Stochastic Geometry-Based Analysis of Cache-Enabled Hybrid Satellite-Aerial-Terrestrial Networks With Non-Orthogonal Multiple Access

Xiaokai Zhang, Bangning Zhang, Kang An, Gan Zheng, *Fellow, IEEE*, Symeon Chatzinotas, *Senior Member, IEEE*, and Daoxing Guo

***Abstract*— Due to the emergence of non-terrestrial platforms with extensive coverage, flexible deployment, and reconfigurable characteristics, the hybrid satellite-aerial-terrestrial networks (HSATNs) can accommodate a great variety of wireless access services in different applications. To effectively reduce the transmission latency and facilitate the frequent update of files with improved spectrum efficiency, we investigate the performance of cache-enabled HSATN, where the user retrieves the required content files from the cache-enabled aerial node (AN) or the satellite with the non-orthogonal multiple access (NOMA) scheme. If the required content files of the user are cached in the AN, the cache-enabled node would serve directly. Otherwise, the user would retrieve the content file from the satellite system, where the satellite system seeks opportunities for proactive content pushing to ANs during the user content delivery phase. Specifically, taking into account the uncertainty of the number and location of ANs, along with the channel fading of terrestrial users, the outage probability and hit probability of the considered network are, respectively, derived based on stochastic geometry. Numerical results unveil the effectiveness of the cache-enabled HSATN with the NOMA scheme and proclaim the influence of key factors on the system performance. The realistic, tractable, and expandable framework, as well as associated methodology, provide both useful guidance and a solid foundation for evolved networks with advanced configurations in the performance of cache-enabled HSATN.**

***Index Terms*— Stochastic geometry, non-orthogonal multiple access, hybrid satellite-aerial-terrestrial networks, content caching.**

I. Introduction

OWING to the inherent advantages in terms of ubiquitous coverage, high throughput, and extreme resilience, hybrid satellite-aerial-terrestrial networks (HSATNs) have been regarded as an emerging architecture by integrating the satellite systems, aerial platforms, and terrestrial infrastructures in the next-generation wireless networks for enhanced mobile broadband (EMBB) or three-dimensional integrated communications (3D-InteCom) [1]–[3]. The emergence of aerial platforms, such as unmanned aerial vehicles (UAVs), airships, and balloons, provides a promising solution for improving the scalability of hybrid networks, where the aerial layer acts as carriers for information acquisition, transmission, and processing [4]–[6]. The satellite and aerial platforms can be seen as non-terrestrial infrastructures, which can be implemented in underserved areas or coexisted with terrestrial infrastructures as a backup transmission routing [7].

There exist a lot of underserved areas, including remote areas, emergency communication, and maritime scenario [8]. And it is difficult to provide telecommunication services by terrestrial networks due to the high construction costs [9]. Hybrid satellite-aerial-terrestrial networks are expected to support the development of seamless mobile communication systems toward a comprehensive sixth-generation (6G) global access infrastructure [10]. Besides, non-terrestrial infrastructures provide downlink connectivity to ground users whose demand cannot be satisfied by terrestrial cell base stations [11]. Non-terrestrial infrastructure is a promising strategy to complement existing terrestrial communication systems, which can quickly establish a flexible access infrastructure on-demand [12].

Manuscript received November 3, 2020; revised February 20, 2021 and June 5, 2021; accepted August 3, 2021. Date of publication August 20, 2021; date of current version February 14, 2022. This work was supported in part by Jiangsu Provincial Natural Science Foundation of China under Grant BK20191328; in part by the National Natural Science Foundation of China under Grant 61901502, Grant 62001517, and Grant 62071352; in part by the National Postdoctoral Program for Innovative Talents under Grant BX20200101; in part by the Research Project of the National University of Defense Technology (NUDT) under Grant ZK18-02-11 and Grant 18-QNCXJ-029; and in part by U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/N007840/1. The associate editor coordinating the review of this article and approving it for publication was W. Saad. *(Corresponding author: Daoxing Guo.)*

Xiaokai Zhang, Bangning Zhang, and Daoxing Guo are with the College of Communications and Engineering, Army Engineering University of PLA, Nanjing 210007, China (e-mail: xiaokaizhang@foxmail.com; aeu_zbn@163.com; xyzgfg@sina.com).

Kang An is with the Sixty-Third Research Institute, National University of Defense Technology (NUDT), Nanjing 210007, China (e-mail: ankang89@nudt.edu.cn).

Gan Zheng is with Wolfson School of Mechanical, Electrical, and Manufacturing Engineering, Loughborough University, Loughborough LE11 3TU, U.K. (e-mail: g.zheng@lboro.ac.uk).

Symeon Chatzinotas is with the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 4365 Luxembourg, Luxembourg (e-mail: symeon.chatzinotas@uni.lu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3103499.

Digital Object Identifier 10.1109/TWC.2021.3103499

In the European FP7 ABSOLUTE program [13], the HSATN is pursued to satisfy the requirement of the network capacity through the combination of satellite, aerial/terrestrial communication links. Although the HSATN has advantages in terms of coverage and ease of rapid deployment, the transmission latency is an unavoidable problem for HSATNs due to the long distance between satellites and aerial or terrestrial terminals [13]. Besides, with an increasing number of devices, the scarcity of spectrum resources has gradually become an important issue [14], [15].

A. Background and Motivation

With advances in the launch technologies as well as miniaturization of the satellite, the academic and industrial communities have put great efforts into developing new platforms offering more capacity to enable richer user cases and applications [16]. The demand for infotainment services has experienced an unprecedented increase over the last decade. Therefore, the effective content distributions in the case of rural areas or underdeveloped countries certainly entail an important engineering effort in future satellite communication trends, especially collaborating with 5G infrastructure [17]. However, the current satellite Internet architecture is founded upon a host-centric communication model, such as Internet Protocol (IP)-based architectures [18]. The host-centric networks cannot guarantee the service continuity, as well as proper handling of different content distribution services with high levels of quality of service (QoS)/quality of experience (QoE) [19], [20].

To address this problem, the advent of information-centric networking (ICN) architectures [21] has prompted the shift of traditional networking towards a content-oriented concept, which can be directly implemented to the existing content delivery networks [22]. By defining information at the network layer, ICN favors the deployment of the in-network content caching mechanism, which facilitates the efficient and timely delivery of information to the users [23]. Content caching is one of the most important approaches in ICN architecture, where the popular contents are proactively cached on nodes with caching capabilities. Users can obtain the required content at the nearest node to reduce transmission latency. Moreover, the excessive propagation delay, as the inevitable long distance can also be effectively alleviated in HSATNs. Hence, by incorporating content caching-enabled aerial nodes (ANs) into HSATN, both wide coverage area and lower transmission delay can be achieved.

As a novel paradigm of multiple access techniques, non-orthogonal multiple access (NOMA) transmits signals simultaneously in the same time/frequency resource block, which can be separated at receivers with successive interference cancellation (SIC) [24]–[27]. It encourages spectrum sharing among multiple users, instead of occupying orthogonal resource blocks as in orthogonal multiple access (OMA) [28]. Therefore, NOMA can significantly improve spectral efficiency, reduce transmission delay, and support massive connectivity. The incorporation of the NOMA scheme into HSATN is a promising direction to address some inherent problems.

B. Related Work

A two-layered hybrid satellite-terrestrial relay network (HSTRN) is designed in [29]–[34] for the seamless integration of satellite with terrestrial components, which has been demonstrated as a simple yet effective architecture with improved reliability and coverage. In the foundation of autonomous working and mutual integration, HSTRNs integrate the advantages of satellite networks which could provide long transmission distance, wide coverage without the limitation of surroundings with terrestrial networks that have large network capacity and high rate. ITU-R S.2222 standard [13] is proposed to a cross-layer design based on HSTRN to provide multimedia services and ensure the user QoS requirements. However, the terrestrial relay can be blocked by obstacles and shadowing. As an extension, HSATN is further proposed, where the high and low altitude platforms (HAPs & LAPs) can provide broadband services complementing the terrestrial networks [35]. Compared with base stations or the relays in the terrestrial network, the aerial network has the features of low cost, easy deployment, and large coverage to offer wireless access services on a regional basis [36]. In recent years, we have witnessed a rapid proliferation of unmanned aerial vehicles (UAVs) in all areas due to the advance in their payload capacity and prolonged battery life. There exist some low cost-cost ways to provide a long-time seamless coverage, for example, unmanned airships, high-altitude balloons, and solar-powered drones [6], [37]. One of the famous tests is Project Loon by Alphabet Inc. [38].

The paper in [3] presented a comprehensive review of recent research works concerning space-air-ground integrated networks (SAGINs) from network design and resource allocation to performance analysis and optimization. The paper in [13] proved that HSATNs could provide better services than traditional infrastructures during the emergency, in which the HSATNs were surveyed and the key technologies were discussed from several aspects. A cooperative satellite-aerial-terrestrial system was considered in [39], considering the randomness of satellite and users and employing stochastic geometry, where the coverage probability in non-interference and interference scenarios was studied, and the outage performance was investigated. The paper in [9] investigated transmission for HSATNs from an energy-efficient perspective, where a multi-antenna UAV was employed as a relay to assist the satellite signal delivery. The authors in [14] were the first to study the cross-layer gateway selection problem for data delivery in the HSATN communication architecture with the practical inter-layer link capacity constraint, which achieved the best data delivery performance across multiple layers.

Wireless caching has been proved to be an efficient approach to enhance the system performance with respect to delay, QoS, and QoE for satellite communication [40], [41]. Particularly, the paper in [41] has justified the significant performance improvement of satellite-based overlay caching for terrestrial content delivery networks. The authors in [17] have analyzed content delivery over satellite integrated cognitive radio networks, which focused on throughput, energy efficiency, and quality as key performance metrics in the heterogeneous system. To alleviate the spectrum shortage and meet

the requirements of improved spectral efficiency, the authors in [42] first proposed the cache-enabled relay in HSTRN, which confirmed the substantial performance improvement through caching capability. However, this paper only considers the uncertainty of the channel fading. As far as we know, there is currently no work to derive performance analysis while considering channel fading, the uncertainty of the number, and the location of relays in HSTRNs or HSATNs.

The NOMA scheme has been diffusely investigated in hybrid satellite-terrestrial networks (HSTN) from increasing the spectrum efficiency perspective [43]–[48], which proves the feasibility and effectiveness of the NOMA scheme in HSTN. The authors in [43] investigated the NOMA in the land mobile satellite system and gave the outage performance of the proposed system model. References [44] and [45] implemented the NOMA scheme into cooperative spectrum sharing in HSTN for single relay and multiple relays, respectively. The paper in [46] and [47] introduced the NOMA scheme in HSTRN. The authors in [48] applied the NOMA-based scheme by integrating terrestrial networks and the satellite cooperatively to provide coverage for ground users while reusing the entire bandwidth. However, there is no prior work analyzing the performance of the NOMA from a stochastic geometry perspective in HSTNs.

The current proactive caching strategy assumes that the contents can be pushed to cache-enabled relays during the off-peak time [49]. However, this content cache strategy is not applicable if caches need to be frequently updated or rapidly changed files, including latest news, electronic commerce promotion with frequent pricing changes, newly released music videos, etc. [40], [41]. Taking full advantage of the features of NOMA, additional files can be pushed to the ANs simultaneously, i.e. the satellite could seek opportunities for content pushing during the content delivery phase by the NOMA scheme. This strategy is important to efficiently use the limited resources reserved for content pushing, which facilitates the frequent update of the files cached at the ANs and improves the cache hit probability. Besides, the utilization of NOMA not only improves the reliability of content delivery but also ensures that more user requests can be served during the content delivery phase.

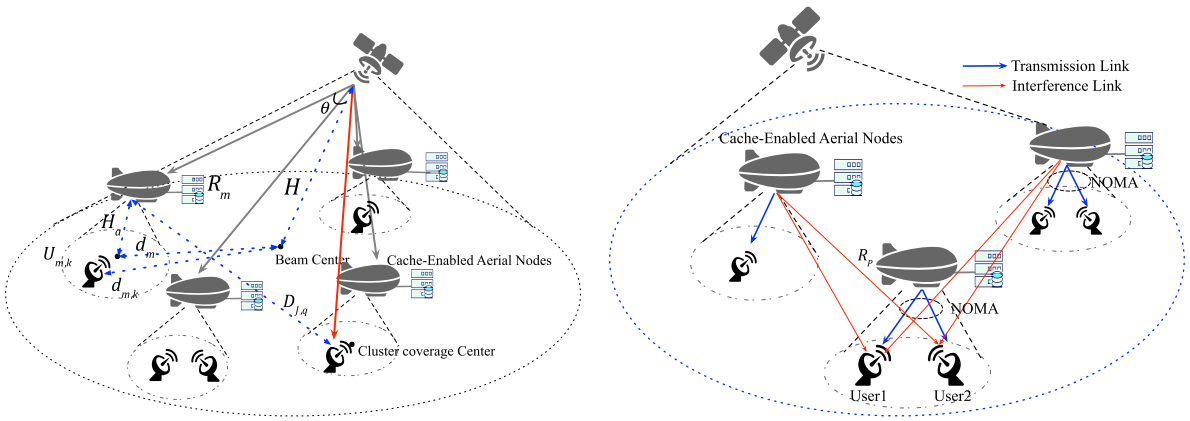
The deterministic and location-based models that have previously been utilized to analyze hybrid satellite-aerial-terrestrial networks (HSATNs) are typically restricted to support simple network topology [39]. However, with the increase in the number of user terminals and the uncertain distribution of locations, the wireless network topology gradually shows the characteristics of randomization and dynamics [24], [50]. To make the theoretical analysis as close to reality as possible, it is no longer possible to accurately characterize the characteristics of random networks based on determining the network topology. In recent years, the use of Stochastic Geometry to describe the randomness of the spatial location for nodes in wireless networks has been recognized by the industry [51], [52], which can depict network interference and analyze average performance from a mathematical perspective. Stochastic geometry plays a vital role in evaluating the system-level performance of wireless networks, where random spatial point patterns describe the nature of wireless systems statistically [53]. Therefore, we take the uncertainty of the channel, the number, and the location into consideration to evaluate the performance of the cache-enabled HSATN in this paper.

C. Contribution and Organization

Motivated by the above observations, this paper investigates the stochastic geometry-based performance of cache-enabled HSATN with the NOMA scheme. The main contributions of this paper are summarized as follows:

- *1) Modeling the Cache-Enabled HSATN With NOMA:* A generic framework of cache-enabled HSATN is first modeled, where the user would retrieve the required content files. If the required content is not cached in the AN, the user tends to retrieve the content file from the satellite, where the satellite simultaneously seeks opportunities for proactive content pushing to the ANs during the user content delivery phase by implementing the NOMA scheme to facilitate the frequent cache updating. Due to the heterogeneity of different layers with various physical mechanisms, the propagation properties and mobility are considered for modeling and constructing the channels between each layer. The proposed framework can achieve enlarged coverage areas with increased spectrum efficiency and reduced latency.
- *2) The Performance Metric:* For retrieving the content files from the satellite, the QoS requirement-based NOMA power allocation strategy is first proposed, which guarantees that the retrieved content files of the user can decode the required files of the user in the AN, which can be seen as a potential relay. Specifically, the outage probability and hit probability of the user and cache-enabled ANs are derived for the proposed power allocation strategy. For retrieving the content files from the cache-enabled ANs, the NOMA scheme is implemented to serve more users, where the co-channel interferences from other ANs are also taken into account. Particularly, a mathematical tractable expression for the outage probability of the user is derived in closed-form.
- *3) System Design Guidelines:* We investigated the impact of the different parameters, including the sequence number of the cache-enabled AN, the transmission power of the satellite, the intensity of ANs, the transmission power of ANs, and the content popularity parameter, etc. Both simulated and analytical results are provided to validate the effectiveness of the cache-enabled HSATN with NOMA and proclaim the influence of key factors on the performance.
- *4) Comparison With OMA:* We investigate the superiority of the NOMA scheme in the HSATRN with the proper NOMA power allocation coefficient in terms of outage probability of the user and hit probability of the cache-enabled AN.

The rest of this paper is organized as follows. Section II presents the stochastic cache-enabled in HSATNs with stochastic geometry and formulates the signal transmission



(a) System model for retrieving the content files from the satellite. (b) System model for retrieving the content files from the cache-enabled node.

Fig. 1. System model.

inspired by the NOMA scheme. Section III and Section IV derive the key performance analysis for retrieving the content files from the satellite and the cache-enabled node, respectively. Section V illustrates the simulation results. Finally, conclusions are drawn in Section VI.

II. Problem Statement

A. System Model

We consider a HSATN with heterogeneous preferences, in which multiple users request contents with the help of the satellite and multiple cache-enabled ANs, where all nodes are equipped with a single antenna and work in the half-duplex mode. If the requested content is not cached in the AN, the user tends to request the content from the satellite as shown in Fig. 1(a), where the satellite can obtain the contents of the entire network from the network control center (NCC) through gateway station connecting to the core network [17]. The NOMA scheme is implemented for delivering content for the user and pushing other currently most popular content to cache-enabled ANs simultaneously,[1] which favors frequently content cache updating. Otherwise, the users request the content from the cache-enabled AN to retrieve the content with alleviated transmission delay, as shown in Fig. 1(b), where the NOMA scheme is implemented in each cluster to increase the spectrum efficiency. The users retrieve the content from the node if the required files exist [42]. This paper mainly considers the downlink of the user content retrieving. The frequency of the transmission is considered in the S-band for two scenarios but on different subchannels.

Without loss of generality, we consider that the distribution of cache-enabled ANs and users obey Poisson Cluster Process (PCP). AN, located in a two-dimensional area, covering a circle with radius $\mathcal{D}$, is considered as homogeneous Poisson point process (HPPP) with an intensity of λ, which is denoted by Φ_c. The height of the ANs is H_a. The m-th neighbor cache-enabled AN from the beam center is denoted by R_m. Moreover, we consider the node as the parent node of the cluster covering a circle whose radius is denoted by $\mathcal{R}$, where the spring users are uniformly distributed in the circle. We assume that there are K users associated with R_m, and the k-th user is denoted by $U_{m,k}$.

[1]The transmission of the ANs for receiving the signal of the satellite and transmitting the signal of the users is divided by different time slots, which avoids the co-channel interference. The transmission of the proposed QoS requirement strategy can be divided into two scenarios. If the user can decode the signal directly, the AN would not retransmit. Otherwise, the AN, acting as a relay, retransmits to the user in the next time slot, which avoids the transmission latency from the satellite in retransmission.

B. Satellite-Link Propagation Model

The propagation factors of satellite links include the free-space path loss, the satellite antenna gain, the receiver antenna gain, the shadowing, and channel fading as well as the other loss. The channel gain G of line-of-sight (LoS) is given by

$$G = \frac{G_R G_S(\theta) f_\varepsilon}{(4\pi H/\lambda_c)^2}, \tag{1}$$

where H is assumed to be the same as the satellite height for simplicity, λ_c denotes the wavelength, f_ε denotes the other loss of the receiver, and the subscript ε denotes the elevation angle from the receiver to the satellite. G_R is the receiving antenna gain. For simplicity of analysis, we consider the receiving antenna gains of the user and the aerial platform are fixed. According to ITU-R S.1528 [54], the LEO satellite antenna gain G_S [dBi] is given by

$$G_S(\theta) = \begin{cases} G_{s,\max} - 3\left(\frac{\theta}{\theta_b}\right)^2 & 0 \leqslant \theta < 1.5\theta_b \\ G_{s,\max} + L_s - 25\log\left(\frac{\theta}{1.5\theta_b}\right) & 1.5\theta_b \leqslant \theta < Y \\ 0 & Y \leqslant \theta < 180^\circ, \end{cases} \tag{2}$$

where $G_{s,\max}$ is the maximum gain at the on-board antenna boresight, $L_N = -6.75dB$, $a = 2.58$, $b = 6.32$, $Y = b\theta_b 10^{0.04(G_{s,\max}+L_N)}$ by ITU-R S.1528, θ_b is one-half of θ_{3dB}, i.e. $\theta_b = 0.5\arctan(L/H)$, L is the half-power beamwidth, θ represents the angular position from the receiver to beam center with respect to the satellite, which can be written as $\theta = \arctan(d/H)$, where d is the distance between the

beam center and the receiver. For $H \gg L$ and $H \gg d$, θ/θ_b is approximately expressed as $\theta/\theta_b \approx 2d/L$. Therefore, the directional beam of the satellite is highly on the position of the AN by d, which[2] can be rewritten as $G_s(d)$.

The shadowing and channel fading of satellite links are considered to follow the Shadowed-Rician fading model [55], which is mathematically tractable and has been widely applied in various fixed and mobile satellite services for a variety of frequency bands. The probability density function (PDF) of $|h|^2$ is shown as [45], [56]

$$f_{|h|^2}(x) = \alpha \exp(-\beta x)\, {}_1F_1(q, 1, \delta x), \tag{3}$$

where h is small scale fading of the satellite link, ${}_1F_1(\cdot,\cdot,\cdot)$ denotes the confluent hypergeometric function and $\alpha = \frac{(2cq)^q}{2c(2cq+\Omega)^q}$, $\delta = \frac{\Omega}{2c(2cq+\Omega)}$, and $\beta = \frac{1}{2c}$, with $2c$ being the average power of the scatter component, Ω being the average power of the LoS component, q being the Nakagami fading parameter. The corresponding cumulative distribution function (CDF) of $|h|^2$ can be obtained as [57]

$$F_{|h|^2}(x) = \alpha \sum_{k=0}^{\infty} \frac{(q)_k \delta^k}{(k!)^2 \beta^{k+1}} \Gamma(k+1, \beta x). \tag{4}$$

We consider that the altitude of the ANs is sufficiently large without the shadowing and obstacles, such as Balloons.[3] Therefore, the channel gain between the satellite and the AN R_m is $G_m = G$. For the user $U_{m,k}$, the LoS satellite communication systems are vulnerable to be blocked by shadowing or obstacles. Therefore, the channel gain between the satellite and user is $G_{m,k} = G|h_{m,k}|^2$ [44].

C. Aerial-Terrestrial Link Propagation Model

The signal transmission of the aerial-terrestrial link is considered to experience the LoS propagation, however, it still is blocked by buildings or bushes. Therefore, the radio signal attenuation probabilistically follows LoS propagation with a lower path-loss exponent and Nakagami-m small-scale fading.

The PDF of the channel power gain between P-th AN R_p and the user q is given as [60]

$$f_{H_{P,q}}(x) = \frac{1}{\Gamma(m^A)(\theta_{P,q})^{m^A}} x^{m^A-1} e^{-\frac{x}{\theta_{P,q}}}, \tag{5}$$

where m^A is the fading severity, which is assumed as integer values in this paper, $\theta_{P,q}$ is the average power, $\Gamma(\cdot)$ denotes the complete gamma function.[4]

[2]Because the height of the LAP is invariant, we can hereby reduce the stochastic model within the aerial layer to two-dimensional. Compared to the distance from the beam center to AN, the height of the LAP can be ignored.

[3]We consider that the Doppler Frequency Shift, caused by the mobility of the satellite, can be estimated and mitigated by the mature pre-compensation method for users [58]. Besides, since the movement of the aerial relay is slow when serving users, the Doppler shift is assumed to be well compensated, which has a negligible impact on the HSATN system [59].

[4]We consider that all channels are assumed to experience independent and identically distributed (i.i.d.) fading and follow quasi-static fading, i.e. the channel gains remain constant within each transmission block but vary independently between different blocks [42], [61].

D. Content Caching Model

The main idea of content placement is that AN cache popular content so that most of the requests from users can be served by ANs directly, reducing the backhaul delay and satellite loading. As for popularity, one of the inherent of those contents is that its request trends tend to follow a "heavy-tailed" distribution [42]. Therefore, we consider that the M files, required to be pushed to ANs, are collected in a finite content catalog $\mathcal{F} = \{f_1, \cdots, f_m, \cdots, f_M\}$. According to the statistic of the popularity of the content, a large fraction of requests occurs for a relatively small fraction of the content, which implies that some contents are popular and requested with high frequency by different users [42]. We consider the popularity of the requested files obeys the *Zipf* distribution for the different ANs. Hence, the popularity of f_m in i-th AN is given by [62]

$$P_i(f_m) = \frac{m^{-\gamma_i}}{\sum_{l=1}^{M} l^{-\gamma_i}}, \tag{6}$$

where $\gamma_i > 0$ denotes the shape parameter defining the content popularity skewness of i-th AN. To be noticed, a large γ_i means the requests on the high popularity files, whereas a small γ_i is related to the requests with heavy-tailed popularity [42]. Notably, the popularity of the content is the probability of the requested file of ANs.[5] Dissimilar from conventional assumptions, we consider that each file has the same length packets, while the information contained in the packets of different files is nonidentical. The reason is that packets, belonging to dissimilar files, have distinct priorities and QoS requirements, which requires the use of different transmission rates for each packet. Specifically, the required data rate for f_m is R^m.

III. Performance Analysis for Retrieving Content From Satellite

If the required content files are not cached in the AN, the satellite would serve the user directly. To improve the spectrum efficiency and increase the frequency of cached file updates, the NOMA scheme is implemented to deliver the content for the user and push other currently most popular contents to cache-enabled ANs simultaneously, which can be seen as "add-on" transmission.

A. NOMA Signal Transmission

By exploiting opportunities for content pushing during the content delivery phase, the satellite would serve the users directly and push the contents to ANs simultaneously by the NOMA scheme. The M most popular files belonging to the same library can be pushed to all cache-enabled ANs. x_0 and x_i denote the signal of the user $U_{m,k}$ and the i-th content cache file, where $\mathrm{E}\left[|x_0|^2\right] = \mathrm{E}\left[|x_i|^2\right] = 1$. The received

[5]The popularity of the content files is statistically calculated, which can be varying timely and written as $\gamma_i(t)$. In this paper, we apply NOMA-based content caching during the content delivery phase to facilitate frequent content updating.

mixed-signal at z can be expressed as

$$y_z = \sqrt{\alpha_0 P_S} G_{S,z} x_0 + \sum_{i=1}^{M} \sqrt{\alpha_i P_S} G_{S,z} x_i + n_0, \quad (7)$$

where $G_{S,z}$ is the power gain between satellite and node z the subscript $z \in \{U_{m,k}, R_m\}$, P_S is the satellite transmitter power, and n_0 is the additive white Gaussian noise (AWGN) with mean zero and variance σ_0^2. α_0 is the NOMA power allocation coefficient of the user, α_i denotes the power allocation coefficient of f_i, where $\sum_{i=0}^{M} \alpha_i = 1$. The receivers implement SIC to separate the multiplexed signal.

According to the property of the NOMA scheme, the power allocation factors and the SIC decoding order would directly determine the performance of the users and ANs [63]. We consider that the decoding order of the received signal is from f_0 to f_M according to the importance and popularity.

B. NOMA Power Allocation Strategy

If f_0 could be successfully decoded on the m-th AN, the m-th AN, seen as a potential relay, can retransmit f_0 to the user $U_{m,k}$ with less delay even the frequent heavy shadowing happens between the satellite and the user $U_{m,k}$. Therefore, the following analysis is based on this strategy, which is named the QoS requirement strategy.

If the $\rho G(d_m) < R^0$, the file f_0 cannot be decoded in the m-th AN, where d_m denotes the distance between m-th AN and the beam center. Therefore, the satellite would allocate all power to the f_0 to decrease the outage probability of the link between the satellite and the user $U_{m,k}$. If the $\rho G(d_m) = R^0$, all power is allocated to the f_0 too. If the $\rho G(d_m) > R^0$, the remaining power could be allocated to other files to increase the spectrum efficiency.

Accordingly, the transmission rate of f_0 in m-th AN is given as

$$R_m^0 = \log_2 \left(1 + \frac{\alpha_0 G(d_m)}{\sum_{i=1}^{M} \alpha_i G(d_m) + \frac{1}{\rho}} \right). \quad (8)$$

Therefore, the power allocation coefficient of f_0 is

$$\alpha_0 = \min \left\{ 1, \frac{\varepsilon_0 \left[\rho G(d_m) + 1 \right]}{\rho G(d_m) (\varepsilon_0 + 1)} \right\}. \quad (9)$$

Furthermore, the remain power coefficient α^r of caching files are given as

$$\alpha^r = \max \left\{ 0, \frac{\rho G(d_m) - \varepsilon_0}{\rho G(d_m) (\varepsilon_0 + 1)} \right\}. \quad (10)$$

To simplify the notation, the new power allocation power coefficient τ_n denotes the ratio of f_n in P_r for $n \in [1, \cdots, M]$, where $\alpha_n = \tau_n \alpha^r$. To reduce the complexity of the system, we consider τ_n is fixed, which can be predefined by the priority or the popularity.

Lemma 1: The distance from m-th AN and n-th AN to the beam center are d_m and d_n, respectively. The joint PDF of d_m and d_n is given by

$$f_{d_m,d_n}(x,y) = 4y(\lambda\pi)^n e^{-\lambda\pi y^2} \frac{x^{2n-1}\left(y^2 - x^2\right)^{m-n-1}}{(m-n-1)!\,(n-1)!}. \quad (11)$$

Proof: See Appendix A. ■

According to the QoS requirement of the m-th AN, we would investigate the performance of the AN first.

C. The Performance of the ANs

In the QoS requirement strategy, f_0 is required to be successfully decoded in the m-th aerial. Therefore, the outage and hit probability are highly related to the probabilities of α_0, and their relation is summarized in *Theorem 1* below. The outage probability is defined as the probability that the instantaneous SINR γ_k falls below a predefined threshold γ_{th}, i.e.,

$$P_{out}(\gamma_{th}) = P(\gamma_k < \gamma_{th}), \quad (12)$$

where $P(\cdot)$ is the probability of the corresponding event. Furthermore, the hit probability of n-th AN for the content category $\mathcal{F}$ is defined as

$$P_n^{hit} = \sum_{i=1}^{M} P_n(f_i) \left(1 - P_n^i\right), \quad (13)$$

where P_n^i is the outage probability of f_i in n-th AN.

Theorem 1: If the power allocation coefficient α_0 makes sure f_0 exactly decodable in the m-th AN, the n-th AN can successfully decode the f_0 for $\forall n < m$ and the p-th AN cannot decode the f_0 for $\forall p > m$.

Proof: The transmission rate of f_0 for k-th AN is rewritten as

$$R_k^0 = \log_2 \left(1 + \frac{\alpha_0}{1 - \alpha_0 + \frac{1}{G(d_k)\rho}} \right). \quad (14)$$

It can be observed in (14) that R_k^0 decreases monotonically with d_k. Due to $n < m < p$, and therefore $d_n < d_m < d_p$, we have $R_n^0 > R_m^0 > R_p^0$. According to the assumption $R_m^0 = R^0$, there exist $R_n^0 > R^0$ and $R_p^0 < R^0$, i.e. the n-th AN can successfully decode f_0 while the p-th relay cannot decode f_0. *Theorem 1* is proved. ■

If f_0 cannot be decoded, the hit probability of each file for p-th AN ($\forall p > m$) would be 0 since SIC cannot achieve. Therefore, we mainly focus on the hit probability for m-th AN and n-th AN ($\forall n < m$).

Lemma 2: The outage probability of m-th cache-enabled AN of the f_0 is given by

$$P_m^0 = e^{-\lambda\pi\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^2} \sum_{k=0}^{m-1} \frac{(\lambda\pi)^k \left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^{2k}}{k!}, \quad (15)$$

and the outage probability of m-th cache-enable node of f_i for ($i \in [1, \cdots, M]$)

$$P_m^i = e^{-\lambda\pi\left(G^{-1}\left(\frac{\varepsilon_0}{\rho} + \frac{(1+\varepsilon_0)\eta_i}{\rho}\right)\right)^2} \times \sum_{k=0}^{m-1} \frac{(\lambda\pi)^k \left(G^{-1}\left(\frac{\varepsilon_0}{\rho} + \frac{(1+\varepsilon_0)\eta_i}{\rho}\right)\right)^{2k}}{k!}, \quad (16)$$

where $\eta_i = \max\left\{\frac{\varepsilon_1}{\omega_1}, \cdots, \frac{\varepsilon_i}{\omega_i}\right\}$, and $\omega_i = \tau_i - \varepsilon_i \sum_{j=i+1}^{M} \tau_j$.

Proof: See Appendix B. ■

Lemma 3: If $\varepsilon_M \geqslant \varepsilon_0$, the outage probability of n-th cache-enabled node of the f_0 is given by

$$P_n^0 = e^{-\lambda\pi\left(G_n^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^2} \sum_{k=0}^{n-1} \frac{(\lambda\pi)^k \left(G_n^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^{2k}}{k!}, \quad (17)$$

and the outage probability of n-th cache-enabled node of the f_i for ($i \in [1, \cdots, M]$) is given by

$$\begin{aligned} P_n^i &= P_m^0 + \frac{4(\lambda\pi)^n e^{-\lambda\pi y^2}}{(m-n-1)!\,(n-1)!} \sum_{k=0}^{t-m-1} (-1)^k \binom{t-m-1}{k} \\ &\quad \times \sum_{l=1}^{N} \frac{\pi(\varphi_2-\varphi_1)}{2N} \sqrt{1-\omega_l^2}\, Z\left(\frac{\varphi_2-\varphi_1}{2}\omega_l + \frac{\varphi_2+\varphi_1}{2}\right), \end{aligned} \quad (18)$$

where $\varphi_1 = G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i+\varepsilon_0}{\rho}\right)$, $\varphi_2 = G^{-1}(\varepsilon_0/\rho)$, $\mu_1 = G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i}{\rho-\frac{\varepsilon_0}{G(y)}}\right)$, $Z(y)$ is given by

$$Z(y) = \frac{e^{-\lambda\pi y^2} y^{2n-2m-2k-2}}{2n+2k}\left(y^{2n+2k} - \mu_1(y)^{2n+2k}\right). \quad (19)$$

N is the Chehysbev-Gauss approximation parameter, and $\omega_l = \cos\left(\frac{2l-1}{2N}\right)$.

Proof: See Appendix C. ∎

Based on the given outage probability P_n^i in *Lemma 3*, the corresponding cache hit probability for a user associated with R_k can be derived by (13).

D. The Outage Probability of the User

The outage probability of the user for the QoS requirement strategy is determined by the NOMA power allocation coefficient. When $\alpha_0 = 1$, the outage probability of the user $U_{m,k}$ is given as

$$\begin{aligned} P_{m,k}^{out} &= P\left(\rho G(d_{m,k})|h_{m,k}|^2 < \varepsilon_0\right) \\ &= P\left(|h_{m,k}|^2 < \frac{\varepsilon_0}{\rho G(d_{m,k})}\right) \\ &= \mathrm{E}\left[F_{|h|^2}\left(\frac{\varepsilon_0}{G(d_{m,k})\rho}\right)\right]. \end{aligned} \quad (20)$$

where $d_{m,k}$ is the distance between the $U_{m,k}$ and the beam center, $h_{m,k}$ denotes the fading between the satellite and $U_{m,k}$.

Conditioned on d_m, the PDF of the $d_{m,k}$ is given by [50]

$$f_{d_{m,k}|d_m}(d) = \begin{cases} \frac{2\pi d}{\pi R^2} & d \leqslant R - d_m \\ \frac{2\pi d}{\pi R^2} - \Psi(d, d_m) & R - d_m < d \leqslant \sqrt{R^2 - d_m^2} \\ \Psi(d, d_m) & \sqrt{R^2 - d_m^2} < d < R + d_m, \end{cases} \quad (21)$$

where $\Psi(d, d_m)$ is defined as

$$\Psi(d, d_m) = \frac{2d \arccos\frac{d_m^2 + d^2 - \mathcal{R}^2}{2dd_m}}{\pi\mathcal{R}^2}. \quad (22)$$

According to the conditional probability of (21), the PDF of $d_{m,k}$ is derived by obtaining the PDF of d_m. To avoid the complexity caused by segmented integration, we consider $d = 0$, i.e., the user is located at the beam center, and no AN is located inside the circle $\Re(o, \kappa\mathcal{R})$, i.e., a circle with the beam center located at its origin and radius $\kappa\mathcal{R}$ with $\kappa > 1$, and $d_m \geq \kappa\mathcal{R}$ for all $m \geq 1$. $\mathcal{A}$ represents the ring with inner radius κR and outer radius d. Therefore, the outage probability of the user $U_{m,k}$ is given by

$$P_{m,k}^{out} = \int_{\kappa R}^{\infty} \int_{y-R}^{y+R} F_{|h|^2}\left(\frac{\varepsilon_0}{G(x)\rho}\right) \Psi(x, y) \tilde{f}_{d_m}(y)\, dxdy, \quad (23)$$

where $\tilde{f}_{d_m}$ denotes the PDF of d_m. Since the cache-enabled ANs follow a HPPP, the complementary cumulative distribution function (CDF) of d_m is the probability that there are less than m ANs closer than d in $\mathcal{A}$, which is given by

$$\begin{aligned} \tilde{F}_{d_m}(d) &= 1 - P(|\mathcal{A}| < m) \\ &= 1 - \sum_{i=1}^{m-1} \frac{\left[\lambda_c\left(\pi d^2 - \pi\delta^2\mathcal{R}^2\right)\right]^i}{i!} e^{-\lambda_c\left(\pi d^2 - \pi\delta^2\mathcal{R}^2\right)}, \end{aligned} \quad (24)$$

where $|*|$ denotes the number of ANs in the corresponding areas. Furthermore, the complementary PDF of d_m can be calculated as[6]

$$\tilde{f}_{d_m}(d) = 2\pi\lambda_c^m d e^{-\lambda_c\left(\pi d^2 - \pi\delta^2\mathcal{R}^2\right)} \frac{\left(\pi d^2 - \pi\delta^2\mathcal{R}^2\right)^{m-1}}{(m-1)!}. \quad (25)$$

Substituting (25) into (23), the outage probability of the user can be obtained for $\alpha_0 = 1$.

When $\alpha_0 < 1$, the outage probability of the user $U_{m,k}$ is given as

$$\begin{aligned} P_{m,k}^{out} &= P\left(\frac{\alpha_0}{1-\alpha_0 + \frac{1}{\rho G(d_{m,k})|h_{m,k}|^2}} < \varepsilon_0\right) \\ &= P\left(\frac{\frac{\varepsilon_0[\rho G(d_m)+1]}{\rho G(d_m)(\varepsilon_0+1)}}{\frac{\rho G(d_m)-\varepsilon_0}{\rho G(d_m)(\varepsilon_0+1)} + \frac{1}{\rho G(d_{m,k})|h_{m,k}|^2}} < \varepsilon_0\right) \\ &= P\left(\frac{G(d_{m,k})|h_{m,k}|^2}{\varepsilon_0\left[G(d_m) - G(d_{m,k})|h_{m,k}|^2\right] + G(d_{m,k})|h_{m,k}|^2} < 1\right) \\ &= P\left(|h_{m,k}|^2 < \frac{G(d_m)}{G(d_{m,k})}\right) \\ &= \mathrm{E}\left[F_{|h|^2}\left(\frac{G(d_m)}{G(d_{m,k})}\right)\right]. \end{aligned} \quad (26)$$

Therefore, the outage probability of the user for $\alpha_0 < 1$ is given as

$$P_{m,k}^{out} = \int_{\kappa R}^{\infty} \int_{y-R}^{y+R} F_{|h|^2}\left(\frac{G(y)}{G(x)}\right) \Psi(x, y) \tilde{f}_{d_m}(y)\, dxdy. \quad (27)$$

Theorem 2: The outage probability of the user $U_{m,k}$ is the same for NOMA and TDMA in the QoS requirement strategy.

[6]The difference between (25) and (36) is due to the assumption that the user is located in the beam center to avoid the trivial piecewise integration.

And the outage probability of the user $U_{m,k}$ is irrelevant to the total transmission power of the satellite when the QoS requirement is satisfied.

Proof: For the situation $\alpha_0 = 1$, all power is allocated into f_0, hence the NOMA scheme is TDMA scheme. For the situation $\alpha_0 < 1$, the outage probability of the TDMA is given by

$$\begin{aligned} P_{m,k}^{out} &= P\left(R^0 \frac{\log_2\left(1+\rho G(d_{m,k})|h_{m,k}|^2\right)}{\log_2(1+\rho G(d_m))} < R^0\right) \\ &= P\left(\frac{G(d_{m,k})|h_{m,k}|^2}{G(d_m)} < 1\right) \\ &= P\left(G(d_m) > G(d_{m,k})|h_{m,k}|^2\right) \\ &= P\left(|h_{m,k}|^2 < \frac{G(d_m)}{G(d_{m,k})}\right) \\ &= \mathrm{E}\left[F_{|h|^2}\left(\frac{G(d_m)}{G(d_{m,k})}\right)\right]. \end{aligned} \tag{28}$$

Therefore, the outage probability of the user $U_{m,k}$ is the same for NOMA and TDMA. According to (26) and (28), the outage probability of the user $U_{m,k}$ is irrelevant to the total transmission power of the satellite.

Hence, *Theorem 2* is proved. ■

IV. Performance Analysis for Retrieving Content From AN

In this section, we consider that the required content file is cached in the AN. Hence, the AN would serve the user to reduce the transmission latency, and the NOMA scheme is implemented to increase the spectrum efficiency.

A. NOMA Signal Transmission

To reduce the transmission latency of retrieving the required content, the user would request the content files from its aerial server. If the required content file is cached in the AN, the AN would serve the user directly. We assume that each node can find at least two users whose requests can be accommodated in its coverage area. By employing the NOMA scheme, the other AN would cause interference to the users.[7] Without loss of generality, we consider that Z_P users are required to be served in one resource block by AN R_P. To improve spectrum efficiency, frequency reuse is implemented among each cluster. Therefore, the received mixed signal can be expressed as

$$y_{P,q} = \underbrace{P_A H_{P,q} \sum_{l_P=1}^{Z_P} \alpha_{l_P}^A x_{l_P}}_{\text{Signals}} + \underbrace{\sum_{J\in\Phi_c\setminus P} P_I H_{J,q} \sum_{l_J=1}^{Z_J} \alpha_{l_J}^A x_{l_J}}_{\text{Interference}} + n_1, \tag{29}$$

where P_A is the transmitted power of P-th AN, P_I is the transmitted power of other ANs, where the transmission power of other nodes is the same to simplify the expression, Φ_c is the HPPP set of AN, $\alpha_{l_X}^A$ is the NOMA power allocation for l-th user of X-th AN, where $\sum_{l_x=1}^{Z_x} \alpha_{l_x}^A = 1$, $H_{x,y}$ is the link between node x and node y, and n_1 is the AWGN with mean zero and variance σ_1^2, x_{l_X} denotes the signal for l-th user of X-th AN, where $\mathrm{E}\left[|x_{l_X}|^2\right] = 1$, and $\mathrm{E}[\cdot]$ means the expectation operation.

[7]Here we consider the general situation, which constitutes the worst case for the reception reliability of users. If no user requests the content for one AN, there exists no interference. If the number of the user is large, the available method is that the users are divided into several clusters by implementing NOMA in different resource blocks.

B. Outage Probability of Users

To obtain tractable analytical results, we consider that the SIC decoding order is from x_{1_P} to x_{Z_P} and the NOMA power allocation coefficient is predefined. In such a general case, the transmission rate of user $U_{P,m}$ is given by

$$R_{U_{P,m}} = \log_2\left(1 + \frac{\rho_A H_{P,m}\alpha_m^A}{\sum_{l=m+1}^{Z_P} \rho_A H_{P,m}\alpha_l^A + \mathbf{I}_{\text{inter}} + 1}\right), \tag{30}$$

where $\rho_A = \frac{P_A}{\sigma_1^2}$, $\rho_I = \frac{P_I}{\sigma_1^2}$ and $\mathbf{I}_{\text{inter}} = \sum_{J\in\Phi_c\setminus P}\rho_I H_{J,m}$.

The successfully decoding probability of user $U_{P,m}$ is give in (31), shown at the bottom of the next page, where $\varepsilon_{P,i}^A = 2^{R_{P,i}^A} - 1$, $R_{P,i}^A$ is the required transmission rate of i-th user in P-th AN coverage area, and $\xi_{P,m} = \max\left\{\frac{\varepsilon_{P,i}^A}{\alpha_i^A - \varepsilon_{P,1}^A \sum_{l=i+1}^{Z_P}\alpha_l^A}, \forall i \in (1,\ldots,m)\right\}$, $\mathcal{L}_{\mathbf{I}}(s)$ is the Laplace transform (LT) of $\mathbf{I}$, and $\mathcal{L}_{\mathbf{I}}^{(i)}(s)$ is i-th derivative of $\mathcal{L}_{\mathbf{I}}(s)$. The radio signal attenuation follows LoS propagation with a lower path-loss exponent and Nakagami-m small-scale fading for the aerial-terrestrial link, and the receive side signal-to-noise ratio (SNR) $\rho H_{P,q}$ follows gamma distributions [60]. According to (5), $\theta_{P,q}$ is defined as $\rho D_{P,q}^{-\alpha_L} \Big/ m_{P,q}$. Therefore, the step (a) in (32), shown at the bottom of the next page, apply the CDF of gamma distribution. The step (b) in (32) is due to $E_{\mathbf{I}}\left[\mathbf{I}^i \exp(-s\mathbf{I})\right] = (-1)^i \mathcal{L}_{\mathbf{I}}^{(i)}(s)$, where $\mathcal{L}_{\mathbf{I}_{\text{inter}}}^{(i)}(s_0) = \left[\frac{d^i}{ds^i} L_{\mathbf{I}_{\text{inter}}}(s)\right]_{s=s_0}$.

To calculate $\mathcal{L}_{\mathbf{I}_{\text{inter}}}^{(i)}(s)$, we utilize the Faa di Bruno formula [60], which is expressed as (33), shown at the bottom of the next page, to deal with the higher order derivatives of the composite function by calculating the LT of the $\mathbf{I}_{\text{inter}}$. $\mathcal{L}_{\mathbf{I}_{\text{inter}}}(s)$ is given as (34), shown at the bottom of the next page, where $D_{J,q}$ is distance between the other AN and the user. The step (c) in (34) utilizes the probability generating functional (PGFL) of the HPPP. The i-th derivative is formulated as (35), shown at the bottom of the next page, where $\mathcal{N}_i = \left\{(n_1 \cdots, n_\tau, \cdots n_i)\,|n_\tau \in \mathbf{N}, \sum_{\tau=1}^{i} \tau n_\tau = i\right\}$.

V. Numerical Simulations

In this section, representative simulation results are presented to confirm the effectiveness of the considered HSATN systems and evaluate the impact of key factors on the performance of the proposed framework. The AN distribution radius $\mathcal{D}$, the height of the AN, and the AN coverage area radius $\mathcal{R}$ are 200 kilometers (Km), 5 Km, and 5 Km,

TABLE I
SIMULATION PARAMETERS

Parameters	Values
Center frequency	4 GHz
Satellite height	800 Km
ANs distribution radius $\mathcal{D}$	200 Km
AN coverage radius $\mathcal{R}$	5 Km
Half-power beamwidth of satellite L	50 Km
Satellite antenna gain G_s	20 dBi
Cluster heads antenna gain G_t	6 dBi
PRs antenna gain G_s	5 dBi
Terrestrial noise temperature	300 K
Satellite noise temperature	350 K

respectively, where the users are randomly distributed in the AN coverage area. The key system parameters are provided in Table I.

A. Simulation Results for Retrieving the Content Files From the Satellite

We consider that the file category $\mathcal{F} = \{f_1, f_2, f_3\}$, and transmission rate requirement are $\{0.75, 0.875, 1\}$ (bps/Hz), respectively.[8] The required rate for f_0 of the user $U_{m,k}$ is 0.6 bps/Hz.

As α_0 is determined by the channel gain of m-th AN, the ratio τ_i of the remaining power for $\{f_1, f_2, f_3\}$ are $\{0.5, 0.3, 0.2\}$, respectively. Besides, the OMA scheme is provided for comparison, where the time slot allocation for f_0 is determined by the QoS requirement of the user $U_{m,k}$,

[8]Due to the existence of interference in NOMA scheme, a single time slot is only available for limited content files. If the number of content files is large, the content files can be divided into different groups for different time slots.

$$P_{U_{P,m}} = P\left(\frac{\rho_A H_{P,q}\alpha_1^A}{\sum_{l=2}^{Z_P}\rho_A H_{P,q}\alpha_l^A + \mathbf{I}_{\text{inter}} + 1} > \varepsilon_{P,1}^A, \cdots, \frac{\rho_A H_{P,q}\alpha_m^A}{\sum_{l=m+1}^{Z_P}\rho_A H_{P,q}\alpha_l^A + \mathbf{I}_{\text{inter}} + 1} > \varepsilon_{P,m}^A\right)$$
$$= P\left(\rho_A H_{P,q} > \xi_{P,m}\left(\mathbf{I}_{\text{inter}} + 1\right)\right) \tag{31}$$

$$P_{U_{P,m}} \overset{a}{=} E_{\mathbf{I}_{\text{inter}}}\left[\exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\left(\mathbf{I}_{\text{inter}}+1\right)\right)\sum_{k=0}^{m^A-1}\tfrac{1}{k!}\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\left(\mathbf{I}_{\text{inter}}+1\right)\right)^k\right]$$
$$= \exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)E_{\mathbf{I}_{\text{inter}}}\left[\exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\mathbf{I}_{\text{inter}}\right)\sum_{k=0}^{m^A-1}\tfrac{1}{k!}\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)^k\sum_{i=0}^{k}\binom{k}{i}\mathbf{I}_{\text{inter}}^i\right]$$
$$= \exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)\sum_{k=0}^{m^A-1}\tfrac{1}{k!}\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)^k\sum_{i=0}^{k}\binom{k}{i}E_{\mathbf{I}_{\text{inter}}}\left[\exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\mathbf{I}_{\text{inter}}\right)\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\mathbf{I}_{\text{inter}}\right)^i\right]$$
$$\overset{b}{=} \exp\left(-\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)\sum_{k=0}^{m^A-1}\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\right)^k\sum_{i=0}^{k}\binom{k}{i}\tfrac{(-1)^i}{k!}L_{\mathbf{I}_{\text{inter}}}^{(i)}\left(\tfrac{\xi_{P,m}}{\theta_{P,m}}\right). \tag{32}$$

$$f^{(\kappa)}\left(g\left(s\right)\right) = \sum_{(n_1,\cdots,n_\kappa)\in\mathcal{N}_\kappa}\frac{\kappa!}{\prod_{\tau=1}^{\kappa}\left(n_\tau!\right)}f^{(n_1+\cdots+n_\kappa)}\left(g\left(s\right)\right)\prod_{\tau=1}^{\kappa}\left(\frac{g^{(\tau)}\left(s\right)}{\tau!}\right)^{n_\tau}. \tag{33}$$

$$\mathcal{L}_{\mathbf{I}_{\text{inter}}}(s) = E\left(\exp(-s\sum_{J\in\Phi_c\setminus P}\rho_I H_{J,q})\right) = E_{\Phi_c\setminus P}\left(\exp(-s\sum_{J\in\Phi_c\setminus P}\rho_I D_{J,q}^{-\alpha^A}\left|h_{J,q}\right|^2)\right)$$
$$\overset{c}{=} E_{\Phi_c\setminus P}\left[\prod_{J\in\Phi_c\setminus P}E\left[\frac{1}{\left(1+\frac{s\rho_I}{m^A}D_{J,q}^{-\alpha^A}\right)^{m^A}}\right]\right] = \exp\left[-2\pi\lambda\int_{H_a}^{\infty}\frac{\sum_{n=1}^{m^A}\binom{m^A}{n}\left(\frac{s\rho_I}{m^A}\right)^n z^{-n\alpha^A}}{\left(1+\frac{s\rho_I}{m^A}D_{J,q}^{-\alpha^A}\right)^{m^A}}zdz\right]$$
$$= \exp\left(-2\pi\lambda\sum_{n=1}^{m^A}\binom{m^A}{n}\left(\tfrac{s\rho_I}{m^A}\right)^n\tfrac{H_a^{2-n\alpha^A}}{n\alpha^A-2}{}_2F_1\left(m^A, n-\tfrac{1}{\alpha^A}, n-\tfrac{1}{\alpha^A}+1, \tfrac{s\rho_I}{m^A H_a^{\alpha^A}}\right)\right). \tag{34}$$

$$\mathcal{L}_{\mathbf{I}_{\text{inter}}}^{(i)}(s) = \sum_{(n_1,\cdots,n_i)\in\mathcal{N}_i}\frac{i!}{\prod_{\tau=1}^{i}\left(n_\tau!\right)}\mathcal{L}_{\mathbf{I}_{\text{inter}}}(s)$$
$$\times\prod_{\tau=1}^{i}\left(2\pi\lambda(-1)^\tau\binom{m^A-1+i}{m^A-1}\left(\frac{\rho_I}{m^A}\right)^\tau\tfrac{H_a^{2-n\alpha^A}}{n\alpha^A-2}{}_2F_1\left(m^A+\tau, \tau-\frac{2}{\alpha^A}, \tau-\frac{2}{\alpha^A}+1, \tfrac{s\rho_I}{m^A H_a^{\alpha^A}}\right)\right)^{n_\tau}. \tag{35}$$

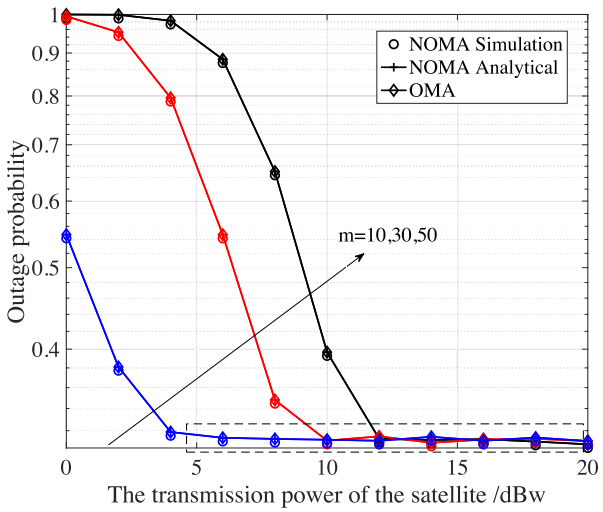


Fig. 2. The outage probability of the user $U_{m,k}$ versus different m, where $\lambda = 1.27 \times 10^{-4}/km^2$.

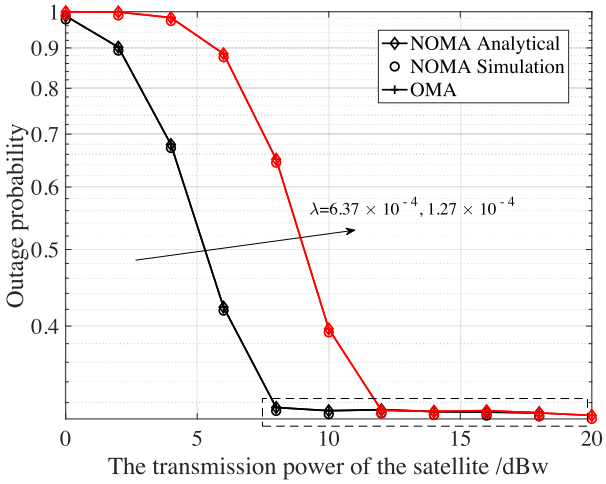


Fig. 3. The outage probability of the user $U_{m,k}$ located in 50-th AN coverage versus different intensity λ, where $\lambda = 1.27 \times 10^{-4}/km^2$ and $\lambda = 6.37 \times 10^{-4}/km^2$, respectively, and the shadowing and channel fading is ILS.

and the rest time slot allocation is proportionally allocated according to the rate requirements of each file.[9]

Fig. 2 illustrates the f_0 outage probability of the user $U_{m,k}$ for the different orders of m. The shadowing and channel fading is considered as ILS, where the corresponding channel coefficient (c, q, Ω) in (3) are given as $(0.158, 19.4, 1.29)$, respectively [45], [55]–[57]. The analytical results match the simulation results, which verifies the accuracy of the outage probability analysis. When the power is insufficient for decoding the f_0 in the AN R_m, i.e. $\alpha_0 = 1$, the outage probability of the user is decreasing with the transmitted power increasing. In this situation, the outage probability of the user achieves superior performance with a shorter distance from the beam center. The main reason is that the antenna gain of the satellite is large for the user closing to the beam center, which decreases the outage probability of the user. In addition, the user's outage probability for f_0 of the NOMA and TDMA is the same in the QoS requirement strategy, which proves the correctness of the *Theorem 2*.

Fig. 3 shows the outage probability of the user under the different intensities of the AN. When the power is not sufficient for decoding the f_0 in the AN R_m, i.e. $\alpha_0 = 1$, the dense intensity would decrease the outage probability. When the QoS of the f_0 is satisfied in R_m, the intensity of the AN has only a limited impact on the outage probability of the user $U_{m,k}$ as shown in the dotted rectangle area of Fig. 2 and Fig. 3.

Fig. 4 depicts the hit probability of 50-th AN for the content category $\mathcal{F}$. Although the outage probability of the user $U_{m,k}$ is the same for NOMA and TDMA, the hit probability of content category $\mathcal{F}$ in the NOMA scheme is superior to the OMA benchmark for R_m, which indicates the effectiveness of the NOMA scheme.

Fig. 5 shows the f_0 outage probability of the aerial R_n within different order n when the f_0 cannot be decoded in

[9]In the TDMA, each user occupies orthogonal specific time blocks in practice system. For a clear comparison with NOMA, we consider the time blocks can be divided into several sub-time slots during the simulation.

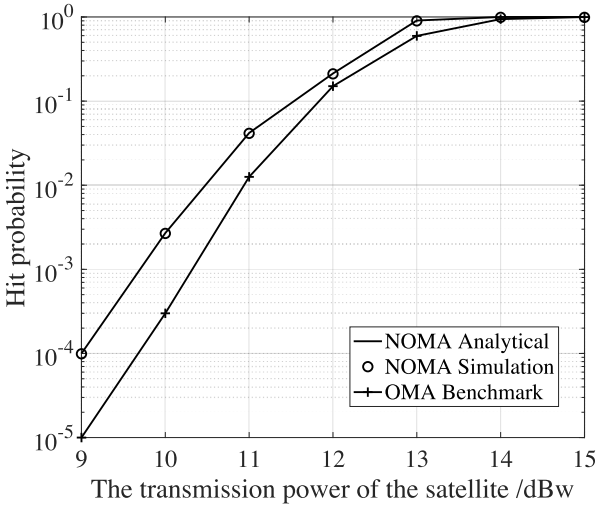


Fig. 4. The hit probability of the AN R_m, where $m = 50$, $\lambda = 7.96 \times 10^{-4}/km^2$ and $\gamma_m = 0.5$.

R_m under the transmitted power. As can be observed from the figure, increasing n will ascend the outage probability due to the larger distance from the beam center. When the f_0 is decoded in R_m, the outage probability of the f_0 in the AN R_n equals zero. The outage probability of f_2 in the AN R_n for different order n is given in Fig. 6. It can be observed that the user achieves superior performance since the shorter distance from the beam center and results in higher beam gain.

Fig. 7 and Fig. 8 depict the hit probability for the content category $\mathcal{F}$ of the AN within the different content popularity parameter γ and order n. Since the expected hit probability of the OMA scheme is the same for different γ, we only plot one curve of TDMA for instance. A large γ means the requests on the high popularity files, whereas a small γ is related to the requests with heavy-tailed popularity. The NOMA scheme decodes the most popular content file at first. Therefore, the larger γ scenario obtains the better hit probability.

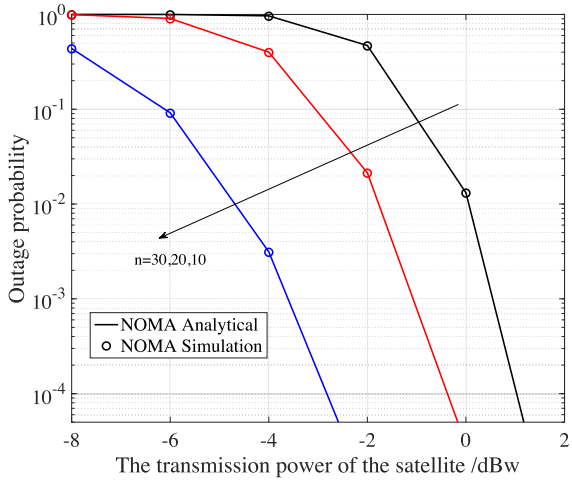


Fig. 5. The outage probability of f_0 in the AN R_n for different n, where $m = 50$ and $\lambda = 7.96 \times 10^{-4}/km^2$.

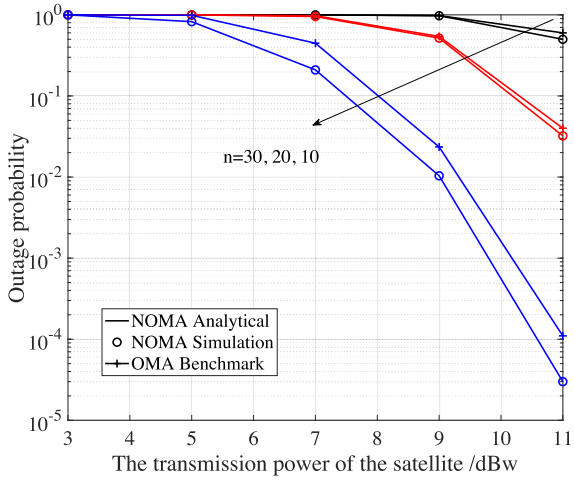


Fig. 6. The outage probability of f_2 in the AN R_n for different n, where $m = 50$ and $\lambda = 7.96 \times 10^{-4}/km^2$.

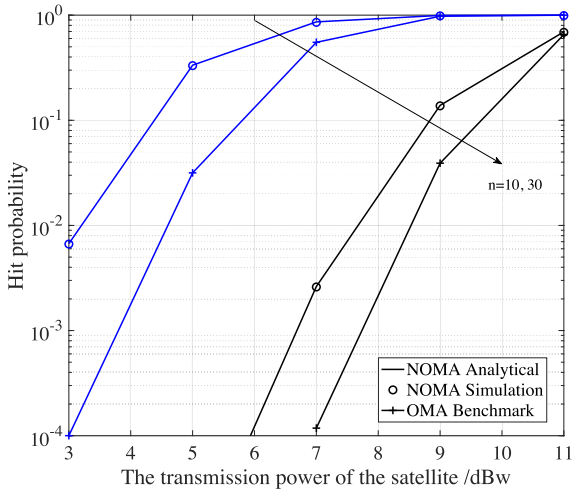


Fig. 7. The hit probability of the AN R_n for the different AN intensity and n, where $\gamma_n = 0.5$, $m = 50$ and $\lambda = 7.96 \times 10^{-4}/km^2$.

B. Simulation Results for Retrieving the Content Files From the Cache-Enabled AN

In this subsection, we evaluate the outage probability of the users for retrieving the content from the AN. We consider

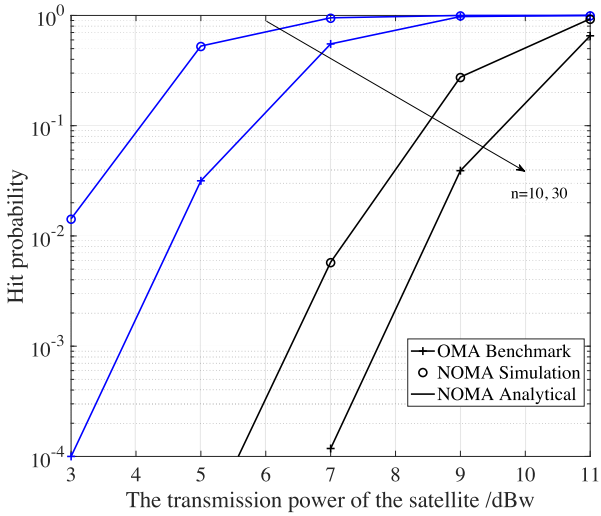


Fig. 8. The hit probability of the AN R_n for the different AN intensity and n, where $\gamma_n = 5$, $m = 50$ and $\lambda = 7.96 \times 10^{-4}/km^2$.

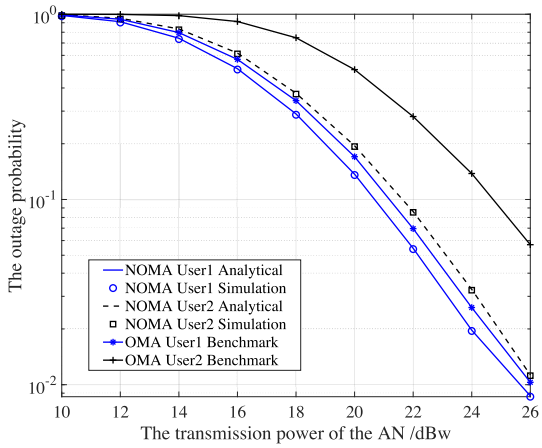


Fig. 9. The outage probability of the two users versus different transmission power of the cache-enabled node, where $\lambda = 1.27 \times 10^{-4}/km^2$.

that the required content files are cached in the AN. Therefore, the AN serves the two users by the NOMA scheme simultaneously. The power allocation coefficient is 0.75 for far user 1 and 0.25 for near user 2. The required transmission rates are $\{1.5, 2\}$ (bps/Hz) respectively. The fading severity m_A of Nakagami-m is set as 3, and the path loss exponent α^A is set as 2 for all aerial-terrestrial link. Besides, the TDMA scheme is provided for comparison, where the same transmission power as the NOMA scheme and the sub-time slot allocation is averagely allocated to two users. The transmission power of the interference nodes is set as 10 W.

Fig. 9 and Fig. 10 depict the outage of the two users in 10-th AN coverage area for the different dense intensities of the AN. Both figures show that the derived outage probability results perfectly match the simulation results for two users, which confirms that our derivations are valid. The outage probability of the NOMA scheme for two users is superior to the OMA scheme, which implies the superiority of the

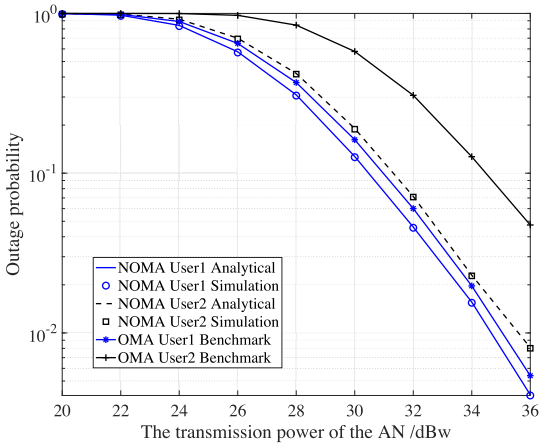


Fig. 10. The outage probability of the two users versus different transmission power of the cache-enabled node, where $\lambda = 1.27 \times 10^{-3}/km^2$.

NOMA scheme. In addition, the higher intensity of the ANs leads to higher outage probability, which is results from more interference from a larger number of ANs.

VI. Conclusion

In this paper, to reduce the transmission latency and facilitate the frequent update of the files cached at the ANs, we investigated the performance of cache-enabled HSATN, where the user retrieved the required content files from the AN or the satellite with the NOMA scheme. If the required content files of the user were cached in the AN, the cache-enabled node would serve the user directly. Otherwise, the user would retrieve the content file from the satellite, where the satellite seeks opportunities for proactive content pushing to the node during the user content delivery phase. Specifically, taking into account the uncertainty of the number and location of ANs, along with the channel fading of users, the outage probability and hit probability of the considered network were derived based on stochastic geometry. Finally, both simulation and analytical results were provided to validate the effectiveness of the cache-enabled HSATN with the NOMA scheme and proclaim the influence of key factors on the system performance.

Appendix A
The Proof of Lemma 1

The number of cache-enabled nodes is followed the HPPP within intensity λ, therefore the PDF of d_m is given by [50]

$$f_{d_m}(y) = \frac{2\lambda^m \pi^m y^{2m-1}}{(m-1)!} e^{-\lambda\pi y^2}. \tag{36}$$

Conditioned on $d_n = x$, the CDF of the d_m is given by

$$\begin{aligned} F_{d_m|d_n}(y) &= P(d_m \le y \,|d_n = x) \\ &= 1 - P(d_m > y \,|d_n = x) \end{aligned} \tag{37}$$

The event $(d_m > y \,|d_n = x)$ indicates that m-th cache-enable node is located in outer of the circle with radius y or inner of the circle with radius x. Without loss of general, we consider $m > n$. Therefore, there exist at most $m-n-1$ cache-enable node inside a ring with inner radius x and outer radius y, where the ring denotes by $\Re(x,y)$. Hence, $F_{d_m|d_n}(y)$ is derived as

$$F_{d_m|d_n}(y) = 1 - \sum_{k=1}^{m-n-1} P(|\Re(x,y)| = k). \tag{38}$$

where $|\Re(x,y)|$ represents the number of the cache-enable located in ring $\Re(x,y)$. According to the property of HPPP, $F_{d_m|d_n}(y)$ is given as

$$F_{d_m|d_n}(y) = 1 - \sum_{k=1}^{m-n-1} e^{-\lambda\pi\left(y^2-x^2\right)} \frac{\lambda^k \pi^k \left(y^2-x^2\right)^k}{k!} \tag{39}$$

We implement $S_k = \frac{\lambda^k \pi^k \left(y^2-x^2\right)^k}{k!}$. The PDF of $f_{d_m|d_n}(y)$ can be obtained by the derivation of CDF $F_{d_m|d_n}(y)$. Based on [50], the conditional PDF for $f_{d_m|d_n}(y)$ is derived as

$$\begin{aligned} &f_{d_m|d_n}(y) \\ &= 2y\lambda\pi e^{-\lambda\pi\left(y^2-x^2\right)} \left(\sum_{k=0}^{m-n-1} S_k - \sum_{k=1}^{m-n-1} S_{k-1} \right) \\ &= 2y(\lambda\pi)^{m-n} e^{-\lambda\pi\left(y^2-x^2\right)} \frac{x^{2m-1}\left(y^2-x^2\right)^{n-m-1}}{(m-n-1)!}. \end{aligned} \tag{40}$$

Therefore, the join PDF for d_m and d_n is obtained by Bayesian formula, which can be given by (11).

Therefore, the *lemma 1* is proved. ■

Appendix B
The Proof of Lemma 2

According to (9), there exists a situation, where $\alpha_0 = 1$, the m-th node still decodes the f_0. The PDF of d_m is given in (36). Therefore, the outage probability of f_0 in m-th node is given by

$$\begin{aligned} P_m^0 &= P\left(G(d_m) < \frac{\varepsilon_0}{\rho}\right) = P\left(d_m > G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right) \\ &= \frac{2(\lambda\pi)^m}{(m-1)!} \int_{G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)}^{\infty} y^{2m-1} e^{-\lambda\pi y^2} dy \\ &= e^{-\lambda\pi\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^2} \sum_{k=0}^{m-1} \frac{(\lambda\pi)^k \left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^{2k}}{k!}. \end{aligned} \tag{41}$$

To be noticed, $G(d_m)\rho < \varepsilon_0$, all power is allocated to the f_0. Hence, $P\left(G(d_m) \le \frac{\varepsilon_0}{\rho}\right)$ and $P(\alpha_0 = 1))$ is the same. The outage probability of f_i in m-th node is given by

$$\begin{aligned} &P_m^i \\ &= P\left(\underbrace{\alpha_0 = 1, G(d_m) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}}_{event1} \right) \\ &\quad + P\left(\underbrace{\alpha_0 < 1, G(d_m) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}}_{event2} \right), \end{aligned} \tag{42}$$

where the $G(d_m) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}$ is always true for $\alpha_0 = 1$ in *event 1*. Therefore $P_{event1} = P(\alpha_0 = 1)$. For the *event 2*, we have

$$\left\{G(d_m) < \frac{\varepsilon_i}{\rho\zeta_i}\right\} = \left\{G(d_m) < \frac{\varepsilon_i}{\rho\omega_i \frac{\rho G(d_m) - \varepsilon_0}{\rho G(d_m)(\varepsilon_0+1)}}\right\} = \left\{G(d_m) < \frac{\varepsilon_0}{\rho} + \frac{(1+\varepsilon_0)\varepsilon_i}{\rho\omega_i}\right\}. \quad (43)$$

Hence, formula (42) can be rewritten as

$$\begin{aligned} P_m^i &= P\left(G(d_m) < \frac{\varepsilon_0}{\rho}\right) \\ &\quad + P\left(G(d_m) < \max\left\{\tfrac{\varepsilon_0}{\rho} + \tfrac{(1+\varepsilon_0)\varepsilon_k}{\rho\omega_k}, \forall k \in [1,\cdots,i]\right\}\right) \\ &= P\left(d_m > G^{-1}\left(\tfrac{\varepsilon_0}{\rho} + \tfrac{(1+\varepsilon_0)\eta_i}{\rho}\right)\right) \\ &= \frac{2(\lambda\pi)^m}{(m-1)!}\int_{G^{-1}\left(\frac{\varepsilon_0}{\rho}+\frac{(1+\varepsilon_0)\eta_i}{\rho}\right)}^{\infty} y^{2m-1}e^{-\lambda\pi y^2}dy \\ &= e^{-\lambda\pi\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}+\frac{(1+\varepsilon_0)\eta_i}{\rho}\right)\right)^2} \\ &\quad \times \sum_{k=0}^{m-1}\frac{(\lambda\pi)^k\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}+\frac{(1+\varepsilon_0)\eta_i}{\rho}\right)\right)^{2k}}{k!}, \end{aligned} \quad (44)$$

where $\eta_i = \max\left\{\frac{\varepsilon_1}{\varsigma_1}, \cdots, \frac{\varepsilon_i}{\varsigma_i}\right\}$, and $\varsigma_i = \tau_i - \varepsilon_i\sum_{j=i+1}^{M}\tau_j$. Hence, *lamma 2* is proved. ■

APPENDIX C
THE PROOF OF LEMMA 3

The premise of n-th node decoding f_0 is that m-th node successfully decode the f_0. Therefore, the outage probability of n-th node successfully decode f_0 is given as

$$\begin{aligned} P_n^0 &= P\left(G(d_m) < \frac{\varepsilon_0}{\rho}, G(d_n) < \frac{\varepsilon_0}{\rho}\right) \\ &= P\left(G(d_n) < \frac{\varepsilon_0}{\rho}\right) \\ &= P\left(d_n > G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right) \\ &= e^{-\lambda\pi\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^2}\sum_{k=0}^{n-1}\frac{(\lambda\pi)^k\left(G^{-1}\left(\frac{\varepsilon_0}{\rho}\right)\right)^{2k}}{k!}. \end{aligned} \quad (45)$$

The outage probability of n-th node for f_i ($\forall i \in [1,\cdots,M]$) is given as

$$\begin{aligned} P_n^i &= P\left(\underbrace{\alpha_0 = 1, G(d_n) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}}_{event3}\right) \\ &\quad + P\left(\underbrace{\alpha_0 < 1, G(d_n) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}}_{event4}\right). \end{aligned} \quad (46)$$

For *event 3*, $\left\{G(d_n) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (0,\cdots,i)\right\}\right\}$ occurs, there must be $\alpha_0 = 1$. Therefore, $P_{event3} = P(\alpha_0 = 1)$. The probability of *event 4* can be rewritten as

$$\begin{aligned} &P\left\{G(d_m) > \frac{\varepsilon_0}{\rho}, G(d_n) < \max\left\{\frac{\varepsilon_k}{\rho\zeta_k}, \forall k \in (1,\cdots,i)\right\}\right\} \\ &\quad = P\left\{G(d_m) > \frac{\varepsilon_0}{\rho}, G(d_n) < \frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G_m(d_m)}}\right\}. \end{aligned} \quad (47)$$

Therefore, the probability of *event 4* is

$$\begin{aligned} P_4 &= P\left(G(d_m) > \frac{\varepsilon_0}{\rho}, G(d_n) < \frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G(d_m)}}\right) \\ &= P\left(d_m < G^{-1}\left(\frac{\varepsilon_0}{\rho}\right), d_n > G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G(d_m)}}\right)\right). \end{aligned} \quad (48)$$

If *event 4* always happens, there must exist

$$G^{-1}\left(\frac{\varepsilon_0}{\rho}\right) > G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G(d_m)}}\right). \quad (49)$$

Since G is monotonically decreasing, G^{-1} is monotonically increasing. We have

$$\frac{\varepsilon_0}{\rho} > \frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G(d_m)}}. \quad (50)$$

Furthermore, (50) is rewritten as

$$\frac{1}{G(d_m)} > \frac{\rho}{\varepsilon_0}\left[1 - \frac{(1+\varepsilon_0)\eta_i}{\varepsilon_0}\right] = \frac{\rho}{\varepsilon_0}[1-\Theta], \quad (51)$$

where Θ can be formulated as

$$\begin{aligned} \Theta &= \frac{(1+\varepsilon_0)\max\left\{\frac{\varepsilon_1}{\varsigma_1}, \cdots, \frac{\varepsilon_M}{\varsigma_M}\right\}}{\varepsilon_0} \geqslant \frac{(1+\varepsilon_0)\varepsilon_M}{\varepsilon_0\varsigma_M} \\ &= \frac{2^{R^0}\varepsilon_M}{\varepsilon_0\varsigma_M}. \end{aligned} \quad (52)$$

The $\varsigma_M < 1$ and $2^{R^0} > 1$ hold. If we consider $\varepsilon_M \geqslant \varepsilon_0$, $\Theta \geq 0$ is obtained. Therefore, the relationship of (51) is always true. According to (48), the *event 4* holds the following constraint

$$d_m > G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i}{\rho - \frac{\varepsilon_0}{G(d_m)}}\right). \quad (53)$$

Further, we have

$$d_m > G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i + \varepsilon_0}{\rho}\right). \quad (54)$$

The joint PDF of d_m and d_n is obtained by *lamma*, the relationship and the range of d_m and d_n are calculated by above observations. Hence, the probability of *event 4* is given as

$$P_{event4} = \int_{\varphi_1}^{\varphi_2}\int_{\mu_1}^{y} f_{d_n,d_m}(x,y)\,dxdy, \quad (55)$$

where $\varphi_1 = G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i+\varepsilon_0}{\rho}\right)$, $\varphi_2 = G^{-1}(\varepsilon_0/\rho)$, $\mu_1 = G^{-1}\left(\frac{(1+\varepsilon_0)\eta_i}{\rho-\frac{\varepsilon_0}{G(y)}}\right)$. The joint PDF $f_{d_n,d_m}(x,y)$ can be expanded by binomial as

$$\begin{aligned} f_{d_n,d_m}(x,y) &= 4y(\lambda\pi)^n e^{-\lambda\pi y^2}\frac{x^{2n-1}\left(y^2-x^2\right)^{m-n-1}}{(m-n-1)!\,(n-1)!} \\ &= \frac{4(\lambda\pi)^n e^{-\lambda\pi y^2}}{(m-n-1)!\,(n-1)!}\sum_{k=0}^{m-n-1}(-1)^k \\ &\quad\times\binom{m-n-1}{k}x^{2n+2k-1}y^{2m-2n-2k-2}. \end{aligned} \tag{56}$$

Substituting (56) to (55), the probability of *event 4* is rewritten as

$$\begin{aligned} P_4 = &\frac{4(\lambda\pi)^n e^{-\lambda\pi y^2}}{(m-n-1)!\,(n-1)!} \\ &\times\sum_{k=0}^{t-m-1}(-1)^k\binom{t-m-1}{k}\int_{\varphi_1}^{\varphi_2}\mathrm{Z}(y)\,dy. \end{aligned} \tag{57}$$

The integral of (57) is burdensome to get a closed-form solution. Hence, Chebyshev-Gauss quadrature is implemented to obtain an approximated solution. The approximated probability of *event 4* is given by

$$\begin{aligned} &P_4 \\ &= \frac{4(\lambda\pi)^n e^{-\lambda\pi y^2}}{(m-n-1)!\,(n-1)!}\sum_{k=0}^{m-n-1}(-1)^k\binom{m-n-1}{k} \\ &\quad\times\sum_{l=1}^{N}\frac{\pi(\varphi_2-\varphi_1)}{2N}\sqrt{1-\omega_l^2}\,Z\left(\frac{\varphi_2-\varphi_1}{2}\omega_l+\frac{\varphi_2+\varphi_1}{2}\right). \end{aligned} \tag{58}$$

where $Z(y)$ is given as

$$Z(y) = \frac{e^{-\lambda\pi y^2}y^{2n-2m-2k-2}}{2n+2k}\left(y^{2n+2k}-\mu_1(y)^{2n+2k}\right). \tag{59}$$

N is the Chebyshev-Gauss approximation parameter, which achieves a tradeoff between accuracy and complexity of the approximation. Specifically, $N \to \infty$, the approximation solution is the same as the precise result. Besides, $\omega_l = \cos\left(\frac{2l-1}{2N}\right)$. Finally, the outage probability of n-th node for f_i ($\forall i \in [1,\cdots,M]$) is given as $P_n^i = P_m^0 + P_4$.

Therefore, *Lemma 3* is proved. ■

REFERENCES

[1] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electron.*, vol. 3, no. 1, pp. 20–29, Jan. 2020.

[2] N. Cheng *et al.*, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[3] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 4th Quart., 2018.

[4] R. Amer, W. Saad, and N. Marchetti, "Mobility in the sky: Performance and mobility analysis for cellular-connected UAVs," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3229–3246, May 2020.

[5] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A survey of channel modeling for UAV communications," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2804–2821, 4th Quart., 2018.

[6] M. T. Dabiri, M. Rezaee, V. Yazdanian, B. Maham, W. Saad, and C. S. Hong, "3D channel characterization and performance analysis of UAV-assisted millimeter wave links," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 110–125, Jan. 2021.

[7] Y. Hu, M. Chen, and W. Saad, "Joint access and backhaul resource management in satellite-drone networks: A competitive market approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3908–3923, Jun. 2020.

[8] K. An, M. Lin, J. Ouyang, and W.-P. Zhu, "Secure transmission in cognitive satellite terrestrial networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 3025–3037, Nov. 2016.

[9] Q. Huang, M. Lin, J. Wang, T. A. Tsiftsis, and J. Wang, "Energy efficient beamforming schemes for satellite-aerial-terrestrial networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3863–3875, Jun. 2020.

[10] N. Kato *et al.*, "Optimizing space-air-ground integrated networks by artificial intelligence," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 140–147, Aug. 2019.

[11] X. You *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 110–301, Jan. 2021.

[12] X. Cao, P. Yang, M. Alzenad, X. Xi, D. Wu, and H. Yanikomeroglu, "Airborne communication networks: A survey," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 1907–1926, Sep. 2018.

[13] Y. Wang, Y. Xu, Y. Zhang, and P. Zhang, "Hybrid satellite-aerial-terrestrial networks in emergency scenarios: A survey," *China Commun.*, vol. 14, no. 7, pp. 204–216, 2017.

[14] Y. Shi, J. Liu, Z. M. Fadlullah, and N. Kato, "Cross-layer data delivery in satellite-aerial-terrestrial communication," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 138–143, Jun. 2018.

[15] P. K. Sharma, D. Deepthi, and D. I. Kim, "Outage probability of 3-D mobile UAV relaying for hybrid satellite-terrestrial networks," *IEEE Commun. Lett.*, vol. 24, no. 2, pp. 418–422, Feb. 2020.

[16] A. J. Roumeliotis, C. I. Kourogiorgas, and A. D. Panagopoulos, "Optimal dynamic capacity allocation for high throughput satellite communications systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 596–599, Apr. 2019.

[17] G. Gür and S. Kafiloğlu, "Layered content delivery over satellite integrated cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 6, no. 3, pp. 390–393, Jun. 2017.

[18] T. De Cola and A. Blanco, "ICN-based protocol architectures for next-generation backhauling over satellite," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[19] C. Fang, H. Yao, Z. Wang, W. Wu, X. Jin, and F. R. Yu, "A survey of mobile information-centric networking: Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2353–2371, 3rd Quart., 2018.

[20] G. Zhong, J. Yan, and L. Kuang, "QoE-driven social aware caching placement for terrestrial-satellite networks," *China Commun.*, vol. 15, no. 10, pp. 60–72, Oct. 2018.

[21] J. Li, K. Xue, J. Liu, Y. Zhang, and Y. Fang, "An ICN/SDN-based network architecture and efficient content retrieval for future satellite-terrestrial integrated networks," *IEEE Netw.*, vol. 34, no. 1, pp. 188–195, Jan. 2020.

[22] G. Gur, "Spectrum sharing and content-centric operation for 5G hybrid satellite networks: Prospects and challenges for space-terrestrial system integration," *IEEE Veh. Technol. Mag.*, vol. 14, no. 4, pp. 38–48, Dec. 2019.

[23] T. de Cola *et al.*, "Network and protocol architectures for future satellite systems," *Found. Trends Netw.*, vol. 12, nos. 1–2, pp. 1–161, 2017.

[24] Z. Ding, P. Fan, G. K. Karagiannidis, R. Schober, and H. V. Poor, "NOMA assisted wireless caching: Strategies and performance analysis," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4854–4876, Oct. 2018.

[25] Z. Ding *et al.*, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

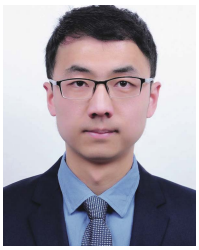
[26] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[27] V. K. Papanikolaou, G. K. Karagiannidis, N. A. Mitsiou, and P. D. Diamantoulakis, "Closed-form analysis for NOMA with randomly deployed users in generalized fading," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1253–1257, Aug. 2020.

[28] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "OTFS-NOMA: An efficient approach for exploiting heterogenous user mobility profiles," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7950–7965, Nov. 2019.
[29] X. Liang, J. Jiao, B. Feng, S. Wu, B. Cao, and Q. Zhang, "Performance analysis of millimeter-wave hybrid satellite-terrestrial relay networks over rain fading channel," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–5.
[30] X. Liang, J. Jiao, S. Wu, and Q. Zhang, "Outage analysis of multirelay multiuser hybrid satellite-terrestrial millimeter-wave networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 1046–1049, Dec. 2018.
[31] J. Du, C. Jiang, H. Zhang, X. Wang, Y. Ren, and M. Debbah, "Secure satellite-terrestrial transmission over incumbent terrestrial networks via cooperative beamforming," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1367–1382, Jul. 2018.
[32] B. Deng, C. Jiang, H. Yao, S. Guo, and S. Zhao, "The next generation heterogeneous satellite communication networks: Integration of resource management and deep reinforcement learning," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 105–111, Apr. 2020.
[33] M. K. Arti and V. Jain, "Relay selection-based hybrid satellite-terrestrial communication systems," *IET Commun.*, vol. 11, no. 17, pp. 2566–2574, Nov. 2017.
[34] P. K. Upadhyay and P. K. Sharma, "Max-max user-relay selection scheme in multiuser and multirelay hybrid satellite-terrestrial relay systems," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 268–271, Feb. 2016.
[35] S. Chandrasekharan *et al.*, "Designing and implementing future aerial communication networks," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 26–34, May 2016.
[36] S. Zhang, W. Quan, J. Li, W. Shi, P. Yang, and X. Shen, "Air-ground integrated vehicular network slicing with content pushing and caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2114–2127, Sep. 2018.
[37] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, Nov. 2016.
[38] A. Fotouhi *et al.*, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 4th Quart., 2019.
[39] G. Pan, J. Ye, Y. Zhang, and M.-S. Alouini, "Performance analysis and optimization of cooperative satellite-aerial-terrestrial systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6693–6707, Oct. 2020.
[40] T. X. Vu, N. Maturo, S. Vuppala, S. Chatzinotas, J. Grotz, and N. Alagha, "Efficient 5G edge caching over satellite," in *Proc. 36th Int. Satell. Commun. Syst. Conf. (ICSSC)*, 2018, pp. 1–5.
[41] C. G. Brinton, E. Aryafar, S. Corda, S. Russo, R. Reinoso, and M. Chiang, "An intelligent satellite multicast and caching overlay for CDNs to improve performance in video applications," in *Proc. 31st AIAA Int. Commun. Satell. Syst. Conf.*, Oct. 2013, pp. 1–9.
[42] K. An, Y. Li, X. Yan, and T. Liang, "On the performance of cache-enabled hybrid satellite-terrestrial relay networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1506–1509, Oct. 2019.
[43] S. A. Tegos, P. D. Diamantoulakis, J. Xia, L. Fan, and G. K. Karagiannidis, "Outage performance of uplink NOMA in land mobile satellite communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1710–1714, Oct. 2020.
[44] X. Zhang, K. An, B. Zhang, Z. Chen, Y. Yan, and D. Guo, "Vickrey auction-based secondary relay selection in cognitive hybrid satellite-terrestrial overlay networks with non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 628–632, May 2020.
[45] X. Zhang *et al.*, "Outage performance of NOMA-based cognitive hybrid satellite-terrestrial overlay networks by amplify-and-forward protocols," *IEEE Access*, vol. 7, pp. 85372–85381, 2019.
[46] X. Yan, H. Xiao, K. An, and C.-X. Wang, "Outage performance of NOMA-based hybrid satellite-terrestrial relay networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 538–541, Aug. 2018.
[47] X. Yan, H. Xiao, K. An, G. Zheng, and W. Tao, "Hybrid satellite terrestrial relay networks with cooperative non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 978–981, May 2018.
[48] X. Zhu, C. Jiang, L. Kuang, N. Ge, and J. Lu, "Non-orthogonal multiple access based integrated terrestrial-satellite networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2253–2267, Oct. 2017.
[49] A. Kalantari, M. Fittipaldi, S. Chatzinotas, T. X. Vu, and B. Ottersten, "Cache-assisted hybrid satellite-terrestrial backhauling for 5G cellular networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.
[50] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3584–3586, Oct. 2005.
[51] N. Okati, T. Riihonen, D. Korpi, I. Angervuori, and R. Wichman, "Downlink coverage and rate analysis of low Earth orbit satellite constellations using stochastic geometry," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5120–5134, Aug. 2020.
[52] A. Talgat, M. A. Kishk, and M.-S. Alouini, "Stochastic geometry-based analysis of LEO satellite communication systems," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2458–2462, Aug. 2021.
[53] J. Tang, G. Chen, J. P. Coon, and D. E. Simmons, "Distance distributions for Matérn cluster processes with application to network performance analysis," in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.
[54] X. Zhang *et al.*, "On the performance of hybrid satellite-terrestrial content delivery networks with non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 454–458, Mar. 2021.
[55] X. Zhang, B. Zhang, K. An, Z. Chen, and D. Guo, "Auction-based secondary relay selection on overlay spectrum sharing in hybrid satellite–terrestrial sensor networks," *Sensors*, vol. 19, no. 22, p. 5039, Nov. 2019.
[56] X. Yue *et al.*, "Outage behaviors of NOMA-based satellite network over Shadowed-Rician fading channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6818–6821, Jun. 2020.
[57] X. Zhang *et al.*, "Performance analysis of NOMA-based cooperative spectrum sharing in hybrid satellite-terrestrial networks," *IEEE Access*, vol. 7, pp. 172321–172329, 2019.
[58] W. Wang, Y. Tong, L. Li, A.-A. Lu, L. You, and X. Gao, "Near optimal timing and frequency offset estimation for 5G integrated leo satellite communication system," *IEEE Access*, vol. 7, pp. 113298–113310, 2019.
[59] T. Z. H. Ernest, A. S Madhukumar, R. P. Sirigina, and A. K. Krishna, "NOMA-aided UAV communications over correlated rician shadowed fading channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 3103–3116, 2020.
[60] X. Wang, H. Zhang, K. J. Kim, Y. Tian, and A. Nallanathan, "Performance analysis of cooperative aerial base station-assisted networks with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5983–5999, Dec. 2019.
[61] K. An, M. Lin, W.-P. Zhu, Y. Huang, and G. Zheng, "Outage performance of cognitive hybrid satellite-terrestrial networks with interference constraint," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 9397–9404, Nov. 2016.
[62] S. M. Azimi, O. Simeone, A. Sengupta, and R. Tandon, "Online edge caching and wireless delivery in fog-aided networks with dynamic content popularity," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1189–1202, Jun. 2018.
[63] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "Simple semi-grant-free transmission strategies assisted by non-orthogonal multiple access," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4464–4478, Jun. 2019.

Xiaokai Zhang received the B.S. degree from Harbin Institute of Technology (HIT), Harbin, China, in 2015, and the M.S. and Ph.D. degrees from the Army Engineering University of PLA, Nanjing, China, in 2017 and 2021, respectively. He is currently a Lecturer with the Army Engineering University of PLA. His research interests focus on satellite communications, hybrid satellite-terrestrial networks, aerial networks, polarization shift keying, physical layer security, non-orthogonal multiple access (NOMA), cognitive radio, auction theory, game theory, reinforcement learning, and reconfigurable intelligent surface.

Bangning Zhang received the B.S. and M.S. degrees from the Institute of Communications Engineering (ICE), Nanjing, China, in 1984 and 1987, respectively. He is currently a Full Professor and the Head of the College of Communications Engineering. He has authored and coauthored more than 80 conference papers and journal articles and has been granted over 20 patents in his research areas. His current research interests include communication anti-jamming technologies, microwave technologies, polarization technologies, satellite communications systems, cooperative communications, and physical layer security. He has served as a reviewer for several journals in communication field.



Kang An received the B.E. degree in electronic engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2011, the M.E. degree in communication engineering from the PLA University of Science and Technology, Nanjing, in 2014, and the Ph.D. degree in communication engineering from the Army Engineering University of PLA, Nanjing, in 2017. Since January 2018, he has been with the National University of Defense Technology, Nanjing, where he is currently a Senior Engineer. His current research interests include satellite communication, B5G/6G wireless communication, reconfigurable intelligent surface, and cognitive radio.

Gan Zheng (Fellow, IEEE) received the B.Eng. and M.Eng. degrees in electronic and information engineering from Tianjin University, Tianjin, China, in 2002 and 2004, respectively, and the Ph.D. degree in electrical and electronic engineering from The University of Hong Kong in 2008.

He is currently a Professor of signal processing for wireless communications with Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, U.K. His research interests include machine learning for communications, UAV communications, mobile edge caching, full-duplex radio, and wireless power transfer. He was a first recipient of the 2013 IEEE SIGNAL PROCESSING LETTERS Best Paper Award, and received the 2015 GLOBECOM Best Paper Award and the 2018 IEEE Technical Committee on Green Communications and Computing Best Paper Award. He was listed as a Highly Cited Researcher by Thomson Reuters/Clarivate Analytics in 2019. He serves as an Associate Editor for IEEE COMMUNICATIONS LETTERS and IEEE WIRELESS COMMUNICATIONS LETTERS.

Symeon Chatzinotas (Senior Member, IEEE) had been a Visiting Professor with the University of Parma, Italy, in the past, where he was lecturing on 5G wireless networks. He is currently a Full Professor/the Chief Scientist I and the Head of the SIGCOM Research Group at SnT, University of Luxembourg. He was involved in numerous research and development projects at NCSR Demokritos, CERTH Hellas, and CCSR, University of Surrey. He is coordinating the research activities on communications and networking, acting as the PI for more than 20 projects and a main representative of 3GPP, ETSI, and DVB. He has (co)authored more than 450 technical papers in refereed international journals, conferences, and scientific books.

Prof. Chatzinotas was a co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award and Best Paper Awards from *EURASIP JWCN*, CROWNCOM, and ICSSC. He is in the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY, and the *International Journal of Satellite Communications and Networking*.

Daoxing Guo received the B.S., M.S., and Ph.D. degrees from the College of Communications Engineering, Nanjing, China, in 1995, 1999, and 2002, respectively. He is currently a Full Professor and a Ph.D. Supervisor with the Army Engineering University of PLA. He has authored or coauthored more than 40 conference papers and journal articles. His current research interests include satellite communications systems and transmission technologies, communication anti-jamming technologies, and communication anti-interception technologies, including physical layer security. He holds over 20 patents in his research areas. He has served as a reviewer for several journals in communication field.