telecom

MDPI

Article

An Unsupervised Machine Learning Approach for UAV-Aided Offloading of 5G Cellular Networks

Lefteris Tsipi, Michail Karavolos and Demosthenes Vouyioukas *

Department of Information and Communication Systems Engineering, School of Engineering, University of the Aegean, 83200 Samos, Greece; ltsipis@aegean.gr (L.T.); mkaravolos@aegean.gr (M.K.)
* Correspondence: dvouyiou@aegean.gr

Abstract: Today's terrestrial cellular communications networks face difficulties in serving coexisting users and devices due to the enormous demands of mass connectivity. Further, natural disasters and unexpected events lead to an unpredictable amount of data traffic, thus causing congestion to the network. In such cases, the addition of on-demand network entities, such as fixed or aerial base stations, has been proposed as a viable solution for managing high data traffic and offloading the existing terrestrial infrastructure. This paper presents an unmanned aerial vehicles (UAVs) aided offloading strategy of the terrestrial network, utilizing an unsupervised machine learning method for the best placement of UAVs in sites with high data traffic. The proposed scheme forms clusters of users located in the affected area using the k-medoid algorithm. Followingly, based on the number of available UAVs, a cluster selection scheme is employed to select the available UAVs that will be deployed to achieve maximum offloading in the system. Comparisons with traditional offloading strategies integrating terrestrial picocells and other UAV-aided schemes show that significant offloading, throughput, spectral efficiency, and sum rate gains can be harvested through the proposed method under a varying number of UAVs.

Keywords: 5G; machine learning; UAV placement; offloading

Citation: Tsipi, L.; Karavolos, M.; Vouyioukas, D. An Unsupervised Machine Learning Approach for UAV-Aided Offloading of 5G Cellular Networks. *Telecom* **2022**, *3*, 86–102. https://doi.org/10.3390/telecom3010005

Academic Editor: Michał Aibin

Received: 6 December 2021
Accepted: 17 January 2022
Published: 20 January 2022

1. Introduction

The new generation of mobile communications networks, e.g., 5G and beyond, along with the abundance of technologies and the support of high transmission rates and reliability, bring a massive increase in mobile devices, sensors, and services ranging from high-throughput multimedia applications to ultra-reliable Internet of Things (IoT). Hence, mobile operators should consider exploiting novel and efficient offloading schemes to alleviate the stress on cellular networks. In this context, the employment of unmanned aerial vehicles (UAVs) as flying base stations, is expected to be a promising solution offering several degrees of freedom and flexibility in the network [1,2].

Traditionally, fixed cells, e.g., picocells and femtocells, are employed to offload the existing terrestrial cellular networks. This approach requires design and analysis to efficiently integrate the fixed small cells in high data traffic demand areas. However, this strategy significantly restricts the degrees of freedom of the network and cannot cope with future on-demand changes in other regions covered by the terrestrial infrastructure [3,4].

In addition, in case of user hot spots formed for a limited period, e.g., sports events and festivals, offloading the terrestrial network would require installing a fixed terrestrial infrastructure, e.g., picocells. Moreover, following this approach, various small base stations will probably be inactive in the long term due to the continuous mobility of the users' hot spots. In other words, if future demand increases in another region within the coverage area provided by the terrestrial infrastructure, it will require re-analysis, design, and installation of fixed cells. Hence, this solution is both cost and time ineffective for the operators. Preferably, it would be better to utilize a network entity capable of moving from region

to region on-demand and efficiently offloading the terrestrial network; therefore, various studies have been proposed in the literature, using UAVs as flying base stations, which, unlike the ground-fixed picocells, can on-demand move to the area of interest and offload the terrestrial network [5].

1.1. Background

The employment of fixed small cells to enhance cellular networks' performance through offloading has been studied in the past. O. Chabbouh et al. [6], have presented an energy-efficient offload decision process for macro base stations using fixed small cells. The proposed technique can provide an adequate user quality of experience by reducing the application response time. Furthermore, simulation results have shown that the proposed algorithm assures the computation of all the applications, subject to both latency and battery lifetime constraints. Moreover, L. Xu et al. [7] have proposed a self-optimized joint traffic offloading (JTO) scheme to offload traffic between picocells and macrocells and achieve mobility load balancing among macrocells. Through computer simulations, it has been proved that the JTO scheme can effectively deal with the cell exacerbation problem regarding handover failure and call dropping.

As an alternative way to offload cellular networks, the exploitation of UAV technology where UAVs act as flying base stations has stimulated many researchers. The work in [8] deals with the 3D UAVs deployment for on-demand offloading in an area where a damaged or overloaded base station operates. To determine the optimal UAVs 3D locations, the authors have proposed the k-means algorithm associated with the pattern search technique to cover the users experiencing a service outage and maximize the operator's profit. The computer simulation results clarified that both operators' profit and load balancing are highly achieved. Furthermore, Y. Qin et al. [9] have studied the performance of a UAV-enabled cellular network, consisting of limited battery drones that stay in operation for a limited time and then must fly back to a dedicated charging station. The authors used stochastic geometry tools to derive the probability of UAV-enabled cellular network coverage as a function of battery size, charging station density, and time required to recharge the battery. Thereinafter in [10], the authors have investigated the spectrum trading problem in the case of a network where UAVs are temporarily deployed for data offloading in regions where macro base stations cover. Considering both the selfish macro base station (MBS) manager and the selfish UAV operators, the authors have modeled the utilities and the costs of spectrum trading and formulated the problem of designing the optimal contract for the MBS manager, utilizing the contract theory. More specifically, they derived the optimal pricing strategy based on fixed bandwidth assignment and proposed a dynamic programming algorithm to calculate the optimal bandwidth with reduced complexity. Computer simulation results revealed that a selfish MBS manager sells less bandwidth to the UAV operators. Lastly, J. Lyu et al. [11] have proposed a hybrid network architecture consisting of UAVs and ground base stations to maximize the minimum throughput of all mobile terminals (MTs) by jointly optimizing the UAVs' trajectory, bandwidth allocation, and user partitioning. The UAVs mainly act as an aerial mobile base station, which flies cyclically along the cell edge to offload data traffic for cell-edge users. Numerical results show that the proposed hybrid network with optimized spectrum sharing and cyclical multiple access designs significantly improves the spatial throughput over the conventional ground base station (GBS) network. Further, performance evaluations showed that the proposed UAV offloading scheme outperforms the traditional small cell offloading strategies in terms of throughput.

UAVs as flying base stations have been considered as the best candidate for offloading the new generation terrestrial networks; however, there are some questions on how many UAVs should be deployed in the affected area and in which location. Utilizing machine-learning-based methods can answer these questions and determine the best possible placement of the UAV base stations (UAV-BSs) to enhance the traffic offloading of the existing macro base station network. The deployment process of UAVs was studied in the

past utilizing unsupervised machine learning methods. More specifically, J. He et al. [12] have proposed an unsupervised machine learning method for emergency drone deployment. The authors utilized a modified k-means clustering algorithm which identifies the number of drones that should be deployed in the area of interest and determines the height and the minimum transmit power of them. The proposed scheme aims to minimize the total transmit power of drones with all users' coverage while maintaining their rate requirements, with constraints on drones' coverage area, capacity, and limited power. Computer simulation results revealed that all users can be served through the proposed technique and the minimum transmit power of drones is ensured. In another approach [13], a variation of the k-means algorithm that considers users' bandwidth requirements is presented for clustering them and determining the total number of UAVs that should be deployed based on the bandwidth that each UAV can offer. Performance evaluation results have shown that, through the proposed scheme, both the number of UAVs and the maximum deployment delay can be reduced. In [14], the application of simultaneous wireless information and power transfer (SWIPT) to millimeter-wave non-orthogonal multiple access (mmWave-NOMA) based aerial networks is studied and, the k-means and k-medoids algorithms are used to group the users into clusters for exploiting the full advantages of NOMA. The simulation results revealed that the proposed unsupervised learning-based clustering framework for mmWave-NOMA enabled aerial SWIPT networks can achieve considerable improvements in terms of the harvested energy compared to conventional aerial SWIPT networks. Moreover, Qi et al. [15], have presented a user-based modified k-means clustering algorithm to solve the offloading issues in heterogeneous networks (HetNets). Simulation results showed that the proposed method could increase the offloading factor from 15% to the required 50%. D. Mandloi et al. [16], have introduced a machine-programming-based approach for building a 5G-enabled UAV-BSs network in mmWave. The unsupervised machine learning (ML) methods such as the k-means, the k-medoids, and the fuzzy cluster means (FCM) algorithm were proposed to determine the optimal 2D placement of UAV-BSs. The three different ML methods were compared in terms of received power, signal-to-interference-plus-noise ratio (SINR), and path loss per active user equipment (UE). Computer simulated results showed that the k-means-based deployment approach outperforms the other two methods when the UEs are uniformly distributed in the region of interest. Finally, F. Tang et al. [17], have presented a novel distributed reinforcement-learning-based traffic offloading system that utilizes a novel network state information gathering protocol, for usage in space–air–ground integrated (SAGIN) networks. Through computer simulations, it has been proved that the proposed method outperforms conventional offloading schemes regarding the signaling overhead, dynamic adaptivity, packet drop rate, and transmission delay.

1.2. Contributions

In this paper, we consider a two-tier 5G heterogeneous cloud radio access network (H-CRAN) [18], which is configured to cope with events of exponential increase in data traffic. The first tier consists of macrocells providing coverage to the whole region of interest, while UAV-BSs form the second tier to maximize macrocells' traffic offloading. The primary objective of this work is to identify the minimum number of UAV-BSs that should be deployed, as well as their best placement within the coverage area of the macrocells, aiming to maximize traffic offloading. Towards this end, we partition the users in the affected area into clusters utilizing the k-medoid algorithm [19]. Then, we identify the number of UAV-BSs that should be deployed, considering both the number of the formed clusters and the maximum number of users that a UAV-BS can serve. Given that the number of the available UAV-BSs may be less than the minimum UAV-BSs required to cover the formed clusters, we also propose a cluster selection process scheme. The proposed cluster selection technique is formulated as an optimization problem, which its objective function is to maximize the traffic offload subject to the constraint that the deployed number of UAV-BSs should be less or equal to the available one. The proposed cluster selection scheme will provide the

clusters that the available UAV-BSs should be deployed to achieve the maximum traffic offload. Furthermore, computer simulations are carried out to evaluate the performance of the proposed scheme in terms of percentage offloading, average received signal strength, throughput, spectral efficiency, sum rate and the effectiveness of localization inaccuracy. Further, we compare the proposed system with a two-tier 5G H-CRAN consisting of either random or planned picocells and other unsupervised machine-learning-based methods for UAV-BS best placement. More specifically, the following contributions are provided:

- A novel UAV-aided offloading framework is proposed, in the context of an overloaded terrestrial 5G network. The offloading procedure can be regarded as a cluster formulation problem that can be dealt with unsupervised machine learning-based methods. To the best of the authors' knowledge, this is the first time that the k-medoid algorithm is utilized in this context.
- A clustering selection scheme is proposed, which is formulated as an optimization problem to further enhance the offloading procedure under a limited number of available UAV-BSs.
- The proposed scheme mainly increases the offloading percentage as well as the spectral efficiency and, at the same time, improves the received signal strength, as compared to random or planned picocells deployment strategies and the k-means algorithm for the cluster formulation problem.
- The impact of the increased localization inaccuracy of the UEs on the proposed framework is evaluated. Also, comparisons concerning the performance of the proposed method with another state-of-the-art unsupervised machine learning method are illustrated and discussed.

1.3. Structure

The rest of this paper is organized as follows. In Section 2 the considered system model is described, while the proposed scheme is presented in Section 3. Performance evaluation results are given in Section 4, followed by conclusions and future directions in Section 5.

2. System Model

As depicted in Figure 1, a typical 5G two-tier heterogeneous cloud radio access network (H-CRAN) is configured in the event of an exponential increase in data traffic. First, as it can be observed in Figure 1a, we consider an area of interest W covered by M macro base stations (MBSs) with their corresponding coordinates $B = \{b_1, b_2, ..., b_M\} \in W$, where $b_i = (x_i^b, y_i^b)$. Due to some frequent events (sports, cultural festivals, etc.), hotspots of users with exponential demand of data traffic are established within the coverage area of these MBSs. The total number of users in the region of interest is N with coordinates $U = \{u_1, u_2, ..., u_M\} \in W$, where $u_i = (x_i^u, y_i^u)$; therefore, in order to maximize the offloading of the existing network for the duration of the event, it is planned to develop a set of unmanned aerial vehicle base stations (UAV-BSs) as wireless access points to further assist the existing terrestrial communications, as depicted in Figure 1b.

From system's point of view, each MBS and UAV-BS consists of a remote radio head (RRH), connected to the baseband unit (BBU) pool through optical fiber link and free-space optical or microwave fronthaul links, respectively [20,21]. Each BBU controls one or multiple RRHs, under the restriction of the maximum data volume limit that a BBU can handle. The virtualization of BBUs in a centralized cloud architecture provides increased flexibility in network upgrades and adaptability to non-uniform traffic. Moreover, direct communication between BBUs is enabled due to the virtualization of BBUs in a centralized cloud architecture. All the MBSs are equipped with an antenna with transmitting gain G_t^m, operating frequency f_m and maximum transmit power P_m. Additionally, we assume that all the UAV-BSs are equipped with an antenna with transmitting gain G_t^d, operating frequency $f_d \neq f_m$, and maximum transmit power P_d. Furthermore, the UEs are equipped with a single antenna with reception gain denoted as G_r.

The UEs can be served only from the MBSs or equally from the MBSs and the UAV-BSs. Before the start of an event, there is not a lot of user traffic, so the demand for services from the users does not burden the existing terrestrial infrastructure; therefore, all UEs can be served from the MBSs, as shown in Figure 1a. On the contrary, in the case of an unexpected event, which increases the mobility in the region of interest and grow user hotspots, the deployment of UAV-BSs over the existing infrastructure is needed to offload the traffic of the existing cellular system. In this case, the UEs can be served from the MBSs or the UAV-BSs, as illustrated in Figure 1b.

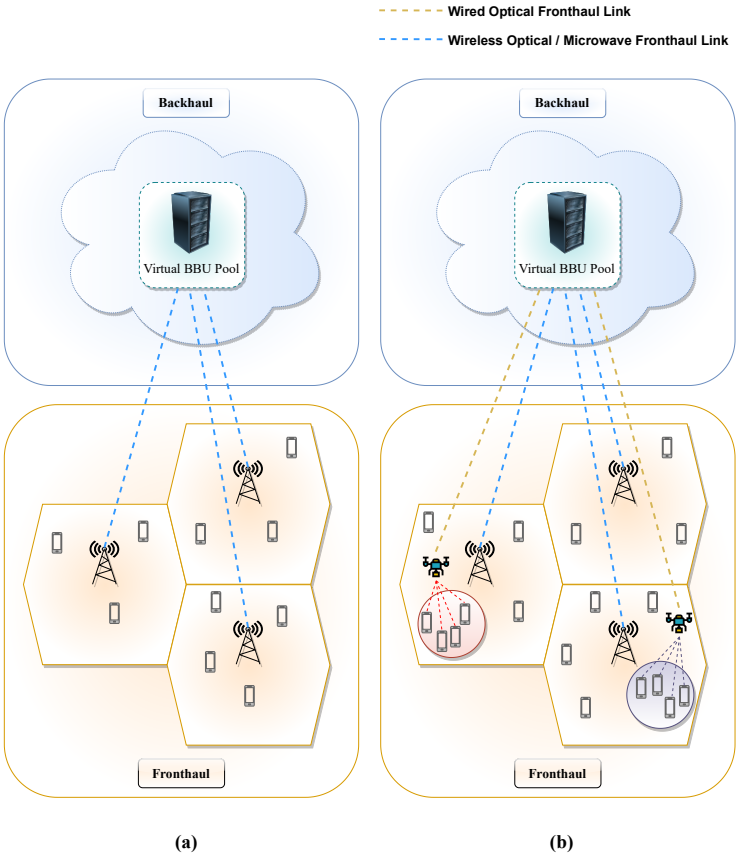

Figure 1. The configuration of the considered 5G two-tier heterogeneous cloud radio access network (H-CRAN): (**a**) Operating under normal traffic load and before the placement of UAV-BSs. (**b**) A heavily loaded scenario caused by an unexpected event and network offloading via the deployment of UAV-BSs.

The deployment of the UAV-BSs is controlled by the UAV-BS deployment process (UDP) executed at the cloud. UDP intelligently integrates the UAV network into the existing macro network to offload the terrestrial infrastructure and improve the quality of service (QoS) of the UEs in the temporary hotspots. Due to the unobstructed vantage position of the UAV-BS over the air, we assume that each UAV-BS maintains its height to ensure line of sight (LOS) condition with the terrestrial UEs; therefore, the challenge that the UDP scheme has to face is the optimal placement of the UAV-BSs under the limitation of the available number of UAV-BSs.

3. UAV-BS Deployment Process (UDP)

The UDP scheme targets to maximize the number of offloaded UEs from the MBSs to the UAV-BSs and it consists of two subprocesses. The first subprocess is named as user clustering process (UCP) and partitions the N UEs into K subsets. Through UCP, the maximum number of required UAV-BSs, as well as their corresponding positions in the 2D plane, are derived. Subsequently, the second subprocess is called the clustering

selection process (CSP), which determines the $D \leq K$ subsets to be offloaded from the M MBSs into the D available UAV-BSs.

3.1. User Clustering Process (UCP)

The UCP subprocess utilizes the k-medoids machine learning algorithm [19,22]. The aim of the k-medoid method is to partition N UEs into K ($1 \leq K \leq N$) disjoint subsets $S = \{S_1, S_2, ..., S_K\}$, such that the within-cluster squared error is minimized. The total within-cluster dissimilarity can be expressed as follows:

$$A = \sum_{k=1}^{K} \sum_{u_i \in S_k \cap U} \|u_i - \mu_k\|^2, \tag{1}$$

where μ_k is the candidate medoid of the cluster S_k. The objective function of the k-medoid algorithm is to minimize (1) across the K clusters and it can be expressed as follows:

$$\min_{S} A = \sum_{k=1}^{K} \sum_{u_i \in S_k \cap U} \|u_i - \mu_k\|^2, \tag{2}$$

The maximum number of the formed clusters K is given by the ceiling function of $K = \left\lceil \frac{N}{C} \right\rceil$, where C represents the maximum number of UEs that a single UAV-BS can serve simultaneously.

The operation of the UCP subprocess is summarized in Algorithm 1.

Algorithm 1 User Clustering Process (UCP).

input: Total number of UEs N, Maximum number of UEs that a single UAV-BS can serve C, and the set of coordinates of UEs U
Convergence tolerance $\epsilon = 10^{-6}$
Calculate the total number of clusters $K = \left\lceil \frac{N}{C} \right\rceil$
Randomly choose initial medoids $\mu_1, \ldots, \mu_K \in U$ for the clusters $S_1, \ldots, S_K \in S$
for $i \leftarrow 1$ to N **do**
  Find the closest medoid μ_c to user u_i via the expression: $c = \arg\min_{j=1...K} \|u_i - \mu_j\|$
  Allocate u_i to the cluster S_c with medoid μ_c
end for
Calculate the total within-cluster dissimilarity using (1) and store it to $A^{[1]}$
repeat
  $A^{[0]} = A^{[1]}$
  for $k \leftarrow 1$ to K **do**
    for $i \leftarrow 1$ to N **do**
      Swap the role of μ_k and u_i
      Allocate each $u_j \in U$ to the cluster with the closest medoid under this new configuration
      Calculate the total within-cluster dissimilarity using (1) and store it to B
      if $B < A^{[1]}$ **then**
        Keep the swap
        $A^{[1]} = B$
      end if
    end for
  end for
until $A^{[0]} - A^{[1]} < \epsilon$
output: The formed clusters $S_1, \ldots, S_K \in S$ along with the corresponding medoids $\mu_1, \ldots, \mu_K \in U$, where the UAV-BSs should be deployed

3.2. Clustering Selection Process (CSP)

At this point, the exact K candidate 2D locations for the placement of UAV-BSs have been identified by the UCP subprocess. The second subprocess of the UDP is called the cluster selection process (CSP). CSP determines which of the D out of K subsets of UEs should be offloaded from the macro-cell infrastructure onto the D available UAV-BSs; therefore, for the best selection of $\binom{K}{D}$, the following weighted score (WS) function is considered:

$$WS_k = an_k + (1-a)d_k \tag{3}$$

where d_k is the distance between the medoid of cluster k and the nearest macro cell, n_k is the number of users belonging to cluster k, and a ($0 \leq a \leq 1$) is the weighting factor [23]. Since d_k and n_k are disparate and in most cases d_k is quite larger than n_k, their corresponding values considered for the WS function are normalized using the following expression:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}, \tag{4}$$

where X is a value of the corresponding feature under normalization, X_{max} and X_{min} are the maximum and the minimum value of this feature, respectively, and $X_{norm} \in [0,1]$ is the final normalized value [24]; therefore, the final expression of WS function used in CSP is:

$$WS^k_{norm} = an^k_{norm} + (1-a)d^k_{norm}. \tag{5}$$

Finally, CSP can be formulated as a maximization problem expressed as follows:

$$\begin{aligned} \max_{Z} \quad & \sum_{k=1}^{K} z_k WS^k_{norm}, \\ \text{s.t.} \quad & \sum_{k=1}^{K} z_k = 1, \\ & z_k \in \{0,1\}, \quad 1 \leq k \leq K, \end{aligned} \tag{6}$$

where $Z = \{z_1, z_2, ..., z_K\}$. Concerning the value of z_k, with $1 \leq k \leq K$, in case a UAV-BS deployed in the cluster k then $z_k = 1$, otherwise $z_k = 0$. The operation of the CSP subprocess, which solves the optimization problem (6), is summarized in Algorithm 2.

Algorithm 2 Clustering Selection Process (CSP).

1: **input**: Available number of UAV-BSs D and the set of formed clusters of users S
2: **for each** $S_k \in S$ **do**
3: Calculate the weight score WS^k_{norm} of S_k using (5)
4: **end for**
5: Sort in descending order the elements of set $S = \{S_1, S_2, \ldots, S_K\}$ based on the corresponding weight score WS^k_{norm} of each cluster S_k
6: From the ordered set S, select the first D clusters $S_1, \ldots, S_D$
7: **output**: The D clusters $S_1, \ldots, S_D$, where the available UAV-BSs should be deployed into the coordinates of the medoids of each cluster to maximize the objective function of (6)

4. Performance Evaluation

This section presents the simulation results obtained from the UDP process. The computer simulation was performed on a custom-made MATLAB© simulator executed on a computer consisting of Windows 10 64-bit operating system, Intel Core i7-8700 CPU, and 8 GB of RAM. For the environment where the whole transmission occurs, we consider an urban region of interest covered by three MBSs, as described in Section 2. Due to the start of an unexpected event, it is essential to integrate the available D UAV-BSs into the existing infrastructure to maximize the offloading of existing MBSs. In this way, we develop

a dynamic UAV-aided offloading scheme for the duration of the event. Regarding the users' handover procedure from MBSs to UAV-BSs, it is triggered for the users who receive a stronger signal from the UAV-BS. As a result, these users are stopped being served by the MBS and continue to receive their services from the UAV-BS.

The users are placed in the coverage area according to the specifications of 3GPP [25], creating H hotspots with an average of N/H users in each hotspot. For comparative purposes, we simulate fixed pico-base stations (PBS) deployments by either placing them randomly in the region of interest or through planned strategy in the center of each hotspot [25]. Hereinafter, we refer to these two schemes as picocell deployment process planned (PDPP) and picocell deployment process randomly (PDPR), respectively. Additionally, performance comparisons are made with another placement technique of UAV-BSs that utilizes the k-means unsupervised machine learning algorithm [23], from now on referred to as k-means deployment process (KDP). Also, for PDPP and PDPR we assume that both LOS and non line of sight (NLOS) channel conditions can exist; therefore, the proposed 3GPP models [25] were used to evaluate the picocell deployment under both LOS and NLOS conditions. The rest of the selected parameters regarding the H-CRAN network configured in the event of an exponential increase in data traffic, are listed in Table 1.

Table 1. Simulation parameters.

Parameters	Values
Simulated frames	100,000
Number of MBSs M	3
Number of UEs N	150
Number of UAV-BSs/PBSs	1–10
Number of UE hotspots H	1–10
MBS downlink operating frequency f_m	2 GHz
UAV-BS downlink operating frequency f_d	1.8 GHz
MBS Transmit power P_m	43 dBm
UAV-BS Transmit power P_d	23 dBm
Area of interest W	4.5 km $\times$ 4.5 km
MBS Cell Radius R_{MBS}	1.27 km
MBS Path loss model (NLOS)	$128.1 + 37.6 \log_{10}(r)$ dB
Path loss model (NLOS) for PBS	$145.4 + 37.5 \log_{10}(r)$ dB
Path loss model (LOS) for PBS	$38 + 30 \log_{10}(r)$ dB
Path loss model (LOS/NLOS) for UAV-BS	Elevation Angle-Based Model [26]
Terrestrial Environment	Urban
UE receive antenna gain G_r	0 dBi
MBS transmit antenna gain G_t^m	15 dBi
UAV-BS transmit antenna gain G_t^d	0 dBi
PBS transmit antenna gain G_t^p	0 dBi
Terrestrial Environment	Urban
Log-Normal Shadowing	6 dB
Bandwidth B_g with $g = \{$MBS, PBS, UAV-BS$\}$	5 MHz

4.1. Comparative Evaluation of UDP in Terms of Offloading Percentage

In this section, we examine the percentage of the UEs offloaded from the MBSs. More specifically, comparisons between the proposed and the compared methods are presented in terms of percentage offload. Furthermore, constraints regarding the maximum number of UEs that a single UAV-BS or PBS can serve, are considered.

First, we determine the best weighting factor a concerning the UDP method. Figure 2 illustrates the offloading percentage achieved by the UDP, depending on the available number of UAV-BSs and the weighing factor a. The red line represents the best values of a, depending on the available number of UAV-BSs, where the offloading percentage is maximized. More specifically, we observe that for 1 to 5 UAV-BSs, we obtain the maximum offloading rates for $a = 0.5$. When the number of UAV-BSs is equal to 6, then the maximum offloading rate is reached for $a = 0.6$. Morevoer, when the number of UAV-BS is greater

than 6, the maximum offloading rates are achieved for $a = 0.7$. Based on this observation, we have set these values in our simulations for the UDP method. Regarding the KDP method, we have considered fixed value of $a = 0.5$, as in [23]. From Figure 2, we can easily conclude that as the number of available UAV-BSs increases, it is more important to consider the number of users in each cluster n^k_{norm} for CSP process instead of the distance d^k_{norm} from the MBSs ($a > 0.5$), in order to achieve the maximum possible offloading.

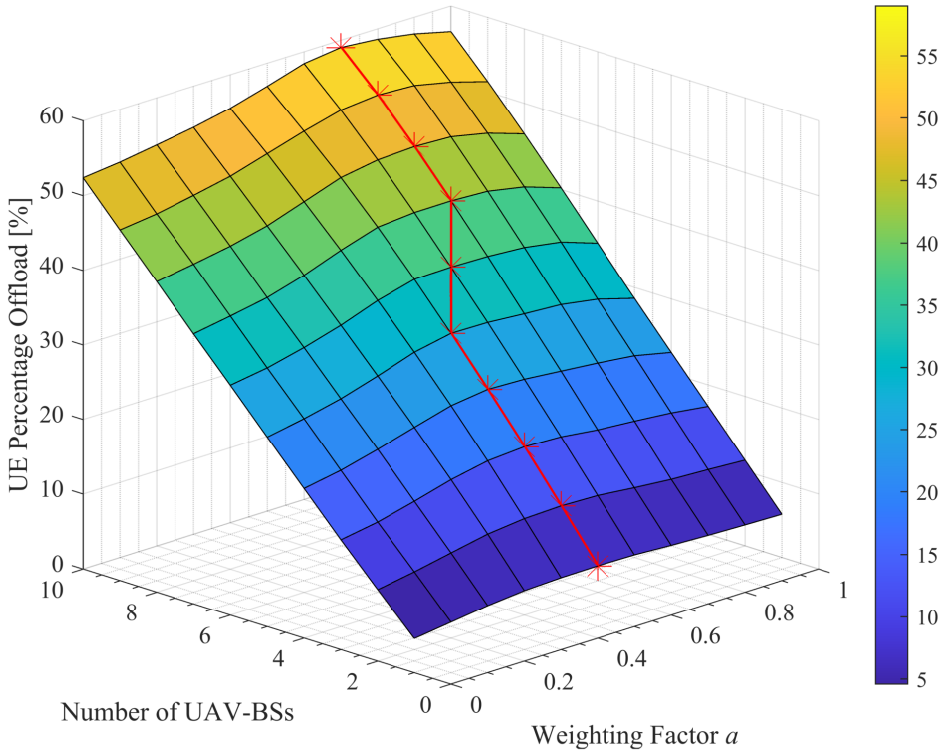

Figure 2. Percentage of UEs offloaded from the MBSs utilizing the UDP method, for different number of UAV-BSs and weighting factor a.

In Figure 3, the percentage of UEs offloaded from the MBSs for different number of UAV-BSs/PBSs is presented, utilizing the proposed scheme, as well as the compared ones. As it can be observed, the UDP process outperforms all the compared methods for all configurations. Furthermore, when the number of the available UAV-BSs or PBSs is less than two, the offloading rate for the KDP and PDPP with LOS channel conditions remains the same. However, when the availability of UAV-BSs/PBSs increases, the methods that utilize ML techniques for the UAV-BSs placement outperform the traditional PDPP and PDPR methods, respectively. The maximum percentage of the UEs offloaded from the MBSs through the PDPP and PDPR is 44% and 3%, respectively. PDPR has quite small offloading percentage since the placement of PBSs is random, and practically, it can not be ensured that the formed hotspots of UEs will be covered. In addition, the maximum offloading percentage achieved by the KDP and UDP methods is 49% and 59%, respectively. Obviously, due to the operation of the UDP and KDP methods that place the UAV-BSs in actual UE concentrations, the maximum offloading rate is achieved, subject to the capacity constraints regarding the maximum number of UEs that each UAV-BS can serve simultaneously.

Concerning the examined techniques that utilize ML methods, from Figure 3, it is observed that UDP provides significant offloading gains as compared to KDP for all configurations. This happens due to the use of the k-medoid algorithm for the UDP method, in contrast to the KDP that utilizes the k-means algorithm to form the candidate clusters. As opposed to k-means, k-medoid is not deceived from the outliers of the cluster and thus is more robust to the extreme positions of the UEs in the hotspots. It should be noted that for 150 UEs and 10 available UAV-BSs, where each UAV-BS can serve 10 UEs

simultaneously, the maximum offloading that can be achieved is 66.7%. In this case, the UDP method achieves 59% offloading, which is equal to 88.5% of the maximum possible offloading percentage. Meanwhile, KDP reaches 73.5% of the maximum possible offloading percentage, which is 15% below the UDP.

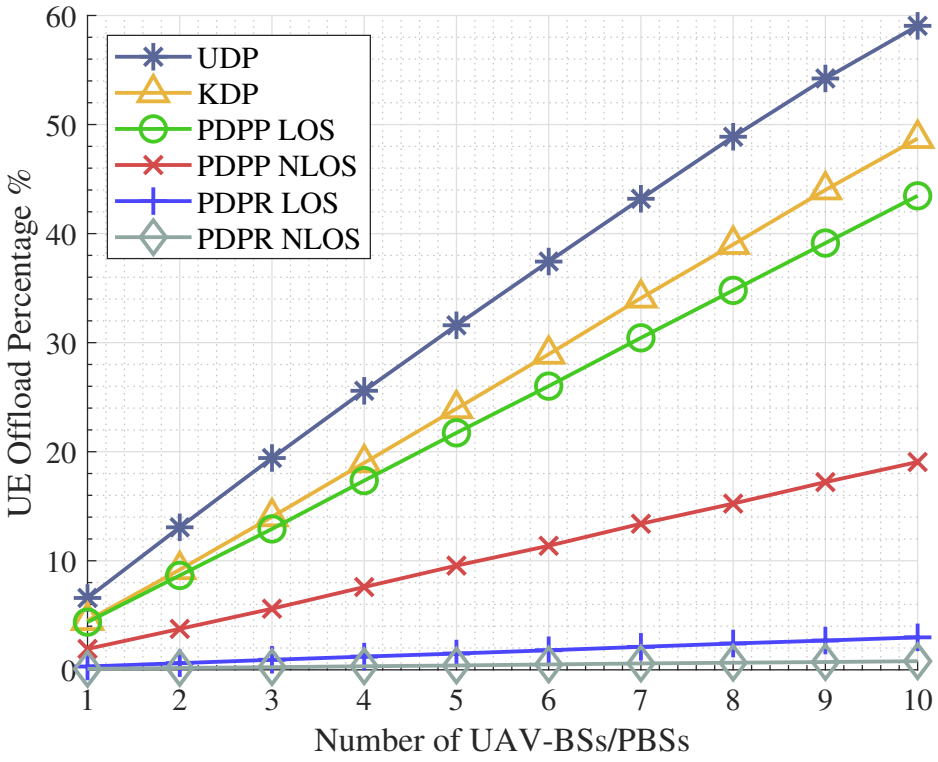

Figure 3. Percentage of UEs offloaded from the MBSs for different number of UAV-BSs/PBSs.

4.2. Comparative Evaluation of UDP in Terms of the Received Signal Strength (RSS)

In Figure 4, we observe users' average downlink received signal strength (RSS) before and after the UDP process. Specifically, the MBS bar refers to users' average RSS when they are served from the MBSs with a value equal to −73.5 dBm. Also, with the optimal placement of 5 UAV-BSs through the UDP process, the average RSS from the UEs increases to −68.8 dBm, which is an expected value, as the available access points introduced to the existing network increases. Additionally, both UDP and KDP methods that employ machine learning algorithms outperform the PDPP and PDPR deployment scenarios under both LOS and NLOS channel conditions. This happens since the picocells are fixed placed at expected points of users' concentration. On the contrary, UAV-BSs are placed in actual concentration points because they consider users' locations in real-time, as well as the on-demand user traffic.

Moreover, UDP outperforms KDP with average received signal strengths of the UEs, −68.8 dBm and −69.8 dBm, respectively. This is due to the nature of k-means and k-medoids algorithms. Specifically, k-medoids clustering is more robust to noise and outliers than k-means and ensures that the UAV-BSs will be adequately positioned, thus resulting in improved average RSS value of the UEs located inside the formed clusters. The robustness of UDP compared to KDP is confirmed in Figure 5, which shows the improvement of the total average received signal power from the UEs by integrating either more PBS or UAV-BSs. As the number of base stations increases, regardless of whether it is PBS or UAV-BS, the quality of the user's reception improves. This is reasonable as the integration of more PBSs or UAV-BSs decreases the average distance between the users and the available base stations. Furthermore, as it can be observed from Figure 5, UDP outperforms KDP at all configurations.

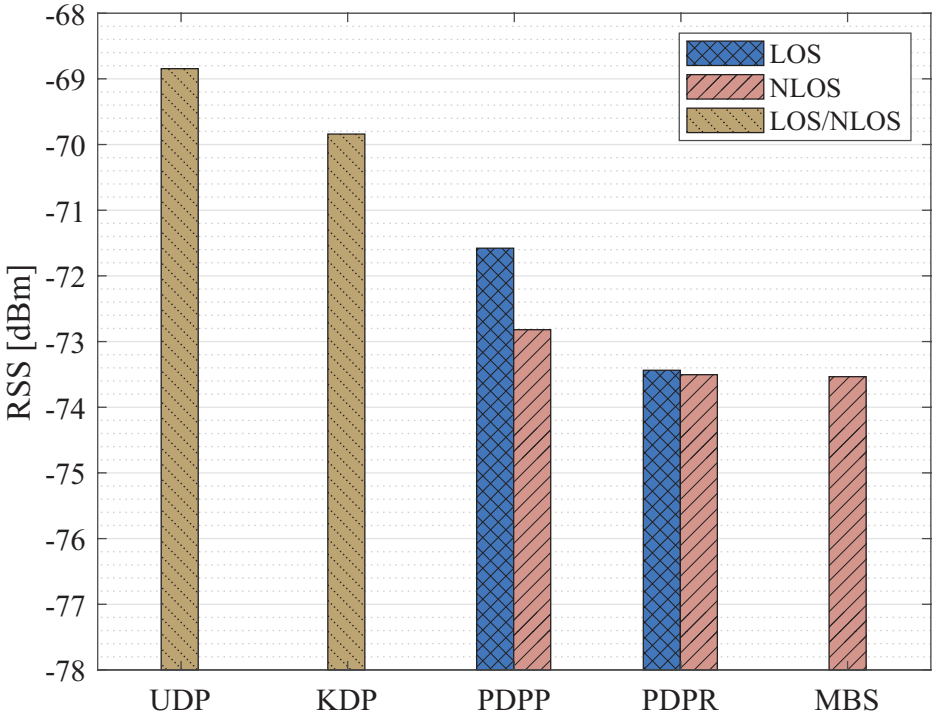

Figure 4. Average downlink RSS with 5 UAV-BSs/PBSs.

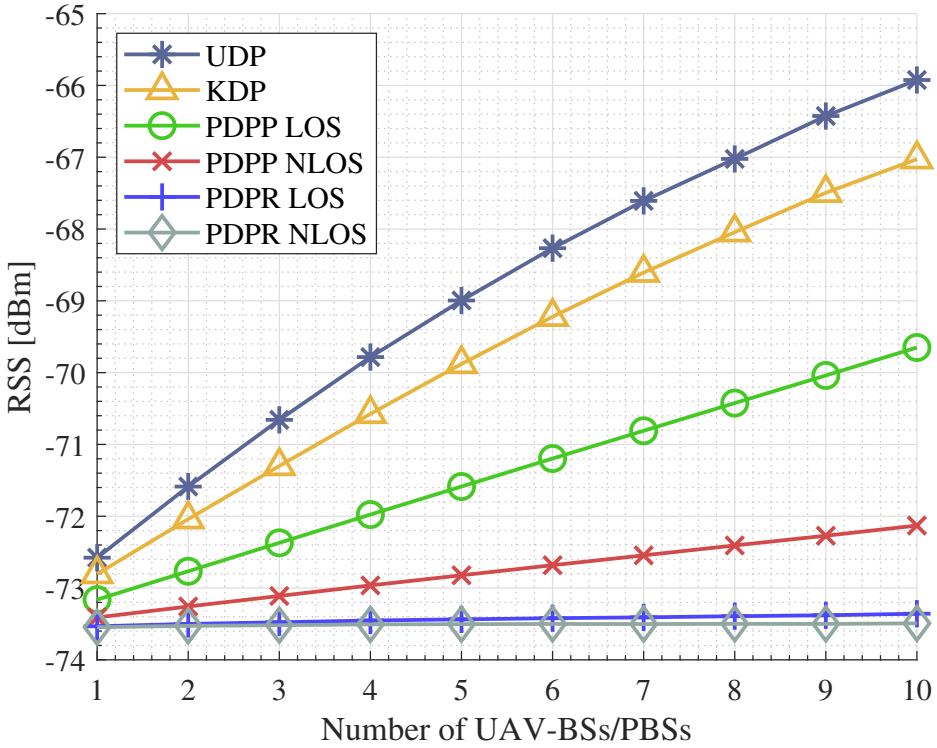

Figure 5. Average downlink RSS for all UEs in the region of interest for different number of UAV-BSs/PBSs.

4.3. Comparative Evaluation of UDP in Terms of System Sum Rate, User Average throughput, and Spectral Efficiency

In this section, simulation results are presented in terms of system sum rate, user average throughput, and spectral efficiency. We consider a 5G H-CRAN system operating under orthogonal frequency division multiplexing (OFDM), based on a standalone orthogonal multiple access (OMA) scheme [27]. Thus, the available spectrum is equally

distributed to the users of the system. In Figure 6, comparisons between the proposed and the compared methods are presented in terms of system sum rate. Before the deployment of either UAV-BSs or PBSs, MBS downlink sum rate is equal to 122 Mbps. We present the MBS downlink sum rate as the lower bound for the system sum rate. After the integration of UAV-BSs / PBSs, we observe that as the number of UAV-BSs / PBSs increases, the system sum-rate increases significantly. Notable system sum-rate gains can be harvested using the UDP and KDP methods, as compared to the PDPP and PDPR under both LOS and NLOS channel conditions, for all configurations. Furthermore, UDP behaves marginally better than KDP in terms of system sum rate, when the number of UAV-BSs is less than 4. For more than 4 UAV- BSs, UDP exceeds KDP in terms of system sum rate. The maximum difference of 28 Mbps between UDP and KDP is presented when the available UAV-BSs are 10. In this case, UDP achieves system sum-rate equal to 857 Mbps, while the KDP method achieves 829 Mbps.

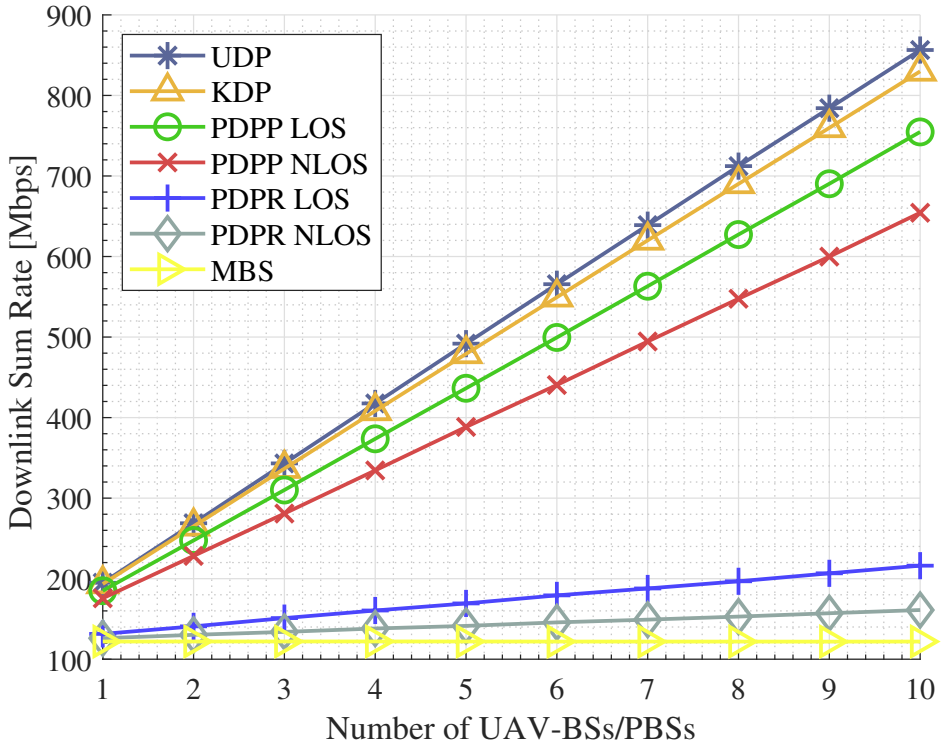

Figure 6. Downlink system sum rate for different number of UAV-BSs/PBSs.

Moreover, as shown in Figure 7, the trend of the average throughput for each UE is the same as the trend shown for the system sum rate regarding the examined methods. The average throughput of each UE utilizing only MBSs is equal to 0.8 Mbps, while with the use of UDP and KDP methods it ranges from 1.3 to 5.7 Mbps and 1.2 to 5.5 Mbps, respectively. The achievable UE average throughput using the PDPP and PDPRR technique under LOS conditions is ranging from 1.2 to 5 Mbps and 0.9 to 1.4 Mbps, respectively.

As discussed in Section 4.1, UDP performance is superior to that of KDP in terms of offloading percentage. This means that the UAV-BSs in the case of the UDP method will be more loaded as compared to the case where UAV-BSs operate under the KDP method. Hence, UAV-BSs will distribute the available spectrum to more users in the case of the UDP method contrary to the KDP method where the available spectrum will be allocated to fewer users. It is important to note that UDP exceeds KDP in terms of system sum-rate and average UE throughput, while the KDP method allocates the available bandwidth to fewer users than the UDP method. This happens because UDP gives a higher average downlink RSS compared to the KDP, as discussed in Section 4.2. Thereby, UDP is twice as profitable as KDP, since it significantly increases the offloading percentage of the terrestrial

network, as compared to KDP and at the same time, provides better sum-rate and average UE throughput performance than KDP. This can be proved through Figure 8, which shows the spectral efficiency of the UDP and KDP methods, for different number of UAV-BSs. It is observed that the UDP method outperforms the KDP method for all configurations. As explained above, since the average downlink received power provided to users using the UDP method is much better than in the case of KDP, the utilization of the spectrum is much higher even for a smaller part of it.

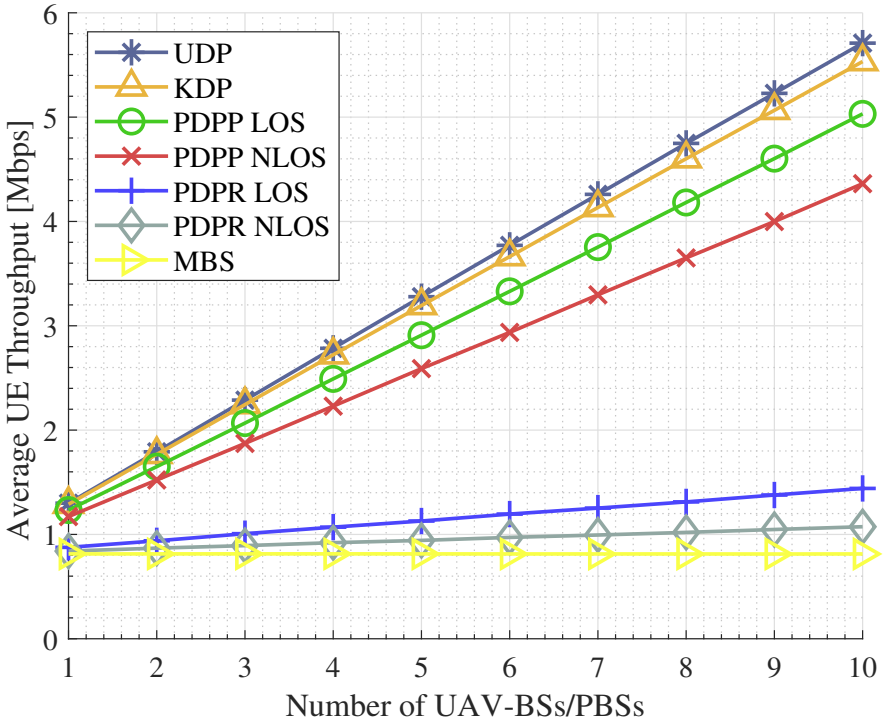

Figure 7. Average UE throughput for different number of UAV-BSs/PBSs.

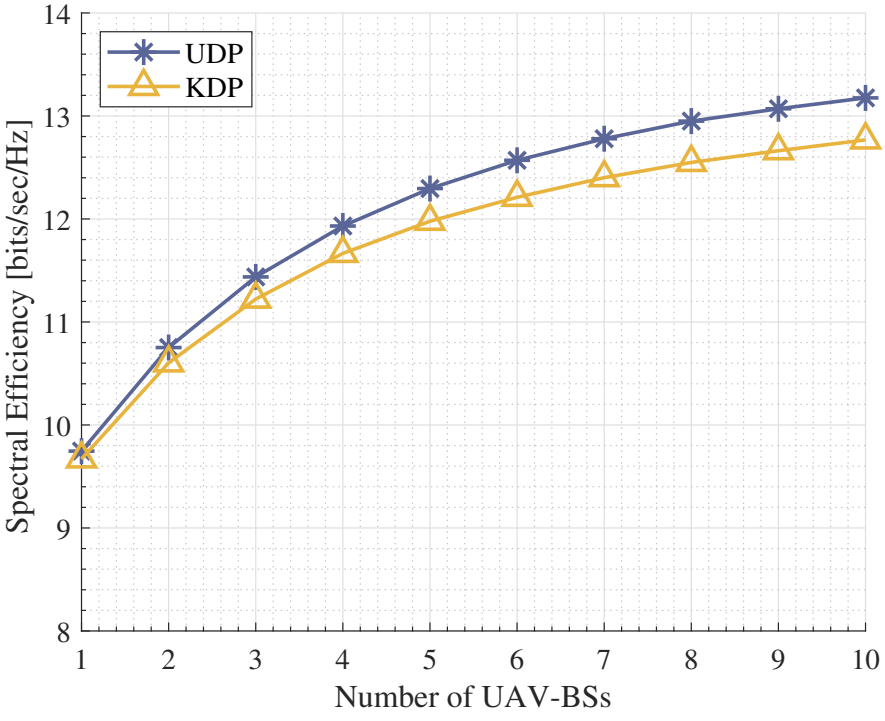

Figure 8. Average UE spectral efficiency for different number of UAV-BSs.

4.4. Comparative Evaluation of UDP under Increased Localization Inaccuracy

The previous results refer to the case where the estimation of user locations in the region of interest is perfect and there are no localization errors. Moreover, the authors in [28] state that 95% of localization errors in LTE systems are less than 20 m. Figure 9 illustrates the impact of the localization error of 5 m, 10 m and 20 m to the average received signal strength of the UEs located in the area of interest for the UDP and KDP methods, respectively. It is noteworthy that these localization errors can be well-approximated via a zero-mean normal distribution with standard deviation σ of 0 m, 5 m, and 10 m [23]. In addition, we evaluate the performance of the proposed UDP method against the KDP in case of an error bigger than 20 m, raising the standard deviation equal to 20, where such a state can be considered as an extreme case. The results show that there is approximately 1 dBm difference regarding the received signal strength between the UDP best scenario (no error) and the worst scenario (σ = 20 m). It is noticed that even with a significant error in user location estimation, UDP development still retains its advantage over the KDP method. This is due to the fact that the algorithm is not sensitive to extreme outliers, which most likely occur with the addition of an error greater than 20 m.

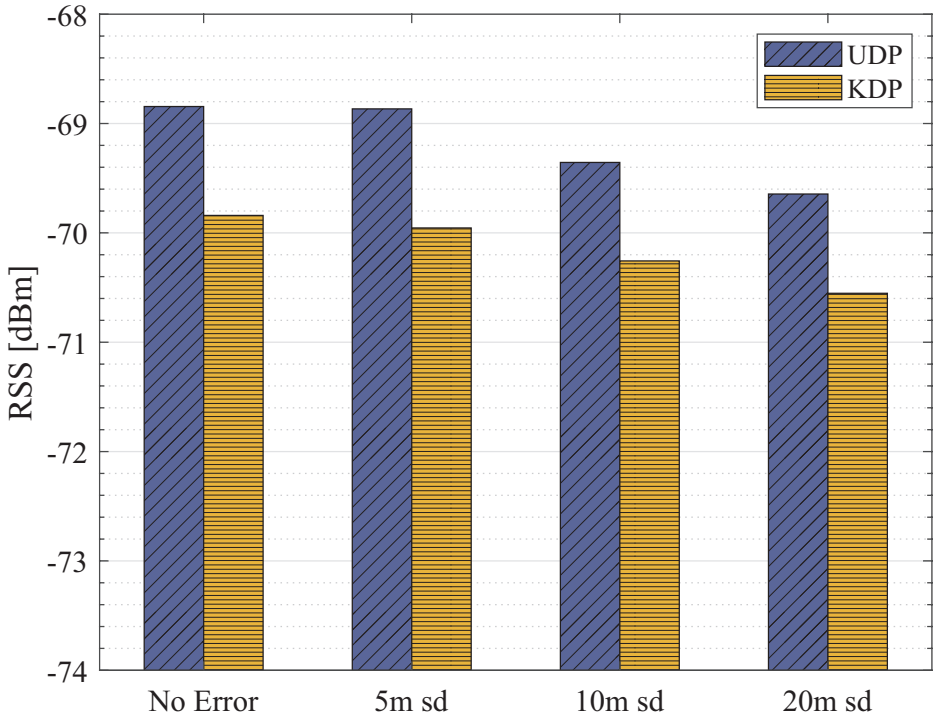

Figure 9. Impact of the increased localization inaccuracy on the achieved average RSS value for the UDP and KDP scheme, given the deployment of 5 UAV-BSs.

5. Conclusion and Future Directions

5.1. Conclusions

The on-demand deployment of aerial base stations can improve the communication quality via offloading the terrestrial cellular network in highly loaded scenarios, where connectivity and service provisioning is threatened. UDP, a UAV-aided offloading scheme for 5th generation terrestrial cellular networks, utilizing the k-medoid, was presented in this context. Through UDP, both the offloading percentage of the terrestrial network and the average received signal strength of the UEs are improved, compared to traditional offloading strategies utilizing pico base stations, PDPP/ PDPR, and to KDP, which is also a UAV-aided offloading scheme utilizing k-means ML technique. Moreover, UDP outperforms the compared methods in terms of system sum rate, user average throughput, and spectral efficiency. Finally, UDP outperforms the compared methods in terms of the

received signal strength of UEs in scenarios where localization inaccuracies exist regarding UEs' locations.

5.2. Future Directions

There are several interesting directions where UDP can be extended and applied. Considering an air–ground integrated network developed for emergency scenarios, the UAV-BSs can be deployed to serve ground users where suffer from communication outages in the disaster area. Specifically, adopting the UDP scheme in emergency networks scenarios will facilitate the fast and unsupervised deployment of a UAV-enabled Emergency Cellular Network (UECN). Furthermore, in scenarios of popular content dissemination to multiple terrestrial users, a preventive caching approximation can be integrated into the UDP for effective content transmission. This approach can achieve significant offloading gains to relieve the existing terrestrial network infrastructure of high-demand-driven data traffic. Another essential aspect of UDP that should be further investigated is the optimization of the handover procedure between the UAV-BSs and the terrestrial base stations, especially in offloading scenarios. In addition to the maximum received signal strength criterion, other network performance metrics should also be considered, such as user rate requirements, delay constraints, and load balancing between network nodes. Finally, of potential interest is the integration of artificial intelligence and reinforcement machine learning methods in the context of aerial–terrestrial networks, which will allow fully autonomous zero-touch UAV trajectory schemes to offload traffic on the edges of the terrestrial base stations.

Author Contributions: Conceptualization, L.T., M.K. and D.V.; methodology, L.T. and M.K.; software, L.T. and M.K.; validation, L.T., M.K. and D.V.; formal analysis, L.T.; investigation, L.T. and M.K.; writing—original draft preparation, L.T. and M.K.; writing—review and editing, M.K. and D.V.; supervision, D.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CSP	Clustering Selection Process
GBS	Ground Base Station
H-CRAN	Heterogeneous Cloud Radio Access Network
IoT	Internet of Things
JTO	Joint Traffic Offloading
KDP	K-means Deployment Process
LOS	Line of Sight
LTE	Long Term Evolution
MBS	Macro Base Station
MmWave-NOMA	Millimeter Wave Non Orthogonal Multiple Access
MTs	Mobile Terminals
NLOS	Non Line of Sight
OFDM	Orthogonal Frequency Division Multiplexing
PBS	Pico Base Station
PDPP	Picocell Deployment Process Planned
PDPR	Picocell Deployment Process Randomly
QoS	Quality of Service
RRH	Remote Radio Head

RSS	Received Signal Strength
SAGIN	Space-Air-Ground Integrated Networks
SWIPT	Simultaneous Wireless Information and Power Transfer
UAV	Unmanned Aerial Vehicle
UAV-BSs	UAV Base Stations
UDP	UAV-BS Deployment Process
UE	User Equipment
UCP	User Clustering Proces
WS	Weighted Score

References

1. Bithas, P.S.; Michailidis, E.T.; Nomikos, N.; Vouyioukas, D.; Kanatas, A.G. A Survey on Machine-Learning Techniques for UAV-Based Communications. *Sensors* **2019**, *19*, 5170. [CrossRef] [PubMed]
2. Cao, X.; Yang, P.; Alzenad, M.; Xi, X.; Wu, D.; Yanikomeroglu, H. Airborne Communication Networks: A Survey. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1907–1926. [CrossRef]
3. Elsawy, H.; Hossain, E.; Kim, D.I. HetNets with cognitive small cells: User offloading and distributed channel access techniques. *IEEE Commun. Mag.* **2013**, *51*, 28–36. [CrossRef]
4. Hoadley, J.; Maveddat, P. Enabling small cell deployment with HetNet. *IEEE Wirel. Commun.* **2012**, *19*, 4–5. [CrossRef]
5. Wang, L.; Chao, Y.; Cheng, S.; Han, Z. An Integrated Affinity Propagation and Machine Learning Approach for Interference Management in Drone Base Stations. *IEEE Trans. Cogn. Commun. Netw.* **2020**, *6*, 83–94. [CrossRef]
6. Chabbouh, O.; Rejeb, S.B.; Choukair, Z.; Agoulmine, N. Offloading decision algorithm for 5G/HetNets cloud RAN. In Proceedings of the 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 22–24 September 2016; pp. 1–5. [CrossRef]
7. Xu, L.; Luan, Y.; Cheng, X.; Xing, H.; Liu, Y.; Jiang, X.; Chen, W.; Chao, K. Self-optimised joint traffic offloading in heterogeneous cellular networks. In Proceedings of the 16th International Symposium on Communications and Information Technologies (ISCIT), Qingdao, China, 26–28 September 2016; pp. 263–267. [CrossRef]
8. Omran, A.; Sboui, L.; Kadoch, M.; Chang, Z.; Lu, J.; Liu, R. 3D Deployment of Multiple UAVs for Emergent On-Demand Offloading. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; pp. 692–696. [CrossRef]
9. Qin, Y.; Kishk, M.A.; Alouini, M.S. Performance evaluation of uav-enabled cellular networks with battery-limited drones. *IEEE Commun. Lett.* **2020**, *24*, 2664–2668. [CrossRef]
10. Hu, Z.; Zheng, Z.; Song, L.; Wang, T.; Li, X. UAV Offloading: Spectrum trading contract design for UAV-Assisted cellular networks. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6093–6107. [CrossRef]
11. Lyu, J.; Zeng, Y.; Zhang, R. UAV-Aided Offloading for Cellular Hotspot. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 3988–4001. [CrossRef]
12. He J.; Wang, J.; Zhu, H.; Gomes, N.J.; Cheng, W.; Yue, P.; Yi, X. Machine Learning based Network Planning in Drone Aided Emergency Communications. In Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 25–28 May 2020. [CrossRef]
13. Qu, H.; Zhang, W.; Zhao, J.; Luan, Z.; Chang, C. Rapid Deployment of UAVs Based on Bandwidth Resources in Emergency Scenarios. In Proceedings of the Rapid Deployment of UAVs Based on Bandwidth Resources in Emergency Scenario, Nanjing, China, 29–31 May 2020; pp. 86–90. [CrossRef]
14. Cui, J.; Khan, M.B.; Deng, Y.; Ding, Z.; NaIlanathan, A. Unsupervised Learning Approaches for User Clustering in NOMA enabled Aerial SWIPT Networks. In Proceedings of the IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, 2–5 July 2019; pp. 1–5. [CrossRef]
15. Qi, W.; Zhang, B.; Chen, B.; Zhang, J. A user-based K-means clustering offloading algorithm for heterogeneous network. In Proceedings of the IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 307–312. [CrossRef]
16. Mandloi, D.; Rajeev, A. Seamless connectivity with 5G enabled unmanned aerial vehicles base station using machine programming approach. *Expert Syst.* **2021**, e12828. [CrossRef]
17. Tang, F.; Hofner, H.; Kato, N.; Kaneko, K.; Yamashita, Y.; Hangai, M. A Deep Reinforcement Learning based Dynamic Traffic Offloading In Space-Air-Ground Integrated Networks (SAGIN). *IEEE J. Sel. Areas Commun.*, **2021**, 40, 276–289. [CrossRef]
18. Ma, Y.; Wang, H.; Xiong, J.; Ma, D.; Zhao, H. Concise and Informative Article Title Throughput Maximization through Joint User Association and Power Allocation for a UAV-Integrated H-CRAN. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 6668756 . [CrossRef]
19. Schubert, E.; Rousseeuw, P.J. *Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms*; Springer: Cham, Switzerland, 2019.
20. Hossain, M.F.; Mahin, A.U.; Debnath, T.; Mosharrof, F.B.; Islam, K.Z. Recent research in cloud radio access network (C-RAN) for 5G cellular systems-A survey. *J. Netw. Comput. Appl.* **2019**, *139*, 31–48. [CrossRef]
21. Alzenad, M.; Shakir, M.Z.; Yanikomeroglu, H.; Alouini, M. FSO-Based Vertical Backhaul/Fronthaul Framework for 5G+ Wireless Networks. *IEEE Commun. Mag.* **2018**, *56*, 218–224. [CrossRef]

22. Wilson, H.; Boots, B.; Millward, A. A comparison of hierarchical and partitional clustering techniques for multispectral image classification. *IEEE Int. Geosci. Remote. Sens. Symp. (IGARSS)* **2002**, *3*, 1624–1626.
23. Galkin, B.; Kibilda, J.; DaSilva, L.A. Deployment of UAV-mounted access points according to spatial user locations in two-tier cellular networks. In Proceedings of the Wireless Days (WD), Toulouse, France, 23–25 March 2016; pp. 1–6. [CrossRef]
24. Moraitis, N.; Tsipi, L.; Vouyioukas, D.; Gkioni, A.; Louvros, S. Performance evaluation of machine learning methods for path loss prediction in rural environment at 3.7 GHz. *Wirel. Netw.* **2021**, *27*, 4169–4188. [CrossRef]
25. 3GPP, Further Advancements for E-UTRA Physical Layer Aspects (Release 9), TR 36.814 V9.2.0 (2017-03). Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2493 (accessed on 5 December 2022).
26. Zhang, H.; Song, L.; Han, Z. *Unmanned Aerial Vehicle Applications over Cellular Networks for 5G and Beyond*; Springer: Cham, Switzerland, 2020.
27. Karavolos, M.; Nomikos, N.; Vouyioukas, D. Enhanced Integrated Satellite-Terrestrial NOMA with Cooperative Device-to-Device Communication. *Telecom* **2020**, *1*, 126–149. [CrossRef]
28. Vaghefi, R.; Buehrer, R. Improving positioning in LTE through collaboration. In Proceedings of the 11th Workshop on Positioning, Navigation and Communication (WPNC), Dresden, Germany, 12–13 March 2014; pp. 1–6.