On-Demand Dynamic Controller Placement in Software Defined Satellite-Terrestrial Networking

Zhenzhen Han, Chuan Xu, *Member, IEEE*, Zhengying Xiong, Guofeng Zhao, *Member, IEEE*, and Shui Yu, *Senior Member, IEEE*

***Abstract*—Software defined satellite-terrestrial networking has been identified as a promising approach to support the diversity of network services. As the fundamental issue to improve the flexibility of network management, the controller placement problem has been attracted increasing attentions for the integration of satellite and terrestrial networking. However, the impact of the dynamic coverage demands on the controller placement have not been well investigated in existed works, which makes them fail to adjust the coverage dynamically according to the actual demands, and leads to an obvious increase of networking response latency to the terminals. Aiming to address this issue, we propose a novel on-demand dynamic controller placement scheme, which can optimize the placement of controllers to improve networking response latency while meeting the dynamic coverage demands. Firstly, to optimize the number of controllers and meet the dynamic coverage demands, we define the coverage redundancy and propose the redundancy-based satellite subnet division method to establish the reliable satellite subnets. Secondly, we quantify the networking response latency of the distributed satellite subnet, and build an optimization mathematical model to optimize the number and location of controllers. Then, we formulate the controller placement problem into the capacitated facility location problem and build the mathematical model for it. Moreover, the on-demand dynamic approximation algorithm is proposed to obtain the approximation solution. Finally, the simulation results demonstrate that the proposed algorithm can effectively optimize the network latency compared with related algorithms.**

***Index Terms*—Software defined satellite-terrestrial networking, controller placement problem, satellite subnet division, networking response latency, approximation algorithm.**

I. Introduction

THE TREMENDOUS growth in network service diversity demands has increased the integration of satellite and terrestrial network [1], [2]. Software defined networking (SDN) can be introduced to construct software defined satellite-terrestrial networking (SDS-TN) to extend the flexibility of network management [3], [4]. Affected by the limitation of on-board storage computing capacities and the scale of the satellite network, multiple controllers need to be deployed to build a distributed control plane to improve the control and management performance of the satellite network. Therefore, the rational placement of controllers has become the bottleneck for improving network management flexibility.

For terrestrial networks, the controller placement problem (CPP) is first proposed in work [5]. Thereafter, many related approaches of controller placement are proposed for increasing the flexibility of the network control plane [6]. Due to the dynamics of satellite network and the limited on-board processing capacity of satellites, the controller placement schemes based on terrestrial network cannot be directly applied to SDS-TN. Therefore, it is still a challenge to deploy appropriate controllers to improve the performance of SDS-TN.

In order to address the CPP in satellite network, some multi-controller placement schemes are proposed [7]–[9]. In these controller placement schemes, the default entire satellite constellation networks are generally selected to deploy the controller. Although these works can obtain placement scheme to optimize network latency, the impact of dynamic coverage demand of terrestrial terminals on the satellite controller placement has not been well investigated. Therefore, the controller placement scheme cannot be adjusted dynamically according to dynamic changes of network nodes, which may make an obvious increase of networking latency when the satellite network responses to the requests from terminals. It is worth mentioning that the idea of network area division has been introduced into the controller placement scheme of the terrestrial network, which can improve network latency and reduce algorithm complexity [10], [11]. However, when the concept of subnet division method is applied into the satellite network to design the controller placement scheme, it is necessary to consider the impact of the relative movement of satellites and terminals on the satellite subnet division method. Otherwise, there will be a problem of coverage vulnerability due to the untimely switching of satellite-to-terminal, which leads the design of the dynamic controller placement scheme of SDS-TN more challenging.

In this paper, we propose a novel on-demand dynamic controller placement scheme for SDS-TN to improve the networking response latency while guaranteeing the dynamic coverage of satellite network. Firstly, we design a method for

Manuscript received November 10, 2020; revised February 4, 2021; accepted February 15, 2021. Date of publication February 22, 2021; date of current version September 9, 2021. This work is supported by the National Key Research and Development Project of China under grant (2018YFB1800301, 2018YFB1800304), the Major Theme Special Project of Chongqing Technology Innovation and Application Development Special Project(cstc2019jscx-zdztzxX0013), the Chongqing Graduate Research and Innovation Project (CYB18175, BYJS201803). The associate editor coordinating the review of this article and approving it for publication was M. F. Zhani. *(Corresponding author: Chuan Xu.)*

Zhenzhen Han, Chuan Xu, Zhengying Xiong, and Guofeng Zhao are with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: xuchuan@cqupt.edu.cn).

Shui Yu is with the School of Computer Science, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: shui.yu@uts.edu.au).

Digital Object Identifier 10.1109/TNSM.2021.3061261

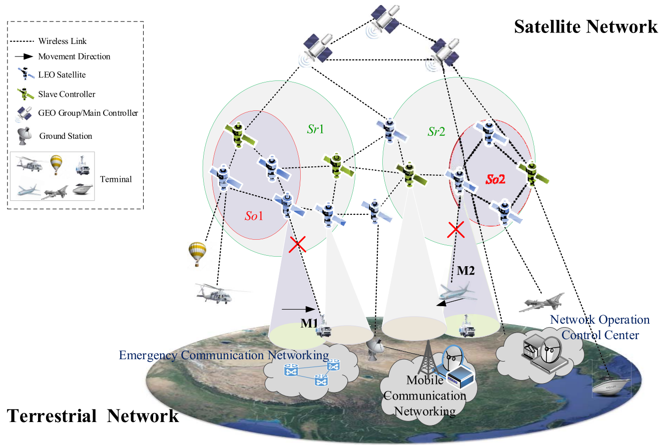


Fig. 1. On-demand controller placement in SDS-TN scenario.

dividing satellite subnets based on redundant coverage, which can optimize the size of satellite subnets while meeting the dynamic coverage demands of terminals, and then optimize the number of controllers. Secondly, in the distributed satellite subnets, we establish the mathematical model for the CPP that takes the networking response latency as the objective function, which can improve the satellite networking latency by optimizing the number and location of controllers. Moreover, we equivalently convert the mathematical model into the capacitated facility location problem (CFLP) and propose an on-demand dynamic approximation algorithm (ODAA) to get the approximate solution of the controller placement. At last, the simulation results show that the proposed ODAA algorithm can optimize the controller placement of SDS-TN, and compared to the most relevant controller placement algorithm, the ODAA algorithm can reduce the networking response latency by approximately 15% and decrease the flow setup time by approximately 25%.

The rest of the paper is organized as follow. SDS-TN scenario and problem statement are described in Section II. Section III reviews the related works. Section IV introduces the on-demand dynamic controller placement modeling. Section V presents the on-demand dynamic approximate algorithm. In Section VI, we conduct experiments to evaluate the proposed controller placement scheme. Finally, we conclude our works in Section VII.

II. SDS-TN Scenario and Problem Statement

As shown in Fig. 1, SDS-TN consists of satellite networks and terrestrial networks. The satellite network includes the satellites with different orbital heights, such as geostationary earth orbit (GEO) group and low earth orbit (LEO) satellites.

In SDS-TN, the control plane and data plane are separated from each other, the control plane can dynamically maintain global network information and is responsible for the control of various forwarding rules. Because the distributed network control architecture can effectively improve network management and control functions, the multi-layer controllers deployed at the satellite orbit layers can support some emergency dynamic networking applications, such as earthquake relief and joint military operations. These service applications require the rapid establishment of a satellite emergency network based on ground terminals coverage requirements to provide low-latency communication services. In this distributed network control architecture where multiple layers of controllers are deployed, the number and location of GEO groups and ground stations are relatively stable. Influenced by geographical and social factors, the GEO group is usually selected as the main controller. Therefore, the key to solving the CPP of SDS-TN is to select appropriate LEO satellites as slave controllers, which can optimize networking response latency between satellite controllers and satellite switches during networking.

In dynamic satellite networking, affected by the topological characteristics of the LEO satellite network constellation, a satellite can establish up to 4 inter satellite links (ISLs) with neighboring satellites, which can affect the matching relationship between the controller satellite and the switch satellite. Therefore, the controller placement scheme in the satellite subnets needs to consider the characteristics of ISLs. As shown in Fig. 1, in order to meet the needs of flexible networking, basic satellite subnets, such as $S_0 1$ and $S_0 2$, need to be established according to the coverage demand. In these basic satellite subnets, it is necessary to select the suitable satellites as controllers, and determine the matching relationship between the controller and the switch to realize the control of the current subnet. If the size of the satellite network and the placement of controllers can be flexibly adjusted, the networking response latency of satellite subnets can be effectively optimized. However, with the relative movement of satellites and terminals, the terminals such as M1 and M2 may be disconnected from the current satellite, and the adjacent satellites cannot provide access services, because these adjacent satellites are not within the coverage area of current subnet. Therefore, the satellite subnet may unable to meet the dynamic coverage demands of the ground terminal.

If the range of the subnet can be expanded in advance, the area surrounded by the red line is expanded to the area surrounded by the green line, such as $Sr1$ and $Sr2$, and controllers are deployed in expanded satellite subnet areas, which can effectively avoid the coverage vulnerability caused by network dynamics. Therefore, the on-demand dynamic placement scheme is more challenging than designing a controller placement scheme based on the entire satellite constellation network. The effectiveness of this on-demand controller placement relies on two key issues:

- How to establish a reliable satellite subnet that can dynamically satisfy the coverage demands of terrestrial terminals. Due to the dynamic impact of network nodes, how to predict the size of the satellite subnet based on the current status information of terminals and satellites will directly affect the coverage performance of the network.
- How to select suitable satellites to deploy controllers to reduce the networking response latency in satellite subnets. Due to the influence of inter-satellite link properties, how to optimize the number and location of the placed controllers according to the current network status

will directly affect the performance of the networking response latency.

III. Related Works

SDN is envisioned as a promising paradigm to enhance the scalability and reliability of control capability for SDS-TN [12], [13], which has the logically centralized and physically distributed network architecture. For emergency communication services, multiple satellites are required to temporarily construct an emergency satellite network to quickly respond to service requests from terminals. In order to increase the flexibility of the satellite network and reduce networking latency, multiple controllers need to be deployed to achieve the network management and control of emergency satellite. Therefore, the problem of controller placement in the satellite network has become a key issue to improve network performance.

A. Controller Placement Problem in SDN Networks

The problem of controller placement has always been an important issue in the study of SDN and it has received extensive attention after it is raised. Heller *et al.* [5] initiated the research of CPP by focusing on two essential questions: how many controllers are needed and where should they go in the topology. Then, the CPP was formulated as a general facility location problem, the metric of the average latency, the worst-case latency, and the bound latency were considered to determine the placement scheme. Motived by this work, more and more controller placement solutions were put forward to satisfy the different requirements of the network, such as shortening the latency [14]–[17], enhancing the reliability [18]–[20], reducing the deployment cost and energy consumption [21]–[23], and optimizing multi-objectives simultaneously [24], [26]. Bari *et al.* [14] first proposed the dynamic controller provisioning problem that required to dynamically adapt the number and location of controllers with changing network conditions. Based on the analysis of switch-to-controller assignment and data synchronization among controllers, the author formulated it as an integer linear program to minimize flow setup time and communication overhead. Aiming to reduce the end-to-end latency, Moazzeni *et al.* [18] presented the clustering-based network partition algorithm to calculate the reliability rate of each subnetwork by taking the load, the number and degree of nodes and the packet loss rate of the link into consideration. In this method, the failure of the controller was detected by the coordinator, then the more appropriate satellite was decided to take over the subnet to undertake the fast recovery scheme though the detection phase which considered reliability and distance between the failed subnet and the assigned new controller. Khorramizadeh and Ahmadi [21] formulated CPP as a location allocation model and proposed the corresponding heuristic algorithms. The controller types, costs and the structure of underlying topologies were considered to minimize the total cost and the fair load distribution function was introduced to solve the location allocation problem and to reduce inter-controller latency. Qin *et al.* [26] provided a proof-of-concept implementation of a multi-controller edge system, and modeled the measure traffic latency and overhead to enhance control capability. It is worth mentioning that network partitioning scheme is introduced to the deployment controller of the large-scale network. The clustering algorithm and cooperative game theory were utilized to divide network into some subnetworks [10], [25] to reduce the end-to-end latency and algorithm complexity.

B. Controller Placement in Software Defined Satellite Networks

Due to the special network topology property of satellite networks, such as the dynamic nature of nodes, the limited on-board processing ability of the satellite and the intermittent interruption of ISLs, the existing research schemes of the terrestrial network cannot be directly applied to satellite space segment networks, but related efforts are being tried to solve the CPP of satellite network. These related solutions can be classified into two main categories according to the geolocation of the deployment controller: The single layer controller placement scheme [7], [12], [27] and the multi-layer controller placement scheme [8], [9], [28], [29].

The single layer controller placement schemes have been investigated, in which controllers are deployed on the terrestrial network or an orbital layer of satellite network. Papa *et al.* [7] proposed a dynamic controller placement scheme for LEO constellation satellite network taking into account user specific metrics, which can guarantee an optimal controller placement by minimizing the average flow setup time. In work [12], SDN and virtualization were enabled into satellite networks for the broadband services, and the controllers were deployed at the ground core network to receive the flexible control of data flows, while the role of satellite is still a bent pipe. This scheme cannot take advantage of satellite networks when designing the control plane. In order to utilize the global network view of the GEO group, Bao *et al.* [27] proposed a novel architecture of the software-defined satellite network, which exploited the GEO group as control plane to support the efficiency and fine-grained control. For the purpose of low networking response latency, controllers were deployed at the LEO layer. However, the above placement strategies cannot comprehensively consider the characteristics of different orbiting satellites when deploying controllers.

The multi-layer controller placement schemes have been identified as a promising approach for improving the reliability and scalability of SDS-TN, in which controllers were deployed at heterogeneous network layers of satellite-terrestrial network with using the mode of master-salve controllers. Wu *et al.* [9] proposed a multi-objective optimization model for the multi-layer control scheme, which took the high dynamic characteristic of satellite network into account to determine the number and location of LEO controllers, and the heuristic algorithm was designed to obtain the feasible solution. To overcome the limitation of slow configuration, inflexible traffic engineering, and coarse-grained quality of service guarantee of the existing satellite communication systems, Zhang *et al.* [28] proposed

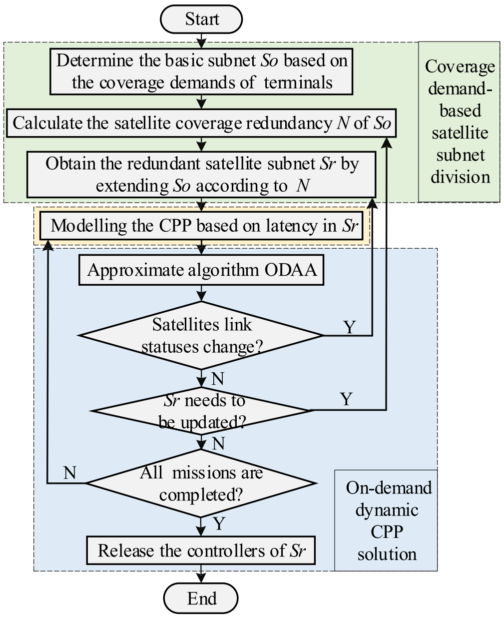


Fig. 2. The overview of the on-demand controller dynamic placement scheme.

a software defined framework where the control plane was made up with the terrestrial controller and the GEO satellites. Nevertheless, the specific controller placement scheme was not put forward. Xu *et al.* [8] proposed a multi-layer control architecture prototype, which consisted of three levels of controllers, super controller at the ground station, the domain controllers at the GEO layer and the slave controllers at the LEO layer. Then, Xu *et al.* [29] presented the slave controller selection strategy to facilitate cost reduction and enhance stability.

C. Overview of On-Demand Controller Placement Scheme

Motivated by the above works, we propose an on-demand dynamic controller placement scheme for CPP of SDS-TN to optimize the number and location of controllers, which can achieve the improvement in the latency performance of the entire network. The biggest difference from the previous work is that the proposed scheme no longer simply selects the default entire constellation system to place the controller, but determines the controller placement scheme according to the actual coverage requirements. In order to design this on-demand controller placement scheme, the concept of subnet division is introduced. Differentiating from the previous subnet division works, we propose a new method for determining redundant subnets based on satellite-to-ground dynamic coverage. In the process of dividing the satellite subnet network, the coverage demand of the dynamic subnet networks need to be guaranteed. When designing the placement scheme of the satellite subnets, the impact of the inter-satellite link characteristic on the placement scheme need to be considered. Therefore, this on-demand dynamic controller placement scheme of SDS-TN needs to be investigated, which is more challenging compared with the existing controller placement schemes. The overview of this scheme is shown in Fig. 2.

Coverage demand-based satellite subnet division: In order to effectively avoid the coverage vulnerabilities caused by the introduction of subnet division, we propose a satellite subnet division scheme based on the redundant coverage, which expands the satellite subnet by predicting the possible handover situation of the terminal. The extended redundancy of the subnet is defined by analyzing the relative movement of the satellite and the terminal.

Modeling of CPP based on latency: In order to solve the problem of controller placement in the satellite subnet, an optimal model of the controller placement scheme is established based on the analysis of the networking response process latency in the satellite distributed control system.

On-demand dynamic CPP solution: We design an approximate algorithm to solve the optimization model. With the dynamic change of satellite link statuses and terminal mission coverage requirements, the size of satellite subnets and the control placement strategy within satellite subnets are dynamically adjusted.

In general, we divide satellite subnet by predicting the redundancy range of satellite subnets, and solve the problem of network coverage vulnerabilities caused by the relative movement of satellites and terminals. At the same time, we optimize the number and location of controllers in the satellite subnet to improve the networking response latency, thus solving the problem of controller selection in SDS-TN.

IV. On-Demand Dynamic Controller Placement Modeling

The on-demand dynamic controller placement scheme is explained in two parts. The first part is the satellite subnet networking method based on the coverage demands of ground terminals. The second part is to select suitable satellites to deploy controllers in the formed satellite subnet.

A. Redundancy-Based Satellite Subnet Division

In order to establish a reliable satellite subnet that can ensure the coverage demand of the satellite subnet to the terminal, where the coverage demand specifically refers to the distribution range of terminals that satellites need to cover, which can be determined by the coverage geometric nature of satellite [30]. We propose a dynamic subnet division method based on coverage redundancy to avoid the issue of coverage vulnerabilities caused by the relative movement of satellites and terminals.

1) Coverage Redundancy Value: The coverage redundancy of satellites is defined as N, which refers to the maximum number of satellites switched by the terminal during a unit window time period of satellite.

The geometric nature of the satellite coverage [30] is used to analyze the dynamic coverage of satellite-to-terminal. As shown in Fig. 3, the geographic altitude of the satellite is H, the satellite running velocity is v_s, the earth radius is R_e, the tangent of satellite to the earth determines the maximum coverage of the satellite, and the geocentric angle is set as α at this moment.

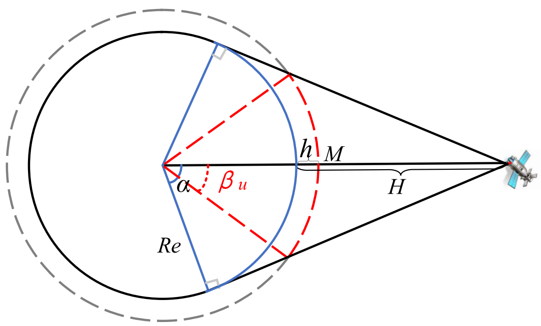


Fig. 3. The coverage geometry of satellite.

According to the geometrical nature of the satellite ground coverage, the arc of satellites rotation α_s in a time T can be calculated by the arc length formula.

First of all, the arc length l_s of the satellite orbiting the orbit at the speed v_s in time T can be calculated by formula (1).

$$l_s = v_s T, \tag{1}$$

At the same time, l_s can also be obtained by the rotation angle α_s of the satellite in time T and the position height of the satellite, as shown in formula (2).

$$l_s = \frac{2\alpha_s}{180°}\pi(R_e + H), \tag{2}$$

Therefore, formula (3) can be derived from formula (1) and formula (2).

$$v_s T = \frac{2\alpha_s}{180°}\pi(R_e + H), \tag{3}$$

According to formula (3), α_s can be calculated by the following formula (4).

$$\alpha_s = \frac{v_s T \cdot 180°}{\pi(R_e + H)}. \tag{4}$$

When $\alpha_s = \alpha$, let ΔT is the maximum coverage time of the satellite-to-terminal, the value of ΔT can be calculated by the following formula (5).

$$\Delta T = \frac{2\alpha\pi(R_e + H)}{v_s \cdot 180°}. \tag{5}$$

When the terminal is any terminal within the coverage of satellite, $\alpha_s \leq \alpha$, then it can be deduced that the coverage time of satellite-to-terminal $T \leq \Delta T$. At this time, ΔT can still represent the coverage longest duration threshold of any terminal. Therefore, ΔT can represent the maximum coverage time of any terminal within the coverage of satellite.

The terminal *M* is a node in the area covered by the satellite, the geographic altitude is h, the velocity is v_u. Similarly, the arc of terminal rotation in a unit time β_u can be calculated by the following formula.

$$\beta_u = \frac{v_u T \cdot 180°}{\pi(R_e + h)}. \tag{6}$$

With relative motion of satellites and terminals, the terminal may exceed the coverage of satellites when the angle difference between α_s and β_u is more than 2α, and the terminal will always maintain a communication connection within the coverage of the current associated satellite when the angle difference between α_s and β_u is less than 2α.

Therefore, the time window threshold value $T_{\max}$ of satellite-to-terminal coverage can be obtained when $\alpha_s - \beta_u = 2\alpha$:

$$T_{\max} = \frac{2\alpha\pi(R_e + h)(R_e + H)}{180°(v_s(R_e + h) - v_u(R_e + H))}. \tag{7}$$

The redundancy coverage indicates the number of satellites possibly switched by terminal at the next moment during a unit satellite-to-ground window time period. Therefore, the redundancy value can be derived from formula (8), where $\lceil\rceil$ indicates rounding up.

$$N = \left\lceil \frac{\Delta T}{T_{\max}} \right\rceil = \left\lceil 1 - \frac{v_u(R_e + H)}{v_s(R_e + h)} \right\rceil. \tag{8}$$

Since *N* is a natural number, the $\frac{\Delta T}{T_{\max}} = 1 - \frac{v_u(R_e+H)}{v_s(R_e+h)} \geq 0$, it can be further deduced that $\frac{v_u(R_e+H)}{v_s(R_e+h)} \in (0, 1)$. Since v_s and v_u are vectors, it can be concluded that the value of *N* is mainly affected by the velocity of the satellite and the terminal. Therefore, the value of *N* can be divided into the following cases.

Case 1: When the terminal velocity is 0, the *N* value is 1. Terminal *M* is a point that is stationary to the ground.

Case 2: The angle between terminal velocity and satellite velocity is an acute angle, $\frac{\Delta T}{T_{\max}} \in (0, 1)$, the value of *N* is taken as 1. The movement direction of the terminal *M* is the same as the movement direction of the satellite.

Case 3: The angle between terminal velocity and satellite velocity is a right angle, $\frac{\Delta T}{T_{\max}} \in (0, 1)$, the value of *N* is 1. The movement direction of the terminal *M* is perpendicular to the movement direction of the satellite.

Case 4: The angle between terminal velocity and satellite velocity is an obtuse angle, $\frac{\Delta T}{T_{\max}} \in (1, 2)$, the value of *N* is taken as 2. The movement direction of the terminal *M* is opposite to the movement direction of the satellite.

2) Redundancy-Based Satellite Subnet Division Method: To ensure the effective coverage of emergency areas, satellite subnets where the terminal located need to meet the possible networks switching requirements at the following moment. The concept of coverage redundancy is introduced to consider the impact of terminal coverage demand changes on the subnet division method.

The specific steps of the satellite subnet division based on coverage redundancy are shown in Fig. 4.

Step 1: According to the distribution of ground terminals, the satellite subnet that meets the basic coverage can be determined. The basic satellite subnet S_0 is determined by the coverage demands of terminals based on the geometric nature of satellites.

Step 2: After obtaining the basic coverage subnet S_0, the *N* values of satellites at the satellite subnet boundary can be calculated according to the status information of satellites and terminals. For the same satellite, the maximum value is chosen as the extended redundancy of this satellite.

Step 3: The coverage range of satellite subnets is expanded based on the obtained satellite expanded redundancy. For the

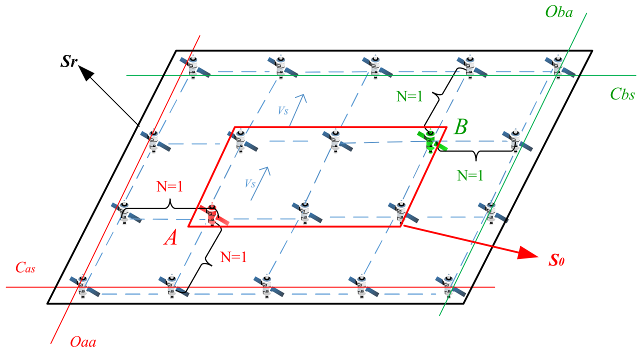


Fig. 4. Redundancy based Satellite subnet division method.

sake of explanation, we choose $N = 1$ as an example. As shown in Fig. 4, O_{aa} is the orbit that interval $N-1$ orbits with satellite A, C_{as} is a circle perpendicular to the A orbital plane and pass through the satellite which has the same orbit and has $N-1$ satellites interval with A. The satellites on the extended boundaries (O_{aa}, C_{as}) include satellites that the terminal of satellite A need to switch to at the next moment.

Step 4: After obtaining the boundary curves of all boundary satellites, all satellites in the area enclosed by these boundary curves form the expanded redundancy subnet area S_r.

In this method, the size of the satellite subnet can be adjusted according to the dynamic demands of ground terminals, which can effectively reduce the network scale and the number of satellites in the subnet, thereby optimizing the number of controllers. At the same time, the current status information of terminals and satellites is used to estimate the possible switching of the network, which allows the expanded satellite subnet to avoid the coverage problem caused by the subnet division.

B. Analysis of Networking Response Latency

As an important indicator of the controller placement problem, latency can effectively measure the effectiveness of the controller placement scheme [14]. Especially for the satellite network, which has the large space-time characteristic of ISLs. Therefore, in the process of selecting suitable satellites to deploy controllers in satellite subnets, we propose the optimization model based on the analysis of networking response latency, and the networking response latency is the time for the network to respond to dynamic changes of nodes.

The network response process of LEO constellation satellite network is shown as Fig. 5. The average network running latency refers to the time consumed to maintain the normal operation of the satellite subnets when the satellite network does not switch, it mainly includes network maintenance time, data flow setup time and control synchronization time. The controller switching time is the time consumed by the network when the switch migrates between different controllers. Affected by the large space-time characteristic and the limitation of on board processing ability of the satellite, the propagation latency and transmission latency are considered when analyzing the networking response latency. The main parameters are listed in Table I.

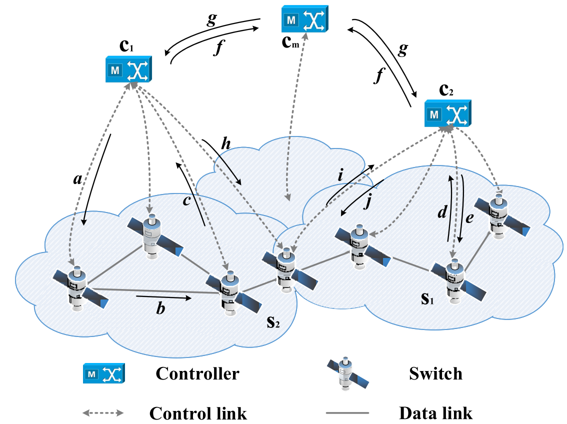


Fig. 5. Response process of distributed control network.

TABLE I
MAIN PARAMETERS USED IN THE SCHEME

Parameter	Definition
C	Set of satellite controllers
V	Set of satellites
p	Switch association matrix in control domain
C'	Set of current satellite controllers
E	Set of ISLs
p'	Switch Association matrix in existing control domain (known)
T_{re}	Interval of algorithm execution
T_s	Interval of network maintenance
d_{ij}	Distance from satellite i to satellite j
F	Switch request routing entries from controller
r	Link transmission rate
r_s	Device port transmission and reception rate
I_m	Network maintenance information
I_p	Packet-In information
I_f	Flow table information
I_{syn}	Controller synchronizes information
I_h	Switch migration request information
I_a	Migration request ack information

(1) Network maintenance time T_m: The process of *a-c* indicates the network maintenance process. Firstly, the controller sends a link detection data packet to the switch satellite managed by it according to link detection protocol. Then, the switch satellite forwards data packet to its neighbor switch satellites. Lastly, the switch satellite that received packet uploads information to the controller connected with it through packet_in method. The network maintenance time can be calculated by the following formula.

$$T_m = \left\lceil \frac{T_{re}}{T_s} \right\rceil \left[\sum_{i \in C} \sum_{j \in V} \left(\frac{2 d_{ij} p_{ij}}{r} + \frac{2 I_m p_{ij}}{r_s} \right) + \sum_{i \in V} \sum_{j \in V, e_{ij} \in E} \left(\frac{d_{ij}}{r} + \frac{I_m}{r_s} \right) \right]. \quad (9)$$

(2) Data flow setup time T_f: The process of *d-e* means the data flow setup process. Take switch S_1 as an example, when a new flow arrives at S_1, S_1 needs to packet_in information to C_2 controller since there is no matching data flow forwarding rule in the S_1 flow table. Then the controller C_2 sends packet_out to S_1 and installs the forwarding rule of flow. The data flow setup time process can be calculated by the following

formula.

$$T_f = \sum_{i\in C}\sum_{j\in V}\left[\frac{2d_{ij}p_{ij}F_{ij}}{r} + \frac{p_{ij}F_{ij}\left(I_p + I_f\right)}{r_s}\right]. \tag{10}$$

(3) Control synchronization time T_{syn}: The process of f-g shows the controller cluster synchronization process, the slave controller sends network information of domain to the master controller C_m, then the master controller sends global network information to the slave controller. The average synchronization time can be calculated by the following formula.

$$T_{syn} = \left\lceil\frac{T_{re}}{T_s}\right\rceil \sum_{j\in C, c_m\neq j}\left(\frac{2d_{c_m j}}{r} + \frac{2|V|I_{syn}}{r_s}\right). \tag{11}$$

Therefore, the average running latency of satellite network is:

$$T_{run} = \frac{1}{|V|}\left(T_m + T_f + T_{syn}\right). \tag{12}$$

(4) Controller switching time T_h: The process of *h*-*j* indicates the switch satellite S_2 migration from the controller C_1 to the controller C_2. Firstly, the network information needs to be synchronized among controller clusters. Then, the original controller sends switching information to migrated satellite switch, and the switch satellite sends a connection request to the new associated controller, and the destination controller replies the request with a confirmation message upon receiving the request. The average latency of state transition caused by switch migration is:

$$T_h = \frac{1}{|V|}\left[\sum_{i\in C'}\sum_{j\in V, p'_{ij}=1, p_{ij}\neq 1}\left(\frac{d_{ij}}{r} + \frac{I_h}{r_s}\right) + \sum_{i\in C} \times \sum_{j\in V, p_{ij}=1, p'_{ij}\neq 1}\left(\frac{2d_{ij}}{r} + \frac{I_h + I_a}{r_s}\right)\right]. \tag{13}$$

(5) Networking response latency: Affected by the dynamic characteristics of satellite networks, when emergency satellite subnets are established to provide communication services, some satellites may be moved in or out the satellite subnet. In order to optimize the management and control capabilities of network, the current placement scheme needs to be updated based on changes in network status. Based on the above analysis of response process of distributed control network, the networking response latency can be obtained by calculating average running latency and controller switching time in satellite subnets.

C. On-Demand Controller Placement Scheme Based on Networking Response Latency Model

In order to meet the requirement of low-latency service in the satellite subnet, the optimization of the number and location of controllers should be considered. Based on the analysis of networking response latency of the satellite subnet, the controller placement scheme can be determined by minimizing the networking response latency, and the following optimization objective function can be established:

$$\min\ T_{run} + T_h. \tag{14}$$

subject to:

$$|C| = \sum_{i\in V} p_{ii} \leq |V|. \tag{15}$$

$$p_{ii} = 1, \forall i \in C. \tag{16}$$

$$p_{ij} = 0, \forall i \notin C, j \in V. \tag{17}$$

$$\sum_{i\in C} p_{ij} = 1, \forall i \in V. \tag{18}$$

$$\sum_{j\in V} F_{ij}p_{ij} \leq U_i, \forall i \in V, j \in V. \tag{19}$$

$$p_{ij} \in \{0,1\}, \forall i \in V, j \in V. \tag{20}$$

The objective function (14) represents the networking response latency of the distributed satellite control system. Constraint (15) indicates that the number of controller satellites is not more than the number of network satellites. Constraint (16) represents that satellite *i* is selected as the controller. Constraint (17) means that there is no control link among switch satellites. Constraint (18) indicates that each switch satellite is only controlled by exactly one controller. Constraint (19) ensures that flow setup requests sent by the switch can be processed by the satellite controller, where U_i represents the maximum capacity of the satellite. Constraint (20) represents that the association relationship of the control link p_{ij} is a binary number, which indicates the matching relationship between the controller and the switch, when satellite *i* is connected to the satellite *j*, $p_{ij} = 1$, otherwise, $p_{ij} = 0$.

V. On-Demand Dynamic Approximation Algorithm

In order to obtain the controller placement scheme of dynamic satellite subnets, we construct an optimization function based on networking response latency, and propose an approximate algorithm to solve this optimization function, which can obtain the stable optimization solution by constraining the validity of the approximate solution. Before introducing the approximation algorithm in detail, we first analyze the concavity and convexity of the objective function (14) to illustrate that the constructed function (14) has the optimal solution.

A. Effectiveness Analysis of the Latency Optimization Function

After the variable *p* is relaxed, the objective function (14) and constraints are expressed as follows:

$$\min\ T_{run} + T_h. \tag{21}$$

subject to:

$$|C| = \sum_{i\in V} p_{ii} \leq |V|. \tag{22}$$

$$p_{ii} = 1, \forall i \in C. \tag{23}$$

$$p_{ij} = 0, \forall i \notin C, j \in V. \tag{24}$$

$$\sum_{i\in C} p_{ij} = 1, \forall i \in V. \tag{25}$$

$$\sum_{j\in V} F_{ij} p_{ij} \le U_i, \forall i \in V, j \in V. \tag{26}$$

$$p_{ij} \ge 0, \forall i \in V, j \in V. \tag{27}$$

Lemma 1: The objective function (14) is a convex function after the linear relaxation of *p*.

Based on the above analysis, it can be seen that the networking response latency function T_{res} of network is a function of *p*, this function can be sorted out as follows:

$$\begin{aligned} T_{res}(p) &= T_{run}(p) + T_h(p) \\ &= \begin{cases} k_1 p_{ij} + \delta(i=j) \\ k_2 p_{ij} + \delta(i \ne j) \end{cases} \end{aligned} \tag{28}$$

where:

$$\begin{aligned} k_1 = &\sum_{i\in V}\sum_{j\in V}\frac{1}{|V|} \\ &\times \left[\frac{T_{re}}{T_s}\left(\frac{2d_{c_m j}}{r} + \frac{2I_s}{r_s}\right) + \frac{\left(I_p + I_f\right)F_{ij}}{r_s}\right. \\ &\quad \left. + \frac{\frac{2T_{re}}{T_s}I_m + \left(1 - P'_{ij}\right)I_a + \left(1 - 2P'_{ij}\right)I_h}{r_s}\right]. \end{aligned} \tag{29}$$

$$\begin{aligned} k_2 = &\sum_{i\in V}\sum_{j\in V}\frac{1}{|V|} \\ &\times \left[\frac{\left(\frac{2T_{re}}{T_s} + 2 + 2F_{ij} - 3P'_{ij}\right)d_{ij}}{r}\right. \\ &\quad + \frac{\left(I_p + I_f\right)F_{ij}}{r_s} \\ &\quad \left. + \frac{\frac{2T_{re}}{T_s}I_m + \left(1 - P'_{ij}\right)I_a + \left(1 - 2P'_{ij}\right)I_h}{r_s}\right]. \end{aligned} \tag{30}$$

$$\begin{aligned} \delta = &\sum_{i\in V}\sum_{j\in V}\frac{1}{|V|} \\ &\times \left[q^{e_{ij}}\left(\frac{2d_{ij}}{r} + \frac{2I_m}{r_s}\right) + \frac{\left(3{P'_{ij}}^2 - 2P'_{ij}\right)d_{ij}}{r}\right. \\ &\quad \left. + \frac{\left(2{P'_{ij}}^2 - P'_{ij}\right)I_h + \left({P'_{ij}}^2 - P'_{ij}\right)I_a}{r_s}\right]. \end{aligned} \tag{31}$$

$$q^{e_{ij}} = \begin{cases} 1, & e_{ij} \in E \\ 0, & e_{ij} \notin E. \end{cases} \tag{32}$$

In the above formula, *p* is variable, the input information parameters are the constant, and the objective function is linear. When $p = E$, *p* is an effective inter-satellite link that simultaneously satisfies the constraints (15)-(20), and the feasible domain is bounded and solvable. Because the constraints (15)-(20) are linear and discrete, and the feasible domain is bounded. Therefore, after the linear relaxation, there is the feasible solution in the feasible domain and the objective function is a convex function.

B. On-Demand Dynamic Approximate Algorithm

Since the objective function is linear and can be relaxed as a convex function, according to the convex optimization theory, the optimal solution of the relaxation problem can be solved. However, this CPP is a non-deterministic polynomial hard problem [5], it is hard to obtain an optimal solution in the polynomial time for high computational complexity. Therefore, we try to achieve an approximate solution using an approximate algorithm, and the objective function is transformed into the optimization function of CFLP.

$$\min \quad \sum_{i\in V} f_i y_i + \sum_{i\in V}\sum_{j\in V} c_{ij} x_{ij} \tag{33}$$

subject to:

$$\sum_{i\in V} x_{ij} = 1, \forall j \in V. \tag{34}$$

$$\sum_{j\in V} x_{ij} \le u_i y_i, \forall i \in V, j \in V. \tag{35}$$

$$x_{ij} \le y_i, \forall i \in V, j \in V. \tag{36}$$

$$x_{ij}, y_i \in \{0,1\}, \forall i \in V, j \in V \tag{37}$$

where:

$$y_i = p_{ii}, x_{ij} = p_{ij}, \forall i \in V, j \in V\ . \tag{38}$$

$$f_i = \frac{2T_{re}}{|V|r} d_{c_m j}\ . \tag{39}$$

$$c_{ij} = \frac{1}{|V|}\left[\frac{\left(\frac{2T_{re}}{T_s} + 2F_{ij} - 3{p'}_{ij} + 2\right)\mathrm{d}_{ij}}{r}\right]. \tag{40}$$

The objective function (33) indicates that controllers are placed with a minimum networking response latency cost, x_{ij} indicates whether there is an association between controller *i* and switch *j*, y_i represents whether satellite *i* is a controller. Constraint (34) indicates that each switch satellite is associated with a controller satellite. Constraint (35) indicates that the number of switch satellites associated with controller satellites is less than the number of ISLs, where u_i represents the number of ISLs. Constraint (36) shows that the number of switch satellites is less than the number of controller satellites. Constraint (37) means that the variables x_{ij} and y_i are binary numbers.

For the CFLP, approximate solutions can be obtained by using approximate algorithms. First, we perform linear relaxation on the constraints of the objective function, then the objective function and the constraint condition are dually converted.

$$\max \quad \sum_{j\in V} \alpha_j \tag{41}$$

subject to:

$$\sum_{j\in V} \max\{\alpha_j - c_{ij}\} \le f_i, \forall i \in V. \tag{42}$$

$$\alpha_j \ge 0, \forall j \in V \tag{43}$$

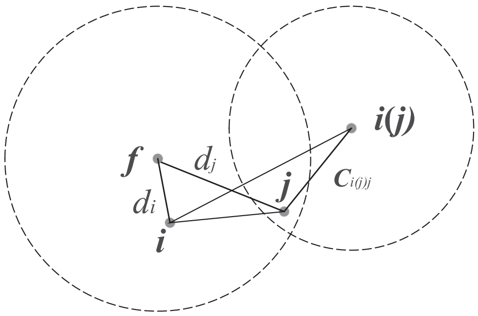


Fig. 6. Triangle inequality.

where, α_j is equivalent to the apportionment of the total cost of coverage by each switch. Then, we mainly prove the effectiveness of the problem conversion through Lemma 2.

Lemma 2: The objective function (33) is a metric facility location problem.

In (33), f_i represents the network latency generated by the controller, including the node synchronization time among controllers. c_{ij} represents the associated latency of switches, including the time cost between the switch and the controller and the time cost among the switches. According to the basic properties of the CFLP, c_{ij} needs to satisfy non-negative, symmetry and triangular inequality [31].

(1) Non-negative: c_{ij} indicates the connection cost of the switch. p'_{ij} is a binary number. F_{ij} is non-negative. If T_{re}/T_s is greater than 0, the first part of c_{ij} is non-negative. There are two cases for the second part.

Case 1: When $p'_{ij} = 0$, the value of (1-2p'_{ij})I_h is non-negative, the second part of c_{ij} is non-negative.

Case 2: When $p'_{ij} = 1$, the configuration relationship between *i* and *j* is unchanged, I_h is 0, and the value of (1-2p'_{ij})I_h is 0, the second part of c_{ij} is non-negative.

Therefore, c_{ij} satisfies the property of non-negativeity.

(2) Symmetry: Since $d_{ij} = d_{ij}, F_{ij} = F_{ji}, p'_{ij} = p'_{ji}, c_{ij} = c_{ji}$ can be obtained from formula (40).

Therefore, c_{ij} satisfies the property of symmetry.

(3) Triangle inequality: As shown in Fig. 6, for $\forall 1 \leq j < i \leq k$, it can be proved that $\alpha_i \leq r_{j,i} + d_i + d_j$ in the following two cases.

Case 1: When $\alpha_j = \alpha_i$, it is obviously satisfied $\alpha_i \leq r_{j,i} + d_i + d_j$.

Case 2: When $\alpha_j < \alpha_i$, because *j* is not connected to another controller node *i(j)* at time *t*, $t = \alpha_i - \varepsilon \leq c_{i(j)j}$, $\alpha_i = t - \varepsilon \leq c_{i(j)i} \leq c_{i(j)j} + d_i + d_j = r_{j,i} + d_i + d_j$.

Therefore, it is proved that the optimization problem of the objective function (33) is a metric facility location problem.

The specific on-demand dynamic approximation algorithm is shown in Algorithm 1. Firstly, the satellite subnet S_0 is obtained according to the terminal set, and the satellite subnet S_r is determined according to the redundancy N of the edge satellite nodes. Then, the controller node is established according to the objective function in the subnet S_r. As the number of controllers increases, the association between the switch and the controller is re-evaluated until the optimal feasible solution result is output. Event 1 indicates that when

Algorithm 1 On-Demand Dynamic Approximate Algorithm (ODAA)

Input: Topology G; Terminal set T_u; $\{f_i, c_{ij}, u_i\}$; $\{d_{ij}\}$
Output: Feasible integer solution (x,y)

According to T_u calculate N to obtain satellite subnet S;
Initialization $C \leftarrow V$, $F \leftarrow V$, $U \leftarrow C$, $\hat{F} \leftarrow \phi$, $f_i \leftarrow \left(1 - \frac{1}{u_i}\right) f_i$, $c_{ij} \leftarrow \frac{f_i}{u_i} + c_{ij}$, $(x, y) \leftarrow 0$, $t \leftarrow 0$.
while $U \neq \phi$ **do**
  for each $j \in C$ **do**
    $\alpha_j = t$, t is time variable;
    the following two events in any order;
    event1:
    if $\alpha_j = c_{ij}$ for $j \in U, i \in \hat{F}$ **then**
      $U = U \backslash \{j\}, i(j) = i$, // switch j connect to i;
    end if
    event2:
    if $\sum_{j \in U} \max\{\alpha_j - c_{ij}, 0\} + \sum_{j \in C \backslash U} \max\{c_{i(j)j} - c_{ij}, 0\} = f_i$, for $i \in F \backslash \hat{F}$ **then**
      $\hat{F} = \hat{F} \cup \{i\}$;
      $N_i^1 = \{\alpha_j \geq c_{ij}, j \in U\}, N_i^2 = \{c_{i(j)j} \geq c_{ij}, j \in C \backslash U\}$;
      $N_i = N_i^1 \cup N_i^2, U = U \backslash N_i^1$;
    end if
    for each $j \in N_i$ **do**
      update $i(j) = i$, // switch j connect to i;
    end for
    for each $i \in \hat{F}$ **do**
      $y_i = 1$;
    end for
    for each $j \in C$ **do**
      $x_{ij} = 1$;
    end for
  end for
end while
return (x, y)

the minimum networking response latency of the system is matched, the switch *j* is associated with the controller *i*. Event 2 indicates that if there is a lower networking response latency when a new controller is added, the new controller is turned on to complete the re-association between the switch and the controller.

Then, we use the Lemma 3 to analyze the approximation of the on-demand dynamic approximate algorithm.

Lemma 3: The CPP corresponding to (33) has an approximation algorithm with the approximation of 2.0.

According to the work [32], the CPP is summarized into a linear-cost facility location problem (LCFLP) in the soft capacitated facility location problem (SCFLP). The cost of the controller facility *i* is:

$$f_i(k) = \begin{cases} \left(1 + \frac{k-1}{u_i}\right) f_i & if\ k \geq 1 \\ 0 & if\ k = 0 \end{cases} \tag{44}$$

when $k \geq 1$,

$$\left\lceil \frac{k}{u_i} \right\rceil f_i \leq \left(1 + \frac{k-1}{u_i}\right) f_i \leq \left\lceil \frac{k}{u_i} \right\rceil f_i. \tag{45}$$

Therefore, the SCFLP is the (2, 1)-metric problem of the LCFLP (a, b, c). Then the problem can be equivalently converted to the corresponding UCFLP (b, a+c).

$$\begin{cases} b = \left(1 - \frac{1}{u_i}\right) f_i \\ a + c = \frac{f_i}{u_i} + c_{ij}. \end{cases} \tag{46}$$

According to the *Inference* 2 of the work [31], the (1, 2)-approximation algorithm of LCFLP (a, b, c) can be obtained by running the above algorithm in UCFLP (b, a+c). Then, according to the Theorem 3 of the work [32], the above SCFLP solving algorithm ODAA is an approximation algorithm with an approximation of 2.0.

VI. Performance Evaluation

The satellite tool kit (STK) is used to create network scenarios and simulate the ISLs of walker δ constellation, then STK and MATLAB are combined to complete the algorithm simulation. When the ISLs are established, there are 2-4 ISLs between the satellite and its adjacent satellites. These four links include two intra plane ISL between two satellites located in the same orbit and two inter plane ISL between two satellites located in the left and right adjacent orbits. These ISLs are intermittent, when the satellite enters the polar region, the inter plane ISL between the satellite and the satellite on its adjacent orbit is broken [33]. Network missions are generated by mobile terminals on the terrestrial, and the size and basic distribution of traffic are mainly [34]. The satellite subnet area ratio γ is the ratio of satellite subnets to the entire network, the simulation time is one day that starts from 4/10/2018 8 AM GMT to 4/11/2018 7 AM GMT. The main parameters are listed in Table II.

In terms of satellite subnet coverage rate, number of controllers, average flow setup time and average networking response latency, we evaluate the performance of the proposed algorithm ODAA, and compare our algorithm with the accelerated particle swarm optimization algorithm (ASPO) [9] and the optimal search algorithm (OptSearch) that can obtain the optimal solution by traversing all feasible solutions.

A. Coverage Redundancy of Satellite Subnets

In order to verify that the proposed subnet division method based on coverage redundancy can effectively guarantee the coverage demands of terrestrial terminals, the coverage demands, we compared the coverage of satellite subnets with different coverage redundancy values under different terrestrial coverage area ratio. The larger the value of N, the larger the size of the satellite subnet, and more controllers need to be deployed. In order to optimize the number of controllers, the satellite subnet needs to be dynamically adjusted according to actual coverage demands. The coverage demands can be calculated by the distribution range of the terminal and the coverage geometric nature of satellite. The range of coverage demand ratio belongs to [0, 1], which refers to the ratio of the mission terminal distribution area to the surface area of entire earth. $N = 0$ means the basic satellite subnet, that is, the satellite subnet is not expanded based on redundancy. $N = 1$ means that the subnet is expanded with redundancy value 1, $N = 2$ means that the subnet is expanded with redundancy value 2. $N = A$ means that the satellite subnet is adaptively expanded according to the dynamic changes of the ground coverage demands.

TABLE II
Main Parameters Used in the Scheme

Parameter	Value
Constellation	48/8/1
Satellite orbit height	1414 Km
Orbit inclination	52°
Step size	60s
Interval of algorithm execution T_{re}	1s
Interval of network maintenance T_s	0.9s
Link transmission rate r	3×10^8 m/s
Device port transmission and reception rate r_s	6 Mbps
Network maintenance information I_m	26 bytes
Packet-In information I_p	32 bytes
Flow table information I_f	56 bytes
Controller synchronizes information I_{syn}	(12+24 e_{ij}) bytes
Switch migration request information I_h	18 bytes
Migration request ack information I_a	13 bytes

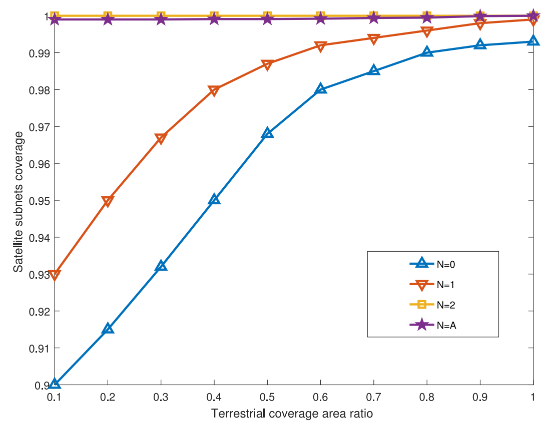


Fig. 7. Satellite subnet coverage in different terrestrial coverage area.

As shown in Fig. 7, as the coverage redundancy value increases, the coverage rate of the satellite subnet gradually increases. When $N = 0$, the subnet division method cannot guarantee the effective coverage of the satellite network to the terminal demand. In other words, the satellite subnet division method has the problem of network coverage vulnerabilities when network dynamics are not considered. As the expansion redundancy value of the satellite subnet increases, the coverage demand change caused by the relative movement of the terminal and the satellite is considered in the process of expanding the satellite subnet, and the coverage of the subnet is effectively guaranteed. When $N = A$, the coverage of the satellite subnet can be guaranteed. Therefore, it is necessary to determine the expansion redundancy of the satellite subnet according to the coverage demand of the terminal. It can be concluded that the proposed subnet division method can effectively ensure the adaptability of the satellite subnet to the coverage demand.

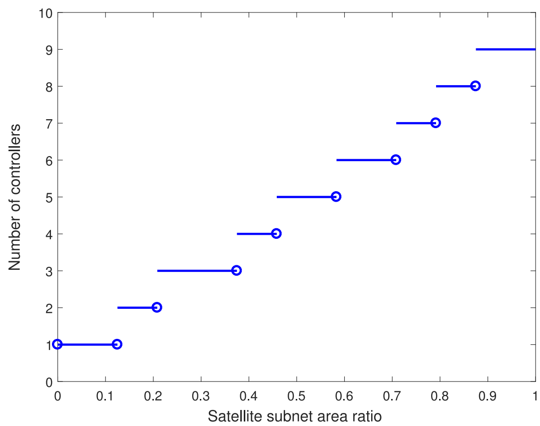


Fig. 8. Number of controllers in different subnet ratios.

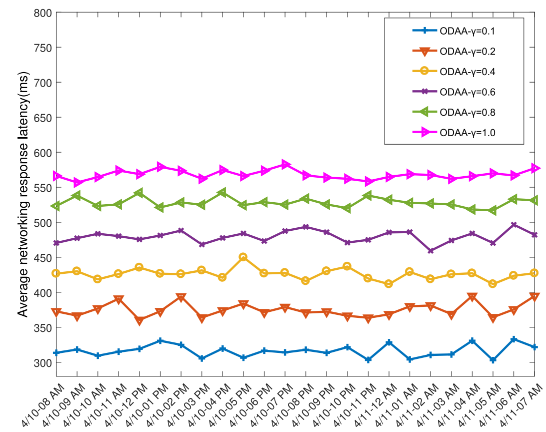


Fig. 9. Average networking response latency of different subnet ratios.

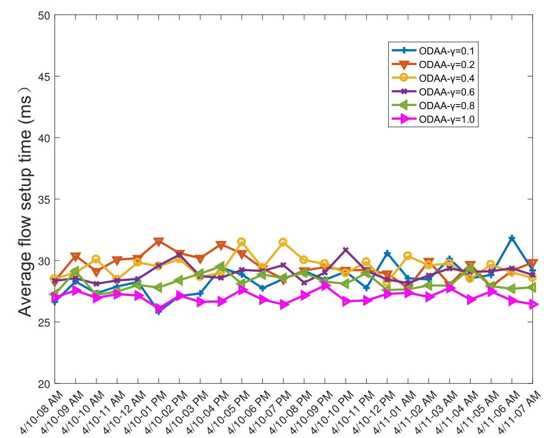


Fig. 10. Average flow setup time of different subnet ratios.

B. Number of Controllers

The number of controllers can reflect the occupation of network resources by the controller placement scheme. Although the larger number of controllers means the higher management and control capabilities, which also takes up too much network resources and means higher computational overhead.

As shown in Fig. 8, as the proportion of satellite subnets continues to increase, the satellite subnet coverage area is expanded, and the number of controllers increases accordingly. When $\gamma = 0.1$, the ratio of the satellite controller and the satellite switch in the subnet is about 1: 4, that is, one control is associated with four switch satellites, which meets the characteristics of ISLs. When the satellite subnet reaches full coverage, about 9 controllers are required to manage the entire network, and the ratio of satellite controllers to satellite switches is slightly less than 1: 4, which satisfies the characteristics of inter-satellite links. Based on the above analysis, in order to meet the coverage requirements of the controller for the switch, the larger the satellite subnets, the more controllers need to be placed. It can be seen that compared with the controller placement scheme based on the entire network, the subnet division method based on coverage demand can deploy controllers according to the network coverage demands, which can optimize the number and location of controllers while meeting network coverage demands.

C. Networking Response Latency

In the distributed control network, the network latency caused by the dynamic deployment of the controller is not only related to the number of controllers, but also to the location where controllers are placed. A reasonable deployment solution can respond to terminal requests in the shortest time. This experiment mainly verifies that the ODAA algorithm can optimize the networking response latency in different size of satellite subnets, and ensure the associated configuration of the switch and the controller to maintain the average flow setup time.

Fig. 9 and Fig. 10 show the change of networking average networking response latency and average flow setup time based on the ODAA algorithm with the different satellite subnet ratio. As the proportion of satellite subnets increases, the network latency increases approximately proportionally, but the average flow setup time is maintained within 25ms-32ms. Because the impact of inter satellite link characteristics on the controller placement is considered, a controller can be associated with up to four switches. As the range of satellite subnets expands, more controllers need to be placed to meet the association with the switch. It can be obtained from formula (14) that the networking response latency increases accordingly. Constrained by characteristics of the inter-satellite link, a satellite node where the controller is placed is associated with up to four satellites of the switching node. The average distance between the switch satellite and the controller satellite is about one hop, and the average flow setup time can be obtained relatively stably from formula (10). The controller placement scheme can be updated dynamically during the simulation time of one day, and the networking response latency and the flow setup time can be kept relatively stable. Because the ODAA algorithm can dynamically adjust the number and location of controllers as the coverage requirements of the ground network change. Therefore, the controller placement algorithm based on the subnet division scheme can optimize the networking response latency and average flow setup time, and optimize the latency of the satellite network controller placement.

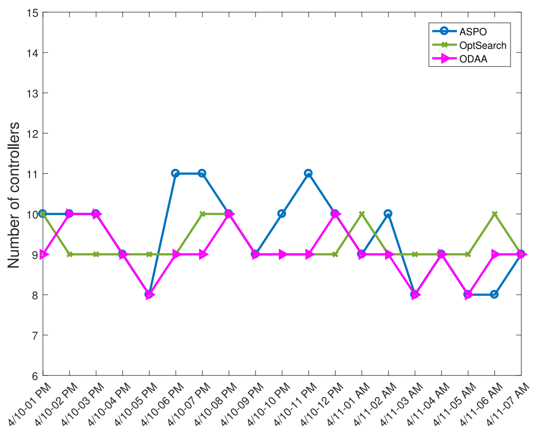


Fig. 11. Number of controllers of differernt algorithms.

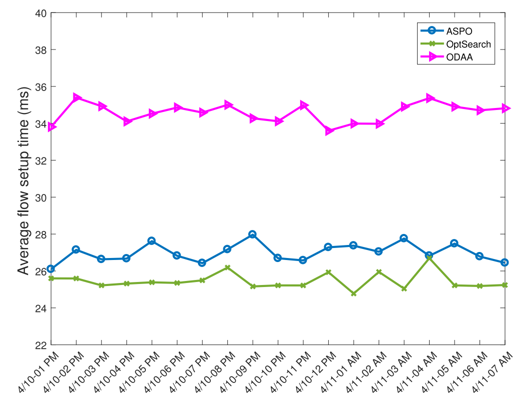


Fig. 12. Average flow setup time of differernt algorithms.

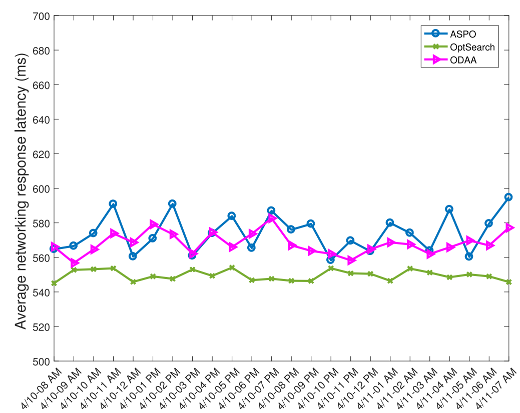


Fig. 13. Average networking response latency of differernt algorithms.

D. Analysis of ODAA Algorithm

In order to verify the effectiveness of the proposed approximation algorithm, based on the entire network system, the proposed algorithm ODAA is compared with the classical heuristic algorithm ASPO and the optimal search algorithm OptSearch.

As shown in Fig. 11, in terms of the number of controllers, three algorithms can effectively obtain the controller to be placed, and the number of controllers can be dynamically updated as the state of network changes. The number of controllers obtained by the OptSearch algorithm is relatively stable. Compared with the ASPO algorithm, the number of controllers obtained by the ODAA algorithm is more inclined to the OptSearch algorithm. Compared to the OptSearch algorithm that has the high algorithm complexity $O(n^2 \cdot 2^n)$, the ODAA algorithm can obtain an approximate solutions and reduce the algorithm complexity to $O(n^2)$. Although the ASPO algorithm can reduce the network complexity to $O(nm)$, this algorithm is a typical heuristic algorithm and cannot guarantee the validity of the solution. In contrast to ASPO, our ODAA as an approximation algorithm has considered finding a feasible solution that is near-optimal. In addition, as the number of satellites in the constellation increases, the running time of the ODAA will not increase significantly. The reason is that ODAA can effectively optimize the size of satellite subnets to reduce the number of controllers, which can reduce the running time of ODAA through narrowing the scale of satellite subnets. Moreover, the approximate algorithm is proposed to obtain the number and location of controllers, which can further reduce the running time of the algorithm. Based on the above analysis, the proposed algorithm can effectively optimize the number of controllers while reducing the computation time complexity.

As shown in Fig. 12, in terms of the average flow setup time, the results obtained by the ODAA algorithm are closer to the optimal solution of the OptSearch algorithm. Compared with the ASPO algorithm, the ODAA algorithm can reduce the average flow setup time by approximately 25%. The reason is that the impact of inter satellite link characteristics on the controller placement scheme is considered when designing the ODAA algorithm, that is, the number of ISLs, its value changes within 2-4 according to the change of the satellite network topology. Therefore, when the matching relationship between the controller and the switch is obtained, the ODAA algorithm can optimize the matching relationship according to the actual number of ISLs, and the position where the controller is placed can be optimized according to the dynamic change of the satellite network topology.

As shown in Fig. 13, in terms of the average networking response latency, the results obtained by the ODAA algorithm tend to the optimal solution of the OptSearch algorithm, and compared with the ASPO algorithm, the ODAA algorithm can reduce the average response latency by approximately 15%. The main reason is that the ODAA approximation algorithm can optimize the number and location of the controller, which can obtain a placement solution that approaches the optimal solution when optimizing the objective function. At the same time, the solution can constrain the stability of the solution. Therefore, when optimizing the networking response latency, compared with the heuristic algorithm, the ODAA approximation algorithm can optimize the networking response latency by optimizing the number and location of controllers, and improve the latency performance of the controller placement scheme.

VII. Conclusion

For the CPP of SDS-TN, we propose a on-demand dynamic controller placement scheme. This scheme can flexibly adjust the size of satellite subnets according to the coverage demand of terminals, which can optimize the size of satellite subnets while ensuring the terminal coverage demand, thereby reducing the number of controllers. Then, the networking response latency is selected as the optimization objective function to improve the latency of satellite networking when designing the controller placement scheme for satellite subnets. In addition, an approximate solution algorithm is designed to solve the objective function, and the optimal controller placement scheme is obtained by optimizing the networking response latency of the satellite subnet. The experiment results show that the proposed algorithm can effectively improve the response latency of satellite networking by optimizing the number and location of controllers. In this paper, we focus on the on-demand dynamic controller placement scheme based on approximate algorithm to improve the latency of network control and management of SDS-TN. In future work, we will try to solve the problem of reliable routing based on this distributed networking control architecture.

REFERENCES

[1] Q. Bi, "Ten trends in the cellular industry and an outlook on 6G," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 31–36, Dec. 2019.

[2] X. Zhu, C. Jiang, L. Kuang, N. Ge, S. Guo, and J. Lu, "Cooperative transmission in integrated terrestrial-satellite networks," *IEEE Netw.*, vol. 33, no. 3, pp. 204–210, May/Jun. 2019.

[3] J. Liu, Y. Shi, L. Zhao, Y. Cao, W. Sun, and N. Kato, "Joint placement of controllers and gateways in SDN-enabled 5G-satellite integrated network," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 2, pp. 221–232, Feb. 2018.

[4] T. Li, H. Zhou, H. Luo, and S. Yu, "SERvICE: A software defined framework for integrated space-terrestrial satellite communication," *IEEE Trans. Mobile Comput.*, vol. 17, no. 3, pp. 703–716, Mar. 2018.

[5] B. Heller, R. Sherwood, and N. McKeown, "The controller placement problem," in *Proc. 1st Workshop Hot Topics Softw. Defined Netw.*, 2012, pp. 7–12.

[6] J. Lu, Z. Zhang, T. Hu, P. Yi, and J. Lan, "A survey of controller placement problem in software-defined networking," *IEEE Access*, vol. 7, pp. 24290–24307, 2019.

[7] A. Papa, T. De Cola, P. Vizarreta, M. He, C. M. Machua, and W. Kellerer, "Dynamic SDN controller placement in a LEO constellation satellite network," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, 2018, pp. 206–212.

[8] S. Xu, X.-W. Wang, and M. Huang, "Software-defined next-generation satellite networks: Architecture, challenges, and solutions," *IEEE Access*, vol. 6, pp. 4027–4041, 2018.

[9] S. Wu, X. Chen, L. Yang, C. Fan, and Y. Zhao, "Dynamic and static controller placement in software-defined satellite networking," *Acta Astronautica*, vol. 152, pp. 49–58, Nov. 2018.

[10] J. Liao, H. Sun, J. Wang, Q. Qi, K. Li, and T. Li, "Density cluster based approach for controller placement problem in large-scale software defined networkings," *Comput. Netw.*, vol. 112, pp. 24–35, Jan. 2017.

[11] K. Yang, B. Zhang, and D. Guo, "Controller and gateway partition placement in SDN-enabled integrated satellite-terrestrial network," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Shanghai, China, 2019, pp. 1–6.

[12] L. Bertaux *et al.*, "Software defined networking and virtualization for broadband satellite networks," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 54–60, Mar. 2015.

[13] Z. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 101–109, Jul. 2017.

[14] M. F. Bari *et al.*, "Dynamic controller provisioning in software defined networks," in *Proc. 9th Int. Conf. Netw. Serv. Manag. (CNSM)*, Zurich, Switzerland, 2013, pp. 18–25.

[15] G. Wang, Y. Zhao, J. Huang, and Y. Wu, "An effective approach to controller placement in software defined wide area networks," *IEEE Trans. Netw. Service Manag.*, vol. 15, no. 1, pp. 344–355, Mar. 2018.

[16] M. T. I. Ul Huque, W. Si, G. Jourjon, and V. Gramoli, "Large-scale dynamic controller placement," *IEEE Trans. Netw. Service Manag.*, vol. 14, no. 1, pp. 63–76, Mar. 2017.

[17] A. Blenk, A. Basta, J. Zerwas, M. Reisslein, and W. Kellerer, "Control plane latency with SDN network hypervisors: The cost of virtualization," *IEEE Trans. Netw. Service Manag.*, vol. 13, no. 3, pp. 366–380, Sep. 2016.

[18] S. Moazzeni, M. R. Khayyambashi, N. Movahhedinia, and F. Callegati, "On reliability improvement of software-defined networks," *Comput. Netw.*, vol. 133, no. 14, pp. 195–211, 2018.

[19] B. P. R. Killi and S. V. Rao, "Optimal model for failure foresight capacitated controller placement in software-defined networks," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1108–1111, Jun. 2016.

[20] M. Tanha, D. Sajjadi, and J. Pan, "Enduring node failures through resilient controller placement for software defined networks," in *Proc. IEEE Global Commun. Conf. (Globecom)*, Washington, DC, USA, 2016, pp. 1–7.

[21] M. Khorramizadeh and V. Ahmadi, "Capacity and load-aware software-defined network controller placement in heterogeneous environments," *Comput. Commun.*, vol. 129, pp. 226–247, Sep. 2018.

[22] A. Ruiz-Rivera, K.-W. Chin, and S. Soh, "GreCo: An energy aware controller association algorithm for software defined networks," *IEEE Commun. Lett.*, vol. 19, no. 4, pp. 541–544, Apr. 2015.

[23] A. Sallahi and M. St-Hilaire, "Optimal model for the controller placement problem in software defined networks," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 30–33, Jan. 2015.

[24] C. Xu, Z. Xiong, Z. Han, G. Zhao, and S. Yu, "Link reliability-based adaptive routing for multilevel vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11771–11785, Oct. 2020.

[25] K. S. Sahoo, S. Sahoo, A. Sarkar, B. Sahoo, and R. Dash, "On the placement of controllers for designing a wide area software defined networks," in *Proc. IEEE Region 10 Conf. (TENCON)*, 2017, pp. 3123–3128.

[26] Q. Qin, K. Poularakis, G. Iosifidis, and L. Tassiulas, "SDN controller placement at the edge: Optimizing delay and overheads," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Honolulu, HI, USA, 2018, pp. 684–692.

[27] J. Bao, B. Zhao, W. Yu, Z. Feng, C. Wu, and Z. Gong, "OpenSAN: A software-defined satellite network architecture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, pp. 347–348, 2014.

[28] Z. Zhang, B. Zhao, Z. Feng, W. Yu, and C. Wu, "MSN: A mobility-enhanced satellite network architecture: Poster," in *Proc. 22nd Annu. Int. Conf. Mobile Comput. Netw.*, 2016, pp. 465–466.

[29] S. Xu, X. Wang, B. Gao, M. Zhang, and M. Huang, "Controller placement in software-defined satellite networks," in *Proc. 14th Int. Conf. Mobile Ad-Hoc Sens. Netw.*, Shenyang, China, 2018, pp. 146–151.

[30] S. Cakaj, B. Kamo, A. Lala, and A. Rakipi, "The coverage analysis for low earth orbiting satellites at low elevation," *Int. J. Adv. Comput. Sci. Appl.*, vol. 5, no. 6, pp. 359–382, 2014.

[31] M. Mahdian, Y. Ye, and J. Zhang, "Approximation algorithms for metric facility location problems," *SIAM J. Comput.*, vol. 36, no. 2, pp. 411–432, 2006.

[32] M. Mahdian, Y. Ye, and J. Zhang, *A 2-Approximation Algorithm for the Soft-Capacitated Facility Location Problem* (Lecture Notes in Computer Science), vol. 2764. Heidelberg, Germany: Springer, 2003, pp. 129–140.

[33] J. Wang, L. Li, and M. Zhou, "Topological dynamics characterization for LEO satellite networks," *Comput. Netw.*, vol. 51, no. 1, pp. 43–53, 2007.

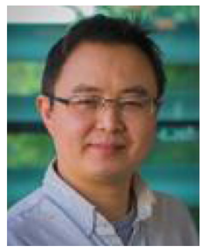
[34] X. Jia, T. Lv, F. He, and H. Huang, "Collaborative data downloading by using inter-satellite links in LEO satellite networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1523–1532, Mar. 2017.

Zhenzhen Han is currently pursuing the Ph.D. degree with the Information and Communication Engineering, Future Networks Research Institute, Chongqing University of Posts and Telecommunications, China. Her research interests are on satellite-terrestrial network, wireless network routing, and software-defined networking.

Chuan Xu (Member, IEEE) received the Ph.D. degree in control theory and engineering from Chongqing University, China, in 2012. He is currently a Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications. He was/is involved in several Chinese Government Projects, including 973 projects and NSFC projects. His research interests are in the areas of wireless communication, software-defined networking, and Internet of Things.

Zhengying Xiong is currently a Research Staff Member with the Future Networks Research Institute, Chongqing University of Posts and Telecommunications, China. Her research interests include the routing in wireless sensor networks and vehicular networks.

Guofeng Zhao (Member, IEEE) received the M.E. degree from Northwestern Polytechnic University, in 1996, and the Ph.D. degree from Chongqing University, in 2003. He currently serves as a Professor with the School of Communication Engineering and the Director of the Research Center of Future Internet, Chongqing University of Posts and Telecommunications. He has undertaken more than 20 projects or program, including the National Basic Research Program of China and the Natural Science Foundation of China. His research interests include the future Internet, the mobile Internet, network management, and network security.



Shui Yu (Senior Member, IEEE) is a Professor of the School of Computer Science, University of Technology Sydney, Australia. He has published two monographs and edited two books, more than 280 technical papers, including top journals and top conferences. He initiated the research field of networking for big data in 2013. His H-index is 46. His research interest includes security and privacy, networking, big data, and mathematical modeling. He is currently serving the editorial boards of IEEE COMMUNICATIONS SURVEYS AND TUTORIALS as Area Editor, *IEEE Communications Magazine*, IEEE INTERNET OF THINGS JOURNAL, IEEE COMMUNICATIONS LETTERS, IEEE ACCESS, and IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS. He is a member of AAAS and ACM, and a Distinguished Lecturer of IEEE Communication Society.