UAV Trajectory Optimization Based on Predicted User Locations

Lester Ho
Wireless Communication Lab
Tyndall National Institute
Dublin, Ireland
lester.ho@tyndall.ie

Sobia Jangsher
School of Electronic Engineering
Dublin City University
Dublin, Ireland
sobia.jangsher@dcu.ie

***Abstract*—Unmanned aerial vehicles (UAVs) can extend the coverage of wireless networks due to their high mobility and the favorable radio propagation characteristics. This paper studies the trajectory optimization of UAV that are acting as radio relays. The optimization is based on predicted user locations (UTO-PUL) to assist communication to the ground users who are unable to get coverage from the base station (BS). The existing work on trajectory design has considered several optimization approaches as well as reinforcement learning (RL) algorithms. All the algorithm takes into consideration the existing state of the network such as the channel conditions, initial positions and computes the destination of the UAV based on it. The proposed algorithm is designed to also consider the predicted user mobility of the future instances. The objective is to ensure the ground users are connected to the BS. The proposed UTO-PUL algorithm's performance is evaluated using simulations in a scenario with challenging terrain, where the proposed algorithm reduced the probability of users having no coverage by between 45% to 85% compared to non-predictive approaches, and achieved gains in median downlink signal power of 14 dB compared with a deep reinforcement learning (DRL) algorithm.**

***Index terms*—** Unmanned aerial vehicle (UAV), trajectory optimization, drone, prediction, predicted user locations.

I. Introduction

Unmanned aerial vehicle (UAV) communications have gained interest for several civilian as well as military applications. The telecommunications community has recognised the potential that UAVs can play in wireless communications applications [1]. Because they operate in the sky, UAVs experience unobstructed Line-of-Sight (LoS) across very long distances, which means that radio signals coming to and from a UAV node is able to avoid signal blockage caused by obstacles such as trees, buildings, and terrain features like mountains. This can allow a relatively small number of UAVs equipped with radio equipment to provide connectivity over a much larger area, compared to an equivalent network of ground-based infrastructure [2].

Because of this, UAV-based radio networks are an attractive solution for organisations that operate in remote areas with unreliable or unavailable infrastructure. Examples of these use cases include organisations performing humanitarian or peacekeeping missions, where vehicle convoys may travel long distances on missions in challenging terrain over several days. Another use case is the support for personnel performing search and rescue operations in mountainous regions. Very High Frequency (VHF) radio communications are often used during their deployments to provide voice and low-bandwidth data transmissions back to base. However, due to the distances and challenging terrain involved, personnel often find themselves out of range for reliable VHF communications, and have to rely on slower High Frequency (HF) communication. The use of long-endurance UAVs to act as radio relays is an ideal solutions to extend the radio communications range for these missions.

Two key requirement for these use cases are a high availability and reliability of the radio links between the users and the base, where any extended breaks in communications can result in loss of life. The optimization of the trajectory of the UAV relays to provide this radio link reliability is therefore an important factor to consider.

There has been several studies on UAV relay placement and trajectory optimization. The UAV trajectory is designed to maximize the transmission rate of the ground users in [3]. A single UAV is considered in this work. The problem is presented as an acyclic directed graph with the users as the vertex and the reward and cost as the edges. The Bellman-Ford algorithm is then applied to graph. In [4], UAVs are employed to relay data in a time division manner. Time slot allocation, power allocation and trajectory design is proposed using an iterative heuristic. Deep reinforcement learning has been used for UAV trajectory design in [5]–[8]. In [5], a joint trajectory design and power allocation problem is solved using deep deterministic policy gradient (DDPG) algorithm of deep reinforcement learning (DRL). A continuous action space is considered to find the optimal policy. In [9], trajectory design of a multiple UAV using DRL with centralized learning and decentralized execution is studied.

In all the studies mentioned above, the resource allocation and trajectory design is studied with the assumption that the users are static, or in the case of [3] based on user demand. They also work on the basis that the radio path loss follows distance-dependent models with probabilistic LoS/non-LoS links, or the assumption that obstruction and shadowing monotonically increases based on distance and acuteness of elevation angles. These assumptions are not a reflection of actual radio conditions that are influenced by user movement and terrain, and would not provide the required level of reliability if used to optimise real systems in the field. For example, a UAV relay needs to consider if a user is about to move behind a hill, and adjusts its location before that happens in order to provide continuous radio coverage to the user.

A. Contribution

In this paper, a novel framework of UAV trajectory optimization based on predicted user locations (UTO-PUL) for multiple relay UAVs to ensure connectivity of the users to the BS located at the base is proposed. UTO-PUL is based on the actual terrain of the area of operation. We consider that the UAV has given a-priori knowledge (or prediction) of the location of the users, as well as information of the terrain. Based on these information, the UAVs decide where they should be to provide coverage to users. UTO-PUL considers not just the final destination of the UAVs, but also the path they should take to move to the final destination to continuously maintain the radio links to the moving users. It also takes into account the potential movement of the users, and optimises the drones based on not just the current user locations, but their likely locations in the future.

B. Structure of the Paper

The paper is structured as follows: Section II presents the system model. Section III studies the problem formulation and proposes the UTO-PUL algorithm. Section IV presents the results of the proposed UTO-PUL algorithm and Section V concludes the paper.

II. System Model

We consider a UAV assisted wireless communication system where UAVs acts as radio relays to assist the communication to the ground users who are unable to achieve the minimum quality-of-service (QoS) from the ground base station (BS) directly as shown in Fig. 1. The set of UAVs is represented as $\mathbb{U} = \{1, 2, \ldots, U\}$ and the set of users is represented as $\mathbb{J} = \{1, 2, \ldots, J\}$. Each user gets associated to a UAV and the UAVs then connects to the ground BS via the wireless backhaul link. The BS-UAV and the UAV-user links are operating on orthogonal frequencies, so co-channel interference between UAV relays does not occur.

For simplification, all UAVs are flying at a height of h and the height of the BS is h_B, where $h_B \ll h$.

The flight time of each UAV is T sec. The flight time is divided into a total of M time frames each of duration $\frac{T}{M}$. The UAV to user association is quasi static over a time frame. Within a time frame the users associated to a particular UAV is assigned a variable time duration to transmit, the time duration is represented as $\tau_{i,j}^{m}$. In a time frame, a UAV moves a total of S steps whereas a user moves a total of L steps. Let $\nu_{i,s} = [x_{i,s}, y_{i,s}]^T \in \mathbb{R}^{2\times 1}$ be the 2D coordinates for the UAV i.

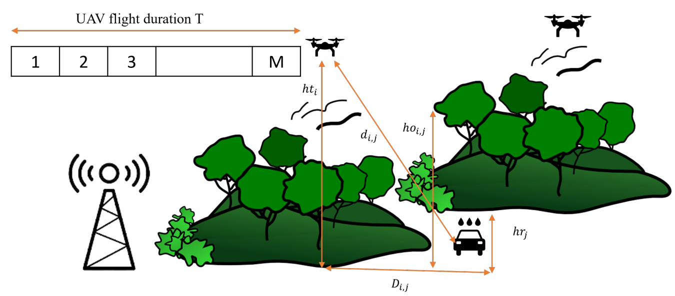


Fig. 1: System model.

A. User Mobility Model

A waypoint-based mobility model is used for the users, where random waypoints in the map are generated for each user, and the users move at a specified velocity between the waypoints in a straight line. Let $\eta_{j,\ell} = [x_{j,\ell}, y_{j,\ell}]^T \in \mathbb{R}^{2\times 1}$ be the 2D coordinates for the user j and ℓ is the index of the number of steps users will move over the time frame.

B. Air-to-Ground Communication Model

The channel gain between UAV i and user j is calculated as:

$$h_{i,j} = PL_{i,j} + KED_{i,j} + F_j \tag{1}$$

where $PL_{i,j}$ is the distance-dependent free-space path loss, $KED_{i,j}$ is the loss due to terrain blockage that is modelled as knife-edge diffraction (KED) losses into the the 1st Fresnel zone, and F_j is the loss due to forest cover for user j. This reflects the targeted use case environment which is in rural, non-urban environments.

$PL_{i,j}$ is calculated using the widely used free-space path loss calculation, with a path-loss exponent of 2.13 [10]:

$$PL_{i,j} = 21.3\log_{10}(d_{i,j}) + 21.3\log_{10}(f_c) - 157.15 \tag{2}$$

where $d_{i,j}$ is the distance in meters between UAV i and user j, and f_c is the carrier frequency in Hz.

$KED_{i,j}$ is used to capture the transition of line-of-sight (LoS) to non line-of-sight (NLoS) conditions when blockage due to the terrain occurs [11], and is given by

$$KED_{i,j} = -20\log_{10}\frac{\sqrt{(1-C(v_{i,j})-S(v_{i,j}))^2+(C(v_{i,j})-S(v_{i,j}))^2}}{2} \tag{3}$$

where $C(v_{i,j})$ and $S(v_{i,j})$ are the real and imaginary parts of the complex Fresnel integral $F_c(v_{i,j}) = int_0^{v}\exp(j\frac{\pi s^2}{2})\,ds$, respectively, and $v_{i,j}$ is calculated as

$$v_{i,j} = \Delta_{i,j}\sqrt{\frac{2}{\lambda}\left(\frac{1}{d1_{i,j}} + \frac{1}{d2_{i,j}}\right)}, \tag{4}$$

$$\Delta_{i,j} = \frac{(ho_{i,j} - ht_i)D_{i,j} + (ht_i - hr_j)d_{i,j}}{\sqrt{D_{i,j}^2 + (ht_i - hr_j)^2}} \tag{5}$$

where $d1_{i,j}$ and $d2_{i,j}$ are the distances in meters between the terrain obstacle and the UAV i and user j, respectively,

$ho_{i,j}$, hr_j and ht_i are the heights in meters of the obstacle, user and UAV respectively, and $D_{i,j}$ is the horizontal distance between UAV i and user j. In order to calculate $KED_{i,j}$, openly available global terrain elevation information from [12] is used to obtain the terrain map of the area of operation. The terrain map is used to determine the presence of any obstacles between the UAV and the user. An obstacle is defined as any peaks in the terrain (e.g. hills, mountains) that is within 60% of the first Fresnel zone radius. If there are more than one obstacle, the sum of all the KED loss for each obstacle is used.

F_j is obtained using geographical forest cover datasets [13], where F_j is set to 5dB if the user j is located in a forested location, and set to zero if the user is in an open area.

III. UAV Trajectory Optimization based on Predicted User Locations

This section formulates the problem and presents the proposed UTO-PUL algorithm. The association of the UAV to user for a time frame is mathematically represented as

$$\alpha_{i,j}^{m} = \begin{cases} 1, & \text{if user } j \text{ is associated to UAV } i \text{ in timeframe } m \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

The total number of association variables for a time frame m is $[\mathcal{A}^{\Updownarrow}]_{U\times J}$. For the trajectory the UAV is considered to fly at a constant altitude above mean sea level (MSL), or 50m above ground level, whichever is higher. In other words, if the UAV encounters terrain that is higher than its default MSL altitude, it will maintain an above ground level (AGL) altitude of 50m. The x-axis and the y-axis of the trajectory is used as the control variable in this problem. We use $x_{i,s}^{m}$ and $y_{i,s}^{m}$ as the x-axis and the y-axis, respectively, of the UAV position in time slot s for a frame m. As already mentioned a UAV moves a total of S steps in a time frame, therefore a total of $2S$ trajectory points are the control variables of a UAV i in a time frame m. The total of for all UAVs are represented as $\mathcal{X}^m = [x_{i,s}^{m}]_{U\times S}$ and $\mathcal{Y}^m = [y_{i,s}^{m}]_{U\times S}$.

A. Problem Formulation

The objective of the problem is to ensure that the connectivity of the ground users to the BS is maximized in the time slots over the time frame. The problem formulation for the joint association and the trajectory design for a time frame m is presented below (For simplicity the variable m is removed):

$$\underset{\mathcal{A},\mathcal{X},\mathcal{Y}}{\text{maximize}} \quad \sum_{i=1}^{U}\sum_{s=1}^{S}\alpha_{i,j}\min_{j\in\mathbb{J}} P_{i,j}^{s} \tag{7a}$$

$$\text{subject to} \quad \sum_{i=1}^{U}\alpha_{i,j} = 1 \qquad \forall j, \tag{7b}$$

$$\alpha_{i,j} \in \{0,1\} \qquad \forall i,j, \tag{7c}$$

$$x_{i,s} \in \{0, X^{max}\} \qquad \forall i,s, \tag{7d}$$

$$y_{i,s} \in \{0, Y^{max}\} \qquad \forall i,s \tag{7e}$$

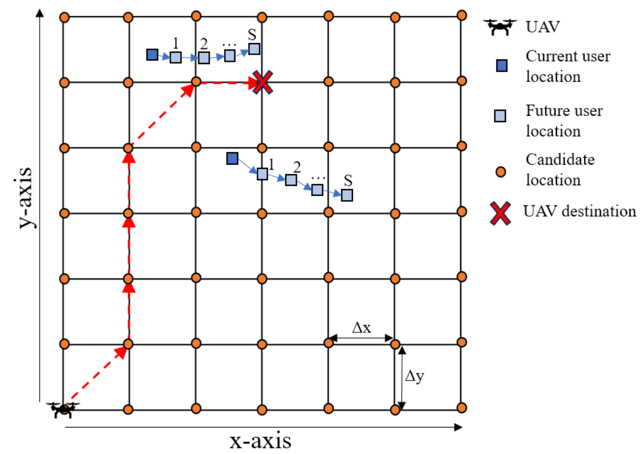


Fig. 2: Trajectory Design

Constraint 7b ensures that a user in a time frame is associated to only one UAV. Constraint 7c, 7d and 7e are the boundary constraints to ensure the value of the control variable is in the limit. The formulated problem is an binary integer problem and optimal solution cannot be computed in polynomial time.

B. UTO-PUL Algorithm

In this section, UTO-PUL algorithm is proposed to solve the formulated problem The basic idea behind the design of the algorithm is to represent the trajectory design as a route selection problem in a directed graph. To solve it we decompose the problem in two parts, a) user-to UAV association, and b) trajectory design of the UAV. Next, we will discuss the two decomposed parts.

1) K-means based user to UAV association: K-means algorithm is a special case of expected maximization algorithm with uniform covariance. The users are divided in a number of clusters equal to the number of UAV present in the network. The K-means algorithm takes number of clusters and the sample points as the input. The total number of clusters is the number of UAVs in the network. The output of the clustering algorithm is the set $\mathbb{J}_i$, $\forall i \in \mathbb{U}$. The cluster are decided based of the variance of the sample points. The sample points in this case are the coordinates of the ground users at time slot 1 represented by the set $\eta_{j,1}$, $\forall j$.

$$[\mathbb{J}_1, \mathbb{J}_2, \ldots \mathbb{J}_U] = \text{K-Means}(U, \eta_{1,1}, \eta_{2,1}, \ldots \eta_{J,1}) \tag{8}$$

The clustering is performed once at the start of a time frame m keeping in view the clustering of the users, the UAV trajectory is designed for a time frame m

2) Predicted user location based trajectory design: As already mentioned, the trajectory design takes into consideration knowledge the location of the users in the future time slots. The exact number of coordinates of the users for which you can look ahead can be varied. The coordinate of user j in slot s is $\eta_{j,s}$. The trajectory of S steps for UAV i in defined using the vector $\eta_j = [\eta_{j,1}, \eta_{j,2}, \ldots \eta_{j,S}], \forall j \in \mathbb{J}_i$. The algorithm for the trajectory of each UAV has the following steps:

1) **UAV Candidate Destination Locations:** The candidate destination location are decided from a grid as shown in Fig. 2. Choosing more number of candidate destination locations will ensure in locating the best possible destination location but it will increase the computation complexity. The candidate location set for a UAV i is represented as $\mathbf{d_i}$. We view this grid of candidate locations as vertices of a lattice graph, where the immediate neighbouring vertices are connected by bi-directional edges, weighted by a score described below.
2) **Score-based Destination Computation:** Each possible candidate location of the UAV is scored based on the received power of the users associated to it. The score of candidate location k, s_k, is a dimensionless metric that is assigned based on the minimum of the received power of the signal from the UAV to the BS and the users connected to the UAV, $R_j = \min_{j \in \mathbb{J}} P^s_{i,j}$. As s_k is the metric used to calculate the UAV trajectory using shortest path calculation, a lower value of s_k corresponds to a better candidate location. Table I shows the mapping of R_j with s_k. If R_j at the candidate location is equal or less than -50 dBm, s_k is set to 1. s_k increases linearly by 1 for every decrease of 5 dBm until the point where R_j is equal to the receive sensitivity threshold of -110 dBm. Below this threshold, the value of s_k is doubled (i.e. $14 \times 2 = 28$) to penalise candidate locations where the UAV cannot provide coverage to users.

s_k	R_j
1	$R_j \leq -50$
2	$-55 \leq R_j < -50$
3	$-60 \leq R_j < -55$
...	...
12	$-105 \leq R_j < -100$
13	$-110 \leq R_j < -105$
28	$R_j < -110$

TABLE I: Mapping of R_j to s_k

The value of s_k for each of the candidate locations are calculated. For a UAV i, the destination is selected as

$$d_i^* = \arg\min_{k \in \mathbf{d_i}} s_k \tag{9}$$

3) **Shortest Path:** The the weight of all the edges coming into the vertices representing the candidate location k is set as s_k. The shortest path between the initial position of the UAV and the destination d_i^* is computed using the Dijkstra's shortest path algorithm [14].

The algorithm essentially treats a user's future location as an additional user location when calculating R_j. The time complexity of the algorithm therefore increases linearly with the number of users and the number of future time slots in which the predicted user locations are considered. There is a trade-off between the gain in user radio performance and the increased amount of computational time needed when considering how far to look ahead. This trade-off is analysed in the next section.

IV. Performance evaluation

In this section, the numerical results of the performance of the proposed UTO-PUL algorithm are presented. We compare the performance with the following two approaches:

1) **Non-Predictive:** The user-to-UAV association is performed at the start of a time frame. The trajectory of the UAV is designed based on the current location of the users. The users are moving but their future locations are not known and not taken into consideration to design the trajectory of the UAV
2) **DRL:** Deep Q-network (DQN) has discrete state and discrete action. The trajectory design of a single UAV is designed using DQN. The state, action and the reward are defined as follows:
   a) **State:** The channel gain information between the UAV and the ground BS, $h_{i,b}$ and UAV to users,$\mathbf{h_i} = [h_{i,1}, h_{i,2}, \ldots h_{i,J}]$, existing UAV, $[x^{uav}_{1,s}, y^{uav}_{1,s}$ and the user coordinates $\mathcal{X}^{user} = [x_{1,s}, x_{2,s}, \ldots x_{J,s}]$ are used as the state information for the DQN. The overall state information is presented as $s = [h_{i,b}, \mathbf{h_i},]$
   b) **Action:** A discrete action space is considered. The agent is trained to move in x and the y plane with a step size Δ. The x coordinate and y coordinate gets updated with step s as follows:

$$[x_{i,s}, y_{i,s}] = [x_{i,s-1}, y_{i,s-1}] + [\delta x, \delta y] \tag{10}$$

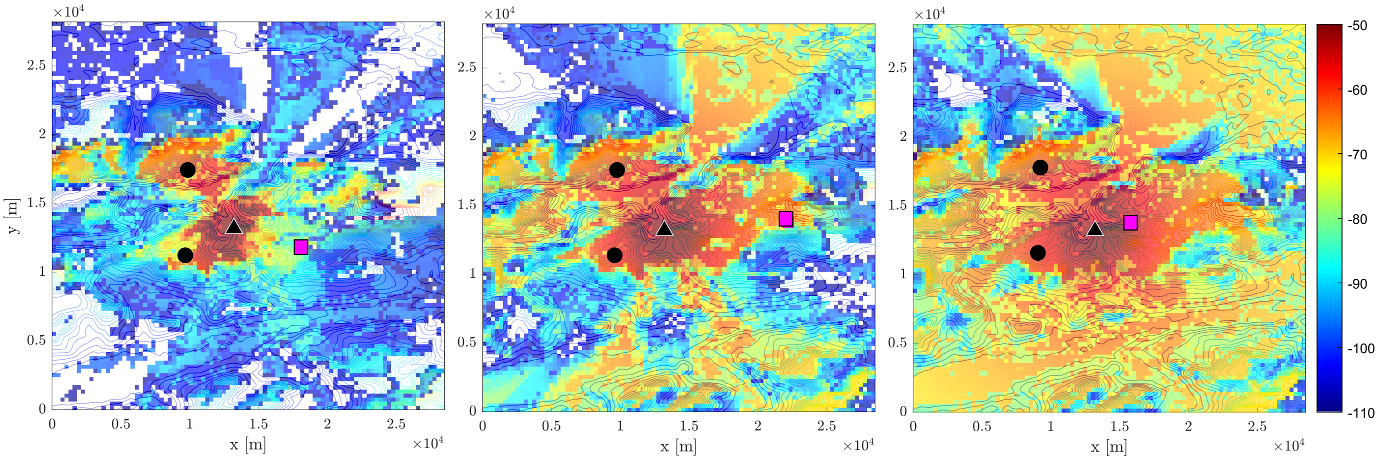
   where $[\delta x, \delta y] = \{0, \Delta\}$
   c) **Reward:**The reward is assigned to the agent on each step s. The reward for a is designed as follows

$$r_s = \min_{j \in \mathbb{J}} P^s_{i,j} \tag{11}$$

   The agent is trained using a DQN on a total of 2000 episodes.

A. Simulation Settings

For our simulations, the UAVs are providing coverage to users operating in a 28 km×28 km, in an area around the Beara Peninsula in Ireland. This location was chosen as it is a rural area with fairly mountainous terrain, which is the targeted use case environment. This area has a median and maximum elevation of 120 m and 905 m above sea level, respectively. A 182×182 grid of candidate locations is used in this map, with a 2100 m distance between the candidate locations. The location of the BS is chosen using a brute-force search approach that maximises the BS to UAV coverage. Other simulation parameters are given in Table II. It should be noted that the wireless network simulated is of a VHF two-way radio network, so all radios are assumed to have the same transmit power. Simulations were performed with varying user prediction window lengths, for different numbers of UAVs and users, and UAV altitudes. A total of 50 simulation runs were performed for each combination of parameters, and the power of the downlink signal for each user is recorded for every user, at every time frame.



(a) (b) (c)

Fig. 3: Snapshot of the simulated scenario showing the radio coverage heatmap provided by a UAV located in the center of the map, flying at (a)120 m, (b)600 m and (c)1200 m altitude.

TABLE II: Simulation parameters

Parameter	Description	Value
P_{tx}	transmit power of radios	37 dBm
f_c	carrier frequency	160 MHz
ht_i	altitude of UAV (above MSL)	120m ,600m, 1200m
(U,J)	no. of UAVs,users	(1,2) (2,4) (3,6)
V_{uav},V_{user}	velocity of UAV and user	25 m/s, 10 m/s
hr_j	height of user antenna (AGL)	1.5 m
hb	height of BS antenna (AGL)	20 m
T	simulated flight time	4000 s
T_f	time frame length	10 s

Figure 3 shows snapshots of simulation scenario, showing the coverage area provided by a UAV whose location is denoted by a black triangle, flying at different altitudes. The magenta square shows the location of the BS, optimised for the UAV's set altitude, and the two black circles are the location of the users. The heatmap shows the received downlink signal power at ground level, with the white regions indicating areas where the received signal is below the receive sensitivity of -110 dBm.

B. Performance Analysis

Impact of prediction: Fig.4 shows the cumulative distribution function (CDF) of the user downlink signal strengths with increasing prediction window lengths when there are 3 UAVs flying at 600 m altitude, serving 6 users. When the non-predictive approach is used, the probability of the user not having radio coverage (i.e. having receive powers of less than -110 dBm) is 3%. With UTO-PUL with prediction windows of 2, 4 and 8 minutes, this reduces to 1.7%, 0.8% and 0.5% respectively. This is an improvement of between 45% to 85%. Prediction also increases the median downlink power significantly from -70 dBm with no prediction to -63 dBm with 2 minute prediction. This increases to -59 dBm when the prediction window is increased to 4 and 8 minutes.

The results show that incorporating future user locations provides significant gains in performance. This is because the UAVs are able to pre-emptively position themselves in better locations when users are about to enter areas that are blocked by terrain. However, there is a limit to the gains achieved with increasingly larger prediction windows, where the benefits of additional predictive information becomes limited. This can be seen by the performance achieved with the 4 minute prediction window, which is similar with the 8 minute window, despite needing just half the computational complexity. The 4 minute window can be viewed as a more optimal configuration with a good trade-off between performance and complexity.

Impact of altitude: In Fig.5, we show the effect of different UAV altitudes on the performance with 2 UAVs serving 4 users for non-predictive and the predictive UTO-PUL algorithm with a 4 minute prediction window. At a 120 m altitude, the UAV will encounter many obstacles between itself and its users, as illustrated Fig.3, and the gains obtained with user prediction is the largest. At higher altitudes, the impact of the terrain obstacles lessens as the UAV would have more LoS connections with the users, and there are less advantages of prediction. At 1200 m, we can see that there is essentially no difference in performance between optimizations with and without prediction, and 100% coverage is provided to all users at all timeslots.

Comparison with DRL approach: Fig.6 compares the performance of the UTO-PUL algorithm with the trained DRL agent in a scenario with a UAV flying at an altitude of 600 m and serving 2 users. The UTO-PUL uses a 4 minute prediction window. We can see that UTO-PUL outperforms the DRL approach, achieving a downlink power that is 14 dB higher.

V. Conclusion and Future Work

In this paper, we considered the use of long-endurance UAVs acting as relays to extend the range of two-way radio systems in peacekeeping and humanitarian use cases. More specifically, we proposed a UAV trajectory optimization based

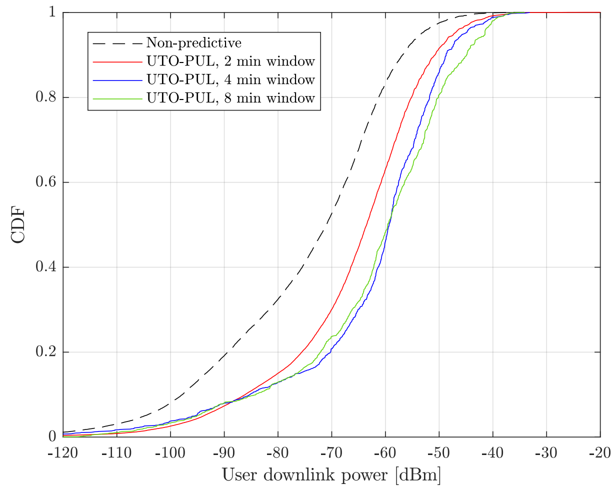


Fig. 4: CDF of user receive power for different prediction windows

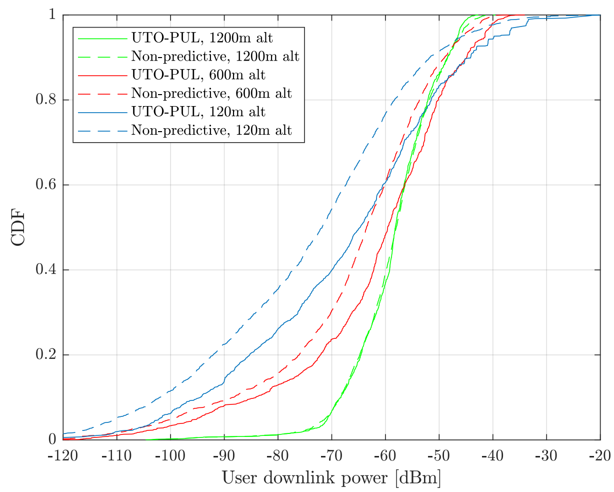


Fig. 5: CDF of user receive power for different UAV altitudes

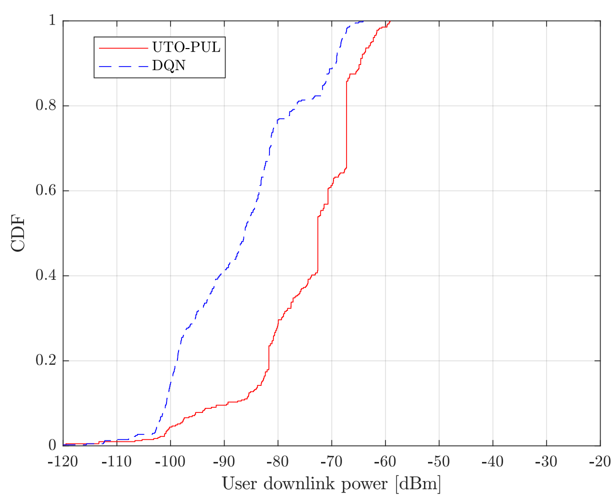


Fig. 6: CDF comparison of the proposed predictive approach with the DQN based approach.

on predicted user locations. Evaluation of the proposed technique using simulations in challenging terrain showed that the user location predictions provided significant gains over DRL-based techniques and optimizations with no predictions. The performance gains of UTO-PUL is limited when the effect of terrain blockage is lower, for example when the UAV is flying at a high altitude. The size of the prediction window also needs to be selected carefully to manage the trade-offs between computational complexity and performance gains.

In the future work, we will consider the joint-optimization of the UAV trajectory and UAV to user association, and extend the use cases to include the provision of coverage and capacity in cellular 5G networks.

Acknowledgements

This work was funded by Science Foundation Ireland under the MISTRAL project, grant number 21/FIP/DO/9949.

References

[1] G. Geraci, A. Garcia-Rodriguez, M. M. Azari, A. Lozano, M. Mezzavilla, S. Chatzinotas, Y. Chen, S. Rangan, and M. D. Renzo, "What will the future of uav cellular communications be? a flight from 5G to 6G," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 3, pp. 1304–1335, 2022.

[2] B. Galkin, L. Ho, K. Lyons, G. Celik, and H. Claussen, "Experimental Evaluation of Air-to-Ground VHF Band Communication for UAV Relays," in *IEEE International Conference on Communications (ICC)*, 2023.

[3] Y. Tang, M. H. Cheung, and T.-M. Lok, "Trajectory design for uav assisted wireless networks," in *IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[4] X. Jiang, Z. Wu, Z. Yin, W. Yang, and Z. Yang, "Trajectory and communication design for uav-relayed wireless networks," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1600–1603, 2019.

[5] N. Zhao, Y. Cheng, Y. Pei, Y.-C. Liang, and D. Niyato, "Deep reinforcement learning for trajectory design and power allocation in UAV networks," in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[6] Y. Liu, J. Yan, and X. Zhao, "Deep reinforcement learning based latency minimization for mobile edge computing with virtualization in maritime UAV communication network," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4225–4236, 2022.

[7] X. Zhong, Y. Huo, X. Dong, and Z. Liang, "Deep Q-network based dynamic movement strategy in a uav-assisted network," in *IEEE Vehicular Technology Conference (VTC2020-Fall)*, 2020, pp. 1–6.

[8] C. Wang, D. Deng, L. Xu, and W. Wang, "Resource scheduling based on deep reinforcement learning in uav assisted emergency communication networks," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 3834–3848, 2022.

[9] Z. Chang *et al.*, "Trajectory design and resource allocation for multi-uav networks: Deep reinforcement learning approaches," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2022.

[10] J. Allred *et al.*, "Sensorflock: An airborne wireless sensor network of micro-air vehicles," in *International Conference on Embedded Networked Sensor Systems*. New York, USA: Association for Computing Machinery, 2007, p. 117–129.

[11] Y. Hmamouche, M. Benjillali, and S. Saoudi, "Fresnel line-of-sight probability with applications in airborne platform-assisted communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5060–5072, 2022.

[12] N. JPL, "NASA Shuttle Radar Topography Mission Global 1 arc second dataset," *NASA EOSDIS Land Processes Distributed Active Archive Center*, 2013.

[13] M. Shimada *et al.*, "New global forest/non-forest maps from ALOS PALSAR data (2007–2010)," *Remote Sensing of Environment*, vol. 155, pp. 13–31, 2014.

[14] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.