Optimizing Cache Content Placement in Integrated Terrestrial and Non-terrestrial Networks

Feng Wang[⋆], Giovanni Geraci[♯], and Tony Q. S. Quek[⋆]

[⋆]*Singapore University of Technology and Design (SUTD), Singapore*

[♯]*Universitat Pompeu Fabra (UPF), Barcelona, Spain*

***Abstract*—Non-terrestrial networks (NTN) offer potential for efficient content broadcast in remote regions, thereby extending the reach of digital services. In this paper, we introduce a novel approach to optimize wireless edge content placement using NTN. Specifically, we dynamically select content for placement via NTN links based on popularity and suitability for delivery through NTN, while considering the orbital motion of LEO satellites. Our comprehensive system-level case studies, based on a practical LEO constellation, demonstrate the significant improvement in placement speed compared to existing methods that neglect network mobility. We further show that the advantages of NTN links over standalone wireless TN solutions are more pronounced in the early stages of content delivery and are amplified by higher content popularity correlation across geographical regions.**

I. Introduction

In our increasingly interconnected world, seamless access to information has become a necessity and the demand for new services continues to surge [1]–[4]. While optical fiber links have long been the backbone of high-speed data transmission, their deployment in remote and underserved areas can be logistically challenging and financially impractical. Enter non-terrestrial networks (NTN), a transformative solution that promises to bridge the digital divide and enable disruptive use cases by leveraging airborne or space-based infrastructure [5]–[9]. With the ability to provide connectivity to even the most remote corners of our planet, NTN represents a paradigm shift in delivering essential services to areas that were previously beyond the reach of traditional terrestrial networks [10]–[12].

Within the realm of NTN, an emerging application with remarkable potential is the utilization of low Earth orbit (LEO) satellite constellations for efficient cache content broadcast to remote areas [13]–[16]. Caching, the process of storing frequently accessed data closer to end-users, has long been vouched for in terrestrial networks (TN) to enhance content delivery speeds and reduce bandwidth requirements [17]–[19]. However, in regions where optical fiber links are absent or inadequate, LEO satellites offer an alternative solution to establish a global network fabric [20]. The use of NTN for caching not only addresses the connectivity gaps but also has the potential to revolutionize content delivery by extending the reach of digital services to previously untapped populations.

This work was supported in part by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications R&D Programme, by the Spanish Research Agency through grants PID2021-123999OB-I00 and CEX2021-001195-M, by the "Ramón y Cajal" program, and by the UPF-Fractus Chair.

Recent work has explored the use of NTN to deliver cache content, while also proposing content placement strategies [21]–[23] and radio resource optimization [24]–[26] aimed at maximizing the delivery throughput and reducing the placement time. Although these contributions have helped in understanding the potential role of NTN in delivering wireless edge caching content, they did not focus on the impact of LEO satellite mobility, producing a highly dynamic coverage pattern on the ground [27]. Indeed, optimal wireless content placement must jointly consider satellite mobility as well as content popularity distribution across the satellites' footprint.

In this paper, we introduce a novel approach to optimize wireless edge cache content placement using NTN. In contrast to existing approaches, our method focuses on dynamically selecting content for placement via NTN links, taking into consideration both the popularity distribution of the content and its suitability for delivery through NTN links. This suitability factor accounts for the orbital motion of NTN base stations (BSs), ensuring efficient and effective content delivery. The main contributions of our work can be summarized as follows:

- We formulate the cache placement optimization problem in an integrated TN-NTN to minimize the content delivery time. To address this problem, we propose a heuristic approach that determines the content to be placed via NTN links by leveraging the content's popularity distribution and the varying NTN coverage pattern.
- We carry out system-level case studies based on a practical LEO constellation to evaluate the performance of our proposed approach in various scenarios. The results demonstrate that our approach significantly improves the placement speed compared to state-of-the-art methods that do not consider network mobility.
- We demonstrate that the advantages offered by NTN links over standalone wireless TN solutions are particularly prominent in the initial stages of content delivery, when the most popular content is placed. We also confirm that the higher the correlation of content popularity across geographical regions, the more significant the advantage of NTN-based broadcast delivery becomes.

II. System Model

We now introduce: (a) the network topology considered, (b) the wireless propagation channel model for TN and NTN links, and (c) the content popularity and placement model.

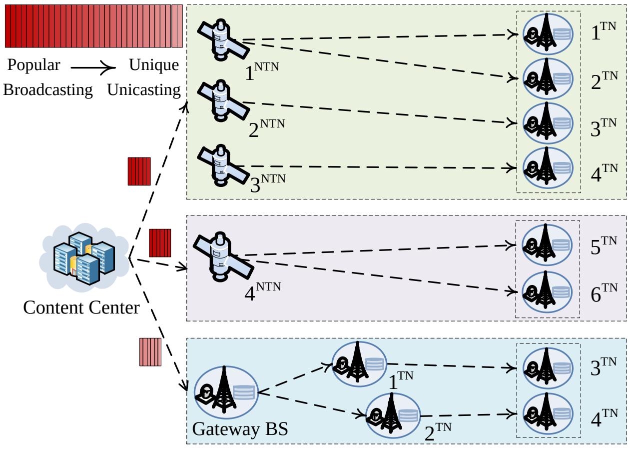


Fig. 1. Illustration of cache content placement in an integrated TN-NTN. More/less popular content is respectively broadcast/unicast via NTN/TN links.

A. Network Topology

The network under consideration is illustrated in Fig. 1 and comprises the following elements, from right to left:

- A terrestrial network (TN) with edge base stations (BSs) $b^{\mathrm{TN}} \in \mathcal{B}^{\mathrm{TN}}$, where all BSs within a certain region r are connected to each other via microwave links forming a tree topology. Each region comprises a gateway BS, connected to a content center via optical fiber links and serving as the root node for content transmission.
- A non-terrestrial network (NTN) consisting of a constellation of low Earth orbit (LEO) satellite BSs $b^{\mathrm{NTN}} \in \mathcal{B}^{\mathrm{NTN}}$. Each satellite may receive files from the content center and broadcast them to some of the edge TN BSs.
- A content center holding files that can be fetched by TN edge BSs for caching. Its role is to determine whether each file should be delivered to TN edge BSs through the TN or NTN segments. The content center oversees a geographical area partitioned into regions r.

In the sequel, we therefore consider two types of links:

- *TN links*, connecting two TN BSs i^{TN} and j^{TN}.
- *NTN links*, between a NTN BS i^{NTN} and a TN BS j^{TN}.

For TN links, we assume each BS to be able to connect to neighboring BSs within a certain coverage distance, so that a TN BS can receive files from the regional gateway via multiple hops. NTN links are affected by the LEO satellite elevation angle, whereby angles closer to 90° yield shorter LEO-to-BS distances and are more likely to be in line-of-sight (LoS), whereas angles below 10° are unable to support data transmission [28], [29]. For convenience, we define an indicator function $\mathbb{1}^{\mathrm{est}}_{i,j}$, whose value is one when an NTN link between BSs i^{NTN} and j^{TN} is established, and zero otherwise. We assume a TN BS to establish a link with at most one NTN BS at a time, namely the nearest at any given time.

B. Propagation Channel and Data Rates

Channel model: All radio links are affected by path loss and lognormal shadow fading, both dependent on the link LoS condition and modeled as in [28]–[30]. Compared to a TN link, the signal travelling on an NTN link undergoes several extra stages of propagation. As a result, the total NTN path loss consists of additional terms accounting for the attenuation due to atmospheric gases and scintillation [31]. We assume the LEO satellite antenna generating seven beams to be a typical reflector with circular aperture, and each TN BS to be equipped with a very small aperture terminal (VSAT). We denote $G_{i,j}$ the total large-scale gain between any two BSs i and j, comprising path loss, shadow fading, and antenna gain at both ends. Similarly, we denote $h_{i,j}$ the small-scale block fading between the two BSs.

Achievable data rates: The signal-to-noise ratio (SNR) on the link connecting BSs i and j is given by

$$\mathtt{SNR}_{i,j} = \frac{P_i \cdot G_{i,j} \cdot |h_{i,j}|^2}{\sigma_j^2}, \tag{1}$$

where P_i is the transmission power of BS i and σ_j^2 is the thermal noise variance at BS j. Assuming that all edge BSs are distributed sparsely enough for the link to be interference-free, the corresponding achievable rate $\mathcal{R}_{i,j}$ can be obtained as

$$\mathcal{R}_{i,j} = B_{i,j} \cdot \mathbb{E}\left[\log_2(1 + \mathtt{SNR}_{i,j})\right], \tag{2}$$

with $B_{i,j}$ denoting the bandwidth allocated to the link and where the expectation is taken over the small-scale fading.

For NTN links, $G_{i,j}$ and $h_{i,j}$ (and in turn, $\mathtt{SNR}_{i,j}$ and $\mathcal{R}_{i,j}$) depend on the distance and elevation angle between BSs i^{NTN} and j^{TN}. Their values vary with time and can be predicted from the satellite orbital information. At any given time, we define the NTN link duration $T_{i,j}$ between BSs i^{NTN} and j^{TN} as the remaining time until the condition $\mathbb{1}^{\mathrm{est}}_{i,j}$ changes its value from 1 to 0. We then partition the time axis into slots $t = 1, \ldots, N_{\mathrm{t}}$, such that all established links (i.e., those with $\mathbb{1}^{\mathrm{est}}_{i,j} = 1$) remain as such within a slot, and at least one changes its value to zero across slots. At any given time, the remaining duration T_t of the current slot t can therefore be computed as

$$T_t = \min_{i,j \,|\, \mathbb{1}^{\mathrm{est}}_{i,j}=1} \{T_{i,j}\}. \tag{3}$$

C. Content Popularity and Placement

Content popularity within a region: The content file popularity is assumed to obey the Zipf distribution [32]. We express the popularity $p_{f,r}$ of file f in region r as

$$p_{f,r} = \left[f^{\alpha_r} \cdot \sum_{g=1}^{N_{\mathrm{f}}} g^{-\alpha_r} \right]^{-1}, \tag{4}$$

where $f = 1, \ldots, N_{\mathrm{f}}$, $\alpha_r \in \mathbb{R}^+$ is a skewness factor that controls the content popularity within r, and $\sum_{f=1}^{N_{\mathrm{f}}} p_{f,r} = 1 \; \forall r$. Let $\mathcal{B}^{\mathrm{TN}}_r \subseteq \mathcal{B}^{\mathrm{TN}}$ be the set of TN BSs in r and let $\|\mathcal{B}^{\mathrm{TN}}_r\|$ denote its cardinality. The expected number of TN BSs requesting file f in region r is then given by $\|\mathcal{B}^{\mathrm{TN}}_r\| \cdot p_{f,r}$.

Content correlation across regions: We assume the file popularity to vary across regions, with a coefficient ρ modeling its correlation. Without loss of generality, we map each region r to geographical coordinates $(\mathrm{x}_r, \mathrm{y}_r) \in \{\mathbb{N}^+\}^2$ and obtain the file popularity correlation between regions r and r' as

$\rho^{|x_r - x_{r'}| + |y_r - y_{r'}|}$. The values of $\{\alpha_r\}$ and ρ thus model the content popularity across the whole geographical area.

We define a binary indicator $\mathbb{1}^{\text{ass}}_{f,b}$ whose value is one if file f is assigned to BS b^{TN} for caching and zero otherwise. Since each TN BS has a finite cache size and the file popularity is probabilistic, we assume that the content center plans for only $\lfloor \|\mathcal{B}^{\text{TN}}_r\| \cdot p_{f,r} \rfloor$ BSs in region r to cache file f in decreasing order of popularity $p_{f,r}$ until the cache of all BSs in r reaches saturation. Such assignments are recorded through $\{\mathbb{1}^{\text{ass}}_{f,b}\}$.

We assume that content placement takes place sequentially, with each file being delivered to the edge BSs through either broadcast (via NTN links) or unicast (via multi-hop TN links). In each time slot t, the content center determines which files to be broadcast over NTN links. We denote each assignment with a binary indicator $\mathbb{1}^{\text{pla}}_{f,t}$, whose value is one if file f is allocated for NTN delivery in time slot t and zero otherwise.

Content placement via TN links: When using TN links, a file f is delivered to a BS b^{TN} in a multi-hop fashion, with d^{TN}_b denoting the distance between the gateway and b^{TN} expressed in number of transmission hops required. The rate $\mathcal{R}^{\text{TN}}_b$ between the gateway and b^{TN} is given by the lowest rate on any of the hops, each of the latter governed by (2). The placement time τ^{TN}_f of file f for the served area is then the maximum across all BSs requiring such file, i.e.,

$$\tau^{\text{TN}}_f = \max_{b^{\text{TN}} \mid \mathbb{1}^{\text{ass}}_{f,b}=1} \left\{ s_f / \mathcal{R}^{\text{TN}}_b \right\}, \tag{5}$$

where s_f denotes the size of file f.

Content placement via NTN links: As the NTN topology remains unchanged in each time slot, we approximate the data rate of an NTN link with its average value. By neglecting the delay incurred on the high-capacity optical link that connects the content center to the NTN BS, we then compute $\tau^{\text{NTN}}_{f,t}$ for a file f in slot t as the transmission time over the NTN link,

$$\tau^{\text{NTN}}_{f,t} = \max_{b^{\text{TN}} \mid \mathbb{1}^{\text{ass}}_{f,b}=1} \left\{ s_f / \mathcal{R}^{\text{NTN}}_b \right\}, \tag{6}$$

where $\mathcal{R}^{\text{NTN}}_b$ denotes the rate on the link between BSs i^{NTN} (delivering the file) and b^{TN} (receiving it). While not shown explicitly for ease of notation, the rate $\mathcal{R}^{\text{NTN}}_b$ varies across time slots and it can be computed via (2).

III. NTN Content Placement Optimization

A. Problem Formulation

Through the following problem, we aim at maximizing the total savings achieved in terms of content delivery time by opportunistically delivering files via either TN or NTN links:

Problem 1. *Content placement optimization*

$$\max_{\mathbb{1}^{\text{pla}}_{f,t}} \sum_{t=1}^{N_\text{t}} \sum_{f=1}^{N_\text{f}} \mathbb{1}^{\text{pla}}_{f,t} \cdot \left(\tau^{\text{TN}}_f - \tau^{\text{NTN}}_{f,t} \right) \tag{7}$$

$$\text{s.t.} \quad \sum_{f=1}^{N_\text{f}} \mathbb{1}^{\text{pla}}_{f,t} \cdot \tau^{\text{NTN}}_{f,t} \leq T_t, \quad t = 1, \ldots, N_\text{t} \tag{7a}$$

$$\sum_{t=1}^{N_\text{t}} \mathbb{1}^{\text{pla}}_{f,t} \leq 1, \quad f = 1, \ldots, N_\text{f}. \tag{7b}$$

Note that the constraint (7a) ensures that the delivery time via NTN links within a slot does not exceed the duration of the slot itself, whereas (7b) ensures that each file is not delivered more than once. Given the delivery times τ^{TN}_f and $\tau^{\text{NTN}}_{f,t}$ of each file f via TN and NTN links, (7) treats the content placement optimization problem as one of determining, for each time slot, the values $\{\mathbb{1}^{\text{pla}}_{f,t}\}$, i.e., which files to be placed via NTN.

Remark 1. *For given t, τ^{TN}_f, and $\tau^{\text{NTN}}_{f,t}$, (7) has a finite solution space. Moreover, (7) is polynomial-time verifiable since the size of every feasible set $\{\mathbb{1}^{\text{pla}}_{f,t}\}$ is polynomially bounded by the number of cache files to be placed, the latter being finite. As a result, (7) is an NP-optimization problem for which exhaustive search is not a practically viable approach.*

B. Optimal Wireless Edge Content Placement

We now propose a heuristic approach to solve (7) by opportunistically leveraging NTN links. The proposed approach is based on two metrics, introduced as follows to capture the advantage of NTN-based broadcast delivery.

NTN participation indicator μ^{part}: We recall the indicators $\mathbb{1}^{\text{est}}_{i,j}$ and $\mathbb{1}^{\text{ass}}_{f,j}$, whose value is one respectively if a link between BSs i^{NTN} and j^{TN} is established and if BS j^{TN} is assigned file f. We can then employ $\mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j}$ to indicate whether BS j^{TN} should receive file f and has a link with BS i^{NTN}. The latter allows us to define an *NTN participation indicator* $\mu^{\text{part}}_{f,t} \in [0, 1]$ as the fraction of NTN BSs suitable to deliver file f out of all NTN BSs providing service in slot t, given by

$$\mu^{\text{part}}_{f,t} = \frac{\sum_{i \in \mathcal{B}^{\text{NTN}}} \left(1 - \left(\prod_{j \in \mathcal{B}^{\text{TN}}} \left(1 - \mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j} \right) \right) \right)}{\sum_{i \in \mathcal{B}^{\text{NTN}}} \left(1 - \left(\prod_{j \in \mathcal{B}^{\text{TN}}} \left(1 - \mathbb{1}^{\text{est}}_{i,j} \right) \right) \right)}. \tag{8}$$

The value of $\mu^{\text{part}}_{f,t}$ can be calculated in each time slot t to infer the degree of participation of NTN BSs in the delivery of file f. Intuitively, higher values of $\mu^{\text{part}}_{f,t}$ indicate a higher suitability for file f to be deployed via NTN broadcast.

NTN superiority indicator μ^{sup}: We recall that d^{TN}_b denotes the number of hops required to deliver content to TN BS b^{TN} via TN links. Then the average number of TN hops $d^{\text{ave}}_{f,t}$ that can be avoided by instead delivering file f via NTN links in slot t can be calculated as

$$d^{\text{ave}}_{f,t} = \frac{\sum\limits_{i \in \mathcal{B}^{\text{NTN}}} \frac{\sum_{j \in \mathcal{B}^{\text{TN}}} \left(\mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j} \cdot d^{\text{TN}}_j \right)}{\sum_{j \in \mathcal{B}^{\text{TN}}} \mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j}} \left(1 - \left(\prod\limits_{j \in \mathcal{B}^{\text{TN}}} \left(1 - \mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j} \right) \right) \right)}{\sum\limits_{i \in \mathcal{B}^{\text{NTN}}} \left(1 - \left(\prod\limits_{j \in \mathcal{B}^{\text{TN}}} \left(1 - \mathbb{1}^{\text{ass}}_{f,j} \cdot \mathbb{1}^{\text{est}}_{i,j} \right) \right) \right)} \tag{9}$$

and it can be normalized to an *NTN superiority indicator* $\mu^{\text{sup}}_{f,t} \in [0, 1]$ through the following rescaling

$$\mu^{\text{sup}}_{f,t} = \frac{d^{\text{ave}}_{f,t} - \min_f \{ d^{\text{ave}}_{f,t} \}}{\max_f \{ d^{\text{ave}}_{f,t} \} - \min_f \{ d^{\text{ave}}_{f,t} \}}. \tag{10}$$

Intuitively, higher values of $\mu^{\text{sup}}_{f,t}$ indicate a higher suitability for a file f to be delivered via NTN links, since delivering the

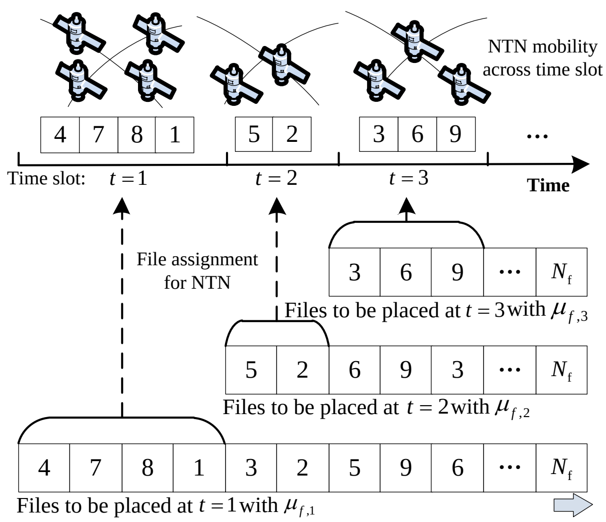


Fig. 2. Illustration of the proposed NTN file assignment approach.

Algorithm 1 Proposed NTN file assignment (NFA).

Input:

BSs $\mathcal{B}^{\mathrm{NTN}}$ and $\mathcal{B}^{\mathrm{TN}}$, files $\mathcal{F}_1$, popularity $\{\mathbb{1}^{\mathrm{ass}}_{f,b}\}$, $t=1$;

Output:

NTN file assignments $\{\mathbb{1}^{\mathrm{pla}}_{f,t}\}$;

```
while F_t ≠ ∅ do
    Compute NTN link conditions 1^est_{i,j};
    Compute τ^TN_f and τ^NTN_{f,t} from (5) and (6);
    for f ∈ F_t do
        Compute μ^part_{f,t} from (8) and μ^sup_{f,t} from (10);
        Compute μ_{f,t} from (11);
    end for
    Sort all files in F_t in decreasing order of μ_{f,t};
    while constraint (7a) is met do
        Set 1^pla_{f,t} ← 1 for ordered files f ∈ F_t;
    end while
    Update F_{t+1} ← F_t − {f | 1_{f,t} = 1};
    t ← t + 1;
end while
```

same file via TN would entail a large number of transmission hops, hence a longer delivery time.

We finally combine indicators $\mu^{\mathrm{part}}_{f,t}$ and $\mu^{\mathrm{sup}}_{f,t}$ into a single metric $\mu_{f,t} \in [0,1]$, capturing the overall suitability of file f to be delivered via NTN links in time slot t and defined as follows

$$\mu_{f,t} = \frac{\left[\mu^{\mathrm{part}}_{f,t}\right]^{\beta} \cdot \left[\mu^{\mathrm{sup}}_{f,t}\right]^{1-\beta}}{\sum\limits_{k=1}^{N_\mathrm{f}} \left\{ \left[\mu^{\mathrm{part}}_{k,t}\right]^{\beta} \cdot \left[\mu^{\mathrm{sup}}_{k,t}\right]^{1-\beta} \right\}}, \tag{11}$$

where the coefficient β is introduced to trade off the participation and superiority criteria in (8) and (10).

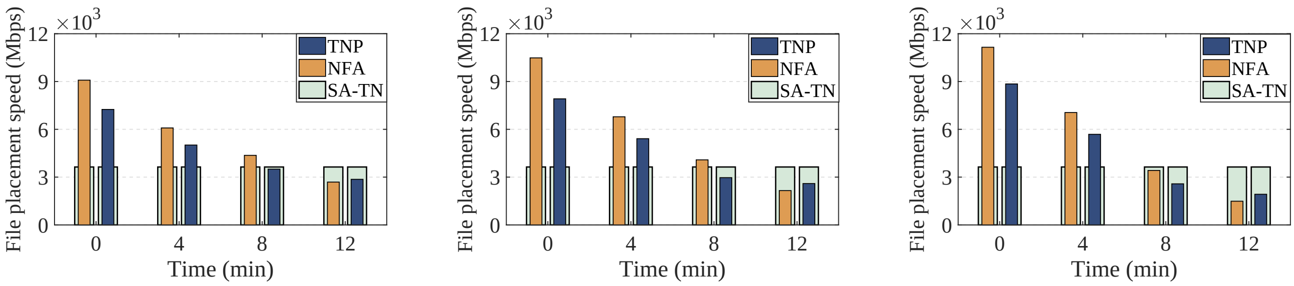
Content placement algorithm: Let $\mathcal{F}_t$ denote the set of files to be placed in a given time slot t. The content center then ranks all files $f \in \mathcal{F}_t$ according to the indicator $\mu_{f,t}$ in (11) and deploys as many files as possible within t via the NTN in decreasing order of $\mu_{f,t}$. Formally, such decision corresponds to setting $\mathbb{1}^{\mathrm{pla}}_{f,t}$ to one when file f is allocated for NTN placement in time slot t. After slot t, the set of files yet to be delivered is calculated as $\mathcal{F}_{t+1} = \mathcal{F}_t - \{f \,|\, \mathbb{1}^{\mathrm{pla}}_{f,t} = 1\}$ and the new values $\mu_{f,t+1}$ are calculated for all files in $\mathcal{F}_{t+1}$. The procedure is repeated until the last file has been deployed. The proposed approach is denoted *NTN file assignment (NFA)* and it is illustrated in Fig. 2 and detailed in Algorithm 1.

IV. Numerical Results

We now evaluate the performance of the proposed file assigment algorithm. For the NTN segment, we employ the System Tool Kit (STK) simulator and consider a LEO satellite constellation as the one deployed by Starlink, consisting of 1584 satellites distributed in 24 orbits with an inclination of 53°, each creating seven beams with 20 km beam diameter [33]. We select the geographical area with coordinates ([33-39°N], [87-93°E]), located in western China, as a typical TN scenario where it is challenging to deploy high-density ground network infrastructure. We establish six rectangular regions evenly distributed across this area, with each region containing 100 randomly placed TN BSs and a gateway BS connecting all regional BSs using the minimum spanning tree algorithm. Each TN BS is configured to cache 50 files, for a total of 30000 files, each with a size of 20 MB. For the content popularity distribution, we set the skewness parameter to $\alpha_r = 1$ for all regions r and the regional file correlation coefficient to $\rho = 0.8$, unless otherwise stated.

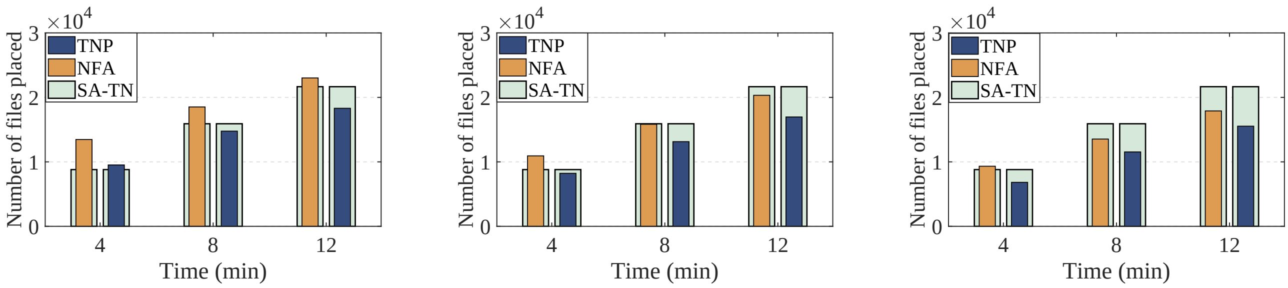
We assume a 100 MHz bandwidth for all links. For NTN links, we set the transmit power to 30 dBW, and the antenna gain at the transmitter (resp. receiver) to 38.5 (resp. 39.7) dBi. We require a minimum elevation angle of 10° to establish an NTN link, resulting in a path loss ranging between 173.4 dB and 184.9 dB [28], [29] and in an NTN link rate ranging from 130 Mbps to 300 Mbps. For TN links, the transmit power is set to 44 dBm and the antenna gain at both transmitter and receiver to 16 dBi [30]. The delivery radius of each TN BS is set to 2 km and the minimum inter-BS distance to 0.5 km [21], yielding achievable data rates ranging between 407 Mbps and 1 Gbps. To assess the performance of our proposed algorithm, we compare the file placement speed of the following four approaches under different file popularity distributions:

- *The proposed NFA approach*, employing the procedure described in Section III and Algorithm I, where we set $\beta = 0.5$ to place equal importance on the NTN participation and superiority indicators μ^{part} and μ^{sup}.
- *A TN popularity (TNP) approach [21]*, using NTN links to deliver files, where the latter are ranked in order of conventional TN file popularity.
- *A maximum backhaul traffic (MBT) approach [24]*, using NTN links to deliver files, prioritizing those required by at least one BS in each NTN BS coverage footprint.
- *A standalone TN (SA-TN) approach*, using only TN links to deploy files, ranked in order of conventional TN file popularity. We regard this as a TN-only baseline.



(a) $\rho = 0.8, \alpha_r = 0.5$ (b) $\rho = 0.8, \alpha_r = 1$ (c) $\rho = 0.8, \alpha_r = 1.5$

Fig. 3. File placement speed over time for the proposed NFA, TNP, and SA-TN, for different values of the file popularity skewness, α_r.



(a) $\alpha_r = 1, \rho = 0.7$ (b) $\alpha_r = 1, \rho = 0.6$ (c) $\alpha_r = 1, \rho = 0.5$

Fig. 4. Files placed over time for the proposed NFA, TNP, and SA-TN, for different values of the content popularity correlation across regions, ρ.

In Fig. 3, we compare the proposed NFA to the TNP and SA-TN approaches in terms of the cumulative data rate provided to all TN BSs versus time. For the time being, we assume NFA and TNP to avail of NTN links only. The data rate provided by SA-TN remains constant over time due to the fixed TN topology. The figure shows such data rates for different values of the file popularity skewness α_r ranging from 0.5 to 1.5. Fig. 3 demonstrates that the proposed NFA scheme exhibit higher file placement speeds than the TNP baseline for all three values of α_r. As the value of α_r increases, so does the proportion of highly popular files and the advantage provided by NTN broadcasting becomes more prominent in the early stages of content delivery. Such advantage vanishes in the later stages, once the most popular content has been delivered, making the TN a better option for content delivery.

In Fig. 4, we show the variation in the number of files placed over time for the proposed NFA and the TNP approach when employing only NTN links, and we compare it to a SA-TN approach. We also evaluate the impact of the content correlation coefficient ρ, with lower values of ρ modeling greater differences in content popularity across regions. As expected, as the value of ρ decreases from 0.7 to 0.5, the relative file placement volume of both NTN approaches compared to SA-TN also does. This is due to a greater heterogeneity in file popularity across regions, which reduces the advantage of NTN broadcasting. For all values of ρ, the proposed NFA scheme consistently outperforms the TNP baseline, also demonstrating its ability to dynamically adapt to the differences in file popularity among regions.

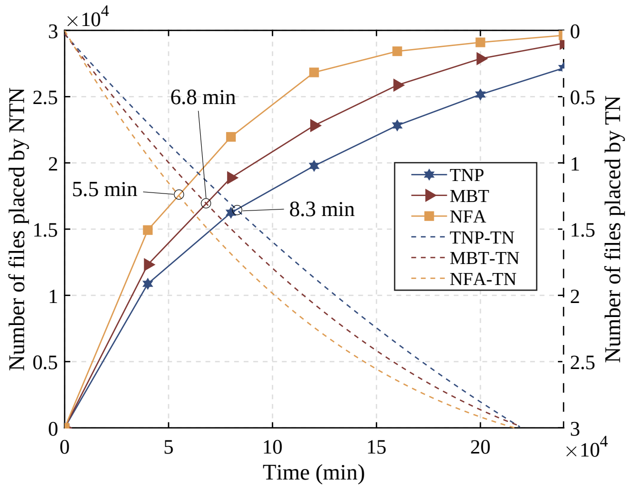


Fig. 5. Number of files placed vs. time and total placement time via NTN and TN for the proposed NFA and for state-of-the art MBT and TNP approaches.

Fig. 5 compares the proposed NFA to the TNP and MBT approaches in terms of file placement time when both NTN and TN links can be used for delivery. In this figure, it is assumed that files are deployed via NTN links in decreasing order of popularity—following the respective criteria for the three approaches—and via TN links in increasing order of popularity. The figure shows the number of files placed over time by each scheme over NTN links (solid) and TN links (dashed). The intersection of solid and dashed curves thus represents the time taken to deploy all 30000 files. For the scenario under consideration, Fig. 5 shows that the proposed NFA strategy completes the file placement in 5.5 minutes, with a reduction of about 20% and 33% when compared to state-of-the-art MBT and TNP approaches.

V. Conclusions

In this paper, we proposed a novel approach to optimize wireless edge content placement using NTN. Our approach dynamically selects content for placement via NTN links based on their popularity and also suitability for NTN delivery, accounting for the orbital motion of NTN BSs. We carried out system-level case studies based on a practical LEO constellation to evaluate the performance of our approach.

Our results demonstrated a significant improvement in placement speed compared to state-of-the-art methods that overlook the dynamic coverage created by LEO satellite mobility. We further observed that the advantages of NTN links over standalone wireless TN solutions are most prominent in the early stages of content placement, diminishing as popular content is delivered. Additionally, the advantage of NTN-based broadcast delivery becomes more significant with higher content popularity correlation across geographical regions.

This paper aims at advancing the understanding of wireless edge content placement optimization via NTN. Our findings emphasize the importance of jointly considering NTN mobility and content popularity distribution for efficient delivery. While our approach optimizes NTN content placement iteratively on a time slot basis, future work should explore joint optimization across time slots to further reduce the content placement time.

References

[1] E. Oughton, G. Geraci, M. Polese, and V. Shah, "Prospective evaluation of next generation wireless broadband technologies: 6G versus Wi-Fi 7/8," *Available at SSRN: https://ssrn.com/abstract=4528119*, 2023.

[2] C.-X. Wang *et al.*, "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2023.

[3] A. Garcia-Rodriguez, D. López-Pérez, L. Galati-Giordano, and G. Geraci, "IEEE 802.11be: Wi-Fi 7 strikes back," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 102–108, 2021.

[4] L. Galati-Giordano, G. Geraci, M. Carrascosa, and B. Bellalta, "What will Wi-Fi 8 be? A primer on IEEE 802.11bn ultra high reliability," *arXiv:2303.10442*, 2023.

[5] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from space: An overview of 3GPP non-terrestrial networks," *IEEE Commun. Standards Mag.*, vol. 5, no. 4, pp. 147–153, 2021.

[6] M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6G era: Challenges and opportunities," *IEEE Network*, vol. 35, no. 2, pp. 244–251, 2021.

[7] G. Geraci, D. López-Pérez, M. Benzaghta, and S. Chatzinotas, "Integrating terrestrial and non-terrestrial networks: 3D opportunities and challenges," *IEEE Commun. Mag.*, pp. 1–7, 2023.

[8] O. Kodheli *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 70–109, 2021.

[9] G. Karabulut Kurt *et al.*, "A vision and framework for the high altitude platform station (HAPS) networks of the future," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 729–779, 2021.

[10] T. Darwish, G. K. Kurt, H. Yanikomeroglu, M. Bellemare, and G. Lamontagne, "LEO satellites in 5G and beyond networks: A review from a standardization perspective," *IEEE Access*, vol. 10, pp. 35 040–35 060, 2022.

[11] M. Benzaghta, G. Geraci, R. Nikbakht, and D. López-Pérez, "UAV communications in integrated terrestrial and non-terrestrial networks," in *Proc. IEEE Globecom*, 2022, pp. 1–6.

[12] G. Geraci *et al.*, "What will the future of UAV cellular communications be? A flight from 5G to 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1304–1335, 2022.

[13] X. Zhang, B. Zhang, K. An, G. Zheng, S. Chatzinotas, and D. Guo, "Stochastic geometry-based analysis of cache-enabled hybrid satellite-aerial-terrestrial networks with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1272–1287, 2022.

[14] F. Rinaldi, H.-L. Määttänen, J. Torsner, S. Pizzi, S. Andreev, A. Iera, Y. Koucheryavy, and G. Araniti, "Broadcasting services over 5G NR enabled multi-beam non-terrestrial networks," *IEEE Trans. Broadcasting*, vol. 67, no. 1, pp. 33–45, 2021.

[15] M. Luglio, S. P. Romano, C. Roseti, and F. Zampognaro, "Service delivery models for converged satellite-terrestrial 5G network deployment: A satellite-assisted CDN use-case," *IEEE Network*, vol. 33, no. 1, pp. 142–150, 2019.

[16] T. Kim, J. Kwak, and J. P. Choi, "Satellite edge computing architecture and network slice scheduling for IoT support," *IEEE IoT Journal*, vol. 9, no. 16, pp. 14 938–14 951, 2022.

[17] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, 2014.

[18] D. Liu, B. Chen, C. Yang, and A. F. Molisch, "Caching at the wireless edge: Design aspects, challenges, and future directions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 22–28, 2016.

[19] T. X. Vu, S. Chatzinotas, and B. Ottersten, "Edge-caching wireless networks: Performance analysis and optimization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2827–2839, 2018.

[20] J. Sedin, L. Feltrin, and X. Lin, "Throughput and capacity evaluation of 5G New Radio non-terrestrial networks with LEO satellites," in *Proc. IEEE Globecom*, 2020.

[21] A. A. Dowhuszko, J. Fraire, M. Shaat, and A. Pérez-Neira, "LEO satellite constellations to offload optical terrestrial networks in placement of popular content in 5G edge nodes," in *Proc. ICTON*, 2020, pp. 1–6.

[22] M. He, C. Zhou, H. Wu, and X. Sherman Shen, "Learning-based cache placement and content delivery for satellite-terrestrial integrated networks," in *Proc. IEEE Globecom*, 2021, pp. 1–6.

[23] Z. Ji, S. Wu, C. Jiang, and W. Wang, "Popularity-driven content placement and multi-hop delivery for terrestrial-satellite networks," *IEEE Commun. Letters*, vol. 24, no. 11, pp. 2574–2578, 2020.

[24] D. Han, W. Liao, H. Peng, H. Wu, W. Wu, and X. Shen, "Joint cache placement and cooperative multicast beamforming in integrated satellite-terrestrial networks," *IEEE Trans. Veh. Tech.*, vol. 71, no. 3, pp. 3131–3143, 2022.

[25] J. Li, K. Xue, D. S. L. Wei, J. Liu, and Y. Zhang, "Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2367–2381, 2020.

[26] C. Ding, J.-B. Wang, H. Zhang, M. Lin, and G. Y. Li, "Joint optimization of transmission and computation resources for satellite and high altitude platform assisted edge computing," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1362–1377, 2022.

[27] F. Wang, D. Jiang, Z. Wang, J. Chen, and T. Q. S. Quek, "Seamless handover in LEO based non-terrestrial networks: Service continuity and optimization," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 1008–1023, 2023.

[28] 3GPP Technical Report 38.811, "Study on New Radio (NR) to support non-terrestrial networks (Release 15)," Sep. 2020.

[29] 3GPP Technical Report 38.821, "Solutions for NR to support non-terrestrial networks (NTN) (Release 16)," Dec. 2019.

[30] 3GPP Technical Report 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)," Dec. 2019.

[31] ITU-R, "Attenuation by atmospheric gases (2016)," available online at https://www.itu.int/dmspubrec/itu-r/rec/p/R-REC-P.676-11-201609-S!!PDF-E.pdf.

[32] Q. Li, Y. Zhang, Y. Li, Y. Xiao, and X. Ge, "Capacity-aware edge caching in fog computing networks," *IEEE Trans. Veh. Tech.*, vol. 69, no. 8, pp. 9244–9248, 2020.

[33] FCC, "SpaceX non-geostationary satellite system (2018)," available online at https://fcc.report/IBFS/SAT-MOD-20181108-00083/1569860.