Adaptive HAPS Offloading: A Strategy for Supporting RAN During High Traffic Load

Adaptive HAPS offloading: A strategy for supporting RAN during high traffic load

Mohamed Amine Mbarek, Michela Meo, Daniela Renga and Greta Vallero
Department of Electronics and Telecommunications
Politecnico di Torino, Italy
first.last@polito.it

***Abstract*—High altitude platform stations (HAPS) have been proposed to support terrestrial mobile networks, offering a sustainable alternative to network densification. With their wide coverage areas and green energy consumption model, HAPS super macro base stations (SMBSs) are well suited to handle the massive and dynamic mobile data traffic demand. This research introduces an adaptive traffic offloading strategy that leverages the capabilities of HAPS to support radio access network (RAN), particularly during periods of high network demand. To enable HAPS to effectively assist the RAN, it is crucial to accurately predict which base stations (BSs) will experience high loads. Precise forecasting of these demands is hence essential to ensure timely and targeted offloading of traffic to the HAPS when and where it is most needed. The proposed approach predicts and manages loads by considering temporal and geographical factors. At the core of this approach is the Q-learning update rule, which is continuously used to refine offloading decisions and flexibly adapt to changing conditions. Our simulation results demonstrate that the proposed HAPS offloading approach is effective in maintaining balanced loads in the terrestrial RAN during peak periods, by dynamically adapting to the typical traffic characteristics of different areas and to their evolution over time.**

***Index Terms*—5G and beyond networks, High Altitude Platform Stations (HAPS), Mobile traffic offloading, Q-learning, Load prediction.**

I. INTRODUCTION

The advent of fifth-generation (5G) technologies has led to a substantial increase in data demand, driven by applications such as gaming and streaming which require high data rates. Furthermore, the proliferation of mobile devices on the Internet of Things (IoT), such as unmanned aerial vehicles (UAV) and autonomous vehicles, contributes to the increase of network demands. To address this, mobile network operators (MNOs) have densified networks by installing more base stations (BSs), particularly in urban areas. However, network densification significantly increases capital and operational expenditures (CapEx and OpEx) and can lead to inefficient use of BSs, raising relevant sustainability concerns. Although various techniques have been introduced to improve network energy savings [1], the over-provisioning of BSs remains a challenge, especially in light of spatial and temporal changes in user demand.

The emergence of High Altitude Platform Stations (HAPSs) to host aerial BSs represents a promising solution as a complement to existing RANs [2, 3]. Furthermore, the operation of HAPS mounted BSs fully rely on renewable energy, hence the integration of HAPS has the potential to promote a more sustainable RAN operation in 5G and beyond scenarios. Finally, these aerial network nodes can effectively function as super macro base stations (SMBSs), offering additional services besides communication capabilities, like content caching and computation functionalities [4]. Various studies have evaluated the benefits of HAPS-SMBS for beyond 5G networks, particularly in offloading mobile traffic from underloaded terrestrial BSs to optimize the network energy efficiency by putting unneeded on-ground nodes into sleep mode [5, 6]. Furthermore, HAPSs yield significant benefits by providing extra capacity to terrestrial RANs, enabling them to flexibly handle sudden surges in mobile traffic demand. These surges, if not managed effectively, can lead to a decrease in the Quality of Experience (QoE) of end users. To this extent, the authors in [7] demonstrate the superior performance of HAPS-SMBS with respect to conventional RAN densification methods in handling peak demand without violating QoE requirements. Their analysis not only highlights the technical advantages of HAPS in handling traffic loads but also underscores the benefits in terms of sustainability derived from using HAPS over traditional methods.

Despite the promising potential of HAPS, effective traffic offloading strategies require accurate estimation of network and BS load dynamics. Relying on perfect knowledge of cell loads is unrealistic and poses a significant challenge for practical implementation. Indeed, without a precise estimation of the cell load, the benefits provided by HAPS integration cannot be fully realized, limiting the practical application of traffic offloading strategies. Various studies from the literature estimate the impact of errors introduced in the prediction of cell loads, confirming that inaccurate traffic forecasts may significantly impair the decisions made by the traffic management optimization algorithm [8].

Our paper aims at investigating the potential of HAPS to support the terrestrial RAN during peak demand periods, by offloading traffic to an HAPS mounted BS to relieve the terrestrial network. The primary objective is to decrease the load within the RAN, ensuring a more balanced and efficient network operation. To achieve this, we propose an adaptive traffic offloading strategy, that relies on an off-policy temporal difference control approach as in Q-learning. Unlike previous studies that assume perfect knowledge of network loads, this Reinforcement Learning (RL) approach estimates loads based on both temporal and geographical factors, leveraging the dynamics of traffic demand within specific time windows and days. These load estimates constitute the basis to take the proper offloading decisions to shift traffic to the HAPS. Furthermore, our approach flexibly adapts to temporal and spatial system variations, allowing to dynamically identify and relieve those overloaded areas that are more likely to benefit from HAPS offloading.

II. A CASE STUDY

We consider a portion of a RAN in a urban scenario, where an HAPS mounted BS provides additional capacity to offload a fraction of the terrestrial mobile traffic during peak demand, as depicted in Fig. 1. We devise an adaptive traffic offloading strategy, that is detailed in Section III, aiming at dynamically handling the RAN peaks and unburden the most overloaded areas of the terrestrial network during high load periods.

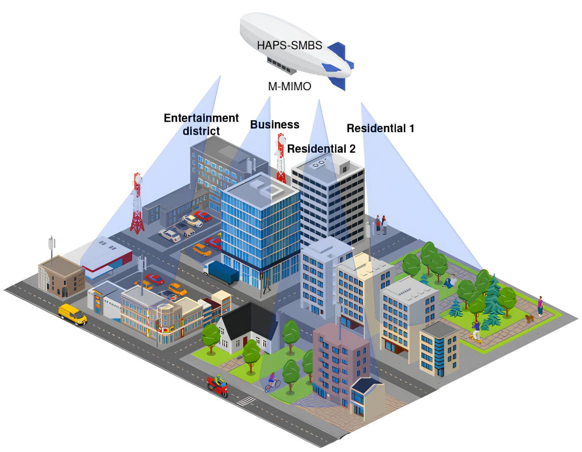

Fig. 1: HAPS-SMBS used to compliment RAN

Our simulation is based on real-world 4G mobile traffic data collected by an Italian mobile network operator (MNO) in Milan during 2015. The dataset includes traffic traces from N=112 BSs, and we select N_{areas}=16 areas of the city, each of which contains a cluster of seven BSs. These areas are chosen as representative samples of the heterogeneous zones that are typically found in an urban environment.

We selected three busy tourist areas: *Duomo*, *Theatre*, and *Touristic*. Additionally, we chose a *business* district, two *residential* areas with typical daily traffic patterns, and an *industrial* area. We included the *FS* train station and *Linate* airport, both busy at the start and end of the workday. Areas with variable traffic due to periodical events include *San Siro* (soccer stadium), *Rho Fiere* (exhibitions), and *Mediolanum Forum* (indoor sports arena). Finally, we selected the *Politecnico di Milano* area (*Polimi*), where a university campus is located, and three rural areas: *Agricultural Park*, *Highway*, and *Monza Park*.

These BS traffic traces report the values of traffic volume collected on an hourly basis, covering a two-month period. To account for the increasing mobile traffic demand observed in recent years, the N available traffic traces from the Italian MNO are scaled up based on various aggregated metrics derived from more recent traffic traces, that are collected from M=1460 BSs in an urban area in China in 2020. The adopted processing methodology, that results similar to the one presented in [6], is detailed hereafter.

First, each traffic profile from the Italian MNO is randomly paired with one of the M recent traffic profiles collected from just as many BSs owned by the Chinese MNO. Second, a set of N new traces are derived, scaling up the original N Italian traces so that each newly derived trace features the same peak and 5^{th} percentile traffic values as the considered paired BS. The shape of these newly derived traces remains similar to the original traffic traces from the Italian MNOs, whereas their peak and 5^{th} percentile are scaled to match the corresponding metrics derived from the most recent traffic profiles, making the traffic profiles considered for our investigation up to date, and thus more realistic.

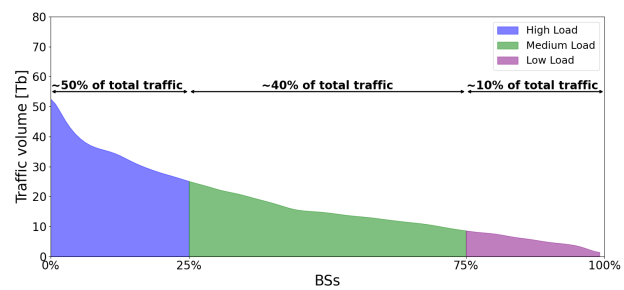

Fig. 2: Traffic distribution across BSs

The distribution of the total traffic volume over two months across all BSs is illustrated in Fig. II. The observed traffic distribution is consistent with typical urban environments, where a relatively small number of BSs handle the majority of traffic [9]. Specifically, 25% of BSs handle approximately 50% of the total traffic, while the remaining 75% of BSs manage the other 50%. Additionally, this graph proposes a classification of BSs into highly loaded (traffic below the 25^{th} percentile - blue color), medium loaded (traffic between the 25^{th} and the 75^{th} percentile - green color), and low loaded (traffic above the 75^{th} percentile - lilac color). The capacity of each BS is determined based on the observed maximum hourly traffic. For high load level BSs, the capacity is set equal to the maximum observed hourly traffic. For medium load level BSs, the capacity is defined assuming that the maximum observed hourly traffic represents 75% of the BS capacity. For low load level BSs, the capacity is assumed to be twice the maximum observed hourly traffic. Although the definition of BS capacity based on the levels of handled traffic volume per BS does not exactly match real data from the literature, our assumption results consistent with the trend observed in the traffic dataset from the Chinese MNO, in which the average BS load tends to increase as the average traffic volume handled by each BS becomes larger. Furthermore, this reasonable assumption allows to better evaluate how the proposed traffic off-loading strategy performs in scenarios characterized by varying traffic intensities.

III. Adaptive HAPS offloading

A key element of the proposed HAPS offloading strategy involves recognizing and utilizing the inherent temporal behaviors associated with different types of areas within the network. Our offloading approach, that is based on the Q-learning update rule, aims to dynamically predict the overloaded areas whose BSs require traffic offloading. Through the proposed methodology, we estimate which areas are expected to experience the highest load. In particular, at each time step, a predefined number of areas, that we denote N_{ca} with $N_{ca} \leq \frac{N_{\text{areas}}}{2}$, are selected to offload to HAPS their aggregated extra traffic that exceeds a given threshold. This selective offloading strategy ensures that HAPS resources are deployed only when significant traffic demands are observed in certain areas.

A crucial component of the proposed strategy involves quantifying the portion of traffic that exceeds a predefined threshold at each BS in a cluster, in order to dynamically identify the areas that are most suitable to offload their traffic to the HAPS. Let us denote C_{th} the operational threshold adopted to determine the overload condition for a BS. C_{th} is defined as a fraction of the BS capacity, that we denote C_{BS}. The offloading decision is performed at every time step t, with each time step being one hour. The volume of traffic to be offloaded from a given BS, denoted TO, at a given time t is hence computed based on how much the actual traffic volume handled by the BS, denoted $R_{\text{BS,t}}$, exceeds the capacity threshold C_{th}, as follows:

$$\text{TO}_{BS,t} = \max\left(R_{BS,t} - C_{th} \cdot C_{BS}, 0\right) \quad (1)$$

Finally, to perform the proposed HAPS offloading strategy, the traffic dataset is split into two periods. Initially, the traffic from the first month is analyzed to identify patterns of peak traffic, which initialize the state-action values, i.e., the Q values in our RL approach, and form the basis to make offloading decisions. These initial values help us to understand the recurring peak traffic patterns and their intensity in different areas. Conversely, traffic time series corresponding to the second month are used to test the offloading strategy, with the initialized values from the first month guiding the decisions. The system continuously updates the state-action values through Q-learning updates after each time step, based on real-time traffic observations during the second month, allowing for adaptive and responsive decision-making. Further details about the HAPS offloading strategy operation are provided hereafter.

A. Incorporating Area-Specific Temporal Behaviors in Traffic Offloading

Focusing on the temporal and spatial distribution of traffic peaks enhances the understanding of peak severity and frequency across different times and areas, allowing to tackle the dynamic nature of network traffic when HAPS offloading is performed. Precise definitions of states and actions are essential to capture the system dynamics and effectively guide the system's response to varying traffic conditions under HAPS offloading. To capture the varying patterns of network usage, the *state* of the system is defined by the current time window, described by two state variables that represent the period of the day and the daytype.

Let P denote the set of periods of the day:

$$P = \{\text{Morning}, \text{Night}\} \quad (2)$$

where *Morning* is the day period spanning from 8 am to 3 pm, that covers the bulk of daytime activities including the start of the business day and early afternoon, whereas *Night* is the day period running from 4 pm to 11 pm, that includes the evening rush hour and extends until late evening.

Let D denote the set of days of the week:

$$D = \{\text{Monday}, \ldots, \text{Saturday}, \text{Sunday}\}. \quad (3)$$

The state $s(t)$ encodes the day d and period p for each area i. Therefore, the state vector can be written as:

$$s(t) = (W_{d,p}(t) \mid d \in D, p \in P) \quad (4)$$

where $W_{d,p}(t)$ represents the period p of day d at the given time t. These time windows are applied consistently on each day of the week, allowing the recognition of daily traffic cycles. Importantly, the system operates with an hourly granularity, meaning that offloading decisions are made on an hourly basis based on the estimated severity of traffic peaks in specific states/time windows.

Specifically, an *action* involves selecting an area for offloading, which is crucial as it directs excess traffic away from high loaded areas to HAPS. In this context, A denote the set of possible actions on areas $\{a_1, \ldots, a_{16}\}$, that can be selected for traffic offloading.

B. Initialization of State-Action Values for Traffic Offloading

To build our understanding of high load conditions, which will inform our future offloading decisions, we consider traffic time series for each area during the first month. This initial dataset is critical as it provides the fundamental understanding of the traffic patterns needed to configure the system. The traffic data is normalized using the Min-Max normalization technique. This normalization adjusts the traffic data to a scale between 0 and 1, facilitating a uniform analysis across different scales of traffic data. Crucially, the capacity of each BS is included in this normalization process to ensure that the analysis accounts for the maximum potential traffic that each area can handle. This inclusion is essential to determine how peak traffic relates to capacity limits, which is a key factor in assessing peak severity.

From these normalized data, a temporally dynamic set of state-action values, corresponding to the values $Q(s,t)$, is created to initially estimate the probability of peak severity within each area. These values represent the expected cumulative reward and serve as the baseline from which the system makes offloading decisions. In our case, the rewards are the expected loads exceeding the predefined capacity threshold C_{th}. Let us introduce the concept of Peak Severity, that we denote PS.

For a period of the day p and BS i belonging to area j, the Peak Severity is defined as follows:

$$PS_{i,p} = \frac{1}{n}\int_{t_a}^{t_b} f(t) - g(t)\,dt \tag{5}$$

where t_a and t_b represent the time limits of a period of a day p (for example, in the morning, from 8 am to 3 pm). Furthermore, $f(t)$ is the linear spline interpolation of the original traffic time series R_{BS}, whereas $g(t)$ is the interpolation of the time series which represents the minimum between $R_{\text{BS,t}}$ and the traffic volume corresponding to the capacity threshold C_{th} (i.e., $C_{th} \cdot C_{\text{BS}}$) in period p. Finally, n is the number of time instant within a period p (i.e., $n = 8$).

The values $Q(s(t), a)$ are calculated by summing the Peak Severity (PS) across all relevant BSs of an area during the morning or at night for a specific day d (e.g., Monday-Morning) across all occurrences $D_{d,m}$ of that day throughout the first month m.

$$Q(s(t), a) = \sum_{i\in A_j}\sum_{d\in D_{d,m}} PS_{i,p} \tag{6}$$

where A_j represents the set of BSs in area j. After computing all values for every area/action and state/time window, the State-Action Value table (Q-table) is initialized. This Q-table contains the initial knowledge based on prior observations, as described above, and is now ready for the second phase of fine-tuning and decision-making.

C. Dynamic adaptation through continual learning

A continual learning approach is essential due to inherent variability in network traffic, particularly during peak periods. By iteratively learning from real-time data after taking an action, the system dynamically adjusts to the observed variations. This process is performed through the following steps:

1) **Action Execution**: The system selects and executes an action, such as offloading traffic from N_{ca} areas to HAPS, based on the current state of the network, i.e., time window and traffic conditions.
2) **RAN Observation**: The system observes the changes in traffic load and the immediate impact on the network of the performed action.
3) **State-Action Value Update**: Using the observed data, the system updates the value estimates for the state-action pairs. This update is done using Q-learning, which adjusts the values based on the difference between expected and actual outcomes.
4) **Go back to step 1**: The process repeats continuously, using new data and insights to better adapt to future traffic trends.

Upon getting the state-action values for every action, N_{ca} actions with the highest values are chosen:

$$\{\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{N_{ca}}\} \tag{7}$$

where $\{\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{16}\}$ is a permutation of $\{a_1, a_2, \ldots, a_{16}\}$ such that $Q(s(t), \hat{a}_i)$ is sorted in descending order. State-Action Value updates occur on an hourly basis after executing an action and observing the resulting traffic in the RAN. This process follows a Q-learning update rule [10]. Following the execution of an action, the values corresponding to the current state and all actions are updated based on the observed traffic in each area. This update ensures that the value estimates remain responsive to recent traffic patterns. The hourly update formula is:

$$Q(s(t), a) = Q(s(t), a) + \alpha\,(\text{HPS} + \gamma\,(\text{HPS} + Q(s(t), a)) - Q(s(t), a)) \tag{8}$$

where α is the learning rate and γ is the discount rate that adjusts the influence of future updates on the current value. The Hourly Peak Severity, denoted HPS, is calculated using the integral from the earlier Eq. 5, with t_a representing the time sample before the current time and t_b representing the current time.

IV. Performance analysis

We now present our performance analysis, which is based on the results of our simulation throughout the second month. The initial investigation focuses on investigating how offloading traffic to HAPS affects the load distribution of the terrestrial network. This analysis is crucial to understand how effective is the offloading mechanism. Subsequently, the investigation focuses on evaluating the utilization of HAPS capacity.

In our study, we assume that the coverage over the considered areas is provided by a beam cell generated by a 5G enabled BS installed on the HAPS, offering a capacity of *120 Mbps* [6]. Notice that the parameter configuration is crucial for an effective performance of the traffic offloading algorithm. The key parameters include the offloading threshold C_{th}, the number of selected areas for traffic offloading N_{ca}, the discount rate γ, and the learning rate α. In our study, C_{th} is set equal to 0.6. The choice of N_{ca} aims to demonstrate how the algorithm prioritizes areas for offloading to HAPS when multiple areas are highly loaded. Selecting this number is critical because if N_{ca} is too small, the offloading would not be effective. Conversely, if it is too large, resources might be unnecessarily used for areas that do not have high loads, leading to inefficiency. After extensive experimentation, the following settings are adopted for the various parameters: N_{ca}=6, γ = 0.95, α=0.7.

A. Impact On Load Distribution Across RAN

The distribution of the load density in different areas, illustrated in Fig. 4, provides valuable insights into the effectiveness of offloading traffic to HAPS. The blue curves represent the load distribution without HAPS offloading, while the red curves show the distribution when HAPS offloading is active. Since various areas share similar load profiles, only representative areas are shown. Notably, areas such as *Theatre*, *FS*, and *Agricultural Park* show a significant shift in load density towards 0.6 after offloading, frequently experiencing high loads up to 0.9. This explains their prioritization for offloading. In contrast, *Residential 1* and *Highway* show minimal changes, with traffic consistently below 0.6, indicating stable traffic levels. *San Siro* and *Rho*, typically stable with loads under 0.6, experience sudden increases during large events, with loads reaching up to 0.9 for some BSs. The increased density around

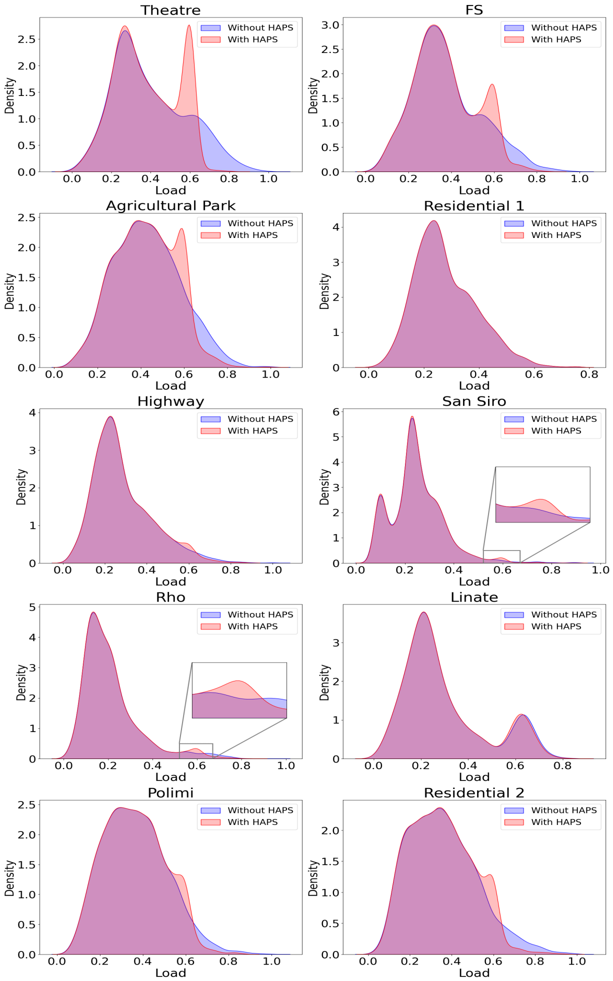

Fig. 3: Load density distribution across different areas before and after offloading to HAPS.

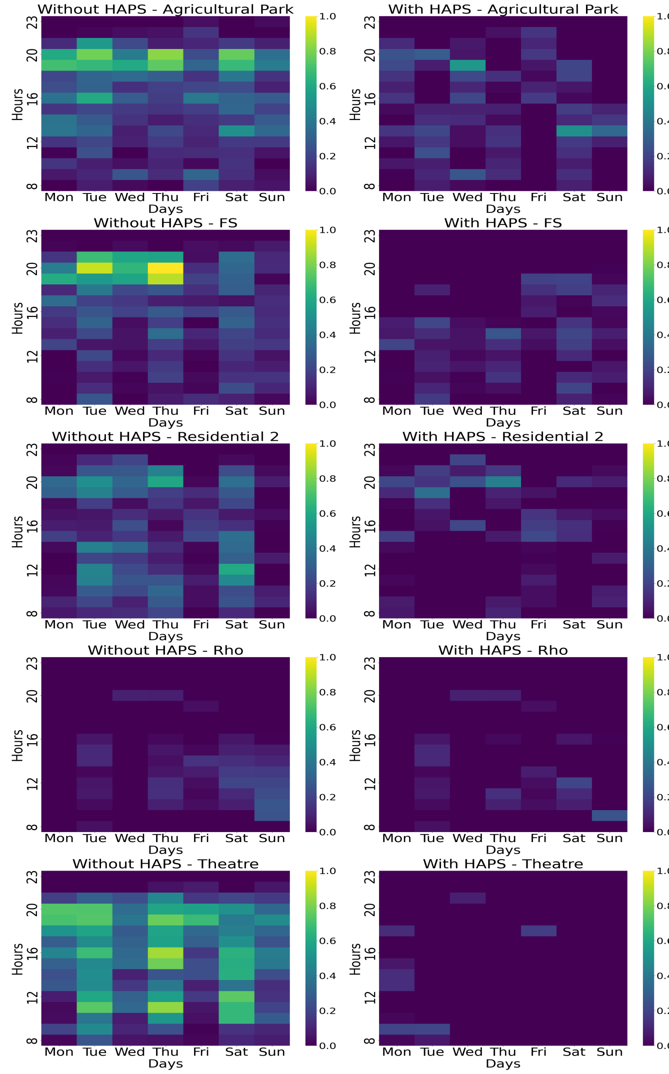

Fig. 4: Values of the portion of load exceeding the capacity threshold, C_{th}, aggregated over the second month for each hour (y-axis) and day of the week (x-axis) and across different areas (one per each pair of plots in a row), before (left column plots) and after (right column plots) offloading traffic to HAPS.

0.6 after offloading indicates successful management of these spikes. *Linate* shows two load ranges: 0.1 to 0.3 and 0.6 to 0.8. Six BSs have very low loads, while one BS consistently has traffic in the 0.6 to 0.8 range, resulting in many hours where traffic exceeds the threshold but rarely goes above 0.8. *Polimi* and *Residential 2* have a uniform load distribution from 0.2 to 0.6, with some instances exceeding the 0.6 threshold. They are selected for offloading when high loads are estimated but are not prioritized as highly as areas like *Theatre*. As a result, they show a small increase around the 0.6 load but not a significant change like other high-loaded areas.

Investigating the hourly and daily impact of HAPS offloading is crucial to understanding traffic distribution and methodological adaptation. To this aim, the heat maps reported in Fig. 4 depict the values of the portion of load exceeding the capacity threshold, C_{th}, aggregated over the second month for each hour (y-axis) and day of the week (x-axis), and across different areas of the RAN (one per each pair of plots in a row), before (left column plots) and after (right column plots) offloading traffic to HAPS. Normalizing the excess load density using Min-Max scaling across all heat maps and areas enables a comparative analysis of load dynamics. Areas with the highest loads, such as *Theatre*, exhibit the most significant reduction in load density after offloading. These areas are characterized by high loads almost every day, resulting in these areas being prioritized for offloading on nearly all days of the week. In contrast, areas like *Agricultural Park*, *Residential 2*, and *FS*, feature less frequent high-load hours with respect to *Theatre*. However, the system successfully identifies these times when traffic is extremely high and effectively offloads the exceeding traffic to the HAPS. For *Agricultural Park*, while not all traffic over the threshold is successfully offloaded, the most significant load surges are correctly identified and effectively offloaded to HAPS. In *Rho*, typically featuring stable traffic, a significant spike is observed on Saturday and Sunday mornings during the second week. However, heat maps show that traffic is not offloaded on Sunday at 9 am and on Saturday at midday,

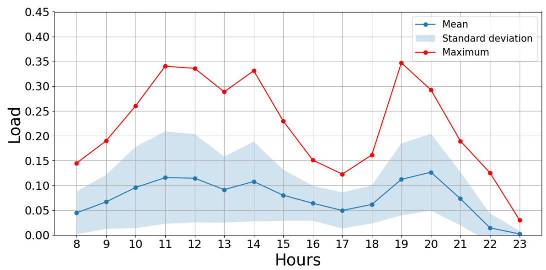

Fig. 5: Daily profile of hourly HAPS capacity utilization.

since the system is not capable to anticipate the spike, as this behavior has not been previously observed. Nevertheless, hourly updates allow the system to quickly adjust, updating the State-Action Value table to improve offloading decisions in the following hours. This reveals a possible limitation of the proposed approach, since the method might miss sudden traffic spikes in case they occur unexpectedly and involve BSs that are usually characterized by stable loads. Nonetheless, considering that only a subset of areas are selected to be offloaded to the HAPS, this phenomenon can be compensated. In fact, even when the detection of areas featuring unexpected traffic surges is missed, still other highly loaded areas can be prioritized to perform traffic offloading, hence allowing to relieve the terrestrial RAN on the whole.

B. HAPS Capacity Utilization

We now evaluate the utilization of the available HAPS capacity under traffic offloading strategy operation. Fig.5 reports the hourly profile of the HAPS-SMBS load during the day, showing its average and maximum values, and the standard deviation, hence providing insight into the hours in which HAPS experiences the highest levels of capacity utilization. On average, the HAPS capacity is significantly underutilized. However, observing the maximum load registered during the day, the highest utilization occurs between 10 am and 2 pm and from 7 pm to 8 pm, when the capacity utilization exceeds 25%, reaching a peak of about 35%. These findings highlight the effectiveness of the proposed approach in providing proper support to relieve the terrestrial RAN in the covered areas right when it is most needed due to peak demand. At the same time, we observe wide margin to further expand the HAPS coverage, exploiting its potential to effectively serve additional traffic areas.

V. CONCLUSION

This paper presents an adaptive traffic offloading strategy exploiting a self-sustainable HAPS-SMBS to support the terrestrial RAN during high mobile traffic demand. Employing a Q-learning approach, the system predicts and manages BS loads, dynamically offloading peak traffic from the most loaded areas to HAPS. The methodology exploits temporal and geographical features of traffic to estimate future loads and make informed offloading decisions for the next time slot. Our simulation results prove the effectiveness of the proposed approach in maintaining balanced loads in the terrestrial RAN across different areas during peak periods, by dynamically adjusting offloading decisions to varying traffic conditions and prioritizing areas that can most benefit from HAPS offloading based on the peak traffic timing.

In this study a fixed capacity threshold ($C_{th} = 0.6$) is uniformly adopted across all BSs to identify peaks. Future work will explore the implementation of offloading strategies based on dynamic thresholds, tailored to each BS cluster based on the varying traffic demand. Furthermore, we aim to investigate the potential of dynamically changing the HAPS position across the covered area to enhance Line of Sight (LoS) conditions, leading to better channel quality and, consequently, improved overall performance.

ACKNOWLEDGMENT

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART", focused project R4R and project 6GWINET).

REFERENCES

[1] D. Lopez-Perez et al. "A survey on 5G radio access network energy efficiency: Massive MIMO, lean carrier design, sleep modes, and machine learning". In: *IEEE Communications Surveys and Tutorials* 24.1 (2022), pp. 653–697.

[2] G. K. Kurt et al. "A vision and framework for the high altitude platform station (HAPS) networks of the future". In: *IEEE Communications Surveys and Tutorials* 23.2 (2021), pp. 729–779.

[3] W. Jaafar and H. Yanikomeroglu. "HAPS-ITS: enabling future ITS services in trans-continental highways". In: *IEEE Communications Magazine* 60.10 (2022), pp. 80–86.

[4] M. S. Alam et al. "High altitude platform station based super macro base station constellations". In: *IEEE Communications Magazine* 59.9 (2021), pp. 103–109.

[5] D. Renga and M. Meo. "Can high altitude platforms make 6G sustainable?" In: *IEEE Communications Magazine* 60.9 (2022), pp. 75–80.

[6] T. Song et al. "High altitude platform stations: the new network energy efficiency enabler in the 6G era". In: *arXiv preprint arXiv:2307.00969* (2023).

[7] Cihan Emre Kement et al. "Sustaining Dynamic Traffic in Dense Urban Areas with High Altitude Platform Stations (HAPS)". In: *arXiv preprint arXiv:2209.05127v2* (2023).

[8] Berk Ciloglu et al. "Cell Switching in HAPS-Aided Networking: How the Obscurity of Traffic Loads Affects the Decision". In: *arXiv preprint arXiv:2405.00387v1* (2024).

[9] *Ericsson Mobility Report*. Technical Report. Ericsson, Nov. 2022.

[10] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. second. The MIT Press, 2018.