IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 2, 2019, accepted February 25, 2019, date of publication February 28, 2019, date of current version March 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902134

Architecture and Network Model of Time-Space Uninterrupted Space Information Network

ZHICHENG QU[1,2], (Student Member, IEEE), GENGXIN ZHANG[1,2], TAO HONG[1,2], HAOTONG CAO[2], (Student Member, IEEE), AND WEI ZHANG[3]

[1]National Engineering Research Center, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]Key Lab of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[3]Department of Electronic and Optical Engineering, Space Engineering University, Beijing 100000 China

Corresponding author: Gengxin Zhang (zgx@njupt.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 91738201, Grant 91338201, and Grant 61801445, and in part by the Open Research Fund of the National Mobile Communications Research Laboratory, Southeast University, under Grant 2018D16.

ABSTRACT Space information network (SIN) plays an extremely important role in civil and military applications. However, a specific definition of SIN still remains a challenging issue. Unlike space communication network (SCN), which merely focuses on information delivery, SIN has the ability to process the delivered information and providing corresponding services based on the information, which makes SIN seem to be able to understand information. This paper offers architecture and corresponding network model for time-space uninterrupted SIN. Initially, both physical and logical constitution of SIN architecture are presented with several corresponding constellation designs, which is the fundamental work for setting nodes in SIN. Based on this SIN architecture, a hierarchical autonomous system (AS)-based network model for SIN is proposed in order to manage SIN more efficiently by separating the whole network into several relatively stable ASs. Furthermore, we analyze topology control schemes and network capacity trend of AS-model based SIN. This paper gives the future research directions in the conclusion.

INDEX TERMS Constellation design, hierarchical autonomous system, network capacity, space information network, topology control.

I. INTRODUCTION

With gradually thorough human space science activities and continuous development of space applied researches, it is becoming difficult for traditional space systems to meet the increasing information service demands. There are many prominent problems in traditional space systems, e.g., limited coverage abilities, difficulties in system extensions, infeasibility of implementing collaborative applications among systems, etc. The inherent reason causing all the above mentioned problems is that traditional space systems are mainly built for specific users(demands), and lose compatibility among each other. Therefore, to overcome those limitations mentioned before, the concept of space information network (SIN) emerges as the times require [1].

SIN is a new type of self-organizing system constituted by various kinds of platforms as network nodes, such as geostationary Earth orbit (GEO) satellites, inclined geostationary Earth orbit (IGSO) satellites, highly elliptical orbit (HEO) satellites, medium Earth orbit (MEO) satellites, low Earth orbit (LEO) satellites, high altitude platform stations (HAPS), manned/unmanned aerial vehicles, etc [2]. Moreover, it is a service-initial system to meet the uninterrupted communication requirement of emergency rescuing, pelagic sailing, air transportation, aerospace telemetry and controlling, ground-based communication service enhancing, etc. Also, it may be able to support relay services provided by remote sensing satellites and deep-space explorer with high dynamic and large transmission delay, respectively.

In recent years, SIN plans called the Space Communications and Navigation (SCaN) program [3] and the Integrated Space Infrastructure for Global Communication (ISICOM) [4] have been proposed by National Aeronautics and Space Administration (NASA) and the Europe,

The associate editor coordinating the review of this manuscript and approving it for publication was Bora Onat.

respectively. SCaN places the three prime NASA space communications networks, Space Network (SN), Near Earth Network (NEN), and the Deep Space Network (DSN), under one management and systems engineering umbrella [3]. Project ISICOM aims to construct an efficient, flexible and reconfigurable satellite infrastructure to provide ubiquitous and secure communications [4]. Besides, in [5], the researchers have presented satellite-assisted localization and communication system for emergency services (SALICE), which hybrids satellite network and HAPS to achieve global coverage of emergency data. Another typical hybrid space network is Chinese BeiDou Navigation Satellite System (BDS), which involves GEO, IGSO and MEO satellites and also intends to provide enhancing navigation services through LEO satellite and HAPS.

So far, current architecture of SIN can be classified into three categories demonstrated as follows:

a) *Space nodes supported by terrestrial network*, which is the most mature architecture represented by Inmarsat, Intelsat, WGS [6], [7], etc. Under this architecture, space nodes of SIN merely serve as transparent transmitting channel, and globally distributed terrestrial network will provide services by processing the receiving data. In this way, system complexity can be reduced by simplifying space equipment, and it is convenient for maintenance and upgrading. However, this architecture has a poor invulnerability since the core elements of SIN are deployed on the ground. Once the ground elements become invalid, the whole network will be paralyzed since space nodes are not able to process data.

b) *Space-based network*, represented by Iridium and AEHF [8], is an architecture to establish an independent SIN through inter-satellite links (ISLs). This architecture highlights the SIN nodes' capabilities of processing, exchanging, and controlling, i.e., signal will be demodulated and decoded on the satellite and then delivered to the destination via suitable inter-satellite routes. On the one hand, it enhances systems' invulnerability comparing to the former architecture. On the other hand, the increasing capabilities may sharply raise the complexity of space equipment as well as the cost of constructing and maintenance.

c) *Space-terrestrial integrated network* is a design that combining advantages of the two aforementioned architectures. Within this framework, the space network is to realize global coverage and the terrestrial network will take charge of most managing and controlling functions. With this combination, SIN can meet the requirements of system invulnerability as well as construction costs. Moreover, this integration design also meats the future trend of the fifth generation (5G) [9].

Recently, space-terrestrial integrated network is evolved to aerospace-terrestrial integrated network with the aerial network consisting of HAPS and manned/unmanned aerial vehicles. Aerial network is an important supplement to SIN with the advantage of flexibility, which can meet emergency requirements promptly and make SIN a truly time-space uninterrupted network. However, the aerospace-terrestrial network architecture of SIN has many outstanding characteristics, e.g., hierarchical layers, heterogeneous nodes, dynamic topologies, various kinds of services, resource constrained, etc. In this field, our team have already made several achievements, e.g. concerning about the architecture and the network model of SIN in [10] and [11], topology control schemes in [10] and [12], and network capacity analysis in [13].

In this article, we summary the previous works and put efforts on two concerning domains of SIN, namely architecture and network model, respectively. First, we analyze the SIN architecture based on aerospace-terrestrial integrated network from different aspects including constitutions and logical functions. Meanwhile, several constellation designs for the GEO backbone network and Services-enhanced satellites are presented. Moreover, we propose a hierarchical autonomous system (AS) based network model for SIN in which the whole SIN is divided into a series of ASs. Within this model, we focus on issues of network topology controlling and varying trend of network capacity.

The rest of this article is organized as follows. A specific design of SIN architecture is discussed in detail. Network model is also analysed followed by topology control schemes and network capacity constraints. We then conclude with final remarks.

II. THE ARCHITECTURE OF TIME-SPACE UNINTERRUPTED SIN

A. SIN AND SCN

The SIN is a service-orient and evolved system, which can be defined as a network to acquire, transmit, and process space information in real-time based on heterogeneous space platforms (e.g. GEO satellites, MEO/LEO satellite constellations, HAPS, etc.) to meet various information service requirements. One of SIN's core aims is to ensure that users can get uninterrupted services at all times and places. According to that, the SIN architecture proposed in this paper is shown in Fig. 1., which consists the GEO backbone network, services-enhanced satellite network, HAPS network, operation and control network, access network, and ground network. Comparing to space communication network (SCN), which only provide pure communicating services, SIN contains almost all kinds of space information sources including communication, navigation, remote sensing, etc. Moreover, from the aspect of protocol stack, SCN merely involves physical, data link and network layers while SIN contains the whole stack. Clearly, the most important difference between SIN and SCN is that SIN not only transmits but also understands information. To be more visualized, in Fig. 2., SCN looks like an information pipe but SIN is a really network. In the architecture of SIN, pre-existing systems are not combined in a simple way but a networking way, which requires a novel design of the architecture and the network model of SIN to ensure that data from different systems can be transmitted in a united network. Besides, the multiple interconnected heterogeneous space platforms and the continuous

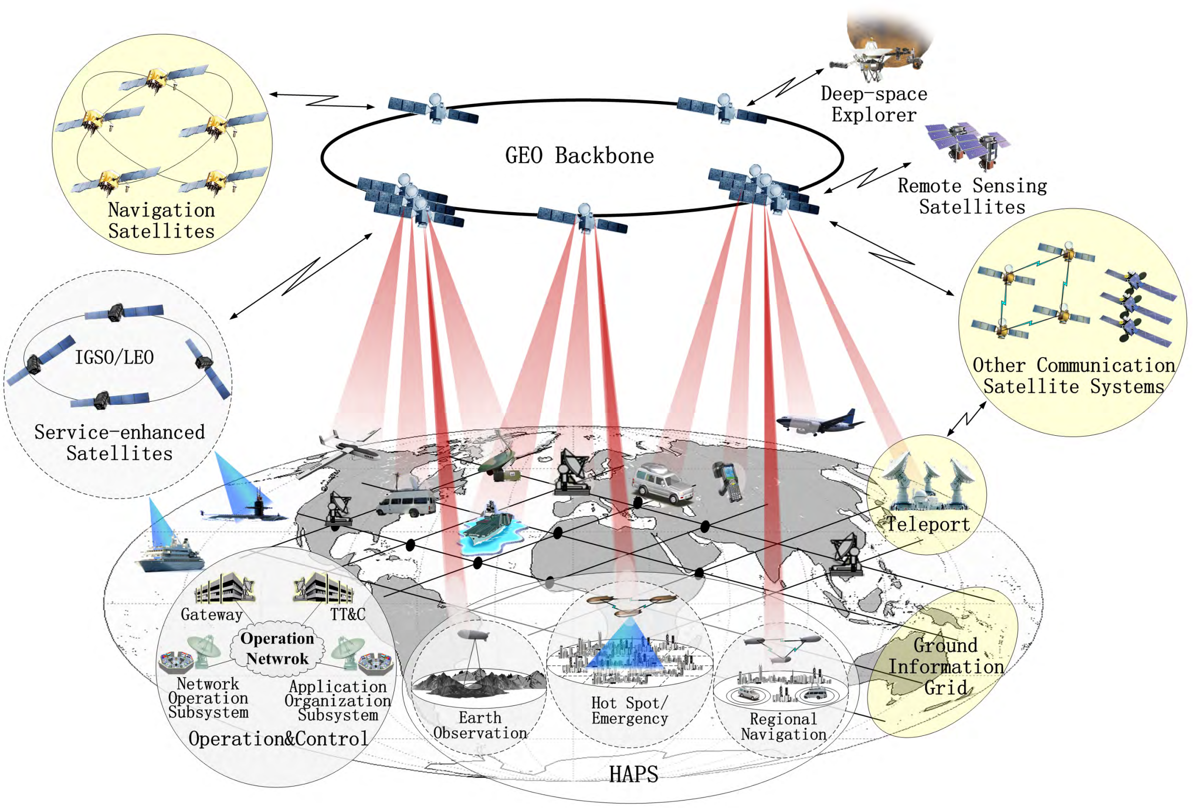


FIGURE 1. **The physical constitution of time-space uninterrupted SIN.**

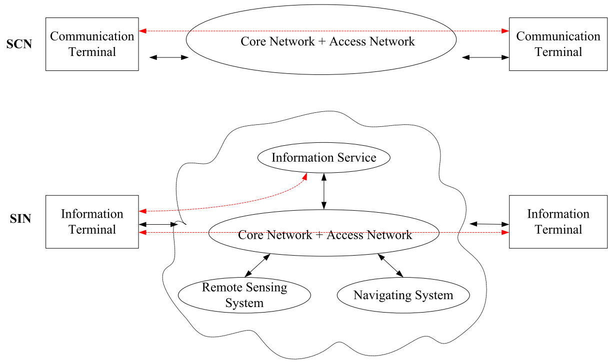


FIGURE 2. **SIN versus SCN.**

coverage analyzed later in Section II-D ensure the time-space uninterrupted characteristic of SIN.

B. SIN APPLICATIONS

SIN is capable for various kinds of applications, and in this article, we mainly discuss four major kinds of them in Table I, namely remote sensing (RS), TT&C, MMTC uRLLC, and communication. RS application in this article refers to space-based RS, which provides one-way data services in numerous fields including geography, land surveying and most Earth Science disciplines. It also has military, intelligence, commercial, economic, planning, and humanitarian applications [14]. Most services of RS application have accuracy requirements of high resolution (image, spatial, spectrum, etc.) leading to huge data amount. With respect to TT&C, it is the abbreviation of telemetry, track and command, which refers to a commanding application orienting from terrestrial TT&C centers to all kinds of space nodes to ensure their controllable orbit motion. Command data of TT&C may be short comparing to other SIN applications, while it requires high level reliability guarantee. Massive Machine Type Communications (MMTC) Ultra Reliable & Low Latency Communications (uRLLC) are two application categories of 5G [15]. Meanwhile, [9] indicated that satellite-terrestrial integrated system is a significant trend of 5G. Hence, several typical services of those two applications, e.g. unmanned vehicles, Supervisory Control and Data Acquisition (SCADA) and Internet of Things (IoT), can be provided via SIN [16], [17]. Last but not least, communication in Table 1 is a main application type in SIN, which involves all nodes in SIN and faces synthetical services (voice, video, internet, etc.) with high capacity.

C. NETWORK ELEMENTS IN SIN

Network elements in SIN can be classified into two categories, namely nodes and links.

1) NODE ELEMENTS

Node elements contain four kinds of logical nodes. a) **Information acquiring nodes (IAN)** are capable of

TABLE 1. Characteristics of applications in SIN.

Application Type	Corresponding Targets	Features
Remote Sensing	Remote sensing satellites	One-way, Large amount of data
Navigation	Navigation satellite constellation	One-way
TT&C	Space nodes, Terrestrial TT&C station	Low data rate, High reliability
MMTC/ uRLLC	user nodes (unmanned vehicles, SCADA users, IoT devices etc.), SIN nodes	Real-time, Accurate, Uninterrupted, Massive-Connections
Communication	All users in SIN	High capacity, Synthetical services

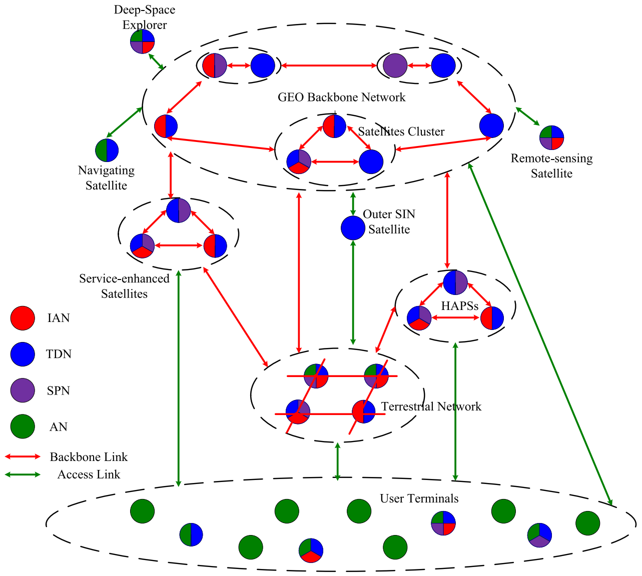


FIGURE 3. Logical version of SIN architecture.

information acquisition and generation, including remote sensing satellites, deep-space explorer, and terrestrial user terminals; b) **Transmitting and distributing nodes (TDN)** are to transmit, distribute and exchange information, including GEO satellites, service-enhanced satellites, HAPSs, and ground information grid nodes; c) **Storage and Processing nodes (SPN)** are to store and process information, also can provide information service, which include GEO satellites, HAPSs, and ground information grid nodes; d) **Application nodes (AN)** include various kind of terrestrial and space-end users which can achieve and apply information.

2) LINK ELEMENTS

There are two types of link elements in SIN, i.e., backbone links and access links. Backbone links offer high speed connections between TDN and SPN through laser or microwave links. Access links cater to connect the other two kinds of nodes with TDN and SPN. The properties of access links are adaptive to the phenomenon they serve.

With classified nodes in SIN, the logical version of SIN architecture is depicted in Fig. 3. In Fig. 3., aforementioned four kinds of SIN nodes (i.e., IAN, TDN, SPN, and AN) refer to red, blue, purple, and green, respectively. In addition, the dash ellipses represent different sub-networks in SIN corresponding to the depiction in Fig. 1, and the solid lines represent backbone and access links, which are discriminated by the color red and green. By combining Fig. 1 and Fig. 3, components of SIN and their functions are clear to be observed. To be noticed, there may be a single satellite or a satellite cluster deployed in a specific orbit position. Within a satellite cluster, each satellite can work both independently and collaboratively. Therefore, physical nodes in SIN generally possess more than one logical node, and this logical architecture clearly prove that SIN has the compatibility to all existing space systems.

D. CONSTELLATION DESIGN IN SIN ARCHITECTURE

Constellation design is a fundamental work for setting nodes in SIN. Generally, current space systems mainly adopt single orbit for constellation design, i.e., GEO, Incline Geosynchronous Orbit (IGSO), highly elliptical orbit (HEO) or LEO. However, single orbit design has two drawbacks: 1) For GEO satellite, the average communication angle in mid-high latitude region is small, and has blind spots in polar areas; 2)For LEO satellite, the satellite amount to achieve global coverage is too large to be cost efficient. For example, Iridium [18] system provides global coverage with a LEO constellation of 66 satellites, and the original plan of OneWeb [19] consists 648 LEO satellites while a mixing orbits constellation of 3 GEO and 24 LEO satellites illustrated next can also realize continuously global coverage. Therefore, mixing orbits constellation design has been taken into consideration due to its better coverage performance and cost efficiency than single orbit design. In the literature, [20] proposed a multilayer architecture of SIN based on LEO satellite backbone. However, in this paper, apart from offering traditional fixed and mobile services, GEO satellites play the role of SIN backbone nodes, and satellites on other orbits play the role of services-enhancing nodes. Comparing to LEO satellites as backbone nodes, GEO satellites have the advantages of relatively wider coverage and stronger processing capability to offer inter-AS (the concept of AS will be further illustrated in Section III) relay services when an AS has no direct connect to the destination and the traffics on other inter-AS paths (e.g., ISLs) are busy. To be noticed, in order to make the whole SIN efficient and reduce the system complexity, only few nodes in each AS (called boundary nodes in Section III) are elected to communicate with the backbone nodes. Moreover, while implementing the concept of Software Defined Network (SDN) and Network Function Virtualization (NFV) in SIN [21]–[25], backbone nodes can realize the broadcasting channel between network controller and SIN nodes. On the other hand, as services-enhancing nodes, satellites in other orbits take their advantages (e.g., enhancing coverage, low

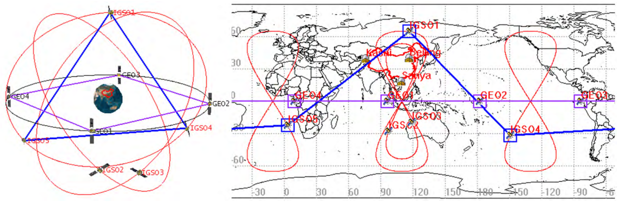


FIGURE 4. 4GEO + 5IGSO constellation.

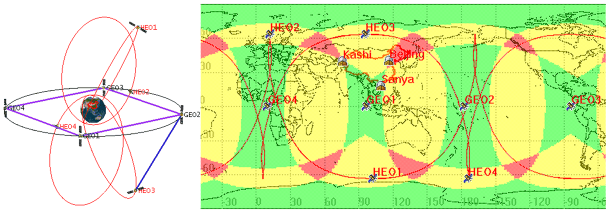


FIGURE 5. Coverage performance of 4GEO + 4HEO constellation.

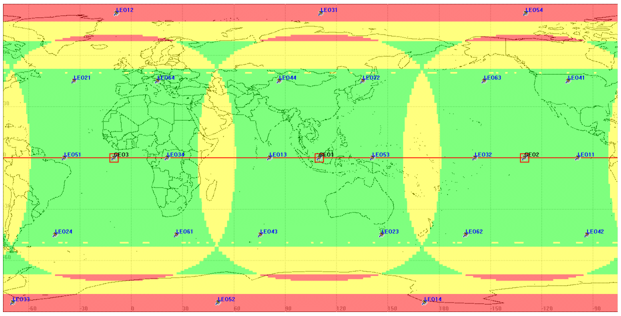


FIGURE 6. Coverage performance of 3GEO + 24LEO constellation.

latency) to provide corresponding SIN applications listed in Table I. Following part of this subsection will illustrate several potential constellation design for SIN, which are all based on the GEO backbone network. All three of the following constellation designs can be utilized in the construction of SIN individually or jointly in order to take advantages of satellites from different orbits and enhance the capability of SIN. In addition, cost effectiveness is also taken into consideration so that all constellation designs are meant to be minimum constellation, which means constellation with minimum number of satellites to realize global coverage.

1) GEO + IGSO CONSTELLATION

Since having the same orbit altitude as GEO satellite, IGSO satellites inherit the advantages of GEO ones. Moreover, IGSO satellites have higher average communication elevation angle (CEA) than GEO ones so that they can provide better service qualities in the areas of cities, canyons and other places where require high CEA. Common ground track and single orbit plane are two methods to construct an IGSO constellation. To ensure efficient and stable operation of SIN nodes, ISLs are not considered between nodes that have relative motion. A mixing orbits constellation design consisted of 4 GEO and 5 IGSO satellites is depicted in Fig. 4, where IGSO 1, 2, 3 are common ground track satellites and IGSO 1,4,5 are in the same orbit plane. In this design, IGSO 1 is the mutual satellite in both common ground track satellite group and same orbit plane satellite group, which connects two groups of IGSO satellites and save an orbit position.

2) GEO + HEO CONSTELLATION

Molniya orbit [26] is considered as a good HEO supplement in SIN, which is a well-known HEO that its apogee's ground projection does not drift by time. This characteristic enables HEO satellites to provide better coverage than GEO ones in mid-high latitude regions when the orbits are well designed. Fig.5. presents a GEO/HEO mixing constellation design, where HEO 1-2 and 3-4 aim at enhancing coverage of northern and southern hemisphere, respectively. In Fig.5., coverage performance is exhibited by different colors. Green represents that the area is completely covered by one satellite. Similarly, yellow and red corresponds to two and three satellites, respectively.

3) GEO + LEO CONSTELLATION

Apart from HEO satellite, LEO satellite constellation [17] can also provide coverage in mid-high latitude regions and polar areas by using near polar orbits. Meanwhile, in order to be convenient for controlling satellites during operation, the adopted orbits are designed as recursive orbit, i.e. satellites will pass the same point after a certain time interval in days. Therefore, the LEO period T_s ought to satisfy

$$\frac{T_s}{T_e} = \frac{k}{n} \tag{1}$$

where T_e is equinoctial day with length of 86164 seconds, and k, n are integers, which represent LEO period in days and recursive cycles, respectively. According to the Kepler's third law, LEO altitude h is calculated by

$$h = \frac{T_s^{\frac{2}{3}} \mu^{\frac{1}{3}}}{(2\pi)^{\frac{2}{3}}} - R \tag{2}$$

In (2), μ is the Kepler constant with $\mu = 3.986 \times 10^{14} m^3/s^2$, and R is the radius of earth with 6371 km. Fig. 6. shows a GEO/LEO mixing constellation with 3 GEO satellites and 24 LEO satellites distributed in 6 orbit planes. The three GEO satellites are uniformly distributed in the equatorial plane with separation of 120°. The legend of coverage performance is the same as that in Fig. 5.

III. NETWORK MODEL FOR SIN BASED ON HIERARCHICAL AUTONOMOUS SYSTEM

As described above, SIN contains various nodes such as satellites, HAPSs, terrestrial terminals, etc. Differences between nodes commonly exist in almost all aspects including deploying altitude, operating environment, major functions, etc. Meanwhile, SIN has a highly dynamic network topology caused by the inherent relative motion among network nodes including satellites, HAPSs, and terrestrial terminals. As a hierarchical and heterogenous network, even a minor change of local part in the SIN such as application situation, topology,

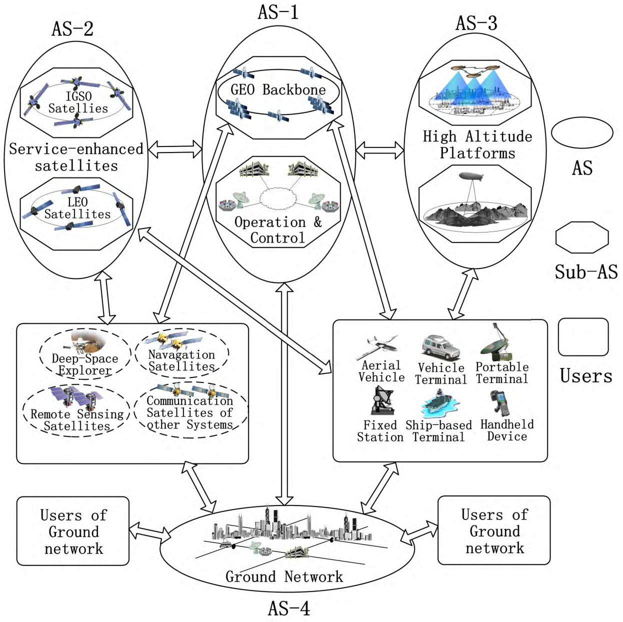


FIGURE 7. Hierarchical autonomous system-based SIN.

and environment of signal propagation, the whole network state will be affected. Therefore, it is unwise to find a unified strategy to manage the whole parts of SIN, which will make inefficiencies to system operation, and even unable to stably handle the overall control information in SIN.

According to that, in this article, SIN is divided into a series of Autonomous Systems (AS) containing nodes and links with similar properties. Within each AS, a relatively independent network and transmission strategy will be adopted. Routing and control information corresponding to different ASs will be exchanged through their boundary nodes (BN). An AS can be further divided into sub-ASs if necessary. In this way, the high dynamic SIN is transformed into several weak dynamic sub networks, and the network model for hierarchical ASs-based SIN is depicted in Fig. 7.

A. AUTONOMOUS SYSTEMS IN SIN

In Fig. 7, the whole SIN is divided into 4 primary AS according to the property, task characteristics, distribution area of the nodes, and link capacities.

1) SPACE AUTONOMOUS SYSTEM

Space AS includes all kinds of space nodes, which are all deployed in outer space with orbit motion and links between them are dynamic and discontinuous. Fortunately, since those nodes have deterministic and controlled motion state, the relative position between different nodes is predictable as well as visibility and usability of links. GEO backbone network and operation & control network have stable network topologies, and form the initial part of SIN. Thus, those two parts are divided into AS-1. AS-2 contains the remaining space nodes that provide space information applications. AS-1 and AS-2 constitute the whole space AS.

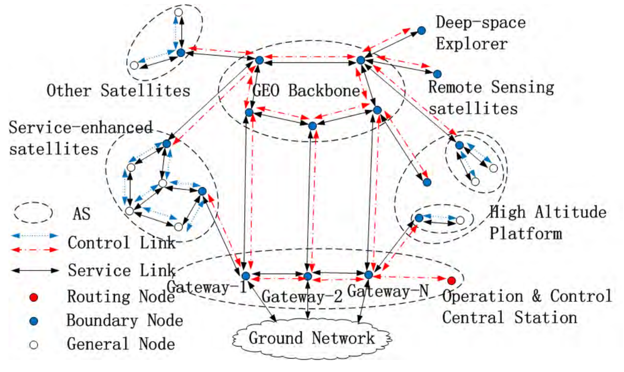


FIGURE 8. Routing information generation and distribution in the AS network model.

2) HAPSs AUTONOMOUS SYSTEM

HAPSs in AS-3 locate in atmosphere with altitude of 10-30 kilometers, and maintain quasi-static to target areas via hovering or gyrating. Comparing to space AS, both motion and link states between nodes in AS-3 have stochastic properties due to the impact of air flow.

3) TERRESTRIAL AUTONOMOUS SYSTEM

AS-4 is terrestrial AS, i.e. ground network. Although AS-4 has extremely large number of nodes and complicated connections comparing to the other ASs, the research of topology control in terrestrial heterogenous network is mature, which can provide guidance for routing and topology optimization in AS-4.

B. LINK SEPARATING STRATEGY

The topology, routing and control information of each AS are exchanged through BNs and finally transmitted to the operation and control central station (OCCS). OCCS then calculates and selects the optimal route according to the topologies and network states of SIN for BN in order to achieve more effective management.

As illustrated in Fig. 8, in the AS network model, the control and service information are transmitted respectively as well as intra-AS and inter-AS routing are separated. This strategy is called link separating strategy, which used to raise the efficiency of network management of SIN. Link separating strategy consists three steps: selecting BNs, separating control and service links, and separating inter-AS and intra-AS routings. In the step of separating control and service links, OCCS will establish stable control links to BNs belong to different ASs and transmit controlling information via those stable links. Within each AS, BN will distribute controlling information to corresponding AS nodes via control links dynamically generated by BN itself. Meanwhile, service links can be oriented from an arbitrary AS node to BN. After this step, controlling and service information are transmitted separately in SIN. As for routing separation, the network combines static and dynamic routing scheme.

1. The topologies of GEO backbone and operation & control network are stable. Thus, OCCS only needs to keep

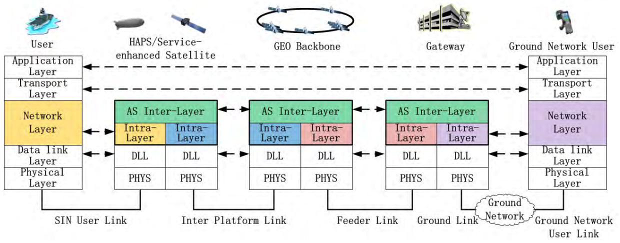


FIGURE 9. The protocol stack model of hierarchical AS SIN.

regularly updating the routing information of this part and distribute this information as static routing to BN of each AS (this control link is illustrated in red dotted line in Fig. 8.).

2. The dynamic routing scheme is generated inside each AS or sub-AS by BNs.

3. Static routing of backbone and control network is distributed to the general nodes through BN (this control link is illustrated in blue dotted line in Fig. 8.). Basing on link separating strategy, SIN is immune to the mutual influence between the massive amount of service information and controlling information. Besides, impact from the variation of dynamic intra-AS routing to the global SIN routing status is isolated via routing separation. Therefore, this strategy will increase the managing and operating efficiency of SIN.

As discussed above, BN has a heavy task to process huge amount of data. Generally, it can be prescribed in advance or be elected by AS nodes via algorithm. No matter in which way, the election of BN is a trade-off among various factors. The most important factors should be considered are summarized as follows:

1. BNs should have strong processing ability to maintain the routing information of an AS and transmit it to BNs belong to higher or lower layer AS. Meanwhile, they have to maintain the service of their own.

2. BNs should be able to maintain a long and stable working state. In the aerospace environment, signal has a long transmitting delay and big propagation loss. Moreover, considering the constrained bandwidth and power resource, an AS should avoid extra cost caused by frequent electing and switching of BNs.

3. Connection between BNs and other nodes in the AS should be convenient, i.e. when BNs exchange control information with other nodes, the hops of data should be as few as possible.

4. BNs should be high reliable. In the AS network model, transmission of control information is independent from the one of service information. BN is the node that fully takes charge of exchanging control information, and maintains the normal operation of an AS, which needs high reliability.

C. *PROTOCOL STACK MODEL OF AS-BASED SIN*

In the AS network model, as the exchanging center of control information, BNs should be compatible with each other on protocol stack and satisfied their own services at the same time. Considering integrating SIN with terrestrial network in the future, protocol stack for SIN should take related protocols such as TCP/IP [27], CCSDS [28] into consideration. In Fig. 9, a protocol stack model of SIN is depicted, which corresponds to a completed SIN application process involving all kinds of ASs. In order to make it clear, protocol stack between the same kinds of platform is omitted. To be noticed, as mentioned before, when nodes in ASs (e.g., sub-ASs in AS-2 and AS-3) have direct connection to the Gateway, they can totally omit the inter-AS transmission to AS-1, i.e. GEO backbone. Obviously, this protocol stack has heterogenous lower layers (Physical Layer and Data Link Layer) for different types of inter-AS links depicted in Fig. 9. In particular, the network layer is divided into two sub-layers, intra-AS sub-layer and inter-AS sub-layer. As shown in Fig. 9, intra-AS sub-layer is only used in the SIN user link and ground link, which corresponds to information exchanging between SIN and terrestrial networks. On the other hand, inter-AS sub-layer, which is identical for all space ASs, plays the role of network layer in the space part of SIN. For inter-AS communications mentioned before, BNs are equipped with inter-AS communicating payload, which enables them to covert intra-AS air interface to inter-AS ones. After matching lower layers of inter-AS links, benefiting from the identical intra-AS sub-layer, SIN information can be transferred through different ASs with same protocol, which enhances the inner connectivity of SIN.

IV. HIERARCHICAL AS BASED TOPOLOGY CONTROL IN SIN

Topology control (TC) algorithms can be divided into two groups, namely centralized and distributed ones. In centralized algorithms, the status of all nodes are known to a central entity (CE), which utilizes these information to optimize topology. The centralized ones, however, are not capable for AS-based SIN since the amount of control messages is too large for the CE to collect. For distributed algorithms, links' preservations autonomously depend on the neighboring nodes' properties collected by each network element. Unfortunately, due to the limited topology information acquired, distributed TC algorithms cannot be solely deployed in SIN for global optimization.

Considering the scale of SIN, the majority of links have extremely long distances. Thus, there will be an extreme impact of SIN's delay and efficiency performance by excessively utilizing long-distance links. In the SIN TC, therefore, end-to-end delay is a much more significant index. Besides long-distance links, via constructing suitable relations between neighbor nodes, SIN nodes will collaboratively determine link properties and define the network topology. That is, unnecessary links will be erased from the SIN topology so that it is insensitive to unpredictable events such as hardware failures.

In SIN, a hybrid AS network topology control (AS-TC) algorithm can be deployed. Briefly, there are three phases in the AS-TC algorithm, namely constructing, intra-sub-AS TC and inter-sub-AS TC. In Phase 1, the whole SIN will be autonomously separated into sub-ASs with sub-AS cores elected by SIN nodes. Then, operations in Phase 2 mainly use centralized method to ensure strong connectivity within each sub-AS. Meanwhile, each core of sub-ASs aims to minimize the overall maximum delay of its sub-AS. Finally, distributed method is utilized in Phase 3 for calculating link delay and exchanging topology information with adjacent sub-AS cores via a set of border nodes selected by core nodes. Those three phases are amply illustrated as follows.

1) PHASE 1: CONSTRUCTING SUB-AS NETWORKS

To construct sub-AS networks, a minimal set of SIN nodes will be selected as core nodes to control its own sub-AS with merely 1-hop. Moreover, operations in Phase 2 and 3 mostly rely on core nodes. Selecting process contains four steps listed as follows:

Step 1 (Broadcasting Process): Initially, to search adjacent nodes, each node will periodically broadcast a hello message that includes the basic properties of a node, i.e. location, connecting degree and transmitting delay.

Step 2 (Selecting Process): After waiting for a predefined time of Step 1, selecting process initiates for every node deciding its capability of being a core node. The standards of core nodes depend on their abilities in local optimality. To be noticed, waiting time should be set long enough to ensure that Step 1 is fully operated among the whole SIN (i.e, a node has received broadcasting messages from all of its neighbors).

Step 3 (Supplementing Process): After Step 2, each node will check cores within their communication range $R_{\max}$. If there exists, the core with the lowest transmitting delay to it will be considered as its parent node.

Step 4 (Optimizing and Maintaining Process): To maintain a minimal set of core nodes, each core will search for the existence of other cores with higher altitude within $R_{\max}$ (from Step 1). If there does not exist, it will be added into a core node set with the highest altitude. Meanwhile, its former member nodes will no longer belong to it, i.e. becoming nodes without parent, and this step will back to Step 3 if such nodes exist in the AS. After iterations, each node becomes either a core or a member node. Due to the relative motion in SIN, this process will keep monitoring the AS network.

2) PHASE 2: INTRA-SUB-AS TC

Due to the relatively small scale of a single sub-AS, centralized method can be deployed intra-sub-AS TC. Briefly speaking, in Phase 2, the core node will calculate all links under given constraints (min-max delay and k-connectivity) within its sub-AS. The algorithm is illustrated in **Algorithm 1**, in which $G = (V, E)$ represents the space AS (V is the set of nodes and E is the set of edges that connect nodes pair of (u_i, v_i)), and let sub-ASs $G_j, j \in (1, n)$ become part of G. **Algorithm 1** ensures the inherent connecting property from G_s to G_k, and minimizes all edges' maximum end to end delay for every sub-AS.

Algorithm 1 Intra-Sub-AS Topology Control

```
Require: Sub-AS $G_s = (V_s, E_s)$
  Required connectivity $k$
Ensure: Sub-AS $G_k = (V_k, E_k)$
  $V_k \leftarrow V_s, E_k \leftarrow \phi$
  Sort all edges in $E_s$ in ascending order of weight
  for all edge $(u_i, v_i)$ in the order do
    if $u_i$ is not k-connected to $v_i$ then
      $E_k \leftarrow E_k \cup (u_i, v_i)$
    end if
  end for
```

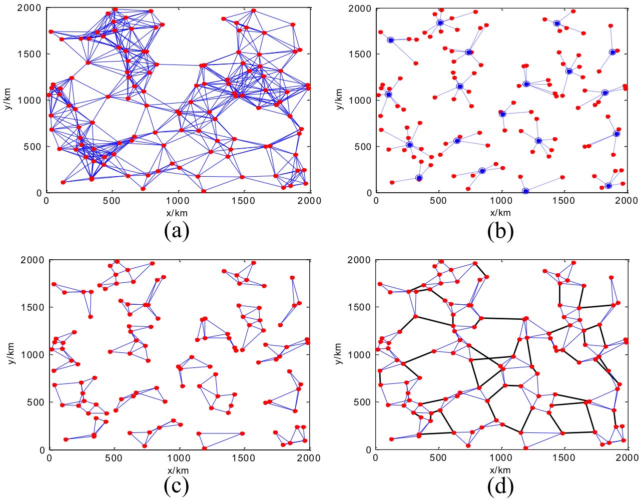
3) PHASE 3: INTER-SUB-AS TC

Phase 3, as its name implies, aims to connect neighboring sub-ASs and do further TC. To make each sub-AS connectable to all adjacent ones, broadcasting process in Phase 1 is repeated by every node. Assuming that node u and v belong to different sub-AS, if u receives broadcasting message from v, v's status information will be added to u's border list. Core nodes will finally be informed of all border lists from their member nodes. Based on border lists, a distributed inter-sub-AS TC algorithm is presented.

In this algorithm, sub-AS A's core c intends to find the existence of k disjoint links directing to each neighboring sub-AS based on the criterion of maximum cardinality matching (MCM) [29] between two sub-ASs using a bipartite graph representation. If k links fulfil the size requirement of MCM and the min-max delay optimal, c will select these links. However, if there only exist k_m ($k_m < k$) disjoint links between A and its neighbor, c is going to maintain k_m-connectivity of the two sub-ASs and minimize the maximum delay between them. To be noticed, though k-connectivity between two sub-ASs may be invalid after this operation, a global k-connectivity can be guaranteed after completing this phase since the establishments of connectivity among adjacent sub-ASs will be finished by then.

A significant parameter in inter-sub-AS TC algorithm is the maximum delay of the selected k links between two sub-ASs, denoted by $D_{IA}(G_1, G_2)$, which is a criterion to judge the necessity of a certain link between two neighboring sub-AS networks (i.e. G_1 and G_2), while maintaining the optimized topology's connectivity. However, when the number of disjoint links between G_1 and G_2 is k_m, $D_{IA}(G_1, G_2)$ tends to ∞. Therefore, when a sub-AS C is adjacent to A and B and has both $D_{IA}(G_A, G_C)$ and $D_{IA}(G_B, G_C)$ less than $D_{IA}(G_A, G_B)$, sub-AS A will not connect to a neighboring sub-AS B directly.

Considering that the nodes in the same AS share similar properties, they can be assumed to be homogeneous. For instance, TC process in HAPS AS is analyzed in Fig. 10. Comparing to other satellite-consisted ASs, HAPS AS has the most random relative position relationships between HAPS components due to their stochastic motion. Therefore, analysis in HAPS AS can be more effective to testify the reliability of TC algorithm. In Fig. 10(a), the original physical topology



(a) (b) (c) (d)

FIGURE 10. **Hierarchical AS based TC for 125 SIN nodes in HAPS AS. (a) Original status. (b) Sub-AS construction finished. (c) Intra-sub-AS TC finished. (d) Inter-sub-AS TC finished.**

with nodes randomly distributed in a $2000 \times 2000 km^2$ region and without topology control is presented for numerical analysis. In this AS, the maximal transmission range $R_{\max}$ for all nodes equals 350 km. Comparing to (a), in Fig. 10(b) sub-ASs were established and nodes of the original AS were divided into 19 sub-AS networks, where the average number of nodes per sub-AS is 6.58. Fig. 10(c), and (d) show the TC results after intra and inter-sub-AS TC, respectively with $k = 2$. In Fig. 10(d), the inter-AS links were drawn in black thick lines.

V. SIN CAPACITY ANALYSIS

Due to the distinguishing characteristics of SIN, making SIN nodes efficiently collaborate with each other is challenging but significant. In terrestrial network, network capacity is a useful index to reflect network efficiency. However, methods utilized for terrestrial network capacity are no longer effective for analyzing the capacity of dynamic, heterogeneous integrated SIN due to lack of universal study approaches that can be used in SIN. To design a suitable analyzing method for SIN is therefore of uppermost priority.

A. SIN CAPACITY MODEL

To establish a general analytical model for SIN, the term of arbitrary space network is used to reflect the arbitrary and deterministic location properties (i.e. not randomly located) of all resident SIN nodes both moving in orbits (LEO, MEO, GEO, etc.) and distributed on earth (terrestrial nodes). Obviously, motion status of those nodes is either stationary (terrestrial nodes) or arbitrarily moving in fixed orbits (satellite nodes). For each node in SIN, there are three roles for them, namely source, relay and destination. A node may play more than one roles, and choose an arbitrary number of destinations, which corresponds to two transmitting types in SIN, unicast and multicast. In this model, SIN is assumed to be a saturated network, where all nodes are capable of generating and processing infinite amount of data.

Let V_N be the node set with $V_N = N$ and let E be the link set. Without loss of generality, each node $n \in V_N$ is assumed to be able to transmit over common channel with the maximum rate $R_n(t)$ bps at time t. With this assumption, unconcerned physical layer details can be ignored and the SIN capacity can be merely analyzed from network topology. Meanwhile, SIN is assumed to be *stable*, which means that if and only if for any fixed N, each node in the network has an infinite queue to transmit, and the queue length of storage packets remaining to be transmitted in any relay node does not grow to infinity as time interval $T \to \infty$.

The transport capacity of arbitrary SIN $G = (V_N, E)$ is denoted by C_G^{χ}, where χ represents the adopted spatial and temporal network scheduling algorithm. C_G^{χ} is defined as

$$C_G^{\chi} = \frac{\sum_{i=1}^{N} A_{i,T}^{\chi}}{T} \tag{3}$$

where $A_{i,T}^{\chi}$ is the amount of bits successfully transmitted by node v_i during T, i.e. the node capacity of v_i. $A_{i,T}^{\chi}$ consists of four components listed as follows:

a) Availability. The availability component of capacity model depends on the existence of line-of-sight of each links in path p. Generally, it is a function with four input parameters: satellite orbits dynamics, terrestrial nodes' position, the minimum constrains in elevation angle, and time.

b) Data rate. Selected data rate mainly rest with the expected channel status, which ought to meet the requirements of the minimum signal-to-noise (SNR). Furthermore, it is adaptive to the network dynamics. χ is a powerful tool for selecting optimized data rate distributions to maximize network throughput.

c) Path selecting scheme. Governed by the network scheduling algorithm χ, a path p may or may not be selected even it is available.

d) Path reliability. Successful data transmitting from v_i to its destinations is influenced by all nodes in path p. Several factors of these nodes may impact the link performances in p including antenna slewing and acquisition maneuvers, unknown noise that degrade the SNR and system failures.

B. MAJOR CONSTRAINTS TO NETWORK CAPACITY

The major constraints that affect network capacity model include nodes' dynamics, long term variations, and practical factors.

1) NODES' DYNAMICS

Nodes' dynamics will impact the availability component of SIN capacity. Space-AS in SIN is a high dynamic network with dramatic relative motion between space and terrestrial nodes. Moreover, new nodes may be added into SIN when invalid nodes is removed from the network. Practically, in designing ISL, though there exist relative motion between satellites in different orbits, ISLs are designed to be stable and reliable. Hence, line-of-sights between satellite and terrestrial nodes mainly affect the availability component in SIN's capacity. This factor highly depends on the orbit design of satellites. For instance, under the network topology constraint, satellite with inclination $i_s = 30°$ cannot cover any ground nodes with latitudes $l_g > 51°$ for an 880 km orbit and $l_g > 58°$ for a 1450 km orbit. For satellites with low inclinations, the percent coverage will expand with the increase of satellite altitude. To compute total visibility duration, satellite's position vector $W(t)$ will be firstly calculated in the Earth centered Earth fixed (ECEF) coordinate with the initial satellite orbit parameters (i.e. semi-major axis, eccentricity, inclination, argument perigee, right ascension of the ascending node and true anomaly). Terrestrial node's position vector U can be also calculated with its longitude and latitude. Hence, the visibility between satellite and terrestrial node can be expressed as

$$\Delta T = \int_{t_0}^{T+t_0} \varepsilon(e_{\min} - \arccos(\frac{W(t)\cdot U}{|W(t)||U|}))\mathrm{d}t \qquad (4)$$

where $\varepsilon(\cdot)$ is unit step function and $e_{\min}$ is the minimum elevation angle of a terrestrial node. This trend can be utilized in optimizing orbit design and terrestrial nodes' distribution in order to maximize path availability and capacity.

2) LONG TERM VARIATIONS

There are two types of variations in satellite orbit trajectory: 1) short term (one or several orbital cycle) variations, and 2) long term (or seasonal) variations. The long term satellite trajectory variations usually have deterministic periods, and so do the visibility duration dominated by the trajectory. The long term visibility duration variations can be largely explained by the perturbations of satellite rotational rates due to Earth's oblateness characterized by the J4 coefficients. The length of visibility duration is directly related to the maximum elevation angle $e_{\max}$ between terrestrial and satellite node, which is defined as

$$e_{\max} = \frac{\pi}{2} - \theta - \arcsin(\frac{r_E \sin\theta}{\sqrt{(r_E^2 + r^2 - 2r_E r\cos\theta)}}) \qquad (5)$$

In (5), θ is the Earth central angle, r_E and r are the earth radius and orbit altitude, respectively. Generally, space missions mainly focus on guaranteeing the satisfaction with thee worst mission scenario and have to sacrifice several potential advantages under better circumstances.

As long term variation trends are periodic in nature, this trend can be modeled to analyze characteristics and take potential advantages of better space cases for maximizing network capacity.

3) PRACTICAL FACTORS

In high fidelity network capacity models, those inefficiencies and constraints that impact the path reliability should be considered, both from the satellite and terrestrial perspectives. ISLs are usually based on high frequency band (Ka, EHF or leaser), and their beam-widths are narrow. In addition, the attitude control of each satellite usually has certain errors, which lead to alignment errors. Due to the alignment errors of the transmitting and receiving antennas, the actual antenna gain is lower than antenna peak gain. Terrestrial-satellite aspect is also complicated with time-varying SNR caused by practical factors such as maximum elevation angle e_{max}, carrier frequency f_c, average rainfall rate R_{rain}, node altitude h_s, etc. Moreover, influences from other SIN nodes including in-line interference between nodes with different operating altitude and adjacent interference between neighboring nodes will affect signal-to-interference ratio (SIR). Those practical factors will further decrease the assessment of SIN capacity, and narrow down the optimization range of χ.

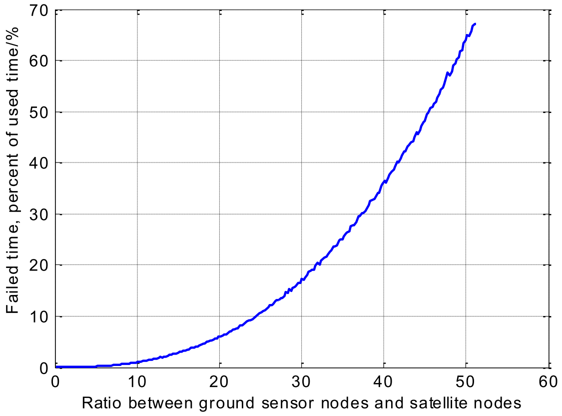


FIGURE 11. Failed access time due to constrains in SIN.

Fig. 11. shows failed access time, defined as the access requests failed time due to aforementioned capacity constrain of SIN. Failed access time increases exponentially with the growth of the ratio between ground nodes and satellite nodes in this analysis. It is obvious that the SIN's capacity does not scale linearly with the size of the network. To be more specific, the whole SIN's capacity depends on the value with the highest ratio between user and service nodes. Thus, intelligent nodes deployment methods is significant to be adopted in order to meet unbalanced service requests in SIN, and to maximize the network capacity of growing SIN with network scheduling algorithm χ.

VI. CONCLUSIONS AND FUTURE WORK

The SIN is a new type of self-organizing space network that integrates numerous kinds of platforms. SIN nodes commonly have complicated behaviours, and the network topology is three-dimensional, dynamic and multi-layered. Therefore, efficiency and reliability of operating SIN are challenging. According to these, SIN architecture based on hierarchical AS is proposed. We have also provided the network model of SIN and analysis on SIN including topology control and capacity trend.

Indeed, the work presented in this article could be extended in many interesting directions, as summarized below.

A. FURTHER RESEARCH ON TIME-VARYING TOPOLOGY MODEL AND TC ALGORITHMS

SIN's dynamic property is one of the most challenging issues in TC. Generally, status of nodes and links are predictable. However, under extreme circumstances such as failure of nodes and links or deliberate interferences, those status will become unpredictable. Therefore, the future direction of TC should combine topology prediction and real-time perception in order to ensure time-space uninterrupted SIN.

B. JOINT OPTIMIZATION OF TOPOLOGY CONTROL AND ROUTING STRATEGY

Current TC methods mainly focus on constructing sparse network topology and reducing link redundancy to save network cost. However, those methods will have retroaction on the performance of routing strategy since lack of link redundancy leads to decrease of available routes. Therefore, combining TC with specific routing strategy to optimize topology and link redundancy may receive a better performance.

REFERENCES

[1] K. Dai and C. Zhu, "A new architecture design of space information networks," in *Proc. 6th Int. Conf. Internet Comput. Sci. Eng.*, Apr. 2012, pp. 217–220.

[2] G. Zhang, W. Zhang, H. Zhang, and Z. Xie, "A novel proposal of architecture and network model for space communication networks," in *Proc. IAF 65th Int. Astron. Congr.*, 2014, pp. 1–7.

[3] *Space Communications and Navigation Program*. Accessed: Jun. 2016. [Online]. Available: https://en.wikipedia.org/wiki/Space_Communications_and_Navigation_Program

[4] A. Vanelli-Coralli, G. E. Corazza, M. Luglio, and S. Cioni, "The isicom architecture," in *Proc. Int. Workshop Satell. Space Commun.*, Sep. 2009, pp. 104–108.

[5] E. D. Re *et al.*, "Salice project: Satellite-assisted localization and communication systems for emergency services," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 9, pp. 4–15, Sep. 2013.

[6] *Intelsat*. Accessed: Feb. 2019. [Online]. Available: http://www.intelsat.com/

[7] E. Morris, L. Levine, C. Meyers, and D. Plakosh, *System of Systems Interoperability (SOSI): Final Report*. Pittsburgh, PA, USA: Carnegie Mellon Univ., 2004.

[8] K. Schroth, N. Burkhardt, T.-S. Che, and D. Pisano, "IP networking over the AEHF MilsatCom system," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Oct./Nov. 2012, pp. 1–6.

[9] M. Jia, X. Gu, Q. Guo, W. Xiang, and N. Zhang, "Broadband hybrid satellite-terrestrial communication systems based on cognitive radio toward 5G," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 96–106, Dec. 2016.

[10] W. Zhang, G. Zhang, Z. Xie, D. Bian, and Y. Li, "A hierarchical autonomous system based space information network architecture and topology control," *J. Commun. Inf. Netw.*, vol. 1, no. 3, pp. 77–89, 2016.

[11] W. Zhang, D. Bian, Z. Xie, and G. Zhang, "A novel space information network architecture based on autonomous system," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[12] W. Zhang, G. Zhang, L. Gou, B. Kong, and D. Bian, "A hierarchical autonomous system based topology control algorithm in space information network," *KSII Trans. Internet Inf. Syst.*, vol. 9, no. 9, pp. 1–22, 2015.

[13] W. Zhang, G. Zhang, F. Dong, Z. Xie, and D. Bian, "Capacity model and constraints analysis for integrated remote wireless sensor and satellite network in emergency scenarios," *Sensors*, vol. 15, no. 11, pp. 29036–29055, 2015.

[14] *Remote Sensing*. Accessed: Aug. 2017. [Online]. Available: https://en.wikipedia.org/wiki/Remote_sensing#Satellites

[15] *IMT Vision: Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond.* Accessed: Jul. 2018. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf

[16] M. de Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite communications supporting Internet of remote things," *IEEE Internet Things J.*, vol. 3, no. 1, pp. 113–123, Feb. 2016.

[17] Z. Qu, G. Zhang, H. Cao, and J. Xie, "Leo satellite constellation for Internet of Things," *IEEE Access*, vol. 5, pp. 18391–18401, 2017.

[18] *Iridium Satellite Communications*. Accessed: May 2016. [Online]. Available: https://www.iridium.com/

[19] *OneWeb*. Accessed: Mar. 2018. [Online]. Available: http://www.oneweb.world/

[20] K. Zhang, L. Xia, S. Zhang, C. Si, and S. Zhou, "Double layer LEO satellite based 'BigMAC' space information network architecture," in *Proc. Int. Conf. Space Inf. Netw.* Kunming, China: Springer, 2016, pp. 217–231.

[21] 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.

[22] 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.

[23] H. Cao, Y. Zhu, G. Zheng, and L. Yang, "A novel optimal mapping algorithm with less computational complexity for virtual network embedding," *IEEE Trans. Netw. Service Manag.*, vol. 15, no. 1, pp. 356–371, Mar. 2018.

[24] H. Cao, H. Hu, Z. Qu, and L. Yang, "Heuristic solutions of virtual network embedding: A survey," *China Commun.*, vol. 15, no. 3, pp. 186–219, Mar. 2018.

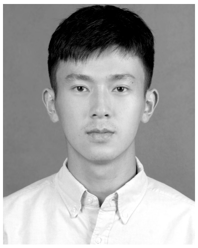
[25] H. Cao, L. Yang, and H. Zhu, "Novel node-ranking approach and multiple topology attributes-based embedding algorithm for single-domain virtual network embedding," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 108–120, Feb. 2018.

[26] S. Q. Kidder, "On the use of satellites in molniya orbit for meteorological and oceanographic observations of the high latitudes," *J. Atmos. Ocean. Technol.*, vol. 7, no. 3, Jun. 1990, Art. no. 517522.

[27] M. Marchese, M. Rossi, and G. Morabito, "PETRA: Performance enhancing transport architecture for Satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 2, pp. 320–332, Feb. 2004.

[28] R. Wang *et al.*, "Unreliable CCSDS file delivery protocol (CFDP) over cislunar communication links," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 147–169, Jan. 2010.

[29] A. Azad, M. Halappanavar, S. Rajamanickam, E. G. Boman, A. Khan, and A. Pothen, "Multithreaded algorithms for maximum matching in bipartite graphs," in *Proc. IEEE 26th Int. Parallel Distrib. Process. Symp.*, May 2012, pp. 860–872.



ZHICHENG QU received the B.S. degree in telecommunication engineering from the Xi'an University of Posts and Telecommunications, Xi'an, China, in 2010. He is currently pursuing the Ph.D. degree in satellite communication with the Institute of Satellite Communication, NUPT.

His research interests include physical and link layer issue for satellite communications and space information networks.

GENGXIN ZHANG received the bachelor's, master's, and Ph.D. degrees from the Department of Radio Communication Engineering, Nanjing Institute of Communication Engineering, Nanjing, China, in 1987, 1990, and 1994, respectively. He is currently a Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China.

His research interests include the satellite communications, deep space communications, and space information networks.

TAO HONG was born in Yangzhou, Jiangsu, China, in 1982. He received the Ph.D. degree in communications and information systems from the Nanjing University of Aeronautics and Astronautics, in 2018. He is currently with the Nanjing University of Posts and Telecommunications, Nanjing, China.

His research interests include satellite communications, antenna array, array synthesis, and directional modulation.

HAOTONG CAO (S'17) received the B.S. and Ph.D. degrees in communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2015, and 2018, respectively. He was a Visiting Scholar with Loughborough University, U.K., in 2017. He has served as the TPC Member of multiple IEEE conferences, such as IEEE INFOCOM, IEEE ICC, IEEE GLOBECOM, and IEEE WCNC. He has published multiple IEEE TRANSACTIONS/Journal/Magazine papers, since 2016. His research interests include next generation networks, wireless communication theory, and resource allocation in wired and wireless networks. He received the 2018 Postgraduate National Scholarship of China. He is also serving as the Reviewer of multiple academic journals, such as the IEEE INTERNET OF THINGS JOURNAL, the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, *Computer Networks* (Elsevier), and the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING.

WEI ZHANG received the B.S. degree (Hons.) in communication engineering from Xidian University, Xi'an, China, in 2009, and the M.S. and Ph.D. degrees from the College of Communication Engineering, PLA University of Science and Technology, Nanjing, China, in 2012 and 2016, respectively. In 2016, he joined the Department of Electronic and Optical Engineering, Space Engineering University, Beijing, where he is currently a Lecturer.

His research interests include signal processing in TT&C systems and space information networks.

• • •