Inter-networking and Function Optimization for Mega-Constellations

Zitong Lin, Hewu Li, Jun Liu, Zeqi Lai, Guojie Fan
Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing, China
Beijing National Research Center for Information Science and Technology (BNRist), Beijing, China
Email:{linzt19,fgj20}@mails.tsinghua.edu.cn; lihewu@cernet.edu.cn;{zeqilai,juneliu}@tsinghua.edu.cn

***Abstract*—Emerging Internet mega-constellations promise to provide broadband network service globally. Typically these constellations follow a *uniform* structure and functionality model, in which the constellation is evenly organized at scale, and each satellite equally serves for both access and forwarding tasks. However, due to the *inherently non-uniform distribution* of terrestrial user demands, the uniform constellation design not only suffers from low network utilization but also involves significant operational costs for constellation operators.**

This paper studies a forward-looking yet important problem for future Internet mega-constellations: *how can we accomplish cost-effective constellation design,* i.e., *satisfying ubiquitous user demands while improving constellation-wide network utilization and optimizing operational cost?* To this end, we propose DIV, an inter-networking mechanism, which attains the cost-effective goal by dynamically adjusting the constellation-wide inter-satellite connectivity and per-satellite functionality based on spatial and temporal varying user demands. Specifically, DIV divides satellites into backbone and access according to the inherent difference between Internet local and inter-regional traffic patterns, and re-organizes the constellation structure to optimize network utilization on-demand. Extensive simulations based on realistic constellation information demonstrate that DIV can achieve cost-effectiveness for various state-of-the-art mega-constellations and increase the network efficiency by up to $2.5\times$ while guaranteeing stable, pervasive network services for geo-distributed terrestrial users.

***Index Terms*—mega-constellation, cost-effectiveness, inter-networking, function optimization**

I. INTRODUCTION

Although satellite networks have been developed for several decades, emerging low earth orbit (LEO) mega-constellation satellite network developed in recent years have undoubtedly brought new opportunities and challenges. The reduction in constellation altitude and deployment of high-speed inter-satellite links allow satellites to achieve lower latency than terrestrial fiber network in long-haul traffic [1]. The increase in the scale of satellites and the improvement of technology have allowed satellites to have higher bandwidth [2], [3]. Mega-constellations today can provide global coverage, low latency, high bandwidth internet services as compared to small constellations decades ago. However, the explosive scale makes the cost of the entire constellation to rise sharply, even though the cost of manufacturing and launching one satellite is decreasing [4]. Worse, the increase in scale and bandwidth leaves more satellites underutilized.

The inefficient use of satellites has a long history for low-Earth orbit (LEO) satellites, not only due to marketing issues [5], [6], but also due to the mobility and structure of satellites. The mobility of satellites drives the constellation to use a uniform structure [7]–[10] to ensure continuity and stability of user service. However, most areas upon the earth's surface have a sparse population (*e.g.,* oceans which cover about 70% of our planet). When applying a uniform structure to serve extremely uneven terrestrial user traffic, typically, only a small percentage of satellites above user areas can be used. This is already severe in the era of small constellations, and the emergence of mega-constellations makes a larger amount of satellites wasteful. Inter-satellite links (ISLs) allow traffic to pass through satellites even if they are not over users, allowing satellites to be used more efficiently. But ISLs also greatly increase the cost of satellite production and launching [11]–[14]. Therefore, some works are devoted to make constellations more cost-efficient. [15]–[18] design new constellations to meet the service demands with fewer satellites. [19]–[22] improve satellite service life to save the cost. New satellite connection schemes [23], [24] and load balancing schemes [25], [26] are also designed to improve the efficiency of ISLs.

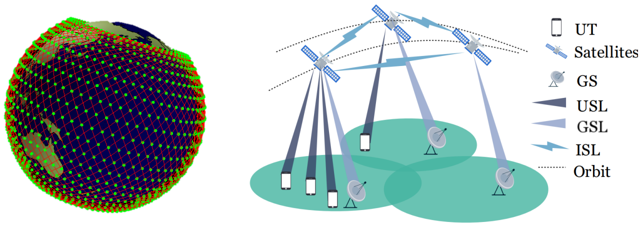
But there are still some critical issues that have not been considered in the existing work. First, these designs imply each satellite is provisioned with the same ISLs [25], [26], resulting in link redundancy. Second, many works do consider the traffic features for constellation design [17], [23], but they are not scalable and adaptive enough for the complex internet traffic. Third, some work simplifies the dynamics of satellites and transforms the problem into sub-problems under multiple snapshots to solve [19], which increases computational complexity and switching overhead. Different from these works, we mainly consider two problems:

◦ *Mobility makes each satellite have the same frequency to serve each area. How to make divergent provisioning in such a network where each satellite is "equivalent" and demands vary widely between regions?*

◦ *Traffic has diverse features and demands, and it changes with spatial and temporal change and constellation development. How to provide a structure that can be adaptive and*

This work is supported by National Key Research and Development Plan of China under Grant 2018YFB1800301, National Natural Science Foundation of China (61902214, 62132009), Tsinghua University-China Mobile Communications Group Co.,Ltd. Joint Institute, and Zhongguancun Laboratory.

ISBN 978-3-903176-48-5 © 2022 IFIP



(a) Orbital model. (b) Communication Scenes.

Fig. 1: Mega-constellation satellite networks.

scalable while ensuring cost-effectiveness?

To address the above issues, this paper conducts research in the following ways.

- Formulating the problem of cost-efficient constellation design and analyzing the complexity of the problem as well as the essential reasons for satellite inefficiency.
- Classifying traffic according to the divergent characteristics of local traffic and inter-regional traffic behaviors on the internet and dividing constellation roles accordingly.
- Combining the constellation structure, operation laws, and traffic characteristics to divide the roles of satellites, and designing the network topology of each role.
- Conducting simulations based on realistic constellation information to verify the effectiveness of DIV under various experimental configurations.

II. COST-EFFECTIVENESS OF CONSTELLATIONS

In this section, we introduce mega-constellation networks including orbit model and link model, and formulate the cost-effectiveness of mega-constellations. As shown in Fig. 1(b), the main components of the network include satellites, user terminals (UTs) and ground stations (GSes). GSes are connected to the terrestrial internet. UTs and GSes are both transmitters and receivers of the network. They communicate with satellites via user-satellite links (USLs) and ground-satellite links (GSLs) respectively. Satellites communicate with each other by ISLs.

A. Orbital Model

There are various types of satellite constellations, such as Polar [27], Walker [28], Rider [29], Rosette [15], *etc.*. The walker (generally refers to the walker-δ) constellation has become a trend. A walker constellation can be represented by $N/P/F : h : i$, where h and i represent altitude and inclination, N is the total number of satellites and P is the number of orbits. F represents the relative phase difference between satellites of adjacent orbits passing through the equatorial plane. It is the relative phase measured by $1/P$ of the interval between adjacent satellites in any orbit, and $0 \leq F \leq S-1$.

For a satellite numbered m in the constellation, the right ascension of ascending node Ω_m and mean anomaly from perigee at epoch u_m can be calculated as follows:

$$\begin{cases} \Omega_m = \frac{2\pi}{P}(P_m-1), P_m = 1,2,...,P; \\ u_m = \frac{2\pi}{S}(N_m-1) + \frac{2\pi}{F}(P_m-1), N_j = 1,2,...,S-1. \end{cases}$$

where S is the number of satellite in an orbit plane, P_m is the orbit number, and N_m is the number in orbit. That is, $S = N/P, P_m = m/S+1, N_m = m-(P_m-1)S$. Several mega-constellations currently deployed consist of multiple *shells*, and each shell is a walker constellation. Fig. 1(a) shown a shell of walker constellation.

B. Link and Traffic Model

The satellites and ISLs in the constellations are expressed as $G\langle V, E\rangle$. The satellite i is represented by v_i,$v_i \in V$. l_{ij},$l_{ij} \in E$ represents the ISL between nodes v_i and v_j. d_i represents the degree of node v_i. That is, satellite i has d_i ISLs.

Due to physical limitations, each satellite has a limit on the maximum number of space communication terminals, so each node has a maximum degree d_i^{max}. Not any two satellites can establish connections because some satellites are blocked by the earth and cannot see each other. For a satellite with an orbital height of h, the maximum link length is D_h.

The traffic carried by the constellation can be represented by the traffic matrix $\mathcal{F}$, and the total amount of the traffic matrix is $|\mathcal{F}|$. For any flow $f_{sd} \in \mathcal{F}$, its bandwidth is r_{sd}. The start and end nodes of this flow are v_s, v_d, respectively. The satellite network transfers traffic with a certain routing algorithm. For each link l_{ij}, its capacity is C_{ij}, of which u_{ij} bandwidth is occupied. Find the link with the highest occupancy ratio $\phi = u_{ij}/C_{ij}$. Then, with the same routing algorithm, the maximum throughput that the entire constellation can carry is $|\mathcal{F}|/\phi$. In fact, a congestion control algorithm can improve the maximum throughput, but here we design from a topology perspective, so we consider the worst-case throughput.

C. Formulate the Cost-Effectiveness

Cost-efficient constellations expect *higher throughput* at *lower cost*.

For cost, a typical communication satellite contains 4 to 5 space communication terminals (SCTs) [30]. However, even the latest laser communication terminals weigh 15-100kg for a single terminal [11]–[14]. At present, a single satellite without SCTs is only a few hundred kilograms($260kg$ for Starlink v1.0, $150kg$ for OneWeb). The addition of SCTs is a huge burden on communication satellites. In the discussion that follows, we use the number of SCTs $\mathcal{N}$ as a proxy for cost analysis.

Throughput is not so easy to quantify. A variety of reasons affect it, including traffic distribution, access control, topology, and routing. In this paper, we use the average hop count $\mathcal{H}$ as a proxy for throughput as [23]. This is justified because, for traffic of the same scale, the fewer hops per flow, the less ISL capacity is occupied, and the more likely the throughput will be improved. Certainly, if the load is extremely uneven, the actual throughput will drop even the capacity usage is small. Therefore, we will discuss the link use differences in § IV-E.

In summary, the cost-effectiveness $\mathcal{U}$ can be denoted as:

$$\mathcal{U} = \frac{1}{\mathcal{N}\mathcal{H}} \tag{1}$$

The objective for us is to maximum $\mathcal{U}$, and the constraints are listed as follows.

1) Each satellite has a maximum of d_i^{max} ISLs.
2) Conservation of flows. The sum of each flow entering a satellite equals the sum of the flows exiting that satellite, except for the source and the sink.
3) For each flow, the sum of the flow exiting at the source satellite and entering at the sink satellite is the total bandwidth of this flow.
4) The total traffic of each link does not exceed capacity.

We use two parameters to formulate the problem. b_{ij} indicates whether there is an edge between nodes v_i and v_j. And w_{ij}^{sd} represents the amount of traffic passing through the edge l_{ij} from v_s to v_d. The objective is to maximum $\mathcal{U}$, which equates to minimized ***cost***, which is the product of the number of ISLs and hop counts, it can be formulated as:

$$min \sum_{l_{ij} \in V} b_{ij} \cdot \sum_{l_{ij} \in V, f_{sd} in \mathcal{F}} b_{ij} w_{ij}^{sd} \tag{2}$$

Subject to

$$\begin{cases} \sum_{i=1}^{N} b_{ij} \leq d_j^{max}, \forall v_j \in V \\ \sum_{i=1}^{N} b_{ij} w_{ij}^{sd} = \sum_{i=1}^{N} b_{ji} w_{ji}^{sd}, \forall f_{sd} \in \mathcal{F}, v_j \in V, j \neq s, d \\ \sum_{i=1}^{N} b_{si} w_{si}^{sd} = \sum_{i=1}^{N} b_{id} w_{id}^{sd} = r_{sd}, \forall f_{sd} \in \mathcal{F} \\ \sum_{f_{sd} \in \mathcal{F}} b_{ij} w_{ij}^{sd} \leq C_{ij}, \forall v_i, v_j \in V \\ b_{ij} \in 0, 1, \forall v_i, v_j \in V \end{cases}$$

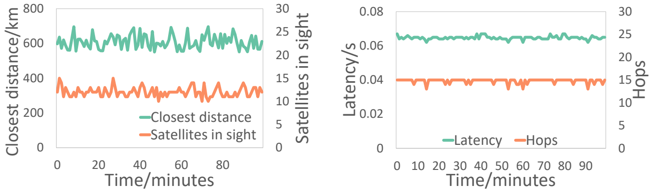
This is a simplification of the problem, which only considers a snapshot, but it is already NP-hard. We have also tried to solve this problem through some heuristic algorithms, such as genetic algorithm, greedy algorithm, random graph algorithm, *etc.*, but these solutions all have the following problems:

1) Strong dependence on flow matrix details, while the precise matrix is not available before deployment.
2) Computational complexity increases exponentially or even factorially with satellite and traffic matrix scale.
3) Extreme differences between different time slices, while the configuration of the satellites is onetime immutable.

The combination of the above three points is also one reason that hinders the cost-efficient design in the preliminary design of the constellation. Therefore, in mass studies [1], [23], [25], [26], [31], [32], each satellite is provisioned with the same number of ISLs by default and connected in the same or similar connection mode. We will analyze the advantages of this design, then give the essential reasons for the inefficiency with complex internet traffic.

D. Opportunities

A monotonous mode is usually used for connection in constellation. Typically, in *+Grid* [33], each satellite is connected to the adjacent satellites in the same orbit, and to the satellites of the same orbit number in the two adjacent orbits. Each satellite repeats the mode. This structure has extremely high stability in the continuous movement of satellites. As shown in Fig. 2(a), for a user in Beijing, the distance between the user and the nearest satellite is stable, and the number of satellites in sight always remains within a range. For two users in Beijing and New York, the latency and the hops are relatively stable as shown in Fig. 2(b). *Motif* [23] expands the definition of the modes and a satellite can connect with any two visible satellites, but each satellite has to ***repeat*** this mode. The expansion of the concept effectively improves the performance of the constellation. However, such monotonous modes show incompatibility when faced with complex internet traffic. As shown in Fig. 3, *+Grid* has very few hops when communication distance is not long. But as the distance increases, the hops also increases proportionally. *Motif* shows much lower hops than *+Grid* when facing long-haul communication, but it loses the performance of short-haul communication. Both of them are less cost-effective.



(a) Single User (b) User Pairs

Fig. 2: Stable service of Starlink phase I.

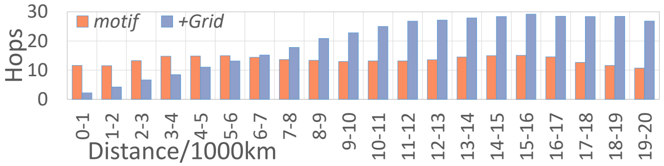


Fig. 3: Communication distance and hop count.

Mega-constellation aims to provide internet service to users. But even for terrestrial networks deployed on-demand, the diverse and time-varying demands of network traffic make network infrastructure less cost-effective, not to mention the evenly deployed satellite constellations that cannot be deployed on-demand due to their dynamics. But this does not mean that satellite constellations have no choice but to use evenly repeated inter-networking patterns and architectures to serve diverse network traffic. Terrestrial internet traffic shows that there are inherent divergences and features between local traffic behavior and inter-regional traffic behavior. *These characteristics do not change due to changes in location and time variation of traffic.* So even for satellite networks, they can be used as inherent properties to help the design of constellation structures. These divergent features prompt us to divide the functions of satellites. Through division, satellites with different functions can undertake different traffic in a more optimized mode. And these designs need to consider the constellation structure and operation laws of satellites, which we will design in the next section.

III. DIV: A COST-EFFECTIVENESS CONSTELLATION INTER-NETWORKING MECHANISM

DIV divides the satellites to backbone satellites and access satellites, as shown in Fig. 4. Users can access any kind of satellite. The focus of the transfer of the two satellites is different. The backbone satellites carry inter-regional traffic, while the access satellites undertake local traffic and inter-regional traffic generated locally. The two networks with

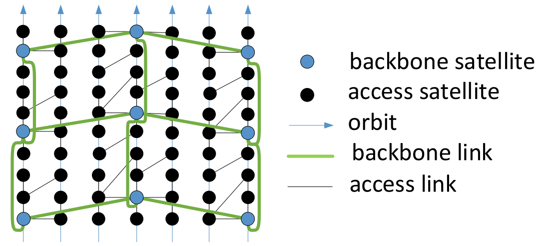


Fig. 4: Constellation structure that divides the roles.

different functions can be optimized according to the traffic features, respectively.

○ The communication distance of inter-regional traffic is long. And due to the terrestrial traffic geographical distribution features, inter-regional traffic has obvious aggregate properties in the direction and scale. To ensure the stability, the backbone satellite needs to be selected uniformly globally. But it can choose longer and more suitable direction links according to the aggregate properties.

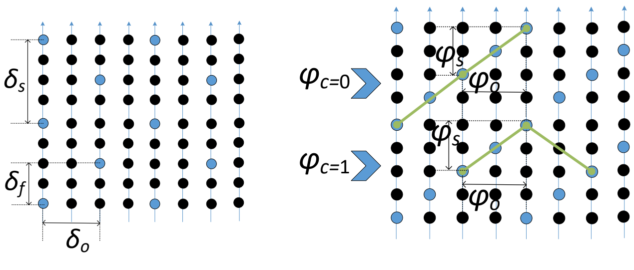
○ The inter-regional traffic generated locally needs to be forwarded to backbone satellites from access satellites, and the local traffic occurs under the same satellite or between several adjacent satellites. Therefore, the access satellites are grouped according to the nearest backbone satellites, and each group is connected according to the characteristics of local traffic.

The more link features matched traffic features, the fewer the hops and the SCTs are required, as in Equ. 1, the more cost-effective the constellation.

A. Backbone

The goal of backbone design is to carry inter-regional traffic with ***fewer satellites***. We select nodes as backbone nodes every few orbits and every few satellites in each orbit. For the selected backbone node set, each node connects others at the same angle and distance. Since there are many possibilities for node selection and connection, we introduce several parameters to describe the methods. After listing all possibilities, we will evaluate each mode and find the most cost-effective one.

Backbone Selection. The selection is crucial to the design of connections and the grouping of access nodes. We introduce three parameters to describe the selection of the backbone nodes: orbit interval δ_o, intra-orbit interval δ_s, and inter-orbit offset δ_f. Its definition is shown in Fig. 5(a). There is a backbone orbit every δ_o orbits, and every δ_s satellite has one backbone node in each backbone orbit. Inter-orbit offset represents the intra-orbit number difference of backbone node for two adjacent backbone orbits is δ_f. δ_o and δ_s determine the number of backbone nodes, while δ_f determines the relative positions between backbone nodes. All satellite nodes form a satellite node set $V = \{v_0, v_1, v_2, ..., v_{N-1}\}$, and each backbone selection method can select a subset in V, we call the i-th backbone node set $\mathcal{B}_i = \{v_0, v_{\delta_s}, v_{2\delta_s}, ..., v_{S\delta_o+\delta_f}, v_{S\delta_o+\delta_f+\delta_s}, ...\}$. Different $\mathcal{B}_i$ has different numbers of backbone nodes. When the throughput and performance are similar, the fewer the number of nodes, the higher the cost-effectiveness of it.



(a) backbone selection (b) backbone connection

Fig. 5: Parameters of backbone design.

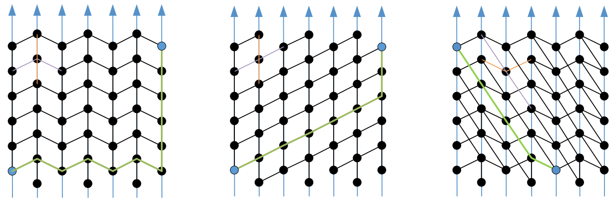
Backbone Connection. Each satellite has multiple ISLs, and we use three parameters to represent *a pair of* ISLs for one satellite: orbit difference φ_o, intra-orbit difference φ_s and symmetry φ_c as shown in 5(b). For the first ISL, the difference between the orbit numbers of two connected satellites is φ_o, and the difference of intra-orbit numbers is φ_s. Symmetry φ_c indicates the symmetry of another ISL and this ISL. If $\varphi_c = 1$, the other ISL is axisymmetric about the orbital axis, and the orbit and intra-orbit number of the two nodes differ by $(-\varphi_o, \varphi_s)$. If $\varphi_c = 0$, the other ISL is centrosymmetric the satellite with the first ISL, and it differs by $(-\varphi_o, -\varphi_s)$. We call the connection mode of each ISL pair a ***mode***, denoted by m. All mods form a *candidate mode set* $\mathcal{M} = \{m_i | m_i = (\varphi_{oi}, \varphi_{si}, \varphi_{ci})\}$. For the backbone node set $\mathcal{B}_i$, we select k (usually $k = 2$) modes in $\mathcal{M}$ to form the network. The modes selected should be valid. That is, in the backbone, the connection peer selected by each backbone node according to all modes should be another backbone node. Any valid *k-mode combinations* θ is added to the *candidate mode combination set* $\Theta_{\mathcal{B}_i} = \{\theta_i | \theta_i = (m_1, m_2, ...m_k), m_i \in \mathcal{M}\}$.

Different mode combination θ will have a large difference in hops. The access nodes are hidden in Fig. 6, and only the backbone nodes are shown. The only difference between θ_1 and θ_2 is the parameter φ_c, but for the same source node and sink node, the hop-count is 11 and 8 respectively. And θ_3 requires 12 hops. But for a special relative position of the source and sink, as shown in Fig. 6(c), θ_3 has a greater advantage. The length and direction of the links greatly affects the hop-count.

Backbone evaluation. For each backbone node set $\mathcal{B}_i$, it is possible to select a variety of candidate mode combinations. With a mode combination θ, we can get a backbone network. We can have inter-regional traffic transmit in the constellation with its specific routing algorithm (such as k-shortest-path). If it cannot satisfy the throughput demands T^{min}, it should be

Algorithm 1: Constellation design for backbone

```
Input: $\mathcal{F}, V, T^{min}$
Output: $\mathcal{B}_{re}, \theta_{re}$
Get mode set $\mathcal{M}$ from $V$; $maxNH = 0$;
for each backbone node set $\mathcal{B}_i$ do
    Get all valid modes $\mathcal{M}_{\mathcal{B}_i}$ from $\mathcal{M}$;
    for each k-mode combination $\theta_j$ in $\mathcal{M}_{\mathcal{B}_i}$ do
        Assess the $T$ and $\mathcal{NH}$ by $\mathcal{F}$;
        if $T > T^{min}$ and $\mathcal{NH} > maxNH$ then
            $\mathcal{B}_{re} = \mathcal{B}_i; \theta_{re} = \theta_j; maxNH = \mathcal{NH}$
```



(a) θ_1 (b) θ_2 (c) θ_3

Fig. 6: Differences in hop counts in different connection modes.

removed from the candidate set. After evaluation of all $\mathcal{B}_i$ and θ, we can choose a node set $\mathcal{B}_{re}$ and mode combination θ_{re} from all candidates with the least cost ($\mathcal{NH}$), so as to achieve high cost-effectiveness of the constellation.

Backbone design. We give an introduction to the algorithm Alg. 1 of the backbone design. The algorithm designs the backbone according to the satellite node set V, the inter-regional traffic model $\mathcal{F}$ and the minimum throughput demands T^{min}.

First, obtain the *mode set* $\mathcal{M}$ of all visible modes in the constellation according to the link length limit. Then we traverse all possible $(\delta_o, \delta_s, \delta_f)$ to generate backbone node sets $\mathcal{B}_i$. For each set, find all valid modes from $\mathcal{M}$ to form $\mathcal{M}_{\mathcal{B}_i}$, and get all k-mode combinations to the candidate set $\Theta_{\mathcal{B}_i}$.

Each θ is first evaluated for throughput, and the backbone with insufficient throughput is eliminated. Then all qualified candidate sets are evaluated against the target $\mathcal{NH}$ to obtain the $\mathcal{B}_{re}$ and θ_{re} with the smallest $\mathcal{NH}$.

B. Access

All satellites other than backbone satellites are access satellites. The goal of the access network is to carry the traffic generated by the local satellites with as ***few ISLs*** as possible. Local satellites generate two types of traffic: ***(a)*** Inter-regional traffic generated locally that needs to be forwarded to other regions. ***(b)*** Local traffic between neighboring satellites. For *(a)*, the access node only needs to forward the traffic to a backbone node. Traffic *(b)* occurs between access nodes. A cost-efficient design should be able to properly provision the SCTs and the connections to undertake these two parts of traffic with fewer SCTs.

We group all access nodes according to their nearest backbone nodes. The number of backbone nodes is $|\mathcal{B}|$. All nodes are divided into $|\mathcal{B}|$ ***groups***. Each group contains several access nodes and one backbone node. We use $G\langle V, E\rangle$ to represent the group of satellites.

SCT number cannot be modified once the satellite is deployed. We use the *local traffic features* rather than real-time traffic matrix. The traffic features here refer to the relationship between the communication distance and the traffic scale of local traffic. That is, the proportion of traffic with a distance of d to all traffic is ρ_d, which can be obtained in the statistics of terrestrial traffic [34], [35]. The proportion of traffic between each pair of satellites can be obtained according to the distance by ρ_d. If the communication distance exceeds this group of satellites, it is set as inter-regional traffic, which is the traffic between the access satellite and the backbone satellite. According to this proportion, a *logical flow matrix* $\mathcal{F}_l$ can be obtained. We get access network according to this matrix, as shown in Alg. 2.

Algorithm 2: Constellation design for access

```
Input: $\mathcal{F}_l, G < V, E >, d^{max}$
Output: $\hat{G} < V, \hat{E} >$
$\hat{E} = E; MinCost = +\infty;\ MinE = \hat{E}$;
while $|\hat{E}| > |V| - 1$ do
    $MinScore = +\infty; deleted = None$ ;
    for each $l_{ij}$ in $\hat{E}$ do
        $s_{ij} = p_{ij}\mathcal{H}^{\hat{E}-l_{ij}}$;
        if $s_{ij} < MinSore$ then
            $deleted = l_{ij}; MinScore = s_{ij}$;
    $\hat{E} = \hat{E} - deleted$;
    $cost = |\hat{E}|\mathcal{H}^{\hat{E}}$;
    if $cost < MinCost$ then
        $MinCost = cost; MinE = \hat{E}$;
$\hat{E} = MinE$;
for each $v$ in $V$ do
    while $d_v^{\hat{E}} > d_v^{max}$ do
        Delete the edge $l_{iv}$ used least by $\mathcal{F}_l$;
        if $\hat{G} < V, \hat{E} >$ is not connected graph then
            Add $l_{ij}$ from $E - \hat{E}$ with the smallest $\mathcal{H}$ among
              edges that make $\hat{G}$ connected;
for each $l_{ij}$ in $E - \hat{E}$ do
    if $d_i^{\hat{E}} < d_i^{max}, d_j^{\hat{E}} < d_j^{max}$ then
        $s_{ij} = (|\hat{E}| + 1)\mathcal{H}^{\hat{E}+l_{ij}}$;
        if $s_{ij} < MinCost$ then
            $\hat{E} = \hat{E} + l_{ij}; MinCost = s_{ij}$.
```

Initially, we get E according to physical constraints. Edges between any two nodes whose distance is less than the maximum distance D_h can be added to E. And due to the limitation of the maximum number of SCTs of each satellite, we initialize its final maximum degree for node i as d_i^{max}.

In each loop, we evaluate the score s_{ij} of each edge and delete the edge with the lowest cost-effectiveness score. The score shows the importance of the edge, which is calculated as $s_{ij} = p_{ij}\mathcal{H}^{\hat{E}-l_{ij}}$. $\mathcal{H}$ is the weighted average hops. When the graph is a disconnected graph, $\mathcal{H} = +\infty$. p_{ij} is the priority of the edge l_{ij}. Smaller p_{ij} means higher priority. When the degree of the two nodes connected by the edge exceeds the limit d^{max}, the edge will have a higher priority to be deleted. p_{ij} can be calculated by the product of the priorities of the two nodes connected by the edge: $p_{ij} = p_i p_j$. The degree of node v_i is d_i, and the maximum allowed degree is d_i^{max}. When $d_i > d_i^{max}$, $p_i = d_i^{max}/d_i$, otherwise $p_i = 1$.

The edge with the smallest s_{ij} is deleted in each loop. And cost of the $\hat{G}$ is calculated after deletion. The graph $\hat{G}$ with the smallest cost in all loops is selected as the target topology.

There may be some nodes where $d_i > d_i^{max}$ in the target topology, so some repairs are needed. For nodes that $d_i > d_i^{max}$, select the one that is least used by $\mathcal{F}_l$ to delete. If the deletion leads to the disconnection of the graph, add an

TABLE I: Primary constellation parameters.

	Planes	*Sat/plane*	*Altitude*	*Inclination*
Telesat [7]	27/**40**	13/**33**	1015/**1325**	98.98/**50.88**
OneWeb [8]	36/**32**/32	49/**72**/72	**1200**	87.9/**55**/40
Kuiper [9]	28/36/**34**	28/36/**34**	590/610/**630**	33/42/**51.9**
Starlink [10]	72/**72**/ 6/4/36	22/**22**/ 58/43/20	540/**550**/560/ 560/570	53.2/**53**/97.6/ 97.6/70

- The **bold** shells are the shells used in the evaluation.

edge from E for supplementation. That added edge should not exceed the degree limit and has the smallest $\mathcal{H}$ among all edges of i. Finally, try to add more edges. If $\mathcal{NH}$ can be improved after the supplement, add the edge.

After the algorithm, we can get the SCT number of each satellite and the connection scheme of the group.

IV. PERFORMANCE EVALUATION

We conduct simulations to evaluate several aspects of DIV, including: **(i)** how can DIV improve cost-effectiveness? **(ii)** can DIV maintain high cost-effectiveness under temporal variation, spatial variability and traffic diversity? **(iii)** how various configurations (constellation parameters, maximum SCT limitations, *etc.*) affect the performance of DIV?

A. Simulated Constellation Network

Constellation parameters. Orbital data for several deployed constellations is given in Table I. We take a shell for each constellation for simulation[1], and the parameters of the chosen shell is marked **in bold**. Among them, *we use **Starlink** data for a more detailed analysis*. We use the astronomy library PyEphem [36] to simulate the track of satellites.

Traffic patterns. According to the report of Cisco [37], by 2021 more than 99.35% percent of traffic is data center (dc) traffic. Amazon is recently deploying GSes to interconnect clouds and satellites. Satellite Internet may also carry a large amount of dc-related traffic in the future. Data center traffic includes dc-to-user traffic(14%), dc-to-dc traffic(9%) and within dc traffic(77%) [37]. We use the traffic from the first two parts to simulate satellite traffic. The traffic model used here refers to the distribution of Amazon dcs [38], [39].

The dc-to-dc traffic that satellites may undertake is the traffic between regions. We obtain the scale r_i and location of region i from [38] and the traffic f_{ij} between regions can be generated:

$$f_{ij} = \frac{\mu r_i r_j}{\sum_{i \in V} \sum_{j \in V, j \neq i} r_i r_j}$$

where μ represents the total size of inter-regional traffic. This traffic is forwarded using the k-shortest path algorithm. According to the population of each $1° \times 1°$latitude and longitude block in [40] and the network penetration rate of each country in [41], the number of Internet users in each block can be obtained. Users of each block communicate with the nearest dc edge node of [39] with the shortest path. The total scale of local traffic is ν, which is weighted and equally divided into each block according to the number of users.

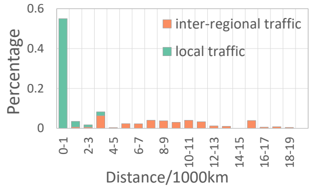


Fig. 7: Histogram of communication distance.

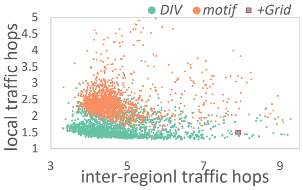


Fig. 8: Hop count for different modes and traffic.

The proportion of these two parts of traffic varies in different scenarios. We discuss that in § IV-D. In other analyses, we mainly use the ratio *local/inter* $= 14/9$ given above. The distribution of the traffic scale at different distances for this traffic model is shown in Fig.7.

Comparisons. The objective of DIV is to achieve higher cost-effectiveness in constellation design. According to Equ. 2, the evaluation of it is composed of two parts: hop count and the number of ISLs. Therefore, in the evaluation of this section, we evaluate the average hop count for traffic and the average number of ISLs per satellite separately, and use the product ***cost*** of the two to map the cost-effectiveness.

We compare the performance of DIV with other two state-of-the-art constellation architecture, *+Grid* [33] and *motif* [23]. There are multiple *motifs* for a constellation. If not specifically stated, the *motif* mentioned below is the optimal one for a traffic model.

B. Cost-Effectiveness Improvement

We find 3412 backbones and 1448 *motifs* in the 72×22 Starlink constellation. Each backbone can generate the access network according to Alg. 2. We call each of these networks a *div*. Their performance in different traffic is shown in Fig. 8. It can be found that *divs* are more friendly to both long-distance traffic and local traffic. The average hop count for all *motifs* for inter-regional traffic is 4.8, while for local traffic it is 2.52. For *divs*, the two values are 4.78 and 1.64 respectively. The difference between them on inter-regional traffic is small because both of them use repeated patterns. The slight advantage of DIV is that there are relatively fewer hops when some *divs* are small in scale. The obvious advantage for local traffic is that DIV uses the access network to optimize local traffic. Compared with *+Grid* (the two values are 7.9 and 1.48) which have great advantages for local traffic, DIV cannot match the average hop count for local traffic, but far exceeds the average hop count for inter-regional traffic. On average, the hop count at all traffic of all DIV (2.87) is better than that of *motif* (3.42) and *+Grid* (4).

There are different optimal *motifs* under different traffic models. We find the optimal *motif-0* for all traffic, the optimal *motif-b* for inter-regional traffic, and the optimal *motif-a* for local traffic to illustrate the superiority of DIV. Table II compares the optimal *motifs*, *div* and *+Grid*. The hop count of DIV is not optimal for both local and inter-regional traffic. But its average hop count is ahead of all other types of optimal solutions. This is because DIV divides the functions so that the satellites of each function can be optimized for the specified traffic characteristics, which can not only improve

---

[1] Although there are constellations that claim not to use ISLs, we also give the results of them to verify the performance under different constellation parameters.

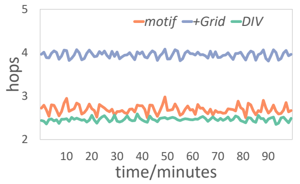


Fig. 9: Fluctuation in hop count over time.

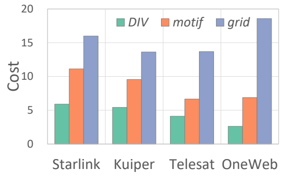


Fig. 10: Constellation parameters and cost.

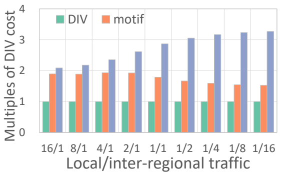


Fig. 11: Traffic proportion and cost.

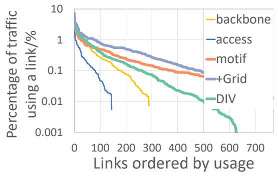
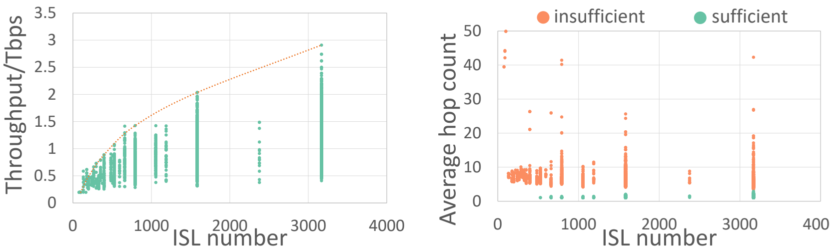


Fig. 12: Differences in usage of all links.



(a) ISL-throughput

(b) ISL-hops

Fig. 13: Backbones with different modes.

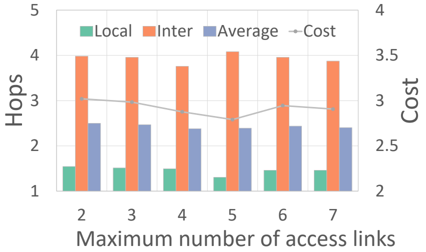


Fig. 14: Efficiency change with maximum access links.

the performance, but also reduce the occupancy of the link. Its cost-effectiveness is 2.7 times that of +*Grid* and 1.9 times that of the optimal *motif*.

TABLE II: Cost-effectiveness comparison.

	inter hops	*local hops*	*avg hops*	*ISL/sate*	*cost*
+Grid	7.90	1.49	4.00	4	15.99
motif-0	3.89	2.07	2.78	4	11.13
motif-b	3.58	2.26	2.78	4	11.13
motif-a	5.58	1.45	3.07	4	12.27
DIV	3.96	1.46	2.44	2.42	5.59

To verify the stability as the constellation moves, we simulate satellite trajectories of Starink for 100 minutes (about one orbital period) and evaluate the performance of the constellation every one minute. Fig. 9 shows the change in the average hop count for each type of topology over time. With the change of time, the hop counts of these schemes all fluctuate slightly. The standard deviations of hop count for *motif*, +*Grid* and DIV are 0.095, 0.072, and 0.048, respectively. For local traffic, +*Grid* fluctuates the least. But the continuous fluctuation of inter-regional traffic makes the overall fluctuation exceed DIV. *Motif* with varying lengths of ISLs make users often experience large hop changes after switching satellites. As for DIV, due to the division of functions, the backbone and access can remain relatively stable and uniform, so the overall stability can be guaranteed.

C. Effect on Various Constellation Configurations

To verify the scalability of the scheme, we evaluate it in different constellations from Table I. Since the maximum ISL length of some constellations may exceed the current technical level, we limit the maximum length to $6000km$. Fig. 10 shows the comparison between the best *motif div*, and +*Grid* for each constellation. The results show that among these different constellations, DIV always achieves the best efficiency. DIV is $2.5\times \sim 7\times$ more cost-efficient than +*Grid* and $1.6\times \sim 2.6\times$ more cost-efficient than *motif*.

D. Different Traffic Proportion

The reason for the improvement of DIV comes from the function division. This is because it can save ISLs and reduce costs. But as the difference in scale of inter-regional and local traffic increases, this boost can gradually diminish. We give a comparison of the effects of these schemes at different ratios, as shown in Fig. 11. The vertical axis indicates how many times the corresponding cost is DIV's. Obviously, +*Grid* costs less when the proportion of local traffic is high, while *motif* costs less when the proportion of inter-regional traffic is high, but no matter what the proportion, the cost of DIV is always the lowest. It's just that when the scale of the two types of traffic is quite different, the advantage of DIV is not so obvious.

E. Link Usage Divergence

Fewer hops do not necessarily mean higher throughput. When the traffic distribution is extremely uneven, congestion in some areas can reduce throughput. The greater the difference between the links, the earlier the congestion occurs and the lower the throughput may be. Therefore, we evaluate the occupancy of different links. We count the usage of different links in a snapshot, then sort all ISLs by frequency of use. As shown in Fig. 12, +*Grid* has the most uneven load, and the heaviest link bears 7.23% of traffic. The most heavily loaded links in *motif* and DIV bear the traffic of 3.04% and 2.85% respectively, and are relatively less prone to congestion.

In DIV, the use of backbone and access links is very different. And almost all the links at the head of utilization in DIV are backbone links. This has important implications for future differentiated designs and deployment. Only the backbone links, which account for 34.5% of all links, require the current high-capacity SCTs. The traffic of the remaining large number of access links will never reach the current capacity, so they can use cheaper SCTs without degrading overall performance.

F. Backbone Scale

We list all possible backbones and estimated their maximum throughput of inter-region traffic as described in section II-B. The routing algorithm used here is the 4-shortest path algorithm. The maximum capacity of a single link is $100Gbps$.

Fig. 13(a) shows the distribution of throughput of all backbones. Intuitively, the larger the number of ISLs, the higher the throughput, and there is indeed such a trend as a whole. However, as the scale increases, the maximum throughput growth rate becomes slower, so the cost-effectiveness also continues to decrease. Setting the backbone throughput target to be $1Tbps$, Fig. 13(b) shows the distribution of all backbones with sufficient and insufficient throughput. There are a total of 3507 types of sufficient backbones, of which 33.7% use only half of the ISLs, and 69 use no more than $1/4$ of the ISLs.

G. Maximum Access Links

Due to physical constraints, the number of SCTs that a single satellite can be equipped with may be limited. We set each backbone node to be equipped with 4 SCTs for the connection between the backbones, and all satellites are equipped with a maximum of x access SCTs. In the previous discussion, we set $x = 6$. In Fig. 14, we evaluate the efficiency for different x. The deviation in efficiency is negligible. This means that even if we only use cheaper satellites with very low configuration, it will not reduce the overall cost-effectiveness.

V. LIMITATIONS AND FUTURE WORK

Limitations. Technology, undisclosed data, and hierarchies lead to the following limitations:

- The launch costs of SCTs and satellites are evaluated, but due to the secretiveness of the industry, more accurate manufacturing and maintenance costs cannot be obtained.
- The choice for backbone design is tiny if the maximum length of ISLs is small due to technical reasons.
- Damage to the backbone node will cause the access satellites in the group to lose contact with other areas unless connected to other backbones.
- The grouping of access nodes causes traffic between nodes that are adjacent but not in the same group to be detoured to the backbone.

Future work. The functional and hierarchical division of the constellation brings opportunities for future design. For routing, hierarchical routing can be designed to improve scalability from the network layer and tolerate local topology adjustments. In the survivable design, backbone satellites have higher priority for backup redundancy. In addition, the structure of this paper is oriented towards single-shell, which can be extended to multi-shells in future.

VI. RELATED WORK

Constellation architecture design. The constellation orbital structure is designed to meet the service demands with fewer satellites. [15] can achieve visibility of k satellites anywhere in the world with a small number of satellites. [16] satisfies the coverage of complex areas with fewer satellites. [17] designs the constellation with the minimal number of satellites to satisfy the backhaul requirement. [18] finds an optimal altitude where the constellation strikes the balance between satellite availability and signal loss. [31] shows that optimizing phase factor can reduce the average hops. These schemes design the optimal constellation orbits according to the trajectory of satellites, but there are special requirements for the orbit, which may be difficult for commercial constellations to meet.

With the emergence of LEO mega-constellations, it needs a hierarchical structure to improve efficiency. The traditional hierarchical structure of the constellation is that satellites of different types act as different layers [42], [43]. [8] uses a hierarchical dynamic design according to the height of the satellite to improve the performance of the network, while [9], [10] designed more from the perspective of coverage. [32] proposes a routing algorithm for IGSO/GEO/MEO hierarchical architecture to combine the advantages of three types of satellites. There are also some studies integrate the space-air-ground networks as shown in [44].

Constellation link and routing design. Constellation can improve the throughput and further improve the utilization through link and routing design. [2], [3] showes that ISLs can significantly improve the overall throughput, and more ground gateways, higher ISL, GSL capacity can bring higher throughput, but it can cause an increase in the cost of a single satellite. [23], [24] reduces hops by designing topology. [25], [26] uses traffic engineering through the cooperation of multiple satellites to improve throughput, but it bring the problem of detour. The storage and computing capabilities of the satellite can be increased to allow the satellite to operate even when the communication function is idle [45], but at the same time the cost is increased.

Extending satellite life. We can reduce cost is by increasing the service life of the satellites. The rate of battery life consumption is positively correlated with the depth of discharge. [19] improves battery life by hibernating some satellites during eclipse. [20], [21] design routing algorithms to transmit traffic over satellites exposed to the sun rather than the eclipsed satellites. [22] proposes an algorithm to maximize the lifetime of a satellite subject to J2 and atmospheric drag perturbations. This type of work can be combined with the work of this paper to jointly improve satellite efficiency.

VII. CONCLUSION

In this paper, we design the cost-effective mechanism DIV for mega-constellations by changing the networking and functions of satellites. Based on the inherent differences and characteristics between local and inter-regional traffic of internet, DIV divide the functions of satellites to backbone and access and design the network structure. The simulations show that for the most deployed constellation Starlink, the number of ISLs can be reduced to 60.5%, and the cost-effectiveness can be increased to 2.86 times. For multiple constellations that are currently being deployed, it can achieve a 2.5-7 times efficiency. Moreover, changes in time and satellite configuration constraints will not significantly reduce this ability. We have verified the robustness and scalability of the scheme from multiple perspectives, hoping that the mechanism can bring practical inspiration to the actual mega-constellation deployment and future networking.

REFERENCES

[1] M. Handley, "Delay is not an option: Low latency routing in space," in *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*, 2018, pp. 85–91.

[2] I. Del Portillo, B. G. Cameron, and E. F. Crawley, "A technical comparison of three low earth orbit satellite constellation systems to provide global broadband," *Acta Astronautica*, vol. 159, pp. 123–135, 2019.

[3] N. Pachler, I. del Portillo, E. F. Crawley, and B. G. Cameron, "An updated comparison of four low earth orbit satellite constellation systems to provide global broadband," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2021, pp. 1–7.

[4] H. Jones, "The recent large reduction in space launch cost," in *International Conference on Environmental Systems*. 48th International Conference on Environmental Systems, 2018.

[5] C. Christensen and S. Beard, "Iridium: failures & successes," *Acta Astronautica*, vol. 48, no. 5-12, pp. 817–825, 2001.

[6] S. Finkelstein and S. H. Sanford, "Learning from corporate mistakes: The rise and fall of iridium," *Organizational Dynamics*, vol. 29, no. 2, pp. 138–148, 2000.

[7] "Application for fixed satellite service by telesat canada," 2020, https://fcc.report/IBFS/SAT-MPL-20200526-00053/.

[8] "Application for fixed satellite service by worldvu satellites limited," 2020, https://fcc.report/IBFS/SAT-LOI-20170301-00031.

[9] "Application for fixed satellite service by kuiper systems llc," 2020, https://fcc.report/IBFS/SAT-LOA-20190704-00057.

[10] "Application for fixed satellite service by space exploration holdings, llc," 2020, https://fcc.report/IBFS/SAT-LOA-20200526-00055.

[11] Hensoldt, "High performance optics," 2021, https://www.hensoldt.net/.

[12] Tesat, "Tesat products conlct," 2021, https://www.tesat.de/products.

[13] Mynaric, "Laser communication in space," 2021, https://mynaric.com/products/space/.

[14] B. Smutny, H. Kaempfner, G. Muehlnikel, U. Sterr, B. Wandernoth, F. Heine, U. Hildebrand, D. Dallmann, M. Reinhardt, A. Freier *et al.*, "5.6 gbps optical intersatellite communication link," in *Free-Space Laser Communication Technologies XXI*, vol. 7199. International Society for Optics and Photonics, 2009, p. 719906.

[15] A. H. Ballard, "Rosette constellations of earth satellites," *IEEE transactions on aerospace and electronic systems*, no. 5, pp. 656–673, 1980.

[16] Y. Ulybyshev, "Satellite constellation design for complex coverage," *Journal of Spacecraft and Rockets*, vol. 45, no. 4, pp. 843–849, 2008.

[17] R. Deng, B. Di, H. Zhang, L. Kuang, and L. Song, "Ultra-dense leo satellite constellations: How many leo satellites do we need?" *IEEE Transactions on Wireless Communications*, 2021.

[18] A. Al-Hourani, "Optimal satellite constellation altitude for maximal coverage," *IEEE Wireless Communications Letters*, 2021.

[19] Y. Yang, M. Xu, D. Wang, and Y. Wang, "Towards energy-efficient routing in satellite networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3869–3886, 2016.

[20] J. Liu, B. Zhao, Q. Xin, J. Su, and W. Ou, "Drl-er: An intelligent energy-aware routing protocol with guaranteed delay bounds in satellite mega-constellations," *IEEE Transactions on Network Science and Engineering*, 2020.

[21] M. Hussein, G. Jakllari, and B. Paillassa, "On routing for extending satellite service life in leo satellite networks," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 2832–2837.

[22] R. A. Zidek and I. V. Kolmanovsky, "Deterministic drift counteraction optimal control and its application to satellite life extension," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 3397–3402.

[23] D. Bhattacherjee and A. Singla, "Network topology design at 27,000 km/hour," in *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, 2019, pp. 341–354.

[24] Z. Liu, W. Guo, C. Deng, W. Hu, and Y. Zhao, "Perfect match model-based link assignment to design topology for satellite constellation system," *International Journal of Satellite Communications and Networking*, vol. 34, no. 2, pp. 263–276, 2016.

[25] T. Lan, H. Li, Q. Wu, Z. Lai, and J. Liu, "Exploiting path diversity to increase system performance in mega-constellations," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2021, pp. 1–7.

[26] X. Li, F. Tang, L. Chen, and J. Li, "A state-aware and load-balanced routing model for leo satellite networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[27] D. C. Beste, "Design of satellite constellations for optimal continuous coverage," *IEEE Transactions on Aerospace and Electronic Systems*, no. 3, pp. 466–473, 1978.

[28] J. G. Walker, "Satellite constellations," *Journal of the British Interplanetary Society*, vol. 37, p. 559, 1984.

[29] L. Rider, "Analytic design of satellite constellations for zonal earth coverage using inclined circular orbits," *Journal of the Astronautical Sciences*, vol. 34, pp. 31–64, 1986.

[30] A. U. Chaudhry and H. Yanikomeroglu, "Free space optics for next-generation satellite networks," *IEEE Consumer Electronics Magazine*, 2020.

[31] Q. Chen, G. Giambene, L. Yang, C. Fan, and X. Chen, "Analysis of inter-satellite link paths for leo mega-constellation networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2743–2755, 2021.

[32] W. Zhou, Y.-f. Zhu, Y.-y. Li, Q. Li, and Q.-Z. Yu, "Research on hierarchical architecture and routing of satellite constellation with igso-geo-meo network," *International Journal of Satellite Communications and Networking*, vol. 38, no. 2, pp. 162–176, 2020.

[33] A. Ferreira, J. Galtier, and P. Penna, "Topological design, routing and handover in satellite networks," *Handbook of wireless networks and mobile computing*, vol. 473, p. 493, 2002.

[34] A. Aral and T. Ovatman, "A decentralized replica placement algorithm for edge computing," *IEEE transactions on network and service management*, vol. 15, no. 2, pp. 516–529, 2018.

[35] G. Liu, H. Shen, and H. Chandler, "Selective data replication for online social networks with distributed datacenters," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 8, pp. 2377–2393, 2015.

[36] "Pyephem," 202, https://rhodesmill.org/pyephem/.

[37] "Cisco global cloud index 2015-2020," 2015, https://www.cisco.com/c/dam/m/en_us/service-provider/ciscoknowledgenetwork/files/622_11_15-16-Cisco_GCI_CKN_2015-2020_AMER_EMEAR_NOV2016.pdf.

[38] "Map of amazon's data centers," 2021, https://wikileaks.org/amazon-atlas/map/.

[39] "Amazon cloudfront," 2021, https://aws.amazon.com/cloudfront/features/.

[40] CIESIN, "Gridded population of the world (gpw), v4," 2020, https://sedac.ciesin.columbia.edu.

[41] M. M. Group, "World internet users statistics and 2020 world population stats," 2021, https://www.internetworldstats.com/stats.htm.

[42] L. Qiao, H. Yan, Y. Zhang, R. Zhang, and W. Jia, "Multilayer satellite network topology design technology based on incomplete igso/meo constellation," in *International Conference on Space Information Network*. Springer, 2019, pp. 28–38.

[43] C. Jiang and X. Zhu, "Reinforcement learning based capacity management in multi-layer satellite networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4685–4699, 2020.

[44] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2714–2741, 2018.

[45] D. Bhattacherjee, S. Kassing, M. Licciardello, and A. Singla, "In-orbit computing: An outlandish thought experiment?" in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks*, 2020, pp. 197–204.

[46] D. Bhattacherjee, W. Aqeel, I. N. Bozkurt, A. Aguirre, B. Chandrasekaran, P. B. Godfrey, G. Laughlin, B. Maggs, and A. Singla, "Gearing up for the 21st century space race," in *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*, 2018, pp. 113–119.

[47] E. Kulu, "Satellite constellations-2021 industry survey and trends," in *35th Annual Small Satellite Conference*, 2021.

[48] T. Butash, P. Garland, and B. Evans, "Non-geostationary satellite orbit communications satellite constellations history," pp. 1–5, 2021.