility management protocols for next-generation all-IP satellite networks," *IEEE Wireless Commun.*, vol. 15, no. 2, pp. 46–54, Apr. 2008.
[21] X. Wang, "Deployment of high altitude platforms in heterogeneous wireless sensor network via MRF-MAP and potential games," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 1446–1451.
[22] X. Wang, X. Gao, and R. Zong, "Energy-efficient deployment of airships for high altitude platforms: A deterministic annealing approach," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–6.
[23] Y. Yang, R. Lu, T. Manzaneque, and S. Gong, "Toward Ka band acoustics: Lithium niobate asymmetrical mode piezoelectric MEMS resonators," in *Proc. IEEE Int. Freq. Control Symp. (IFCS)*, May 2018, pp. 1–5.
[24] Y. Zeng, G. Xiao, and L. Qiu, "Ka-band satellite communication antenna dome using dual-band frequency selective surfaces," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2019, pp. 1–3.
[25] M. W. Atwood, G. N. Marcoux, and W. P. Craig, "Demonstration of two-way extremely high frequency (EHF) satellite communication (SATCOM) using submarine-survivable phased arrays," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2008, pp. 1–7.
[26] E. Cianca, T. Rossi, A. Yahalom, Y. Pinhasi, J. Farserotu, and C. Sacchi, "EHF for satellite communications: The new broadband frontier," *Proc. IEEE*, vol. 99, no. 11, pp. 1858–1881, Nov. 2011.
[27] Z. Sodnik, B. Furch, and H. Lutz, "Free-space laser communication activities in Europe: SILEX and beyond," in *Proc. 19th Annu. Meeting IEEE Lasers Electro-Opt. Soc. (LEOS)*, Oct. 2006, pp. 78–79.
[28] A. Ibrahim and A. S. Alfa, "Using Lagrangian relaxation for radio resource allocation in high altitude platforms," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5823–5835, Oct. 2015.
[29] S. Wang, Y. Li, Q. Wang, M. Su, and W. Zhou, "Dynamic downlink resource allocation based on imperfect estimation in LEO-HAP cognitive system," in *Proc. 11th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2019, pp. 1–6.
[30] W. Wang, A. Liu, Q. Zhang, L. You, X. Gao, and G. Zheng, "Robust multigroup multicast transmission for frame-based multi-beam satellite systems," *IEEE Access*, vol. 6, pp. 46074–46083, 2018.
[31] L. Li, N. Deng, W. Ren, B. Kou, W. Zhou, and S. Yu, "Multi-service resource allocation in future network with wireless virtualization," *IEEE Access*, vol. 6, pp. 53854–53868, 2018.
[32] Y. Albagory, M. Nofal, and A. Ghoneim, "Handover performance of unstable-yaw stratospheric high-altitude stations," *Wireless Pers. Commun.*, vol. 84, no. 4, pp. 2651–2663, Oct. 2015.
[33] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
[34] C. Zhang, L. Zheng, Z. Zhang, L. Shi, and A. J. Armstrong, "The allocation of berths and quay cranes by using a sub-gradient optimization technique," *Comput. Ind. Eng.*, vol. 58, no. 1, pp. 40–50, Feb. 2010.

YITAO LI received the B.S. degree in electronic engineering from the Department of Electronic Engineering and Information Science, University of Science and Technology of China (USTC), Hefei, China, in 2014, where he is currently pursuing the Ph.D. degree. His research interests include low earth orbit satellite mobile networks, geostationary earth orbit satellite relay networks, space-air-ground integrated networks, 5G networks, radio resource management, handover management, and RRC and MAC protocols of wireless access networks.



NA DENG (Member, IEEE) received the B.S. and Ph.D. degrees in information and communication engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2010 and 2015, respectively. From 2013 to 2014, she was a Visiting Student with the Prof. Martin Haenggi's Group, University of Notre Dame, Notre Dame, IN, USA. From 2015 to 2016, she was a Senior Engineer with Huawei Technologies Co., Ltd., Shanghai, China. She is currently an Associate Professor with the Dalian University of Technology, Dalian, China. Her scientific interests include networking and wireless communications, green communications, and network design based on wireless big data.

WUYANG ZHOU (Member, IEEE) received the B.S. and M.S. degrees from Xidian University, Xi'an, China, in 1993 and 1996, respectively, and the Ph.D. degree from the University of Science and Technology of China (USTC), Hefei, China, in 2000. He is currently a Professor with the Department of Electronic Engineering and Information Science, USTC. His research interests include space-air-ground integrated networks, 5G networks, interference cancellation in satellite networks, intelligent computing, and positioning technologies.