Journal of Turkish

Operations Management

Cell switching in 6G networks for improved sustainability and handover management

Metin Öztürk[1*],
[1]Ankara Yıldırım Beyazıt University, Ankara, Türkiye
metin.ozturk@aybu.edu.tr, ORCID No: http://orcid.org/0000-0001-8665-5291

*Corresponding Author

Article Info

Article History:
Received: 28.04.2025
Revised: 15.05.2025
Accepted: 03.06.2025

Keywords

Telecommunications,
6G Networks,
Cell switching,
Handover,
Optimization

Abstract

Sustainability and latency are two critical parameters for future generations of cellular communication networks, such as the sixth generation (6G). To this end, this work targets both sustainability and latency by strategically switching off idle or lightly loaded base stations to maximize network energy savings while minimizing the number of handovers, which are considered a source of latency in this study. In order to perform this optimization, a mixed-integer programming (MIP) problem is modeled, and a heuristic-based solution algorithm is developed. Moreover, high-altitude platform stations (HAPS), a non-terrestrial network (NTN) component, are integrated into the network architecture to provide additional coverage—with their large footprints—and capacity for cell switching
and traffic offloading purposes. Therefore, the key components of this study are energy efficiency and latency optimization through strategic cell switching operations, as well as the inclusion of HAPS as an offloading entity. Additionally, the developed MIP problem and heuristic-based solutions are also primary components of this work. The efficacy of the developed optimization problem and heuristic-based solution is validated through simulation studies, and the results confirm that the proposed methodology effectively reduces both energy consumption and the number of handovers, with performance strongly influenced by the handover penalty and the number of users in the network. Overall, the findings suggest that the outcomes of this research can enable more efficient and sustainable industrial operations and management through minimized energy consumption and handovers, along with the extensive coverage provided by HIBSs.

1. Introduction

1.1. Evolution of Cellular Networks

Wireless cellular communication networks have been evolving in a positive direction quite rapidly, and such improvement can be attributed to their generation-based structure. For example, the first generation of cellular communication technology (1G), introduced in the 1980s, was voice-only and analog, whereas in the 1990s, the second generation (2G) introduced digital technology and messaging services alongside voice communications. Data communication became widely available in the 2000s with the third generation (3G), along with full roaming capabilities due to global standardization through the 3rd Generation Partnership Project (3GPP). Furthermore, broadband communication, all-IP systems, and the early stages of Internet of things (IoT) technology emerged with the fourth generation (4G) starting in the 2010s (Goldsmith, 2005). The beginning of the 2020s saw the rollout of the fifth generation (5G) with hard-hitting promises, including remote surgery, augmented/virtual reality, smart concepts (cities, agriculture, infrastructure, ports, etc.), and key enabling technologies such as millimeter-wave communication, massive multiple-input multiple-output (mMIMO) systems, and pervasive IoT (Gupta& Jha, 2015; Chettri & Bera, 2020). Additionally, 5G introduced novel scenarios, namely enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive machine-type communications (mMTC), all of which target distinctive and numerous use cases. eMBB, for example, is aimed at enhancing data rates for bandwidth-hungry applications such as high-definition (HD) video streaming, online gaming, augmented/virtual reality, smart healthcare (e.g., real-time transmission of large files), and remote work and collaboration, to name a few, while URLLC supports critical applications, including remote surgery, autonomous vehicles, remote control of drones, and manufacturing and process control in industrial operations. Asset tracking, environmental monitoring, fleet management, and industrial IoT (IIoT) are among the use cases falling within the scope of the mMTC scenario. Therefore, 5G has been popularized with its verticals — that is, its support for a diverse set of fields — which are directly related to operational research and management. Among all, use cases such as remote work and collaboration, industrial operations, smart cities, and asset tracking are directly within the scope of operational research. Put another way, wireless communication technologies are among the key enablers of operational management, making it more autonomous, intelligent, and efficient (Banchs et al.,2019).

From this perspective, work related to shaping the sixth generation of cellular communication technology (6G) has already begun, and 6G is expected to be more efficient and powerful compared to legacy networks. The International Telecommunication Union (ITU) has published its vision for 6G networks, introducing novel components never seen in previous generations of cellular communication technology, along with multiple improvements to 5G networks (International Telecommunication Union [ITU], 2023). Starting with the enhancements to 5G: eMBB, URLLC, and mMTC are being transformed into immersive communication, hyper-reliable and low-latency communication, and massive communication, respectively, elevating the promising offerings of 5G networks to an even higher level. This can be interpreted as more efficient and powerful support for the aforementioned verticals, enabling industrial operations and management in a more effective and intelligent manner. Transforming eMBB into immersive communication is expected to significantly boost the data rates perceived by users, enabling technologies that require ultra-fast connectivity, while applications that are highly sensitive to reliability and latency issues will benefit from the shift from URLLC to its "hyper" version. Lastly, still within the context of 5G enhancements, massive communication will connect an even greater number of devices to the Internet, making operational management more widely available and accessible.

1.2. Sustainability in 6G Networks

When it comes to the new attributes in 6G networks (i.e., the components that have never been addressed in previous generations), the story becomes even more interesting and catchy. First of all, four different design principles (or overarching aspects) are introduced for 6G networks by the ITU: namely, sustainability, connecting the unconnected, ubiquitous intelligence, and security and resilience. These are crucial and highly challenging components, as they are very prone to conflict with other components. For example, sustainability — which has been defined for the first time in the history of wireless cellular communication technology — is challenging to achieve while improving the data rate of users through immersive communications. In order to increase the data rate, the use of more antennas (e.g., through ultra-massive MIMO), more base station (BS) deployments due to the shorter link ranges of millimeter waves and terahertz frequencies (which are discussed for use in 6G networks), and increasing the transmit power are among the potential solutions, and they all conflict with sustainability goals because of the increased network energy consumption. In addition, the design principles also conflict with each other, such that the connecting the unconnected principle again conflicts with sustainability due to greater infrastructure requirements, which, in turn, increase network energy consumption. Therefore, sustainability indeed necessitates special attention, not only to achieve the full potential of 6G networks, as it conflicts with various other 6G components, but also for environmental reasons. Additionally, there is an economic pillar of sustainability, and the increased energy consumption of cellular communication networks is reflected as energy costs for mobile network operators (MNOs), putting their business at risk. To this end, network energy consumption should be minimized while maintaining the other objectives of 6G networks, which is a challenging task and poses a non-trivial problem for the operation of wireless cellular communication networks.

In providing sustainability to cellular communication networks in the form of minimized network energy consumption, cell switching techniques are regarded as one of the primary solutions. Cell switching, in its most fundamental definition, is a set of techniques aimed at switching off BSs that are idle (i.e., unused) or lightly used (i.e., with minimal traffic load) in order to reduce the overall network power consumption. In this concept, when a BS is switched off, if there are users associated with it, they are offloaded to other BSs in the network. It has already been proven in the literature that cell switching is a successful approach for reducing network energy consumption, since the overall energy consumption of cellular communication networks is dominated by the radio access network (RAN) part, in which BSs are the main contributors. Therefore, the cell switching approach hypothesises that if the energy consumption of BSs can be reduced by putting them into a sleep mode (i.e., switching them off), it can have a considerable impact on decreasing the total energy consumption of a cellular network. However, switching off BSs is not a "free lunch" concept, as it can adversely affect the quality of service (QoS), since users are removed from their original BSs, which can subsequently reduce their achieved data rates. That is, a user is usually associated with a BS that provides the best signal strength, and once the user is offloaded from that BS, the signal strength is expected to degrade. Hence, a proper and efficient switching-off solution is required to avoid such adverse impacts.

The offloading destination, in this regard, plays a vital role in orchestrating the cell switching operations. In the literature, a control-data separation architecture (CDSA) is usually considered, where small BSs (SBSs) are responsible for high data provision, while the macro BS (MBS) is mainly responsible for control signaling. In this architecture, the traffic load of SBSs that are switched off is typically offloaded to the MBS. However, the capacity of the MBS is crucial in being able to accommodate the offloaded users. In other words, in maximizing energy savings through cell switching, capacity becomes the primary constraint. Nonetheless, capacity is often fully exploited in dense and ultra-dense urban network conditions, meaning that there is no room for additional traffic to be offloaded, and hence no room for cell switching. Moreover, this scenario is highly reliant on the availability of MBSs, which is not always the case, particularly in rural and remote environments.

1.3. The Role of Non-Terrestrial Networks

Therefore, at this point in the story, non-terrestrial networks (NTN) come onto the stage and act as a helping hand in enhancing network capacity and offloading destination availability, regardless of the scenario (i.e., urban, rural, or remote) (Iqbal et al., 2023). The integration of NTN into terrestrial networks is no longer just a vision, as standardization efforts related to such integration have already started from 3GPP Release 17, meaning that NTN is a player even in 5G networks and will be a key component in 6G networks, especially when considering the "connecting the unconnected" principle, which is impossible to achieve without NTN components in the loop (Saad et al., 2024). Among various NTN components, including uncrewed aerial vehicles (UAVs), high-altitude platform stations (HAPS), and satellite networks, HAPS has a special place due to its promising characteristics. Positioned at around 20 km above the Earth's surface, they provide a huge footprint when equipped with super macro base stations (SMBS) (Gunes et al., 2021). Compared to UAVs, they offer significantly more service time and coverage, and compared to satellites, they result in lower latency (Omid et al., 2024). This is the reason why the ITU has already considered the inclusion of HAPS in cellular communication networks in a harmonized manner and assigned frequency bands to it at the 2023 World Radiocommunication Conference (WRC-23). In the ITU's terminology, HAPS SMBS is regarded as HAPS as an IMT base station (HIBS), to indicate that they will be a part of cellular network operations.

1.4. Contributions of This Work

In this work, a cell switching approach is considered with the presence of HIBS as an offloading destination. Moreover, handover, which is defined as the change of BSs for a user equipment (UE) while being in an active mode (e.g., during ongoing data transmission or a voice call), is also included in the problem formulation, because the cell switching and handover concepts are highly correlated. In particular, users originally connected to BSs that are switched off are offloaded to HIBS, which can be regarded as a handover for users in active mode. Minimizing the number of handovers is one of the objectives for cellular communication networks, as each handover incurs additional latency due to the required signaling between network entities (i.e., UE, BSs, core network), thereby undermining the low-latency goals of 6G networks. On the other hand, the minimization of energy consumption through cell switching and the minimization of handover events become conflicting objectives in this setup, because whenever a BS is switched off (energy saved), its users are offloaded to HIBS (increased number of handovers). Hence, in this work, a multi-objective optimization problem is modelled, aiming to minimize both network energy consumption and the total number of handover events. The optimization variable is considered to be the combinations of cell switching operations (i.e., which BSs should be active/inactive), and therefore, the formulated problem is a combinatorial optimization problem. This is an important approach, as considering energy minimization alone is not an appropriate solution for users that are highly intolerant to latency (as in the hyper-reliable and low-latency communication usage scenario of 6G networks), and handover must therefore be included in the formulation.

The ON and OFF status of BSs is captured as a binary variable, where ON and OFF are represented by binary 1 and 0, respectively. Thus, the optimization problem is formulated as a mixed-integer programming (MIP) problem. In order to solve this multi-objective problem, a genetic algorithm-based solution is developed with an objective function that captures both the energy and handover costs, i.e., minimizing both. The contributions of this work are summarized as:

- A multi-objective optimization model is developed for the minimization of both energy and handover costs, in order to achieve the sustainability and latency goals of 6G networks.
- A greedy heuristic algorithm-based solution is designed for the optimization problem, with both the energy and handover costs included in the objective function.
- The impacts of the optimization outcomes are discussed from the perspectives of vertical industries, i.e., the support of 6G networks for operational research and optimization.
- A simulation environment is developed, incorporating cellular communication network components such as UEs, BSs, and HIBSs, to test the efficiency of the developed optimization model and genetic algorithm-based solution.

The rest of this paper is organized as follows: The related work is identified and thoroughly discussed in Section 2, and the detailed system modeling is introduced in Section 3. The problem formulation and methodology are presented in Section 4, followed by performance evaluations with the results and discussion in Section 5. Lastly, Section 6 concludes the paper. The list of key acronyms used throughout the paper is given in Table 1.

Table 1: List of Acronyms for the Key Terminology.

Acronym	Definition
3GPP	3rd Generation Partnership Project
BS	Base station
CDSA	Control-data separation architecture
eMBB	Enhanced mobile broadband
HAPS	High-altitude platform stations
HIBS	HAPS as an IMT base station
IIoT	Industrial Internet of things
IoT	Internet of things
ITU	International Telecommunication Union
MBS	Macro base station
MIP	Mixed-integer programming
mMIMO	Massive multiple-input multiple-output
mMTC	Massive machine-type communications
MNO	Mobile network operator
NTN	Non-terrestrial networks
RAN	Radio access network
SBS	Small base station
SMBS	Super macro base stations
UAV	Uncrewed aerial vehicle
UE	User equipment
URLLC	Ultra-reliable low-latency communications

2. Literature Review

The literature on cell switching can be categorized into two areas within the scope of this work: terrestrial cell switching and cell switching involving NTN components. The former refers to cases where only terrestrial BSs are considered in the network architecture, while the latter involves NTN components, specifically HAPS, in their cell switching implementations.

Starting with terrestrial cell switching, in (Han et al., 2024), the authors employed a deep reinforcement learning framework to determine the optimum set of BSs to be switched off. However, since the action space increases exponentially with the number of BSs included in the network, optimization of the action space was also considered. Their results reveal significant energy savings compared to off-the-shelf benchmark algorithms. The combination of deep reinforcement learning and federated learning was utilized in (Movahedkor & Shahbazian, 2024), wherein a reinforcement learning framework is used to learn the optimal cell switching strategy, while federated learning is primarily adopted to share data between agents to improve learning performance. Federated learning offers a secure way of data transmission between network entities, as only model parameters are shared rather than raw and sensitive data; thus, privacy is also maintained in their model. A more comprehensive and detailed cell switching study was conducted in (Tan et al., 2023), wherein the authors sought to optimize not only the operational status of BSs, but also beamforming vectors and user-BS connections. To achieve this, the transmit power constraint of BSs was taken into consideration along with the QoS requirements from users and the profit objectives for MNOs. The authors developed two different algorithms: to optimize the BS ON/OFF status and UE-BS connection state, they developed a roaming-cost-based BS switching-off algorithm (EE-BSOA), while to optimize the beamforming vector, an energy efficiency maximization beamforming algorithm (EE-BFA) was employed. Additionally, in (Lin et al., 2024), a joint problem modelling of long-term cell switching control and short-term beamforming control was presented. The idea in their work is that by controlling the ON/OFF status of BSs over the long term, stable UE-BS connectivity is maintained, and the energy cost of frequent cell switching is minimized. Simultaneously, the beamforming vectors are updated rapidly to minimise transmit power consumption, resulting in a holistic energy-efficient framework. The work in (Yahya & Stanica, 2024) investigates the impact of cell switching implementation in an urban area using a real-life dataset collected from Lyon, France. A heuristic cell switching approach was employed, and their results confirmed that cell switching has significant potential in reducing network energy consumption. The authors in (Abubakar et al., 2025) proposed a secure cell switching approach with the help of federated learning. A real-life dataset collected from Milan, Italy, was used, and the traffic loads of BSs were predicted to achieve a proactive cell switching mechanism. For the traffic prediction phase, a bi-directional long short-term memory (LSTM) algorithm was developed, followed by optimization of the ON/OFF status of BSs in the network using a modified version of the actor-critic deep reinforcement learning method. In summary, the literature on cell switching applied in terrestrial networks is quite rich, with a diverse set of studies addressing the cell switching problem from various perspectives.

On the other hand, this is not the case for cell switching implementations using NTN components, especially considering HAPS technology as one of the main focuses of this current work. The authors in (Salamatmoghadasi et al., 2024) considered a HAPS-SMBS in their network architecture, serving as an offloading destination for the traffic of switched-off terrestrial BSs. In this regard, a sorting-based cell switching algorithm was developed, wherein the terrestrial BSs are sorted based on their traffic loads in ascending order, followed by switching them off until the capacity of the HAPS-SMBS is fully utilized. Through this approach, significant energy savings are achieved in the network. Although not directly focused on cell switching, the study in (Kement et al., 2023) investigates the impacts of deploying HAPS-SMBSs in networks where additional capacity is required. Such deployment was compared to terrestrial BS deployments (i.e., network densification), and it was observed that deploying a HAPS-SMBS is equivalent to the deployment of multiple terrestrial BSs. HAPS-SMBS is also used in (Song et al., 2024) to switch off terrestrial BSs and offload their traffic. The authors utilized a real-life dataset collected from Milan, Italy, and processed it to reflect the current scale of data traffic using another real-life dataset collected from China. Similar to (Salamatmoghadasi et al., 2024), a sorting-based least-traffic offloading algorithm was developed to facilitate the cell switching and traffic offloading framework. The cell traffic load estimation problem of sleeping BSs was considered in (Ciloglu et al., 2024), where the impact of errors during the estimation of cell traffic loads was investigated in a HIBS-enabled cell switching scheme. Moreover, two different Q-learning models were designed: one full-scale model to prioritize cell switching performance, and one lightweight model to balance computation cost. The study by Mbarek et al. (2024) follows a predictive cell switching framework, wherein the future traffic loads of BSs are predicted in advance to enable timely cell switching decisions. Overloaded BSs were identified to determine which ones required traffic offloading, and a Q-learning algorithm was employed for this purpose. The authors incorporated HAPS into their network architecture, using it as an offloading entity for the traffic of switched-off BSs. HAPS-based system architecture was also considered in (Nauman et al., 2024), where a three-layer resource optimization approach was developed. The authors considered a non-orthogonal multiple access (NOMA)-based vehicular-aided heterogeneous network. User associations, bandwidth allocation, and transmit power are optimized to maximize the network utilization while respecting the QoS constraints. An access point switching was considered as a second-layer solution to further enhance the network resource utilization. A very comprehensive and novel multi-layer cell switching framework was proposed in (Ozturk et al., 2025). In their work, in addition to terrestrial MBSs, UAV-BSs, HIBS, and satellites were considered as offloading destinations for the traffic of switched off BSs. Since latency is a key consideration when it comes to satellite networks, a latency-aware cell switching approach was also developed in addition to conventional energy-focused cell switching. Their results reveal the potential of multi-tier cell switching when all the components of NTN are included.

The literature on HIBS/HAPS-supported cell switching is still in its infancy, and this work is an important attempt to fill a gap, as, to the best of the authors' knowledge, there is no existing study that jointly optimizes energy and handover costs in a HIBS-assisted terrestrial network architecture. Since NTN, sustainability, and latency are three crucial factors shaping 6G networks, this work, positioned at the intersection of these domains, is expected to pave the way for more efficient network management, making operational management in vertical industries more effective, accessible, and intelligent.

3. System Modeling

In this section, the network, user mobility, user-BS association, energy consumption, and handover models are presented in detail.

3.1. Network Model

As seen in Figure 1, an $S \times S$ m^2 area is considered as an environment, wherein users, denoted by u_k for $k = 1, 2, \cdots, K$, are initially (at $t = 0$, where $t = 1, 2, \cdots, T$ denotes the time slot) distributed around the environment following a uniform distribution, such that $x_k(0) \sim \mathcal{U}(0, S)$ and $y_k(0) \sim \mathcal{U}(0, S)$, where x_k and y_k are the x and y coordinates of user k, and S is the lengths of the square-shaped simulation area. $\mathcal{U}(a, b)$ denotes a uniform distribution over the interval ($[a, b]$), and $x_k(t)$ and $y_k(t)$ are independent variables, such that their distributions are not correlated to each other. The set of K mobile users is defined as $U = \{u_1, u_2, \cdots, u_K\}$.

Additionally, there are terrestrial BSs, denoted by λ_i for $i = 1, 2, \cdots, \Lambda$, distributed around the environment following a uniform distribution, such that $x_\lambda \sim \mathcal{U}(0, S)$ and $y_\lambda \sim \mathcal{U}(0, S)$, where x_λ and y_λ are the x and y coordinates of the terrestrial BS λ. The set of Λ terrestrial BSs is defined as $B = \{\lambda_1, \lambda_2, \cdots, \lambda_\Lambda\}$. Similarly, there are HIBSs, denoted by h_j for $j = 1, 2, \cdots, M$, distributed around the two-dimensional (2D) environment following a uniform distribution, such that $x_h \sim \mathcal{U}(0, S)$ and $y_h \sim \mathcal{U}(0, S)$, where x_h and y_h are the x and y coordinates of HIBS h. The set of M HIBSs is given as $H = \{h_1, h_2, \cdots, h_M\}$, and the third dimension (i.e., altitude) of all the HIBSs is assumed to be fixed and set to 20 km.

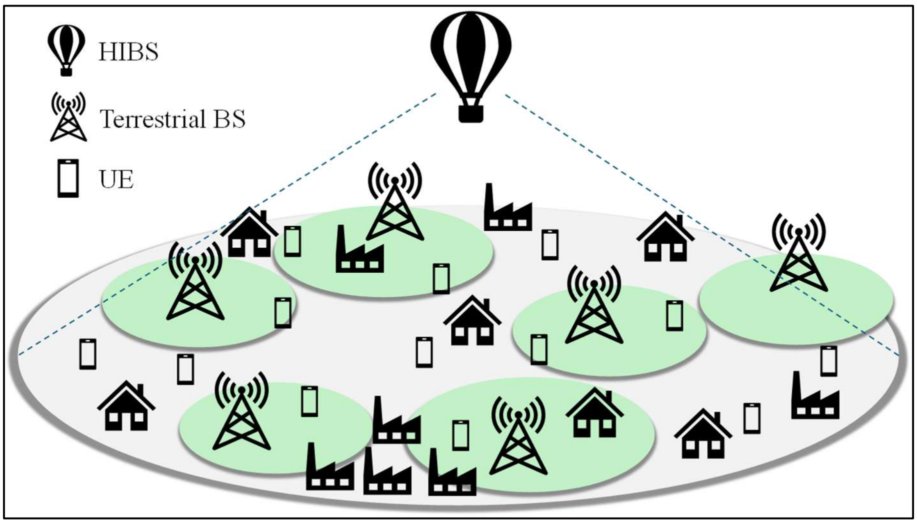

Figure 1: The network model consisting of UEs, terrestrial BSs, and HIBSs.

3.2. User Mobility Model

After the initial distribution at $t = 0$, all the users move following a Gaussian random walk, such that

$$\mathbf{p}_k(t) = \mathbf{p}_k(t-1) + \Delta\mathbf{p}_k(t), \tag{1}$$

where $\mathbf{p}_k(t) = (x_k(t), y_k(t))$ is the position of user k at time t, and $\Delta\mathbf{p}_k(t) = (\Delta x_k(t), \Delta y_k(t))$ is the displacement vector at time t. The displacement components (i.e., $\Delta x_k(t)$ and $\Delta y_k(t)$) are independently and normal distributed as $\Delta x_k(t) \sim N(0, \sigma^2)$ and $\Delta y_k(t) \sim \mathcal{N}(0, \sigma^2)$, where $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian/normal distribution with mean μ and standard distribution σ. In order to keep the users within the environment considered, the updated user positions are clamped as

$$\mathbf{p}_k(t) = (\max(0, \min(x_k(t), S)), \max(0, \min(y_k(t), S))). \tag{2}$$

Note that the terrestrial BSs and HIBSs are assumed to be static, i.e., stay at their fixed location during the entire simulation.

3.3. User Association Model

In this work, users are associated with a BS or HIBS based on the Euclidian distance between the users and BSs or HIBS (i.e., proximity) and availability of BSs, such that the association function a_k for user u_k at time t is defined as

$$a_k(t) = \begin{cases} \underset{\lambda_i \in B,\ \beta_{\lambda_i}(t)=1}{\operatorname{argmin}} \ d(u_k(t), \lambda_i) & \text{if} \quad \exists\, \lambda_i \in B \text{ s.t. } \beta_{\lambda_i}(t) = 1, \\ \underset{h_j \in H,\ \beta_{h_j}(t)=1}{\operatorname{argmin}} \ d(u_k(t), h_j) & \text{if} \quad \nexists\, \lambda_i \in B \text{ s.t. } \beta_{\lambda_i}(t) = 1 \text{ and } \exists\, h_j \in H \text{ s.t. } \beta_{h_j}(t) = 1 \\ \underset{\lambda_i \in B}{\operatorname{argmin}} \, d(u_k(t), \lambda_i) & \text{otherwise,} \end{cases} \tag{3}$$

where $d(u_k(t), \lambda_i)$ and $d(u_k(t), h_j)$ denote the Eurclidian distance between user u_k and terrestrial BS λ_i, and between user u_k and HIBS h_j, respectively. Here, let $N = B \cup H = \{b_1, b_2, \cdots, b_{\Lambda+M}\}$ denote the set of all types of BSs (terrestrial and HIBS), where b_n denotes any BS n that is either a terrestrial BS or a HIBS. Then, let $\beta_{b_n}(t)$ is the ON/OFF status of a BS n such that

$$\beta_{b_n}(t) = \begin{cases} 1 & \text{if} \quad b_n \text{ is ON} \\ 0 & \text{if} \quad b_n \text{ is OFF.} \end{cases} \tag{4}$$

Here, $\boldsymbol{\beta}(t) = [\beta_{b_1}, \beta_{b_2}, \cdots, \beta_{b_{\Lambda+M}}] \in \{0,1\}^{\Lambda+M}$ is a decision variable, defining the ON/OFF status of the all the BSs included in the network model.

3.4. Energy Model

The energy consumption of all the BSs (terrestrial and HIBS) are considered to be constant in case being active and zero otherwise. In this regard, the total energy consumption model accounting for all the BSs are

$$E_t(t) = \sum_{n=1}^{\Lambda+\mathrm{M}} \beta_{b_n}(t) E_o, \tag{5}$$

where E_o is the operational energy consumptions. The overall network energy consumption during the simulation time frame is calculated as

$$E_T = \sum_{t=1}^{T} E_t(t). \tag{6}$$

3.5. Handover Model

In this work, the handover event is triggered when a UE changes its serving BS, such that handover is modeled as a binary variable for each UE at time t

$$\varrho_k(t) = \begin{cases} 1 & \text{if} \quad a_k(t) \neq a_k(t-1) \\ 0 & \text{if} \quad a_k(t) = a_k(t-1), \end{cases} \tag{7}$$

and the total number of handovers at time t is obtained as

$$\rho(t) = \sum_{u_k \in U} \varrho_k(t), \tag{8}$$

which makes the overall handover count during the simulation time frame as

$$\rho_T = \sum_{t=1}^{T} \rho(t). \tag{9}$$

3.6. The Overview of Assumptions

The assumptions considered in this work are given as follows:

- The simulation environment is a square area of size $S \times S$ m², with users, terrestrial BSs, and HIBSs initially distributed uniformly.
- HIBSs are positioned at a fixed altitude of 20 km.
- Users move according to a Gaussian random walk with normally distributed displacements, and their positions are clamped within the simulation area.
- Terrestrial BSs and HIBSs are static, remaining at fixed locations throughout the simulation.
- Users associate with the nearest active BS or HIBS based on Euclidean distance; if no BSs are active, users are not associated.
- All BSs (terrestrial and HIBSs) consume a constant energy (E_o) when ON and zero energy when OFF.

- A handover occurs when a user's serving BS changes between time slots, modeled as a binary variable.
- The optimization problem is solved independently at each time step.
- At least one BS must be active at all times.
- Inactive BSs have a fixed probability (p_r) of reactivation at each time step.
- Energy and handover costs are non-negative.

4. Problem Formulation and Methodology

In this section, the problem is presented and formulized as an optimization problem, followed by introducing the methodology employed in this work,

4.1. Problem Formulation

In this work, network energy consumption and the total number of handovers are jointly minimized by modeling a multi-objective optimization problem. In this regard, the optimization problem is considered to be a minimization problem and run at each time step t independently, such that

$$\underset{\alpha(t)}{\text{minimize}}\ E_t(t) + \gamma\rho(t) \tag{10}$$

subject to (10a)

$\beta_{b_n}(t) \in \{0, 1\} \quad \forall b_n \in N, \forall t$ (10b)

$a_k(t) \in \{b_n \in N: \beta_{b_n} = 1\} \cup \{0\} \quad \forall u_k \in U, \forall t$ (10c)

$\sum_{b_n \in N} \beta_{b_n}(t) \geq 1, \quad \forall t$ (10d)

$P(\beta_{b_n} = 1 | \beta_{b_n} = 0) = p_r$ (10e)

$0 \leq x_k(t), y_k(t) \leq S, \quad \forall u_k \in U, \forall t$ (10f)

$E_t(t) \geq 0, \rho(t) \geq 0.$ (10g)

Here, p_r indicates a fixed reactivation probability for all the BSs in the network. This problem is a combinatorial and MIP problem, and the solution time increases exponentially with the number of BSs in the network: i.e., the search space equals to $2^{\Lambda+\mathrm{M}}$. In the optimization, the variable α is strategically selected, as it impacts both total energy consumption, E_t, and the number of handovers, ρ. First, Equation (5) indicates that network energy consumption is a function of the number of active BSs in the network. When the number of active BSs is reduced through cell switching (via α\alphaα), the network's energy consumption can be minimized, satisfying the objective function in Equation (10). Second, a handover occurs when a user moves from one BS to another, as described in Equation (7). When the BS density in the network—defined as the number of active BSs per unit area—increases, the probability that a user requires a handover also increases, as it becomes more likely for the user to enter the footprint of another BS. With this in mind, switching off some BSs in the network reduces BS density, and thus lowers the probability of handovers.

4.2. Methodology

In this work, a greedy heuristic algorithm is employed, and the algorithmic details are given in Algorithm 1.

Algorithm 1: Greedy Heuristic Algorithm for BS ON/OFF Status Optimization

Inputs:
Positions of UEs: $\mathbf{p}_k, \ \forall u_k \in U$
Parameters: Λ, M, K, γ, p_r, and E_o
BS ON/OFF status: $\boldsymbol{\beta}(t-1)$
Outputs:
BS ON/OFF status: $\boldsymbol{\beta}(t)$
UE-BS associations: $a_k(t), \ \forall u_k \in U$
Energy cost: $E_t(t)$
Handover cost: $\rho(t)$
1. *% Initialize*
$\boldsymbol{\beta}(t) = \boldsymbol{\beta}(t-1)$
$a_k(t) = 0, \ \forall u_k \in U$
2. *% User mobility*
for each u_k in U
    $\mathbf{p}_k(t) = \mathbf{p}_k(t-1) + \Delta p_k(t)$
3. *% User-BS association*
for each u_k in U

$d_n = ||\mathbf{p}_k(t) - \mathbf{q}_m||, \forall b_n \in N$
if $\sum_{b_n \in N} \beta_{b_n}(t) \geq 1, \quad \forall t$ *% active BS exists*
$a_k(t) = \underset{b_n \in N,\ \beta_n(t)=1}{\text{argmin}} \ d(u_k(t), b_n)$

else $a_k(t) = 0$
if $a_k(t) \neq a_k(t-1)$ *% handover is triggered*
$\varrho_k(t) = \varrho_k(t) + 1$ *% increase the handover count*

4. *% Switch off BSs*
for each b_n in N
$\eta_n(t) = |\{u_k : a_k(t) = b_n\}|$ *% compute the user count*
if $\beta_{b_n}(t) = 1$ and $\eta_n(t) < \frac{1}{\gamma}$
Set $\beta_{b_n}(t) = 0$
for each u_k in $\{a_k(t) = b_n\}$
Recompute $a_k(t) = \underset{b'_n \in N,\ \beta'_{b_n}(t)=1}{\text{argmin}} \ d(u_k(t), b'_n)$
if $a_k(t) \neq a_k(t-1)$ *% handover is triggered*
$\varrho_k(t) = \varrho_k(t) + 1$ *% increase the handover count*
5. *% Reactivate BSs*
for each b_n in N
if $\beta_{b_n}(t) = 0$ and rand() $< p_r$
Set $\beta_{b_n}(t) = 1$
for each u_k in $\{a_k(t) = b_n\}$
Recompute $a_k(t) = \underset{b'_n \in N,\ \beta'_{b_n}(t)=1}{\text{argmin}} \ d(u_k(t), b'_n)$
if $a_k(t) \neq a_k(t-1)$ *% handover is triggered*
$\varrho_k(t) = \varrho_k(t) + 1$ *% increase the handover count*
6. *% Compute energy cost and handover cost*

$$E_t(t) = \sum_{n=1}^{\Lambda+\mathrm{M}} \beta_{b_n}(t) E_o$$

$$\rho(t) = \sum_{u_k \in U} \varrho_k(t),$$

7. Return:
$\boldsymbol{\beta}(t)$, $\{a_k(t)\}_{u_k \in U}$, $E_t(t)$, and $\rho(t)$

5. Performance Evaluation

In order to measure the performance of the proposed methodology in the developed problem formulation, a set of simulation campaigns are conducted. The simulation parameters are given in Table 1.

Table 2: Simulation Parameters

Parameter	Description	Value
K	Number of UEs considered in the network	{50, 100, 150, 200}
Λ	Number of terrestrial BSs	3
M	Number of HIBSs	2
S	Dimension of the simulation area	1000 m
T	Number of time slots	100
E_o	Operational energy consumption	1 unit
γ	Handover penalty	{0.01, 0.1, 1, 10, 100}
p_r	Probability of reactivating an inactive BS per time step	{0.01, 0.05, 0.1}
R	Number of simulation runs for averaging purposes	10
σ	Standard deviation of Gaussian displacement in UE mobility model	10 (units per time step)

5.1. Results and Discussions

The average energy cost for various values of handover penalty (γ) is illustrated in Figure 2 for the scenario where $\Lambda = 3$, $M = 2$, and $p_r = 0.01$. The results in Figure 2 demonstrate that the average energy cost increases with growing values of handover penalty, because when the handover penalty is higher, the algorithm attempts fewer handovers, which, in turn, increases the energy cost. In other words, at lower handover penalty values, there are more opportunities to switch off BSs. Another takeaway from Figure 2 is that the average energy cost saturates at some point, such that it does not increase further beyond a certain value of handover penalty (mainly after $\gamma = 10$). This behavior can be attributed to the fact that when the handover penalty is set to a high value, no BSs in the network are switched off, which acts as a saturation point. Additionally, the results confirm that the average energy cost scales with the number of UEs considered in the network, as more UEs result in fewer switching opportunities, considering the cell switching criteria in Algorithm 1.

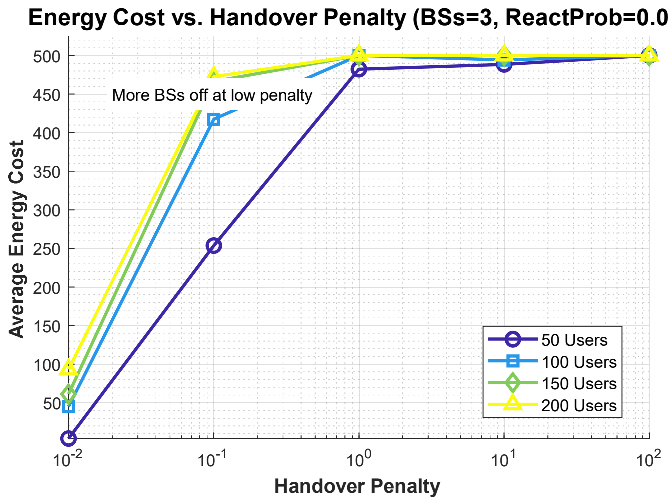

Figure 2: Average energy cost for sweeping values of handover penalty ($\Lambda = 3$, $M = 2$, $p_r = 0.01$).

For the same setup (i.e., ($\Lambda = 3$, $\mathrm{M} = 2$, and $\mathrm{p_r} = 0.01$), the handover cost is depicted in Figure 3 across varying values of handover penalty. The results in Figure 3 show that the average handover cost (i.e., the total number of handovers) is highly dependent on both the handover penalty and the number of UEs in the network. First, the average handover cost decreases with increasing values of handover penalty. This is because a higher penalty is incurred for each handover, which subsequently impacts the objective function given in Equation (10). Put another way, Equation (10) includes γ as a multiplier of the number of handovers; therefore, when γ is high, the heuristic algorithm tries to reduce the number of handovers (ρ) in order to comply with the minimization objective. Second, the average number of handovers increases with the number of UEs in the network (K), as a higher number of users raises the probability of handovers, given that the environment area is fixed. In other words, when there are more users in the network, it becomes more likely for handovers to occur due to the increasing network density.

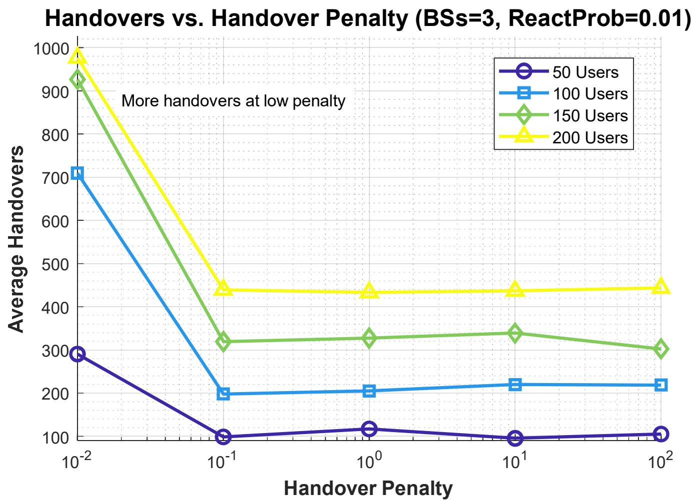

Figure 3: Average handovers for sweeping values of handover penalty ($\Lambda = 3$, $M = 2$, $p_r = 0.01$).

In Figure 4, the energy cost results are presented for an increased number of terrestrial BSs, i.e., $\Lambda = 7$, in order to observe the system behaviour as the network size grows. This set of results reveals that the energy cost follows the same trend as in Figure 2, such that the energy cost increases with increasing handover costs because the switching-off opportunities narrow with growing handover penalties. However, unlike the results in Figure 2, the magnitude of the energy cost is higher in Figure 4. This is attributed to the energy model given in Equation (5), where the network energy consumption increases with the number of BSs included in the model. Additionally, the average energy cost does not saturate when the number of users is low, i.e., $K = 50$. This is because when the number of users is low, the number of users per BS is also low, leading to more switching-off opportunities even with increased handover penalty values.

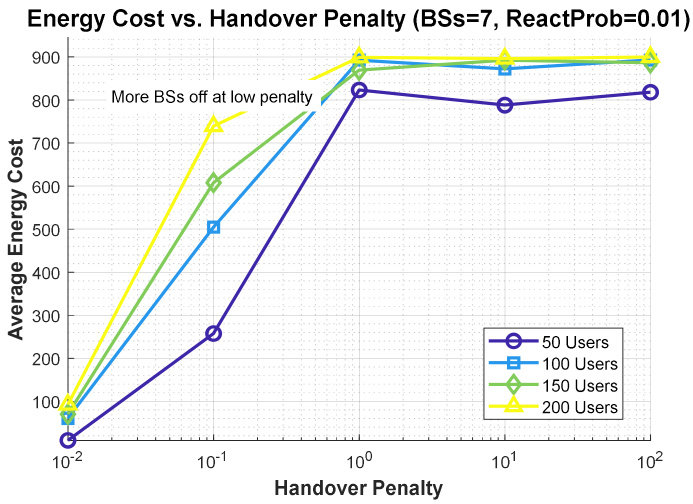

Figure 4: Average energy cost for sweeping values of handover penalty ($\Lambda = 7, M = 2,\ p_r = 0.01$).

Figure 5 illustrates the average number of handovers when the network size increases with more terrestrial BSs ($\Lambda = 7$). Unlike the difference observed between Figures 2 and 4, no significant change is observed in the average number of handovers as the network size grows. The major difference lies in the scale of the number of handovers, which significantly increases when the number of terrestrial BSs increases from 3 to 7. This is because a higher number of BSs leads to increased BS density, while the dimensions of the environment remain the same. The increased BS density raises the probability of a user encountering a BS during movement, which in turn increases the total handover count. Additionally, the number of handovers stabilizes at higher handover penalty values when the number of users is 50. Similar to the results in Figure 4, this behavior is attributed to the increased number of switching-off opportunities when the user density is lower, which subsequently reduces the number of handovers as there are fewer active BSs in the network.

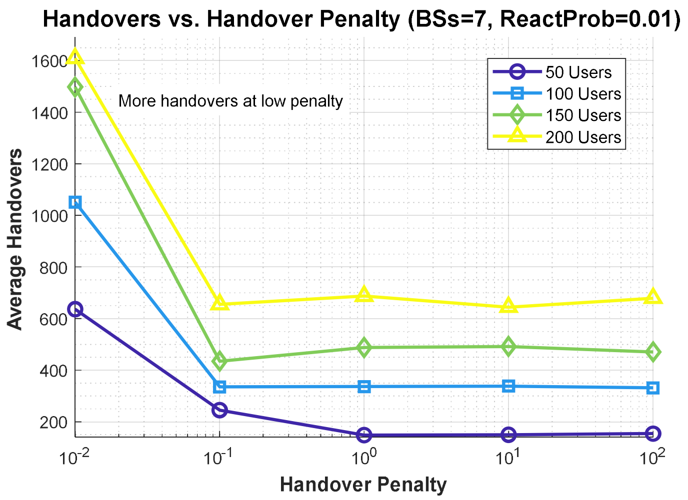

Figure 5: Average handovers for sweeping values of handover penalty ($\Lambda = 7, M = 2,\ p_r = 0.01$).

With the results presented in Figures 2, 3, 4, and 5, scalability is a real concern. Since cell switching is a binary problem, it usually has $O(2^{(\Lambda+M)})$ complexity, as the number of switching options increases exponentially with the number of BSs in the network. The heuristic algorithm developed in this paper, on the other hand, has a near-linear complexity ($\mathrm{O}(TK(\Lambda + M))$), meaning that it remains scalable up to medium-sized networks. For comparison, exhaustive search, the optimum solution for cell switching, is not scalable at all because its complexity is also exponential. Having said that, for ultra-dense deployments, more scalable and effective solutions will be needed.

Through this work, 6G networks are made more efficient and feasible, enhancing their suitability for the operations of vertical industries. First, sustainability in 6G networks is promoted through the problem formulation and heuristic-based solution proposed in this work. This, in turn, leads to environmentally positive outcomes, as network energy consumption is minimized. Additionally, economic sustainability is also addressed, as reduced energy consumption translates into lower energy costs for MNOs, making their businesses more sustainable. Second, by emphasizing the minimization of the number of handovers, this work enables more feasible cellular network operations and helps meet latency-related requirements, given that each handover incurs additional time during the process. Furthermore, since latency is minimized, various industrial applications—including IIoT, industrial automation, and extended reality use cases—can be supported. Third, incorporating HIBS not only provides additional capacity and flexibility to 6G networks but also maximizes network coverage, enabling ubiquitous connectivity. This, in turn, supports various operations such as asset management and environmental monitoring, to name a few. Therefore, through the outcomes of this research study:

- 6G networks are made more efficient and manageable;
- standardization requirements for 6G can be achieved; and
- industrial operations and management for vertical industries are made more effective, feasible, and accessible.

6. Conclusion

In this research, cell switching was employed to jointly optimize energy efficiency (i.e., sustainability) and latency in 6G networks. In contrast to existing literature, where cell switching is primarily utilized for energy-saving purposes, this work considers cell switching as an enabler for latency minimization through strategic handover management. The underlying idea is that by reducing the number of base stations (BSs) in the network via cell switching, the BS density decreases, thereby reducing the likelihood of handover events per user movement. In this work, handovers are considered a source of latency. To simultaneously optimize energy efficiency and handover management, a MIP and combinatorial optimization problem was modelled, and a greedy heuristic-based solution algorithm was developed. The proposed framework minimizes the energy consumption of 6G networks while also reducing the number of handovers, achieved through the strategic design of the objective function. This function incorporates both the total energy cost and the product of the handover penalty and the number of handovers. Additionally, HIBS were considered as NTN components to provide additional capacity. This enables more extensive cell switching operations, thereby increasing energy savings and further reducing handovers. The inclusion of HIBSs also enhances network coverage and supports the "connecting the unconnected" vision of 6G networks. Simulation studies were conducted for a specific set of parameters by sweeping the handover penalty and the number of users in the network. The results revealed that both the energy cost and the number of handovers are functions of these parameters, and an increasing handover penalty leads to higher energy costs but fewer handovers, while a higher number of UEs results in increases in both energy cost and handover events. In summary, this work enables more sustainable operation of 6G networks and supports more efficient orchestration with a reduced number of handovers. Furthermore, the outcomes of this work can enable a set of use cases, significantly impacting operations and management across various sectors. Future work will focus on optimizing additional handover parameters such as hysteresis, time-to-trigger, and handover margin, within the cell switching framework. Moreover, the integration of additional NTN components, including UAVs and satellites, is also planned for future studies.

Conflict of Interest

No conflict of interest was declared by the author.

References

Abubakar, A. I., Mollel, M. S., Ozturk, M., & Ramzan, N. (2025). A secured energy saving with federated assisted modified actor-critic framework for 6G networks. *IEEE Transactions on Vehicular Technology*. https://doi.org/10.1109/TVT.2025.3540641

Abbasi, O., Yadav, A., Yanikomeroglu, H., Đào, N. -D., Senarath, G., & Zhu, P. (2024). HAPS for 6G networks: Potential use cases, open challenges, and possible solutions. *IEEE Wireless Communications, 31*(3), 324–331. https://doi.org/10.1109/MWC.012.2200365

Banchs, A., Gutierrez-Estevez, D. M., Fuentes, M., Boldi, M., & Provvedi, S. (2019). A 5G mobile network architecture to support vertical industries. *IEEE Communications Magazine, 57*(12), 38–44. https://doi.org/10.1109/MCOM.001.1900258

Çiloğlu, B., Koç, G. B., Ozturk, M., & Yanikomeroglu, H. (2024). Cell switching in HAPS-aided networking: How the obscurity of traffic loads affects the decision. *IEEE Transactions on Vehicular Technology, 73*(11), 17782–17787. https://doi.org/10.1109/TVT.2024.3420245

Chettri, L., & Bera, R. (2020). A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. *IEEE Internet of Things Journal, 7*(1), 16–32. https://doi.org/10.1109/JIOT.2019.2948888

Goldsmith, A. (2005). *Wireless communications*. Cambridge University Press. https://www.cambridge.org/core/books/wireless-communications/800BA8A8211FBECB133A7BB77CD2E2BD

Gupta, A., & Jha, R. K. (2015). A survey of 5G network: Architecture and emerging technologies. *IEEE Access, 3*, 1206–1232. https://doi.org/10.1109/ACCESS.2015.2461602

Han, Z., et al. (2024). Energy consumption optimization for 5G base stations based on deep reinforcement learning. In *2024 10th International Conference on Computer and Communications (ICCC)* (pp. 1625–1629). https://doi.org/10.1109/ICCC62609.2024.10942221

Iqbal, A., et al. (2023). Empowering non-terrestrial networks with artificial intelligence: A survey. *IEEE Access, 11*, 100986–101006. https://doi.org/10.1109/ACCESS.2023.3314732

Kement, C. E., et al. (2023). Sustaining dynamic traffic in dense urban areas with high altitude platform stations (HAPS). *IEEE Communications Magazine, 61*(7), 150–156. https://doi.org/10.1109/MCOM.001.2200584

Lin, J., Ma, M., & Wang, X. (2024). Joint design of long-term base station activation and short-term beamforming for green wireless networks. *IEEE Transactions on Wireless Communications, 23*(1), 185–200. https://doi.org/10.1109/TWC.2023.3276639

Mbarek, M. A., Meo, M., Renga, D., & Vallero, G. (2024, October). Adaptive HAPS Offloading: A Strategy for Supporting RAN During High Traffic Load. In *2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)* (pp. 1-6). https://doi.org/10.1109/WiMob61911.2024.10770387

Movahedkor, N., & Shahbazian, R. (2024). Decentralized federated deep reinforcement learning framework for energy-efficient base station switching control. In *2024 11th International Symposium on Telecommunications (IST)* (pp. 455–460). https://doi.org/10.1109/IST64061.2024.10843518

Nauman, A., Maashi, M., Alkahtani, H. K., Al-Wesabi, F. N., Aljehane, N. O., Assiri, M., ... & Khan, W. U. (2024). Efficient resource allocation and user association in NOMA-enabled vehicular-aided HetNets with high altitude platforms. *Computer Communications*, *216*, 374-386. https://doi.org/10.1016/j.comcom.2024.01.021

Omid, O., Yadav, A., Yanikomeroglu, H., Đào, N. -D., Senarath, G., & Zhu, P. (2024). HAPS for 6G networks: Potential use cases, open challenges, and possible solutions. *IEEE Wireless Communications, 31*(3), 324–331. https://doi.org/10.1109/MWC.012.2200365

Ozturk, M., Salamatmoghadasi, M., & Yanikomeroglu, H. (2025). Integrating Terrestrial and Non-Terrestrial Networks for Sustainable 6G Operations: A Latency-Aware Multi-Tier Cell-Switching Approach. *IEEE Network*. https://doi.org/ 10.1109/MNET.2025.3554393

Saad, M. M., Tariq, M. A., Khan, M. T. R., & Kim, D. (2024). Non-terrestrial networks: An overview of 3GPP release 17 & 18. *IEEE Internet of Things Magazine, 7*(1), 20–26. https://doi.org/10.1109/IOTM.001.2300154

Salamatmoghadasi, M., Mehrabian, A., & Yanikomeroglu, H. (2024). Energy sustainability in dense radio access networks via high altitude platform stations. *IEEE Networking Letters, 6*(1), 21–25. https://doi.org/10.1109/LNET.2023.3328918

Song, T., Lopez, D., Meo, M., Piovesan, N., & Renga, D. (2024). High altitude platform stations: The new network energy efficiency enabler in the 6G era. In *2024 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1–6). https://doi.org/10.1109/WCNC57260.2024.10571153

Yahya, S., & Stanica, R. (2024). Assessing the energy impact of cell switch off at an urban scale. In *2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)* (pp. 36–41). https://doi.org/10.1109/WiMob61911.2024.10770444