Space-Air-Ground Integrated Networks

Space-Air-Ground Integrated Networks:

Design and Performance Analysis

By

Dongqing Li, Shaohua Wu,
Weiqiang Wu and Qinyu Zhang

Cambridge
Scholars
Publishing

Space-Air-Ground Integrated Networks:
Design and Performance Analysis

By Dongqing Li, Shaohua Wu, Weiqiang Wu and Qinyu Zhang

This book first published 2025

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-0364-1725-3
ISBN (Ebook): 978-1-0364-1726-0

CONTENTS

PREFACE

The development trends of Space-Air-Ground-Sea Integrated Networks (SAGSINs) represent a convergence of advanced wireless communication technologies at the forefront of space and ocean sciences. These trends can be encapsulated by the emergence of high-capacity broadband satellites, the interconnectivity of satellites across various orbits, the diversification of application scenarios, and the heterogeneous integration of satellite and terrestrial networks. SAGSINs are characterized by their vast scales, high dynamics, and diverse resource compositions, which present considerable challenges for the real-time perception, processing, and application of information. Furthermore, the increasing prevalence of mission-critical applications, such as precise remote control, situational awareness, and real-time command decisions in battlefield contexts has resulted in a heightened demand for timely and effective dissemination of information. The efficient transmission of information is essential for successfully executing mission-critical services within SAGSINs. However, the intrinsic characteristics of SAGSINs constrain the applicability of conventional transmission technologies, thereby complicating efforts to ensure continuous and timely information delivery across communication links that are extensive in scale and subject to rapid fluctuations. Furthermore, providing precise support for mission-critical services with diverse requirements and attributes remains a significant challenge. As a result, there is an urgent demand for innovative research focused on time-efficient transmission technologies specifically designed for SAGSINs.

This monograph is dedicated to addressing the temporal requirements associated with information transmission and application within SAGSINs, aiming to enhance the applicability of transmission technologies considering the constraints imposed by these networks. By analyzing link characteristics, node capabilities, and specific network tasks and scenarios, this study conducts innovative research on transmission theory and technology relevant to SAGSINs. The primary technical objective is to improve the efficiency, adaptability, and practical utility of transmission technologies. To achieve this goal, the monograph aims to elucidate fundamental factors constraining information timeliness, develop high-efficiency transmission strategies suitable for large-scale aerospace links with a dynamic nature, and propose a coordinated task scheduling and resource management framework for marine Internet of Things (IoT) scenarios. In doing so, it ensures both the timeliness and efficiency of transmission technologies within SAGSINs.

This monograph presents a thorough review and detailed examination of transmission strategy design and rigorous performance analysis for age-critical applications within SAGSINs. It encompasses advanced Hybrid Automatic Repeat reQuest (HARQ) techniques, two-layer coding methodologies, hybrid satellite networks, and on-orbit processing strategies. Given the significant propagation delay between transmitters and receivers, traditional HARQ methods that rely on frequent feedback may not be advantageous for age-critical applications in SAGSINs. In response to this challenge, the monograph investigates the specific threshold of propagation delay within which retransmission proves beneficial for minimizing age. To improve the timeliness of transmission strategies within this threshold, we propose a fast Incremental Redundancy HARQ (IR-HARQ) protocol that enhances timeliness without compromising reliability. Additionally, we introduce a two-layer coding strategy to improve the timeliness of age-critical applications that fall outside the identified threshold. For time-sensitive marine IoT applications, we propose a joint task scheduling and resource allocation scheme based on a hybrid satellite network. The monograph concludes with several proposed future research directions in related fields.

We look forward to this monograph shedding light on the practical implementation of SAGSINs, offering significant insights that will be of great value. The systematic principles outlined within these pages provide not only a roadmap but also critical guidance for the establishment and optimization of future SAGSINs. We are deeply grateful to all the members of the Wush group for their invaluable contributions, including their enlightening discussions, insightful suggestions, and constructive feedback, which have greatly enriched the content of this work. Special recognition and heartfelt thanks are extended to the dedicated staff at Cambridge Scholars Publishing, with a particular mention of Alison Duffy, whose unwavering support and assistance throughout the publication preparation process have been instrumental in bringing this monograph to fruition.

Dongqing Li
Shenzhen, Guangdong, China

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

In the context of digital transformation, the demand for communication networks that are seamless, ubiquitous, and reliable has reached unprecedented levels. Although traditional terrestrial communication networks have demonstrated effectiveness across various applications, they are increasingly encountering limitations related to coverage, capacity, and resilience. This situation has catalyzed the development of Space-Air-Ground-Sea Integrated Networks (SAGSINs), which present a promising solution to these challenges and facilitate genuine global connectivity. SAGSINs integrate a diverse array of communication technologies and platforms across space, air, ground, and sea, thereby establishing a comprehensive, interconnected, and dynamic network infrastructure. Central to this paradigm shift is the concept of SAGSINs, which possesses the potential to transform the dissemination, processing, and utilization of information on a global scale. This innovative integration of various communication domains represents a significant advancement in the ongoing quest for seamless connectivity and universal access to information. This monograph explores the complexities of SAGSINs, emphasizing their primary drivers, architectural framework, distinctive characteristics, encountered challenges, and advancements in transmission technology, with a special emphasis on the timely delivery of information.

1.1 Space-Air-Ground-Sea Integrated Networks

SAGSINs represent a comprehensive and innovative communication architecture that effectively integrates satellite systems (space-based), aerial platforms (air-based), terrestrial networks (ground-based), and maritime communications (sea-based) into a cohesive framework. As illustrated in Fig. 1-1, this integrated approach aims to provide ubiquitous, intelligent, collaborative, and highly efficient information connectivity across extensive geographical regions, thereby facilitating a diverse array of applications and services. By leveraging the unique strengths of each component network, SAGSINs offer global coverage, resilience against disruptions, and optimized resource utilization. They are specifically designed to address the increasing demand for seamless connectivity in remote, challenging, or dynamic environments, such as disaster relief, military operations, maritime navigation, and everyday consumer services. The successful integration of these diverse networks requires advanced technologies, including network slicing, efficient transmission protocols, and intelligent management systems, to ensure seamless interoperability, optimal performance, and scalability. Ultimately, SAGSINs represent the future of communication infrastructure, promoting a truly connected world where information can flow freely across all domains.

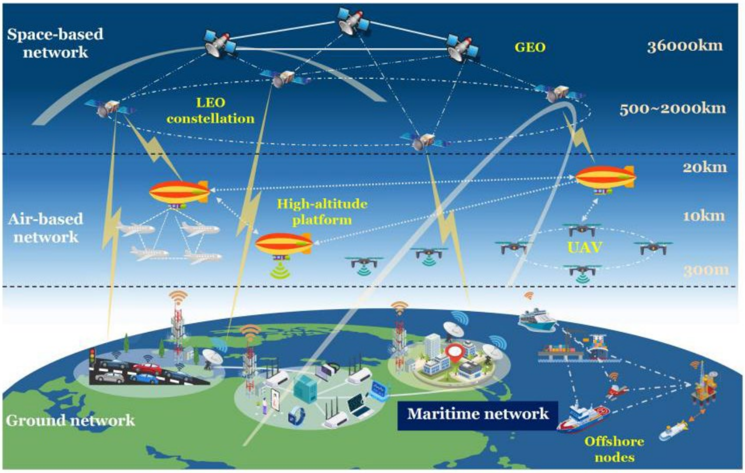

Fig. 1-1 Architecture of Space-Air-Ground-Sea integrated network

1.1.1 Key Drivers of SAGSINs

The fundamental drivers behind the development of SAGSINs are closely tied to the evolving demands of an interconnected global society that necessitates continuous communication and data exchange beyond traditional constraints. Essentially, SAGSINs strive to enhance worldwide connectivity with an unwavering dedication in order to facilitate seamless and reliable information transmission across diverse platforms such as space, air, land, and maritime environments.

This comprehensive framework is driven by the imperative for resilience, as contemporary societies increasingly depend on continuous communication infrastructures for essential services and emergency management. SAGSINs present a viable solution by incorporating multiple layers of redundancy and adaptability, ensuring that in the event of a disruption to one communication channel, alternative pathways can seamlessly assume responsibility, thereby preserving the continuity of information flow. The impetus for advancing SAGSINs arises from several interrelated factors.

Firstly, the surge in data-intensive applications, particularly within the realms of IoT, autonomous vehicles, and remote sensing technologies has generated an insatiable demand for bandwidth and coverage. Traditional communication networks often fall short of meeting these escalating demands due to their intrinsic limitations regarding reach, capacity, and resilience.

Secondly, sustaining continuous communication capabilities across varied environments is strategically significant. This is evident in military operations that necessitate secure real-time communication over extensive distances; disaster relief initiatives that rely on dependable connectivity in emergencies; and routine activities that benefit from ubiquitous access to information. SAGSINs provide distinct advantages in these scenarios.

Lastly, advancements in technology such as satellite constellations have rendered it feasible to integrate different domains into a singular network. These technological progressions along with innovations in software-defined networking (SDN), artificial intelligence (AI), and edge computing have facilitated the realization of SAGSINs.

In conclusion, the principal drivers of SAGSINs represent a synergistic amalgamation of the relentless pursuit for global connectivity, resilience against disruptions, efficient resource utilization, technological innovation, and support for the emerging requirements of a digitalized society. This integrated network paradigm signifies the future of communication by fostering a genuinely connected world where information can flow freely and seamlessly across all domains.

1.1.2 Architecture of SAGSINs

The architecture of SAGSINs is characterized by its multi-layered and heterogeneous nature, encompassing space-based networks (e.g., satellites), air-based networks (such as unmanned aerial vehicles, UAVs), terrestrial networks (including cellular and fiber-optic networks), and maritime networks. These layers are interconnected through advanced networking protocols and interfaces, enabling seamless data transmission and information sharing across different domains.

In the context of SAGSINs, the space segment is characterized by satellites equipped on different orbits, thereby facilitating extensive coverage and long-range communication capabilities. These satellites play an instrumental role in relaying data between geographically isolated locations, enhancing global connectivity and supporting various services including navigation, broadcasting, and remote sensing.

Complementing the space segment, the air segment incorporates aerial platforms such as drones, high-altitude platforms, and commercial aircraft. These aerial vehicles function as relay nodes to extend the communication network's reach while providing on-demand connectivity for critical operations or events. The ground segment constitutes the foundational infrastructure of SAGSINs comprising terrestrial networks that span land areas connecting to undersea cables. These networks employ a diverse array of technologies including fiber optics, microwave links, and cellular systems to deliver high-speed and low-latency communication for terrestrial users. Additionally, the sea segment encompasses underwater communication systems facilitating data exchange between submerged platforms such as submarines or underwater drones with surface vessels or shore-based stations. This is accomplished through specialized acoustic or optical technologies enabling uninterrupted connectivity within the underwater domain.

The segments collectively establish a robust and adaptive architecture that enables SAGSINs to effectively respond to evolving conditions and requirements. The architecture is designed for interoperability, allowing seamless integration among different networks and technologies, ensuring efficient and unrestricted data flow across all domains. This comprehensive approach not only enhances global connectivity but also promotes innovation and technological progress, driving the development of new applications and services that require high-bandwidth and low-latency communication networks.

1.1.3 Characteristics of SAGSINs

SAGSINs represent an advanced paradigm in communication technology, distinguished by their exceptional connectivity and adaptability. These networks facilitate the seamless integration of diverse communication platforms, encompassing satellites in orbit, UAVs in the atmosphere, terrestrial infrastructure on land, and underwater systems, thereby establishing a comprehensive communication ecosystem.

A notable attribute of SAGSINs is their extensive coverage. By capitalizing on the broad reach of satellites and the complementary capabilities of UAVs, ground stations, and undersea cables, these networks guarantee that communication services are accessible in virtually all locations on Earth, effectively bridging the digital divide between urban and remote regions. Additionally, SAGSINs are characterized by their dynamic adaptability. As network nodes, including satellites, UAVs, and mobile ground stations, are in constant motion and environmental conditions vary, SAGSINs utilize advanced adaptive routing and resource management algorithms. These algorithms enable the network to rapidly modify its topology and resource distribution, ensuring stable and reliable connections despite the challenges associated with high mobility and environmental variability.

Furthermore, SAGSINs demonstrate significant resilience and fault tolerance. The integration of multiple communication layers and the capacity to reroute data through alternative pathways in the event of disruptions ensure that critical services remain operational, even amidst natural disasters, technical malfunctions, or deliberate interference. In conclusion, SAGSINs signify the future of communication technology, providing unmatched coverage, dynamic adaptability, and robust resilience. They are positioned to transform the manner in which we connect, communicate, and access information, facilitating seamless and efficient communication across a variety of environments and scenarios.

1.2 Challenges of Timely Information Delivery in SAGSINs

The pursuit of seamless and ubiquitous connectivity through SAGSINs has generated numerous opportunities for enhancing communication and data exchange across various sectors. However, the prompt delivery of information within these intricate, multi-layered systems encounters a multitude of interconnected challenges that warrant thorough examination.

Primarily, the extensive range and diverse environments navigated by SAGSINs inherently present challenges associated with distance and signal propagation. The considerable distances separating satellites from ground stations, along with the complexities of underwater communication channels, result in significant delays in information transmission. These delays may be further intensified by fluctuations in signal strength and quality influenced by weather conditions, interference, and other environmental factors.

Secondly, the dynamic characteristics of the network's components, particularly within the aerial and spatial segments, represent a significant barrier to timely information delivery. The continuous movement of satellites in their orbits coupled with agile maneuvers of UAVs leads to a perpetually changing network topology. This situation necessitates advanced handoff and re-routing mechanisms to sustain connectivity which can introduce additional delays and complicate ensuring timely data transmission.

Moreover, the heterogeneity of communication technologies and protocols employed across various layers of SAGSINs adds another dimension of complexity. Achieving seamless interoperability among these disparate systems requires sophisticated gateways and translation mechanisms capable effectively bridging gaps between different communication protocols; however, these interfaces may introduce supplementary processing delays potentially jeopardizing timeliness of information delivery.

Another critical challenge pertains to the unpredictability and variability of environmental conditions, which can disrupt satellite signals and impair communication links due to weather phenomena such as storms and atmospheric disturbances. Similarly, underwater communication channels are particularly vulnerable to interference from ocean currents, temperature fluctuations, and other environmental factors, significantly affecting the reliability and promptness of information delivery.

Furthermore, the security of SAGSINs is an essential concern that must not be overlooked because these networks span extensive geographical areas involving multiple stakeholders. They are susceptible to a wide array of cyber threats including data breaches, denial-of-service attacks, and potential injection of false information into the network. Such attacks not only disrupt communication links but also compromise the integrity and authenticity of transmitted information, further obstructing timely delivery.

In conclusion, addressing these challenges will require innovative solutions and a comprehensive understanding of the complex interactions among various factors including the vastness and diversity of network environments, dynamic nature of components within it, heterogeneity in communication technologies used as well as unpredictable environmental conditions that need to be considered alongside implementing robust security measures.

1.3 Development of Transmission Technology in SAGSINs

Investigations into the timeliness of systems, examined through a network-level lens utilizing queuing theory, have historically operated under the presumption of optimal channel conditions. However, in an integrated Space-Air-Ground-Sea communication network framework, channels are inherently flawed and frequently exposed to challenging environmental conditions that can result in inaccurate updates at the receiving end. Therefore, adopting error control mechanisms is essential to increase the likelihood of successful packet reception and enhance state update timeliness. Two widely used error control strategies are Forward Error Correction (FEC) and Automatic Repeat reQuest (ARQ).

1.3.1 Forward Error-Correction Coding

FEC coding is a critical methodology employed to enhance the reliability of data transmission in SAGSINs. By incorporating redundant information into the transmitted data, FEC facilitates the receiver's ability to identify and rectify errors without necessitating retransmission. This capability is particularly vital in high-error-rate environments, such as those encountered in space-based communication systems. Recent research, as exemplified in **[1]**, has applied information-theoretic frameworks to ascertain optimal encoding strategies aimed at minimizing the Average Age of Information (AoI) in erasure channels, both with and without feedback. These studies have identified Maximum Distance Separable (MDS) codes as optimal for reducing AoI under specific conditions. Additionally, **[2]** explores the design of source coding techniques to minimize average AoI through the use of prefix-free codes, while **[3]** introduces a differential coding scheme that capitalizes on temporal correlations within source messages. This scheme alternates between transmitting actual state information and differential information, demonstrating enhanced timeliness when codeword lengths for both differential and actual messages are judiciously selected, even in the absence of feedback. Furthermore, **[4]** examines the efficacy of adaptive coding for packet-based traffic, utilizing timeliness metrics to derive cLoSed-form expressions for both average and peak AoI in a communication network modeled as an erasure channel with discrete time. The results indicate that the transmission of linear combinations of multiple original packets can substantially enhance AoI.

However, these investigations operate under the assumption of negligible propagation delays, wherein the receiver receives state updates instantaneously and the transmitter promptly acknowledges feedback. This means that in these idealized scenarios, there is no significant time lag between the moment a piece of information is sent and the moment it is received and acknowledged by the other end. In contrast, Space-Air-Ground-Sea communication scenarios, particularly those beyond near-Earth communications, typically experience propagation delays that are significantly greater than transmission delays. These propagation delays occur because the signals have to travel much longer distances, often involving various mediums such as space, air, and ground, each with its own characteristics that can slow down the transmission of data. Consequently, in the context of space-air communications, ensuring reliable transmission within a single propagation delay necessitates the elimination of feedback and retransmission mechanisms. This is because the time it would take for a signal to travel back and forth between the transmitter and receiver could be prohibitively long, making traditional methods of error correction and confirmation inefficient or even impractical. Therefore, alternative strategies must be employed to ensure that the data transmitted is received correctly the first time, without relying on the conventional back-and-forth exchange of information.

Traditional error control coding, which primarily addresses bit-level errors at the physical layer, may prove insufficient for the highly dynamic and time-varying nature of space communication channels. Furthermore, excessive redundancy at the physical layer can lead to processing delays and necessitate considerable computational and storage resources, which may be impractical for resource-constrained satellite nodes. As a result, recent research has proposed two-layer forward error correction (FEC) mechanisms that utilize cross-layer coding to improve the reliability of communication systems. Luby **[5]** emphasizes the importance of cross-layer design in fading channels, contending that reliance solely on physical layer coding for error correction is less effective than employing cross-layer coding. Vehkapera investigates the maximization of system throughput under a constraint on total codeword length by jointly optimizing the distribution of redundancy between layers in non-fading channels **[6]**. Xiao explores the adjustment of two-layer code rates to minimize delays in network coding systems **[7]**, while Berger examines cross-layer rate optimization in fading channels characterized by constant fading values per codeword **[8]**.

Space communication links are characterized by large-scale, highly dynamic, and time-varying properties, as well as asymmetric uplink and downlink connections. Courtade et al. analyze the optimal allocation of redundancy between layers for block fading channels with fixed fade diversity per codeword, as well as channels where fade diversity increases linearly with codeword length, aiming to minimize power consumption while adhering to reliability constraints **[9]**. ahmood et al. concentrate on the joint optimization of cross-layer encoding to enhance reliability under energy limitations **[10]**. Although much of the current research emphasizes traditional performance metrics such as throughput, reliability, and transmission power, there remains a notable gap in the literature regarding two-layer coding transmission mechanisms specifically designed to address timeliness in space communication contexts.

1.3.2 Hybrid Automatic Repeat Request

Although propagation delay typically surpasses processing delay in most aerospace communication contexts, it is noteworthy that in near-Earth communication scenarios, propagation delay can be comparable to or even less than transmission delay. In these situations, the implementation of ARQ technology may facilitate a more favorable balance between timeliness and reliability. Conventional ARQ methods increase the probability of successful transmission by repeatedly sending identical update packets, which, while enhancing success rates, may result in the reception of outdated information. This, in turn, compromises the freshness of the received data and adversely impacts system timeliness.

To mitigate this challenge, researchers have introduced the HARQ mechanism, which combines traditional ARQ technology with Forward Error Correction (FEC) techniques. The current body of literature identifies three principal HARQ mechanisms employed in communication systems: Fixed Redundancy HARQ (FR-HARQ) **[11]**, Chase Combining HARQ (CC-HARQ) **[12],** and Incremental Redundancy HARQ (IR-HARQ) **[13**]. FR-HARQ utilizes fixed-rate codes, wherein the bits of information within each packet are encoded into a packet containing symbols. Successful decoding is presumed when the receiver acquires a sufficient number of unerased symbols. Conversely, if the number of unerased symbols received is insufficient, the packet is retransmitted until successful decoding occurs at the receiver. In contrast, both CC-HARQ and IR-HARQ are categorized as incremental redundancy HARQ transmission schemes, differing primarily in the nature of the redundant information that is retransmitted. Specifically, CC-HARQ retransmits the identical information as originally sent, whereas IR-HARQ retransmits the same information but incorporates various forms of redundant bits. Compared to FR-HARQ, both CC-HARQ and IR-HARQ mechanisms enhance error correction capabilities by integrating additional redundant bit information, thereby improving overall system performance.

Recent studies have investigated the implementation of Automatic Repeat reQuest (ARQ) and HARQ mechanisms in scenarios involving cLoSe-proximity communication. For example, reference **[14]** evaluates the AoI performance of IIR and FR schemes for status updates transmitted over a BEC. The findings indicate that by adjusting redundancy in accordance with channel erasure rates, the performance of FR can be aligned with that of IIR in single-receiver systems, while FR demonstrates superior performance in multi-receiver contexts. The research presented in **[15]** examines the energy costs associated with LDPC-based status update systems, revealing that a reduction in AoI is achieved at the cost of increased energy consumption due to retransmissions. The authors of **[16]** focus on the transmission of state information over block-fading channels, employing Turbo codes and IR-HARQ to address channel errors. In **[17]**, a comparison is made between the timeliness of CC-HARQ and IR-HARQ strategies in Rayleigh fading channels, utilizing finite blocklength coding theory. Additionally, the authors in **[18]** propose an early HARQ strategy that anticipates decoding outcomes, thereby minimizing feedback and retransmission delays. Reference **[19]** introduces a network coding-based HARQ strategy for satellite communications, which achieves a lower average AoI in dual-hop transmissions characterized by limited or absent feedback.

Nonetheless, it is important to note that HARQ is frequently regarded as a high-latency approach due to its dependence on feedback and retransmissions to enhance reliability. In the varied communication environments encountered in aerospace, maritime, and satellite networks, HARQ strategies may not be universally applicable across all mission scenarios, as indicated in **[20]**. At present, there exists a notable gap in the literature concerning the design of HARQ strategies specifically tailored for aerospace, maritime, and satellite communication contexts.

1.3.3 On-orbit Processing

The extensive temporal and spatial dimensions of networks, combined with the high dimensionality of information, present considerable challenges for the integrated Space-Air-Ground-Sea network in terms of processing and transmitting substantial volumes of data. For instance, in remote sensing operations, a variety of terrestrial and maritime observation images are collected through spatial nodes. These include remote sensing images, monitoring images, reconnaissance images, aerial photographs, and radar images obtained from devices such as cameras, infrared spectrometers, imaging radars, and LiDAR systems. Such image data constitute the foundational source for analysis across numerous applications. However, the limited transmission capacity of space-air-sea wireless links often necessitates traditional methods that rely on "compression-storage-delay/opportunistic transmission," thereby exerting significant strain on communication links **[21][22]**. In many emerging applications that highlight national and urban competitiveness–such as natural disaster monitoring and assessment, urban transportation, fisheries monitoring, and emergency response–the value of information diminishes rapidly as delays occur in relaying data to ground-based decision-making centers. As a result, conventional transmission methods frequently fail to ensure the timely delivery of information. **[23]**.

Therefore, it is imperative to utilize the processing capabilities of nodes along the return path to facilitate on-orbit processing and the efficient transmission fusion of spatial information. This approach is essential for maintaining high levels of timeliness and accuracy in task responses within the integrated Space-Air-Ground-Sea network. The design phiLoSophy