Space-Air-Ground Integrated Network (SAGIN) for 6G: Requirements, Architecture and Challenges

Huanxi Cui[1], Jun Zhang[2], Yuhui Geng[1], Zhenyu Xiao[1,*],Tao Sun[3], Ning Zhang[4], Jiajia Liu[5], Qihui Wu [6], Xianbin Cao [1]

[1] School of Electronic and Information Engineering, Beijing Key Laboratory for Network-Based Cooperative Air application Management, Beihang University, Beijing 100191, China
[2] Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing 100081, China.
[3] China Mobile Research Institute, Beijing 100053, China
[4] Department of Electrical and Computer Engineering, University of Windsor, ON, Canada, N9B 3P4
[5] National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, the School of Cybersecurity, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, China
[6] College of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210007, China
[*] The corresponding author, email: xiaozy@buaa.edu.cn

Abstract: As the fifth-generation (5G) mobile communication network may not meet the requirements of emerging technologies and applications, including ubiquitous coverage, industrial internet of things (IIoT), ubiquitous artificial intelligence (AI), digital twins (DT), etc., this paper aims to explore a novel space-air-ground integrated network (SAGIN) architecture to support these new requirements for the sixth-generation (6G) mobile communication network in a flexible, low-latency and efficient manner. Specifically, we first review the evolution of the mobile communication network, followed by the application and technology requirements of 6G. Then the current 5G non-terrestrial network (NTN) architecture in supporting the new requirements is deeply analyzed. After that, we proposes a new flexible, low-latency and flat SAGIN architecture, and presents corresponding use cases. Finally, the future research directions are discussed.

Keywords: 6G; AI; DT; SAGIN; NTN; Network architecture.

Received: May 31, 2021
Revised: Aug. 29, 2021
Editor: Zhenyu Xiao

I. INTRODUCTION

The 5th generation mobile communication network (5G) was proposed to support three typical application scenarios including enhanced mobile broadband (eMBB), massive machine type of communication (mMTC) and ultra-reliability, and ultra-low latency (URLLC) [1]. With the commercialization of 5G, researchers have begun preliminary explorations on the new emerging application and technology requirements for the 6th generation mobile communication network (6G). It is expected that 6G should have ubiquitous coverage and ultra-wide-area broadband access capabilities anytime and anywhere to support the future Internet of Things (IoT), remote area coverage, emergency communications, ecological remote sensing, and so on [2–7]. Moreover, ubiquitous artificial intelligence (AI), digital twins (DT), etc. are considered to be the key technologies to improve the operation, maintenance and management for 6G. However, the current 5G network architecture is difficult to support those services and applications.

As an emerging network architecture, space-air-ground integrated network (SAGIN) holds great potential to meet the needs of the aforementioned services and applications [8, 9]. SAGIN is an integrated mobile communication network that integrates satellites, airborne platforms (various types of aerial vehicles) and terrestrial networks [10]. The terrestrial networks achieve normalized coverage of urban hot spots. Satellite and airborne realize ubiquitous coverage in remote areas, sea and airspace. SAGIN is not a simple interconnection among different communication networks, but deep integration in terms of system, technology and application. SAGIN includes a unified terminal and air interface protocol, and it builds a unified control plane and data plane through the integration of radio access network (RAN) and core network (CN). Intelligent deployment technology of control network element is used to build a robust control plane. Besides, intelligent mobility management technology is used to realize seamless handover. These space-air QoS provisioning, policy routing, resource allocation and other network operation management functions are realized by differentiated micro-service network elements with space-air characteristics.

In the literature, a flurry of research works have been done on SAGIN to harvest its benefits. In [11], a software-defined SAGIN architecture was proposed to support diverse vehicular services. The authors in [12] developed a cross-domain soft defined network (SDN) architecture for multi-layer space-ground integrated network (SGIN). The key technologies such as hybrid SGIN architecture, radio resource management, and transparent handover for emergency scenarios were investigated in detail in [13]. In [14], a 5G-satellite network architecture based on SDN/network function virtualization (NFV) was proposed. Based on this architecture and network coding, a multi-path concurrent transmission technology and analysis model were proposed to achieve optimal application splitting in multipath between space and ground segments. Moreover, 3GPP set up a special non-terrestrial network (NTN) group on SAGIN. In 3GPP Rel-16 [15], standardization work of 5G NTN on interfaces, network architecture, protocols and resource control was carried out. In the 3GPP Rel-17 [16], different 5G NTN architectures were defined, namely transparent bent-pipe and regenerative network architectures, respectively. Besides, three types of network management requirements were proposed, including slice management, satellite management, and network indicator monitoring for 5G NTN. The above works are summarized and compared in Table 1.

Although many advantages have been brought in, the above-mentioned network architectures still face the problems of high delay and low flexibility. The main reasons for that are as follows. First, the access flexibility of 5G NTN is restricted by the current physical infrastructures, and RAN cannot be fully virtualized and softwareized in the current SAGIN architecture. Therefore, RAN may not perform flexible and effective management according to user requirements. Second, in the existing SAGIN architecture, CN is completely separated from RAN and located on the ground, causing large round-trip time (RTT). Thus, the existing SAGIN architecture may not meet the requirments of delay-sensitive services. Finally, different SAGIN components use different protocols, which leads to highly-complex proptocol stack.

To address the aforementioned issues, we design a novel flexible, and flat SAGIN architecture to support 6G new services. First, CN and RAN are integrated into a single network, which consists of a service-based control plane (SBCP), a service-based user plane (SBUP), and radio units (RU). Second, the interfaces, including interface between user and control plane, and that between user and data plane are redefined to share a unified and simple protocol stack.

The main contributions are summarized as follows:

1. We review the state-of-art of mobile communication network architecture and SAGIN architecture. Moreover, we analyze the problems of the 5G NTN architecture in terms of latency and flexibility.
2. We propose a lite and holistic service-based SAGIN architecture for 6G. First, RAN and CN adopt a holistic service-based concept to split network element functions into more fine-grained service-based network elements, which may be freely integrated and evolved independently. Second, CN and RAN are integrated into a single network to simplify the signaling process and interfaces, which consists of a SBCP, a SBUP, and RU. Third, the interfaces, including interface between user and control plane, and that between user and data plane, are redefined to share a unified and simple protocol stack.
3. The deployment cases of the three application scenarios of the new architecture are given.
4. We discuss the technical challenges faced in this

Table 1. *Comparison of existing works.*

Architecture	Data plane composition	Control mode	Network composition
HetNet [13]	Satellite, Router, Switch, Gateway	Distributed	Satellite-ground network
FRBSN [14]	GEO, LEO	Centralized	Satellite-ground network
SSAGV [11]	Satellite, Airborne network, Ground network	Centralized	Satellite-air-ground network
MLSTIN [12]	MEO, LEO, LAP, Ground network	Distributed	Satellite-air-ground network
NTN [16]	GEO, MEO, LEO, Ground network	Centralized	Satellite-ground network

new architecture and possible future research directions.

The rest of this paper is organized as follows. Section II reviews the evolution of mobile communication network architecture. Section III discusses the new emerging requirements of 6G. Section IV analyzes and summarizes the problems of 5G NTN. Section V presents a novel SAGIN architecture for 6G. Finally, Section VI summarizes the future research directions. The abbreviations used in the paper are summarized in Table 2.

II. EVOLUTION OF MOBILE COMMUNICATION NETWORK ARCHITECTURE

The 2nd-generation (2G) mobile networks, as a digital communication system, has a wider range of services compared to the 1st generation (1G) analog communication system [17]. As shown in Figure. 1, 2G adopts a three-level architecture namely base transceiver station (BTS), base station controller (BSC) and CN. BTS and BSC both belong to RAN, and CN is the core network of 2G. Notably, after the appearance of general packet radio service (GPRS), the mobile communication network can support circuit switching (CS) and packet switching (PS) simultaneously [18]. To reduce expenditure, 3G employs a three-level architecture similar to that of 2G [19] with only minor changes. As shown in Figure. 1, 3G applies NodeB and radio network controller (RNC) to replace BTS and BSC, respectively. 3G basically shares the same CN as that of GPRS.

In the 4G era, as depicted in Figure.1, a two-level all-internet protocol (IP) network architecture was proposed [20], including eNodeB (corresponding to RAN) and evolved packet core (EPC, corresponding to CN). Partial functions of RNC and NodeB are reorganized to form the eNodeB. The main function entities of EPC include mobility management entity (MME), home subscriber server (HSS), serving gateway (S-GW), and packet gateway (P-GW). As user plane (UP) is separated from control plane (CP), end-to-end delay of 4G system is largely reduced compared to 3G.

5G also applies a two-level architecture, including next generation-RAN (NG-RAN) and CN (termed as 5GC). In NG-RAN, a base station is separated into central unit (CU), distributed unit (DU) and active antenna unit (AAU) to support on-demand placement. There are three remarkable changes in 5GC architecture, namely service based architecture (SBA) [21], forwarding and control separation, and network slicing. SBA separates network functions (NFs) of 4G's CN into finer grained units or network function services (NFSs) [22], and redefines them as access and mobility function (AMF), session management function (SMF) and user plane function (UPF), respectively. Network slicing is able to satisfy various QoS requirements by creating logically isolated end-to-end subnetworks over a common hardware platform [23].

III. THE NEW REQUIREMENTS OF 6G

Generally, the next mobile communication era is driven by two main factors, namely application and technology [24]. In the following, we will present the needs for a new network architecture, from the perspectives of applications and technologies.

3.1 Application Requirements

There are some typical new applications in the future 6G, such as ubiquitous coverage, IIoT, Holographic communication, etc., which are distilled in some detail as follows.

3.1.1 Ubiquitous coverage

Ubiquitous coverage means the wireless access coverage is available anytime and anywhere, whose characteristics can be described as all-terrain and all-space 3D coverage. In the future, the requirement of IIoT

Table 2. *List Of abbreviation.*

6G	6th generation mobile communication technology	AI	Artifical intelligence
eMBB	enhanced mobile broadband	mMTC	massive machine type communication
URLLC	Ultra-reliable and low latency communication	VR	Virtual reality
SDN	Soft define network	NFV	Network function virtualization
AR	Augmented reality	5G	5th generation mobile communication technology
SAGIN	Space-air-ground integrated network	NTN	Non-terrestrial networks
SGIN	Space-ground integrated network	IoT	Internet of things
IIoT	Industrial internet of things	QoS	Quality of service
DRL	Deep reinforcement learning	MILP	Mixed integer programming
AMF	Access and mobility management function	SMF	Session management function
AUSF	Authentication server function	UDM	Unified data management
UPF	User plane function	PCF	Policy control function
NSSF	Network slice selection function	UDR	Unified data repository
RAN	Radio access network	DU	Distribute unit
CU	Centralized unit	AAU	Active antenna unit
SBCP	Service based control plane	SBUP	Service based user plane
ISL	Inter-satellite-link	SRI	Satellite radio interface
DN	Data center	CN	Core network
FFT	Fast fourier transform	IFFT	Inverse fast fourier transform
PDUSF	Packet data unit and session function	SDU	Session date unit
RRBMCF	Radio resource and bear management control function	NB-IoT	Narrow-band IoT
BTS	Base transceiver station	BSC	Base station controller
MSC	Mobile service switching	VLR	Visit location register
CS	Circuit switching	PS	Packet switching
GPRS	General packet radio service	UAV	Unmanned aerial vehicle
PS	Packet switching	RNC	Radio network controller
eMTC	enhanced machine-type communications	V2X	vehicle to everything
GMSC	Gateway mobile service switching	HLR	Home location register
mLLMT	massive low latency machine type	IoE	Internet of everything
LAN	Local Area Network	M2M	machine to machine
PCU	Packet control unit	SGSN	Service GPRS supported node
GGSN	Gateway GPRS supported node	EPC	Evolved packet core
MME	mobility management entity	HSS	Home subscriber server
S-GW	Serving Gateway	P-GW	Packet gateway
MEC	multi-access edge computing	VNFs	Virtualized network functions
NFS	Network function service	SBA	Service based architecture
MA	Microservice architecture	NF	Network functions

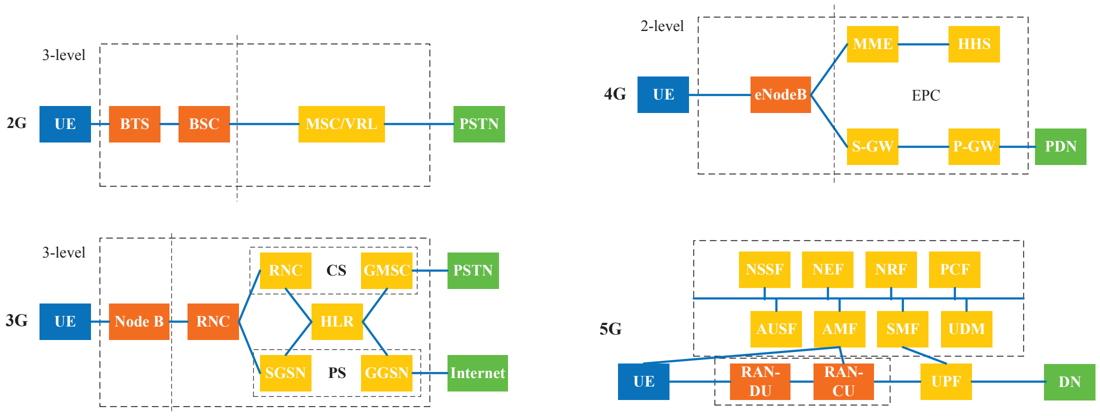


Figure 1. *The evolution of mobile communication network architecture.*

is more complicated. Besides, human activities will get much more involved in polar, desert, ocean, and so on. Therefore, the requirement of IoT equipment will be significantly expanded, including deep space, deep earth, and deep sea [25]. However, the current 5G mobile communication has limited coverage, which re-

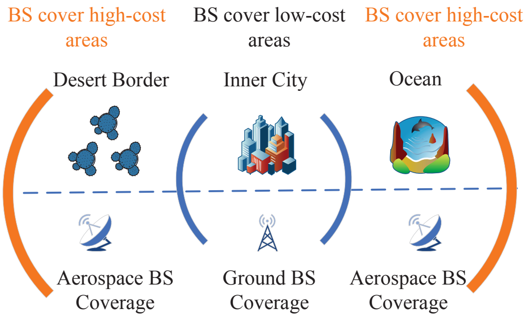


Figure 2. *Ubiquitous coverage.*

quires a new SAGIN architecture[25].

In the following, several typical service scenarios of ubiquitous coverage are listed:

1 ***Emergency Communications***: Emergency telecommunication means utilizing all kinds of communication resources to guarantee rescuing, emergency relief, and necessary communication when encountering natural or man-made emergencies. If emergencies occur in remote regions, or extreme environments, the current 5G network cannot provide emergency telecommunication service. Therefore, in the future, the next-generation communication network needs to provide wide and stable coverage.

2 ***IoT in wide area***: IoT in wide area means a kind of Internet of everything with wide coverage, low cost, simple deployment, and extensive connections. At present, there are some limitations for 5G IoT, such as the shortage of licensed spectrum, the interference of unlicensed spectrum, and difficulty of realizing large-scale IoT. At the same time, the current 5G technologies such as UAV communication technology, enhanced machine-type communications (eMTC), and narrow-band IoT (NB-IoT) are far from meeting the requirement of ubiquitous connectivity of everything [26]. Therefore, symbiotic radio technologies and new SAGIN architectures are needed to provide better support for IoT in wide area [26].

3 ***Ecological remote sensing monitoring***: Ecological remote sensing monitoring means utilizing various kinds of remote sensing technologies to monitor the dynamic changes of the ecosystem. Currently, the monitoring ecosystems are almost sparsely populated natural areas. The 5G network has difficulty covering and monitoring. Therefore, the SAGIN architecture pulls in the satellite network, in which several remote sensing satellites can realize effective monitoring.

3.1.2 IIoT

IIoT, which contains machine to machine (M2M), industry communication technologies, and automation application field, is a subset and an evolution of IoT. Compared with IoT, the IIoT aims to integrate and connect the independent factories or machines, thus providing more effective production and new services [27].

However, IIoT may face problems that 5G network can not solve, such as energy efficiency, real-time performance, and so on [28]. Because the mobile communication system is people-centered, the latency depends on the human reaction time. Although $1ms$ delay can be achieved in 5G, it is long for IIoT. With the popularization of industrial automation and the appearance of unmanned applications, the latency does not depend on the human reaction time anymore. The target latency of 6G is below $1ms$ and even undetectable [29].

In 6G vision, massive low latency machine type (mLLMT) is one of the core service scenarios. Applying mLLMT to large-scale IIoT needs low-latency communication and interaction among numerous personnels, sensors, and machines [30]. In the commonly accepted model of IIoT, edge, platform, and enterprise layer are connected according to the proximity, access, and service networks. The edge consists of sensors, controllers, and actuators, which are connected to edge gateways by separating LAN. The edge gateways connect to the platform layers to realize the global coverage [27]. However, it's challenging to achieve the large-scale secure, efficient IIOT in the wide-area only depending on the ground mobile communication network. Therefore, a new SAGIN architecture is needed to support the large-scale IIoT all over the world [30].

As one of the typical application examples of IIoT, motion control is highly challenging in industrial automation. As for the refined motions such as remote surgery and real-time remote control, they require a communication system with ultra-high reliability, low latency, and high certainty. The 5G network is hard to

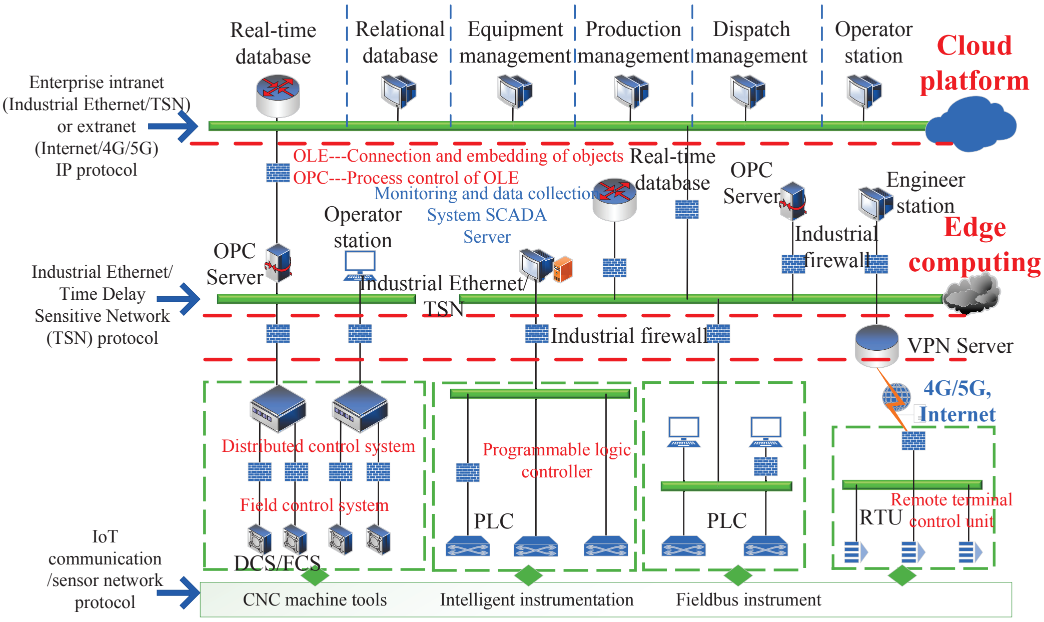


Figure 3. *Schematic diagram of IIoT.*

support those applications well. Thus, they are considered the initial case for 6G [31].

3.1.3 Other application requirements

Besides the above application requirements, there are other important applications as follows.

1 ***Holographic communication***: Holographic communication is a new presentation form, where virtual and reality are deeply integrated. Empowering the true three-dimensional display capabilities of the virtual world will enable people to enjoy a fully immersive holographic interactive experience without being restricted by time and space. 6G has shaped new life forms in many fields such as holographic intelligent communication, efficient learning and education, medical health, intelligent display, free entertainment, and industrial intelligence. By realizing the holographic intelligent connection of everything, communication system will face higher technical challenges to meet the transmission requirements of holographic communication.

2 ***Detection and perception***: Detection and perception include multi-dimensional and in-depth perception services such as precise positioning, 4D imaging, and material feature recognition. Detection and perception will be an important service capability provided by the 6G and the foundation of intelligent network applications. In addition, network performance can be improved through precise shaping, interference coordination, flow control, and network resource optimization. At the same time, with the help of detection and perception, 6G will provide people with more accurate and personalized data information in a series of applications such as smart factories, smart transportation, and smart medical care. Communication perception puts forward new requirements for 6G performance, including perception accuracy, perception capacity, perception delay, and perception range.

3 ***Medical health***: Based on the 6G network, health care will be further developed in the direction of intelligence, personalization and ubiquity. Currently, artificial intelligence, big data, sensor and haptic network technologies are gradually maturing. A wide range of new medical and health

forms will be formed based on AI-aided intelligent human body parameter acquisition and disease prediction, personalized health supervision and targeted therapy based on digital twins, multidimensional in-depth research on medical samples based on holography, remote diagnosis and surgery. It has strengthened the social medical service system and escorted the healthy life of the people. In order to realize cloud storage and big data intelligent analysis in medical and health, the network is required to have ultra-high cloud storage and computing capabilities as well as the efficient protection of data provacy and security.

3.2 Technology Requirements

Technology is the other key driving force to push the next generation mobile communication network. In the past several decades, the technologies represented by the Internet, AI, cloud computing, and so o, have developed rapidly. In the future 6G, under the technology-driven such as ubiquitous intelligence and digital twins [32], there are some new requirements.

3.2.1 Ubiquitous AI

The ubiquitous intelligence has two meanings, namely the ubiquity of technologies as well as capabilities, and the ubiquity of application of AI [33]. In the future, AI technology may become a universal technology as common as water and electricity [33]. What's more, not only the basic industries such as medical treatment, agriculture, and urban management are associated with AI, but also the new services such as emergency telecommunications, IIoT, etc., are gradually inseparable from AI [33]. Therefore, to provide ubiquitous intelligence globally, a new SAGIN architecture is needed to provide computing capabilities and coverage anytime and anywhere.

Traditional plug-in AI and big data solutions lead to complex network deployment and increase the cost of network operation and maintenance, while the capabilities of AI and big data may not bring essential improvements to the network. Moreover, the application of traditional plug-in AI and big data solutions has also caused the security of data leakage within the network, which in turn further restricts the drive of AI and big data to improve network capabilities. The native AI is a revolutionary technology which aims to enhance transmission rate with powerful performance, restructure management mode with intelligent algorithms, and activate the underlying value with comprehensive innovation. The architecture of native AI is expected to consist of three parts, namely the AI case self generator, the QoS AI interpreter and estimator, as well as the AI super brain and cerebellum. The AI case self generator can support the automatic generation of required intelligence. The QoS AI interpreter and estimator can support the translation, decomposition, and evaluation of native AI QoS. AI super brain and cerebellum aim to realize the distributed and centralized native AI architecture. Therefore, a new 6G architecture is needed to carry the entire life cycle to promote the widespread application of native AI.

3.2.2 Digital twins

Digital twins means that the whole world generates a digital virtual world based on physical world. All the information and communication in physical world can be transmitted in digital virtual world [32]. Several applications of the 6G, including internet of everything (IoE), IIoT, etc., have different performance requirements. Therefore, a new architecture is needed to manage, operate, and optimize the 6G communication system and underlying IoE service. Digital twins can be a potential tecnology to the new architecture. The digital twins utilizes the virtual representation of 6G physical system. Besides, the relevant algorithms such as machine learning, communication technologies, computing system such as MEC, and privacy and security technologies such as blockchain are also used by digital twins to support 6G network architecture [34]. Additionally, as a technology driving, the mobile communication application scenario will emerge new features, including SAGIN 3D coverage [32].

3.2.3 Other technology requirements

In addition to the above technical requirements, there are some other technical requirements for the future SAGIN as follows.

1 ***Programmable Network Technology***: The on-demand deployment, dynamic orchestration, and on-demand resource scheduling of 6G network functions or services are realized through programmable network technology to meet the requirements of dynamic changes in 6G application

and flexible network expansion.

2 ***Native Security***: Native security of 6G network should have the following characteristics. First, active immune function of network infrastructure, software, etc. is realized through trusted technology. Second, dynamic orchestration of security capabilities, flexible deployment, and improvement of network resilience are realized according to users and industry applications. Third, digital twin technology realizes the unification of physical network and virtual twin network security. Finally, intelligent collaboration of end, edge, network, and cloud accurately perceives the security situation of the network to deal with security risks agilely.

3 ***Computing Power Network Technology***: Computing power network technology is the key technology of computing network integration. Computing power network technology mainly includes computing power measurement and modeling, computing power resource perception, computing power routing, and in-network computing (INC). This technology can meet the requirements of new services in the future network, lite and dynamic computing.

IV. 5G NTN SUPPORTING THE NEW REQUIREMENTS

This section will introduce the NTN architecture and corresponding issues to support 6G network.

4.1 Architecture of NTN

Network architecture of NTN based on 5G NR is mainly divided into two categories: transparent bent-pipe architecture and regenerative architecture. This section details the main components and communication processes of these two types of NTN architectures.

4.1.1 Bent-pipe architecture of NTN

In the bent-pipe architecture based on the 5G NR, as shown in Figure. 4, the satellite only serves as a node of signal amplification and forwarding. The data from the 5G core network is first sent to the gNB, then transmitted to the satellite through the satellite gateway, and finally forwarded by the satellite to the users on the ground. In order to support delay-sensitive application, multi-connection is used to increase the transmission bandwidth of users to increase throughput and reduce the handover delay. However, this architecture increases the signaling overhead of the system, especially under the massive-connectivety scenario.

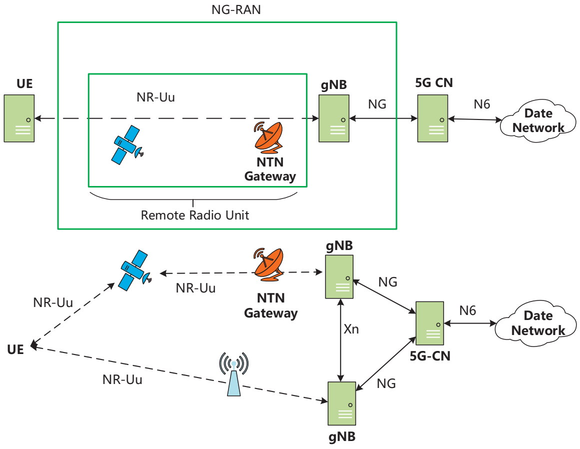


Figure 4. *The bent-pipe architecture.*

4.1.2 Regenerative architecture of NTN

The architecture of regenerative NTN mainly consists of two categories. As shown Figure. 5, the first type is a gNB that is completely equipped on the satellite in 5G NR. The second type is the DU equipped in the satellite, while CU is located on the ground.

In the first type, the complete gNB is mounted on the satellite, where the message on the NG interface at the gNB and 5G CN is transparently transmitted on the satellite interface SRI. The inter-satellite-link (ISL) could be developed among the satellites. In this architecture, the user can connect to 5G CN located on the ground via ISL. In the second type, the DU is mounted on the satellite, where the message on the interface F1 of the gNB-CU and 5G CN is transparently transmitted on the satellite interface SRI. If a satellite payload is equipped with more than one DU, all the corresponding F1 interface instances can be transported by the same SRI.

4.2 Disadvantages of NTN Architecture

The above two NTN architectures can guarantee the requirements of some services, while there are still certain problems with respect to delay and flexibility.

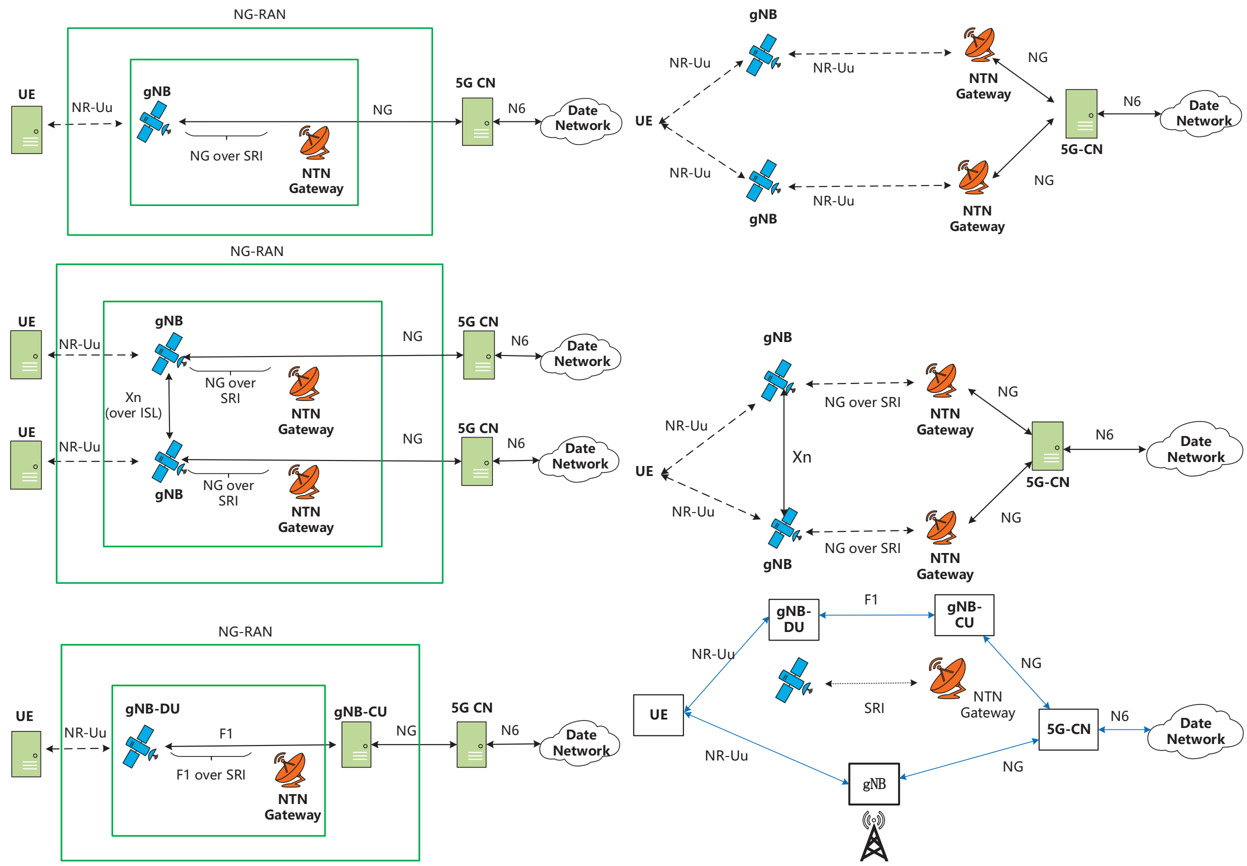


Figure 5. *The regenerative NTN architecture (The left-side of picture is single connection and the right-side is multi-connection).*

4.2.1 High Latency

The control plane, the forwarding plane, and the data center of the bent-pipe architecture are all located on the ground. As a result, the transmission delay of user data packets mainly comes from two aspects. One is the propagation delay caused by the distance of the space-ground link, and the other is the round-trip delay caused by the control plane located on the ground. Here, we give a simple example to illustrate the above-mentioned problems. When the satellite is moving at a high speed, it will cause frequent handover to users. However, some the gNB is located on the ground, the handover requests of the user need to be transmitted to the ground station via satellite, and then forwarded to the gNB on the ground. Next, the CU on the ground starts the handover process. As the commands such as user admission control and connection release during the handover process need to be transmitted between the user and the ground gNB, the processing delay of the handover is very high.

Similar problems also exist in the regenerative architecture. Regardless of whether the satellite is equipped with a complete gNB or gNB-DU, when users need to perform inter-satellite handover in the satellite edge area, the handover between different base stations requires AMF located on 5G CN according to the handover process between 5G NR base stations. Therefore, the handover command of user also needs to be transmitted to the AMF of 5G CN on the ground via the satellite gNB or gNB-DU to complete the handover process. It can be observed that the processing delay of the control plane cannot be reduced, whether the satellite is equipped with gNB or gNB-DU. Detailed analysis of mobility management of NTN architecture based on 5G NR can refer to [35].

In summary, all functions of the core network are located on the ground. All user traffic data and control messages need to be transmitted back to the ground for forwarding and processing. Therefore, the data forwarding delay and the control message transmission delay are espected to be extreme high.

4.2.2 Poor Flexibility

In the transparent bent-pipe architecture, network slicing is used to meet different application requirements, while the slicing operation ends at the gNB on the ground and the forwarding node on the satellite is a black box. This kind of architecture cannot be flexibly managed and operated. Therefore, physical resources such as power, spectrum, storage, and calculation on the satellite cannot be flexibly allocated. Besides, end-to-end slicing cannot be performed when providing differentiated services.

In the regenerative architecture, although the satellite is equipped with a complete gNB or part function of gNB, the AAU function in the gNB is still a black box. When the network slicing technology is used to meet different application requirements, the functions in the AAU cannot be sliced, as such the baseband processing functions cannot be flexibly allocated. For example, when performing ecological monitoring, strong baseband processing functions, bandwidth, power and other resources are not required. Therefore, the traditional slicing method will cause a great waste of resources.

V. PROPOSED SAGIN ARCHITECTURE

To solve the aforementioned problems of inflexibility and high latency in the 5G NTN architecture, we propose a new SAGIN architecture for 6G in this section.

5.1 Lite and Holistic Service-oriented SAGIN Architecture for 6G

Due to the reason that delay requirement, some core functions of CN of the new SAGIN architecture should be located on LEO to reduce latency. The main reason of high latency in 5G NTN is that CN is deployed on the ground, which leads to large user-satellite-ground CN round-trip delay. The necessary core functions of CN can be moved from ground to satellite, such that the round-trip delay can be reduced to user-satellite CN. To this end, the current centralized CN needs to be updated to distributed autonomous architecture, i.e., some of the elements of CN can be deployed closer to RAN rather than located together in the data center, thereby shaping a local autonomy cellular system. However, unlike data centers, satellite platforms have carry limitations and high-speed mobility. For some core functions, such as UPF, are complex and resource-consuming, they are required to achieve lightened design to be mounted on the satellite platforms. Meanwhile, some functions of CN are only used for terrestrial networks, while the main algorithms of CN need to be adapted to the characteristics of satellites before they can be transplanted to satellite platforms. Therefore, in the future architecture, we need to design a lightened and customizable CN to meet the requirements of SAGIN.

Second, the existing RAN need to be holistic service-oriented. The core components of 5G NTN RAN, such as DU and CU, adopt dedicated hardware platforms, which leads to low flexibility. For SAGIN, the types and the capabilities of satellites are different. Deploying RAN on satellites using existing proprietary equipment will make the deployment and update maintenance costs unaffordable. Therefore, we apply the design concept of 5G SBA to make the functions of RAN holistic service-oriented, which means that the functions of DU and CU are split the in fine-grained manner. In this design, these functions can be evolved independently and the interfaces between network elements are simplified. In addition, the holistic service-oriented RAN also promotes the application of NFV on satellite platforms.

Considering the above two principles, this paper proposes a novel lite and holistic service-based SAGIN architecture, composing of SBCP, SBUP, and RU, as shown in Figure. 6. SBCP is formed by the integration of the control plane of the CN and that of the RAN, which is mainly used for mobility management, session management, etc. SBUP is formed by the integration of the data plane of the CN and the data plane of the RAN, which is mainly used for the routing of user data, QoS flow management and other functions. RU is composed of the partial functions of the AAU, and the rest of AAU is combined with the DU. The proposed SAGIN architecture is a deeply integrated network of CN and RAN, which no longer distinguishes CN and RAN.

Compared with the 5G NTN, the SAGIN architecture proposed has the following main advantages.

1 *Lower latency*. In 5G NTN architecture, all functions of the CN are located on the ground, where users' data and control signaling need to be transmitted to the ground for forwarding and process-

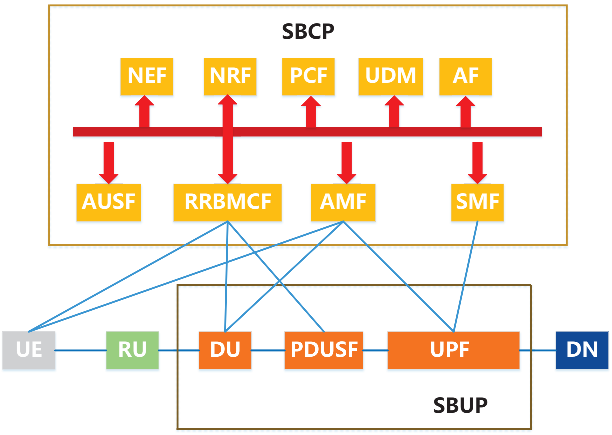


Figure 6. *The architecture of proposed SAGIN.*

ing. As such, the data forwarding delay and the control signaling transmission delay become high. In contrast, the proposed SAGIN architecture is a lite and holistic-service architecture. Thus, the functions of CN can be flexibily and on-demand deployed on the satellite to reduce the transmission delay. In addition, the deep integration of the CN and the RAN simplifies the signaling processing flow, thereby further reducing the network forwarding and processing delay.

2 *More flexible and customizable*. In 5G NTN architecture, all the functions of RAN and some functions of CN are customized by hardware. Therefore, network element functions of RAN cannot be flexibly designed for requirements and characteristics of platforms and application scenario. In addition, the costs of deployment and update maintenance costs are high. In contrast, the proposed SAGIN architecture is service-oriented, customizable and lighten. RAN no longer uses proprietary devices, but instead uses service-based technology to design. Therefore, each function of RAN can be split into more fine-grained network elements. Specific functions for network elements can be customized according to different application scenarios and characteristics, and even perform independent evolution. Moreover, these network elements can be deployed in a distributed manner on satellite networks, airborne networks, and ground networks to form a distributed autonomous system.

5.2 The Integration of CN and RAN

The proposed solution includes virtualization of the AAU based on NFV technology, only leaving AD/DA/RF functions, and merging the functions with DU to form a new network element. The further virtualization of AAU is to integrate with the CN, so that the network can be sliced more flexibly to meet various application requirements. The CU is separated into CU-CP and CU-UP. The CU-CP and AMF are integrated, and CU-UP and UPF are integrated. The CN and RAN are integrated together. CU-CP and the control plane of 5G CN from a new network control plane, and CU-UP, UPF and DU form a new network data plane. This paper redefines the interface between the user and the control plane. Besides, the interface and control of the data plane is redefined. The internal interface, which is based on the HTTP protocol, is SBI. The internal interface of the data plane adopts an improved interface, which between the data plane and the control plane adopts an improved interface. Thereby, the customization of the network is realized and the signaling interaction process is simplified to reduce the time delay of data transmission. In the following, we will introduce the integration method of proposed in detail.

5.2.1 Evolution of AAU

As shown in Figure. 7, the FFT/IFFT in the PHY lower layer of AAU is separated and virtualized, leaving only the AD/DA and RF functions to form RU. The functions of FFT and IFFT are virtualized as VNF on a general programmable platform through SDN technology, which can be integrated with the DU according to the user requirements. The integration scheme improves the flexibility of network slicing.

5.2.2 Integrating of CU-CP and AMF

As shown in Figure. 8, the 5G CU protocol stack includes radio resource control (RRC), packet data control protocol (PDCP) and service data adaptation protocol (SDAP). This paper uses SDN technology to separate the control plane and user plane of CU into CU-CP and CU-UP. The CU-CP protocol stack is composed of RRC and PDCP-CP, which mainly implements radio resource control, mobility management, and radio bearer control of PDUs. The CU-UP protocol stack is composed of PDCP-UP and SDAP, which

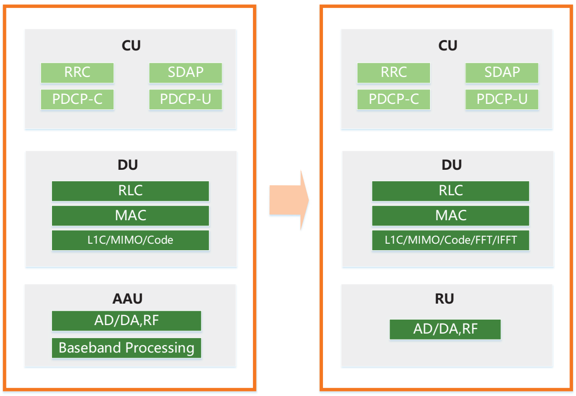


Figure 7. *Evolution for AAU.*

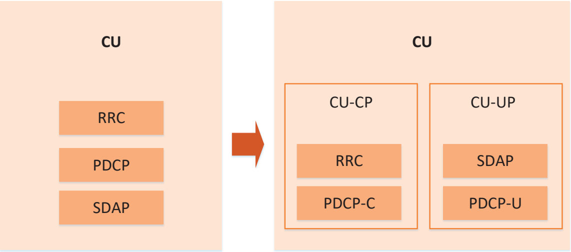


Figure 8. *Evolution for CU.*

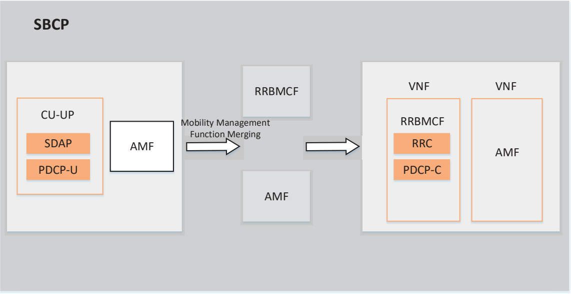


Figure 9. *Evolution for CU-CP.*

mainly implements QoS flow routing, PDU header compression, and data packet reconstructing and numbering.

As shown in Figure. 9, the functions related to mobility management and admission control in CU-CP are integrated with AMF to form a new AMF, and the functions related to radio resource management and radio bearer control of CU-CP are integrated into a new network element namely radio resource and bear management control function (RRBMCF). The protocol stack of the integrated control plane is shown in the Figure. 11.

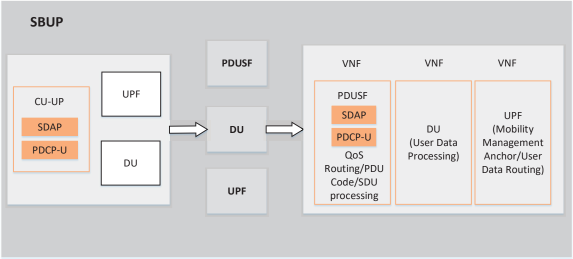


Figure 10. *Evolution for CU-UP.*

5.2.3 Integrating of CU-UP and UPF

As shown in Figure. 10, the CU-UP protocol stack includes PDCP-UP and SDAP. The function of CU-UP is mainly related to the packet data unit. This paper defines the function as a new network element namely packet data unit and session function (PDUSF), specially used to process services related to PDU and SDU. The protocol stack of the integrated user plane is shown in the Figure. 12.

5.3 New Interface and Protocol Stack

The proposed architecture is based on micro-service, so it is the interface of RRBMCF with the interior network element of the control plane that is based on HTTP. In the 5G-CU, the interface of the CU with the AMF is the N2, while the interface with UPF is the N3. In the architecture proposed, the interface of N2 for the control signaling is removed, while the NAS signaling message of user can be transmitted by the new interface of NRS1 directly. Second, the interface is removed.

5.3.1 Interface between user and control plane

In the proposed architecture, the interface of RRBMCF with the DU of date plane is defined as C1. The interface of RRBMCF with the PDUSF is defined as C2. RRBMCF is connected to the DU through the C1 to transmit the control signaling of the resource allocation and radio bearer control of user. When the user is connected to the network, RRBMCF performs resource allocation (bandwidth, frequency, power, radio resource block etc.) according to the individual requirements, and then assigns radio bearer for each PDU of the user. The PDUSF is informed via the C2 to perform a PDU number, encryption, decryption, reordering, repeated detection, and routing of

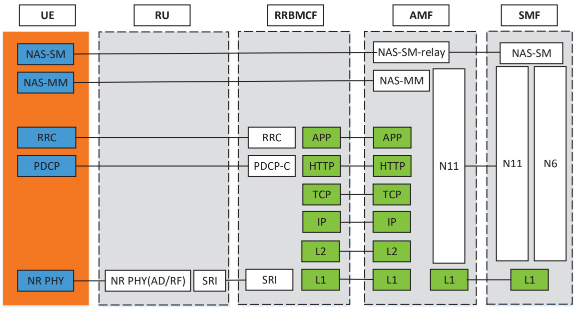


Figure 11. *Proposed control plane protocol stack.*

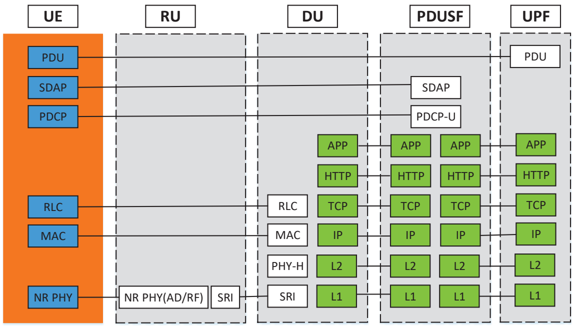


Figure 12. *Proposed user plane protocol stack.*

QoS streams after the RRBMCF processing is completed.

Both RRBMCF and AMF are service based, so the interface between them uses a unified HTTP. In order to support the interaction of the user and the control plane in SAGIN, it is necessary to be compatible with the SRI interface in satellite network. Therefore, this paper defines the interface of the new user and the control plane to NRS1, which includes NR-CP, N1, and SRI interface protocol stacks, and the control plane protocol stack as shown in Figure. 11. In Figure. 13, it is shown that the NAS signaling message is transmitted to the RRBMCF, AMF, SMF and other network elements through the NRS1 interface.

5.3.2 Interface between the user and user plane

In the novel architecture, the DU is responsible for tasks related to user radio signal processing and user data transmission to the PDUSF for encryption, numbering. In this architecture, the interface of PDUSF and UPF uses an SBI interface, and its protocol is HTTP.

5.4 Networking Cases with the Proposed SAGIN Architecture

The proposed SAGIN architecture can be flexibly deployed according to the requirements of different application scenarios to meet QoS constraints such as delay, bandwidth, and reliability. In the following, we give specific deployment cases and methods for two typical application scenarios.

5.4.1 Network deployment solution in mobile broadband communication scenarios

Typical applications of mobile broadband communication include ultra-high-definition video, virtual reality, voice communication, emergency communications, and augmented reality, etc. These scenarios require wide bandwidth and low latency. SAGIN is expected to provide enhancement in terms of coverage, access bandwidth, and service quality.

A possible deployment case for 5G NTN is shown in Figure. 14. The gNB or gNB-DU is deployed on LEO, and connects to the NTN gateway on the ground, and then connects to the 5GC through the NTN gateway. Due to the fast movement of LEO, handover occurs frequently. In each handover procedure, many-time handshakes are required. The control signaling of each handshake needs to be transmitted to the ground 5GC, and then transmitted back to the LEO. Therefore, the round-trip delay is large, and data transmission service may be interrupted, which reduces the users' QoS. The main reason is that 5GC is located on ground in the 5G NTN architecture.

In contrast, the proposed SAGIN has a more flexible deployment option for this scenario. Due to service-oriented RAN, lightened and autonomically distributed CN, and flattened network, some key CN elements, together with RAN, can be deployed on LEO, to improve users' QoS, as shown in Figure. 14. Therefore, personalized functions can be customized according to the application requirements and the capabilities of the LEO. To fit this scenario, AMF, SMF, UPF, RRBMCF, and PDUSF are deployed on LEO. AMF is used for access and mobility management, SMF for session management, UPF for user plane forwarding, RRBMCF for radio resource management and bearer control, and PDUSF for packet data unit and session. The core algorithms of these network elements are customized according to the characteristics

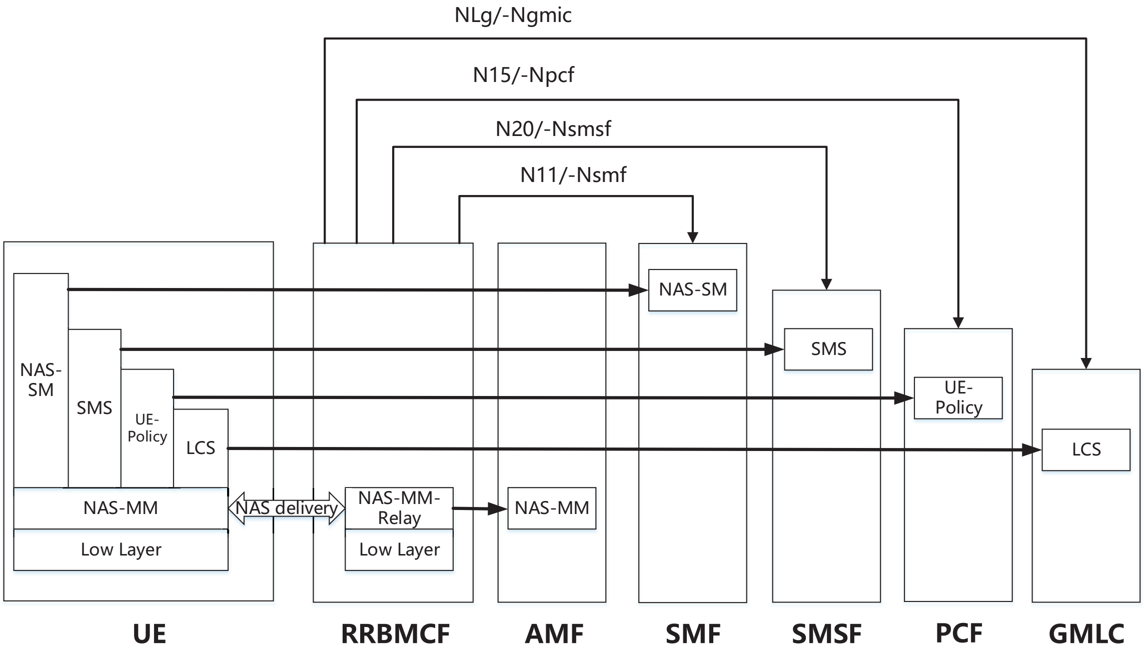
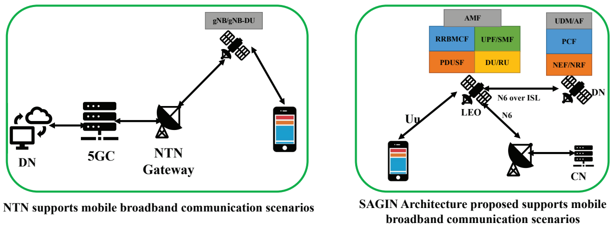


Figure 13. *The NAS signaling transmission process through NRS1.*



NTN supports mobile broadband communication scenarios

SAGIN Architecture proposed supports mobile broadband communication scenarios

Figure 14. *Network element deployment for mobile broadband communication scenarios.*

of satellite. In the handover procedure, control signaling only needs to be transmitted to the LEO, but does not need to be transmitted to the ground, which reduces handover delay and maintains continuity of service. Besides, routing decision can be made by the SMF deployed on the LEO, and can be directly distributed to the UPF located on the same LEO. In this way, forwarding and processing delay can be degraded.

5.4.2 Network deployment solution in large-scale IoT scenarios

Typical applications of large-scale IoT include smart cities, smart homes, remote area coverage, ecological remote sensing, and navigation, etc. These scenarios require high connection density, mass data processing and transmission capabilities, as well as the capability to support diverse and differentiated applications. SAGIN is expected to increase coverage, improve massive data processing and transmission capabilities, as well as the number of connections and service quality, etc.

If 5G NTN is exploited to support the scenario, a possible deployment is shown in Figure. 15. The gNB or gNB-DU is deployed on LEO, and connects to the NTN gateway on the ground, and then connects to the 5GC through the NTN gateway. It is noteworthy that to process large amounts of data transmitted by satellites, MEC is deployed on the NTN gateway. However, these large data needs to be transmitted to the ground for processing, which causes high delay and increases bandwidth and cost of backhaul. Due to fast moving of LEO, the problem of frequent handover still exists, therefore, the round-trip delay of signaling is still large. In addition, due to the huge number of sensor users in the scenario, the handover signaling can be large and even causes a signaling storm. The main reason is that the MEC and CN are located on the ground, and there lacks specific mobility management scheme for massive connections in the 5G NTN architecture.

In contrast, the proposed SAGIN can provide a more efficient deployment scheme for this scenario, because, in the proposed SAGIN architecture, RAN is service-oriented, CN is lightened and autonomically

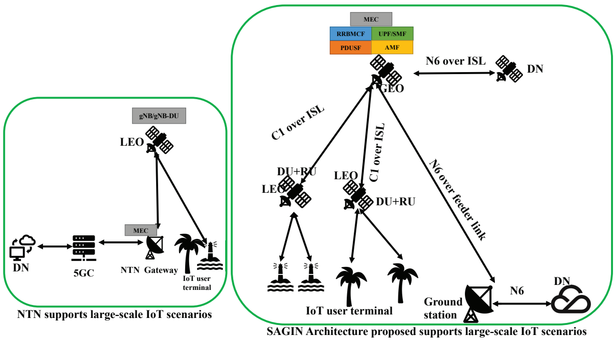


Figure 15. *Network element deployment for large-scale IoT scenarios.*

distributed, and network is flattened. Thus, MEC, some important CN elements, together with RAN, can be deployed on LEO to improve massive data processing and enhance users' QoS, as shown in Figure. 15. To meet this scenario, MEC, AMF, UPF, SMF, RRBMCF and PDUSF are deployed on the LEO. As the MEC is deployed on the LEO, large data generated by large-scale IoT terminals can be directly processed on the satellite, rather than being transmitted to the ground. As a result, the bandwidth and cost of backhaul are reduced. Moreover, in our proposed SAGIN architecture, the functions of network elements can be customized according to the characteristics of different application scenarios. Therefore, specific mobility management scheme for the AMF can be designed to eliminate signaling storms and reduce handover overhead in the large-scale IoT.

VI. FUTURE RESEARCH DIRECTIONS AND CHALLENGES

To successfully apply our proposed network architecture, there are some important research directions that should be investigated and some key challenges need to be overcome in the future.

6.1 Intelligent Deployment of Network Elements

The application scenarios of 6G are diversified, while different application scenarios have various requirements for network capabilities. To achieve differentiated network capabilities to support differentiated application scenarios, intelligent network element deployment is an important research direction in the future.

Besides the benefits of lite and customizable network elements that can be deployed in satellite and airborne networks, many challenging issues are exposed when deploying network elements of the proposed SAGIN architecture. In the following, from a perspective of deployment and design, we present several challenges that require further study.

1 *Backup mechanisms and recovery strategies.* Satellite networks are susceptible to the physical environment. When the satellite platform is maintained or energy-saving, connection will be terminated. As such the network will lose part of the control and forwarding function, and even cause the destruction of the satellite network. In order to solve this challenge, multiple network element backup mechanisms, fault detection and recovery strategies are indispensable.

2 *Coordinated load balancing.* In the proposed SAGIN architecture, the diversity of user density and QoS requirements in different regions will result in uneven traffic distribution and unbalanced network element deployment. Therefore, a flexible network element management strategy and a coordinated load balancing solution based on distributed deployment are required.

6.2 Mobility Management

In SAGIN, user terminals can make the best of the ubiquitous coverage advantages of non-terrestrial networks and low-latency transmission advantages of terrestrial networks to handover freely between satellite and ground networks as required. For LEO, since the satellite moves fast relative to the ground, special mobility management scheme will be an important research direction in the future.

When designing a mobility management solution for LEO communications, there are still the following challenges that need to be solved.

1 *Prediction-based handover.* To reduce the handover delay, the handover timing of the user terminal can be predicted in advance. The handover technology based on ephemeris information, according to the characteristics of the satellite moving track, etc., enables the accurate beam tracking. However, the prediction problem is usually

modeled as a MILP problem, which is a very difficult NP-hard problem to deal with.

2 *User Position based handover.* As the transmission distance between the satellite and the ground is relatively long, the distance effect in the satellite coverage cell is not obvious. Therefore, the user terminal position is very important for the handover decision. Together with the L1/L3 measurement value, user Position is used as the handover decision condition to improve the consistency of the service quality of the edge users. However, user's location information involves the user's privacy and security. How to obtain the location information and protect the user's privacy and security is another key challenge.

6.3 Intelligent Session Management

Future network will support more vertical services such as holographic communication, while large-scale mobile terminals may simultaneously access multiple network. Highly concurrent data application management and multi-connection routing algorithms are important research directions in the future.

To meet the consistent and seamless perception of user services, multiple connections and intelligent session management technologies play a vital role. However, there are still some open challenges.

1 *Intelligent routing of concurrent transmission.* SAGIN has multiple heterogeneous access networks. Bandwidth, delay, packet loss rate, and network congestion control status of each access network are different. The subflow routing rules of multi-path will affect the data disorder and the transmission delay at the receiver. However, in a satellite network with highly dynamic topology, efficient routing algorithm design is challenging.

2 *Intelligent data scheduling technology for concurrent transmission of multiple connections.* A group of subflow of users is selected to maximize the average throughput for concurrent transmission from a set of available subflows. However, the optimized average throughput will increase the receiving buffer. As, the problem is a mixed integer programming (MILP). Solving this NP-hard problem is significantly challenging.

6.4 Air Interface

In order to support the transmission of multiple services, the terminal is extremely simple to access the most suitable heaven and earth network node. Therefore, the unified air interface design scheme is an important research direction in the future.

Unified air interface mechanism supports the transmission of different services. Unified air interface mechanism efficiently implements mobile broadband data transmission, voice transmission, Internet of Things service transmission (URLLC and mMTC service), and broadcast service transmission. Moreover, it avoids a closed chimney development model for each business. Below, we present the following challenge that require further study.

1 *Novel air interface mechanism.* The link transmission of the space-air network is quite different from that of the ground network. Compared with terrestrial networks, distance scales of access links and feeder links have increased by orders of magnitude, resulting in high propagation loss and delay. Fast moving of the space-air platform results in significant Doppler shift. Air interface of the space-air network should be based on the advanced technology of the existing mobile communication network, combined with the characteristics of the space-air network, such as strong mobility, long transmission delay, and limited load. In this way, adaptive improvement can be archived in terms of frequency planning, transmission waveforms, modulation methods, channels bandwidth, and antenna configuration, etc.

VII. CONCLUSION

This paper first explored the requirements of new emerging technologies and applications for 6G, and discussed the limitations of using 5G NTN to support the requirements. After deep analysis, it was found that 5G NTN has a long data forwarding delay, because its CN is located on the ground. Besides, 5G NTN lacks flexibility of network management, due to the inadequate virtualization and softwarization. To overcome the above problems, a novel SAGIN architecture is proposed. In the new architecture, with sufficient virtualization and softwarization, CN and RAN are integrated into a single network, which consists of

SBCP, SBUP and RU. The interfaces, including interface between user and control plane, and that between user and data plane, have been redefined to share a unified and simple protocol stack. The proposed architecture brings many advantages. First, SBCP and SBUP can be deployed on satellite and airborne platforms, which greatly reduces the forwarding delay. Second, high degree of softwarization and virtualization makes the network fulfill on-demand services in a significantly flexible manner. Moreover, this paper studied two different deployment scenarios to illustrate the advantages of low delay and high flexibility. Finally, future research directions were discussed to further enhance the architecture, so as to make it a competitive candidate for 6G.

ACKNOWLEDGEMENT

This work was supported in part by the National Key Research and Development Program under grant number 2020YFB1806800, the Beijing Natural Science Foundation under grant number L212003, and the National Natural Science Foundation of China (NSFC) under grant numbers 62171010 and 61827901.

References

[1] N. Alliance, “5g white paper -executive version,” 2014.

[2] W. Saad, M. Bennis, *et al.*, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE network*, vol. 34, no. 3, 2019, pp. 134–142.

[3] N. Zhao, W. Lu, *et al.*, “Uav-assisted emergency networks in disasters,” *IEEE Wireless Communications*, vol. 26, no. 1, 2019, pp. 45–51.

[4] X. Chen, D. Li, *et al.*, “Securing aerial-ground transmission for noma-uav networks,” *IEEE Network*, vol. 34, no. 6, 2020, pp. 171–177.

[5] Z. Xiao, L. Zhu, *et al.*, “Uav communications with millimeter-wave beamforming: Potentials, scenarios, and challenges,” *China Communications*, vol. 17, no. 9, 2020, pp. 147–166.

[6] L. Zhu, J. Zhang, *et al.*, “Millimeter-wave full-duplex uav relay: Joint positioning, beamforming, and power control,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, 2020, pp. 2057–2073.

[7] L. Zhu, J. Zhang, *et al.*, “Millimeter-wave noma with user grouping, power allocation and hybrid beamforming,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, 2019, pp. 5065–5079.

[8] E. Yaacoub and M.-S. Alouini, “A key 6G challenge and opportunity—connecting the base of the pyramid: A survey on rural connectivity,” *Proceedings of the IEEE*, vol. 108, no. 4, 2020, pp. 533–582.

[9] K. David and H. Berndt, “6G vision and requirements: Is there any need for beyond 5G?” *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, 2018, pp. 72–80.

[10] S. Dang, O. Amin, *et al.*, “What should 6G be?” *Nature Electronics*, vol. 3, no. 1, 2020, pp. 20–29.

[11] N. Zhang, S. Zhang, *et al.*, “Software defined space-air-ground integrated vehicular networks: Challenges and solutions,” *IEEE Communications Magazine*, vol. 55, no. 7, 2017, pp. 101–109.

[12] Y. Shi, Y. Cao, *et al.*, “A cross-domain SDN architecture for multi-layered space-terrestrial integrated networks,” *IEEE Network*, vol. 33, no. 1, 2019, pp. 29–35.

[13] Y. Wang, Y. Xu, *et al.*, “Hybrid satellite-aerial-terrestrial networks in emergency scenarios: a survey,” *China Communications*, vol. 14, no. 7, 2017, pp. 1–13.

[14] G. Giambene, S. Kota, *et al.*, “Satellite-5G integration: A network perspective,” *IEEE Network*, vol. 32, no. 5, 2018, pp. 25–31.

[15] 3GPP, *Solutions for NR to support Non-Terrestrial Networks (NTN) (Release 16)*, Std., TR 38.821 V1.0.0 Dec. 2018.

[16] 3GPP, *Solutions for NR to support Non-Terrestrial Networks (NTN) (Release 17)*, Std., RP 193234 3GPP TSG RAN Meeting 86 Dec.2019.

[17] A. U. Gawas, “An overview on evolution of mobile wireless communication networks: 1G-6G,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 3, no. 5, 2015, pp. 3130–3133.

[18] B. H. Walke, “The roots of GPRS: the first system for mobile packet-based global internet access,” *IEEE Wireless Communications*, vol. 20, no. 5, 2013, pp. 12–23.

[19] J. Korhonen, *Introduction to 3G mobile communications*. Artech House, 2003.

[20] S. Chen, S. Sun, *et al.*, “A comprehensive survey of TDD-based mobile communication systems from TD-SCDMA 3G to TD-LTE (A) 4G and 5G directions,” *China Communications*, vol. 12, no. 2, 2015, pp. 40–60.

[21] K. Samdanis and T. Taleb, “The road beyond 5G: A vision and insight of the key technologies,” *IEEE Network*, vol. 34, no. 2, 2020, pp. 135–141.

[22] A. Ghosh, A. Maeder, *et al.*, “5G evolution: A view on 5G cellular technology beyond 3GPP release 15,” *IEEE Access*, vol. 7, 2019, pp. 127 639–127 651.

[23] M. Chahbar, G. Diaz, *et al.*, “A comprehensive survey on the E2E 5G network slicing model,” *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, 2021, pp. 49–62.

[24] E. C. Strinati, S. Barbarossa, *et al.*, “6G: The next frontier,” *IEEE Vehicular Technology Magazine*, vol. PP, no. 99, 2019, pp. 1–1.

[25] Y. Zhao, G. Yu, *et al.*, “6G mobile communication network: vision, challenges and key technologies,” *arXiv preprint arXiv:1905.04983*, 2019.

[26] L. Zhang, Y.-C. Liang, *et al.*, “6G visions: Mobile ultra-broadband, super internet-of-things, and artificial intelligence,” *China Communications*, vol. 16, no. 8, 2019, pp. 1–14.

[27] L. Da Xu, W. He, *et al.*, “Internet of things in industries: A survey,” *IEEE Transactions on industrial informatics*, vol. 10, no. 4, 2014, pp. 2233–2243.

[28] E. Sisinni, A. Saifullah, *et al.*, “Industrial internet of things:

Challenges, opportunities, and directions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, 2018, pp. 4724–4734.

[29] S. Fu, B. Wu, *et al.*, "Multi-resources management in 6g-oriented terrestrial-satellite network," *China Communications*, vol. 18, no. 9, 2021, pp. 24–36.

[30] Y. Liu, X. Yuan, *et al.*, "Federated learning for 6g communications: Challenges, methods, and future directions," *China Communications*, vol. 17, no. 9, 2020, pp. 105–118.

[31] X. Tang, C. Cao, *et al.*, "Computing power network: The architecture of convergence of computing and networking towards 6g requirement," *China Communications*, vol. 18, no. 2, 2021, pp. 175–185.

[32] G. Liu, Y. Huang, *et al.*, "Vision, requirements and network architecture of 6g mobile network beyond 2030," *China Communications*, vol. 17, no. 9, 2020, pp. 92–104.

[33] "2020 tencent artificial intelligence white paper." https://ishare.iask.sina.com.cn/f/1f84jXu7NeJf.html, 2020.

[34] L. U. Khan, W. Saad, *et al.*, "Digital-twin-enabled 6G: Vision, architectural trends, and future directions," *arXiv preprint arXiv:2102.12169*, 2021.

[35] E. Juan, M. Lauridsen, *et al.*, "5g new radio mobility performance in leo-based non-terrestrial networks," in *2020 IEEE Globecom Workshops (GC Wkshps*, 2020, pp. 1–6.

Biographies

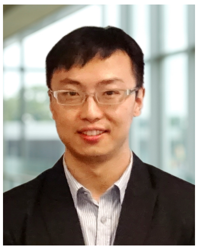
Huanxi Cui is currently pursuing the Ph.D. degree at Beihang University (BUAA). His main research interests include the mobility management of heterogeneous network, space-air-ground integrated network architecture, network slicing, optimization theory, deep reinforcement learning.

Jun Zhang received the B.S., M.S. and Ph.D. degrees in communications and electronic systems from Beihang University, Beijing, China, in 1987, 1991, and 2001, respectively. He used to be a professor in Beihang University, and has served as the dean of the school of Electronic and Information Engineering, the vice president, and the secretary of the Party Committee of Beihang University. Now he is a professor in Beijing Institute of Technology, and also the president of Beijing Institute of Technology. His research interests are networked and collaborative air traffic management systems, covering signal processing, integrated and heterogeneous networks, and wireless communications. He has won the awards for science and technology in China many times, and he is a member of Chinese Academy of Engineering.

Yuhui Geng is pursuing her phD in the school of Electronic and Information Engineering of Beihang University. Her research interests include intelligent placement of VNF in SAGIN, SDN and SAGIN communications.

Zhenyu Xiao (M'11, SM'17) received the B.E. degree with the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2006, and the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2011. From 2011 to 2013, he held a post-doctorial position with the Electronic Engineering Department, Tsinghua University. From 2013 to 2016 he was a Lecturer with the Department of Electronic and Information Engineering, Beihang University, Beijing, China, where he is currently an associate professor. Dr. Xiao has published over 60 papers, and served as reviewers for IEEE Transactions on Signal Processing, IEEE Transactions on Wireless Communications, IEEE Transactions on Vehicular Technology, IEEE Communications Letters, etc. He has been TPC members of IEEE GLOBECOM'12, IEEE WCSP'12, IEEE ICC'15, etc. His research interests are communication signal processing and practical system implementation for wideband communication systems, which cover synchronization, multipath signal processing, diversity, multiple antenna technology, etc. Currently his is dedicated in millimeter-wave 5G and airborne communications.

Tao Sun received his Ph.D. degree in control science and engineering from Tsinghua University in 2008. His research interest covers new mobile network architectures and AI enabled networks.



Ning Zhang (Senior Member, IEEE) received the Ph.D. degree from the University of Waterloo, Canada, in 2015. He was a Post-Doctoral Research Fellow with the University of Waterloo and also with the University of Toronto, Canada. He is currently an Associate Professor with the University of Windsor, Canada. He received the Best Paper Awards from IEEE Globecom in 2014, IEEE WCSP in 2015, and the Journal of Communications and Information Networks in 2018, IEEE ICC in 2019, the IEEE Technical

Committee on Transmission Access and Optical Systems in 2019, and IEEE ICCC in 2019, respectively. He also serves/served as a track chair for several international conferences and a co-chair for several international workshops. He serves as an Associate Editor for the IEEE Internet of Things Journal, the IEEE Transactions on Cognitive Communications and Networking, IEEE Access, IET Communications, and Vehicular Communications. He was a Guest Editor of several international journals, such as the IEEE Wireless Communications, the IEEE Transactions on Industrial Informatics, and the IEEE Transactions on Cognitive Communications and Networking.

Jiajia Liu (Senior Member, IEEE) was a Full Professor at the School of Cyber Engineering, Xidian University, from 2013 to 2018. Since Jan. 2019, he has been a Full Professor at the School of Cybersecurity, Northwestern Polytechnical University. He has published more than 180 peer-reviewed papers in many high quality publications, including prestigious IEEE journals and conferences. He received IEEE VTS Early Career Award in 2019, IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2019, Niwa Yasujiro Outstanding Paper Award in 2012, the Best Paper Awards from many international conferences including IEEE flagship events, such as IEEE GLOBECOM in 2016 and 2019, IEEE WCNC in 2012 and 2014, IEEE WiMob in 2019, IEEE IC-NIDC in 2018. He was also a recipient of the Tohoku University President Award 2013. His current research interests include a wide range of areas including intelligent and connected vehicles, mobile/edge/cloud computing and storage, Internet of things security, wireless and mobile ad hoc networks, and space-air-ground integrated networks. He has been actively joining the society activities, like serving as Associate Editors for IEEE Transactions on Wireless Communications (May 2018–present), IEEE Transactions on Computers (Oct. 2015–Jun. 2017) and IEEE Transactions on Vehicular Technology (Jan. 2016–present), Editor for IEEE Network (July 2015–present), Editor for IEEE Transactions on Cognitive Communications and Networking (January 2019–present), Guest Editors of top ranking international journals like IEEE Transactions on Emerging Topics in Computing, IEEE Vehicular Technology Magazine, IEEE Internet of Things Journal, etc., and serving as technical program committees of numerous international conferences like the leading symposium Co-Chair of AHSN symposium for GLOBECOM 2017, CRN symposium for ICC 2018, AHSN symposium for ICC 2019. He is the Vice Chair of IEEE IOT-AHSN TC, and is a Distinguished Lecturer of IEEE Communications Society and Vehicular Technology Society.

Qihui Wu (SM'13) received the B.S. degree in communications engineering and the M.S. and Ph.D. degrees in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1994, 1997, and 2000, respectively. From 2003 to 2005, he was a Postdoctoral Research Associate with Southeast University, Nanjing. From 2005 to 2007, he was an Associate Professor with the College of Communications Engineering, PLA University of Science and Technology, Nanjing, where he was a Full Professor from 2008 to 2016. Since May 2016, he has been a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. From March 2011 to September 2011, he was an Advanced Visiting Scholar with the Stevens Institute of Technology, Hoboken, NJ, USA. His current research interests include the areas of wireless communications and statistical signal processing, with emphasis on system design of software defined radio, cognitive radio, and smart radio.

Xianbin Cao (M'08-SM'10) is Dean and a Professor at the School of Electronic and Information Engineering, Beihang University, Beijing, China. His current research interests include intelligent transportation systems, airspace transportation management, and intelligent computation.