This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

Satellite Computing: Vision and Challenges

Shangguang Wang, *Senior Member, IEEE,* Qing Li

***Abstract*—The space industry experiences a rise in low-earth-orbit satellite mega-constellations to achieve universal connectivity. At the same time, cloud firms (such as Google, Microsoft, and AWS) also have ambitions for computing in space to offer public cloud services in orbit. Satellite computing, as a new emerging concept, is promising to enable new paradigms in on-orbit autonomy, remote sensing, edge computing, and other areas by empowering satellites with computing resources. However, LEO mega-constellations bring inherent challenges for satellite computing in satellite networking, computing, and others due to the moving core infrastructure, reduced system power budget, and harsh space environment. This paper presents vision and challenges for satellite computing based on a brief survey of the very recent literature in the "NewSpace" era and gives a case study of an open research platform on real satellites named Tiansuan constellation. This paper aims to call researchers to collaboratively undertake the research of satellite computing and provide some insights for the research community.**

***Keywords*-Satellite Computing, Satellite Internet, Testbed**

I. INTRODUCTION

Recently, multiple commercial powers are gearing up to deploy LEO mega-constellations (with hundreds to tens of thousands of LEO satellites) to provide global low-latency high-bandwidth Internet. Dated to July 2023, SpaceX has launched more than 4700 Starlinks and London-based OneWeb has launched 620 internet satellites.[1] With the exciting development of satellite networks, satellites are equipped with more powerful computing resources to supports in-orbit data processing [1]. Cloud firms (Google, Microsoft, and AWS) also have ambitions for computing in space by forging significant alliances with leading satellite companies. Mega-constellations offer public cloud services in orbit so they are another hybrid cloud option for global users.

As satellite computing is still a concept, it deploys computing resources at the satellites to enable new paradigms in on-orbit autonomy, remote sensing, edge computing, and other areas [2]. First, traditional satellite systems adopt a bent-pipe architecture, where ground stations send human-operated commands to orbit and satellites reply with raw data (just like a bent pipe). This human-in-the-loop architecture breaks down in face of the mega-constellation scale [2]. Satellite computing can enable on-orbit self-operation and reduce the reliance on the ground segment. Second, space-native raw data are increasing explosively with the constellation size and cannot be downloaded in time due to the limited satellite-ground link bandwidth. Satellites with computing power can process the raw data, identify the features of interest, and transmit only the interesting data, improving the transmission efficiency and reducing the ground infrastructure costs [1]. Third, advanced computing platforms onboard transform constellations into sophisticated data processing infrastructure and enable public cloud services in space as they exist today on the ground [3].

The LEO mega-constellation is a new type of infrastructure and brings inherent challenges for satellite computing which is fundamentally different from cloud computing and edge computing on the ground. LEO mega-constellations consist of thousands of satellites and each moves at high speed relative to the Earth and other satellites. For example, a satellite at an altitude of 550km must maintain a speed of 27,000km/h to maintain its orbit [4]. Besides, satellites in mega-constellations usually have very limited weight and volume due to the reduced manufacturing and launching cost. Moreover, the space environment is harsh due to deep vacuum conditions, radiation, strong vibrations, and higher temperature ranges [5]. These factors make networking, computing, and other research in satellite computing challenging. First, satellite network has generated tremendous interest among networking researchers. They have highlighted the new opportunities and challenges of LEO mega-constellations [6]–[8] and explored network topology [4], intra-constellation routing [9], inter-domain routing [10]. But, it is unclear how to integrate the frequently changed satellite networks with the Internet's Border Gateway Protocol and how to achieve congestion control. At the same time, some researchers propose orbit edge computing to process the space-native data [1] and provide public service for ground users [11]. However, it is still unsettled how to design space environment adaptive computing hardware, construct computing platforms, and develop organization paradigms for applications due to the moving service infrastructure, reduced system power budget, and harsh space environment. There are also many open questions about how to realize reliable and secure satellite computing.

This paper aims to present vision and challenges for satellite computing in the ”NewSpace” era. We first take a quick look at the current development of satellite constellations (Section II). We then review the recent work in satellite computing (Section III), discuss the challenges, and present several potential research topics (Section IV). We also give a case study of an open research platform on real satellites named Tiansuan constellation and show some experiments deployed on Tiansuan constellation (Section V). Finally, we conclude the paper and call researchers to collaboratively undertake the research on satellite computing (Section VI).

II. A GLANCE AT THE EXISTING CONSTELLATIONS

Recent years witness the surge of LEO mega-constellations, where hundreds or thousands of LEO satellites move around

Shangguang Wang is with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China, 100876. E-mail: sgwang@bupt.edu.cn.

Qing Li is with the School of Computer Science, Peking University, Beijing, China, 100871. E-mail: liqingpostdoc@pku.edu.cn. Qing Li is the corresponding author.

---

[1] https://en.wikipedia.org/wiki/List_of_Starlink_and_Starshield_launches

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

the earth at the altitude of about 500-2000 km over the sea level at a high speed. Most LEO satellites are miniaturized at lower cost, such as minisatellites, microsatellites, and nanosatellites. Their orbits are usually determined by several key parameters, such as orbit plane number, satellite number of each plane, inclination, altitude, and phase shift. For example, the planned Starlink Gen1 comprises 4,425 satellites. The first phase uses 24 orbits, each with 66 satellites, for a total of 1,584 satellites. Orbit inclination is 53° and the altitude is about 550 km.

In this section, we show the top 20 constellations (in terms of the launched satellite number) in TABLE I mainly according to the statistics from NewSpace Index[2] and update the data dated by July 2023. Only about 4% of the constellations recorded by NewSpace Index have been fully launched and 8% are currently being launched, so the information of the top 20 constellations tells a lot. From the launched satellite number, we can observe that the large constellations are dominated by only several organizations. This gap may become even larger in the future. From the first launch date, 19 of the 20 constellations are launched after 2010, which implies the surge of the NewSapce era. From the form factor of satellites, we can tell that small satellites (weight lower than 500Kg) dominate the LEO mega-constellations. From the applications, constellations are widely applied for Internet, Earth observation, IoT/M2M, AIS (Automatic Identification System), ADS-B (Automatic Dependent Surveillance-Broadcast), GNSS (Global Navigation Satellite System), PTN (Positioning, Timing, Navigation) and so on.

The industry has a new expectation of mega-constellations in application fields of the Internet, Earth observation, and IoT fields. Satellite Internet is promising to provide high-speed, low-latency broadband internet across the globe. For example, Starlink is available for Internet service for well over 1.5 million users globally[3]. Crisis on the earth such as the war in Ukraine, climate change, and the covid pandemic fuel the demand for earth observation. For example, Earth observation companies like BlackSky or Planet use their constellation to monitor the activities interesting for customers such as investors and financial firms[4]. Satellite IoT helps modern-day businesses and organizations to track, monitor, and manage assets anywhere on the Earth, improving remote operations. For example, the Iridium network is uniquely qualified to provide global satellite IoT services by 66 LEO satellites blanketing the earth with reliable and ubiquitous coverage[5].

III. LITERATURE REVIEW

While the industry is gearing up to deploy LEO mega-constellations, the research community is also interested in the new satellite infrastructure. Recent work has explored the new research filed from many aspects. We summarize the related work into three categories: networking, computing, and others as follows.

A. Satellite Network

Back in the 1990s, Geosynchronous Orbit (GEO) and small LEO satellites networks were heavily researched, but more recently, position papers have shed light on the exciting new possibilities and obstacles that come with LEO mega-constellations. These papers [6]–[8] have pointed out that satellite networks have immense potential for low-latency long-distance communication, and that there is still a lot of unexplored territory when it comes to network topology, intra-constellation routing, inter-domain routing, reordering, and congestion control.

1) Physical Layer: Compared with bent-pipe architecture without inter-satellite links, Singla *et al.* [12] analyze the benefit of inter-satellite links in improving satellite network latency, throughput, and reliability. Besides, to enable the research in the NewSpace era, they develop a packet-level satellite network simulation tool for simulating and visualizing the network behavior of mega-constellations [9]. They design a new network topology to maximize network bandwidth and minimize latency after analyzing the limitation of the existing +Grid topology [4]. The network topology is generalized from +Grid pattern by repeating a motif for each satellite in the network, all connected in exactly the same way, where the connection remains stable over time. Kur *et al.* [13] conduct a strategic evaluation of a simulated inter-satellite link in Galileo satellites, utilizing variance component estimation of varying complexity to test performance. The results of the evaluation indicate that incorporating different estimation methods in inter-satellite link measurements could enhance link accuracy. However, the paper does not provide clarity on the deviation range and calibration method of the inter-satellite link. Arslan *et al.* [14] use signal attenuation of satellite links to estimate rainfall accumulation. The results of their study, which focused on Ku-band link measurements, indicate that satellite link attenuation is highly sensitive to heavy rainfall. Therefore, it is necessary to improve the sensitivity and prediction accuracy to small rainfall.

2) Network Layer: Routing is a fundamental function and core task of communication networks, and satellite networks are no exception. As such, the development of routing technologies in satellite networks is a major focus, aimed at enhancing user service quality and improving business efficiency. In order to deal with the high cost of two-point communication in satellite networks, Lan *et al.* [15] propose a traffic scheduling strategy based on multiple routing planes, which flexibly utilizes the diversity of network links to simplify the path selection problem. The results show that this method can significantly improve network throughput and reduce latency. Highly dynamic and time-varying topology are typical characteristics of LEO satellite networks. To address this issue, Lai *et al.* [16] propose a hybrid data transmission architecture that combines LEO constellations and ground distributed base stations. This architecture aims to achieve stable communication and routing in satellite networks. The authors develop a

---

[2] https://www.newspace.im

[3] https://www.starlink.com

[4] https://spacenews.com/a-boom-in-earth-observation-satellites-creating-new-demands-for-intelligence/

[5] https://www.iridium.com/blog/2021/04/01/what-is-satellite-iot-and-how-is-it-used/

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

TABLE I: Global Top 20 Constellations Information (Dated by July 2023).

Index	Organization	Launched /Planned Number	First Launch	Form Factor	Application
1	SpaceX (Starlink Gen1)	4700 / 4408	2018	Smallsat	Internet
2	OneWeb	620 / 648	2019	Smallsat	Internet
3	Planet (Flock/ Dove/SuperDove)	555 / 150	2013	CubeSat	Earth Observation
4	Swarm Technologies	189 / 150	2018	CubeSat	IoT / M2M
5	Spire (Lemur / Minas)	166 / 150	2013	CubeSat	Weather, AIS, ADS-B
6	Chang Guang (Jilin-1)	89 / 300	2015	Satellite	Earth Observation
7	Aireon	75 / 75	2017	Hosted	ADS-B
8	DDK Positioning	75 / 75	2017	Hosted	GNSS
9	Iridium (NEXT)	75 / 75	2017	Satellite	Internet, IoT / M2M
10	ExactEarth	68 / 67	2008	Hosted, Microsat	AIS
11	Satelles	66 / 66	2017	Satellite	GNSS, PNT
12	Orbcomm (OG2)	51 / 52	2012	Microsat, Smallsat, CubeSat	IoT / M2M, AIS
13	Satellogic	39 / 90	2016	Microsat	Earth Observation
14	Globalstar (Second-Generation)	25 / 42	2010	Satellite	Internet, IoT / M2M
15	SES (O3b / mPOWER)	24 / 70	2013	Satellite	Internet
16	BlackSky	20 / 16	2016	Microsat	Earth Observation
17	Astrocast	20 / 80	2018	CubeSat	IoT / M2M
18	Kepler Communications	19 / 140	2018	CubeSat	IoT / M2M, Internet
19	Spacety	18 / 480	2018	CubeSat	Earth Observation
20	Planet (Terra Bella / Skybox)	15 / 24	2013	Smallsat	Earth Observation

simulation platform, driven by public constellation information, to verify the performance of the proposed scheme in reducing latency. However, the data transmission architecture is unable to preprocess remote sensing data, which increases the pressure on data transmission. The instability of satellite links poses a challenge to the direct application of the shortest path algorithm on the ground to satellite networks. Zhang *et al.* [17] propose a scalable two-layer routing architecture, which supports the implementation of two algorithms: delay-bounded routing and delay-aware routing. The results indicate that both algorithms can significantly optimize the average network latency. However, the authors do not combine the advantages of the two algorithms, leading to the requirement of a larger forwarding table. Handly *et al.* [18] suggest using ground relays instead of inter-satellite links to reduce access latency. They implement an enhanced routing algorithm for large networks, with the result that lower latency can be achieved. But it ignores the case where the link changes due to satellite movement. Wang *et al.* [19] investigate the routing issue of link failure in LEO satellite networks to enhance network survivability. In satellite network design, ensuring routing security is of paramount importance. Due to the resource-constrained and rapidly changing nature of satellite networks, implementing simple encryption algorithms to meet demand can be challenging. Zhao *et al.* [20] introduce a lightweight risk-averse routing algorithm to mitigate routing risks, which proves to be a successful endeavor in improving satellite network routing security. However, packet forwarding that avoids high-risk areas comes at the cost of increased link overhead.

3) Transport Layer: Satellite communication systems are crucial in providing broadband services globally and are becoming increasingly significant in terms of strategic importance. However, congestion control remains a concern when it comes to satellite transmissions. Page *et al.* [21] suggest a distributed probabilistic congestion control scheme that utilizes a datagram routing algorithm to compute the minimum delay path between any two satellite nodes. The scheme's implementation assumes a fixed topology of the constellation. However, since satellites are in constant motion, the satellite network topology is subject to continuous changes. As a result, the proposed scheme lacks practicality. Bui *et al.* [22] introduce a congestion control strategy for high-throughput satellite communication that is based on power control. By utilizing a multi-objective optimization framework, a trade-off is made between the system speed and the number of users satisfying the QoS requirement. Liu *et al.* [23] explore the impact of various TCP congestion control algorithms on commercial satellite networks, highlighting how performance enhancing proxies can significantly improve throughput. Meanwhile, Claypool *et al.* [24] compare different congestion control algorithms for satellite networks and not only find differences in throughput but also in round-trip times. However, the unique conditions of the satellite network may lead to a decline in TCP performance, particularly when dealing with network jitter and signal interference. Therefore, it is important to consider alternative protocols to support satellite communications. In this regard, Dai *et al.* [25] suggest a distributed congestion control routing protocol for LEO satellite network traffic classification that achieves distributed congestion control performance through

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

traffic classification. Nevertheless, the artificially divided traffic types may not be standardized.

4) Core Network on Satellites: Deploying a core network in aircraft and even in space is a cutting-edge topic that presents many unknown challenges. However, extending the mobile core function to remote areas can help avoid the high costs of ground station construction and operation, as well as prevent various unpredictable geological disasters. Qazi *et al.* [26] initially propose extending the evolved packet core for deployment in IoT and designing an evolved packet core architecture to achieve optimal performance and scalability. However, there has been no indication of any trend towards deployment at high altitudes yet. Moradi *et al.* [27] have made significant progress in deploying the core network on non-terrestrial networks, successfully decomposing the evolved packet core into independent, lightweight entities to address the challenges of deploying in resource-constrained and fast-moving scenarios such as UAVs. The UAV-based core network can communicate effectively with LTE base stations and smartphones. Li *et al.* [28] conduct a comprehensive study of the onboard core network, focusing on the challenges posed by the extreme mobility of satellites to the mobile core. The authors highlight issues such as signaling storms, intermittent failures, and malicious attacks on the stateful core. To address the inherent contradiction between mobility and stateful core, they develop a decoupling mechanism that separates the state and functions of the core network by saving the core network's state in UEs. They also design and verify corresponding strategies through typical core network processes. Finally, the efficacy of the space core is validated through extensive simulations.

One limitation of the aforementioned studies is the lack of verification of the proposed schemes or architectures on actual satellites. However, we have made significant strides by successfully deploying a 5G core network on the Tiansuan constellation satellite, making history as the first instance of this achievement. Through thorough discussions, we demonstrate the feasibility of deploying the core network on satellites. Finally, we validate the proper functioning of the primary functions of the onboard core network [29].

5) Satellite-Ground Integration Network: Scholars and industries have carried out work on satellite-ground integration under 5G and B5G [30], [31]. Li *et al.* [32] propose to use a handover synchronization method to reduce the network convergence time and improve the availability of space-ground integrated networks. However, the authors do not discuss the suitability of different routing protocols for the current architecture of satellite-ground integrated networks. Ji *et al.* [33] design the network control structure by constructing an optimal spatial control network to achieve the best arrangement of controllers in mega satellite networks. With the aim of enhancing the effectiveness and reliability of wireless communication services, the authors utilize geometric topology analysis to determine the optimal deployment conditions for spatial control networks, ultimately striking a balance between network size, number of controllers, and transmission latency. Although the study takes into account satellite coverage and latency, it does not account for satellite mobility. Chen *et al.* [34] propose an IP address management scheme for satellites in mega-constellations. The authors incorporate feature information and geographic locations of the two satellites into IPv6 addresses. The geographic address management mechanism only conducts duplicate address detection on the corresponding satellite routers, which can significantly prolong the lifespan of the satellite's IPv6 addresses. However, the study does not include the development of trust mechanisms for IPv6 addresses or address security issues such as identity authorization. Furthermore, Lai *et al.* [35] develop a cost-effective content distribution framework based on mega constellations. The authors store a copy of the content on the LEO satellite or cloud and dynamically allocate user requests to a cache server based on relevant information, significantly reducing content access latency. They also create a simulation platform geared towards mega constellations for performance testing [36]. The tool's effectiveness is demonstrated by evaluating and comparing the performance of several typical LEO constellations. However, the simulation platform lacks access to traffic data from real satellites, limiting its ability to assess the actual benefits of performance. The LEO satellite network architecture based on software-defined networks (SDN) is a promising network deployment architecture that has garnered attention from researchers. Tang *et al.* [37] propose an SDN-based satellite-ground integrated network architecture and investigate the dynamic cooperative transmission problem. The study yields optimal network transmission characteristics. The deployment and assignment of controllers becomes more difficult due to the highly dynamic and topological randomness of the LEO satellite networks. In addition, Chen *et al.* [38] define an adaptive controller configuration problem and then propose to use the control relation graph to measure the control overhead of the LEO satellite networks. The controller assignment and placement algorithm based on the control relation graph can effectively reduce the cost of network management and shorten the response time of satellite networks. However, the model does not take into account factors such as satellite movement and topology changes. Klenze *et al.* [39] discuss business and interconnection models for space-operating ISPs and study satellite-ground integrated routing [10] to integrate the satellite networks into today's Internet backbone. To ensure consistent high performance between the satellite and the ground, congestion control is necessary for the highly heterogeneous satellite-ground integrated network. Li *et al.* [40] propose an adaptive congestion control scheme based on multi-objective reinforcement learning, training the reinforcement learning agent to adapt to the network environment and balance congestion control objectives. However, as the scale of the satellite-ground network increases, the performance of the proposed scheme requires improvement.

The satellite ground station is a crucial element of the satellite network system architecture, making it possible to optimize the satellite network by improving the satellite ground station. However, real-time reception of satellite imagery from LEO satellite ground receivers is challenging due to receiver scarcity. Singh *et al.* [41] propose a satellite receiver system that maximizes the diversity of LEO satellites by stitching together noisy images to create a clear image of the Earth.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

RF signals can be transformed and aligned based on different tracks, viewing angles, and wireless channel quality. However, the system does not currently support high-frequency real-time access to satellite images. Additionally, the authors propose a community-driven distributed reception scheme for LEO satellite signals to address the high cost of renting ground station infrastructure [3]. The scheme synergistically leverages satellite trajectories and other environmental signals to achieve a new receiver synchronous orientation technique, which can recover signals required by handheld receivers. Vasisht *et al.* [42] propose a hybrid distributed satellite ground station design to improve the download process of satellite images, utilizing low-cost commodity hardware with a hybrid deploy model to achieve low latency and stable downlinks. However, the study does not account for the effects of weather, ground station failures, and other factors on data transmission performance. The authors of [43] take into account the traffic characteristics and LEO satellite topology to develop an iterative satellite ground station deployment scheme that maximizes revenue. This solution deploys each satellite ground station at the geographical location that has the largest marginal benefit, significantly improving system throughput. However, the global traffic demand is unbalanced, and the authors do not consider the system throughput when selecting the ground station deployment location. Guo *et al.* [44] employ UAVs as repeaters to transmit signals to satellites, taking hardware impairments into consideration. They use a closed expression of the downtime probability to assess the effects of key parameters on the system. Xia *et al.* [45] propose an SDN-based LEO satellite network communication architecture. Since a large number of routing requests need to be forwarded by satellites with limited processing power, the authors design a layered ground controller architecture. The architecture reuses satellite ground stations and adopts online network view integration to avoid the shortcomings of small ground station coverage. The routing method based on this architecture can realize load awareness and improve system flexibility. The study discusses GEO-based SDN deployments and does not consider the more prevalent LEO satellites.

Challenges: The inherent dynamics of LEO satellites create many challenges at different network layers, such as the network addressing at the physical layer [46], routing at the network layer [47], congestion control at the transport layer [48], and task scheduling at the application layer [49]. It can be found that there are still deficiencies in the research on satellite networking. Firstly, the mobility modeling of satellites is insufficient, and some even ignore the mobility. Addressing, routing, and congestion control are completely transformed into static problems, which do not conform to real space scenarios. Secondly, there is a lack of real and complete satellite data sets to support experiments. The experimental results obtained through the simulation environment and artificially manufactured data sets often deviate from the actual application. Additionally, various network architectures proposed in the above studies are verified in different simulation environments. Although ideal performance is achieved, there is a lack of deployment experiments in actual satellite networks, i.e., no real satellite platform to support the conclusion. Finally, there is no complete and effective protocol system to support the integration of network segments in space, air, ground and even ocean.

B. Computing

1) Satellite Edge Computing: Traditional satellite systems typically have limited computing capabilities for specific missions and rely on a bent-pipe architecture, which can lead to excessive energy loss, system failures, and other issues when the number of constellations increases. Enhancing the computing power of satellites has garnered significant attention as a potential solution. Edge computing offers services that are located close to the user, addressing the flexibility demands of satellite network deployment while reducing latency and providing robust computing capabilities. Cassara *et al.* [50] explore the potential advantages of integrating LEO satellite constellations with edge computing, such as enhancing system performance by delegating computing tasks. They also underscore the significance of machine learning in ensuring the quality of satellite services. Denby *et al.* [2] suggest leveraging edge computing to overcome the challenges posed by traditional centralized architecture. They present a novel approach for spatially parallel computing and illustrate that an edge computing-based architecture for satellites can substantially reduce ground infrastructure. However, their simulation platform is only applicable to satellite remote sensing services and cannot be utilized for other applications. Pfandzelter *et al.* [51] investigate the feasibility of implementing edge computing in LEO satellite networks. Given the characteristics of LEO satellites, a serverless deployment approach is a promising solution, as determined through theoretical analysis. However, the authors do not compare the performance of the serverless paradigm with other application paradigms. Furthermore, the team presents a microVM-based LEO edge virtual platform to simulate the practical effects of deploying edge computing on LEO satellites [52]. Nevertheless, the platform may be subject to hardware device performance limitations when simulating large LEO constellations. Additionally, simulating the impact of environmental changes on satellite-ground links and inter-satellite links is challenging. Moreover, the authors delve into service placement in satellite edge computing, leveraging the topological characteristics of LEO satellite networks to meet the QoS requirements of service placement [53].

Empowering satellites with computing capabilities will improve the processing efficiency onboard, while it is nontrivial to design edge computing architecture and schedule heterogeneous network resources. Xie *et al.* [54] present a satellite-ground edge computing framework that deploys computing resources in multi-layer heterogeneous edge clusters. The authors highlight that efficient and trustworthy protocol design, mass user access, and reliable reception of concurrent signals are potential research concerns. Meanwhile, Tong *et al.* [55] propose a satellite-ground network resource allocation and computation offloading decision-making approach based on mobile edge computing. In this scheme, resource allocation and computation offloading decisions are decomposable optimization problems, and techniques such as potential games and the Lagrange multiplier method are employed to obtain optimal decisions. However, the authors do not consider computational offloading between satellites. Additionally, Gao *et al.* [56] propose a potential game-based virtual network function placement method for satellite edge computing to optimize request deployment costs. The decentralized resource allocation algorithm based on potential games solves the virtual network function placement problem by searching for Nash equilibrium. However, this paper does not examine the consequences and countermeasures of satellite topology changes on the service function chain. Tang *et al.* [57] propose a hybrid cloud edge computing LEO satellite network based on a three-layer computing architecture and aim to provide diverse computing services for ground users. The authors design a distributed algorithm to transform a non-convex problem into a linear programming problem using a binary-variable relaxation method, optimizing computational offloading decisions. Nonetheless, this paper also does not account for the impact of satellite movement on computing offload, and security during computational offloading also warrants attention.

Satellite-ground IoT represents a crucial service paradigm for satellite edge computing [58]–[62]. Wei *et al.* [63] examine the impact of edge computing and machine learning on target image detection in satellite IoT and suggest an intelligent application strategy for edge computing in satellite physical networks. However, this approach is only tested in a small-scale satellite network and necessitates verification on a larger scale with more satellites and satellite data. Song *et al.* [64] separate the computing offloading of IoT mobile devices and LEO satellites into two components: ground and space. Using a unique edge computing framework, the authors perform computing offloading and resource allocation for satellite-ground IoT from two perspectives, including minimizing space segment delay via Lagrangian dual decomposition methods. The findings demonstrate that the proposed approach can significantly lower energy consumption. Nonetheless, the authors do not take into account the involvement of multiple terrestrial satellite terminals in computing offloading. Zhou *et al.* [65] present a handover algorithm for satellite edge computing that enables flexible handover and scheduling of services for fast mobile terminals. The authors suggest that utilizing machine learning-based methods to predict satellite trajectories can effectively enhance the algorithm's efficiency. Wang *et al.* [66] develop a profit maximization model for satellite-ground edge computing systems to ensure the QoE of IoT devices. The authors optimize the offloading strategy and resource allocation from multiple perspectives, and the results confirm that the service provider's profit improves while maintaining the QoE. However, the authors overlook the possibility that satellite movement may cause computing offloading terminals, thereby impacting the quality of service. Gost *et al.* [67] propose a collaborative optimization algorithm for joint communication and edge computing resource management. The algorithm enables satellites to select edge or cloud server computing tasks, reducing satellite system energy consumption while satisfying end-to-end latency constraints.

We have explored satellite edge computing in terms of service coverage and satellite energy consumption optimization. In the first work, we investigate how to deploy services flexibly and efficiently at satellite edge nodes to provide service coverage and propose an online service placement algorithm using Lyapunov optimization and Gibbs sampling [11]. The results show that it can significantly improve service coverage. Since the energy-consuming computing component deepens battery discharge and impacts battery life, we attempt to reduce energy consumption by coordinating the sensing, computing, and communication processes for an earth observation mission [68]. Specifically, an energy scheduling algorithm based on online convex optimization is proposed to reduce the depth of battery discharge.

2) AI in Space: With the significant advantages that people have reaped from various AI applications on the ground, there has been a growing interest in extending AI capabilities into space. Kothari *et al.* [69] concentrate on the implementation of deep learning in satellites and the advantages it offers for space data processing. However, their comparative experiments may not accurately reflect the real satellite network environment due to the excessive restrictions that are artificially imposed.

Given that the majority of satellites currently in orbit are remote sensing satellites, image processing is a fundamental task in satellite computing. The primary focus of mainstream research on AI in space also centers around target detection or image processing. While satellite image data sets are undoubtedly crucial, they also present a bottleneck in this field. Ding *et al.* [70] create a satellite data set containing 11,268 images annotated with 18 categories. The authors use this data set to establish a baseline that consists of 10 advanced algorithms and 70 configurations. This data set provides researchers with extensive opportunities for further research. Hoeser *et al.* [71] conduct research on the impact of convolutional neural networks on earth observation applications, providing a theoretical foundation for the application of deep learning in satellite image processing. The efficiency of image processing can be affected by the computing, memory, and power constraints of satellites. Lofqvist *et al.* [72] study the performance of convolutional neural networks using various image compression methods. Furthermore, they examine the inference time and memory consumption of tasks performed on different hardware. In addition to deep learning, deep reinforcement learning has also been utilized in the object detection of large satellite images. Uzkent *et al.* [73] propose an adaptive detection method for image resolution based on the reinforcement learning agent, which significantly enhances operational efficiency. Cube satellites have gained increasing attention from manufacturers and commercial companies in recent years. However, they are severely limited in terms of power and downlink capabilities. Maskey *et al.* [74] propose a downlink image selection scheme based on the lightweight convolutional neural network for classification before transmitting the data to the ground station. Their research has shown that even deploying a small neural network model on the cube satellites can improve data reception quality.

Challenges: It is obvious that the current research on satellite edge computing mainly focuses on computing offloading [75], resource allocation [76], service placement [38], and service optimization. It is also because most of them study satellite edge computing from the perspective of simulation, so they often ignore the inherent characteristics of satellites. In particular, the satellite network topology is set to a fixed state according to the time slot, which will hide many practical problems. On the other hand, many of the proposed optimization algorithms and applications lack a real satellite platform for verification, which is also an urgent problem facing the academic community.

Satellite computing faces several significant challenges. Firstly, the computing components in LEO satellites typically operate under extreme power conditions due to the limited energy generation and storage onboard. This necessitates energy-efficient data processing, as well as position control and communication. Satellite computing requires new power and energy management systems, computer architecture, and hardware specialization. Secondly, satellites in mega-constellations have limited computing capabilities and require collaboration to distribute computing tasks for multiple application workloads. The constellation can be viewed as a moving distributed computing infrastructure, and it requires cluster orchestration and resource management. Thirdly, the organizational paradigm of applications on LEO constellations is still unsettled. Lastly, it is urgent and necessary for satellite computing to be deployed on the real satellite test environment, particularly for AI algorithms. However, the above studies do not involve deployment on real satellites as LEO edge computing testbeds on real satellites are still missing.

C. Others

There are several new hardware designs for LEO small satellites, including antennas, GPS receivers, and FPGA accelerators. Global positioning is a crucial service provided by LEO satellites. However, the high power consumption and mobility characteristics of satellites present many challenges for satellite positioning services. Delamotte *et al.* [77] discuss the application of various differentiated multi-antenna access schemes in satellites to support anytime, anywhere connectivity. The time to first fix is significantly increased due to the relative movement and Doppler shift between GPS satellites and small satellites. The cold start of GPS satellites will cause a lot of extra energy consumption. Narayana *et al.* [78] design a low-power GPS receiver suitable for small satellites and propose an energy optimization algorithm to shorten the satellite's first positioning time. The CloudScout project [79] demonstrated by the European Space Agency on Earth observation satellites, claims that deep learning can bring many advantages to satellite autonomy, including alleviating downlink pressure and reducing operating costs. The first instance of CloudScout deployment can be optimized for cost, size, and power efficiency.

Challenges: The new computing demands have brought about a revolution in satellite systems, requiring new system architecture and hardware (such as sensors, antennas, and ground stations) to support the deployment and operation of LEO mega-constellations. Reducing costs is vital for the advent of the New Space era, while also presenting strict design requirements for satellite hardware and software. Moreover, it is extremely urgent to design and implement a real satellite experiment platform to test new payloads and devices.

IV. RESEARCH OPPORTUNITES

In this section, we envision several main research directions in satellite computing as follows.

A. Networking

It has become a consensus in the industry that satellite constellations provide ubiquitous Internet services. Therefore, it is necessary to design a new satellite network architecture that can provide efficient, flexible, and customized user services, including the physical layer, network layer, transport layer, and even control plane. In our previous work [5], we propose the idea of deploying a 5G or even a 6G core network on satellites, which is an essential and significant step forward.

Deploying the core network on satellites would be of great benefit. On the one hand, LEO satellites are made of general-purpose hardware and utilize onboard computing, network, and storage resources based on custom software. By interconnecting the core network with the access network and data network, real-time and reliable customized services can be provided to users. On the other hand, given the alarming rate at which the number of satellites in space is growing, a flexible and efficient unified management method for satellites is urgently needed. By deploying a lightweight core network on satellites, onboard resource scheduling and satellite unified management can become more convenient and flexible.

The development of the satellite core network and the decision-making process on which essential network elements are to be deployed on the satellite are crucial to the successful implementation of the satellite network. Thus, building satellite networks comes with numerous challenges. Nevertheless, these obstacles indicate promising avenues for research in communication [80], [81], virtualization, lightweight core network deployment, and other aspects of satellite network construction.

1) Satellite Communication: An important difference between satellite networks and ground networks is the lack of stable communication links and channels. Satellite links are vulnerable to various factors such as weather, signal-to-noise ratio, power, and antenna gain. Therefore, reducing path loss is a crucial consideration in satellite communication. Beamforming is a common method used to improve the quality of service and expand coverage. However, inter-satellite and satellite-ground communications face challenges such as long distance and interference, and path loss is more significant. Therefore, more effective access solutions are urgently needed to reduce path loss.

Another important difference between satellite and ground networks is the limited resources. Onboard computing, network, and storage resources are often limited by the satellite's design and operation. Effectively utilizing the limited onboard resources is key to improving the performance of satellite networks, in which spectrum resources play a critical role.

Spectrum is a limited natural resource, and several fixed frequency bands have been allocated for satellite communications [82]. However, in the era of 5G, B5G, and even 6G, massive terminal access, ultra-high data speed, and significant capacity requirements are bound to put more pressure on the existing spectrum. Cognitive radio and millimeter waves are effective technologies to improve the efficiency of spectrum utilization and deserve further research.

Communication delay is a crucial factor affecting user service experience. Satellite communication is affected by the objective environment and channel quality, and the delay is usually substantial. Therefore, it is of great significance to reduce the delay of satellite networks further. Satellite communication delay is usually caused by transmission, propagation, queuing, and processing. The high dynamic and time-varying characteristics of satellite topology increase the instability of communication delay. Improving the onboard computing power will enable user services to be processed on the satellite, significantly reducing the delay.

The frequent movement of satellites raises the problem of service migration. The ground area covered by LEO satellites is not fixed, and the high-speed movement of LEO satellites results in frequent switching of satellite networks. Therefore, mobility management is essential to ensure the continuity of satellite communication and meet the QoS requirements of users. Computing offloading is also a problem caused by satellite movement and limited onboard resources. Satellites need to offload some computing tasks to other eligible satellites or ground base stations to improve QoS and nearby service capabilities. Dynamic unloading schemes are the focus of research on satellite communication. In addition, congestion control and delay are also challenges faced in the computation offloading process.

Access to LEO satellites by a large number of users raises security concerns. Some user data and information are security-sensitive and private, which brings inevitable security problems. The design of the security mechanism has to face the challenges of an open satellite environment and highly dynamic topology. Inter-satellite communications are prone to interference or eavesdropping, which is also the main security challenge facing satellite communications today. Mature data encryption techniques have been developed. However, in order to further adapt to satellite communication scenarios and avoid the increase in delay caused by complex keys, distributed key calculation and management have attracted the attention of researchers. Additionally, the authentication mechanism and the security routing mechanism should be redesigned [83].

2) Virtualization of Satellite Networks: As mentioned above, onboard resources and energy are typically limited. To fully utilize the potential of satellites, onboard resources and network functions should be arranged flexibly and efficiently. Network virtualization technology is an effective means to improve network flexibility and points towards the direction for the construction of future network architecture. Network function virtualization (NFV) and software-defined networking (SDN) are typical applications of network virtualization, which can provide a reference for the architecture design of satellite constellations [84]. However, the development of satellite networks is not keeping up with the trend of network virtualization.

The separation of the SDN control plane and data plane helps to achieve efficient and flexible network management. The SDN-based LEO satellite constellation can adopt the design concept of a control plane and data plane. Deploy network controller in the control plane to realize data processing and transmission. Deploy infrastructure such as satellites, gateways, and switches on the data plane. The control plane performs integrated management of the whole satellite network, and the data plane is used to collect various basic equipment and terminal data. The separation of the control plane and the data plane of the LEO satellite constellation makes computing offloading, routing decisions, and resource allocation more independent and flexible.

Since satellites are made of general-purpose hardware, the invocation and implementation of onboard functions depend on the dedicated interfaces of specific suppliers, which limits the flexibility of satellite architecture and the programmability of onboard resources. NFV shifts the realization of network functions from hardware dependence to a software programming paradigm. One view is that SDN can effectively address some of the challenges associated with NFV. The extensive application of virtualization technology in cloud computing platforms can inspire the construction of NFV-based satellite networks.

Network resource allocation is a core business of network virtualization, called virtual network embedding satellite networks [85]. Differences in the design, structure, and characteristics of satellite networks and ground networks determine that the management of satellite network resources is not easy. Onboard computing resources are extremely limited. Network resources such as storage, spectrum, and energy also need to be properly orchestrated. SDN/NFV-based satellite network resource management is an effective solution to the above challenges and deserves further research.

3) Lightweight Core Network [29]: The integration and evolution of 5G and satellite networks have become an inevitable development trend [86]. To improve the access capability of satellite networks, the core network functions can be deployed on satellites. However, directly deploying the huge and complex core network functions on the satellite may limit the flexibility of the satellite and increase the difficulty of management. Additionally, the large amount of signaling that cannot be handled directly over the satellite networks increases network latency. As mentioned above, it is necessary to deploy core network functions on satellites and tailor them appropriately to provide flexible management control services, as shown in Fig. 1. The reason for tailoring the core network is that the ground core network architecture is too large and complex, and it is unrealistic to directly deploy it on resource-constrained satellites. Deploying the lightweight core network directly on the satellite can endow the satellite network with basic core capabilities such as access management, mobility management, and session management, thereby optimizing network management and reducing latency [87], [88]. In addition, the onboard core network is expected to provide more key technologies and solutions for the vertical industry [89]. It

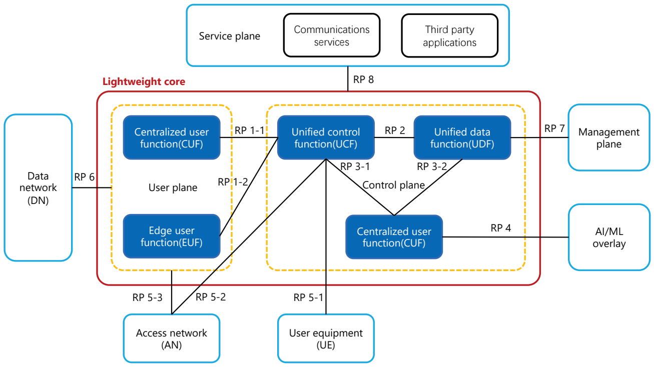

Fig. 1: Architecture Reference Model of Lightweight Core Network.

cannot be ignored that the research on core network satellites is still in its infancy. Which network functions need to be deployed on satellites? How to simplify network functions? What impact will the deployment of the core network on satellites bring to the satellite network's architecture? These questions need to be answered.

4) Others: The combination of satellite constellations and other fields provides researchers with numerous research options. Hardware platforms and distributed system design are all starting points for future satellite network construction. Considering the inherent characteristics of satellite networks, the optimization of satellite network performance will be more complex than ground network optimization, including resource management, cost, energy efficiency, etc. For example, inter-satellite or inter-satellite-ground communication links are susceptible to interference and become unstable. The commonly used method is to combine beamforming and distributed control strategy to suppress communication interference. Another issue worth discussing is that with the wide deployment of satellite constellations, it is necessary to carry out experiments in real space environments based on satellite platforms, not just at the simulation level. Therefore, it is necessary to strengthen the hardware platform construction of satellite constellations, such as intelligent reflecting surfaces [90]. The satellite network also needs to provide services such as network spatiotemporal change information (data traffic, location, ephemeris) and target positioning.

B. AI in Space

With the construction of Starlink and OneWeb, our focus has shifted towards space exploration [2]. Remote sensing satellites capture vast amounts of space images from various angles to explore unknown information in space. However, traditional methods of transmitting these images to the ground for analysis face two challenges [42], [91]: limited uplink and downlink bandwidth, especially since the uplink bandwidth is only in the range of tens to hundreds of kb/s and the transmission process has a large latency, and the transmission process is relatively fragile, and the transmission may be interrupted, leading to loss of data.

Satellites have become more powerful with the development of satellite technology, and they can process information onboard. Therefore, deploying AI models on satellites can save communication costs and improve responsiveness by processing data onboard instead of transmitting it to the ground station. However, due to the limited size of a single satellite, the computing ability of the deployed models is also limited. Processing high-resolution earth imagery in real-time is a typical application scenario in LEO satellite applications, but it is challenging to process such large resolution images with a small model. Therefore, deploying a single AI model in a single satellite is unrealistic, and transmitting all images to the ground station is not feasible. Instead, a distributed learning framework deployment on LEO is promising [92].

However, traditional distributed learning deployment is not practical in LEO satellites because inter-satellite communication is currently not possible, and even if satellites could communicate in the future, the amount of memory and power required to do so would quickly become the bottleneck. Therefore, Federated Learning (FL) [93]–[95] as a special distributed learning paradigm has great prospects for widespread deployment in satellites. In this way, we can directly process and analyze the collected information on the satellite while protecting the privacy of each satellite's data [96]. We can also deploy a large model in the ground station and use this model to assist a shallow model deployed on the satellite. However, FL, like traditional deep learning systems, often demonstrate incorrect or unexpected corner-case behaviors, especially in the harsh space environment. Therefore, a systematic testing tool is necessary for automatically detecting erroneous behaviors of FL-driven satellites that can potentially lead to data invalidation analysis [97]. This tool can automatically generate test cases leveraging real-world changes in the space environment (such as meteorites, lighting conditions, shooting angle, etc.) to retrain the corresponding FL model and improve the model's robustness.

C. Satellite and Smart City

With the increasing availability of remote sensing (RS) data from satellites, urban analysis and management have seen significant progress. For instance, meteorological satellites can capture climate change, improving weather prediction accuracy, while synthetic aperture radar imaging satellites can identify urban buildings and roads, aiding in urban planning. However, utilizing RS data effectively and efficiently for urban computing faces several challenges. Firstly, the single view/modality of RS data cannot fully represent the entire city, necessitating fusion with data from other sources that has not been extensively studied. Secondly, the low spatio-temporal resolution of RS images limits their applicability in real-time scenarios such as urban traffic perception and prediction. We discuss research opportunities that address these issues using AI technologies and data management.

1) Urban Satellite-Ground Data Fusion: Ground data primarily collected from sensors and GPS devices do not cover all the spatial and temporal dimensions of a city. Satellite data can, therefore, enrich the available information. For instance, Wu *et al.* [98] used the attention mechanism to fuse RS images and GPS trajectories for automatic road network generation. However, consolidating satellite-ground data faces two main

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

challenges. Firstly, the quality of urban satellite-ground data may not meet application requirements due to data errors, deficiencies, and noise. Hence, data cleaning is a primary problem to be solved. To do this, we can utilize satellite data to verify ground data and vice versa, and leverage advanced data management technologies such as active learning and crowdsourcing to correct errors. Secondly, urban satellite-ground data have different modalities such as texts, images, and videos, which requires different models for encoding data in different modes. Additionally, there are two ways to fuse them based on the order of data integration in the whole process. Pre-fusion involves concatenating encoding results (e.g., vector representations) for downstream applications, while post-fusion independently applies each mode of data for each task, then fuses task results as the final result of downstream applications.

2) Data Augmentation for Urban RS: The low spatio-temporal resolution of urban RS images limits their applicability in real-time scenarios. Enhancing RS data with data augmentation technology presents a promising solution. Existing methods mainly focus on mapping from low spatial resolution to high spatial resolution, such as detecting changes via mapping multi-spatial-resolution RS images [99]. However, mapping-based methods require sufficient low-high resolution pairs, making them impractical in urban applications. We can convert the data augmentation problem into a generation problem solved using generative models like generative adversarial networks [100]. Prior knowledge, such as the road network topology and the first theorem of geography, should be considered in the generation process. Additionally, we can utilize urban data collected from other domains to enhance RS data. For example, detecting road lines in street view photos is useful when the related RS image has low resolution.

D. Satellite Operation System

Satellite software is advancing rapidly, and it needs innovative hardware to support it. The satellite components are becoming more generalized. They no longer rely solely on radiation-resistant devices like the CPU, DRAM, and FPGA. The use of commercial computer devices to build the satellite platform has provided a solid foundation for the development of satellite operating systems [101]. However, over the past few decades, existing satellite operating systems have mainly focused on stability and real-time performance, with no changes to basic functions and structures. As a result, these systems are unable to keep up with the new trend of satellite software upgrades and hardware generalization.

Currently, there has been some industry research on the new generation of satellite operating systems. Lockheed Martin has come up with SmartSat, an operating system that is CubeSat-based [6]. Satellites that are equipped with SmartSat will create a space cloud computing platform with on-board data processing capabilities. The satellite capabilities are updated through software that is uploaded from the ground. SmartSat is centered around the application store and provides a relatively comprehensive solution for the application development process. However, SmartSat does not take into account the use of commercial components for assembling the on-board computer. Additionally, it does not provide any solutions for fault tolerance, improving operating system performance, and scheduling resources.

HPE has equipped Space-born Computer [7] with a server operating system that considers the performance improvement and fault tolerance of servers using commercial devices. However, this system is not specifically designed for the satellite environment, and the cost of fault tolerance is quite high, which puts it far behind the actual deployment of satellite operating systems.

In the past, SpaceX has launched Starlink satellites that carry a stripped-down version of the Linux operating system. This solution takes into account multiple factors such as compatibility with the ecosystem, improving the performance of the operating system, and maintaining its real-time performance. However, the system relies on radiation-resistant CPUs and does not use commercial devices, resulting in a relatively limited system flexibility.

We believe that the future satellite operating system should possess the following characteristics. Firstly, in the past, on-board computers were designed for stability and real-time performance. However, with the new requirements for computing power and the use of commercial devices to build the on-board computer, the satellite operating system needs to consider performance factors as well. It should also be compatible with the existing Linux software ecosystem, allowing for smooth migration of existing software to satellites. Secondly, the on-board computers are built with radiation-resistant special devices, and the operating system is designed to be reliable in that environment. However, after using commercial devices to improve performance, the hardware can no longer provide security and reliability guarantees. Therefore, providing a fault-tolerant mechanism at the software level of the operating system is a key issue [102]. Additionally, the traditional satellite operating system has a long application development and deployment cycle. Once an application is deployed, it is difficult to update, and development between different satellites cannot be migrated, leading to extremely high development costs. Therefore, providing a set of mechanisms for development, testing, deployment, and dynamic adjustment on the satellite operating system is of great significance. Lastly, in the scenario where a multitude of tasks are performed on satellites, such as AI image inference, video transmission, and resource monitoring programs, it is also necessary to consider ensuring the isolation between different tasks and providing scheduling capabilities based on the importance of each task.

V. CASE STUDY

In order to promote research on satellite computing, we established the Tiansuan constellation as an open research platform [5], illustrated in Fig. 2. Tiansuan supports on-board computing services, satellite operating systems, 6G core network systems, and federated learning as well as AI acceleration. A total of 300 satellites will be launched in

---

[6] https://www.lockheedmartin.com/

[7] https://www.hpe.com/us/en/home.html

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

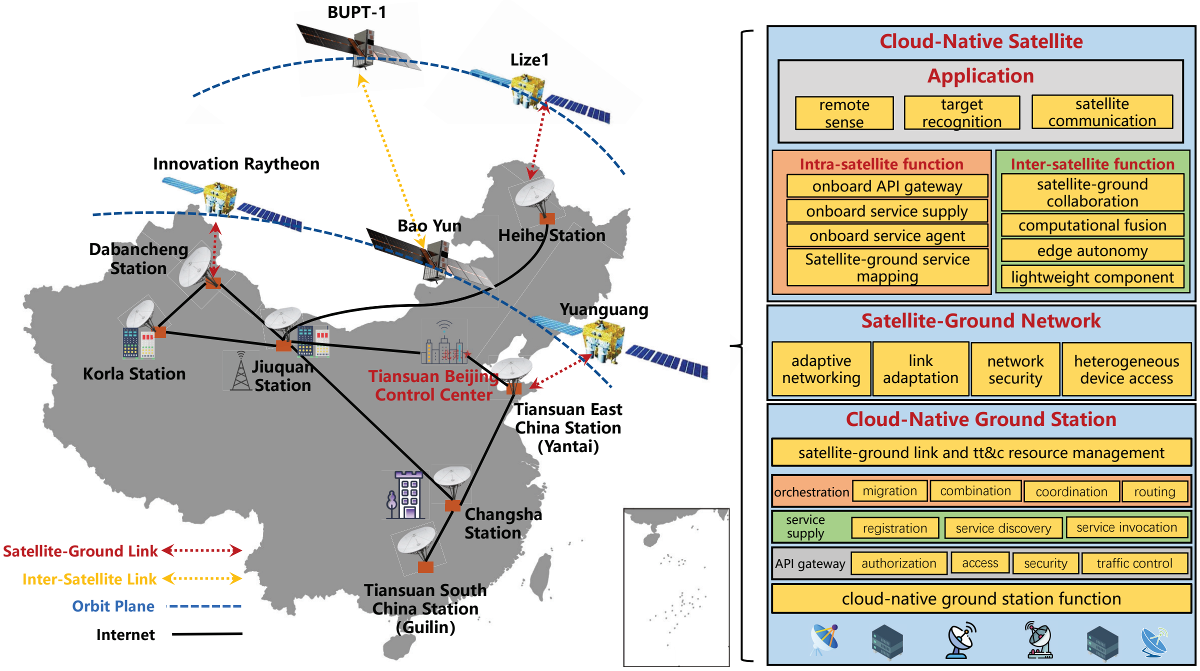

Fig. 2: Tiansuan Constellation Platform Architecture.

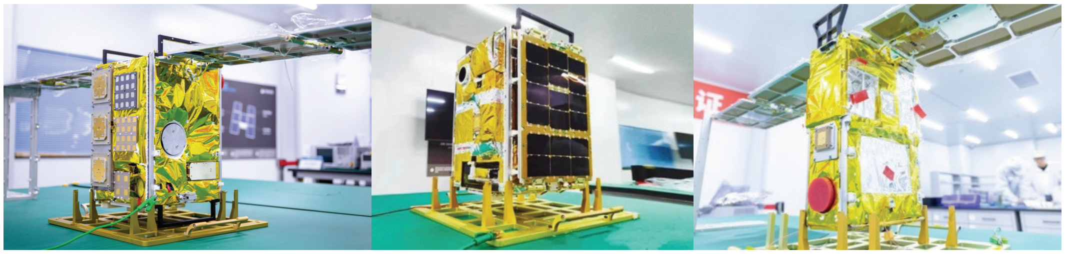

Fig. 3: The picture of Baoyun, BUPT-1, and Innovation Raytheon (from left to right).

TABLE II: Main parameters of Tiansuan Constellation Phase 1.

Number	Orbital Altitude	Mass	Battery Capacity	Spectrum	Uplink Rate	Downlink Rate	ISLs	Processors
1	500±50km	$\leq$ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
2	500±50km	$\leq$ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
3	500±50km	$\leq$ 30kg	118Wh – 236Wh	X-band	0.1Mbps – 1Mbps	100Mbps – 600Mbps	NO	CPU/NPU
4	$>$ 500km	$>$ 50kg	$>$ 360Wh	X, Ku, Ka bands	$\geq$ 200Mbps	$\geq$ 1Gbps	YES	CPU/NPU/GPU
5	$>$ 500km	$>$ 50kg	$>$ 360Wh	X, Ku, Ka bands	$\geq$ 200Mbps	$\geq$ 1Gbps	YES	CPU/NPU/GPU
6	$>$ 500km	$>$ 50kg	$>$ 360Wh	X, Ku, Ka bands	$\geq$ 200Mbps	$\geq$ 1Gbps	YES	CPU/NPU/GPU

three phases. This will include the first phase with 6 satellites, the second phase with 24 satellites, and the third phase with 300 satellites. Table II lists the parameters of satellites in the first phase. As of July 2023, the Tiansuan constellation has launched five satellites, Baoyun, Innovation Raytheon, Lize1, Yuanguang, and BUPT-1 as shown in Fig. 3. Another satellite, Wangqizhou, will also be launched this year. In Tiansuan, satellites are manufactured according to [103]. Most satellites will be placed in sun-synchronous orbit. For the first three satellites in the first phase, edge computing capabilities with remote sensing applications are tested. Communication capabilities with inter-satellite links will be explored with the last three satellites. The on-board payloads provide the majority of the computing power as listed in TABLE II. As shown in Fig. 2, ground stations are gateways for satellites to transmit data. By integrating with the private data centers, ground stations can offer the received data through the Internet to users. We have deployed many ground stations, distributed in Hunan, Xinjiang, and Heilongjiang Provinces. Next, we introduce several use cases deployed in the Tiansuan constel-

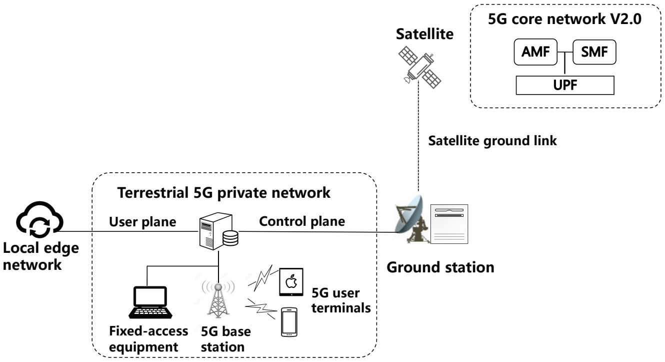

Fig. 4: Satellite-Borne B5G Core Network.

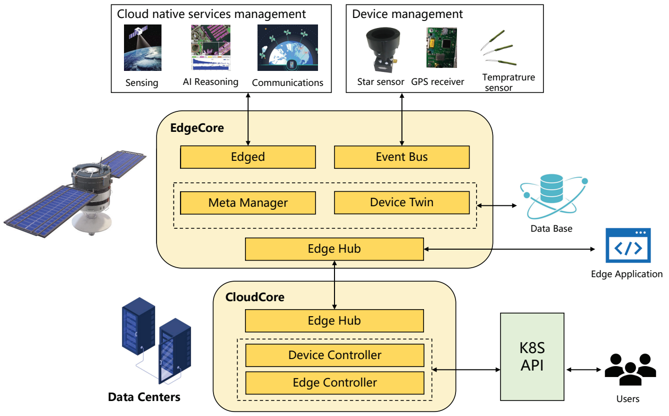

Fig. 5: Satellite-Ground Collaboration of the Cloud-Native Satellite.

lation.

A. Satellite-borne B5G Core Network

On August 9, 2021, we conducted the first deployment test of a 5G core network system on the experimental satellite (TY20) of Tiansuan. The experimental network consists of the onboard B5G core network and the terrestrial private 5G network. The on-board B5G core network implements three network functions (AMF, SMF, and UPF). The network functions support essential system procedures such as user registration and session establishment. By incorporating the terrestrial full-fledged 5G network, we test the signaling interaction between the control plane and the user plane. Downlink telemetry showed that the three network functions were operating normally. It also showed that the control signaling was generated correctly. The control signaling was then transmitted to the terrestrial private 5G network. It triggered the local computation offloading controlled by satellites. We conducted further tests such as video calls. As shown in Fig. 4, we have successfully deployed lightweight B5G core networks on both the Baoyun and Innovation Raytheon satellites. This core network is the updated version of the former 5G core network. It improves signaling interactions and can be used to set up video calls based on the Session Initiation Protocol. After the two satellites are launched, functional and performance tests will be conducted. More details about this work can refer to [29].

B. Cloud-Native Satellite

Baoyun was successfully launched on 7 December 2021, which carried the computing payload of Tiansuan Constellation. Fig. 5 shows the first cloud-native satellite that integrates satellite and ground computing capabilities to compete the space tasks. The satellites run smoothly and provide service in orbit. In addition, Huawei Cloud is the first time to make “Cloud-Edge Synergy” come true in space by cloud-native satellite. In the future, the “Cloud-Edge Synergy” scheme will be deployed in 6 satellites in the first phase of the Tiansuan constellation to form a unified computing network collaboratively. We expect cloud-native satellites to create new capabilities for emergency communication, ecological monitoring, disaster prevention and mitigation, urban construction, etc.

As the first cloud-native satellite, Baoyun leverages the space computing resources and central cloud on the ground station and consequently they have AI and multi-tasking capabilities. For on-orbit AI inference application, low-quality image data may be discarded to lower transmission overhead due to over 50% coverage by clouds. After useful image data is transmitted back, a high-precision model deployed at the ground station with abundant computing resources is used to the following computing, which further shortens the processing time from usual about one day to even one hour. Besides, these models in space can be updated when needed. Cloud-native satellites will be an integral part in our daily lives. For example, we compare the images before and after rainstorms to identify the risk of mountain collapse through on-orbit AI inference, which guides us to discover hidden danger and prevent catastrophe in advance.

BUPT-1, which is the first primary satellite of Tiansuan, was successfully launched and operated on January 15th, 2023. Over a period of 15 days, we conducted several verifications to confirm and quantify the performance gains of our cloud-native system on BUPT-1. During this time, we tested remote sensing image AI inference, real-time video streaming transmission between the satellite and ground, and other key functionalities. Based on telemetry from BUPT-1, we found that the average response time for the end-to-end measurement and control service between satellite and ground was 8 seconds. Additionally, telemetry data collected from the on-board program running status field revealed that a significant improvement in the utilization rate of onboard services compared to non-cloud-native platforms. We package the service programs and relevant dependencies into containers, and deploy them uniformly on satellites. Containerized deployment ensures the elasticity and scalability of the services, allowing for dynamic resource adjustments based on actual needs and automatic addition or removal of instances when necessary. On the ground, the required services can be flexibly launched on demand through remote control commands. In contrast, non-cloud-native platforms can only provide customized services, occupying limited on-board resources for extended periods during program runtime, resulting in a significant decrease in service utilization. More details about this work can refer to

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

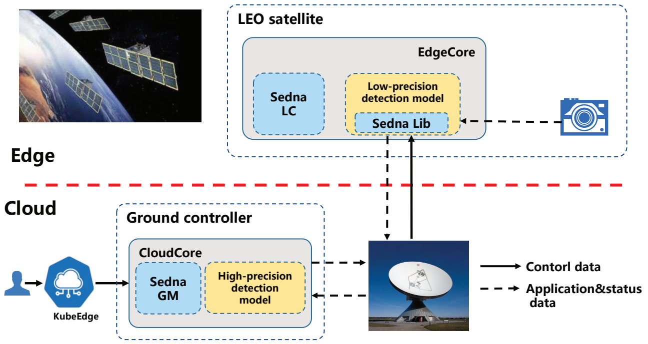

Fig. 6: Satellite-Ground Collaborative Image Inference based on KubeEdge.

[104].

C. Satellite-Ground Collaborative Image Inference

We have implemented a satellite-ground collaborative inference application based on KubeEdge and its AI extension, Sedna, on both Baoyun and Innovation Raytheon as shown in Fig. 6. To communicate with the satellite, we modified the KubeEdge runtime (CloudCore) to build a central ground controller upon a Linux server, which communicates with the satellite intermittently according to the position of the satellite. We later deployed two image detection models with different precisions, a low-precision model and a high-precision model. The low-precision image detection model is deployed inside the EdgeCore of the satellite and used to detect whether the captured image is of interest. Once the detection on the low-precision model achieves high confidence, the satellite will use the concrete result to facilitate later processing (e.g., calculate total the area of agricultural lands). When the confidence is low, the satellite will download the captured image to the ground controller. The ground controller will detect the objects using the high-precision model and then upload the exact result to the satellite. By fully leveraging the computation capacity on the satellite, we can reduce the image detection latency from days to minutes in such a collaborative computing paradigm without accuracy cost. In addition, we propose an on-board image segmentation solution to mitigate the occurrence of redundant images due to cloud occlusion. Performing image segmentation before on-orbit inference helps reduce the computational load and save bandwidth.

D. Quic Protocol based Real-Time Transmission of the Satellite-Ground Link

On August 25, 2022, Tiansuan successfully completed the world's first QUIC (Quick UDP Internet connections) protocol based real-time transmission test of the satellite-ground link. This test is mainly based on the satellite-ground integrated distributed network verification platform jointly developed by Huawei Cloud and our scientific research team. The test was started on May 18, 2022. At 11:23 on June 18, 2022, multiple real-time transmission tests of the QUIC protocol of the satellite-ground link were completed. On August 24, 2022, the data analysis was completed. This test realized end-to-end data transmission based on the QUIC protocol. The remote control command was initiated by the ground station in Changsha, Hunan Province, and the onboard client started to initiate transmission. The ground station in Tongchuan, Shanxi Province received the data at high transmission speed and analyzed the QUIC traffic. For the QUIC protocol based real-time transmission of the satellite-ground link, we developed a customized QUIC server logic to achieve compatibility and a data frame analysis system following the AOS (Advanced On-orbit System) protocol issued by the CCSDS (Consultative Committee for Space Data Systems) to achieve QUIC protocol parsing enhancement.

VI. CONCLUSION

LEO satellite constellations are experiencing rapid development and satellite computing is promising to address the limitation of traditional satellite bent-pipe architecture and provide computing services for ground users. In this paper, we first analyze the current development of LEO constellations. Then, we survey recent work in the field of satellite computing and discuss the research challenges. We also put forward research opportunities that are worth working on, naming networking, computing, smart city, satellite operating system. Finally, we introduce our open research platform on real satellites named Tiansuan constellation and several experiments deployed on it. Our next step is to establish a larger open-source satellite constellation, allowing more people to access and study satellite technology. Our ultimate aim is to serve the entire human population by democratizing access to satellite computing.

REFERENCES

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

[2] B. Denby and B. Lucia, "Orbital edge computing: nanosatellite constellations as a new class of computer system," in *Proceedings of International Conference on Architectural Support for Programming Languages and Operating Systems*, 2020, pp. 939–954.

[3] V. Singh, A. Prabhakara, D. Zhang, O. Yağan, and S. Kumar, "A community-driven approach to democratize access to satellite ground stations," *GetMobile: Mobile Computing and Communications*, vol. 26, no. 1, pp. 35–38, 2022.

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

[5] S. Wang, Q. Li, M. Xu, X. Ma, A. Zhou, and Q. Sun, "Tiansuan constellation: An open research platform," in *Proceedings of IEEE International Conference on Edge Computing*, 2021, pp. 94–101.

[6] 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 ACM Workshop on Hot Topics in Networks*, 2018, pp. 113–119.

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

[8] D. Bhattacherjee, W. Aqeel, G. Laughlin, B. M. Maggs, and A. Singla, "A bird's eye view of the world's fastest networks," in *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 521–527.

[9] S. Kassing, D. Bhattacherjee, A. B. Águas, J. E. Saethre, and A. Singla, "Exploring the" internet from space" with hypatia," in *Proceedings of ACM Internet Measurement Conference*, 2020, pp. 214–229.

[10] G. Giuliari, T. Klenze, M. Legner, D. Basin, A. Perrig, and A. Singla, "Internet backbones in space," *ACM SIGCOMM Computer Communication Review*, vol. 50, no. 1, pp. 25–37, 2020.

[11] Q. Li, S. Wang, X. Ma, Q. Sun, H. Wang, S. Cao, and F. Yang, "Service coverage for satellite edge computing," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 695–705, 2021.

[12] Y. Hauri, D. Bhattacherjee, M. Grossmann, and A. Singla, ""internet from space" without inter-satellite links," in *Proceedings of ACM Workshop on Hot Topics in Networks*, 2020, pp. 205–211.

[13] T. Kur and T. Liwosz, "Simulation of the use of variance component estimation in relative weighting of inter-satellite links and gnss measurements," *Remote Sensing*, vol. 14, no. 24, p. 6387, 2022.

[14] C. H. Arslan, K. Aydin, J. V. Urbina, and L. Dyrud, "Satellite-link attenuation measurement technique for estimating rainfall accumulation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 681–693, 2018.

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

[16] Z. Lai, Q. Wu, H. Li, M. Lv, and J. Wu, "Orbitcast: exploiting mega-constellations for low-latency earth observation," in *Proceedings of IEEE International Conference on Network Protocols*, 2021, pp. 1–12.

[17] S. Zhang and K. L. Yeung, "Scalable routing in low-earth orbit satellite constellations: architecture and algorithms," *Computer Communications*, vol. 188, pp. 26–38, 2022.

[18] M. Handley, "Using ground relays for low-latency wide-area routing in megaconstellations," in *Proceedings of the 18th ACM Workshop on Hot Topics in Networks*, 2019, pp. 125–132.

[19] S. Wang, Y. Zhao, and H. Xie, "Improving survivability of leo satellite network with guaranteed based approaches," in *Proceedings of IEEE Symposium on Computers and Communications*, 2020, pp. 1–6.

[20] Z. Zhao, Q. Wu, H. Li, Z. Lai, and J. Liu, "Lrar: a lightweight risk-avoidance routing algorithm for leo satellite networks," in *2021 International Wireless Communications and Mobile Computing*, 2021, pp. 223–228.

[21] P. S. Page, K. S. Bhargao, H. V. Baviskar, and G. S. Kasbekar, "Distributed probabilistic congestion control in leo satellite networks," in *Proceedings of the International Conference on COMmunication Systems and NETworks*, 2023, pp. 335–339.

[22] V.-P. Bui, T. V. Chien, E. Lagunas, J. Grotz, S. Chatzinotas, and B. Ottersten, "Robust congestion control for demand-based optimization in precoded multi-beam high throughput satellite communications," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6918–6937, 2022.

[23] M. Liu, Y. Liu, Z. Ma, Z. Porter, J. Chung, S. Claypool, F. Li, J. Tutlis, and M. Claypool, "The effects of a performance enhancing proxy on tcp congestion control over a satellite network," in *Proceedings of the IEEE International Performance, Computing, and Communications Conference*, 2022, pp. 325–331.

[24] S. Claypool, J. Chung, and M. Claypool, "Comparison of tcp congestion control performance over a satellite network," in *Proceedings of the International Conference on Passive and Active Network Measurement*, 2021, pp. 499–512.

[25] S. Dai, L. Rui, S. Chen, and X. Qiu, "A distributed congestion control routing protocol based on traffic classification in leo satellite networks," in *Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management*, 2021, pp. 523–529.

[26] Z. A. Qazi, M. Walls, A. Panda, V. Sekar, S. Ratnasamy, and S. Shenker, "A high performance packet core for next generation cellular networks," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017, pp. 348–361.

[27] M. Moradi, K. Sundaresan, E. Chai, S. Rangarajan, and Z. M. Mao, "Skycore: moving core to the edge for untethered and reliable uav-based lte networks," in *Proceedings of the Annual International Conference on Mobile Computing and Networking*, 2018, pp. 35–49.

[28] Y. Li, H. Li, W. Liu, L. Liu, Y. Chen, J. Wu, Q. Wu, J. Liu, and Z. Lai, "A case for stateless mobile core network functions in space," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2022, pp. 298–313.

[29] R. Xing, X. Ma, A. Zhou, S. Dustdar, and S. Wang, "From earth to space: A first deployment of 5g core network on satellite," *China Communications*, vol. 20, no. 4, pp. 315–325, 2023.

[30] T. Darwish, G. K. Kurt, H. Yanikomeroglu, M. Bellemare, and G. Lamontagne, "Leo satellites in 5g and beyond networks: a review from a standardization perspective," *IEEE Access*, vol. 10, pp. 35 040–35 060, 2021.

[31] B. Al Homssi, A. Al-Hourani, K. Wang, P. Conder, S. Kandeepan, J. Choi, B. Allen, and B. Moores, "Next generation mega satellite networks for access equality: opportunities, challenges, and performance," *Proceedings of IEEE Communications Magazine*, vol. 60, no. 4, pp. 18–24, 2022.

[32] J. Li, H. Li, J. Liu, Z. Lai, Q. Wu, and X. Wang, "A timeslot division strategy for availability in integrated satellite and terrestrial network," in *Proceedings of IEEE Wireless Communications and Networking Conference*, 2021, pp. 1–7.

[33] S. Ji, D. Zhou, M. Sheng, and J. Li, "Mega satellite constellation system optimization: from a network control structure perspective," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 913–927, 2022.

[34] Y. Chen, H. Li, J. Liu, Q. Wu, and Z. Lai, "Gams: an ip address management mechanism in satellite mega-constellation networks," in *Proceedings of International Wireless Communications and Mobile Computing*, 2021, pp. 229–234.

[35] Z. Lai, H. Li, Q. Zhang, Q. Wu, and J. Wu, "Cooperatively constructing cost-effective content distribution networks upon emerging low earth orbit satellites and clouds," in *Proceedings of International Conference on Network Protocols*, 2021, pp. 1–12.

[36] Z. Lai, H. Li, and J. Li, "Starperf: characterizing network performance for emerging mega-constellations," in *Proceedings of International Conference on Network Protocols*, 2020, pp. 1–11.

[37] F. Tang, "Dynamically adaptive cooperation transmission among satellite-ground integrated networks," in *Proceedings of IEEE International Conference on Computer Communications*, 2020, pp. 1559–1568.

[38] L. Chen, F. Tang, and X. Li, "Mobility- and load-adaptive controller placement and assignment in leo satellite networks," in *Proceedings of IEEE International Conference on Computer Communications*, 2021.

[39] T. Klenze, G. Giuliari, C. Pappas, A. Perrig, and D. Basin, "Networking in heaven as on earth," in *Proceedings of ACM Workshop on Hot Topics in Networks*, 2018, pp. 22–28.

[40] X. Li, F. Tang, J. Liu, L. T. Yang, L. Fu, and L. Chen, "Auto: adaptive congestion control based on multi-objective reinforcement learning for the satellite-ground integrated network," in *Proceedings of the USENIX Annual Technical Conference*, 2021, pp. 611–624.

[41] V. Singh, O. Yagan, and S. Kumar, "Selfiestick: towards earth imaging from a low-cost ground module using leo satellites," in *Proceedings of the ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2022, pp. 220–232.

[42] D. Vasisht, J. Shenoy, and R. Chandra, "L2d2: low latency distributed downlink for leo satellites," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, 2021, pp. 151–164.

[43] S. Liu, T. Wu, Y. Hu, Y. Xiao, D. Wang, and L. Liu, "Throughput evaluation and ground station planning for leo satellite constellation networks," in *Proceedings of International Conference on Space Information Network*, 2020, pp. 3–15.

[44] K. Guo and K. An, "On the performance of ris-assisted integrated satellite-uav-terrestrial networks with hardware impairments and interference," *IEEE Wireless Communications Letters*, vol. 11, no. 1, pp. 131–135, 2021.

[45] D. Xia, X. Zheng, P. Duan, C. Wang, L. Liu, and H. Ma, "Ground-station based software-defined leo satellite networks," in *Proceedings of IEEE International Conference on Parallel and Distributed Systems*, 2019, pp. 687–694.

[46] L. Han, A. Retana, C. Westphal, R. Li, T. Jiang, and M. Chen, "New ip based semantic addressing and routing for leo satellite networks," in *Proceedings of the IEEE International Conference on Network Protocols*, 2022, pp. 1–6.

[47] H. Qi, Y. Guo, D. Hou, Z. Xing, W. Ren, L. Cong, and X. Di, "Sdn-based dynamic multi-path routing strategy for satellite networks," *Future Generation Computer Systems*, vol. 133, pp. 254–265, 2022.

[48] P. Zhao, B. Peters, J. Chung, and M. Claypool, "Competing tcp congestion control algorithms over a satellite network," in *Proceedings of the IEEE Annual Consumer Communications and Networking Conference*, 2022, pp. 132–138.

[49] X. Cao, B. Yang, Y. Shen, C. Yuen, Y. Zhang, Z. Han, H. V. Poor, and L. Hanzo, "Edge-assisted multi-layer offloading optimization of leo satellite-terrestrial integrated networks," *EEE Journal on Selected Areas in Communications*, vol. 41, no. 2, pp. 381–398, 2023.

[50] P. Cassará, A. Gotta, M. Marchese, and F. Patrone, "Orbital edge offloading on mega-leo satellite constellations for equal access to computing," *IEEE Communications Magazine*, vol. 60, no. 4, pp. 32–36, 2022.

[51] T. Pfandzelter, J. Hasenburg, and D. Bermbach, "Towards a computing platform for the leo edge," in *Proceedings of International Workshop on Edge Systems, Analytics and Networking*, 2021, pp. 43–48.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

[52] T. Pfandzelter and D. Bermbach, “Celestial: virtual software system testbeds for the leo edge,” in *Proceedings of the ACM/IFIP International Middleware Conference*, 2022, pp. 69–81.

[53] Pfandzelter, Tobias and Bermbach, David, “Qos-aware resource placement for leo satellite edge computing,” in *Proceedings of International Conference on Fog and Edge Computing*, 2022, pp. 66–72.

[54] R. Xie, Q. Tang, Q. Wang, X. Liu, F. R. Yu, and T. Huang, “Satellite-terrestrial integrated edge computing networks: architecture, challenges, and open issues,” *IEEE Network*, vol. 34, no. 3, pp. 224–231, 2020.

[55] M. Tong, X. Wang, S. Li, and L. Peng, “Joint offloading decision and resource allocation in mobile edge computing-enabled satellite-terrestrial network,” *Symmetry*, vol. 14, no. 3, p. 564, 2022.

[56] X. Gao, R. Liu, and A. Kaushik, “Virtual network function placement in satellite edge computing with a potential game approach,” *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 1243–1259, 2022.

[57] Q. Tang, Z. Fei, B. Li, and Z. Han, “Computation offloading in leo satellite networks with hybrid cloud and edge computing,” *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9164–9176, 2021.

[58] T. Kim, J. Kwak, and J. P. Choi, “Satellite edge computing architecture and network slice scheduling for iot support,” *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14 938–14 951, 2021.

[59] X. Fang, W. Feng, T. Wei, Y. Chen, N. Ge, and C.-X. Wang, “5g embraces satellites for 6g ubiquitous iot: Basic models for integrated satellite terrestrial networks,” *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14 399–14 417, 2021.

[60] J. Chu and X. Chen, “Robust design for integrated satellite–terrestrial internet of things,” *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9072–9083, 2021.

[61] X. Zhu and C. Jiang, “Integrated satellite-terrestrial networks toward 6g: Architectures, applications, and challenges,” *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 437–461, 2022.

[62] L. Zhen, A. K. Bashir, K. Yu, Y. D. Al-Otaibi, C. H. Foh, and P. Xiao, “Energy-efficient random access for leo satellite-assisted 6g internet of remote things,” *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5114–5128, 2021.

[63] J. Wei and S. Cao, “Application of edge intelligent computing in satellite internet of things,” in *Proceedings of IEEE International Conference on Smart Internet of Things*, 2019, pp. 85–91.

[64] Z. Song, Y. Hao, Y. Liu, and X. Sun, “Energy-efficient multiaccess edge computing for terrestrial-satellite internet of things,” *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14 202–14 218, 2021.

[65] P. Zhou, Y. Liu, L. Guo, N. Lu, J. Wu, and H. Jiang, “Handoff of satellite network for high-speed mobile terminals based on edge computing,” in *Proceedings of IEEE International Conference on Cyber Security and Cloud Computing*, 2021, pp. 167–175.

[66] B. Wang, X. Li, D. Huang, and J. Xie, “A profit maximization strategy of mec resource provider in the satellite-terrestrial double edge computing system,” in *Proceedings of IEEE International Conference on Communication Technology*, 2021, pp. 906–912.

[67] M. M. Gost, I. Leyva-Mayorga, A. Perez-Neira, M. A. Vazquez, B. Soret, Moretti, and Marco, “Edge computing and communication for energy-efficient earth surveillance with leo satellites,” in *Proceedings of the IEEE International Conference on Communications Workshops*, 2022, pp. 556–561.

[68] Q. Li, S. Wang, X. Ma, A. Zhou, and F. Yang, “Towards sustainable satellite edge computing,” in *Proceedings of IEEE International Conference on Edge Computing*, 2021, pp. 1–8.

[69] V. Kothari, E. Liberis, and N. D. Lane, “The final frontier: deep learning in space,” in *Proceedings of the International Workshop on Mobile Computing Systems and Applications*, 2020, pp. 45–49.

[70] J. Ding, N. Xue, G.-S. Xia, X. Bai, W. Yang, M. Y. Yang, S. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang, “Object detection in aerial images: a large-scale benchmark and challenges,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7778–7796, 2022.

[71] T. Hoeser, F. Bachofer, and C. Kuenzer, “Object detection and image segmentation with deep learning on earth observation data: a review - part ii: applications,” *Remote Sensing*, vol. 12, no. 10, p. 1667, 2020.

[72] M. Lofqvist and J. Cano, “Accelerating deep learning applications in space,” *arXiv preprint arXiv:2007.11089*, 2020.

[73] B. Uzkent, C. Yeh, and S. Ermon, “Efficient object detection in large images using deep reinforcement learning,” in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, 2020, pp. 1824–1833.

[74] A. Maskey and M. Cho, “Cubesatnet: ultralight convolutional neural network designed for on-orbit binary image classification on a 1u cubesat,” *Engineering Applications of Artificial Intelligence*, vol. 96, p. 103952, 2020.

[75] R. Wang, W. Zhu, G. Liu, R. Ma, D. Zhang, S. Mumtaz, and S. Cherkaoui, “Collaborative computation offloading and resource allocation in satellite edge computing,” in *Proceedings of the IEEE Global Communications Conference*, 2022, pp. 5625–5630.

[76] P. Li, Y. Wang, and Z. Wang, “A game-based joint task offloading and computation resource allocation strategy for hybrid edgy-cloud and cloudy-edge enabled leo satellite networks,” in *Proceedings of the IEEE/CIC International Conference on Communications in China*, 2022, pp. 868–873.

[77] T. Delamotte, M. G. Schraml, R. T. Schwarz, K.-U. Storek, and A. Knopp, “Multi-antenna-enabled 6g satellite systems: roadmap, challenges and opportunities,” *International ITG Workshop on Smart Antennas*, pp. 1–6, 2021.

[78] S. Narayana, R. V. Prasad, V. Rao, L. Mottola, and T. V. Prabhakar, “Hummingbird: Energy efficient gps receiver for small satellites,” pp. 1–13, 2020.

[79] E. Rapuano, G. Meoni, T. Pacini, G. Dinelli, G. Furano, G. Giuffrida, and L. Fanucci, “An fpga-based hardware accelerator for cnns inference on board satellites: benchmarking with myriad 2-based solution for the cloudscout case study,” *Remote Sensing*, vol. 13, no. 8, p. 1518, 2021.

[80] A. Gaber, M. A. ElBahaay, A. M. Mohamed, M. M. Zaki, A. S. Abdo, and N. AbdelBaki, “5g and satellite network convergence: Survey for opportunities, challenges and enabler technologies,” in *Proceedings of Novel Intelligent and Leading Emerging Sciences Conference*, 2022, pp. 366–373.

[81] X. Zhu and C. Jiang, “Integrated satellite-terrestrial networks toward 6g: Architectures, applications, and challenges,” *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 437–461, 2021.

[82] ESA, https://www.esa.int/Applications/Telecommunications Integrated Applications/Satellite frequency bands, 2022.

[83] K. Gai, Y. Wu, L. Zhu, K.-K. R. Choo, and B. Xiao, “Blockchain-enabled trustworthy group communications in uav networks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4118–4130, 2021.

[84] G. Gardikis, H. Koumaras, C. Sakkas, and V. Koumaras, “Towards sdn/nfv-enabled satellite networks,” *Telecommunication Systems*, vol. 66, pp. 615–628, 2017.

[85] P. Zhang, H. Yao, and Y. Liu, “Virtual network embedding based on computing, network, and storage resource constraints,” *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3298–3304, 2017.

[86] L. Ni, B. Hu, C. Wang, S. Gu, M. Meng, and J. Zhang, “Research on the evolution of 5g and satellite network integration,” *Mobile Communication*, vol. 46, no. 01, pp. 51–57, 2022.

[87] Z. Han, C. Xu, K. Liu, L. Lu, G. Zhao, and S. Yu, “A novel mobile core network architecture for satellite-terrestrial integrated network,” in *Proceddings of the IEEE Global Communications Conference*, 2021, pp. 01–06.

[88] J. Kim, J. Lee, H. Ko, T. Kim, and S. Pack, “Space mobile networks: Satellite as core and access networks for b5g,” *IEEE Communications Magazine*, vol. 60, no. 4, pp. 58–64, 2022.

[89] X. Zhu and X. Qu, “Research on 5g lightweight core network technology for vertical industries,” in *Proceddings of the International Wireless Communications and Mobile Computing*, 2021, pp. 1499–1505.

[90] S. Alfattani, W. Jaafar, Y. Hmamouche, H. Yanikomeroglu, A. Yongaçoglu, N. D. DJào, and P. Zhu, “Aerial platforms with reconfigurable smart surfaces for 5g and beyond,” *IEEE Communications Magazine*, vol. 59, no. 1, pp. 96–102, 2021.

[91] Y. Li, H. Li, L. Liu, W. Liu, J. Liu, J. Wu, Q. Wu, J. Liu, and Z. Lai, “”internet in space” for terrestrial users via cyber-physical convergence,” in *Proceedings of the ACM Workshop on Hot Topics in Networks*, 2021, pp. 163–170.

[92] J. So, K. Hsieh, B. Arzani, S. Noghabi, S. Avestimehr, and R. Chandra, “Fedspace: An efficient federated learning framework at satellites and ground stations,” *arXiv preprint arXiv:2202.01267*, 2022.

[93] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Proceedings of Artificial intelligence and Statistics*, 2017, pp. 1273–1282.

[94] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” *Proceedings of the Conference on Machine Learning and Systems*, vol. 2, pp. 429–450, 2018.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3303346

[95] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security*, 2017.

[96] M. Yurochkin, M. Agarwal, S. Ghosh, K. Greenewald, N. Hoang, and Y. Khazaeni, "Bayesian nonparametric federated learning of neural networks," in *Proceedings of International Conference on Machine Learning*, 2019, pp. 7252–7261.

[97] Q. Yang, "Advances and open problems in federated learning," *Foundations and Trends in Machine Learning*, 2021.

[98] H. Wu, H. Zhang, X. Zhang, W. Sun, B. Zheng, and Y. Jiang, "Deepdualmapper: A gated fusion network for automatic map extraction using aerial images and trajectories," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 1037–1045.

[99] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 24–41, 2016.

[100] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *Proceedings of the International conference on machine learning*, 2019, pp. 7354–7363.

[101] V. Bhosale, K. Bhardwaj, and A. Gavrilovska, "Toward loosely coupled orchestration for the leo satellite edge," in *Proceedings of Workshop on Hot Topics in Edge Computing*, 2020.

[102] Y. Kimoto and H. Matsumoto, "Evaluation of radiation effects on commercial-off-the-shelf (cots) parts for use on low-orbit satellite," *COSPAR Scientific Assembly*, vol. 43, p. 542, 2021.

[103] *Cubesat design specification*, California Polytechnic State University, 2014, rev. 13. Technical report.

[104] C. Wang, Y. Zhang, Q. Li, A. Zhou, and S. Wang, "Satellite computing: A case study of cloud-native satellites," url-https://arxiv.org/pdf/2307.08530.pdf, 2023.

Shangguang Wang is a Professor at the School of Computer Science and Engineering, Beijing University of Posts and Telecommunications, China. He received his Ph.D. degree at Beijing University of Posts and Telecommunications in 2011. He has published more than 150 papers. His research interests include service computing, mobile edge computing, and satellite computing. He is currently serving as Chair of IEEE Technical Committee on Services Computing (2022-2023), and Vice-Chair of IEEE Technical Committee on Cloud Computing (2020-). He also served as General Chairs or Program Chairs of 10+ IEEE conferences. He is a Fellow of the IET, and Senior Member of the IEEE. For further information on Dr. Wang, please visit: http://www.sguangwang.com.

Qing Li received her Ph.D. degree in Beijing University of Posts and Telecommunications, Beijing, China, in 2022. She is currently a research assistant in the School of Computer Science at Peking University, Beijing, China. Her research interests include mobile edge computing and satellite computing.