

# Learning-Assisted Dynamic VNF Selection and Chaining for 6G Satellite-Ground Integrated Networks

Jianxin Zhang, Qiang (John) Ye, *Senior Member, IEEE*, Kaige Qu, *Member, IEEE*, Yanglong Sun, *Member, IEEE*, Yuliang Tang, *Member, IEEE*, Dongmei Zhao, *Member, IEEE*, Tong Ye

**Abstract**—The sixth generation (6G) mobile communication system is expected to provide global seamless network coverages, where a satellite-ground integrated network (SGIN) is seen as one of the typical 6G networking paradigms. In this paper, a dynamic virtual network function (VNF) selection and chaining (DVSC) problem in an SGIN is investigated. We aim to balance the network resource provisioning and VNF migration costs with service performance gain to maximize the long-term network profit. Specifically, we formulate the DVSC problem as a Markov decision process (MDP), by taking into consideration the heterogeneity and time-varying nature of SGINs. A novel VNF selection and chaining scheme is proposed, where a deep *Q*-learning (DQL) algorithm is designed to dynamically determine a set of VNF selection and chaining policies (VSCPs) based on the evolving network states (e.g., network resources, network topology, and network traffic load). Furthermore, to elaborate the level of computing resource sharing of VSCP sets, a new sharing ratio (SR) is proposed. To efficiently allocate heterogeneous network resources, the action space is built by clustering the historical records of the network load and selecting the VSCP set for each cluster in a greedy manner. Extensive simulation results are presented to demonstrate the effectiveness of the proposed framework in comparison with the state-of-the-art schemes.

**Index Terms**—6G, satellite-ground integrated networks, service function chaining, NFV, SDN, deep reinforcement learning.

## I. INTRODUCTION

THE sixth generation (6G) mobile communication systems are envisioned to extend wireless communication coverages from cities to remote areas, such as oceans, deserts, and forests, providing global seamless coverages and ubiquitous connectivities. However, existing terrestrial communication networks cover about 20 percent of all terrestrial areas over the globe, less than 6 percent of the entire earth surface, leaving a huge gap for achieving global coverages [1]. Satellite

Jianxin Zhang, Yuliang Tang and Tong Ye are with the School of Informatics, Xiamen University, Xiamen, 361001, China (email: zjxxmu@stu.xmu.edu.cn; tyl@xmu.edu.cn; tongyy@stu.xmu.edu.cn).

Qiang (John) Ye is with the Department of Electrical and Software Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: qiang.ye@ucalgary.ca).

Kaige Qu is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (email: k2qu@uwaterloo.ca).

Yanglong Sun is with the Navigation College, Jimei University, Xiamen, 361021, China (email: ylsun4work@gmail.com).

Dongmei Zhao is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: dzhao@mcmaster.ca).

communication remains a good option to extend the communication coverage, and has been regarded as one of the important components for 6G. One promising networking architecture for 6G is to integrate ground networks with satellites to form a satellite-ground integrated network (SGIN), which can provide extended communication services towards users in remote areas and meet responsive user requirements in emerging 6G applications, such as immersive online education, logistics tracking, and remote data collection [2].

In an SGIN, a flexible and efficient resource management framework is expected to coordinate the allocation of heterogeneous network resources (e.g., computing and communication resources) from different network segments (i.e., satellite network segment and ground network segment). Network function virtualization (NFV) and software-defined networking (SDN) are recognized as two potential technologies to improve resource allocation efficiency and realize the integration of satellites with ground networks [3]–[5]. Through NFV, network functions, e.g., intrusion detection system (IDS) and network address translation (NAT), are decoupled from conventional function-specific middleboxes and softwarized as virtual network functions (VNFs) hosted at generic commodity servers. To ensure efficient end-to-end (E2E) data delivery, SDN is leveraged to decouple the control plane from the data plane to realize a centralized network control, with packet forwarding rules configured at each software programmable switch via the Openflow protocol [6]. Based on NFV and SDN, an E2E service can be flexibly supported in an SGIN by forming a service function chain (SFC) which comprises a sequence of VNFs interconnected by virtual links in a predefined order. These VNFs can be flexibly instantiated and executed on heterogeneous network elements (e.g., satellites and ground nodes) in the SGIN by an NFV orchestrator (NFVO), which is often integrated with the SDN controller and placed in a logically centralized location for VNF placement and flow routing.

To improve the service performance in a cost-effective way, we need to determine the optimal VNF selection and chaining policy (VSCP) for each SFC which consists of VNF selection and virtual link mapping decisions. The VNF selection is to map VNFs in an SFC to the nodes installed with corresponding VNF instances to process traffic, and the virtual link mapping is to allocate physical transmission resources to virtual links to interconnect the selected VNF instances. A VSCP can steer a traffic flow through multiple VNFs in a predefined order to

fulfil a composite service in E2E data delivery.

The dynamic VNF selection and chaining (DVSC) problem in the fifth generation (5G) mobile communication systems has been extensively investigated [7]–[12]. These works mainly study an SFC mapping problem in traditional wired networks, where ground networks are operated independently to support the 5G services. Different from the 5G networks, in SGINs, VNFs and virtual links are mapped to heterogeneous network nodes (e.g., satellites and ground nodes) and communication links. Satellites can steer traffic flows to earth station (ES) for using the VNF instances installed at ground nodes, and improve network computing capacity by deploying VNFs to satellites. Ground nodes can utilize abundant computing resources to host computation intensive VNFs (e.g., image rendering) and process packets from multiple flows on shared VNF instances. Meanwhile, satellite movements may lead to SFC mapping decision changes which also needs to be taken into consideration. Therefore, an SFC mapping problem in 6G SGINs should be studied by taking into consideration the heterogeneity and time-varying features of the SGINs, which would lead to new technical challenges for dynamic VNF selection and chaining.

First, an SGIN is heterogeneous consisting of different network segments with complementary characteristics. Specifically, a satellite network segment can provide a global communication coverage with limited computation capacity and communication bandwidth, whereas a ground network segment may have sufficient computing and communication resources but with a restricted communication coverage. To save the computing resource provisioning cost, existing works focus on how to execute VNF instances at shared ground nodes, and on leveraging the satellite networks as a complementary solution to balance the network-wide computation load [13], [14]. To efficiently allocate heterogeneous network resources, existing studies extract differentiated network attributes (e.g., node computing resources, communication coverages, and link bandwidth resources) for different network segments and obtain VSCPs via matching between network resources and service demands [15], [16]. To improve service performance, more VNF instances may need to be executed. Satellite networks can leverage global communication coverages to route flows through VNF instances installed at different network locations [17], [18]. Given traffic rates of flows and network topology, the above works can obtain a near-optimal SFC mapping scheme for SGINs in terms of resource utilization efficiency or service performance. However, SFCs are remapped when traffic rates of flows or network topology change, which may lead to frequent VNF migrations. Considering a more practical scenario with diverse traffic flows and dynamic traffic rates, we need to investigate how a VSCP set can be dynamically determined over sequential time slots to make use of the complementary advantages of different network segments in SGINs and how to balance the network resource provisioning and VNF migration costs with service performance.

Second, SGIN topology varies with time due to satellite movements. It leads to the connection may need to be switched among different satellites and variations in the traffic routing path over time [19]. In [20]–[22], authors use the time-

varying graphs to describe the topology evolution of the satellite network where a sequence of discrete time intervals is considered, with a fixed topology assumed in each interval, and VNF placement and routing decisions are calculated at each topology snapshot. As satellite movements, the above approach requires many topology snapshots to describe the changing network topology and increases the complexity of algorithms. The VN approach is used to divide the Earth's surface into cells and assign an accessible satellite for each cell based on geodetic positions of the subsatellite points, and the routing path is calculated based on the segment routing approach [23]. Nevertheless, the VNF selection and chaining decision updates still need to be further studied when topological changes. In this paper, we characterize the topological changes between the satellite network segment and the ground network segment over sequential time slots, and design VSCPs consisting of multiple SFC mapping decisions to adapt to the changing network topology.

To deal with the aforementioned issues for dynamic service provisioning in SGINs, in this paper, we present a deep  $Q$ -learning (DQL) based dynamic VNF selection and chaining framework for an SGIN. We investigate how to determine the VSCP set in a heterogeneous and time-varying SGIN with dynamic network traffic load. The main contributions are as follows:

- An SDN/NFV-based SGIN architecture is established to support ubiquitous services such as immersive online education for users in remote areas and cities, in which a sequence of VNFs in an SFC are traversed in a predefined order based on a VSCP in a time-varying SGIN topology.
- A DVSC problem in an SGIN is formulated as a Markov decision process (MDP) to capture the relation between network state and VSCP set. The objective is to balance the network resource provisioning and VNF migration costs with service performance gain to maximize the long-term network profit.
- We use a VSCP set to represent a joint VNF selection and chaining policy of multiple flows. Furthermore, we propose a new *sharing ratio* (*SR*) to elaborate the level of computing resource sharing of VSCP sets. To efficiently allocate heterogeneous network resources, VNF selection and chaining policies are selected by the following steps: 1) distinguish network load levels by clustering the historical records of the network load, and 2) search the optimal VSCP set for different network loads in a greedy manner according to the network resource provisioning cost, service performance gain, and the value of SR. These selected VSCP sets constitute the action space, and a deep  $Q$  network (DQN) is trained to determine the VSCP set based on the evolving network states to maximize the accumulative network profit.
- The satellites covering the ground network segment are identified based on the VN approach. To guarantee the selected VSCP set is feasible with satellite movements, VSCP sets consisting of a sequence of VNF selection and chaining decisions are delivered to satellites covering the ground network segment.
- Extensive simulation results demonstrate that our pro-

posed DQL based VNF selection and chaining algorithm outperforms baseline algorithms and approaches the performance upper bound. The proposed framework balances the resource provisioning and VNF migration costs with service performance gain in the presence of a dynamic network load.

The rest of the paper is organized as follows. Section II provides a literature review. The system model is presented in Section III. The VNF selection and chaining problem in SGINs is formulated in Section IV. In Section V, a DQL-based VNF selection and chaining algorithm is proposed. In Section VI, simulations are conducted to demonstrate the effectiveness of the proposed scheme, followed by the conclusions given in Section VII.

## II. RELATED WORK

SFC technology enables elastic service provisioning and reduces network operation cost, and has been widely used in 5G networks. SFC orchestration in NFV-based wireline networks is usually formulated as an integer linear programming (ILP) [7], [8], or mixed integer linear programming (MILP) problem [9], [10], and heuristic algorithms are proposed to optimize several parameters, such as embedding cost, network throughput, QoS, energy efficiency, etc. Besides, the SFC embedding problem is formulated as an MDP model, and deep neural network (DNN)-based [11] or graph-neural network (GNN)-based [12] approaches are utilized to efficiently make embedding decisions. The existing works focus on wired networks, mostly a single network segment consisting of ground nodes and wired links. SGINs, as one of the typical 6G network paradigms, consist of an integration of different network segments, including a satellite network segment and a ground network segment, where satellites and ground nodes have different features in coverage, dynamics, communication, and computing capacities [24]–[26].

There are some research works on VNF selection and chaining problems in SGINs in terms of supporting future wireless applications (e.g., immersive online education, remote driving and remote sensing). In [13], [14], the service chain instantiation problem in an SGIN is studied to improve resource utilization efficiency. A heuristic greedy algorithm is presented in [13], where a higher priority is allocated to ground networks to deploy VNFs to make full use of their computing resources. In [14], a related approach considering VNF multiplexing and SFC merging is proposed to save the network resource provisioning cost. Zhang *et al.* [15], [16] model the heterogeneous resource allocation in a space-air-ground integrated network (SAGIN) as a multi-domain SFC embedding problem. In [15], a deep reinforcement learning (DRL) method is proposed where the agent is a multi-layer policy network used to determine the SFC embedding strategy and the feature matrix is built by extracting the SAGIN resource attributes to provide an environment for agent training. In [16], an SAGIN architecture with service function chaining is established, and an SFC mapping algorithm based on delay prediction is proposed to meet the requirements for delay-sensitive services. Both works improve the efficiency and rationality of the allocation of heterogeneous network

resources in SAGIN. To improve service performance, more computing and communication resources may need to be used. Satellite networks are seen as a good complement to ground networks in terms of providing additional communication and computing resources to increase the network capacity and coverage [17], [18]. In [17], a VNF instance scaling problem for SGINs is investigated, and an algorithm based on traffic forecasting is proposed to scale VNF instances where satellites can host some VNF instances to increase the network capacity. In [18], the authors study a bidirectional mission offloading problem for SDN/NFV-enabled SAGIN. A heuristic routing algorithm is proposed, where satellites act as relay nodes to route flows through VNF instances installed at different network locations to reduce the communication delay with fewer hops. The existing works can efficiently map SFCs under SGINs where the network traffic load and the network topology are considered relatively stable. However, the SFC mapping decisions would need to be re-determined when network state transitions happen (e.g., a link connection or network traffic load changes), which may result in VNF migrations. As a result, when the network environment changes more frequently, it will lead to an increase in VNF migration cost and a decrease in service performance. Different from the existing works, we aim to balance network resource provisioning and VNF migration costs with performance gain to maximize the long-term network profit.

Satellite movements result in SGIN topology variations, and the VNF selection and chaining decisions should be changed to adapt to the time-varying network topology. In [20]–[22], a time-varying graphs method is used to describe the changing network topology. In [20], a dynamic evolution of topology is described over discrete time slots, and a tabu search-based algorithm is proposed to determine the SFC mapping scheme for each static network topology. Further, considering the storage capacity of nodes, storage links are introduced to connect adjacent network topology instances, and the VNF deployment and routing policy can be performed across multiple time slots [21], [22]. More snapshots are needed to characterize the topological changes over time, which lead to more storage overhead and increased algorithm complexity. In [23], the VN method is used to divide the Earth's surface into cells, and assign a satellite to each VN, where the routing path is divided into multiple segments by middle points, and satellites need only forward packets to the middle points. However, if a user changes its accessing satellite (i.e., satellite handover), the midpoints and routes need to be recalculated which increases the computational load on the controller. The dynamic VSCP that can adapt to SGIN topological changes still needs to be investigated.

## III. SYSTEM MODEL

We consider an SDN/NFV-enabled SGIN, as shown in Fig. 1. The satellite network segment is formed by interconnecting satellites, very-small-aperture-terminals (VSATs), and ESs, where VSATs provide bidirectional communication between end users and satellites using satellite-to-ground links (SGLs), ESs provide bidirectional communication between terrestrial infrastructures and satellites using SGLs, and satellites are

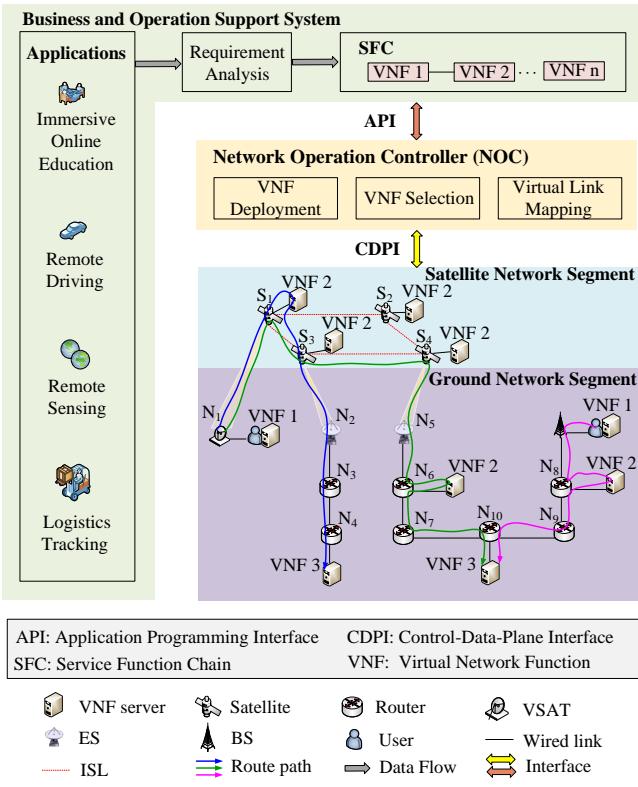


Fig. 1: An SDN/NFV enabled SGIN architecture.

interconnected via inter-satellite links (ISLs). The satellite network segment of an SGIN is modeled by using the satellite constellation for phase I of Starlink [27]. We consider two types of ISLs based on satellite locations within a constellation: *intra-plane ISLs*, which are established by satellites in the same orbit plane, and *inter-plane ISLs*, which are formed by satellites in adjacent orbit planes. Each satellite establishes four permanent ISLs with its neighboring satellites, including two intra-plane ISLs and two inter-plane ISLs [19]. The ground network segment is established through interconnecting base stations (BSs) and routers using wired links, where BSs provide network access for local end users, and routers are responsible for packet forwarding via wired links. Based on NFV/SDN technologies, VNFs can be flexibly deployed at the satellites, routers and end users, and the network operation controller (NOC) can monitor the network state and service requirements with a global view and enforce flow entries to programmable network nodes to improve traffic steering efficiency.

There can be multiple paths traversing a sequence of VNFs in a predefined order to provide service for end users at different locations (e.g., the blue arrow, the green arrow, and the pink arrow in Fig. 1). Note that due to the wide communication coverage of the satellite network segment, satellites can route flows to any location where an ES is deployed (e.g., the blue solid arrow and the green arrow in Fig. 1). Table I and II summarize the important parameters, variables and acronyms.

TABLE I: SUMMARY OF IMPORTANT PARAMETERS

Notations	Descriptions
$V_t$	The set of physical nodes at slot $t$
$E_t$	The set of physical links at slot $t$
$\mathcal{F}$	The set of VNFs required by the service
$\mathcal{L}$	The set of flows in network
$\pi_l$	The SFC of flow $l$
$\mathcal{F}_l$	The set of VNFs for SFC $\pi_l$
$E_l$	The set of virtual links for SFC $\pi_l$
$a_{f_i}^v$	The binary parameter that indicates whether VNF $f_i$ is deployed at node $v$
$c_{f_i}^v$	The packet processing rate of node $v$ for VNF $f_i$
$B_t^{(v,u)}$	The link transmission rate of link $(v,u)$ at slot $t$
$f_{i,l}^{(j)}$	The $j^{th}$ VNF in SFC $\pi_l$ and its function type is $i$
$\lambda_{l,t}$	The packet arrival rate of flow $l$ at slot $t$
$x_t^v(f_{i,l}^j)$	The binary variable that indicates whether the VNF $f_{i,l}^j$ is mapped to physical node $v$ at slot $t$
$y_t^{(v,u)}(f_{i,l}^j, f_{i',l}^{j+1})$	The binary variable that indicates whether the virtual link $(f_{i,l}^j, f_{i',l}^{j+1})$ is mapped to physical link $(v,u)$ at slot $t$
$D$	The end-to-end delay requirement of the service
$D_{l,t}$	The end-to-end delay of flow $l$ at slot $t$
$\phi_{1,t}$	The normalized computing resource provisioning cost at slot $t$
$\phi_{2,t}$	The normalized communication resource provisioning cost at slot $t$
$\phi_{3,t}$	The VNF migration cost at slot $t$
$R_t$	The service performance gain at slot $t$
$\chi_t$	The total network profit at slot $t$

TABLE II: SUMMARY OF IMPORTANT ACRONYMS

Acronyms	Descriptions
ANRPC	Accumulative network resource provisioning cost
ASPG	Accumulative service performance gain
CDPI	Control-data-plane interface
DDVSC	DRL-based dynamic VNF selection and chaining
DRL	Deep reinforcement learning
DVSC	Dynamic VNF selection and chaining
ISL	Inter-satellite link
NOC	Network operation controller
SFC	Service function chain
SGIN	Satellite-ground integrated network
SGL	Satellite-to-ground link
SR	Sharing ratio
VNF	Virtual network function
VSAT	Very-small-aperture-terminals
VSCP	VNF selection and chaining policies

#### A. Network Model

In an SGIN, a VSAT can communicate with an ES through a satellite network segment. The accessing satellite from a VSAT is determined according to the longest link duration rule [28] and a traffic routing path is selected with the smallest hop-count [19]. In this regard, we focus on the satellites which are used to cover the ground network segment. Satellite movements result in a time-varying SGIN topology. Based on the VN approach [23], the satellite network segment is described as a virtual lattice grid consisting of interconnecting VNs. The size of a virtual lattice grid is determined according to the communication coverage of the satellite network segment. The satellites covering the ground network segment are identified as VNs in the virtual lattice grid. The actual satellite that is

associated with a particular VN keeps changing, and the virtual lattice grid itself remains static to the ground network segment.

In addition, the NOC is deployed at the ground network segment which can steer traffic, activate/deactivate VNF instances, and migrate VNFs. For the satellites in the virtual lattice grid, the satellites periodically (e.g., every second) report their updated resource states (i.e., usage of links and satellites) and locations to the ESs, and the ESs forward this information to the NOC [29]. Furthermore, when the VNs are assigned to other satellites or NOC changes VSOPs, flow tables for traffic routing among satellites are updated and forwarded to the satellites via ESs. Meanwhile, NOC can monitor the resource states of the ground network segment (i.e., usage of links, BSs and routers) and deliver flow tables for traffic routing among ground nodes to routers and BSs by control-data-plane interfaces (CDPIs). Based on Fig. 1, the service establishment and maintenance process is described as follows: 1) users order the service from the business and operation support system (BOSS), which analyzes the user's service requirements (e.g., required network functions, delay requirements, etc.) to form SFCs; 2) SFCs are submitted to the NOC via application programming interfaces (APIs), and the NOC deploys the required VNFs to routers and satellites based on the monitored the resource states of the SGIN, the network topology, and the network function type; 3) the NOC calculates flow tables for VNF selection and virtual link mapping, and delivers flow tables to satellites and routers via CDPIs.

Consider time  $\mathcal{T}$  is divided into  $T$  time slots indexed by  $1, 2, \dots, T$  with fixed and identical length. Within a time slot  $t$ , the connections among nodes are assumed unchanged and the network topology is static. The SGIN at time slot  $t$  is modeled as an undirected graph  $\mathcal{G}_t = (V_t, E_t)$ , where  $V_t$  represents the set of physical nodes and  $E_t$  represents the set of physical links interconnecting the nodes. We have  $V_t = V^G \cup V_t^S$ , where  $V^G$  is the set of ground nodes of SGIN including BSs, routers, VSATs and ESs, and  $V_t^S$  is the set of satellites in the virtual lattice grid at  $t$ . We also have  $E_t = E^G \cup E_t^{SG} \cup E_t^S$ , where  $E^G$  is the set of wired links connecting ground nodes,  $E_t^S$  are the ISLs among satellites at  $t$ , and  $E_t^{SG}$  is the set of SGLs at  $t$ . Define the link pointing from node  $v$  to node  $u$  as  $(v, u) \in E_t$ , where we have  $v, u \in V_t$ . Due to satellite mobility,  $V_t^S$ ,  $E_t^{SG}$  and  $E_t^S$  can change over consecutive time slots.

### B. Service Model

We consider an SGIN providing an immersive online education service which requires four VNFs [2], [30]: *movement capture*, *information fusion*, *logic programming*, and *image rendering*. These VNFs are deployed at the heterogeneous network nodes in the SGIN. Specifically, *movement capture* function is responsible for capturing users' movements, emotions, and facial expressions via immersion devices (e.g., XR headsets and motion sensors), which is deployed at the user terminals. *Information fusion* function is a lightweight network function deployed at the satellites and ground nodes and responsible for collecting cognitive data from immersion devices. *Logic programming* function is responsible for human-environment interaction computing that guarantees the virtual

environment to react to users' actions. *Logic programming* function needs to consume numerous computing resources, which is suitable to be deployed at the ground nodes. *Image rendering* function is responsible for modeling and refreshing the virtual environment, which needs to be deployed at the ground nodes with a high-performance processor.

The set of VNFs required by the education service is denoted by  $\mathcal{F} = \{f_1, f_2, \dots, f_n\}$ , where  $f_i$  ( $i = 1, 2, \dots, n$ ) denotes a function of type  $i$ . The binary parameter  $a_{f_i}^v$  indicates whether VNF  $f_i \in \mathcal{F}$  is deployed at node  $v \in V_t$ , given by

$$a_{f_i}^v = \begin{cases} 1, & \text{VNF } f_i \text{ is deployed at node } v \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

To simplify resource management, we consider cross server pipelined SFC model to deploy VNFs where a node can only host at most one type of VNF [31], i.e.,

$$\sum_{f_i \in \mathcal{F}} a_{f_i}^v \leq 1, \quad \forall v \in V_t. \quad (2)$$

The VNF deployment scheme can be determined according to node computing resources, link bandwidth resources, network topology, and network function types [32]. Moreover, to ensure continuous service, consider that *information fusion* as a lightweight and basic function is deployed at all satellites. There is an example of VNF deployment for an immersive online education service, as shown in Fig. 1, where VNF1 is the *movement capture* function, VNF 2 is the *information fusion* function, VNF 3 is the *logic programming* function and VNF 4 is the *image rendering* function.

The aggregate traffic of end users accessing to the same BS or VSAT forms a flow. Suppose that there are  $L$  flows in the SGIN, denoted by  $\mathcal{L} = \{1, 2, \dots, L\}$ . Suppose flow  $l \in \mathcal{L}$  traverse SFC  $\pi_l$ , represented by a sequence of VNFs,  $f_{i,l}^j \in \mathcal{F}_l$ , in a predefined order, where  $\mathcal{F}_l$  indicates the set of VNFs for  $\pi_l$ ,  $j$  indicates the sequence number of a specific VNF in  $\pi_l$ ,  $i$  indicates the function type. We call the logical abstraction of all mapped physical paths between two consecutive VNFs of a flow as a *virtual link* and use  $E_l = (f_{i,l}^j, f_{i',l}^{j+1})$  to represent the virtual link from VNF  $f_{i,l}^j$  to VNF  $f_{i',l}^{j+1}$  in  $\pi_l$ .

We consider an immersive online education service, with E2E delay requirement  $D$  in ms and average packet size  $\sigma$  in bits. The packet arrival process of flow  $l$  over sequential time slots is modeled as a Poisson process with rate parameter  $\lambda_{l,t}$  in slot  $t$ , and we assume that the packet arrival processes of flows are independent and identically distributed (i.i.d).

### C. Computation and Communication Model

The processing capacity  $p^v$  of node  $v \in V_t$  is denoted by its maximum CPU processing rate in cycles/s. We define the computation intensity of VNF  $f_i$  as  $\zeta_{f_i}$  which is the number of CPU cycles to process one bit of information for VNF  $f_i$ . The packet processing rate (in packets/s) at node  $v$  for VNF  $f_i$  is

$$c_{f_i}^v = \frac{p^v}{\zeta_{f_i} \sigma} \quad (3)$$

where  $c_{f_i}^v$  represents the number of packets that can be processed by node  $v$  per second for VNF  $f_i$ . The packet processing delay on node  $v$  for VNF  $f_i$  is given by

$$D_{f_i}^v = \frac{1}{c_{f_i}^v}. \quad (4)$$

The link transmission rate (in bits/s) of  $(v, u) \in E_t$  at slot  $t$  is denoted by  $B_t^{(v,u)}$ , and the transmission delay per packet over link  $(v, u)$  is

$$D_t^{(v,u)} = \frac{\sigma}{B_t^{(v,u)}}. \quad (5)$$

For a link  $(v, u) \in E_t^G$ , i.e., a wired link, the link transmission rate (in bits/s)  $B_t^{(v,u)}$  is constant. For a link  $(v, u) \in E_t^{SG} \cup E_t^S$ , i.e., an SGL or an ISL, the link transmission rate (in bits/s) is given by

$$B_t^{(v,u)} = b \times \eta_t \quad (6)$$

where  $b$  is the communication frequency (in Hz) of link  $(v, u)$  and  $\eta_t$  is the spectral efficiency (in bps/Hz) at slot  $t$ . According to the DVB-S2X standard [33],  $\eta_t$  can be determined according to communication signal to noise ratio (SNR). The SNR at slot  $t$  is given by [22], [34]

$$SNR_t = EIRP + C - A_{t,1} - A_2 - \kappa - b \quad (7)$$

where  $EIRP$  is the effective isotropic radiated power of a sender (i.e., VSAT or satellite),  $C$  is the quality factor of a receiver (satellite or ES),  $A_1$  is the free space loss,  $A_2$  is other signal transmission losses,  $\kappa$  is the Boltzmann's constant (i.e.,  $1.38 \times 10^{-23}$  kB), and  $b$  is the communication bandwidth of link  $(v, u)$ . Specifically,  $A_1$  is given by

$$A_1 = \varsigma_1 + 10\varsigma_2 \times \log_{10} d_t^{(v,u)} + 10\varsigma_3 \times \log_{10} M \quad (8)$$

where  $\varsigma_1, \varsigma_2$  and  $\varsigma_3$  are constants obtained by the actual measurement in the specific environment,  $d_t^{(v,u)}$  is the communication distance (in km) of link  $(v, u)$  at slot  $t$ , and  $M$  is the communication center frequency (in GHz). For an SGL,  $A_2$  mainly includes atmospheric absorption loss and rain attenuation, and for an ISL,  $A_2$  mainly includes beam pointing loss. In clear weather conditions,  $A_2$  can be assumed to be a small positive constant.

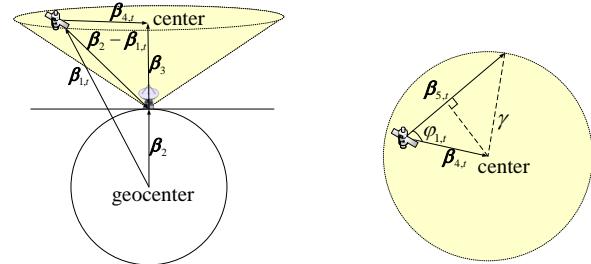
#### D. Satellite-to-Ground Link Handover Mechanism

An ES is located on the Earth's surface and a satellite flies in space, as shown in Fig. 2(a). When the satellite is visible to the ES, the SGL can be established and maintained. Otherwise, the ES hands the communication link over to another satellite according to the longest link duration rule [28].

The duration of SGLs is estimated based on satellite trajectories. Assume that the satellite moves in a circular cross-section of the observation area of the ES, as shown in Fig. 2(a). The distance between the satellite and the center of the circular cross-section  $|\beta_{4,t}|$  is given by

$$|\beta_{4,t}| = |\beta_2 - \beta_{1,t} + \beta_3| \quad (9)$$

where  $\beta_{1,t}$  and  $\beta_2$ , respectively, are the position vectors of the satellite and the ES,  $\beta_3$  is the central visual vector of the ES,



(a) SGL visibility analysis  
(b) SGL duration analysis

Fig. 2: An illustration on SGL handover mechanism.

which points from the ES to the center of the circular cross-section of the observation area of the ES. Due to the limited viewing field of the ES, the satellite trajectory vector  $\beta_{5,t}$  is assumed to be straight, as shown in Fig. 2(b). The length of satellite trajectory  $|\beta_{5,t}|$  depends on the radius of the circular cross-section  $\gamma$ , the vector  $\beta_{4,t}$  and the angle  $\varphi_{1,t}$  formed by the vector  $\beta_{4,t}$  and the satellite trajectory vector  $\beta_{5,t}$ , given by

$$|\beta_{5,t}| = \left( |\beta_{4,t}| \cos \varphi_{1,t} + \sqrt{\gamma^2 - |\beta_{4,t}|^2 \sin^2 \varphi_{1,t}} \right). \quad (10)$$

The duration of SGL  $T_{SG}$  is given by

$$T_{SG} = \frac{1}{\vartheta} \times |\beta_{5,t}| \quad (11)$$

where  $\vartheta$  is the satellite velocity.

#### E. Service Request

To complete service requests, a flow  $l$  starting from the user terminal needs to sequentially traverse the VNFs in SFC  $\pi_l$  to reach the destination node, which consists of VNF selection and virtual link mapping. Let binary variable  $x_t^v(f_{i,l}^j)$  represent the mapping of VNF  $f_{i,l}^j$  to node  $v$  at slot  $t$ . We have

$$x_t^v(f_{i,l}^j) = \begin{cases} 1, & \text{VNF } f_{i,l}^j \text{ of flow } l \text{ is mapped} \\ & \text{to node } v \text{ at slot } t \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

The VNF  $f_{i,l}^j$  can map to only nodes that are installed with the corresponding VNF instance, represented as

$$a_{f_i}^v \geq x_t^v(f_{i,l}^j), \quad \forall v \in V_t, \forall f_{i,l}^j \in \mathcal{F}_l, \forall t \in \mathcal{T}. \quad (13)$$

If VNF  $f_i$  is not deployed at node  $v$ , we have  $a_{f_i}^v = 0$ .

The VNF  $f_{i,l}^j$  can only be mapped to one physical node. That is

$$\sum_{v \in V_t} x_t^v(f_{i,l}^j) = 1, \quad \forall f_{i,l}^j \in \mathcal{F}_l, \forall t \in \mathcal{T}. \quad (14)$$

For an immersive online education service, the first VNF *movement capture* is mapped to the user terminal which is the source node of flow  $l$  denoted as  $v_{l,s}$ , given by

$$x_t^{v_{l,s}}(f_{i,l}^1) = 1, \quad \forall l \in \mathcal{L}. \quad (15)$$

If flow  $l$  uses the second VNF *information fusion* deployed at satellites, we consider the VNF *information fusion* is mapped to the accessing satellite of flow  $l$  denoted as  $v_{l,a}$ , and the following relationship holds

$$x_t^{v_{l,a}}(f_{i,l}^2) + \sum_{v \in E^G} x_t^v(f_{i,l}^2) = 1, \forall l \in \mathcal{L}. \quad (16)$$

The refreshed virtual environment can be obtained through the last VNF *Image rendering*, which is mapped to the destination node of flow  $l$  denoted as  $v_{l,d}$ , given by

$$x_t^{v_{l,d}}(f_{i,l}^n) = 1, \forall l \in \mathcal{L}. \quad (17)$$

Let binary variable  $y_t^{(v,u)}(f_{i,l}^j, f_{i',l}^{j+1})$  indicate the mapping of virtual link  $(f_{i,l}^j, f_{i',l}^{j+1})$  to link  $(v, u)$  at slot  $t$ . We have

$$y_t^{(v,u)}(f_{i,l}^j, f_{i',l}^{j+1}) = \begin{cases} 1, & \text{virtual link } (f_{i,l}^j, f_{i',l}^{j+1}) \text{ of flow } l \\ & \text{is mapped to link } (v, u) \text{ at slot } t \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

#### F. Cost Model

The network resource provisioning cost includes communication resource provisioning cost and computing resource provisioning cost. At slot  $t$ ,  $\phi_t^v$  is a binary variable with  $\phi_t^v = 1$  if at least one VNF is mapped to node  $v$ , and  $\phi_t^v = 0$  otherwise. The binary variable  $\phi_t^v$  can be expressed as

$$\phi_t^v = \mathbb{I} \left[ \sum_{l \in \mathcal{L}} \sum_{f_{i,l}^j \in \mathcal{F}_l} x_t^v(f_{i,l}^j) \right] \quad (19)$$

where  $\mathbb{I}(\cdot)$  is the indicator function, given by

$$\mathbb{I}(x) = \begin{cases} 1, & x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

Define the computing resource provisioning cost as the computing resource utilization in the network, given by

$$\phi_{1,t} = \frac{\sum_{v \in V_t} \phi_t^v p^v}{\sum_{v \in V_t} p^v} \quad (21)$$

where  $\sum_{v \in V_t} \phi_t^v p^v$  is the total computing resource usage by all VNFs, and  $\sum_{v \in V_t} p^v$  is the total computing resource capacity of all nodes.

At slot  $t$ ,  $\phi_t^{(v,u)}$  is a binary variable with  $\phi_t^{(v,u)} = 1$  if at least one virtual link is mapped to link  $(v, u)$ , and  $\phi_t^{(v,u)} = 0$  otherwise. The binary variable  $\phi_t^{(v,u)}$  can be expressed as

$$\phi_t^{(v,u)} = \mathbb{I} \left[ \sum_{l \in \mathcal{L}} \sum_{(f_{i,l}^j, f_{i',l}^{j+1}) \in E_l} y_t^{(v,u)}(f_{i,l}^j, f_{i',l}^{j+1}) \right]. \quad (22)$$

Define the communication resource provisioning cost as the communication resource utilization in the network, given by

$$\phi_{2,t} = \frac{\sum_{(v,u) \in E^G} \phi_t^{(v,u)} B^{(v,u)} + \sum_{(v,u) \in E_t^S \cup E_t^{SG}} \phi_t^{(v,u)} B_t^{(v,u)}}{\sum_{(v,u) \in E^G} B^{(v,u)} + \sum_{(v,u) \in E_t^S \cup E_t^{SG}} B_t^{(v,u)}} \quad (23)$$

where  $\sum_{(v,u) \in E^G} \phi_t^{(v,u)} B^{(v,u)} + \sum_{(v,u) \in E_t^S \cup E_t^{SG}} \phi_t^{(v,u)} B_t^{(v,u)}$  is the total communication resource usage by all virtual links,  $\sum_{(v,u) \in E^G} B^{(v,u)} + \sum_{(v,u) \in E_t^S \cup E_t^{SG}} B_t^{(v,u)}$  is the total communication resource of all links.

We also consider the VNF migration cost. For flow  $l$ , let integer variable  $z_{l,t}$  indicate the number of VNFs that are remapped at slot  $t$ . The integer variable  $z_{l,t}$  can be expressed as

$$z_{l,t} = \sum_{f_{i,l}^j \in \mathcal{F}_l} \sum_{v \in V_t} [1 - x_{t-1}^v(f_{i,l}^j)] x_t^v(f_{i,l}^j) \quad (24)$$

where  $[1 - x_{t-1}^v(f_{i,l}^j)] x_t^v(f_{i,l}^j)$  is equal to 1 if VNF  $f_{i,l}^j$  is remapped to node  $v$  at slot  $t$ , and equal to 0 otherwise. Different from ground networks, there are two reasons for VNF migration. First, NOC determines a new VSCP that leads to a change in VNF selection decision; and second satellite movements lead to a change in VNF selection decision to adapt to the time-varying network topology. Define the VNF migrations cost as the number of VNF migrations, given by

$$\phi_{3,t} = \sum_{l \in \mathcal{L}} z_{l,t}. \quad (25)$$

## IV. PROBLEM FORMULATION

**Problem statement:** Consider an SDN/NFV-enabled SGIN where VNFs can be flexibly deployed at both the satellites and ground nodes. A VNF selection and chaining problem is to determine the VNF mapping decision and virtual link mapping decision for flow  $l$ ,  $l \in \mathcal{L}$  at each time slot, based on the previous mapping decision and the current network state. The objective is to balance the network resource provisioning and VNF migration costs with service performance gain to maximize the long-term network profit, taking into consideration service request constraints, capacity constraints, delay constraints, and flow conservation constraints.

### A. Capacity Constraints

For a node  $v \in V_t$ , the traffic rate of flows  $\sum_{l \in \mathcal{L}} \sum_{f_{i,l}^j \in \mathcal{F}_l} \lambda_{l,t} x_t^v(f_{i,l}^j)$  should be upper bounded due to the limited processing capacity, given by

$$\sum_{l \in \mathcal{L}} \sum_{f_{i,l}^j \in \mathcal{F}_l} \lambda_{l,t} x_t^v(f_{i,l}^j) < c_{f_i}^v. \quad (26)$$

For a link  $(v, u) \in E_t$ , the traffic rate of flows  $\sum_{l \in \mathcal{L}} \sum_{(f_{i,l}^j, f_{i',l}^{j+1}) \in E_l} \lambda_{l,t} y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1})$  should be upper bounded due to the limited link transmission rate, given by

$$\sum_{l \in \mathcal{L}} \sum_{(f_{i,l}^j, f_{i',l}^{j+1}) \in E_l} \lambda_{l,t} y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1}) < B_t^{(v,u)}. \quad (27)$$

### B. Delay Constraints

Let  $D_{l,t}$  represent the E2E delay of flow  $l$ , which should not exceed the E2E delay requirement  $D$ , given by

$$D_{l,t} \leq D, \forall l \in \mathcal{L}, \forall t \in \mathcal{T}. \quad (28)$$

For E2E delay  $D_{l,t}$ , we consider processing delay, queuing delay, transmission delay and propagation delay. With Poisson packet arrival and deterministic packet processing time, the average packet delay can be found approximately using an M/D/1 queueing model, which has been proven to be a more accurate upper bound than that using a G/D/1 queueing model [35].

Let  $D_{q,t}^v$  denote queuing delay before a packet is processed by node  $v$ .  $D_{q,t}^v$  can be expressed as

$$D_{q,t}^v = \frac{K_{q,t}^v}{\sum_{l \in \mathcal{L}} \sum_{f_{i,l}^j \in \mathcal{F}_l} \lambda_{l,t} x_t^v (f_{i,l}^j)} \quad (29)$$

where  $K_{q,t}^v$  represents the packet queue length at node  $v$  at slot  $t$ , given by

$$K_{q,t}^v = \frac{(\rho_t^v)^2}{2(1 - \rho_t^v)}. \quad (30)$$

In (30),  $\rho_t^v$  represents the processing intensity of node  $v$ , given by

$$\rho_t^v = \frac{\sum_{l \in \mathcal{L}} \sum_{f_{i,l}^j \in \mathcal{F}_l} \lambda_{l,t} x_t^v (f_{i,l}^j)}{c_{f_i}^v}. \quad (31)$$

We have  $\rho_t^v < 1$  to keep queue length stable.

For flow  $l$ , let  $D_{1,l,t}$  denote the total packet processing and queuing delay at all processing nodes, which can be expressed as

$$D_{1,l,t} = \sum_{f_{i,l}^j \in \mathcal{F}_l} \sum_{v \in V_t} (D_{f_i}^v + D_{q,t}^v) x_t^v (f_{i,l}^j) \quad (32)$$

where  $D_{f_i}^v$  is given by (4).

Let  $D_{2,l,t}$  denote the total packet transmission delay in all transmitting links, which can be expressed as

$$D_{2,l,t} = \sum_{(f_{i,l}^j, f_{i',l}^{j+1}) \in E_l} \sum_{(v,u) \in E_t} D_t^{(v,u)} y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1}) \quad (33)$$

where  $D_t^{(v,u)}$  is given by (5).

Let  $D_{3,l,t}$  denote the total propagation delay in SGLs and ISLs, which can be expressed as

$$D_{3,l,t} = \sum_{(f_{i,l}^j, f_{i',l}^{j+1}) \in E_l} \sum_{(v,u) \in E_t^{SG} \cup E_t^S} \frac{d_t^{(v,u)} y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1})}{\nu} \quad (34)$$

where  $d_t^{(v,u)}$  is the communication distance of link  $(v, u)$  at slot  $t$ ,  $\nu$  is speed of signal propagation.

The end-to-end delay  $D_{l,t}$  can be expressed as

$$D_{l,t} = D_{1,l,t} + D_{2,l,t} + D_{3,l,t}. \quad (35)$$

### C. Flow Conservation Constraints

For flow  $l$ , the flow conservation is given by (37). In (37), if  $v = v_{l,s}$ , i.e.,  $v$  is the source node, the number of flows out is one more than the number of flows in. If  $v = v_{l,d}$ , i.e.,  $v$  is the destination node, the number of flows in is one more than the number of flows out. If  $v \in V_t \setminus \{v_{l,s}, v_d\}$ , the number of flows in must be equal to the number of flows out.

### D. Optimization Problem

Let  $R_t$  denote the service performance gain of all flows. For an immersive online education service, E2E delay can impact the user experience. Thus,  $R_t$  is expressed as

$$R_t = \sum_{l \in \mathcal{L}} \lambda_{l,t} (1 - D_{l,t}/D) \quad (38)$$

where  $D$  is the E2E delay requirement of the service,  $\lambda_{l,t} (1 - D_{l,t}/D)$  is the service performance gain of flow  $l$  at slot  $t$ .

From the perspective of the service provider, the objective is to maximize the service performance gain. Meanwhile, physical resource cost and VNF migration cost should be minimized. Thus, the network profit at slot  $t$  is designed as follows

$$\chi_t = -\alpha_1 \phi_{1,t} - \alpha_2 \phi_{2,t} - \alpha_3 \phi_{3,t} + \alpha_4 R_t \quad (39)$$

where  $\alpha_1, \alpha_2, \alpha_3$  and  $\alpha_4$  are weights.

To maximize the long-term network profit, the dynamic virtual network function selection and chaining problem in an SGIN can be formulated as

$$(P1) : \begin{aligned} & \max_{\mathbf{x}_t, \mathbf{y}_t} \sum_{t=1}^T \chi_t \\ & \text{s.t. } (13) - (17), (26) - (28), (37) \\ & \quad x_t^v (f_{i,l}^j), y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1}) \in \{0, 1\}. \end{aligned} \quad (40)$$

In (P1), the vector  $\mathbf{x}_t$  represents a sequence of VNF selection decisions at slot  $t$ , which can be expressed as

$$\mathbf{x}_t = [x_t^v (f_{i,l}^j)]_{|l \in \mathcal{L}, f_{i,l}^j \in \mathcal{F}_l, v \in V_t}. \quad (41)$$

The vector  $\mathbf{y}_t$  represents a sequence of virtual link mapping decisions at slot  $t$ , which can be expressed as

$$\mathbf{y}_t = [y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1})]_{|l \in \mathcal{L}, (f_{i,l}^j, f_{i',l}^{j+1}) \in E_l, (v,u) \in E_t}. \quad (42)$$

The proposed VNF selection and chaining problem can be reduced to the capacitated plant location problem with single source constraints, which is NP-hard [36]. We can obtain the exact solutions only when the problem size is small, using the conventional optimization-based schemes. However, with an increase of problem size, it becomes computationally complex to solve for the optimal VNF selection and

$$\sum_{u \in V_t} y_t^{(v,u)} (f_{i,l}^j, f_{i',l}^{j+1}) - \sum_{u \in V_t} y_t^{(u,v)} (f_{i,l}^j, f_{i',l}^{j+1}) = \begin{cases} 1, & v = v_{l,s} \\ -1, & v = v_{l,d} \\ 0, & \text{else} \end{cases}, \forall l \in \mathcal{L}, \forall (f_{i,l}^j, f_{i',l}^{j+1}) \in E_l, \forall v \in V_t. \quad (37)$$

chaining pattern using the conventional optimization-based methods (e.g., dynamic programming). Moreover, to capture the relation between network state and VSCP set over time, we describe the VNF selection and chaining problem as an MDP formulation. A DRL-based algorithm can be a less computational complex option to solve the MDP problem with large problem size (e.g., state spaces), which can also capture the network dynamics with state-action transitions to optimize a long-term network profit. After the learning model training is converged, a DRL-based algorithm can provide an efficient SFC mapping solution across time slots through online implementation, which is more computationally tractable than solving per-time-slot optimization.

### E. Problem Analysis

A VSCP set represents a joint VNF selection and chaining policy of multiple flows. To save computing resources, VNF instances can be shared among multiple flows. The SR is defined as the ratio of computing resources that are shared to the total computing resources to evaluate the level of computing resource sharing of a VSCP set, given by

$$SR = \frac{H_2}{H_1} \quad (43)$$

where  $H_1$  is the total computing resources and  $H_2$  is the shared computing resources by flows.

Although computing resources can be saved through VNF instances sharing, service performance may reduce. This is because VNF instances shared by multiple flows will lead to longer queue latency, and the route of flows might be longer to reach the shared VNF instances, which increases the transmission delay. Thus, the SR can reflect the tradeoff relationship between network resource provisioning cost and service performance gain.

## V. DRL-BASED DYNAMIC VNF SELECTION AND CHAINING ALGORITHM FOR SGINS

We propose a DRL-based dynamic VNF selection and chaining (DDVSC) algorithm for SGINs, which chooses a VSCP set based on the network resources, network traffic load, network topology and the previous VSCP set. The objective is to balance the network resource provisioning and VNF migration costs with service performance gain to maximize the long-term network profit.

The optimization period  $\mathcal{T}$  is divided into  $K$  time intervals indexed by  $1, 2, \dots, K$ , and each time interval contains  $N$  time slots, where  $k^{\text{th}}$  interval is indexed by  $[N(k-1)+1, N(k-1)+2, \dots, Nk]$ . The NOC observes the network state in each time slot and determines a VSCP set at the start point of each time interval. The current decision of the NOC is only dependent on the latest state, so the VNF

selection and chaining decision process can be formulated as an MDP. The key elements in MDP are listed as follows:

**State Space:** Let  $\lambda_{\mathcal{L}}^{(k)} = [\lambda_l^{(k)}]_{l \in \mathcal{L}}$  denote the forecast traffic rate of flows in interval  $k$  based on the traffic rate of flows in interval  $k-1$  [31]. The state  $s^{(k)}$  includes traffic rate of flows  $\lambda_{\mathcal{L}}^{(k)}$ , network state parameter  $\varpi^{(k)}$  and the previous VSCP set  $w^{(k-1)}$ , i.e.,

$$s^{(k)} = [\lambda_{\mathcal{L}}^{(k)}, \varpi^{(k)}, w^{(k-1)}] \quad (44)$$

where  $\varpi^{(k)}$  includes network topology parameter  $\mathcal{G}_{[n(k-1)+1]}$ , VNF deployment information  $a_{f_i}^v$ , packet processing rate of node  $c_{f_i}^v$ , and transmission rate of link  $B_t^{(v,u)}$ .

**Action Space:** The action space  $\mathcal{W}$  consists of a sequence of actions,  $w_m \in \mathcal{W}$ , where  $w_m$  is a VSCP set that represents a joint VNF selection and chaining policy of multiple flows:

$$w_m = [p_1^{(m)}, p_2^{(m)}, \dots, p_L^{(m)}] \quad (45)$$

where  $m$  indicates the sequence number of a specific action in action space and  $p_l^{(m)}$  indicates the VSCP of flow  $l$  in set  $w_m$ . The VSCP  $p_l^{(m)}$  can be expressed as

$$p_l^{(m)} = [x_{l,S}^{(m)}, x_{l,G}^{(m)}, y_{l,S}^{(m)}, y_{l,G}^{(m)}] \quad (46)$$

where  $x_{l,S}^{(m)}$  and  $x_{l,G}^{(m)}$  represent the VNF selection policy in the satellite network segment and ground network segment, and  $y_{l,S}^{(m)}$  and  $y_{l,G}^{(m)}$  represent the link mapping policy in the satellite network segment and ground network segment.  $x_{l,S}^{(m)}$  is a binary variable with  $x_{l,S}^{(m)} = 1$  if the VNF instance installed at the accessing satellite of flow  $l$  is used, and  $x_{l,S}^{(m)} = 0$  otherwise. Due to satellite movements,  $x_{l,S}^{(m)}$  can make different VNF selection decisions across time slots.  $x_{l,G}^{(m)}$  can be expressed as

$$x_{l,G}^{(m)} = [x^v(f_{i,l}^j)]_{|f_{i,l}^j \in \mathcal{F}_l, v \in V^G} \quad (47)$$

where  $x^v(f_{i,l}^j)$  is a binary variable with  $x^v(f_{i,l}^j) = 1$  if VNF  $f_{i,l}^j$  is mapped to the ground node  $v$ , and  $x^v(f_{i,l}^j) = 0$  otherwise.  $y_{l,S}^{(m)}$  can be expressed as

$$y_{l,S}^{(m)} = \begin{cases} [v_{l,s}, v_{l,ES}], & \text{flow } l \text{ accesses to the satellite } v_{l,s} \text{ and is routed to the node } v_{l,ES} \\ [0, 0], & \text{otherwise} \end{cases} \quad (48)$$

where  $v_{l,s}$  is the source node, which is equipped with a VSAT to connect to satellites, and  $v_{l,ES}$  is the ES that routes the traffic from the satellite network to the ground network. If flow  $l$  does not traverse the satellite network segment,  $y_{l,S}^{(m)} = [0, 0]$ . The path of flow  $l$  in the satellite network segment is searched according to the smallest hop-count [19].

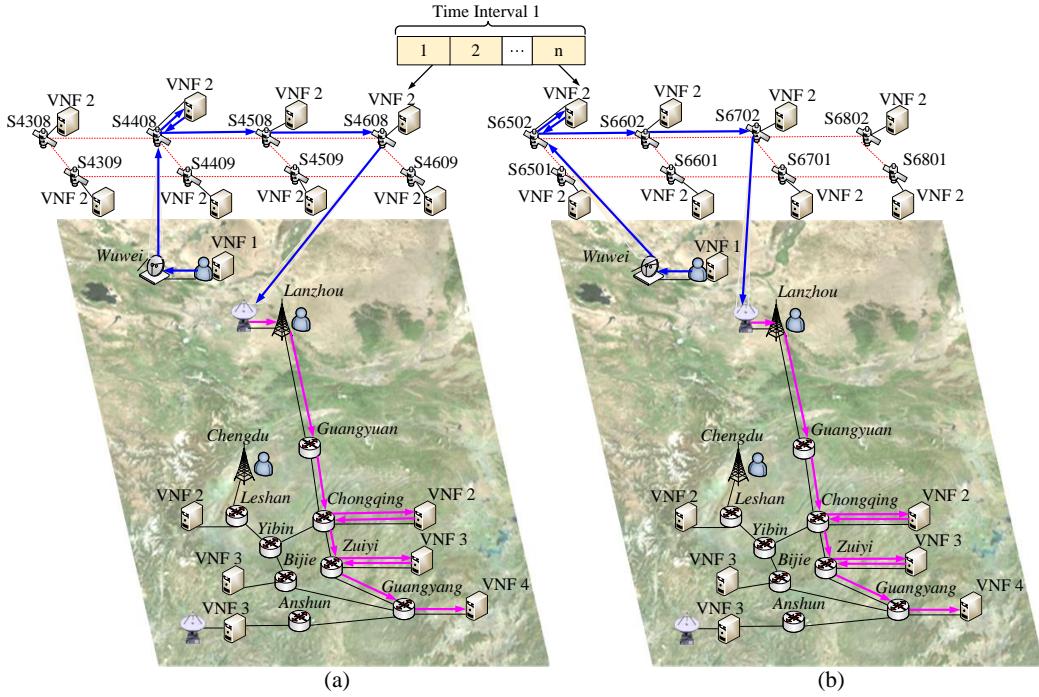


Fig. 3: A VSCP across time slots to adapt to the time-varying network topology.

Due to satellite movements,  $\mathbf{y}_{l,S}^{(m)}$  can make different virtual link mapping decisions across slots.  $\mathbf{y}_{l,G}^{(m)}$  can be expressed as

$$\mathbf{y}_{l,G}^{(m)} = \left[ y^{(v,u)} \left( f_{i,l}^j, f_{i',l}^{j+1} \right) \right]_{|(v,u) \in E^G, \left( f_{i,l}^j, f_{i',l}^{j+1} \right) \in E_l} \quad (49)$$

where  $y^{(v,u)} \left( f_{i,l}^j, f_{i',l}^{j+1} \right)$  is a binary variable with  $y^{(v,u)} \left( f_{i,l}^j, f_{i',l}^{j+1} \right) = 1$  if virtual link  $\left( f_{i,l}^j, f_{i',l}^{j+1} \right)$  is mapped to the ground link  $(v, u)$ , and  $y^{(v,u)} \left( f_{i,l}^j, f_{i',l}^{j+1} \right) = 0$  otherwise.

There is an example to describe how a VSCP makes a sequence of VNF selection and changes decisions across time slots, as shown in Fig. 3. In the ground network segment in Western China, a user in *Wuwei* city accesses services via a VSAT. The satellite network segment includes 8 satellites to cover the ground network segment and VNF 2 is deployed at each satellite. This VSCP sets  $x_{l,S}^{(m)} = 1$  and  $\mathbf{y}_{l,S}^{(m)} = [v_1, v_2]$ , where  $v_1$  is the user terminal at *Wuwei* city and  $v_2$  is the ES at *Lanzhou* city.  $\mathbf{x}_{l,G}^{(m)}$  and  $\mathbf{y}_{l,G}^{(m)}$  determine a VNF selection decision and routing path in the ground network segment, and they remain unchanged in a time interval (e.g., the pink solid arrow in Fig. 3(a) and Fig. 3(b)). At time slot 1,  $x_{l,S}^{(m)}$  lets VNF 2 map to satellite S4408, i.e., the 8<sup>th</sup> satellite located at the 44<sup>th</sup> orbit, and  $\mathbf{y}_{l,S}^{(m)}$  lets virtual links (VNF 1, VNF 2) and (VNF 2, VNF 3) map to links with the blue solid arrow to connect the VSAT at *Wuwei* city to the ES at *Lanzhou* city, as shown in Fig. 3(a). Due to satellite movements, satellite S4408 can not communicate with the VSAT at time slot  $n$ . Thus,  $x_{l,S}^{(m)}$  lets VNF 2 remap to satellite S6502, and  $\mathbf{y}_{l,S}^{(m)}$  lets virtual links (VNF 1, VNF 2) and (VNF 2, VNF 3) remap to other physical links with the blue solid arrow in Fig. 3(b).

With the number of traffic flow increasing, the action space suffers from the curse of dimensionality issue. To reduce the action space, the following steps are performed: (1) unitizing the  $k$ -means clustering method to cluster the historical records of the network load; and (2) selecting a VSCP set for each cluster to build action space  $\mathcal{W}$ .

The historical records of the network load in optimization period  $\mathcal{T}$  is denoted by  $\lambda_{\mathcal{L},\mathcal{T}}^H = \{\lambda_{\mathcal{L},1}^H, \lambda_{\mathcal{L},2}^H, \dots, \lambda_{\mathcal{L},T}^H\}$ , where  $\lambda_{\mathcal{L},t}^H \in \lambda_{\mathcal{L},\mathcal{T}}^H$  represents historical record of the network load at slot  $t$ , given by

$$\lambda_{\mathcal{L},t}^H = (\lambda_{1,t}^H, \lambda_{2,t}^H, \dots, \lambda_{L,t}^H) \quad (50)$$

where  $\lambda_{l,t}^H (l = 1, 2, \dots, L)$  represents the historical record of traffic rate of flow  $l$  at slot  $t$ .

The clustering result of the historical records of the network load is expressed as  $\mathcal{H} = (h_1, h_2, \dots, h_{|\mathcal{W}|})$ , where  $h_j, h_j \in \mathcal{H}$  is a cluster consisting of several historical records of the network load at different time slots. The cluster center of  $h_j$  is denoted as  $\lambda_{\mathcal{L},j}^C$ , which can be expressed as

$$\lambda_{\mathcal{L},j}^C = (\lambda_{1,j}^C, \lambda_{2,j}^C, \dots, \lambda_{L,j}^C) \quad (51)$$

where  $\lambda_{l,j}^C \in \lambda_{\mathcal{L},j}^C$  is the average of traffic rate of flow  $l$  in  $j^{th}$  cluster.

Then, we need to select the optimal VSCP set for each cluster to build action space, which should satisfy the service request constraints, capacity constraints, delay constraints and flow conservation constraints. Furthermore, the balance of network resource provisioning cost and service performance gain must be considered. The estimated network profit of a VSCP set at slot  $t$  can be expressed as

$$\chi_t' = -\alpha_1 \phi_{1,t} - \alpha_2 \phi_{2,t} + \alpha_4 R_t. \quad (52)$$

---

**Algorithm 1** The Historical Records of the Network Load Clustering Algorithm

---

**Require:** Number of desired clusters  $|\mathcal{W}|$ , the historical records of the network load  $\lambda_{\mathcal{L},\mathcal{T}}^H$ , the number of iterations  $C$ .

**Ensure:** The clustering result of the historical records of the network load  $\mathcal{H} = (h_1, h_2, \dots, h_{|\mathcal{W}|})$ .

- 1: Randomly select  $|\mathcal{W}|$  elements from the historical records of the network load  $\lambda_{\mathcal{L},\mathcal{T}}^H$  as initial cluster centers, and denote  $i = 1$ ;
- 2: **while**  $i \leq C$  **do**
- 3: Calculate the Euclidean distances between each element  $\lambda_{\mathcal{L},t}^H$  and all cluster centers  $\|\lambda_{\mathcal{L},t}^H - \lambda_{\mathcal{L},j}^C\|_{\lambda_{\mathcal{L},t}^H \in \lambda_{\mathcal{L},\mathcal{T}}^H, 1 \leq j \leq |\mathcal{W}|}$ , and assign  $\lambda_{\mathcal{L},t}^H$  to the nearest cluster;
- 4: **for** each cluster  $h_j$  ( $1 \leq j \leq |\mathcal{W}|$ ) **do**
- 5: Recalculate the cluster center by averaging the elements in the cluster;
- 6: **end for**
- 7:  $i = i + 1$ ;
- 8: **end while**

---

Note that  $\chi_t'$  neglects VNF migration cost because we select the optimal VSCP set for different network traffic loads, the temporal correlation of network traffic loads is broken.

Algorithm 2 finds the optimal VSCP set for different network loads in a greedy manner. The algorithm firstly searches all VSCPs for each flow based on a breadth-first search method. For each cluster, it searches the VSCP set with the highest value of SR and calculates the estimated network profit of the VSCP set. This VSCP set saves network resource cost while the service performance gain may not be optimal. Then, algorithm 2 searches the new VSCP sets by reducing SR and calculates their estimated network profits. It then compares the estimated network profit of the searched VSCP sets, and selects the set with the highest estimated network profit to add the action space  $\mathcal{W}$ . The search process is terminated until the estimated network profit of new searched VSCP sets does not increase or all VSCP sets are traversed.

**Reward:** The reward obtained in time interval  $k$  is evaluated as the total network profit, which is expressed as

$$r^{(k)} = \sum_{t=N(k-1)+1}^{Nk} \chi_t. \quad (53)$$

The DDVSC algorithm is shown in Algorithm 3 and Fig. 4. The NOC receives the current network state  $s^{(k)}$ , and the evaluation network and the target network estimate the  $Q$ -values of each action  $Q(s^{(k)}, w; \theta)$  and the target  $Q$ -value  $\max_{w \in \mathcal{W}} Q'(s^{(k+1)}, w; \theta')$ , where  $\theta$  and  $\theta'$  are the corresponding parameters of these two neural networks. Then, the NOC selects the VSCP set  $w^{(k)}$  via  $\epsilon$ -greedy exploration and executes the corresponding VNF selection and virtual link mapping decisions to obtain the reward  $r^{(k)}$  and next state  $s^{(k+1)}$ . A transition  $[s^{(k)}, w^{(k)}, r^{(k)}, s^{(k+1)}]$ , including the current network state, VSCP set, network profit and the next network state, is stored into a replay buffer  $\mathbf{E}$ . The minibatch

---

**Algorithm 2** Optimal VSCP set Searching for Different Network Loads

---

**Require:** The clustering results of the historical records of the network load  $\mathcal{H} = (h_1, \dots, h_{|\mathcal{W}|})$ .

**Ensure:** The action space  $\mathcal{W}$ .

- 1: Search all VSCPs for each flow based on a breadth-first search method;
- 2: **for** each cluster  $h_i$  ( $1 \leq i \leq |\mathcal{W}|$ ) **do**
- 3: Initialize searching step  $\delta_1$  ( $0 < \delta_1 < 1$ ), the estimated network profit  $p_0 = 0$ , the optimal VSCP set  $w = \emptyset$ , and set iteration number  $i = 0$ ;
- 4:  $i = i + 1$ ;
- 5: Set sharing ratio of computing resource  $1 - i \times \delta_1 \leq SR \leq 1 - (i - 1) \times \delta_1$ ;
- 6: **while**  $1 - (i - 1) \times \delta_1 < 0$  **do**
- 7:   **if**  $w = \emptyset$  **then**
- 8:     There is no VSCP set can meet service demands;
- 9:   **else**
- 10:     The optimal VSCP set  $w$  is added to the action space  $\mathcal{W}$ ;
- 11:   **end if**
- 12:   Break loop;
- 13: **end while**
- 14: Search the VSCP sets that satisfy the sharing ratio of computing resource and problem constraints (13) – (17), (26) – (28), (37);
- 15: **if** no new VSCP set is searched **then**
- 16:   Return to step 4;
- 17: **else**
- 18:   Calculate the estimated network profit of the searched VSCP sets using (52);
- 19:   **end if**
- 20: Select the VSCP set  $w'$  with the highest estimated network profit  $p_1$  from the new searched sets;
- 21: **if**  $p_1 > p_0$  **then**
- 22:    $w = w'$ ,  $p_0 = p_1$ ;
- 23:   Return to step 4;
- 24: **else**
- 25:   Break loop;
- 26: **end if**
- 27: **end for**

---

$\mathcal{B}$  is formulated with  $G$  experiences randomly sampled from  $\mathbf{E}$ , given by

$$\mathcal{B} = \left\{ \left( s^{(g(j))}, w^{(g(j))}, r^{(g(j))}, s^{(g(j))+1} \right) \in \mathbf{E} \mid 1 \leq j \leq G \right\}. \quad (54)$$

Then, the weights of the evaluate NN,  $\theta$ , are updated by minimizing the loss function given by [37]

$$\mathcal{L}(\theta) = \mathbb{E} \left[ \mu^{(k)} - Q(s^{(k)}, w^{(k)}; \theta) \right]^2 \quad (55)$$

where  $\mu^{(k)}$  is the target value given by

$$\mu^{(k)} = r^{(k)} + \delta_2 \max_{w \in \mathcal{W}} Q'(s^{(k+1)}, w; \theta') \quad (56)$$

where  $\delta_2 \in [0, 1]$  is the discount factor and represents the weight of uncertainty about the future utility in the learning

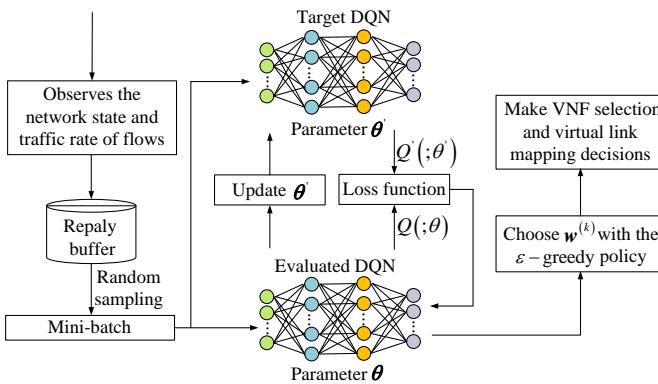


Fig. 4: An illustration of the DRL-based dynamic VNF selection and chaining algorithm.

process. The target NN copies evaluation NN weights to update its weights  $\theta'$  in every  $C$  steps.

In the VNF selection and chaining process, the NOC first observes the satellite segment covering the ground network segment to identify available satellites. The SGIN is monitored by the NOC in each time slot, where the topology (i.e., location and connectivity of nodes), resource states (i.e., usage of links and nodes) of the SGIN, and traffic rates of flows can be obtained via CDPIs. Then, NOC uses the DDVSC algorithm to determine the VSCP set for flows at the start point of each time interval, and flow tables for VNF selection and chaining are delivered to the satellites and ground nodes via SGLs and wired links. When users switch accessing satellites, satellites in the virtual lattice grid are changed and their flow tables are updated by control signaling based on the selected VSCP set. As the number of users increases, more flow tables may need to be updated. To solve this problem, cluster-based flow tables update approaches [38] are proposed where users in similar areas are partitioned into groups and relevant flow tables are updated by control signaling in groups.

The computational complexity of the DDVSC algorithm mainly depends on the number of multiplications in the two NNs [39]. The evaluation NN executes both the forward propagation and backpropagation, while the target NN only executes the forward propagation. We consider that the evaluated NN consists of two FC layers, each with  $m_2$  and  $m_3$  neural cells, respectively. The input size of the evaluated NN is denoted by  $L + m_1 + p$ .  $L$  parameters are used to represent traffic rates of flows,  $m_1$  parameters are used to represent the network configuration information, and  $p$  parameters are used to represent the previous VSCP set. The evaluated NN outputs  $|\mathcal{W}|$  values representing the value function of each action in the current state.  $G$  experiences are sampled from the replay buffer.

The forward propagation of the evaluated NN requires  $\mathcal{O}_1$  multiplications, given by

$$\mathcal{O}_1 = G [m_2(L + m_1 + m_3 + p) + m_3 + |\mathcal{W}|(m_3 + p)]. \quad (57)$$

The backpropagation of the evaluated NN has  $\mathcal{O}_2$  multiplica-

### Algorithm 3 DRL-Based Dynamic VNF Selection and Chaining Algorithm for SGINs

```

1: Initialize the learning rate of the evaluation network, the discount factor  $\delta_2$ , the maximum learning episode  $EP$ , the existing VSCP set  $w^{(0)}$ , the maximum training steps  $n$  per episode, the replay buffer  $E$ , the evaluation network with random weights  $\theta$ , the target network with weights  $\theta' = \theta$ .
2: for episode = 1 :  $EP$  do
3:   The NOC receives the traffic rate of flows  $\lambda_L^{(1)}$ , the network state parameter  $\varpi^{(1)}$  and the existing VSCP set  $w^{(0)}$ ;
4:    $s^{(1)} = [\lambda_L^{(1)}, \varpi^{(1)}, w^{(0)}]$ ;
5:   for  $k = 1 : n$  do
6:     Select action  $w^{(k)}$  according to  $\varepsilon$ -greedy algorithm ;
7:     Execute the action  $w^{(k)}$  in the SGIN, obtain the reward  $r^{(k)}$ , next state  $s^{(k+1)}$  ;
8:     Formulate a memory transition  $e^{(k)} = [s^{(k)}, w^{(k)}, r^{(k)}, s^{(k+1)}]$  ;
9:     Store the transition into replay buffer:  $E \leftarrow E \cup e^{(k)}$  ;
10:    Obtain a minibatch  $B$  by uniformly and randomly sampling  $G$  experiences from the replay buffer  $E$  ;
11:    Set  $\mu^{(k)}$  according to (56) ;
12:    Perform a gradient descent step on (55) with respect to the network parameters  $\theta$  ;
13:    Update  $\theta'$  with  $\theta$  every  $C$  steps ;
14:  end for
15: end for

```

tions, with

$$\mathcal{O}_2 = G [2m_2(L + m_1 + m_3 + p) + 2m_3 + 3|\mathcal{W}|(m_3 + p)]. \quad (58)$$

The target NN has  $\mathcal{O}_3$  multiplications given by

$$\mathcal{O}_3 = G [m_2(L + m_1 + m_3 + p) + m_3 + |\mathcal{W}|(m_3 + p)]. \quad (59)$$

The computational complexity of DDVSC algorithm  $\mathcal{O}_4$  depends on  $\mathcal{O}_1$ ,  $\mathcal{O}_2$  and  $\mathcal{O}_3$ , given by

$$\mathcal{O}_4 = \mathcal{O} [G(Lm_2 + m_2m_3 + m_1m_2 + |\mathcal{W}|m_3 + m_2p + |\mathcal{W}|p)]. \quad (60)$$

## VI. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed algorithm based on an important 6G case study on immersive online education service.

### A. Simulation Setup

We consider an SGIN scenario as Fig. 3. There are three flows in an SGIN. The users in *Lanzhou* city and *Chengdu* city access an immersive online education service via local BS, and the users in *Wuwei* city access an immersive online education service via a VSAT. The two ESs are located in *Chengdu* city and *Anshun* city, and they are responsible for communication with satellites. The whole time horizon is 95

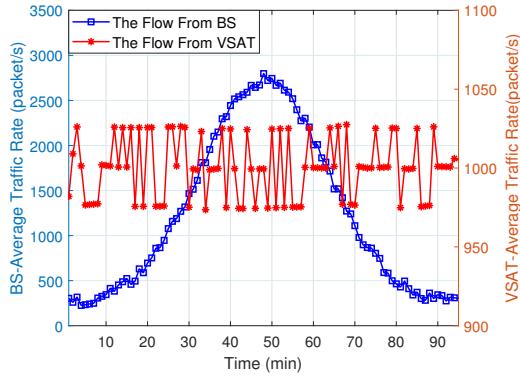


Fig. 5: Average traffic rate of flows.

TABLE III: SIMULATION PARAMETERS [2], [27]

Parameter	Value
Number of orbits planes	72
Number of satellite per plane	22
Height of orbits	550 km
Orbit inclination angle	53 deg
Minimum elevation angle of ES	20 deg
EIRP of VSATs/satellites	34.5 dBW
Quality Factor of Satellite/ES	11.1/22.7 dBi/K
$\varsigma_1, \varsigma_2, \varsigma_3$	92.5, 20, 20
Other signal transmission losses of SGLs/ISLs	1 dB
Packet size	10 KB
Delay requirement	10 ms
Simulation time	5700 s
$\alpha_1, \alpha_2, \alpha_3, \alpha_4$	1000, 5000, 125, 1

min from 14 Sep 2023 04:00:00.000 UTCG to 14 Sep 2023 05:35:00.000 UTCG. The time slot length is set to 1 second and the time interval length is set to 5 min, thus the whole time horizon consists of 19 time intervals.

The CPU processing rate of satellites is 2 GHz. The ground node in *Guiyang* city has a professional processor to render images, and its CPU processing rate is 9 GHz. The CPU processing rate is 3 GHz for the other ground nodes with a server. The computation intensity of physical nodes for VNFs is 10 cycles/bit and the packet size is 10 KB. The delay requirement of the service is 10 ms [2]. The transmission rate of wired links is 0.4 Gbps and the communication bandwidth of SGLs and ISLs is 45 MHz. The EIRP of VSAT and satellite is 34.5 dBW, and the quality factor of ES and satellite is 22.7 dBi/K and 11.1 dBi/K [27]. We consider the packet arrival process of flows according to Poisson distribution, and the two flows accessing BSs have the same traffic rate. The average traffic rate of flows is shown in Fig. 5. The main simulation parameters are listed in Table III.

### B. Results and Analysis

The NOC selects a VSCP set at each time interval. The feasible VSCPs sets must not violate the physical resource constraints, service request constraints, delay constraints and flow conservation constraints. The number of feasible VSCP sets selected in the whole time horizon is shown in Fig. 6, which is calculated using the average value over 10 episodes. Fig. 6 proves the proposed DDVSC algorithm can always

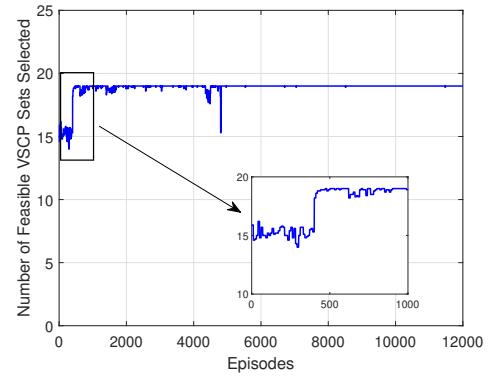


Fig. 6: Number of feasible VSCP sets selected in whole time horizon.

choose the feasible VSCP sets in each time interval after 4800 episodes.

The performance of the DDVSC algorithm is evaluated and compared with that of three existing algorithms [13], [40], [41]. The work in [13] proposes a greedy minimum cost algorithm that prefers to use and share the VNF instances installed at ground nodes to save the computing resources. When the service performance can not be guaranteed solely by the ground network segment, the satellites are leveraged to balance the network-wide computation load. The work in [40] proposes a greedy best availability algorithm that uses multiple VNF instances at different network locations to improve service performance. In [41], a dynamic VNF selection and chaining algorithm is designed based on the actor-critic network to maximize the long-term network profit.

The comparison of the four algorithms in terms of accumulative network profit is shown in Fig. 7. Moreover, the upper bound of accumulative network profit is also given in Fig. 7. Compared with the three existing algorithms, the proposed DDVSC algorithm increases the accumulative network profit by 9.5 %, 32.9 % and 8.9 %, respectively. Meanwhile, the DDVSC algorithm is only 3.4 % lower than the upper bound of accumulative network profit. The reason is that the DDVSC algorithm can select an appropriate VSCP set in different network traffic loads. Specifically, when the network traffic load is low, the DDVSC algorithm lets multiple flows share the VNF instances to save the computing resource provisioning cost. When the network traffic load is high, the DDVSC algorithm lets flows use more VNF instances to obtain more service performance gain. Note that the greedy best availability algorithm has negative network profits in some time intervals. This is because when the network traffic load is low, the network resource provisioning cost outweighs the service performance gain.

The comparison of the four algorithms in terms of accumulative network resource provisioning cost (ANRPC) is shown in Fig. 8. It can be observed that among the four algorithms, the greedy minimum cost algorithm has the lowest ANRPC and the proposed DDVSC algorithm has the second lowest ANRPC. This is because the greedy minimum cost algorithm aims to minimize the network resource provisioning cost. However, the proposed DDVSC algorithm can obtain

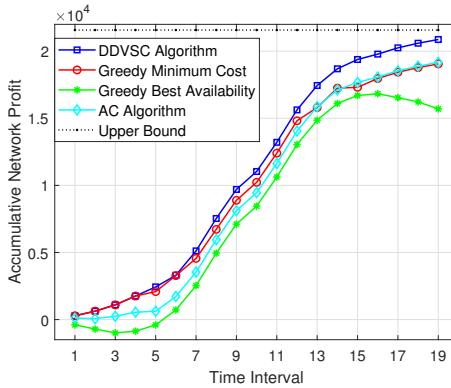


Fig. 7: Comparison of accumulative network profit between the three existing algorithms and the proposed DDVSC algorithm.

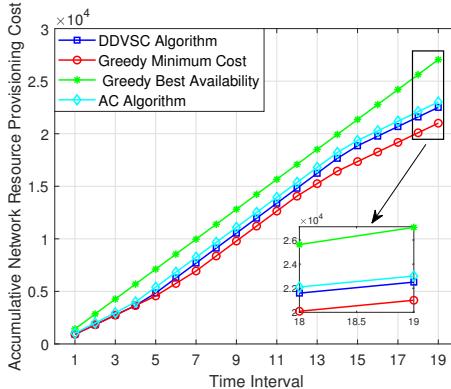


Fig. 8: Comparison of accumulative network resource provisioning cost between the three existing algorithms and the proposed DDVSC algorithm.

more service performance gain than the greedy minimum cost algorithm. In addition, the greedy best availability algorithm has the highest ANRPC. This is because the greedy best availability algorithm uses more computing resources to obtain the best service performance.

Figure 9 shows the comparison among the four algorithms in terms of the accumulative service performance gain (ASPG). We can see that the greedy best availability algorithm has the highest ASPG but it also pays the highest network resource provisioning cost. The ASPG of the AC algorithm is slightly higher than the DDVSC algorithm, this is because the AC algorithm provides more computing resources for the service which leads to increasing the network resource provisioning cost. Note that the greedy minimum cost algorithm and the greedy best availability algorithm can only obtain optimality in terms of network cost or service performance, while the proposed DDVSC algorithm balances network cost and service performance to maximize the total network profit.

In SGINs, the satellite network segment can provide additional computing resources to balance the network-wide computation load and improve service performance. Figure 10 shows the comparison among the four algorithms in terms of the number of VNF instances executed at satellites in each time interval. We can see that the greedy best availability algorithm uses the VNF instances installed at satellites in each

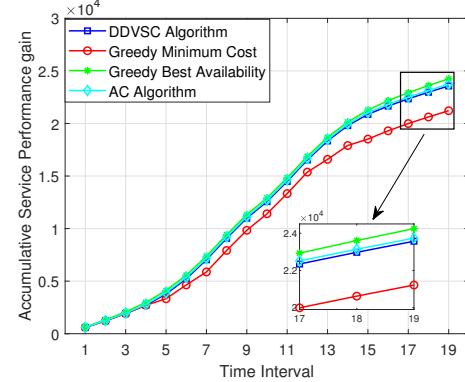


Fig. 9: Comparison of accumulative service performance gain between the three existing algorithms and the proposed DDVSC algorithm.

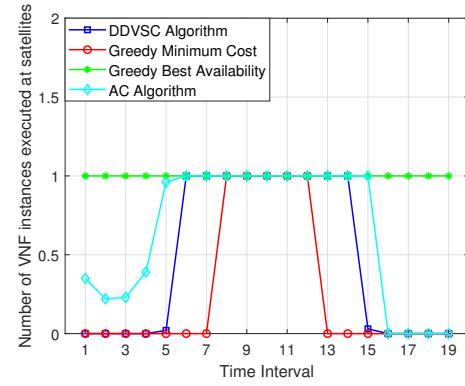


Fig. 10: Number of VNF instances executed at satellites in each time interval.

time interval to obtain more computing resources to improve service performance. The greedy minimum cost algorithm minimally utilizes the VNF instances installed at satellites because the algorithm prefers to use the VNF instances deployed at ground nodes to improve resource utilization efficiency. Both the DDVSC algorithm and AC algorithm can use VNF instances installed at satellites based on the evolving network states, and the proposed DDVSC algorithm is more robust to varying traffic flows as shown in Fig. 10 compared with the AC algorithm. This is because the action space of the DDVSC algorithm is compressed by the clustering approach.

When a VNF is mapped to a satellite, satellite movements lead to a VNF migration. For the proposed DDVSC algorithm, Fig. 11 shows the number of VNF instances executed at satellites in each time interval in different VNF migration costs. We set  $\alpha_3 = 0, 500, 2000$  to represent different migration costs. We can see that the number of VNF instances executed at satellites is decreasing as the VNF migration cost increases. This is because when migration cost is high, the NOC prefers to use the VNF instances deployed at ground nodes to reduce VNF migrations due to satellite movements.

## VII. CONCLUSION

In this paper, we investigate a DVSC problem in an SGIN which is described as an MDP formulation. The objective is to

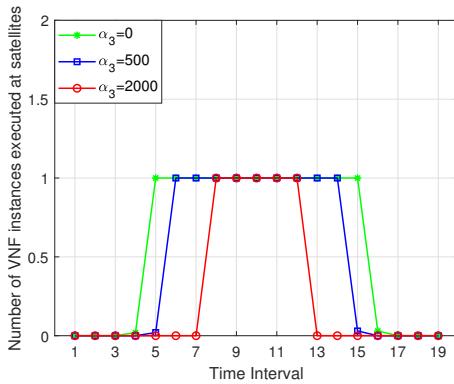


Fig. 11: Number of VNF instances executed at satellites in each time interval in different VNF migration costs.

maximize the accumulative network profit in a heterogeneous and time-varying SGIN with a dynamic network traffic load. To solve the DVSC problem, a novel VNF selection and chaining algorithm based on DRL is proposed to dynamically determine the VSCP set. Due to satellite movements, each VSCP consisting of a sequence of VNF selection and virtual link mapping decisions adapts to the changing network topology. Moreover, SR is proposed to capture the relation between the network resource provisioning cost and the service performance gain. To efficiently allocate heterogeneous network resources, we cluster the historical records of the network load and search the optimal VSCP sets for different network loads in a greedy manner. Extensive simulation results demonstrate the advantages of the proposed DDVSC algorithm, compared with existing works, in terms of maximizing the accumulative network profit and service performance gain with minimal network resource provisioning cost. In SGINs, the satellite network segment can route flows to ESs for accessing VNF servers in the ground network segment. For future work, we will investigate the impact of the density and location of ES deployment on the VNF selection and chaining problem.

#### ACKNOWLEDGMENTS

The work presented in this paper is partially supported by National Natural Science Foundation of China under Grant NO. 91638204.

#### REFERENCES

- [1] S. Chen, Y. -C. Liang, S. Sun, S. Kang, W. Cheng and M. Peng, "Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed," *IEEE Wirel. Commun.*, vol. 27, no. 2, pp. 218-228, Apr. 2020.
- [2] *6G Applications and Use Cases*, Next G Alliance, May. 2022.
- [3] X. Qin, T. Ma, Z. Tang, X. Zhang, H. Zhou and L. Zhao, "Service-aware resource orchestration in ultra-dense LEO satellite-terrestrial integrated 6G: A service function chain approach," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 9, pp. 6003-6017, Sept. 2023.
- [4] X. Shen, J. Gao, W. Wu, K. Lyu, M. Li, W. Zhuang, X. Li, and J. Rao, "AI-assisted network-slicing based next-generation wireless networks," *IEEE Open Journal of Vehicular Technology (OJVT)*, vol. 1, no. 1, pp. 45-66, Jan. 2020.
- [5] X. Shen, J. Gao, W. Wu, M. Li, C. Zhou, and W. Zhuang, "Holistic Network Virtualization and Pervasive Network Intelligence for 6G," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 1, pp. 1-30, 1st Quart. 2022.
- [6] W. Zhuang, Q. Ye, F. Lyu, N. Cheng, and J. Ren, "SDN/NFV empowered future IoV with enhanced communication, computing, and caching," *Proceedings of the IEEE (JPROC)*, vol. 108, no. 2, pp. 274-291, Feb. 2020.
- [7] A. Varasteh, B. Madiwalar, A. Van Bemten, W. Kellerer and C. Mas-Machuca, "Holu: Power-aware and delay-constrained VNF placement and chaining," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 2, pp. 1524-1539, Jun. 2021.
- [8] X. Shang, Y. Huang, Z. Liu and Y. Yang, "Reducing the service function chain backup cost over the edge and cloud by a self-adapting scheme," *IEEE Trans. Mobile Comput.*, vol. 21, no. 8, pp. 2994-3008, Aug. 2022.
- [9] H. Hawilo, M. Jammal and A. Shami, "Network function virtualization-aware orchestrator for service function chaining placement in the cloud," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 3, pp. 643-655, Mar. 2019.
- [10] M. Karimzadeh-Farshbafan, V. Shah-Mansouri and D. Niyyato, "Reliability aware service placement using a viterbi-based algorithm," *IEEE Trans. Netw. Serv. Manag.*, vol. 17, no. 1, pp. 622-636, Mar. 2020.
- [11] Y. Liu, Y. Lu, X. Li, Z. Yao and D. Zhao, "On dynamic service function chain reconfiguration in IoT networks," *IEEE Internet Things J.*, vol. 7, no. 11, pp. 10969-10984, Nov. 2020.
- [12] X. Fu, F. R. Yu, J. Wang, Q. Qi and J. Liao, "Dynamic service function chain embedding for NFV-enabled IoT: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 507-519, Jan. 2020.
- [13] G. Wang, S. Zhou, S. Zhang, Z. Niu and X. Shen, "SFC-based service provisioning for reconfigurable space-air-ground integrated networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 7, pp. 1478-1489, Jul. 2020.
- [14] B. Feng, G. Li, G. Li, H. Zhou, H. Zhang and S. Yu, "Efficient mappings of service function chains at terrestrial-satellite hybrid cloud networks," in *Proc. IEEE GLOBECOM*, Abu Dhabi, Dec. 2018, pp. 1-6.
- [15] P. Zhang, C. Wang, N. Kumar and L. Liu, "Space-air-ground integrated multi-domain network resource orchestration based on virtual network architecture: A DRL method," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2798-2808, Mar. 2022.
- [16] P. Zhang, P. Yang, N. Kumar and M. Guizani, "Space-air-ground integrated network resource allocation based on service function chain," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7730-7738, Jul. 2022.
- [17] J. Zhang, Y. Tang, T. Ye and Y. Sun, "SFC-based service provisioning for 6G satellite-ground integrated networks," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Xiamen, Jul. 2021, pp. 951-956.
- [18] S. Zhou, G. Wang, S. Zhang, Z. Niu and X. S. Shen, "Bidirectional mission offloading for agile space-air-ground integrated networks," *IEEE Wirel. Commun.*, vol. 26, no. 2, pp. 38-45, Apr. 2019.
- [19] Q. Chen, G. Giambene, L. Yang, C. Fan and X. Chen, "Analysis of inter-satellite link paths for LEO mega-constellation networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2743-2755, Mar. 2021.
- [20] J. Li, W. Shi, H. Wu, S. Zhang and X. Shen, "Cost-aware dynamic SFC mapping and scheduling in SDN/NFV-enabled space-air-ground integrated networks for internet of vehicles," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5824-5838, Apr. 2022.
- [21] H. Yang, W. Liu, X. Wang and J. Li, "Group sparse space information network with joint virtual network function deployment and maximum flow routing strategy," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5291-5305, Aug. 2023.
- [22] H. Yang, W. Liu, J. Li and T. Q. S. Quek, "Space information network with joint virtual network function deployment and flow routing strategy with QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1737-1756, Jun. 2023.
- [23] A. Kak and I. F. Akyildiz, "Towards automatic network slicing for the internet of space things," *IEEE Trans. Netw. Serv. Manag.*, vol. 19, no. 1, pp. 392-412, Mar. 2022.
- [24] Z. Niu, X. S. Shen, Q. Zhang and Y. Tang, "Space-air-ground integrated vehicular network for connected and automated vehicles: Challenges and solutions," *Intelligent and Converged Networks*, vol. 1, no. 2, pp. 142-169, Sept. 2020.
- [25] T. Chen *et al.*, "Learning-based computation offloading for IoT through Ka/Q-band satellite-terrestrial integrated networks," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12056-12070, Jul. 2022.
- [26] T. Wang, P. Li, Y. Wu, L. Qian, Z. Su and R. Lu, "Quantum-empowered federated learning in space-air-ground integrated networks," *IEEE Network*, doi: 10.1109/MNET.2023.3318083.
- [27] *Space exploration technologies. SpaceX non-geostationary satellite system attachment: schedule stечnical report*. Accessed: Nov. 2023. [Online]. Available: <https://www.fcc.report/IBFS/SATMOD-20181108-00083/1569860.pdf>.
- [28] F. Wang, D. Jiang, Z. Wang, J. Chen and T. Q. S. Quek, "Seamless handover in LEO based non-terrestrial networks: Service continuity and

- optimization," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 1008-1023, Feb. 2023.
- [29] M. Sheng, Y. Wang, J. Li, R. Liu, D. Zhou and L. He, "Toward a flexible and reconfigurable broadband satellite network: Resource management architecture and strategies," *IEEE Wirel. Commun.*, vol. 24, no. 4, pp. 127-133, Aug. 2017.
- [30] M. Salehi, K. Hooli, J. Hulkko and A. Tölli, "Enhancing next-generation extended reality applications with coded caching," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 1371-1382, Jun. 2023.
- [31] H. Tang, D. Zhou and D. Chen, "Dynamic network function instance scaling based on traffic forecasting and VNF placement in operator data centers," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 3, pp. 530-543, Mar. 2019.
- [32] D. Li, P. Hong, K. Xue and J. Pei, "Availability aware VNF deployment in datacenter through shared redundancy and multi-tenancy," *IEEE Trans. Netw. Serv. Manag.*, vol. 16, no. 4, pp. 1651-1664, Dec. 2019.
- [33] "Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications; Part II (DVB-S2X)," Digital Video Broadcasting (DVB), ETSI EN 302307-2, Oct. 2014.
- [34] S. Nie and I. F. Akyildiz, "Channel modeling and analysis of inter-small-satellite links in terahertz band space networks," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8585-8599, Dec. 2021.
- [35] Q. Ye, W. Zhuang, X. Li and J. Rao, "End-to-end delay modeling for embedded VNF chains in 5G core networks," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 692-704, Feb. 2019.
- [36] F. Baris, S. R. Chowdhury, R. Ahmed, R. Boutaba, and O. C. M. B. Duarte, "Orchestrating virtualized network functions," *IEEE Trans. Netw. Serv. Manag.*, vol. 13, no. 4, pp. 725-739, Dec. 2016.
- [37] Q. Ye, W. Shi, K. Qu, H. He, W. Zhuang and X. Shen, "Joint ran slicing and computation offloading for autonomous vehicular networks: A learning-assisted hierarchical approach," *IEEE Open Journal of Vehicular Technology (OJVT)*, vol. 2, pp. 272-288, Jun. 2021.
- [38] Q. Liu et al., "Cluster-based flow control in hybrid software-defined wireless sensor networks," *Comput. Netw.*, vol. 187, 2021.
- [39] Y. Xiao, L. Xiao, X. Lu, H. Zhang, S. Yu and H. V. Poor, "Deep-reinforcement-learning-based user profile perturbation for privacy-aware recommendation," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4560-4568, Mar. 2021.
- [40] R. Mijumbi, J. Serrat, J. -L. Gorracho, N. Bouten, F. De Turck and S. Davy, "Design and evaluation of algorithms for mapping and scheduling of virtual network functions," in *Proc. IEEE NetSoft*, London, Apr. 2015, pp. 1-9.
- [41] R. Wang, J. Li, K. Wang, X. Liu and X. Lit, "Service function chaining in NFV-enabled edge networks with natural actor-critic deep reinforcement learning," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Xiamen, Jul. 2021, pp. 1095-1100.

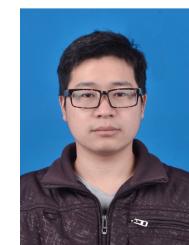


**Qiang (John) Ye** (Senior Member, IEEE) received the PhD degree in Electrical and Computer Engineering from the University of Waterloo, ON, Canada, in 2016. Since Sept. 2023, he has been an Assistant Professor with the Department of Electrical and Software Engineering, Schulich School of Engineering, University of Calgary, AB, Canada. Before joining UCalgary, he worked as an Assistant Professor with the Department of Computer Science, Memorial University of Newfoundland, NL, Canada from Sept. 2021 to Aug. 2023 and with the Department of Electrical and Computer Engineering and Technology, Minnesota State University, Mankato, USA, from Sept. 2019 to Aug. 2021, respectively. He was with the Department of Electrical and Computer Engineering, University of Waterloo as a Postdoctoral Fellow and then a Research Associate from Dec. 2016 to Sept. 2019. He has published over 70 research articles on top-ranked journals and conference proceedings. He is/was the General, Publication, Program Co-chairs for different reputable international conferences and workshops, and serves/served as Associate Editors of prestigious international journals, e.g., IEEE Transactions on Vehicular Technology and IEEE Transactions on Cognitive Communications and Networking. He also serves/served as the IEEE Vehicular Technology Society (VTS) Region 7 Chapter Coordinator (2024) and the Regions 1-7 Chapters Coordinator(2022-2023). Dr. Ye received the Best Paper Award in the IEEE/CIC International Conference on Communications in China (ICCC) in 2024 and the IEEE Transactions on Cognitive Communications and Networking Exemplary Editor Award in 2023.



**Kaige Qu** (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2021. She received the B.Sc. degree in communication engineering from Shandong University, Jinan, China, in 2013, and M.Sc. degrees in integrated circuits engineering and electrical engineering from Tsinghua University, Beijing, China, and KU Leuven, Leuven, Belgium, respectively, in 2016. From February 2021 to December 2023, she was a Post-doctoral Fellow and then a Research Associate

with the Department of Electrical and Computer Engineering, University of Waterloo. Her research interests include connected and autonomous vehicles, network intelligence, network virtualization, and digital twin assisted network automation.



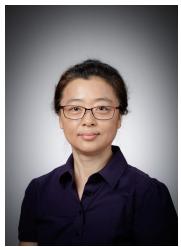
**Yanglong Sun** (Member, IEEE) received the B.S. and M.S. degrees from Zhengzhou University, China, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree in communication engineering, Xiamen University, China. He was a visiting Ph.D. student at BBCRlab, Department of Electrical and Computer Engineering, University of Waterloo, Canada, from 2018 to 2019. His research is focus on vehicular adhoc networks, resource allocation in IoV, mobile 5G networks, and intelligent transportation systems.



**Yuliang Tang** (Member, IEEE) received the M.S. degree from the Beijing University of Posts and Telecommunications, China, in 1996, and the Ph.D. degree in information and communication engineering from Xiamen University in 2009. He is currently a Professor with the Department of Information and Communication Engineering, Xiamen University. He has published more than 100 papers in journals and international conferences. His current research interests include wireless communication, 5G and beyond, and vehicular ad hoc networks. He has been granted more than 20 patents in his research areas.



**Jianxin Zhang** received the B.S. degree in communication engineering from Henan Normal University, Xinxiang, China, in 2016, the M.S. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 2019. He is currently pursuing the Ph.D. degree in communication engineering with Xiamen University, Xiamen, China. His research interests include space-air-ground integrated networks, SDN/NFV and network slicing.



**Dongmei Zhao** (Senior Member, IEEE) received a Ph.D degree in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada in June 2002. In July 2002 she joined the Department of Electrical and Computer Engineering at McMaster University, where she is a full professor. Dr. Zhao is an associate editor of the IEEE Internet of Things Journal. She served as an associate editor of the IEEE Transactions on Vehicular Technology between 2007 and 2017. She serves as IEEE VTS distinguished lecturer. Dr. Zhao

a Professional Engineer of Ontario. Her current research areas are mainly in mobile wireless networking, energy efficient networking, vehicular networks, and digital twins.



**Tong Ye** received the B.S. degree in communication engineering from Fuzhou University, Fuzhou, China, in 2020, the M.S. degree in communication engineering from Xiamen University, Xiamen, China, in 2023. Her research interests include service function chain and network virtualization .