



The Application of Mobile Edge Computing in the Space-Air-Ground Integrated Network

Peiying ZHANG^{1,2,†}, Guilong WANG³, Shengpeng CHEN^{1,2}, Yihong YU^{1,2}, Lijuan CHEN^{1,2}

¹College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China

²Shandong Key Laboratory of Intelligent Oil & Gas Industrial Software, Qingdao 266580, China

³State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

[†]E-mail: zhangpeiying@upc.edu.cn

Abstract: The rapid evolution of next-generation communication technologies has made the integration of Multi-access Edge Computing (MEC) into Space-Air-Ground Integrated Networks (SAGIN) a critical research frontier. This integration offers global seamless coverage, ultra-low latency, enhanced reliability, and optimized use of computational and communication resources. This paper systematically investigates the fundamental principles, technological architectures, and recent advancements of MEC-enabled SAGIN, with particular emphasis on computation offloading methodologies, resource allocation and management strategies, network architecture optimization, and communication protocol standardization. Critical application scenarios—including intelligent transportation and autonomous driving, emergency communications and disaster mitigation, telemedicine and remote education, as well as aerial and maritime mobile services—are analyzed to illustrate the transformative potential of integrated MEC–SAGIN frameworks. Concurrently, this study examines key technical challenges arising from heterogeneous dynamic network conditions, constrained resources, computational offloading complexities, security threats, and privacy concerns. To overcome these challenges, future research directions are proposed, highlighting the integration of advanced technologies such as digital twins for precise network modeling and optimization, artificial intelligence for intelligent resource orchestration, blockchain for secure and transparent resource sharing, and quantum computing and communication for enhanced network security and performance. This comprehensive survey provides a foundational reference for theoretical advancement and practical deployment of MEC-integrated SAGIN systems, facilitating the evolution toward intelligent, efficient, and globally interconnected communication infrastructures.

Keywords: SAGIN; MEC; computing service; 6G; internet of things

<https://doi.org/10.64509/jicn.11.13>

1 Introduction

1.1 Research Background

Rapid advancements in information technology are reshaping global communication networks. The full deployment of the fifth generation of mobile communication technology (5G) has not only greatly improved data transmission rates and network capacity, but also promoted the rapid development of emerging applications such as the Internet of Things (IoT) [1], automated driving and telemedicine. Looking ahead, the research and test of the sixth generation mobile communication technology (6G) [2] has kicked off, with the goal of building a communication network system that seamlessly covers the globe and supports a variety of heterogeneous accesses

while realizing higher data rates, lower latency and greater connection density.

Against the backdrop of limited coverage, insufficient network capacity and high latency of traditional terrestrial communication networks, Space-Air-Ground Integrated Networks (SAGIN) has emerged as a new type of network architecture that integrates the satellite network (space segment), drone and high-altitude platform network (air segment), and terrestrial communication network (ground segment). By integrating satellite networks (space segment), drones and high-altitude platforms (air segment), and terrestrial communication networks (ground segment), SAGIN aims to realize seamless coverage and efficient communication services on a global scale. Satellite networks offer wide coverage but suffer from high transmission delays, limited bandwidth, and scarce

[†] Corresponding author: Peiying Zhang

* Academic Editor: Chunxiao Jiang

© 2025 The authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

resources. In contrast, aerial networks composed of UAVs and high-altitude platforms provide flexible deployment, rapid responsiveness, and supplementary communications. Terrestrial networks are characterized by high capacity, low latency, and stable performance.

Mobile Edge Computing (MEC), also known in recent years as Multi-Access Edge Computing, is an emerging distributed computing paradigm [3]. MEC meets end-users' needs for high-bandwidth, low-latency, and real-time services by deploying compute, storage, and network resources at network edge nodes to push data processing capabilities closer to the data source. This technology effectively reduces data transmission latency, improves network performance and efficiency, and becomes a key technology to support real-time and interactive applications in 5G and future 6G networks.

With the proposal of SAGIN architecture and the development of MEC technology, the convergence of the two has become an inevitable trend [4]. Figure 1 shows the system architecture diagram of the combination of SAGIN and MEC. The multilayered and heterogeneous nature of SAGIN puts forward higher requirements for the efficient utilization of computation, communication, and storage resources, while MEC, as a key enabling technology, is able to push computation power to the edge of the network to satisfy the needs of low latency and high bandwidth for various applications in SAGIN. However, the convergence of SAGIN and MEC faces many challenges, including dynamic topology changes, diverse computing node capabilities, unstable communication links, and complex resource management.

1.2 Research Motivation and Significance

The rapid advancement of next-generation communication technologies has underscored the critical need for integrating MEC into SAGIN [5]. Existing research has predominantly explored MEC and SAGIN in isolation or focused on narrow intersections, failing to address the complex synergies and challenges arising from their integration. This gap is particularly pronounced given SAGIN's unique architecture, which combines satellite, aerial, and terrestrial networks to enable global connectivity [6], while MEC offers low-latency computation capabilities at the network edge. The fragmented nature of prior studies has left unresolved questions about how to effectively leverage MEC's computational resources within SAGIN's multi-tiered framework, hindering progress toward seamless, efficient, and robust next-generation communication systems [7]. The dynamic topology, diverse node capabilities, and unstable links in SAGIN further exacerbate the need for a systematic exploration of MEC integration, as traditional approaches prove insufficient for optimizing performance in such heterogeneous environments.

The significance of this research lies in its potential to bridge theoretical and practical divides in communication and computing infrastructures [8]. The integration of MEC into SAGIN holds transformative promise for addressing long-standing challenges in network coverage, latency, and resource efficiency. The study's comprehensive framework for analyzing key technologies, application scenarios, and future directions not only advances the theoretical understanding of how edge computing can enhance SAGIN's

capabilities but also provides tangible guidelines for industry implementations. By identifying technological limitations and proposing innovative solutions—such as leveraging digital twins, blockchain, and advanced AI methodologies—this research paves the way for more resilient and intelligent networks [9]. These insights are invaluable for enabling real-time services in critical domains like emergency communications, autonomous transportation, and industrial IoT, where the combination of SAGIN's extensive coverage and MEC's low-latency processing is essential. Ultimately, this work contributes to shaping the technological foundation for 6G and beyond, ensuring that future networks can meet the escalating demands of latency-sensitive, data-intensive applications at scale.

1.3 Contributions and Structure

The integration of MEC with SAGIN represents a promising yet challenging frontier for next-generation communication and computing infrastructures. Existing research efforts often examine MEC and SAGIN in isolation or focus only on narrow intersections between the two. Although some prior works, such as [10], have attempted to survey the integration of MEC and SAGIN, they tend to emphasize specific aspects, such as MEC deployment strategies, resource management techniques, optimization algorithms, or the design of particular network architectures and service frameworks.

In contrast, this survey provides a comprehensive and systematic investigation of MEC in the context of SAGIN. It offers an integrated perspective that encompasses key enabling technologies, representative application scenarios, and future research directions. Specifically, the main contributions of this work are summarized as follows:

Firstly, we establish a clear and unified foundational framework by comprehensively introducing core concepts and technological backgrounds of both SAGIN and MEC, explicitly highlighting the unique characteristics and challenges posed by their integration. This integrated perspective clarifies the complex interdependencies and opportunities arising from combining MEC's low-latency computation capabilities with SAGIN's extensive coverage and multi-tiered network architecture.

Secondly, we provide an in-depth and critical analysis of the state-of-the-art MEC technologies tailored specifically for SAGIN scenarios, covering computation offloading strategies, resource optimization mechanisms, network architecture design, and standardized communication protocols. Through systematic categorization, rigorous evaluation, and identification of existing technological limitations, we offer valuable insights into the practical deployment and performance enhancement of MEC-enabled SAGIN systems.

Thirdly, we present a comprehensive overview of current challenges and provide innovative, forward-looking research directions for effectively addressing these issues. Our proposed research pathways integrate multidisciplinary cutting-edge technologies—including digital twins, blockchain, quantum security, energy harvesting, and advanced AI methodologies—thus laying a clear and practical roadmap for future academic studies and industry implementations.

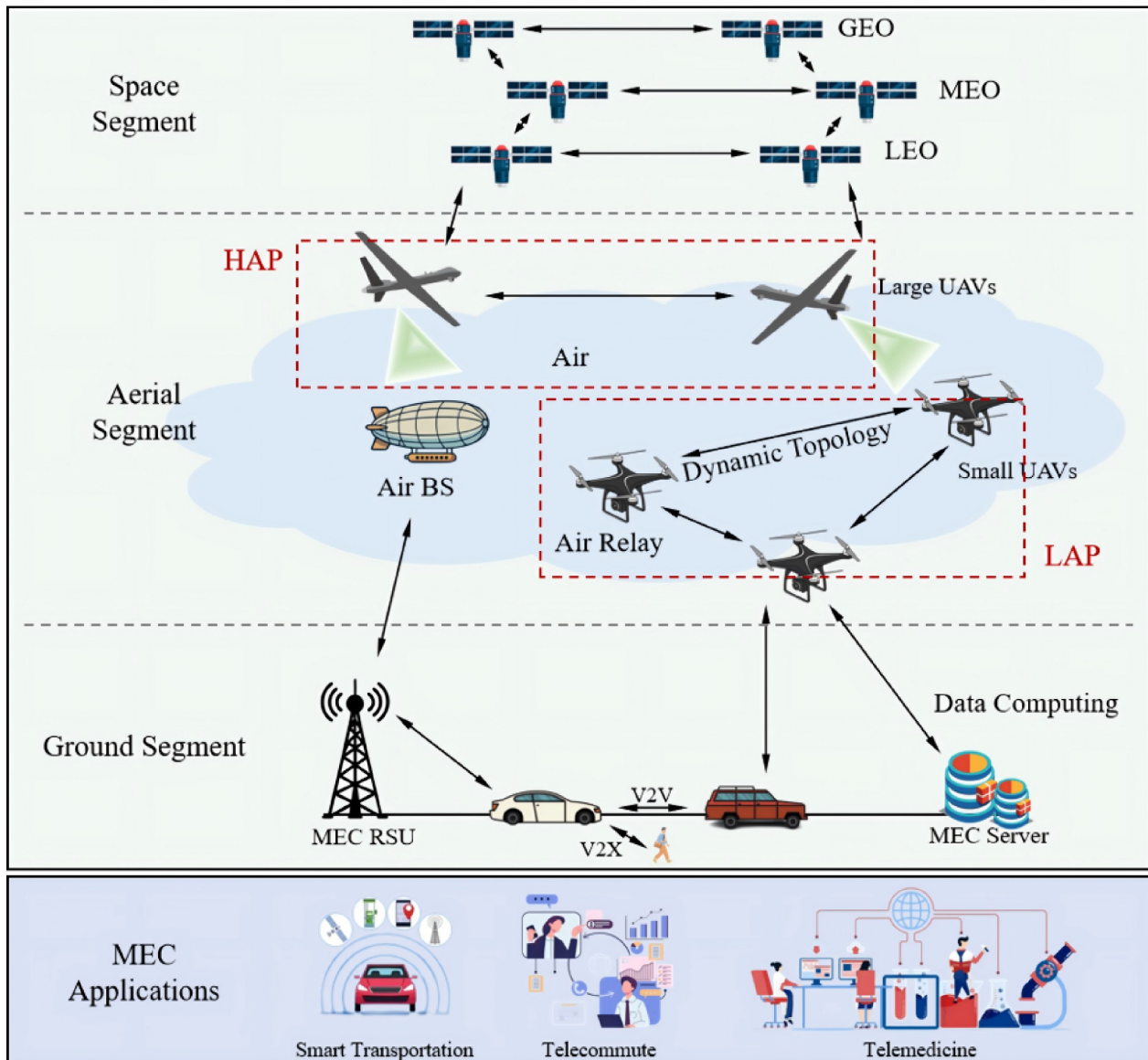


Figure 1: System Architecture of MEC-enabled SAGIN.

Figure 2 shows the overall structure of this article. The remainder of this paper is structured as follows. Section 2 introduces fundamental concepts and technological backgrounds of SAGIN and MEC. Section 3 comprehensively discusses key MEC technologies in SAGIN. Section 4 presents representative practical application scenarios. Section 5 critically analyzes the open challenges and unresolved issues. Section 6 outlines promising future research directions, and Section 7 concludes this paper by summarizing key insights, contributions, and implications.

2 Fundamental Concepts and Technological Background

2.1 Space-Air-Ground Integrated Networks (SAGIN)

The SAGIN is a novel network architecture that integrates satellites, aerial platforms, and terrestrial networks into a

unified architecture. Its primary objective is to provide seamless global connectivity and efficient communication services. With the advent of 5G and the forthcoming 6G communication technologies, SAGIN has emerged as one of the key trends in future network development and is widely regarded as an effective means to overcome the traditional issues of limited coverage, insufficient network capacity, and high latency [9, 11].

2.1.1 Space Segment

The space segment primarily consists of satellites positioned at varying orbital altitudes, including Geostationary Earth Orbit (GEO), Medium Earth Orbit (MEO), and Low Earth Orbit (LEO) satellites. Each of these categories exhibits distinct communication characteristics and caters to different application scenarios:

- **GEO Satellites:** GEO satellites, located at an altitude of approximately 35,786 km, offer extensive coverage but incur relatively long communication delays (around 250 ms).

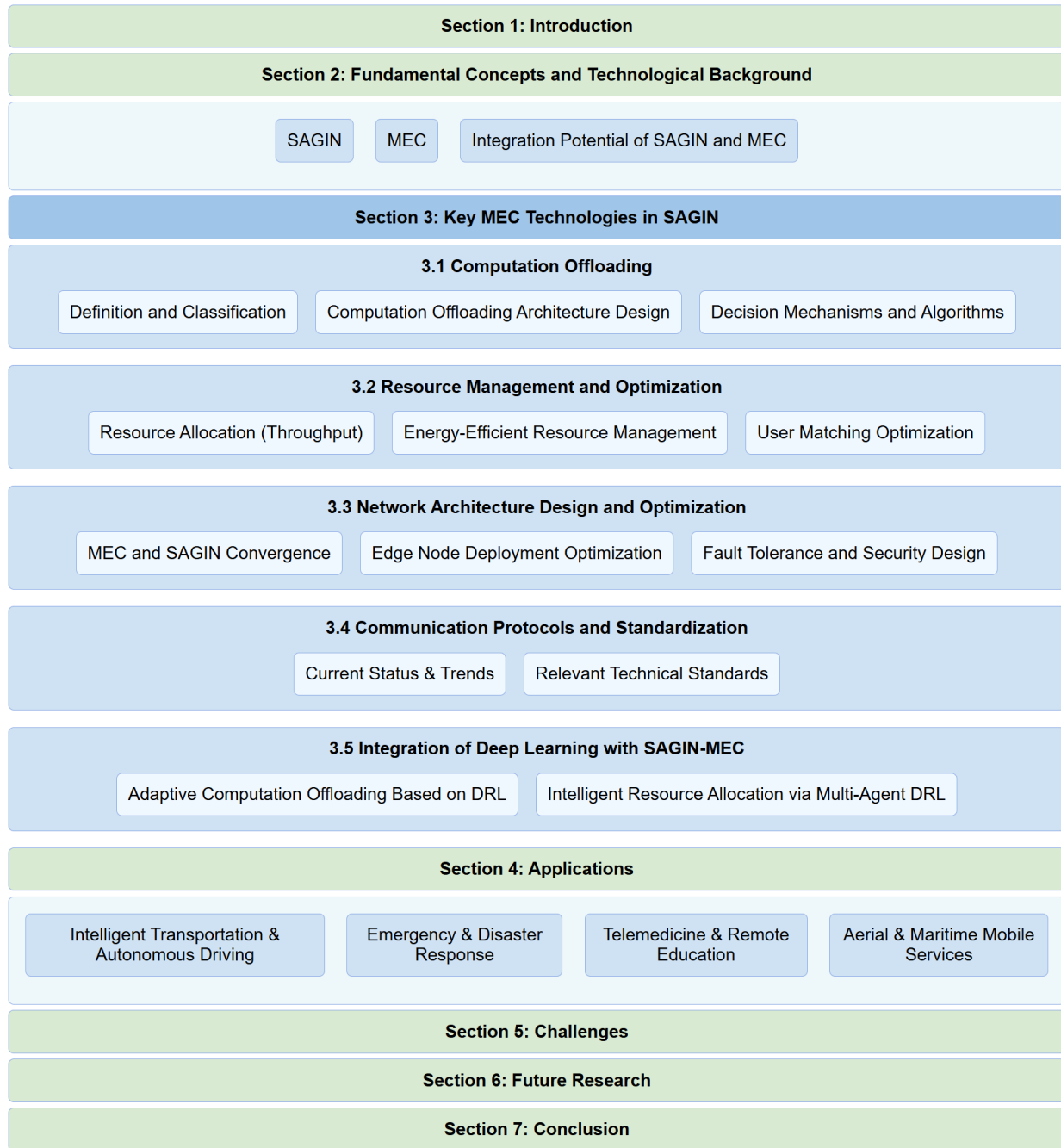


Figure 2: Overall Structure and Logical Flow of the Paper.

one-way latency). These satellites are typically employed for broadcast television, fixed communication services, and ensuring coverage in remote areas.

- **MEO Satellites:** Positioned at altitudes ranging from 2,000 to 20,000 km, MEO satellites achieve moderate latency (approximately 50–150 ms) and are well-suited for navigation and positioning services, such as those provided by the Global Positioning System (GPS) and the Galileo system.
- **LEO Satellites:** LEO satellites operate at altitudes between 200 km and 2,000 km. They are characterized by lower communication latency (around 10–50 ms) and higher data transmission rates, making them ideal for broadband internet access and real-time communication

services. Notable examples include satellite constellation projects such as SpaceX’s Starlink and OneWeb [12].

Satellite networks possess the inherent advantage of wide coverage and are less affected by terrestrial environmental factors, rendering them crucial in remote, maritime, mountainous, and emergency disaster scenarios [13]. However, challenges such as high transmission delays, limited bandwidth, and the high cost as well as scarcity of satellite resources persist. Consequently, researchers have been actively working to optimize satellite communication protocols and develop effective resource management strategies to enhance the performance and efficiency of satellite networks [14].

2.1.2 Air Segment

The air segment mainly encompasses networks of unmanned aerial vehicles (UAVs) and high-altitude platforms (HAPs), which serve as vital bridges connecting the space segment with the ground segment.

- **UAV Networks:** UAVs, typically operating at altitudes ranging from tens to a few hundred meters, provide benefits such as flexible deployment, cost-effectiveness, and rapid response. These attributes make them suitable for applications like temporary communication coverage, post-disaster emergency communications, military operations, and IoT data collection. Recent research has focused extensively on UAV-assisted network architecture design, trajectory optimization, energy management, and resource allocation strategies to improve overall network performance and energy efficiency [15].
- **High-Altitude Platforms:** High-altitude platforms generally function within the stratosphere, at altitudes between 17 km and 22 km, and include vehicles such as high-altitude airships and unmanned high-altitude aircraft. They are capable of providing extensive area coverage and relatively stable communication links, while offering lower latency than satellite networks and broader coverage than terrestrial networks. Notable HAP projects include initiatives like Google Loon [16].

The air segment effectively compensates for the higher latency of satellite networks and the coverage limitations of ground networks by providing flexible and high-performance supplementary communication solutions. Nevertheless, UAVs and high-altitude platforms face challenges including limited endurance, susceptibility to adverse wind conditions, and potential instability in communication links [17].

2.1.3 Ground Segment

The ground segment comprises terrestrial cellular networks (e.g., 4G/5G base stations, Wi-Fi access points) and terminal devices (such as smartphones, IoT devices, and vehicular networks). This segment forms the fundamental infrastructure and serves as the direct access point for communication terminals within SAGIN. Its key characteristics include:

- **Cellular Networks:** These networks are marked by high capacity, low latency, and robust stability, rendering them suitable for the communication demands of urban and densely populated areas. The integration of 5G technology has significantly enhanced the performance of ground communication networks, enabling higher data rates, reduced latency, and a considerably higher connection density [18].
- **IoT Devices:** The rapid proliferation of IoT devices has catalyzed the expanded application of ground networks, particularly in areas such as smart cities, precision agriculture, environmental monitoring, and industrial automation. Low-power wide-area network (LPWAN) technologies, including LoRa and NB-IoT, deliver low-cost, energy-efficient, and wide-ranging communication solutions tailored for IoT applications [19].

In remote areas, over the sea, in mountainous regions, or following disasters when terrestrial networks may struggle to provide reliable support, the synchronization of the space and air segments becomes paramount. The cooperative functionality of SAGIN enables comprehensive network coverage and ensures efficient communication even in challenging environments.

Table 1 illustrates a comparative analysis of the key attributes for the Space, Air, and Ground segments in SAGIN, as described in the literature.

In summary, by integrating the strengths of the space, air, and ground segments, SAGIN effectively addresses the challenges associated with network coverage, capacity, and latency. This holistic approach positions SAGIN as a critical development direction for next-generation communication systems.

2.2 Mobile Edge Computing (MEC)

2.2.1 Basic Principles and Architecture of MEC Technology

MEC is an emerging distributed computing paradigm. Its primary objective is to deploy computing, storage, and networking resources at network edge nodes that are in close proximity to end devices, thereby meeting the demands for high bandwidth, low latency, and real-time services by end users [20]. By leveraging network devices—such as cellular base stations, routers, and access points—to provide computational capabilities, MEC brings data processing closer to the data source, reducing transmission latency and enhancing overall network performance and efficiency [21].

Typically, the MEC architecture comprises three layers:

- **Terminal Device Layer:** Consists of various mobile terminals, Internet of Things (IoT) devices, sensors, unmanned aerial vehicles (UAVs), etc.
- **Edge Computing Node Layer:** Deployed at the network periphery (e.g., 5G base stations, small-scale data centers, routers), this layer is tasked with real-time processing and analysis of data from terminal devices to enable prompt responses.
- **Central Cloud Data Center Layer:** Primarily responsible for large-scale data storage, deep analytics, and long-term maintenance, this layer provides services for massive data analysis and storage [22].

2.2.2 Differences and Interrelationships Between MEC and Conventional Cloud Computing

MEC and conventional cloud computing exhibit both distinct differences and close interrelationships. Traditional cloud computing is typically deployed in remote, centralized data centers that offer powerful computing and storage resources. These centers are well-suited for large-scale data analysis, offline processing, and non-real-time tasks; however, they encounter significant latency and bandwidth constraints when handling real-time tasks [23].

In contrast, MEC deploys these resources at the network edge, thereby shortening data transmission paths and significantly reducing latency—advantages that are especially beneficial for highly interactive, real-time applications. Importantly, MEC does not aim to replace conventional cloud

Table 1: Comparative Analysis of Space, Air, and Ground Segments in SAGIN

Attributes	Space Segment	Air Segment	Ground Segment
Coverage Scope	Global; extensive reach including remote, maritime, mountainous, and disaster-affected areas	Regional; flexible, rapidly deployable for temporary and emergency scenarios	Localized; primarily urban and densely populated areas
Communication Latency	High; GEO ~250 ms, MEO ~50–150 ms, LEO ~10–50 ms	Moderate; typically less than satellites, altitude-dependent	Low; typically below 10 ms with advanced terrestrial networks (4G/5G)
Data Capacity	Moderate; constrained bandwidth, improved by LEO constellations (e.g., Starlink, OneWeb)	High; flexible, high-rate links achievable, but endurance and link stability constraints exist	High; robust infrastructure, significantly boosted by 5G and dense deployments
Deployment Flexibility	Limited; fixed orbits, long-term planning, slow adjustments	High; rapid deployment, repositioning, suitable for dynamic or emergency scenarios	Limited; static, permanent installations, modifications require substantial investments
Network-Stability	High stability in terms of coverage persistence; susceptible to long delays	Moderate; susceptible to weather conditions, limited flight endurance, and dynamic link quality	High; robust and stable within coverage areas, vulnerability outside dense deployments
Application Scenarios	Remote connectivity, maritime communications, emergency backup, global broadcasting	Emergency response, temporary coverage, military communications, IoT data collection	Urban communications, industrial IoT, smart cities, precision agriculture, vehicular networks
Operational Challenges	High cost, limited bandwidth, resource scarcity, significant propagation delays	Limited operational endurance, weather sensitivity, potential link instability	Coverage gaps in rural and remote regions, vulnerability in disaster or extreme conditions

computing; rather, it complements it. MEC is generally used as an extension of cloud computing by processing high real-time or sensitive data tasks locally at edge nodes, while delegating storage-intensive and large-scale analytical tasks to core cloud infrastructures. This collaborative approach facilitates more efficient distributed processing and storage [24]. Together, MEC and cloud computing constitute a multi-tiered architectural paradigm, allowing for dynamic resource allocation and optimization based on the specific requirements of various tasks.

2.2.3 Application Scenarios of MEC

MEC technology demonstrates promising application prospects across various domains. Typical application scenarios include:

- **Real-Time Video Analysis and Processing:** In contexts such as smart city surveillance, security monitoring, and UAV inspections, MEC technology enables the local, real-time processing and analysis of video streams to facilitate rapid responses to anomalous events. For example, during UAV inspections, MEC platforms can process high-definition images and video data in real time to accurately locate faults and trigger appropriate responses, thereby significantly enhancing both efficiency and safety [25].
- **Vehicle-to-Everything (V2X) Communication and Autonomous Driving:** In intelligent transportation systems,

MEC processes data from vehicles and traffic infrastructure in real time via edge nodes, offering low-latency decision feedback to support collaborative interactions between vehicles and between vehicles and infrastructure. The European 5G Automotive Association (5GAA), for instance, has identified MEC as a key technology in advancing autonomous driving and intelligent transportation systems.

- **Augmented Reality (AR) and Virtual Reality (VR):** Given the extreme sensitivity of AR/VR applications to delays, MEC mitigates network latency and reduces bandwidth requirements by deploying computing resources close to the end user, thereby significantly enhancing the overall user experience. Typical applications include real-time AR gaming, virtual exhibition spaces, and industrial remote maintenance [26].
- **Industrial Internet of Things (IIoT) and Smart Manufacturing:** In industrial settings, MEC provides local real-time data processing capabilities that support intelligent, edge-based analytics and decision-making. This minimizes the need to transmit data to a central cloud, thereby improving both security and real-time performance. Notably, both Germany's Industry 4.0 initiative and China's Intelligent Manufacturing 2025 strategy emphasize the critical role of edge computing in industrial transformation.

In summary, by decentralizing computing capabilities to the network edge, MEC technology provides critical support

for applications with high real-time and interactivity requirements. It has thereby emerged as a key technological enabler in the evolution of SAGIN.

2.3 Trends, Demands, and Integration Potential of SAGIN and MEC

In recent years, the rapid advancement of digital infrastructure and emerging intelligent applications has driven both academic and industrial sectors to intensify their focus on SAGIN and MEC. These two paradigms, though originally evolving along distinct technological paths, are increasingly seen as complementary solutions to the complex demands of future communication networks. This section reviews the respective development trends and challenges of SAGIN and MEC, and highlights the strategic value of their convergence.

2.3.1 Development Trends and Focus Areas

In the SAGIN domain, the emergence of mega-constellation satellite systems (e.g., SpaceX Starlink, Amazon Kuiper) has marked a shift toward low-latency, high-throughput satellite communications[27]. Research efforts have increasingly concentrated on dynamic topology management[28], inter-segment routing protocols[29], and adaptive resource allocation under heterogeneous constraints[30]. In the industry, SAGIN is being actively explored for its potential to support global broadband services, emergency communication, and ubiquitous connectivity in 6G networks[5].

In the MEC domain, the evolution from single-access edge computing to multi-access and federated edge computing has attracted considerable attention. Academic studies are delving into lightweight virtualization, edge AI, and cross-domain orchestration mechanisms[22]. On the industrial front, major players such as Huawei, Intel, and Nokia are deploying MEC-enabled solutions in fields ranging from autonomous driving and industrial automation to immersive media services. The integration of MEC with 5G infrastructures is also advancing rapidly, enabling tighter coordination between computing and communication resources.

2.3.2 Real-World Demands and Persistent Challenges

Despite significant progress, SAGIN faces several operational challenges, including high dynamicity and heterogeneity across segments, making seamless handover and routing complex, limited onboard computing and storage resources on satellites and UAVs, restricting real-time service capabilities and latency and reliability issues in inter-segment communication, especially in highly mobile or remote environments. Similarly, MEC systems confront their own limitations:

- Geographic deployment constraints, with MEC nodes primarily concentrated in urban or populated regions, leaving coverage gaps in remote or disaster-stricken areas[29].
- Resource limitations at the edge, including computational bottlenecks and energy constraints[5].
- Scalability and interoperability issues, particularly when managing multi-tenant environments or cross-operator deployments[31].

These challenges underscore a pressing demand for more adaptive, intelligent, and widely accessible computing and communication frameworks.

2.3.3 Necessity and Prospects of SAGIN–MEC Integration

The convergence of SAGIN and MEC emerges as a promising approach to address the above limitations. By deploying MEC functionalities across the space-air-ground segments, such as equipping LEO satellites or UAVs with edge computing modules, the system can achieve:

- Expanded service coverage: MEC capabilities can be extended to previously underserved or unreachable areas (e.g., oceanic zones, remote villages, or post-disaster regions), enabled by the ubiquitous presence of SAGIN nodes[32].
- Reduced end-to-end latency: Computing tasks can be offloaded and executed directly at or near airborne or satellite nodes, significantly shortening data transmission paths[33].
- Improved real-time responsiveness: Time-critical applications such as autonomous UAV coordination, aerial surveillance, and cross-domain IoT services can benefit from faster local processing and decision-making.
- Enhanced resource coordination: The layered topology of SAGIN allows for hierarchical MEC deployment, supporting adaptive task migration and resource balancing between segments[5].

Furthermore, the integration fosters context-aware and intelligent services, whereby global network awareness from satellites, local environmental data from UAVs, and user-centric demands from ground terminals can be jointly processed to optimize application performance.

2.3.4 Strategic Value for Future Networks and Intelligent Applications

The fusion of SAGIN and MEC not only enhances technical performance but also holds transformative potential for the architecture of future networks. It paves the way for:

- Highly resilient and adaptive communication infrastructures, capable of maintaining service continuity in volatile environments.
- Intelligent network control and orchestration, leveraging distributed edge intelligence and cross-layer optimization.
- Scalable support for emerging applications, including space-based IoT, aerial-edge collaborative sensing, and 6G-enabled metaverse experiences.

In summary, the joint evolution of SAGIN and MEC is both a necessary response to current technological constraints and a strategic move toward next-generation, ubiquitous intelligent networks. Their integration sets a solid foundation for subsequent discussions on resource allocation strategies, architectural designs, and intelligent control mechanisms in complex, multi-domain network environments.

3 Key MEC Technologies in SAGIN

3.1 Computation Offloading

3.1.1 Definition and Classification

With the rapid development of mobile communication and Internet of Things (IoT) technologies, the vast amount of data generated by various terminal devices (such as smart phones, in-vehicle terminals, unmanned aerial vehicles, sensor devices, etc.) makes computing and communication necessary to be carried out anytime and anywhere. However, limited by the computing power, battery capacity and memory resources of terminal devices, relying solely on local computing of terminal devices has become difficult to meet the requirements of real-time and energy-sensitive tasks. To solve this contradiction, the technology of partially or fully migrating the computationally intensive tasks on terminal devices to edge nodes or the cloud with higher computing capabilities for processing is called Computation Offloading [34].

Specifically, computing offloading refers to the process in which terminal devices migrate computing tasks that are difficult for them to handle or not suitable for local execution to other nodes in the network (such as edge servers, air nodes, satellite nodes or cloud servers) through wireless communication technology for efficient processing, and return the results to the terminal devices after the task execution is completed [35]. According to the granularity and mode of task migration, computing offloading is usually classified into the following categories [36]:

- **Full Offloading:** In this mode, the entire computational workload is migrated from the terminal device to remote nodes, leaving the terminal primarily responsible for data sensing, basic input/output operations, and result presentation. This approach is beneficial when terminal devices have extremely limited computing resources or severe energy constraints. Web-app is a typical example of complete uninstallation [37]. The client is only responsible for collecting input and displaying output.
- **Partial Offloading:** In contrast to full offloading, partial offloading splits the computational workload into multiple sub-tasks or functional components. Only selected sub-tasks, especially those with high computational complexity, are offloaded, while the remaining tasks are processed locally. This mode offers a balance between reducing latency and saving local computing resources, thus being suitable for situations with moderate computing capacities or tasks with mixed latency and computational intensity requirements [38].
- **Granularity-based Offloading:** Depending on how finely tasks are partitioned for offloading decision-making, granularity-based offloading can further be divided into:
 - *Task-level Offloading:* Tasks are treated as indivisible units when making offloading decisions. This simplifies offloading decisions and is suitable for relatively simple and highly independent tasks.

- *Component-level Offloading:* This method partitions applications into multiple functional components. Developers analyze the execution flow and divide computationally intensive components for offloading, leading to more optimized resource utilization.
- *Method-level Offloading:* At this fine granularity, individual methods or functions within an application are selectively offloaded. Such detailed partitioning requires sophisticated analysis of program structures and execution patterns, making it suitable for scenarios requiring highly precise optimization. In the research work of [39], the character recognition (OCR) of AR applications was offloaded for calculation, which is a typical function offloading.

Due to the inherent complexity and heterogeneity of SAGIN, computation offloading faces unique challenges, including dynamic topology changes, diverse computing node capabilities, and unstable communication links. Therefore, exploring effective offloading strategies tailored specifically to SAGIN scenarios has significant theoretical importance and practical value.

3.1.2 Computation Offloading Architecture Design in Space-Air-Ground Integrated Networks

Due to the multi-layered and heterogeneous characteristics of SAGIN, the architecture design for computation offloading is more complex compared to traditional terrestrial networks. According to the number of involved nodes, their distribution, and the coordination mechanism, this paper categorizes computation offloading architectures in SAGIN into single-tier offloading architecture and multi-tier offloading architecture.

(1) Single-tier Offloading Architecture

The single-tier offloading architecture represents the most fundamental structure, in which terminal devices directly offload computational tasks to terrestrial edge computing nodes, such as base station servers or micro data centers, via wireless communication networks [40]. The primary advantages of single-tier architecture include simplicity, ease of deployment, and rapid task offloading and result feedback. It is especially suitable for urban scenarios with well-developed communication infrastructures, relatively low user mobility, and latency-sensitive applications, such as intelligent transportation, video analytics, and virtual reality (VR) services [21]. However, its limitations are also evident, including high dependency on terrestrial network coverage and stability, making it challenging to provide consistent services in remote areas or under special circumstances like natural disasters.

(2) Multi-tier Offloading Architecture

Compared to single-tier architecture, the multi-tier offloading architecture fully leverages the unique capabilities of SAGIN by coordinating resources among ground-based nodes, aerial nodes, satellite nodes, and cloud nodes, thus achieving more efficient collaborative computation offloading [41]. A typical multi-tier offloading architecture includes terminal devices, aerial nodes, satellite nodes, terrestrial edge nodes, and cloud data centers. These nodes collaboratively optimize offloading decisions and paths dynamically based on task characteristics, network conditions, and resource availability.

Specifically, the main features and advantages of multi-tier offloading architecture are as follows:

- **Aerial-assisted Offloading:** Aerial nodes (such as Unmanned Aerial Vehicles (UAVs), airships, etc.) positioned between terrestrial nodes and satellite nodes possess advantages like flexible deployment, rapid responsiveness, and adjustable positioning [42]. Appropriately deploying aerial nodes can effectively bridge terrestrial coverage gaps, reduce link latency, and enhance real-time task processing, making them particularly beneficial in emergency communications, disaster recovery, and rural or remote area scenarios. Additionally, aerial nodes can dynamically adjust their positions to further optimize service quality and computational performance based on user demands [43].
- **Satellite-assisted Offloading:** Satellite nodes, including low-earth-orbit (LEO), medium-earth-orbit (MEO), and geostationary-earth-orbit (GEO) satellites, can provide wide geographical coverage, strong communication capability, and certain computing capabilities [44]. Among them, LEO satellites are suitable for latency-sensitive applications due to their lower orbital altitude and shorter communication delays. In contrast, MEO and GEO satellites, positioned at higher orbits with broader coverage areas, are more suitable for latency-tolerant tasks such as wide-area data broadcasting and multimedia content distribution [45]. Satellite nodes are particularly advantageous in remote, oceanic, mountainous, or disaster-affected areas where terrestrial and aerial coverage may be insufficient or unavailable.
- **Terrestrial Edge and Cloud Collaboration:** The collaboration between terrestrial edge nodes and cloud data centers achieves optimized hierarchical allocation of computational resources. Due to their proximity to end-users, terrestrial edge nodes can efficiently process latency-sensitive tasks, such as real-time video analytics and augmented reality (AR) applications. In contrast, cloud data centers, possessing extensive computational and storage resources, are better suited for large-scale, computationally intensive, but latency-tolerant tasks, such as big data analytics and machine learning model training [46]. Through effective collaboration, edge and cloud nodes can allocate tasks appropriately based on their specific requirements, thereby optimizing overall system performance.

(3) Applicability and Performance Comparison of Different Architectures

In summary, both single-tier and multi-tier architectures have their distinct advantages and are suitable for different application scenarios:

- The single-tier architecture is applicable to scenarios with well-developed terrestrial network infrastructures, low user mobility, and latency-sensitive applications, such as intelligent transportation in urban areas, indoor localization, and AR/VR services.
- The multi-tier architecture is particularly advantageous in complex environments characterized by inadequate network coverage, geographical constraints, high node mobility, or heterogeneous task requirements, such as disaster

relief, emergency communications, rural or mountainous areas, and maritime communications.

Regarding performance, single-tier architecture offers lower complexity and shorter task latency but is limited in adaptability. In contrast, multi-tier architecture demonstrates greater flexibility, coverage, and robustness but involves higher implementation complexity and more sophisticated dynamic scheduling algorithms.

Therefore, when practical deployment and applications are considered, it is essential to comprehensively evaluate network conditions, task characteristics, deployment costs, and user requirements to select the most appropriate offloading architecture. Such careful selection ensures an optimal balance between resource efficiency and user experience.

3.1.3 Computation Offloading Decision Mechanisms and Algorithms in Space-Air-Ground Integrated Networks

Computation offloading decision mechanisms and algorithms in SAGIN are crucial for achieving efficient, stable, and intelligent task processing. Due to the numerous, heterogeneous, and highly dynamic nature of SAGIN nodes, computation offloading presents abundant opportunities but also significant challenges for optimization decisions [47]. This subsection systematically analyzes four aspects: influential factors for offloading decisions, optimization theory-based methods, game-theoretic approaches, and machine learning methods.

(1) Analysis of Influential Factors in Offloading Decisions

Network Conditions. In SAGIN, satellite links typically exhibit propagation delays significantly greater than traditional terrestrial links. Bandwidth resources are scarce and costly, and link stability fluctuates markedly due to weather conditions, obstructions, or the mobility of nodes [48].

Node Status. Node status includes computing resources, energy constraints, and storage capabilities. Aerial nodes (e.g., UAVs, airships) have limited battery capacities, making energy management critical. Satellite node lifespans are constrained by energy and equipment longevity, further limiting computational resource usage. Additionally, terrestrial edge nodes, although abundant in computing resources, may face load fluctuations [49].

User Mobility and Task Characteristics. High-speed user mobility significantly increases the frequency and intensity of network topology changes. Task characteristics, such as task size, computational complexity, and deadlines, significantly affect the selection and optimization of offloading decisions [50].

(2) Optimization Theory-Based Computation Offloading Methods

Convex optimization methods, favored for theoretical maturity and rapid solutions, were widely adopted in early studies. Recently, researchers introduce non-convex optimization methods into SAGIN computation offloading scenarios to better address complex network characteristics. For example, Sardellitti et al. [51] addressed computation offloading in MIMO multi-cell systems, achieving global optimal solutions via convex optimization in single-user scenarios and iterative approaches based on successive convex approximation for multi-user non-convex optimization scenarios,

thereby reducing overall energy consumption under delay constraints. Yu et al. [52] explored transforming the task offloading problem in mobile edge computing into a convex optimization problem by jointly considering completion time and energy consumption, designing a distributed algorithm encompassing offloading strategy selection, clock frequency configuration, transmit power allocation, and channel rate scheduling to ensure energy-efficient offloading while maintaining user experience. Mei et al. [53] deployed edge servers on satellites and HAP UAVs, utilizing NOMA technology for spectrum sharing, and formulated a weighted energy minimization problem jointly considering power control, computational frequency allocation, and offloading decisions. They adopted block coordinate descent (BCD), solving sub-problems via convex optimization and penalty-based CCP methods. Mixed-integer linear programming (MILP) methods have been widely employed in offloading problems with discrete decision variables. For instance, Bi et al. [54] modeled service cache placement, computation offloading decisions, and resource allocation as a MINLP problem under limited cache resources, which was further transformed into a pure 0-1 ILP. They iteratively updated caching and offloading decisions using alternating minimization techniques. Khan et al. [55] proposed a computation offloading algorithm based on integer linear optimization, enabling mobile devices to choose among local execution, offloading execution, or task dropping, effectively addressing energy saving and computational capacity enhancement challenges introduced by energy harvesting and MEC technologies in IoT environments.

(3) Game Theory-Based Computation Offloading Methods

With the continuous advancement of SAGIN technologies, mobile edge computing has become increasingly critical in supporting massive data processing and real-time service requirements. Game theory is widely applied in this domain due to its precise modeling of interactions among participants in resource allocation and task offloading. Liu et al. [56] formulated the task offloading problem as an NP-hard binary integer linear programming (BILP) problem minimizing delay and energy consumption, subsequently transforming it into a non-cooperative strategic game. They proved the existence of Nash equilibrium through potential games and proposed the Nash equilibrium iterative offloading algorithm (NEIO-G), significantly reducing delay and energy consumption costs in simulations. Gao et al. [57] proposed a game theory-based SAGIN computational resource offloading mechanism involving LEO satellites, UAVs, and users, simultaneously addressing delay and energy consumption by leveraging multi-level game equilibrium combined with resource allocation and pricing strategies, maximizing system profits and demonstrating superior performance. Zhang et al. [58] proposed a distributed dynamic task offloading mechanism based on multi-agent stochastic learning from a game theory perspective. Addressing dynamic challenges arising from varying task performance weights, sizes, and processing needs in SAGIN IoT devices, each device was modeled as a game player aiming to minimize weighted delay and energy costs. A multi-agent entropy-enhanced stochastic learning

(MESL) algorithm was developed without requiring device-to-device information exchange, significantly reducing total costs and substantially improving convergence speed.

(4) Reinforcement Learning Methods

Deep reinforcement learning (DRL), due to its adaptive decision-making capabilities, is widely applied to computation offloading in dynamic SAGIN environments. Hwang et al. [59] addressed the joint optimization of task offloading, resource allocation, and UAV mobility from a DRL perspective. For decentralized implementation, they proposed a multi-agent DRL method where autonomous UAVs collaboratively determined computation and communication strategies without central coordination. Li et al. [60] formulated a mixed-integer nonlinear programming model considering communication, computation, and caching cost constraints, jointly optimizing computation offloading and resource management problems. Employing a DRL strategy, tasks were adaptively offloaded and resources allocated, achieving favorable outcomes across varying terminal counts and reducing reliance on complete formulations and prior information. Xie et al. [61] proposed a vehicle-based SAGIN computation offloading framework, coordinating space, aerial, and terrestrial resources for real-time offloading decisions. They leveraged DRL combined with a ranking-based prioritized experience replay method, substantially outperforming traditional methods in latency fairness and resource utilization. Tang et al. [62] proposed a reinforcement learning-based traffic offloading scheme for highly dynamic SAGIN environments. They adopted an improved delay-sensitive replay memory algorithm (DSRPM) combined with double Q-learning, enabling nodes to intelligently make offloading decisions based on local and neighboring historical information. Through hello packets and offline training mechanisms, their approach outperformed traditional algorithms in reducing signaling overhead, enhancing dynamic adaptability, reducing packet loss rates, and lowering transmission delays.

Table 2 is a summary and comparison of the above-mentioned various computing offloading algorithms.

3.2 Resource Management and Optimization

In the context of SAGIN, the heterogeneous and dynamic nature of network components poses significant challenges to the efficient utilization of computing, communication, and storage resources. MEC, as a key enabler, brings computation closer to end users across all network layers, making resource management and optimization a critical issue. Effective strategies are required to dynamically allocate limited resources, adapt to changing network topologies, and ensure low-latency and high-reliability services. This section reviews the state-of-the-art approaches in resource management within MEC-enabled SAGIN, highlighting key methods, optimization objectives, and open challenges.

3.2.1 Resource Allocation for Throughput Maximization

Huang et al. [63] proposed a DRL-based multi-agent algorithm to optimize task offloading and resource allocation in SAGIN, aiming to improve system throughput while reducing energy consumption and latency. Zhu et al. [64] proposed a SAGIN-MEC framework and decomposed the resource

Table 2: Summary of Different Computation Offloading Methods

Method Category	Ref.	Methods & Contributions
Optimization Theory-Based	[51]	Employed convex optimization for single-user scenarios and iterative successive convex approximation for multi-user non-convex scenarios, minimizing energy under delay constraints.
	[52]	Transformed task offloading into convex optimization, proposed distributed algorithm for energy-efficient offloading and user experience enhancement.
	[53]	Developed weighted energy minimization using NOMA, solved via BCD and penalty-based CCP for joint power control and offloading.
	[54]	Modeled offloading as MINLP, transformed into 0-1 ILP, and solved iteratively via alternating minimization.
	[55]	Proposed integer linear optimization-based offloading for IoT, enabling local execution, offloading, or dropping tasks.
Game Theory-Based	[56]	Formulated task offloading as potential game, proved Nash equilibrium existence, proposed NEIO-G algorithm significantly reducing delay and energy costs.
	[57]	Proposed multi-level game-theoretic mechanism involving LEO satellites and UAVs, optimizing resource allocation and pricing for system profits.
	[58]	Proposed a distributed dynamic task offloading mechanism based on multi-agent stochastic learning from a game theory perspective.
Reinforcement Learning-Based	[59]	Proposed decentralized multi-agent DRL approach for joint offloading, resource allocation, and UAV mobility.
	[60]	Applied DRL for joint optimization of caching, computation offloading, and resource management, achieving adaptability without complete prior information.
	[61]	Employed DRL with prioritized experience replay for SAGIN offloading, improving latency fairness and resource utilization.
	[62]	Combined delay-sensitive replay memory with double Q-learning, reducing signaling overhead, packet loss, and transmission delays.

allocation problem into sub-problems for UAV computing, satellite computing, and task offloading, which were solved using deep reinforcement learning and convex optimization. Nguyen et al. [65] studied resource allocation in hybrid edge-cloud SAGIN by jointly optimizing task offloading, UAV trajectory, user scheduling, and bandwidth allocation using alternating optimization and successive convex approximation methods. Yang et al. [66] proposed a reconfigurable SDN-based SAGIN architecture and formulated a joint UAV trajectory and link selection problem, solving it with a K-means-DDPG algorithm to optimize downlink throughput. Fan et al. [67] proposed a scalable scheduling framework and a joint task scheduling and resource allocation algorithm based on deep reinforcement learning for SAGIN.

These studies in SAGIN resource management and optimization focus on leveraging deep reinforcement learning, multi-agent algorithms, and convex optimization techniques to jointly optimize task offloading, computing resources, and link scheduling, aiming to maximize system throughput while accommodating diverse task requirements. However, due to the network's heterogeneity, dynamic nature, and security threats, designing efficient and robust resource allocation mechanisms remains a critical and unresolved challenge.

3.2.2 Energy-Efficient Resource Management in SAGIN

redQin et al. [68] proposed a NOMA-enabled SAGIN-IPIoT model and formulated an energy efficiency maximization problem, jointly optimizing subchannel assignment and terminal power using matching game and Lagrange dual methods. Chen et al. [69] proposed a UAV-assisted SAGIN architecture and designed a TD3-based DRL algorithm to jointly optimize energy-efficient and fair resource scheduling under jamming constraints. Wei et al. [70] formulated a joint caching and user selection problem in SAGIN and proposed a primal decomposition-based algorithm to optimize UAV caching and user access for energy-efficient resource management. Zhou et al. [71] proposed an algorithm to jointly optimize offloading, task allocation, computing frequency, and UAV deployment for energy-efficient resource management in SAGIN. Jiang et al. [72] formulated a joint optimization problem of offloading, scheduling, computation, and UAV trajectory to minimize weighted energy consumption in SAGIN, solved via block coordinate descent. Liu et al. [3] formulated an energy efficiency maximization problem in SAGIN-supported edge computing by jointly optimizing association, trajectory, offloading, computing frequency, and power, and solved it via alternating optimization.

Recent studies on energy-efficient resource management in SAGIN focus on formulating joint optimization problems involving subchannel allocation, task offloading, computing frequency, UAV deployment, and trajectory planning. Core methods include decomposition techniques, matching theory, and DRL-based algorithms such as TD3. While these approaches enhance energy efficiency under dynamic and constrained conditions, challenges remain in handling large-scale coupling, balancing energy-fairness trade-offs, and achieving real-time adaptability in highly heterogeneous and time-varying environments.

3.2.3 User Matching Optimization

Onyekwelu et al. [73] proposed a user clustering algorithm and decomposed beam-hopping LEO resource allocation into subproblems solved with heuristic and dynamic optimization methods. Sharif et al. [74] formulated a binary linear programming problem to jointly optimize energy efficiency, resource utilization, and task-priority-based user association, solved via BBA, IPM, and BSA methods. Jain et al. [75] proposed a MILP-based AURA-5G framework to jointly optimize bandwidth allocation and access point selection for user association across multiple application profiles. Mao et al. [76] proposed a user tracking and beamforming scheme in SAGIN-ISAC with space- and ISAC-assisted location acquisition for energy-efficient user association. A collaborative multi-agent DRL algorithm is proposed to jointly optimize user association and power allocation in SAGINs using only local information for global energy efficiency [77]. Chen et al. [78] proposed an energy-efficient data access scheme for CAE-SAGIN using AANET and reinforcement learning to jointly optimize resource allocation and request distribution.

User Matching Optimization in integrated networks has seen diverse approaches including clustering, mathematical programming (e.g., MILP, binary linear models), and reinforcement learning. Key optimization goals focus on improving energy efficiency, resource utilization, and user association under complex constraints such as task priority, mobility, and limited information. Recent trends emphasize decentralized and learning-based methods (e.g., DRL, multi-agent collaboration) to enable real-time, scalable solutions. However, challenges remain in achieving globally optimal decisions under dynamic environments, limited CSI, and multi-dimensional resource coupling.

Table 3 provides a comprehensive summary of the resource management and optimization methods discussed in this section, categorizing them by their primary focus areas and highlighting the key techniques employed in each approach.

3.3 Network Architecture Design and Optimization

3.3.1 MEC and SAGIN Convergence

With the wide application of IoT devices and the increasing demand for low-latency, high-bandwidth, and high-reliability communications, the convergence of SAGINs and MECs has become an important development direction for future communications networks. This converged architecture can fully utilize the wide coverage and high capacity of SAGIN and the

low-latency computing capability of MEC to provide users with better quality services.

Literature [4] proposes a SAGIN architecture that incorporates blockchain and MEC, where LEO satellites and UAVs act as edge nodes to provide computation and storage services. Trust in task offloading and wireless data transmission is guaranteed through blockchain technology, and MEC technology is utilized to reduce task execution delay and system energy consumption. The architecture minimizes network energy through intelligent task segmentation, bandwidth allocation and computational resource allocation, and real-time online decision making through DDPG algorithm.

Literature [10] describes the network architecture, key technologies, and challenges of MEC in SAGIN. The article proposes a three-tier SAGIN network architecture, including a space tier, consisting of LEO, MEO, and GEO satellites, an airborne tier, consisting of UAVs and HAPs, and a ground tier, consisting of ground base stations and edge data centers. In this architecture, MEC technology is widely used in edge nodes at all levels to provide low latency and high bandwidth computing services. By analyzing the characteristics and advantages of edge nodes at different levels, the article proposes a variety of MEC deployment models, including single-edge computing, dual-edge computing, and multi-edge computing, to meet the needs of different application scenarios. Meanwhile, the article also deeply discusses the key technologies of MEC in SAGIN, such as network resource scheduling, edge intelligence, optimization objectives and key algorithms, and points out the current challenges and future research directions, such as high uncertainty, random access demand, task migration and load balancing, network security and reliability. Through these analyses and discussions, the article provides comprehensive guidance and reference for the development of MEC technology in SAGIN.

Literature [6] provides a comprehensive overview of resource allocation strategies in SAGIN. The complexities and challenges of resource allocation in SAGIN are discussed in depth, especially in the context of 6G communication technology, where resource optimization becomes more complex and critical. By classifying and summarizing existing resource allocation strategies, gaps in current research are identified and future research directions are discussed. Applications of mathematical optimization, dynamic optimization, game theory, and artificial intelligence, including machine learning, deep learning, and deep reinforcement learning in SAGIN resource allocation are covered. By analyzing in detail the application of these methods in different network scenarios (e.g., IoT, Telematics, Edge Computing Networks, and Emergency Response Networks), it is shown how to optimize the performance of SAGIN through intelligent task allocation, bandwidth allocation, and computational resource allocation. The importance of resource allocation in guaranteeing quality of service, improving network efficiency, and achieving seamless global coverage is also emphasized, providing a comprehensive guide and reference for research in the area of SAGIN resource allocation.

Table 3: Summary of Resource Management and Optimization Methods in SAGIN

Subsection	Ref.	Methods Used
Resource Allocation for Throughput Maximization	[63]	Multi-agent DRL for joint task offloading and resource allocation
	[64]	DRL and convex optimization for decomposed SAGIN-MEC sub-problems
	[65]	Alternating optimization and SCA for edge-cloud joint allocation
	[66]	K-means-DDPG to jointly optimize UAV trajectory and link selection
	[67]	DRL-based joint scheduling and resource allocation framework
Energy-Efficient Resource Management in SAGIN	[68]	Matching game and Lagrangian dual method for energy-efficient NOMA-SAGIN
	[69]	TD3-based DRL for anti-jamming and energy-fair scheduling
	[70]	Primal decomposition for UAV caching and user selection
	[71]	Joint optimization of task allocation, UAV deployment and computing frequency
	[72]	Block coordinate descent to minimize weighted energy consumption
User Matching Optimization	[3]	Alternating optimization of association, trajectory, offloading, and power
	[73]	Heuristic and dynamic optimization for beam-hopping resource allocation
	[74]	BBA, IPM, and BSA for priority-aware user association and utilization
	[75]	MILP-based AURA-5G for bandwidth and AP selection
	[76]	ISAC-assisted user tracking and beamforming for energy efficiency
	[77]	Multi-agent DRL using local information for energy-efficient association
	[78]	RL-based data access scheme for CAE-SAGIN

3.3.2 Edge Node Deployment Location Optimization in SAGIN

In SAGIN, the deployment location of edge nodes has a significant impact on network performance. A reasonable deployment location can reduce data transmission delay, improve computational efficiency, and reduce energy consumption. Therefore, optimizing the deployment location of edge nodes becomes an important issue in the design of SAGIN network architecture.

Relevant chapters in literature [79] explore task offloading, resource allocation, and optimization methods in SAGIN. A DRL-based task offloading and resource allocation optimization method is proposed to be applied in an end-edge-cloud collaborative computing environment. The method models the task offloading and resource allocation problem as a Markov Decision Process (MDP) by constructing a multi-user, multi-server system model and solves it using the Game-PPO algorithm. Through the intelligent task offloading decision and resource allocation strategies, the task processing time and energy consumption of mobile devices are effectively reduced and the system performance is optimized. In addition, the literature also covers research on data enhancement methods based on Generative Adversarial Networks (GANs), blockchain-based data traceability optimization, and task offloading and resource allocation in UAV-assisted vehicular edge computing networks, demonstrating the application of a variety of advanced technologies and optimization strategies in the field of SAGIN.

Literature [80] investigates the computational offloading, UAV deployment, and resource allocation optimization

problems in SAGIN. For the nonconvex mixed integer non-linear optimization problem in SAGIN, the article proposes an alternating optimization method that optimizes user association, partial offloading control, computational resource and bandwidth allocation, and UAV deployment through iterations until convergence. Specifically, the article uses the successive convex approximation (SCA) method to convexify the non-convex bandwidth allocation and UAV deployment subproblems, and verifies the effectiveness of the proposed design under different network settings through numerical studies. Simulation results show that the proposed design improves about 35%-40% in terms of weighted energy consumption compared to the benchmark scheme. In addition, the article provides insights into the importance of UAV deployment and computational offloading optimization for achievable weighted energy consumption, providing new ideas and methods for resource allocation and optimization in SAGIN.

3.3.3 Fault Tolerance and Security Design for SAGIN Network Architecture

Due to the complexity and diversity of SAGIN, the fault tolerance and security design of its network architecture is crucial. In SAGIN, nodes such as satellites, UAVs, and terrestrial base stations may be affected by various factors, such as natural environment, hardware failures, and network attacks. Therefore, the design of fault-tolerant and secure network architecture is essential to ensure the stable operation of the network and the security of user data.

Literature [9] proposes a flexible, low-latency and flat SAGIN architecture for 6G. Addressing the problems of existing 5G non-terrestrial network (NTN) architectures in terms of

Table 4: Comparison of Literature Methods

Key Method	Application Scenario	Advantage
SAGIN architecture combining blockchain and MEC, utilizing LEO satellites and UAVs as edge nodes	Task offloading and wireless data transmission in SAGIN networks	Ensures trust in task offloading and wireless data transmission through blockchain technology; reduces task execution delay and system energy consumption using MEC technology.
Proposing a three-layer SAGIN network architecture (space layer, air layer, ground layer) with widespread application of MEC technology at various edge nodes	SAGIN network architecture design, key technologies, and challenge analysis	Provides comprehensive architectural design guidance; analyzes multiple MEC deployment models (single-edge, dual-edge, multi-edge computing), meeting diverse application scenario requirements.
Classifying and summarizing existing resource allocation strategies, including mathematical optimization, dynamic optimization, game theory, and artificial intelligence methods	Resource allocation strategies in SAGIN	Offers a comprehensive overview of resource allocation strategies; identifies future research directions, particularly in resource optimization under 6G communication technology.
Task offloading and resource allocation optimization method based on Deep Reinforcement Learning (DRL), modeled as a Markov Decision Process (MDP)	Task offloading and resource allocation in edge-cloud collaborative computing environments	Effectively reduces task processing time and energy consumption of mobile devices through intelligent task offloading decisions and resource allocation strategies, optimizing overall system performance.
Proposing an alternating optimization approach to optimize user association, partial offloading control, computing resources and bandwidth allocation, and UAV deployment	Computing offloading, UAV deployment, and resource allocation optimization in SAGIN	Demonstrates effectiveness through numerical studies, significantly improving weighted energy efficiency by approximately 35%-40% compared to baseline schemes.

latency and flexibility, the article designs a novel SAGIN architecture that integrates the core network (CN) and radio access network (RAN) into a single network consisting of a service-based control plane (SBCP), a service-based user plane (SBUP), and radio units (RUs). With this design, data transmission latency is reduced, network management flexibility is increased, and the protocol stack is simplified. The integration methodology of the new architecture, the design of new interfaces and protocol stacks are elaborated in detail, and the superiority of the new architecture is verified by analyzing the application requirements and technical needs of future 6G. In addition, research challenges and directions for future intelligent deployment of network elements, mobility management, and intelligent session management are discussed, providing comprehensive guidance and reference for the development of the 6G SAGIN architecture.

Literature [81] investigates SatBFT, a blockchain-based consensus protocol in SAGIN, aiming at efficient and secure spectrum sharing. Aiming at the challenge of reaching consensus on spectrum resource allocation strategies in SAGIN, a multi-layered architecture of the SatBFT consensus protocol is proposed, which combines radio environment sensing, dynamic behavioral assessment, and an efficient block generation mechanism. Specifically, the SatBFT protocol consists

of three layers: the link layer, the block generation layer, and the application layer, each of which is responsible for a specific blockchain spectrum management function. The effectiveness of the SatBFT protocol in improving spectrum utilization, reducing delay and enhancing security is verified through simulation. The simulation results show that the SatBFT protocol can significantly improve the overall performance of the SAGIN network and provide a new solution for the dynamic allocation and management of spectrum resources.

To better understand the diversity of approaches in network architecture design and optimization, Table 4 presents a comparative analysis of the key methods discussed in recent literature, highlighting their respective application scenarios and advantages.

3.4 Communication Protocols and Standardization

3.4.1 Current status and development trend of MEC and SAGIN related communication protocols

In the field of MEC (Mobile Edge Computing), the current major communication protocols include HTTP/2 for data transmission, gRPC, etc. The HTTP/2 protocol improves the data transmission efficiency through binary frame splitting,

header compression, etc., and can better support the fast data interaction between edge devices and servers in the MEC scenarios. gRPC, as a high-performance RPC framework It supports multiple programming languages and can realize efficient communication between edge devices and edge servers [82]. With the development of MEC technology, communication protocols are also evolving. On the one hand, protocols will focus more on low latency and high reliability to meet edge computing applications with high real-time requirements, such as industrial automation and autonomous driving. For example, URLLC (Ultra-Reliable Low-Latency Communication) technology in 5G networks provides MEC with latency guarantees as low as 1 ms, and the related communication protocols are being optimized to adapt to this low-latency demand. On the other hand, the security of the protocols will be further enhanced to cope with potential security threats in the edge computing environment. For example, the Transport Layer Security (TLS) protocol is used to encrypt the communication data to ensure the confidentiality and integrity of the data during transmission [83].

In SAGIN, a variety of communication protocols are involved. In the space segment, satellite communications mainly use protocols such as DVB-S2 and CCSDS, which are widely used in satellite digital video broadcasting and can provide highly efficient data transmission and good anti-jamming capability [84, 85], while the CCSDS protocol is mainly used for communications between space data systems to ensure reliable data transmission between satellites and ground stations [86]. In the air segment, UAV communication often uses variants of the 802.11 protocol, such as 802.11ad, 802.11ay, etc. These protocols support communication in high-frequency bands, and are capable of realizing high data transmission rates. In the ground segment, it mainly relies on cellular network protocols such as 4G/5G. In the future, SAGIN communication protocols will develop in the direction of convergence and intelligence. On the one hand, the protocol will pay more attention to the seamless integration between different network segments to realize the efficient synergy of space, air and ground networks. For example, by developing a unified network interface protocol, it enables satellites, drones and ground base stations to seamlessly interface and form an organic whole. On the other hand, the protocol will introduce artificial intelligence technology to realize intelligent resource allocation and traffic scheduling. For example, machine learning-based traffic prediction algorithms can predict changes in network traffic in advance, thereby optimizing the parameter configuration of the communication protocol and improving the overall performance of the network.

3.4.2 Relevant technical standards and progress by international organizations

The Third Generation Partnership Project (3GPP) is one of the important international organizations in the development of MEC technical standards. 3GPP introduced MEC-related standards in version R16, defining the architecture and interfaces of MEC systems[87]. European Telecommunication Standardisation Institute (ETSI) is also actively involved in the development of MEC standards, and its MEC Industry Specification Group (ISG) has released a number

of MEC-related specifications and guidelines. International Telecommunication Union (ITU) plays an important role in the development of SAGIN technical standards, and its satellite communication standards provide the basis for SAGIN space segment communications. ITU plays an important role in the development of SAGIN technical standards, and its satellite communication standards provide the basis for the space segment communication of SAGIN. In addition, 3GPP is also actively promoting the development of SAGIN-related standards, especially in non-terrestrial network (NTN).

At present, MEC technical standards are evolving and improving. On the one hand, standards organizations are promoting the deep integration of MEC with 5G networks and formulating relevant technical specifications to achieve efficient deployment and operation of MEC in 5G networks. For example, 3GPP is studying how to better support the computational offload function of MEC in 5G networks. On the other hand, standards are also exploring the integration of MEC with other emerging technologies, such as artificial intelligence and blockchain. For example, ETSI is studying how to apply blockchain technology to the security and trust management of MEC. SAGIN technology standards are moving toward a more comprehensive and systematic approach. On the one hand, standards organizations are formulating unified network architecture standards to achieve organic integration of space, air, and ground networks. For example, 3GPP is studying how to incorporate satellite networks and drone networks into the overall architecture of 5G networks. On the other hand, standards are constantly optimizing communication protocols to improve the performance and efficiency of SAGIN. For example, ITU is studying new satellite communication protocols to support higher data rates and lower latency.

Table 5 summarizes and compares the key methods and protocols discussed in this section, providing insights into their specific applications and technical advantages in SAGIN environments.

3.5 Integration of Deep Learning with SAGIN-MEC

In recent years, the integration of deep learning (DL) methods with SAGIN-MEC systems has significantly enhanced their adaptability, robustness, and intelligence. Building upon the fundamental SAGIN-MEC components discussed in previous sections, this section systematically reviews the latest advances at the intersection of DL and SAGIN-MEC, critically examining recent progress, unique contributions, and existing limitations.

3.5.1 Adaptive Computation Offloading Based on Deep Reinforcement Learning

Adaptive computation offloading strategies play a crucial role in improving the overall performance of SAGIN-MEC systems under dynamic and heterogeneous environments. In recent years, deep reinforcement learning (DRL) has been widely adopted in computation offloading due to its ability to autonomously learn optimal policies through continuous interaction with uncertain and highly dynamic environments.

Ke et al. [88] addressed the task offloading problem in heterogeneous vehicular networks by proposing an adaptive

Table 5: Comparison of Key Methods in References

Key Method	Scenario	Advantage
Integrating the core network (CN) and the radio access network (RAN) into a single network in 6G-SAGIN architecture design	6G-SAGIN architecture design	Reduces data transmission delay; increases network management flexibility; simplifies the protocol stack; enhances overall system performance.
SatBFT consensus protocol based on blockchain	Spectrum sharing in SAGIN	Improves spectrum utilization; reduces communication delay; enhances security; effectiveness verified through simulation experiments.
Mentioning HTTP/2 and gRPC as MEC communication protocols	Data transmission and communication in MEC scenarios	HTTP/2 increases data transmission efficiency via binary framing and header compression; gRPC supports multiple programming languages, facilitating efficient communication between edge devices and servers.
DVB-S2, CCSDS for satellite communication, and IEEE 802.11 protocol variants for UAV communication	Communication in different network segments of SAGIN	Provides standardized protocols for effective data transmission; robust anti-interference capabilities; ensures reliable communication among satellites, UAVs, and ground stations.

computation offloading method (ACORL) based on DRL, which leverages the Ornstein-Uhlenbeck process for noise generation to effectively enhance exploration in continuous action spaces.

Qiu et al. [89] designed an end-edge-cloud collaborative offloading model in dynamic heterogeneous vehicular edge computing (VEC) networks, and proposed an adaptive computation offloading and power allocation (DDPG-ACOPA) scheme based on the deep deterministic policy gradient (DDPG) algorithm. By formulating the problem as a Markov decision process, their scheme jointly considers stochastic task arrivals, channel variation, and vehicle mobility, achieving minimization of system-wide latency and energy consumption.

Considering task heterogeneity, Wang et al. [90] modeled mobile applications as directed acyclic graphs (DAGs) and adopted a meta-reinforcement learning (Meta-RL) based offloading method. Their approach employs a customized Seq2Seq neural network, combined with first-order approximation and truncated surrogate objectives, which significantly accelerates adaptation and improves sample efficiency in new environments.

Despite notable theoretical and simulation advances achieved by DRL-based adaptive computation offloading, practical deployment still faces numerous challenges, such as high computational resource consumption, slow training convergence, and limited model interpretability. Future research should focus on developing efficient, lightweight, and interpretable DRL models, especially suitable for resource-constrained aerial and satellite platforms, to facilitate the practical application and deployment of SAGIN-MEC systems.

3.5.2 Intelligent Resource Allocation via Multi-Agent DRL

Resource allocation and management in SAGIN-MEC face tremendous challenges due to the heterogeneity across integrated segments and the dynamic availability of resources.

Liao et al. [91] proposed a cooperative multi-agent deep reinforcement learning (CMDRL) method, modeling the bandwidth allocation problem as a multi-agent cooperative task. This approach not only significantly improves transmission efficiency but also exhibits strong adaptability to objectives and low implementation complexity.

In vehicular network scenarios, unmanned aerial vehicles (UAVs) serve as aerial base stations to provide network access and edge computing services for vehicles, especially in the absence of road-side units (RSUs). For multi-UAV cooperative service to vehicular terminals, Zhang et al. [92] introduced an enhanced multi-agent deep deterministic policy gradient (MADDPG) algorithm, which optimizes trajectory planning, spectrum allocation, and dynamic data offloading, thereby significantly reducing task completion latency.

Furthermore, in IoT edge networks, given the computational and energy constraints of end devices, efficient task offloading and resource allocation are critical. Seid et al. [93] put forward a clustered multi-UAV MADRL method, jointly formulating resource allocation and computation offloading as a stochastic game in the Markov decision process (MDP) framework. The approach fully considers UAV channel time-variability and dynamic resource demands, achieving a 38.6%–55.6% reduction in average cost and a 58.3%–85.3% increase in reward compared to single-agent DRL and traditional heuristic methods.

Overall, the MADRL framework effectively enables collaborative resource optimization in complex, multi-domain SAGIN-MEC systems, significantly improving resource utilization and allocation fairness. Nevertheless, as system scale

increases, key challenges remain, including scalability, convergence stability, and inter-agent coordination overhead. Future research may further explore hierarchical or distributed DRL architectures and lightweight neural network models to balance system performance and deployment feasibility.

In summary, the integration of deep learning with SAGIN-MEC has brought transformative solutions for adaptive offloading, cooperative resource management, network optimization, proactive management, and secure collaboration. By effectively addressing the inherent dynamics, heterogeneity, and complexity of SAGIN-MEC, these DL-driven methods clearly outperform traditional approaches. However, several key challenges remain unresolved, including computational complexity, scalability, data scarcity, robustness, security vulnerabilities, and interpretability constraints. Further exploration of lightweight network architectures, efficient training paradigms (e.g., meta-learning, transfer learning), robust and interpretable DL frameworks, and comprehensive hybrid deep learning optimization methods will be indispensable.

4 Application

4.1 Intelligent Transportation and Automatic Driving

In typical vehicular networks (V2X) architectures, sensor data collected by vehicles—including cameras, LiDAR, and millimeter-wave radar—is traditionally uploaded to cloud servers or regional computing centers for processing. The resulting decisions are then transmitted back to the vehicles [94]. However, this approach suffers from considerable uplink latency and may result in blind spots or perception dead zones.

To address these limitations, MEC nodes can perform real-time inference. MEC servers are deployed at roadside units (RSUs) such as curbsides, toll stations, and charging piles. Once a vehicle establishes a direct V2I (Vehicle-to-Infrastructure) connection with a nearby base station, sensor data can be locally processed within a few milliseconds [95, 96].

SAGIN further extend the perception capabilities by leveraging mobile satellite platforms and airborne platforms (APs). For example, long-endurance UAVs equipped with high-resolution cameras or millimeter-wave radars can be deployed along major transport hubs or highway loops to monitor vehicle traffic across tens of square kilometers in real time [97]. The UAVs transmit image data to underlying MEC nodes, which perform tasks such as vehicle detection and trajectory prediction. The results are then aggregated at ground-based command centers.

To mitigate communication blind spots, particularly in suburban areas, mountainous regions, and remote highway segments, where terrestrial 4G/5G coverage is often weak or intermittent LEO/MEO satellite links within the SAGIN framework provide uninterrupted connectivity. This ensures consistent communication support for autonomous vehicles, even under extreme weather conditions or in challenging environments such as tunnels and mountainous roads. Through

dynamic satellite link switching, vehicles can access traffic information from virtually any location [98].

Compared to conventional architectures that rely solely on cloud computing or ground-based perception, the integration of MEC with SAGIN enables millisecond-level response times during critical moments. This hybrid approach overcomes the limitations of terrestrial networks and offers robust support for intelligent transportation and autonomous driving, particularly in remote and complex terrains.

4.2 Emergency Communication and Disaster Response

Terrestrial communication infrastructure is highly susceptible to damage or congestion during natural disasters such as earthquakes, floods, and wildfires, as well as unexpected emergencies. Traditional emergency communication strategies often rely on temporary base stations or direct satellite communication, both of which face significant challenges. Base station repairs are costly and time-consuming [99], while satellite links often provide limited bandwidth and high latency, hindering the transmission of large-scale data essential for disaster assessment and rescue coordination [100, 101].

SAGIN offers a superior alternative by enabling the rapid deployment of multi-tiered emergency communication networks [102]. When ground base stations are completely disabled, LEO satellite links can be immediately activated to establish a backhaul network for the affected area. Concurrently, UAVs can form high-density communication meshes over disaster zones, linking on-site terminals—including tablets, mobile phones, and ambulance communication devices—to satellite backhaul networks.

Under the SAGIN paradigm, once the terrestrial network is partially restored, it collaborates with satellite and UAV relays to form a resilient, multi-path redundant communication infrastructure. Edge computing tasks can dynamically switch between available links to optimize performance and reliability [29, 103]. For example, when UAVs experience power depletion, the system can automatically reroute data through MEO satellites. Similarly, once 5G base stations resume operation, communication can be preferentially shifted to terrestrial networks to reduce overhead.

4.3 Telemedicine and Remote Education

In remote or resource-limited areas, traditional telemedicine and education typically rely on limited terrestrial internet connectivity, where fiber-optic or public LTE/5G networks are difficult to deploy, and sometimes ADSL or satellite connections are used [104]. However, fiber deployment is expensive and time-consuming, while direct satellite links have limited bandwidth and high latency (round-trip delay of approximately 250 ms–600 ms) [105, 106], resulting in severe video stuttering and image lag. Due to centralized computing bottlenecks, a large volume of medical images, such as CT, MRI, and ultrasound scans, must be transmitted to central hospitals for analysis by remote experts or cloud-based AI models, which is often time-consuming and unfavorable for urgent diagnosis and decision-making. In addition, most schools in remote mountainous or island regions rely on pre-recorded

courses, making real-time interaction difficult [107]. A substantial amount of multimedia educational content must be retrieved from urban central education platforms, leading to high data costs and vulnerability to network jitter.

Edge AI models are redefining the paradigm of medical diagnosis. By deploying MEC servers in county-level hospitals or township clinics [108, 109], pre-trained medical image segmentation and classification models can locally accomplish tasks such as lung CT lesion detection and key feature extraction from fetal ultrasound images. Only critical information, such as “lesion location” and “lesion grade,” is uploaded to municipal or provincial centers, thus conserving bandwidth and significantly reducing diagnostic time. Furthermore, compared to traditional educational networks, MEC nodes in remote schools can pre-cache popular courses—such as mathematics, English, and physics demonstration videos—from municipal education cloud platforms or county-level live broadcast centers. When students request these resources, educational terminals access them locally, thereby avoiding classroom stutter or waiting caused by network jitter. Additionally, the combination of LEO satellite links and terrestrial MEC can provide stable video conferencing and multimedia transmission capabilities.

4.4 Aerial and Maritime Mobile Services

On high-speed mobile platforms such as airplanes, high-speed rail, and ships, passengers often experience limited network quality [9, 110]. For example, high-speed trains may encounter signal blind spots in tunnels, mountainous regions, or bridges; after takeoff, aircraft cannot connect to ground networks for a short period and must rely on satellite links; and ocean-going vessels lose all terrestrial signals after leaving port. Traditional onboard Wi-Fi or shipboard networks mostly use terrestrial 4G/5G relay, which can suffer packet loss when traversing different base stations. Although direct satellite links provide global coverage, they feature high round-trip latency (about 250 ms–600 ms) and expensive, limited bandwidth. Airborne MEC nodes can effectively alleviate resource constraints; when network latency fluctuates, users can swiftly switch to locally cached offline content in MEC, thus avoiding prolonged waiting. Satellite remote sensing enables rapid acquisition of large-scale meteorological data, such as cloud cover, wind speed, and atmospheric pressure [111]; HAPs are responsible for three-dimensional regional meteorological detection; airborne MEC, combined with aircraft sensors (e.g., wing accelerometers, barometers, thermometers), can perform local real-time meteorological inference to assess risks such as turbulence and icing [112]. Each vessel can deploy a compact MEC node; near the sea surface, HAPs serve as aerial relays or connect via LEO satellite links, forming a multi-link mesh within the fleet. Ships can locally complete intelligence fusion, situational assessment, and coordinated decision-making, thereby reducing dependence on shore-based centers.

Table 6 provides a comprehensive comparative analysis of the application scenarios discussed in this section, summarizing the main challenges, technical solutions, and remaining issues for each scenario in MEC-enabled SAGIN systems.

5 Challenges

The integration of SAGIN with MEC presents numerous open challenges that require detailed analysis and targeted solutions. In this section, we explicitly clarify each challenge, providing their detailed manifestations and underlying causes.

5.1 High Mobility and Dynamic Network Conditions

Due to the high mobility of aerial nodes (UAVs, HAPs) and satellites (particularly LEO satellites), SAGIN exhibits frequent topology variations and intermittent connectivity. This leads to unstable communication links and unpredictable network states, severely complicating the timely offloading of computational tasks and the effective allocation of resources. Traditional static or semi-static scheduling approaches lack adaptability in rapidly changing SAGIN environments, resulting in increased latency and degraded user experience [6, 8].

5.2 Complexity of Joint Computation and Communication Optimization

In SAGIN-MEC environments, computation and communication resources are strongly coupled and interdependent. This joint optimization complexity arises from the necessity to simultaneously consider multiple conflicting objectives—including latency reduction, throughput maximization, and energy efficiency—under continuously fluctuating network conditions and service demands. Existing optimization methods often fail to adequately balance these objectives, resulting in suboptimal overall performance and inefficient resource utilization [5].

5.3 Challenges in AI Integration and Data Availability

Integrating artificial intelligence into SAGIN-MEC systems faces practical difficulties due to limited availability of labeled datasets that accurately represent dynamic, multi-layered SAGIN scenarios. Additionally, common AI models are computationally intensive, making deployment challenging on resource-constrained aerial and satellite platforms. Furthermore, the absence of unified cross-layer AI frameworks constrains collaborative decision-making and limits the effectiveness and scalability of current AI solutions in heterogeneous SAGIN environments [113, 114].

5.4 Security Threats and Privacy Preservation Issues

The integration of heterogeneous network segments (space, air, ground) expands the potential vulnerability surface, exposing SAGIN-MEC systems to threats such as satellite communication interference, UAV hijacking, and data breaches. Privacy-preserving mechanisms in task offloading and data processing face additional challenges from varying privacy requirements, dynamic link-level threats, and the computational overhead associated with encryption methods. Most existing approaches cannot simultaneously satisfy stringent

Table 6: Comparative Analysis of Typical Application Scenarios in MEC-enabled SAGIN

Scenario	Main Issues	Key Technical Solutions	Remaining Challenges
Intelligent Transportation	Network blind spots, latency-sensitive control	UAV-MEC deployments, LEO satellite integration, edge AI models	Seamless cross-tier handover, real-time dynamic topology management
Emergency Communication	Infrastructure damage, limited bandwidth	UAV relay networks, satellite backhaul, MEC-based local processing	Efficient bandwidth allocation, energy-limited UAV endurance
Telemedicine & Remote Education	Limited terrestrial coverage, high latency	MEC-based local caching, satellite uplink optimization	Bandwidth-constrained satellite links, standardized data sharing protocols
Aerial & Maritime Mobile Services	High mobility, intermittent connectivity	Satellite links, MEC nodes on vehicles, local caching	High latency of GEO satellite links, dynamic edge resource allocation

latency, security, and privacy requirements in dynamic SAGIN environments [115].

5.5 UAV-Assisted MEC Deployment Difficulties

Deploying MEC platforms on UAVs introduces several practical constraints, including limited onboard computing capacities, strict energy constraints, and complex topology management due to UAV mobility. Ensuring stable, efficient, and adaptive task offloading between UAVs and terrestrial nodes remains challenging. Moreover, the lack of standardized and unified MEC-UAV coordination frameworks limits practical scalability and real-world deployment feasibility [116].

5.6 Hybrid Optimization Problem Complexity

The integration of discrete offloading decisions (such as choosing which tasks to offload) and continuous resource allocations (such as bandwidth and computing frequency) creates complex, non-convex optimization problems. Current Multi-Agent Reinforcement Learning (MARL) algorithms experience significant challenges handling these hybrid discrete-continuous action spaces, restricting their effectiveness in large-scale SAGIN resource management scenarios [117].

5.7 Application Placement Challenges

Efficient application placement across SAGIN-MEC systems is complicated by heterogeneous and dynamically changing resource availabilities at different network layers. Diverse real-time application requirements, fluctuating network conditions, and stringent latency constraints lead to imbalanced resource usage and increased computational overhead. Existing methods generally fail to achieve real-time adaptability and scalability in handling multi-objective placement optimization problems [118].

5.8 Robustness and Security Issues in Reinforcement Learning Approaches

Applying reinforcement learning (RL) to SAGIN-MEC faces challenges due to node mobility, intermittent connectivity, and diverse hardware constraints, all of which complicate the design of generalized and scalable RL models. Furthermore, ensuring robust, secure, and privacy-preserving RL becomes essential when sensitive data is processed across distributed SAGIN nodes, posing additional practical challenges to existing RL methods [119].

6 Future Research

Building upon the challenges and limitations identified in previous sections, we propose the following detailed and actionable research directions to advance the integration of MEC and SAGIN.

6.1 Intelligent and Adaptive Resource Scheduling

Due to the dynamic topology induced by high mobility of aerial and satellite nodes, traditional static scheduling schemes are inadequate. Specific research questions include:

- How can real-time predictive algorithms be developed to forecast topology changes and resource availability effectively?
- What lightweight AI methods can efficiently predict Inter-Satellite Link (ISL) conditions, enabling stable and efficient task offloading?

Potential research approaches include integrating digital twin technology with Multi-Agent Deep Reinforcement Learning (MADRL) methods for proactive resource adaptation. Federated learning frameworks could also facilitate distributed and adaptive decision-making across heterogeneous SAGIN nodes [120].

6.2 Robust Security and Privacy Mechanisms

Security risks across heterogeneous SAGIN domains remain significant. Critical future research directions include:

- How can lightweight blockchain protocols, such as enhanced SatBFT, ensure decentralized authentication and security auditing?
- How can Quantum Key Distribution (QKD) techniques be integrated to secure sensitive data transmissions between satellites and MEC nodes?
- How can satellite-based global threat intelligence dynamically enhance MEC security policies?

Promising approaches include blockchain-based decentralized frameworks and quantum cryptography advancements to enhance secure communication channels and mitigate emerging threats [121].

6.3 Energy-Efficient MEC Architectures

Energy efficiency remains critical for resource-constrained aerial and satellite MEC nodes. Specific research questions include:

- How can renewable energy harvesting technologies effectively extend UAV and satellite operational lifetimes?
- What scheduling algorithms can optimize satellite intermittent computing capabilities while ensuring service continuity?
- How can satellite-aided renewable energy forecasting enhance UAV deployment and task scheduling efficiency?

Future research could focus on solar-powered UAV MEC platforms, intermittent-computing-aware scheduling methods, and dynamic workload migration strategies favoring renewable-powered terrestrial data centers [122].

6.4 6G-Enabled Converged Network Architectures

Emerging 6G technologies present significant opportunities to enhance SAGIN-MEC integration. Critical research questions include:

- How can Reconfigurable Intelligent Surfaces (RIS) deployed on UAVs dynamically enhance wireless communication channels?
- What semantic communication techniques can reduce bandwidth demands by extracting and transmitting essential data features?
- How can holographic beamforming optimize multi-hop MEC links across space-air-ground layers?

Potential paths include developing RIS-enhanced UAV platforms, semantic-aware communication protocols, and advanced holographic beamforming solutions [9, 123].

6.5 Resilient MEC Architectures in Extreme Scenarios

Enhancing SAGIN-MEC resilience under extreme conditions is crucial. Specific research topics include:

- How can autonomous UAV swarms and satellite-edge cached containers quickly form self-organizing emergency MEC networks?
- What unified resilience metrics (e.g., Service Survival Time per Recovery Energy) can standardize robustness evaluation?
- How can simulation platforms such as OMNeT++ effectively validate resilient SAGIN-MEC architectures?

Future studies could explore autonomous UAV-swarm algorithms, edge-caching strategies, and advanced simulation-based evaluation frameworks [124].

6.6 Unified Standardization and Protocol Innovation

Standardization is essential for seamless MEC-SAGIN integration. Key research considerations include:

- How can a unified API framework based on gRPC facilitate seamless task migration across heterogeneous nodes?
- What Delay-Tolerant Networking (DTN) protocols can reliably ensure computation and communication in intermittently connected networks?
- How can international organizations (e.g., 3GPP, ITU) collaboratively develop comprehensive MEC standards tailored for Non-Terrestrial Networks (NTN)?

Future research could focus on unified API definitions, reliable DTN protocols, and collaborative standardization efforts [125].

7 Conclusion

The integration of MEC into SAGIN is positioned as a pivotal approach to addressing the growing demands for seamless global coverage, low latency, and high computational efficiency in next-generation networks. This survey provides a comprehensive analysis and synthesis of existing literature, emphasizing the technological foundations, recent advancements, practical applications, critical challenges, and future research directions related to MEC-enabled SAGIN.

Through an extensive review, this article elaborates on core MEC functionalities within SAGIN, such as computation offloading, resource management, network architecture optimization, and protocol standardization. The detailed exploration of these components underscores the complexity and heterogeneity inherent in SAGIN, highlighting the necessity for sophisticated and adaptive technological solutions.

The presented application scenarios, including intelligent transportation, emergency communications, telemedicine, remote education, and aerial and maritime services, clearly illustrate the transformative potential of MEC-SAGIN integrations. These scenarios demonstrate how MEC significantly enhances system responsiveness, resilience, and efficiency, particularly in environments characterized by limited terrestrial coverage or stringent latency requirements.

However, several critical challenges persist, notably dynamic resource scheduling, energy constraints, security and privacy vulnerabilities, and seamless cross-tier coordination. These issues require further investigation and innovative

solutions, particularly through advanced methodologies incorporating artificial intelligence, blockchain, and quantum technologies.

Finally, this survey identifies promising research directions, advocating for advancements in intelligent resource scheduling using digital twins and reinforcement learning, robust security frameworks leveraging blockchain and quantum cryptography, energy-efficient operations via renewable energy integration, and unified standardization efforts for seamless interoperability across SAGIN layers.

In summary, MEC and SAGIN integration represents a compelling and necessary evolution toward ubiquitous, intelligent, and resilient communication and computation infrastructures. Continued research and development in this field will undoubtedly play a critical role in shaping future telecommunications landscapes, ultimately delivering unprecedented levels of connectivity, intelligence, and reliability on a global scale.

Funding

This work is supported by the National Natural Science Foundation of China under Grant 62471493, and partially supported by the Natural Science Foundation of Shandong Province under Grant ZR2023LZH017, ZR2024MF066.

Author Contributions

Conceptualization, P.Z. and A.V.V.; methodology, P.Z. and S.C.; software, S.C. and Y.Y.; validation, P.Z., G.W. and L.C.; formal analysis, P.Z. and G.W.; investigation, P.Z., S.C. and Y.Y.; resources, A.V.V. and L.C.; data curation, S.C. and Y.Y.; writing—original draft preparation, P.Z. and S.C.; writing—review and editing, P.Z., G.W. and A.V.V.; visualization, S.C. and Y.Y.; supervision, A.V.V.; project administration, P.Z.; funding acquisition, P.Z. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data Available

The data and materials used in this study are available upon request from the corresponding author.

References

- [1] Rose, K., Eldridge, S., Chapin, L.: The internet of things: An overview. Technical report, Internet Society (ISOC) (2015). <https://www.internetsociety.org/resources/doc/2015/iot-overview/> Accessed 2025-05-15
- [2] Salahdine, F., Han, T., Zhang, N.: 5g, 6g, and beyond: Recent advances and future challenges. *Annals of Telecommunications* **78**(9), 525–549 (2023)
- [3] Liu, W., Zhang, X., Ren, J., Shen, Y., Wang, S., Yang, B., Guan, X., Cui, S.: Energy-efficient multi-uav-enabled mec systems with space-air-ground integrated networks. *arXiv preprint arXiv:2409.14782* (2024)
- [4] Du, J., Wang, J., Sun, A., Qu, J., Zhang, J., Wu, C., Niyato, D.: Joint optimization in blockchain and mec enabled space-air-ground integrated networks. *IEEE Internet of Things Journal* **11**(19), 31862–31877 (2024) <https://doi.org/10.1109/JIOT.2024.3421529>
- [5] Chen, Q., Guo, Z., Meng, W., Han, S., Li, C., Quek, T.Q.: A survey on resource management in joint communication and computing-embedded sagin. *IEEE Communications Surveys & Tutorials* **27**(3), 1911–1954 (2025) <https://doi.org/10.1109/COMST.2024.3421523>
- [6] Liang, H., Yang, Z., Zhang, G., Hou, H.: Resource allocation for space-air-ground integrated networks: A comprehensive review. *Journal of Communications and Information Networks* **9**(1), 1–23 (2024)
- [7] Chen, J., Zhang, H., Xie, Z.: Space-air-ground integrated network (sagin): A survey. *arXiv preprint arXiv:2307.14697* (2023)
- [8] Xiao, Y., Ye, Z., Wu, M., Li, H., Xiao, M., Alouini, M.-S., Al-Hourani, A., Cioni, S.: Space-air-ground integrated wireless networks for 6g: Basics, key technologies and future trends. *IEEE Journal on Selected Areas in Communications* **42**(12), 3327–3354 (2024) <https://doi.org/10.1109/JSAC.2024.3492720>
- [9] Cui, H., Zhang, J., Geng, Y., Xiao, Z., Sun, T., Zhang, N., Liu, J., Wu, Q., Cao, X.: Space-air-ground integrated network (sagin) for 6g: Requirements, architecture and challenges. *China Communications* **19**(2), 90–108 (2022)
- [10] Qiu, Y., Niu, J., Zhu, X., Zhu, K., Yao, Y., Ren, B., Ren, T.: Mobile edge computing in space-air-ground integrated networks: Architectures, key technologies and challenges. *Journal of Sensor and Actuator Networks* **11**(4), 57 (2022)
- [11] Zhang, P., Chen, S., Zheng, X., Li, P., Wang, G., Wang, R., Wang, J., Tan, L.: Uav communication in space-air-ground integrated networks (sagins): Technologies, applications, and challenges. *Drones* **9**(2) (2025) <https://doi.org/10.3390/drones9020108>
- [12] Del Portillo, I., Cameron, B.G., Crawley, E.F.: A technical comparison of three low earth orbit satellite constellation systems to provide global broadband. *Acta astronautica* **159**, 123–135 (2019)
- [13] Kodheli, O., Lagunas, E., Maturo, N., Sharma, S.K., Shankar, B., Montoya, J.F.M., Duncan, J.C.M., Spano, D., Chatzinotas, S., Kisseleff, S., *et al.*: Satellite communications in the new space era: A survey and future

- challenges. *IEEE Communications Surveys & Tutorials* **23**(1), 70–109 (2020)
- [14] Jia, Z., Sheng, M., Li, J., Zhou, D., Han, Z.: Joint hap access and leo satellite backhaul in 6g: Matching game-based approaches. *IEEE Journal on Selected Areas in Communications* **39**(4), 1147–1159 (2020)
- [15] Zeng, Y., Wu, Q., Zhang, R.: Accessing from the sky: A tutorial on uav communications for 5g and beyond. *Proceedings of the IEEE* **107**(12), 2327–2375 (2019) <https://doi.org/10.1109/JPROC.2019.2952892>
- [16] Alam, M.S., Kurt, G.K., Yanikomeroglu, H., Zhu, P., Đào, N.D.: High altitude platform station based super macro base station constellations. *IEEE Communications Magazine* **59**(1), 103–109 (2021) <https://doi.org/10.1109/MCOM.001.2000542>
- [17] Giordani, M., Polese, M., Mezzavilla, M., Rangan, S., Zorzi, M.: Toward 6g networks: Use cases and technologies. *IEEE communications magazine* **58**(3), 55–61 (2020)
- [18] Nuriev, M., Kalyashina, A., Smirnov, Y., Gumerova, G., Gadzhieva, G.: The 5g revolution transforming connectivity and powering innovations. In *International Scientific Conference Transport Technologies in the 21st Century*, pp. 04008–10 (2024). EDP Sciences
- [19] Raza, U., Kulkarni, P., Sooriyabandara, M.: Low power wide area networks: An overview. *IEEE communications surveys & tutorials* **19**(2), 855–873 (2017)
- [20] Veeramachaneni, V.: Edge computing: Architecture, applications, and future challenges in a decentralized era. *Recent Trends in Computer Graphics and Multimedia Technology* **7**(1), 8–23 (2025)
- [21] Abbas, N., Zhang, Y., Taherkordi, A., Skeie, T.: Mobile edge computing: A survey. *IEEE Internet of Things Journal* **5**(1), 450–465 (2017)
- [22] Mao, Y., You, C., Zhang, J., Huang, K., Letaief, K.B.: A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys Tutorials* **19**(4), 2322–2358 (2017) <https://doi.org/10.1109/COMST.2017.2745201>
- [23] Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L.: Edge computing: Vision and challenges. *IEEE internet of things journal* **3**(5), 637–646 (2016)
- [24] Mach, P., Becvar, Z.: Mobile edge computing: A survey on architecture and computation offloading. *IEEE communications surveys & tutorials* **19**(3), 1628–1656 (2017)
- [25] Wang, X., Han, Y., Leung, V.C., Niyato, D., Yan, X., Chen, X.: Convergence of edge computing and deep learning: A comprehensive survey. *IEEE communications surveys & tutorials* **22**(2), 869–904 (2020)
- [26] Tran, T.X., Hajisami, A., Pandey, P., Pompili, D.: Collaborative mobile edge computing in 5g networks: New paradigms, scenarios, and challenges. *IEEE Communications Magazine* **55**(4), 54–61 (2017)
- [27] McDowell, J.C.: The low earth orbit satellite population and impacts of the spacex starlink constellation. *The Astrophysical Journal Letters* **892**(2), 36 (2020) <https://doi.org/10.3847/2041-8213/ab8016>
- [28] Wang, X., Shen, T., Zhang, Y., Chen, X.: An efficient topology emulation technology for the space-air-ground integrated network. In *IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–8 (2023). <https://doi.org/10.1109/INFOCOMWKSHPS57453.2023.10225913>
- [29] Marinho, R.P., Vieira, L.F.M., Vieira, M.A.M., Loureiro, A.A.F.: Sagin-cain: A 3d routing protocol for post-disaster sagin 6g network. In *2024 20th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*, pp. 685–692 (2024). <https://doi.org/10.1109/DCOSS-IoT61029.2024.00106>
- [30] Jia, H., Wang, Y., Wu, W.: Dynamic resource allocation for remote iot data collection in sagin. *IEEE Internet of Things Journal* **11**(11), 20575–20589 (2024) <https://doi.org/10.1109/JIOT.2024.3371486>
- [31] Wang, C., Pang, M., Wu, T., Gao, F., Zhao, L., Chen, J., Wang, W., Wang, D., Zhang, Z., Zhang, P.: Resilient massive access for sagin: A deep reinforcement learning approach. *IEEE Journal on Selected Areas in Communications* **43**(1), 297–313 (2025) <https://doi.org/10.1109/JSAC.2024.3460030>
- [32] Du, J., Wang, J., Sun, A., Qu, J., Zhang, J., Wu, C., Niyato, D.: Joint optimization in blockchain- and mec-enabled space-air-ground integrated networks. *IEEE Internet of Things Journal* **11**(19), 31862–31877 (2024) <https://doi.org/10.1109/JIOT.2024.3421529>
- [33] Qin, P., Li, H., Fu, Y., Hu, J., Wu, X., Zhang, X.: Learning-based noma-enabled queue-aware task offloading and uav 3d trajectory planning for sagin. *IEEE Transactions on Vehicular Technology* **74**(8), 12364–12375 (2025) <https://doi.org/10.1109/TVT.2025.3552807>
- [34] Zhang, S., Yi, N., Ma, Y.: A survey of computation offloading with task types. *IEEE Transactions on Intelligent Transportation Systems* **25**(8), 8313–8333 (2024) <https://doi.org/10.1109/TITS.2024.3410896>
- [35] Feng, C., Han, P., Zhang, X., Yang, B., Liu, Y., Guo, L.: Computation offloading in mobile edge computing

- networks: A survey. *Journal of Network and Computer Applications* **202**, 103366 (2022) <https://doi.org/10.1016/j.jnca.2022.103366>
- [36] Lin, L., Liao, X., Jin, H., Li, P.: Computation offloading toward edge computing. *Proceedings of the IEEE* **107**(8), 1584–1607 (2019)
- [37] Yang, F., Gupta, N., Gerner, N., Qi, X., Demers, A., Gehrke, J., Shanmugasundaram, J.: A unified platform for data driven web applications with automatic client-server partitioning. In *Proceedings of the 16th International Conference on World Wide Web*, pp. 341–350 (2007)
- [38] Zhao, L., Wang, S., Ding, X.: Optimization method of task uninstallation in mobile edge computing environment combining improved deep q-learning and transmission learning. *Discover Applied Sciences* **7**(1), 9 (2024)
- [39] Kwon, Y.-W., Tilevich, E.: Energy-efficient and fault-tolerant distributed mobile execution. In *2012 IEEE 32nd International Conference on Distributed Computing Systems*, pp. 586–595 (2012). IEEE
- [40] Dong, S., Tang, J., Abbas, K., Hou, R., Kamruzzaman, J., Rutkowski, L., Buyya, R.: Task offloading strategies for mobile edge computing: A survey. *Computer Networks* **254**, 110791 (2024)
- [41] Cheng, N., Lyu, F., Quan, W., Zhou, C., He, H., Shi, W., Shen, X.: Space/aerial-assisted computing offloading for iot applications: A learning-based approach. *IEEE Journal on Selected Areas in Communications* **37**(5), 1117–1129 (2019)
- [42] He, P., Lei, H., Wu, D., Wang, R., Cui, Y., Zhu, Y., Ying, Z.: Non-terrestrial network technologies: Applications and future prospects. *IEEE Internet of Things Journal* **12**(6), 6275–6299 (2025) <https://doi.org/10.1109/JIOT.2024.3522912>
- [43] Zhou, F., Hu, R.Q., Li, Z., Wang, Y.: Mobile edge computing in unmanned aerial vehicle networks. *IEEE Wireless Communications* **27**(1), 140–146 (2020) <https://doi.org/10.1109/MWC.001.1800594>
- [44] Wei, J., Cao, S.: Application of edge intelligent computing in satellite internet of things. In *2019 IEEE International Conference on Smart Internet of Things (SmartIoT)*, pp. 85–91 (2019). <https://doi.org/10.1109/SmartIoT.2019.00022>
- [45] Zhang, Z., Zhang, W., Tseng, F.-H.: Satellite mobile edge computing: Improving qos of high-speed satellite-terrestrial networks using edge computing techniques. *IEEE Network* **33**(1), 70–76 (2019) <https://doi.org/10.1109/MNET.2018.1800172>
- [46] Edge computing: from standard to actual infrastructure deployment and software development. Intel White Paper (2021). <https://builders.intel.com/solutionslibrary/edge-computing-from-standard-to-actual-infrastructure-deployment-and-software-development>
- [47] Shen, Z., Jin, J., Tan, C., Tagami, A., Wang, S., Li, Q., Zheng, Q., Yuan, J.: A survey of next-generation computing technologies in space-air-ground integrated networks. *ACM Computing Surveys* **56**(1), 1–40 (2023)
- [48] Liu, J., Shi, Y., Fadlullah, Z.M., Kato, N.: Space-air-ground integrated network: A survey. *IEEE Communications Surveys & Tutorials* **20**(4), 2714–2741 (2018)
- [49] Kato, N., Fadlullah, Z.M., Tang, F., Mao, B., Tani, S., Okamura, A., Liu, J.: Optimizing space-air-ground integrated networks by artificial intelligence. *IEEE Wireless Communications* **26**(4), 140–147 (2019)
- [50] Chen, X., Zhang, H., Wu, C., Mao, S., Ji, Y., Bennis, M.: Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning. *IEEE Internet of Things Journal* **6**(3), 4005–4018 (2018)
- [51] Sardellitti, S., Scutari, G., Barbarossa, S.: Joint optimization of radio and computational resources for multicell mobile-edge computing. *IEEE Transactions on Signal and Information Processing over Networks* **1**(2), 89–103 (2015) <https://doi.org/10.1109/TSIPN.2015.2448520>
- [52] Yu, H., Wang, Q., Guo, S.: Energy-efficient task offloading and resource scheduling for mobile edge computing. In *2018 IEEE International Conference on Networking, Architecture and Storage (NAS)*, pp. 1–4 (2018). <https://doi.org/10.1109/NAS.2018.8515731>
- [53] Mei, C., Gao, C., Wang, H., Xing, Y., Ju, N., Hu, B.: Joint task offloading and resource allocation for space-air-ground collaborative network. *Drones* **7**(7) (2023) <https://doi.org/10.3390/drones7070482>
- [54] Bi, S., Huang, L., Zhang, Y.-J.A.: Joint optimization of service caching placement and computation offloading in mobile edge computing systems. *IEEE Transactions on Wireless Communications* **19**(7), 4947–4963 (2020)
- [55] Khan, P.W., Abbas, K., Shaiba, H., Muthanna, A., Abuarqoub, A., Khayyat, M.: Energy efficient computation offloading mechanism in multi-server mobile edge computing—an integer linear optimization approach. *Electronics* **9**(6) (2020) <https://doi.org/10.3390/electronics9061010>
- [56] Liu, L., Mao, W., Li, W., Duan, J., Liu, G., Guo, B.: Edge computing offloading strategy for space-air-ground integrated network based on game theory. *Computer Networks* **243**, 110331 (2024) <https://doi.org/10.1016/j.comnet.2024.110331>

[g/10.1016/j.comnet.2024.110331](https://doi.org/10.1016/j.comnet.2024.110331)

- [57] Gao, Y., Liu, J., Geng, S., Zhao, X., Chen, Z., Zhou, H.: Edge computing task offloading based on game theory for space-air-ground integrated network. In 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), pp. 627–631 (2024). <https://doi.org/10.1109/CISCE62493.2024.10653097>
- [58] Zhang, J., Zhang, J., Shen, F., Yan, F., Bu, Z.: Dogs: Dynamic task offloading in space-air-ground integrated networks with game-theoretic stochastic learning. *IEEE Internet of Things Journal* **12**(2), 1655–1672 (2025) <https://doi.org/10.1109/JIOT.2024.3457855>
- [59] Hwang, S., Lee, H., Park, J., Lee, I.: Decentralized computation offloading with cooperative uavs: Multi-agent deep reinforcement learning perspective. *IEEE Wireless Communications* **29**(4), 24–31 (2022)
- [60] Li, F., Qu, K., Liu, M., Li, N., Sun, T.: Collaborative computation offloading and resource management in space-air-ground integrated networking: A deep reinforcement learning approach. *Electronics* **13**(10), 1804 (2024)
- [61] Xie, W., Chen, C., Ju, Y., Shen, J., Pei, Q., Song, H.: Deep reinforcement learning-based computation offloading for space-air-ground integrated vehicle networks. *IEEE Transactions on Intelligent Transportation Systems* **26**(5), 5804–5815 (2025) <https://doi.org/10.1109/TITS.2025.3551636>
- [62] Tang, F., Hofner, H., Kato, N., Kaneko, K., Yamashita, Y., Hangai, M.: A deep reinforcement learning-based dynamic traffic offloading in space-air-ground integrated networks (sagin). *IEEE Journal on Selected Areas in Communications* **40**(1), 276–289 (2022) <https://doi.org/10.1109/JSAC.2021.3126073>
- [63] Huang, C., Chen, G., Xiao, P., Xiao, Y., Han, Z., Chambers, J.A.: Joint offloading and resource allocation for hybrid cloud and edge computing in sagins: A decision assisted hybrid action space deep reinforcement learning approach. *IEEE Journal on Selected Areas in Communications* **42**(5), 1029–1043 (2024) <https://doi.org/10.1109/JSAC.2024.3365899>
- [64] Zhu, W., Deng, X., Gui, J., Zhang, H., Min, G.: Cost-effective task offloading and resource scheduling for mobile edge computing in 6g space-air-ground integrated network. *IEEE Internet of Things Journal* **12**(12), 19428–19442 (2025) <https://doi.org/10.1109/JIOT.2025.3541082>
- [65] Nguyen, M.D., Le, L.B., Girard, A.: Integrated computation offloading, uav trajectory control, edge-cloud and radio resource allocation in sagin. *IEEE Transactions on Cloud Computing* **12**(1), 100–115 (2023)
- [66] Yang, Y., Li, B., Liu, Y., Dai, J., He, J., Wang, T., He, X.: Joint delay and throughput network resource orchestration in sagin based on sdn. In 2023 9th International Conference on Computer and Communications (ICCC), pp. 232–236 (2023). IEEE
- [67] Fan, K., Feng, B., Zhang, X., Zhang, Q.: Demand-driven task scheduling and resource allocation in space-air-ground integrated network: A deep reinforcement learning approach. *IEEE Transactions on Wireless Communications* **23**(10), 13053–13067 (2024) <https://doi.org/10.1109/TWC.2024.3398199>
- [68] Qin, P., Zhao, H., Fu, Y., Geng, S., Chen, Z., Zhou, H., Zhao, X.: Energy-efficient resource allocation for space-air-ground integrated industrial power internet of things network. *IEEE Transactions on Industrial Informatics* **20**(4), 5274–5284 (2023)
- [69] Chen, S., Yang, H., Xiao, L., Xu, C., Xie, X., Yang, W., Xiong, Z.: Intelligent energy-efficient and fair resource scheduling for uav-assisted space-air-ground integrated networks under jamming attacks. In 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring), pp. 01–05 (2024). IEEE
- [70] Wei, Q., Chen, Y., Jia, Z., Bai, W., Pei, T., Wu, Q.: Energy-efficient caching and user selection for resource-limited sagins in emergency communications. *IEEE Transactions on Communications* **73**(6), 4121–4136 (2025) <https://doi.org/10.1109/TCOMM.2024.3511958>
- [71] Zhou, Q., Yao, Y., Liu, Q., Huang, S., Yue, X.: Energy consumption minimization of multi-uav assisted mobile edge computing in sagin. In 2024 IEEE/CIC International Conference on Communications in China (ICCC Workshops), pp. 499–504 (2024). IEEE
- [72] Jiang, Y., Tang, X., Li, B., Zhang, R., Liu, J., Liu, N.: Energy-efficient uav edge computing for space-air-ground integrated networks. In 2025 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2025). IEEE
- [73] Eberechukwu N., P., Onyekwelu, M., Yoon, D.: Extended dl coverage in sagin: Cell association and resource allocation with beam hopping leo. *IEEE Internet of Things Journal* **12**(5), 6014–6028 (2025) <https://doi.org/10.1109/JIOT.2024.3490668>
- [74] Sharif, S., Manzoor, S., Ejaz, W.: Priority-based resource optimisation and user association in integrated networks. *IET Networks* **14**(1), 12140 (2025)
- [75] Jain, A., Lopez-Aguilera, E., Demirkol, I.: User association and resource allocation in 5g (aura-5g): A joint optimization framework. *Computer Networks* **192**, 108063 (2021)

- [76] Mao, W., Lu, Y., Pan, G., Ai, B.: Uav-assisted communications in sagin-isac: Mobile user tracking and robust beamforming. *IEEE Journal on Selected Areas in Communications* **43**(1), 186–200 (2025) <https://doi.org/10.1109/JSAC.2024.3460065>
- [77] Wang, Z., Zhang, L., Feng, D., Wu, G., Yang, L.: Intelligent cloud-edge collaborations for energy-efficient user association and power allocation in space-air-ground integrated networks. *IEEE Journal on Selected Areas in Communications* **42**(12), 3659–3673 (2024) <https://doi.org/10.1109/JSAC.2024.3459089>
- [78] Chen, Q., Meng, W., Han, S., Li, C., Chen, H.-H.: Reinforcement learning-based energy-efficient data access for airborne users in civil aircrafts-enabled sagin. *IEEE Transactions on Green Communications and Networking* **5**(2), 934–949 (2021)
- [79] Qiu, M. (ed.): Algorithms and architectures for parallel processing. Springer (2020) <https://doi.org/10.1007/978-3-030-60245-1>
- [80] Nguyen, M.D., Le, L.B., Girard, A.: Computation offloading, uav placement, and resource allocation in sagin. In 2022 IEEE Globecom Workshops (GC Wkshps), pp. 1413–1418 (2022). IEEE
- [81] Zhang, H., Zhao, Y.: Satbft: An efficient and scalable consensus protocol for blockchain-enabled space-air-ground integrated network. *IEEE Transactions on Cognitive Communications and Networking*, 1–1 (2025) <https://doi.org/10.1109/TCCN.2025.3547786>
- [82] Gao, Z., Zhang, J., Zhang, D., Wang, A., Pan, C., Li, X.: Blockchain based secure relay scheme for space-terrestrial integrated networks. *China Communications* **20**(5), 170–181 (2023)
- [83] Wang, X., Sun, T., Duan, X., Wang, D., Li, Y., Zhao, M., Tian, Z.: Holistic service-based architecture for space-air-ground integrated network for 5g-advanced and beyond. *China Communications* **19**(1), 14–28 (2022)
- [84] Li, Y., Feng, R., Gao, R., Wang, J.: Fountain coded streaming for sagin with learning-based pause-and-listen. *IEEE Networking Letters* **5**(1), 36–40 (2022)
- [85] Jiang, W., Zhan, Y., Xiao, X., Sha, G.: Network simulators for satellite-terrestrial integrated networks: A survey. *IEEE Access* **11**, 98269–98292 (2023)
- [86] Wang, H., Xia, X., Song, T., Xing, Y.: Survey on space-air-ground integrated networks in 6g. In 2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops), pp. 315–320 (2021). IEEE
- [87] Damnjanovic, A., Montojo, J., Wei, Y., Ji, T., Luo, T., Vajapeyam, M., Yoo, T., Song, O., Malladi, D.: A survey on 3gpp heterogeneous networks. *IEEE Wireless communications* **18**(3), 10–21 (2011)
- [88] Ke, H., Wang, J., Deng, L., Ge, Y., Wang, H.: Deep reinforcement learning-based adaptive computation offloading for mec in heterogeneous vehicular networks. *IEEE Transactions on Vehicular Technology* **69**(7), 7916–7929 (2020) <https://doi.org/10.1109/TVT.2020.2993849>
- [89] Qiu, B., Wang, Y., Xiao, H., Zhang, Z.: Deep reinforcement learning-based adaptive computation offloading and power allocation in vehicular edge computing networks. *IEEE Transactions on Intelligent Transportation Systems* **25**(10), 13339–13349 (2024) <https://doi.org/10.1109/TITS.2024.3391831>
- [90] Wang, J., Hu, J., Min, G., Zomaya, A.Y., Georgalas, N.: Fast adaptive task offloading in edge computing based on meta reinforcement learning. *IEEE Transactions on Parallel and Distributed Systems* **32**(1), 242–253 (2021) <https://doi.org/10.1109/TPDS.2020.3014896>
- [91] Liao, X., Hu, X., Liu, Z., Ma, S., Xu, L., Li, X., Wang, W., Ghannouchi, F.M.: Distributed intelligence: A verification for multi-agent drl-based multibeam satellite resource allocation. *IEEE Communications Letters* **24**(12), 2785–2789 (2020) <https://doi.org/10.1109/LCOMM.2020.3019437>
- [92] Zhang, W., Tan, L., Huang, T., Huang, X., Huang, M., Zhang, G.: Resource allocation and trajectory optimization in multi-uav collaborative vehicular networks: An extended multiagent drl approach. *IEEE Internet of Things Journal* **12**(8), 9391–9404 (2025) <https://doi.org/10.1109/JIOT.2024.3492953>
- [93] Seid, A.M., Boateng, G.O., Mareri, B., Sun, G., Jiang, W.: Multi-agent drl for task offloading and resource allocation in multi-uav enabled iot edge network. *IEEE Transactions on Network and Service Management* **18**(4), 4531–4547 (2021) <https://doi.org/10.1109/TNSM.2021.3096673>
- [94] Gyawali, S., Xu, S., Qian, Y., Hu, R.Q.: Challenges and solutions for cellular based v2x communications. *IEEE Communications Surveys Tutorials* **23**(1), 222–255 (2021) <https://doi.org/10.1109/COMST.2020.3029723>
- [95] Kovalenko, A., Hussain, R.F., Semiari, O., Salehi, M.A.: Robust resource allocation using edge computing for vehicle to infrastructure (v2i) networks. In 2019 IEEE 3rd International Conference on Fog and Edge Computing (ICFEC), pp. 1–6 (2019). <https://doi.org/10.1109/ICFEC.2019.8733151>
- [96] Nezami, Z., Chaniotakis, E., Pournaras, E.: When

- computing follows vehicles: Decentralized mobility-aware resource allocation for edge-to-cloud continuum. *IEEE Internet of Things Journal* **12**(13), 23324–23340 (2025) <https://doi.org/10.1109/JIOT.2025.3552504>
- [97] Cao, B., Zhang, J., Liu, X., Sun, Z., Cao, W., Nowak, R.M., Lv, Z.: Edge–cloud resource scheduling in space–air–ground-integrated networks for internet of vehicles. *IEEE Internet of Things Journal* **9**(8), 5765–5772 (2022) <https://doi.org/10.1109/JIOT.2021.3065583>
- [98] Tang, F., Wen, C., Zhao, M., Kato, N.: Machine learning for space–air–ground integrated network assisted vehicular network: A novel network architecture for vehicles. *IEEE Vehicular Technology Magazine* **17**(3), 34–44 (2022) <https://doi.org/10.1109/MVT.2022.3188405>
- [99] Matracia, M., Saeed, N., Kishk, M.A., Alouini, M.-S.: Post-disaster communications: Enabling technologies, architectures, and open challenges. *IEEE Open Journal of the Communications Society* **3**, 1177–1205 (2022) <https://doi.org/10.1109/OJCOMS.2022.3192040>
- [100] Do-Duy, T., Nguyen, L.D., Duong, T.Q., Khosravi-rad, S.R., Claussen, H.: Joint optimisation of real-time deployment and resource allocation for uav-aided disaster emergency communications. *IEEE Journal on Selected Areas in Communications* **39**(11), 3411–3424 (2021) <https://doi.org/10.1109/JSAC.2021.3088662>
- [101] Shang, S., Xu, D.: Beamforming design for double-ris assisted uav communication with limited feedback in disaster scenarios. In 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), pp. 1–5 (2023). <https://doi.org/10.1109/VTC2023-Spring57618.2023.10200825>
- [102] Okoumassoun, T.P., Ridhawi, I.A., Abbas, A., Al-Oqily, I.: Blockchain-enabled sagin communication for disaster prediction and management. In 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech), pp. 0678–0683 (2023). <https://doi.org/10.1109/DASC/PiCom/CBDCCom/Cy59711.2023.10361502>
- [103] Tan, J., Tang, F., Zhao, M., Kato, N.: Outage probability, performance, and fairness analysis of space–air–ground integrated network (sagin): Uav altitude and position angle. *IEEE Transactions on Wireless Communications* **24**(2), 940–954 (2025) <https://doi.org/10.1109/TWC.2024.3503060>
- [104] Haleem, A., Javaid, M., Singh, R.P., Suman, R.: Telemedicine for healthcare: Capabilities, features, barriers, and applications. *Sensors International* **2**, 100117 (2021) <https://doi.org/10.1016/j.sintl.2021.100117>
- [105] Lawal, L.S., Aibinu, A.M., Chris, C.R., Udoyen, U., Iheanacho, T., Jaafar, A., Ibrahim, I.A.: Overview of satellite communications and its applications in telemedicine for the underserved in nigeria: A case study. In 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1–10 (2022). <https://doi.org/10.1109/ICECCME55909.2022.9988286>
- [106] Panicacci, S., Giuffrida, G., Donati, M., Lubrano, A., Olivelli, M., Ruiiu, A., Fanucci, L.: Enhancing a telemedicine platform with global navigation satellite system technology and clustering algorithms for supporting epidemiological analysis. In 2021 IEEE International Conference on Electronic Technology, Communication and Information (ICETCI), pp. 410–414 (2021). <https://doi.org/10.1109/ICETCI53161.2021.9563621>
- [107] Pregowska, A., Masztalerz, K., Garlińska, M., Osial, M.: A worldwide journey through distance education—from the post office to virtual, augmented and mixed realities, and education during the covid-19 pandemic. *Education Sciences* **11**(3) (2021) <https://doi.org/10.3390/educsci11030118>
- [108] Guo, Y., Wang, S.: Challenges and opportunities in space service computing. In 2021 IEEE International Conference on Services Computing (SCC), pp. 44–51 (2021). <https://doi.org/10.1109/SCC53864.2021.00016>
- [109] Zhang, H., Jiang, M., Ma, L., Xiang, Z., Zhuo, S.: Computing offloading strategy for internet of medical things in space–air–ground integrated network. In 2023 IEEE International Conference on E-health Networking, Application Services (Healthcom), pp. 177–182 (2023). <https://doi.org/10.1109/Healthcom56612.2023.10472389>
- [110] Han, C., An, K., Lin, Z., Chatzinotas, S., Wang, J.: Endogenous anti-jamming communications for sagin: A network perspective. *IEEE Network*, 1–1 (2025) <https://doi.org/10.1109/MNET.2025.3551248>
- [111] Yang, K., Wang, Y., Gao, X., Shi, C., Huang, Y., Yuan, H., Shi, M.: Communications in space–air–ground integrated networks: An overview. *Space: Science & Technology* **5**, 0199 (2025) <https://doi.org/10.34133/space.0199> <https://spj.science.org/doi/pdf/10.34133/space.0199>
- [112] Tan, J., Tang, F., Zhao, M., Kato, N.: Performance analysis of space–air–ground integrated network (sagin): Uav altitude and position angle. In 2023 IEEE/CIC International Conference on Communications in China (ICCC), pp. 1–6 (2023). <https://doi.org/10.1109/ICCC57788.2023.10233368>

- [113] Bakambekova, A., Kouzayha, N., Al-Naffouri, T.: On the interplay of artificial intelligence and space-air-ground integrated networks: A survey. *IEEE Open Journal of the Communications Society* **5**, 4613–4673 (2024) <https://doi.org/10.1109/OJCOMS.2024.3429198>
- [114] Ali, J., Shan, G., Mohsin, A.R., Khalid, M., Chandroth, J., Roh, B.-h.: *Software-Defined Space-Air-Ground IoT: Motivation, Architecture and Research Challenges*, pp. 59–76. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-63705-6_4
- [115] Cai, Y., Yao, H., Gong, Y., Wang, F., Zhang, N., Guizani, M.: Privacy-driven security-aware task scheduling mechanism for space-air-ground integrated networks. *IEEE Transactions on Network Science and Engineering* **11**(5), 4704–4718 (2024) <https://doi.org/10.1109/TNSE.2024.3392389>
- [116] Li, S., Liu, G., Li, L., Zhang, Z., Fei, W., Xiang, H.: A review on air-ground coordination in mobile edge computing: Key technologies, applications and future directions. *Tsinghua Science and Technology* **30**(3), 1359–1386 (2024)
- [117] Cheng, L., Li, X., Feng, G., Peng, Y., Qin, S., Quek, T.Q.: Cooperative transmission for space-air-ground integrated networks: A multi-agent cooperation method. *IEEE Transactions on Vehicular Technology* **74**(8), 12879–12894 (2025) <https://doi.org/10.1109/TVT.2025.3553625>
- [118] Akhter, N., Mahmud, R., Jin, J., But, J., Ahmad, I., Xiang, Y.: Configurable harris hawks optimisation for application placement in space-air-ground integrated networks. *IEEE Transactions on Network and Service Management* **21**(2), 1724–1736 (2024)
- [119] Yang, N., Chen, S., Zhang, H., Berry, R.: Beyond the edge: An advanced exploration of reinforcement learning for mobile edge computing, its applications, and future research trajectories. *IEEE Communications Surveys & Tutorials* **27**(1), 546–594 (2025) <https://doi.org/10.1109/COMST.2024.3405075>
- [120] Yang, H., Guo, B., Xue, X., Deng, X., Zhao, Y., Cui, X., Pang, C., Ren, H., Huang, S.: Interruption tolerance strategy for leo constellation with optical inter-satellite link. *IEEE Transactions on Network and Service Management* **20**(4), 4815–4830 (2023)
- [121] Sabella, D., Vaillant, A., Kuure, P., Rauschenbach, U., Giust, F.: Mobile-edge computing architecture: The role of mec in the internet of things. *IEEE Consumer Electronics Magazine* **5**(4), 84–91 (2016)
- [122] Chen, Q., Giambene, G., Yang, L., Fan, C., Chen, X.: Analysis of inter-satellite link paths for leo mega-constellation networks. *IEEE Transactions on Vehicular Technology* **70**(3), 2743–2755 (2021)
- [123] Pan, C., Zhou, G., Zhi, K., Hong, S., Wu, T., Pan, Y., Ren, H., Di Renzo, M., Swindlehurst, A.L., Zhang, R., *et al.*: An overview of signal processing techniques for ris/irs-aided wireless systems. *IEEE Journal of Selected Topics in Signal Processing* **16**(5), 883–917 (2022)
- [124] Mehrabi, M., You, D., Latzko, V., Salah, H., Reisslein, M., Fitzek, F.H.: Device-enhanced mec: Multi-access edge computing (mec) aided by end device computation and caching: A survey. *IEEE Access* **7**, 166079–166108 (2019)
- [125] Madni, M.A.A., Iranmanesh, S., Raad, R.: Dtn and non-dtn routing protocols for inter-cubesat communications: A comprehensive survey. *Electronics* **9**(3), 482 (2020)