Joint Optimization of Access and Backhaul Links for UAVs Based on Reinforcement Learning

Azade Fotouhi*, Ming Ding†, Lorenzo Galati Giordano§, Mahbub Hassan*†, Jun Li‡, and Zihuai Lin¶,
*University of New South Wales (UNSW), Sydney, Australia {a.fotouhi, mahbub.hassan}@unsw.edu.au
†Data61, CSIRO, Australia {ming.ding, mahbub.hassan}@data61.csiro.au
§Nokia Bell Labs, Dublin, Ireland {lorenzo.galati_giordano@nokia-bell-labs.com}
‡Nanjing University of Science and Technology, China {jun.li@njust.edu.cn}
¶University of Sydney, Australia {zihuai.lin@sydney.edu.au}

***Abstract*—In this paper, we study the application of unmanned aerial vehicle (UAV) base stations (BSs) in order to improve the cellular network capacity. We consider flying BSs where BS-equipments are mounted on UAVs, making it possible to move BSs freely in space. We study the optimization of UAVs' trajectory in a network with mobile users to improve the system throughput. We consider practical two-hop communications, i.e., the access link between a user and the UAV BS, and the backhaul link between the UAV BS and a macrocell BS plugged into the core network. We propose a reinforcement learning based algorithm to control the UAVs' mobility. Additionally, the proposed algorithm is subject to physical constraints of UAV mobility. Simulation results show that considering both the backhaul and access links in the UAV mobility optimization is highly effective in improving the system performance than only focusing on the access link.**

***Index Terms*—Drone Base Station, Reinforcement Learning, Trajectory Optimization, UAV.**

I. Introduction

With the recent advances in telecommunications and unmanned aerial vehicles (UAVs), it becomes feasible to develop flying BSs that could bring considerable benefits for wireless networks. For example, these days UAV BSs/Drone BSs are emerging in a broad range of applications such as agriculture [1], telecommunication [2], [3], traffic surveillance and transportation [4], [5], and etc.

Thanks to their flexibility, degree of freedom, and being autonomous, they are emerging as a potential solution to improve the performance of cellular networks, especially in the emergency scenarios. In particular, the cellular industry has recognized the importance of providing support to low-altitude drones for establishing a reliable communication. The Third Generation Partnership Project (3GPP) is studying the LTE support for aerial vehicles and exploring the challenges and the opportunities of their usage [6].

Although there are still many challenges to be solved before integrating drones into cellular networks [7], researchers are exploring their advantages in numerous use-cases. The drones can be used to optimize the coverage of mobile users in cellular networks [8], improve the network throughput by their movements [9], or complement the traditional cellular network in a multi-tier architecture [10].

Providing backhaul link and its performance is one of the main challenges in employing drone networks. Millimeter-wave [11], LTE [12], and free space optical [13] are among the proposed solutions for backhaul of drone communications. Then the drones' deployment must be designed in a way to utilize the backhaul link efficiently. Both the backhaul link and the end user link performance are considered in [14], where the coverage of a typical user in a UAV network is optimized. By considering end-to-end communication links, the UAVs adjust their heights in order to improve the probability of coverage. Zhang et al. [15] addressed a network of multiple UAVs acting as relays to help the transmission of information from a ground source to a ground destination. The end-to-end communication is considered in this work to find the optimal trajectory for the UAV relays.

The importance of both backhaul and access links in UAV networks motivates us to improve the system performance using a joint optimization model. In this paper, we consider multiple UAVs/drones as base station, serving mobile users. A reinforcement learning based algorithm is proposed to control the heading direction of drones, while considering both access and backhaul link. In the proposed model, each drone uses the information of other drones BS in the network to choose a direction that improves the system performance. Additionally, a simpler algorithm that only uses the local information in each BS is developed, where the simulation results show that adopting algorithms that are more intelligent will increase the performance. To understand the impact of joint optimization, we review the performance of the system through a simple algorithm that only focuses on maximizing the access link throughput.

This paper is structured as follows. In Section II, we present our system model and the problem formulation. In Section III, we describe our proposed reinforcement learning algorithm as well as a heuristic algorithm based on the Signal to Noise Ratio metric. Section IV presents the simulation model and discusses the performance results. Finally, our conclusion is drawn in Section V.

II. System Model

We assume that there is a set of Drone Base Stations (DBSs) denoted by $\mathcal{N}$, which are employed to improve the two-hop communications between users and a macro cell. The access link is established between users and the drone base stations, which are connected by a wireless backhaul link to a terrestrial

978-1-7281-0960-2/19/$31.00 ©2019 IEEE

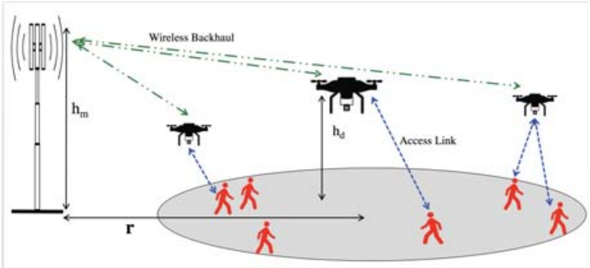

Fig. 1: A two-hop communications system model

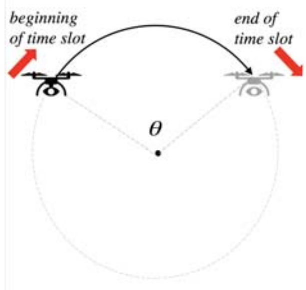

Fig. 2: Drone path while taking a turn of θ rad

macro base station (MBS). The drones are initially distributed randomly over the area of interest, at a fixed height of h_d. In the considered network, we assume that all drones fly with a constant speed of v, equipped with a single omnidirectional antenna, and transmit with a power of p_d. Therefore, a 2D mobility model is considered for drones, where they are capable of changing their heading directions while serving users. Note that the constant speed model is compatible with both rotary and fixed wing UAVs. Additionally, moving with a constant speed avoid extra energy consumption as a result of subsequent stopping and starting the flight. In this paper, f_d denotes the transmission frequency of the DBSs.

The MBS, located at a height of h_m, is only for backhaul purposes. We assume all users communicate directly with their associated DBSs, and the traffic will be offloaded to MBS through DBSs. We assume that the MBS communicates with DBSs using a fixed transmission power of p_m, at a central carrier frequency of f_m. Additionally, the macro cell is located at a distance of r from the center of the service area.

We consider a square service area with a width of w, where there are U users randomly moving according to a Random Way Point model (RWP) [16]. In the RWP model, each user selects a random destination within the area independent of other users, and moves there following a straight trajectory with a constant speed selected randomly from a given range. Upon reaching the destination, users may pause for a while before continuing to move to another destination. The pause duration is also chosen randomly from a given range. Each user follows a full buffer traffic model with continuous downloads.

Each user selects a DBS with the highest received signal strength (RSS), and it is assumed that each DBS allocates its resource uniformly to the associated users during transmission. Additionally, there is no limitation on the number of users that can be associated to a specific DBS. The set of associated users to a drone n is denoted by $\mathcal{Q}_n$. The system model is illustrated in Figure 1.

The path loss associated with distance of d for a communication link can be modeled as [17]

$$\eta(d) = \begin{cases} A^L d^{-\alpha^L}, & \text{with a probability of } P^L(d) \\ A^{NL} d^{-\alpha^{NL}}, & \text{with a probability of } P^{NL}(d) \end{cases} \quad (1)$$

where d is the 3D distance between a transmitter and a receiver pair. A^L and A^{NL} are the path losses at a reference distance ($d = 1$ m) for the Line of Sight (LoS) and the Non Line of Sight (NLoS) cases, respectively. α^L and α^{NL} are the path loss exponents for the LoS and the NLoS cases, respectively. These are all obtainable from field tests [18].

The main goal of the drone movements is to provide high quality services for the mobile users, taking into account the two-hop communications. Therefore, the trajectories of the DBSs need to be optimized.

To this end, we discretized the time into equally spaced time slots, in which the best heading direction of each drone should be found. By following the best heading directions at each time slot, the optimized trajectory of each DBS can be achieved. The length of each time slot is chosen to be sufficiently small to react to the users' movement and network changes, and sufficiently large so that drones can follow a heading direction and move forward. Note that choosing a heading direction affects the distances from the DBS to both the users and the MBS. It indicates that the location of each DBS at any time slot is a function of its previous location.

It is important to note that when the drone is turning to change its direction with a constant linear speed v, the drone‘s path will follow an arc of a circle as shown in Figure 2 [19]. Given the 2D location of a drone n at the start of current time slot denoted by $[x_n^t, y_n^t]$, then by taking a turn of θ_n rad, the drone location at the end of time slot (t), starting the following time slot $(t+1)$, can be calculated by

$$\begin{gathered} \begin{bmatrix} x_n^{t+1} \\ y_n^{t+1} \end{bmatrix} = R \times \begin{bmatrix} x_n^t - X_n^t \\ y_n^t - Y_n^t \end{bmatrix} + \begin{bmatrix} X_n^t \\ Y_n^t \end{bmatrix}; \\ R = \begin{bmatrix} \cos(\theta_n) & -\sin(\theta_n) \\ \sin(\theta_n) & \cos(\theta_n) \end{bmatrix}, \end{gathered} \quad (2)$$

where R is the rotation matrix, and $[X_n^t, Y_n^t]$ denotes the coordinates of the circle centre of the segment. The drone location at $t+1$ is denoted by $[x_n^{t+1}, y_n^{t+1}]$. Note that θ_n needs to satisfy the drone maximum turning angle.

To simplify the presentation of the above equation, we denote the location of drone at any time as a function of heading direction and the its previous location, as follows:

$$\mathcal{L}_n^{t+1} = f(\mathcal{L}_n^t, \theta_n^t), \quad (3)$$

where $\mathcal{L}_n^t = (x_n^t, y_n^t)$ represents the 2D location of DBS n at time slot t, and θ_n^t shows its heading direction which is taken at the start of time slot t. As a result, the capacity of each link highly depends on the drones' decision of mobility. In the rest of this paper, we drop the time slot indicator for simplicity.

The capacity can be formulated according to the Shannon Theorem as

$$\mathcal{C} = \log_2(1 + SINR) = \log_2(1 + \frac{S}{I + N_0}), \quad (4)$$

where N_0 denotes the additive white Gaussian noise power at each user. The received signal power, S, varies based on the drone decision, as follows

$$S = p_{tx}\eta(d)g, \quad (5)$$

where g is the channel gain, modeled as an exponential random variable with a mean of one, and p_{tx} is the sender's transmission power. Additionally, I in equation (4) represents the aggregated interference from other transmitters. We consider orthogonal frequency allocation between the backhaul and the access links, which means that we do not have any interference between these two links. All drones, however, use the same frequency band, thus creating the potential for interference. As a result, I for the access link between a DBS n and a user can be modeled as:

$$\sum_{i \in \mathcal{N}, i \neq n} p_d \eta(d) g \quad (6)$$

where d is the distance from the user to the interfering drone base stations. The above equations show that the movements of each drone not only impacts the signal strength of their associated users, but also it has a considerable influence on the interference signal on other users in the network.

Denoting the capacity of access link by $\mathcal{C}_a$, and the capacity of backhaul link by $\mathcal{C}_b$, then the overall capacity that can be delivered to the associated users is defined as the minimum of them.

Therefore, the problem of finding the best heading directions for drones at each time slot can be formulated as:

$$\begin{aligned} \theta^* &= \arg\max \big(\min(\mathcal{C}_a, \mathcal{C}_b) \big) \\ s.t. \quad \theta^* &\in [-\theta_{max} \quad \theta_{max}] \end{aligned} \quad (7)$$

where, without loss of generality, we assume that the range of heading directions that a drone can select is bounded between $-\theta_{max}$ and θ_{max} due to physical limitations of UAVs [20], [19].

III. Proposed Algorithms

In this section, we propose algorithms to optimize the system capacity in a network including several mobile users, and mobile DBSs.

A. Reinforcement Learning Method

Reinforcement learning is a model in which the learner or the agent must discover which actions yield the most reward by trying them. The agent can use its experience to improve its performance over time. The agents are connected to the environments and make decisions by taking actions. Despite the uncertainty in the environment, the agents always pursue achieving a goal [21].

In our scenario, the DBSs which are moving are the agents. The environment consists of locations of users, the MBS, and neighboring DBSs. And finally, the DBSs moving directions are the actions taken in the reinforcement learning model. For any DBS decision on heading direction, there is a corresponding reward.

The trade-off between exploration and exploitation is the most important challenge of reinforcement learning models. An agent can earn more rewards by exploiting previous experiences, but exploration will help make better decisions in the future. The estimated average associated users' performance of each DBS is defined as the reward of an agent action, hence, the policy of each agent is to maximize the average user performance. Given the state of users, MBS and other DBSs, each DBS makes a decision to improve its performance. To this end, each DBS considers the history of other DBSs to make more intelligent decision. We suppose that the drones can communicates with other drones and send their heading direction when there is a change. In this model, each DBS assumes that there is a high probability for other agents to follow their previous directions. It can be expressed that a DBS estimates the location of other drones for the next time slot, using their previous direction:

$$\mathcal{L}_n^{t+1} = f(\mathcal{L}_n^t, \theta_n^{t-1}). \quad (8)$$

Note that considering possibilities of changing the previous direction and taking a new heading direction is left for future work. That will help each drone decide more wisely, however, the time complexity of the reinforcement learning algorithm will increase. In the presented model in this paper, each drone assumes the probability of zero for taking a new direction for the other DBSs.

Given the location of other DBSs following their previous decisions, the best direction that optimizes the average user performance is calculated through brute force search for each DBS. All DBSs then follow their selected directions to achieve the estimated performance, while they have no information of other drones' decisions. The actual user performance then could be different from the estimated one, as there is a possibility that drones won't follow their previous directions. The actual performance can be computed according to the final decisions of all drones. The difference between the actual performance and the reward of selected policy (estimated performance) is defined as *regret.*

As a result of DBSs' movements, and the mobility of users, the state of environment changes for the following time slot. DBSs share their last taken heading directions with each other.

The reward of DBS n can be defined as follows

$$
\begin{aligned}
\mathcal{R}_n =& \arg\max_{\theta_n} \frac{\sum_{u\in\mathcal{Q}_n} min(\mathcal{C}_a,\mathcal{C}_b)}{|\mathcal{Q}_n|}, \quad \forall\ n \in \mathcal{N} \\
\mathcal{C}_a =& \log_2(1+\frac{S_n}{N_0+\sum_{i\in\mathcal{N} i\neq n} p_d g \eta(d(\theta^{t-1}))}) \\
\mathcal{C}_b =& \frac{1}{|\mathcal{N}|}\log_2(1+\frac{S_m}{N_0}) \\
s.t. \quad & \theta_n \in [-\theta_{max} \quad \theta_{max}]
\end{aligned} \tag{9}
$$

where $d(\theta^{t-1})$ denotes the distance of a DBS from users assuming that it moves forward the previous direction (θ^{t-1}). Additionally, the actual system performance, will be calculated according to the final decision of each DBS at the end of time slot t. It means that we need to replace $d(\theta^{t-1})$ by $d(\theta^t)$ in equation (9).

Given the above description, an agent requires the following information to calculate the reward, and make decision:

- The location of MBS for calculating the backhaul performance.
- The location of the associated users to evaluate the signal power.
- The location of other DBSs and their previous directions, in order to compute their estimated location and the interference signal power

The detail of the reinforcement learning model is described by **Algorithm 1.**

Algorithm 1 RL Algorithm for DBS n

procedure
  for each timleslot t **do**
    $\mathcal{L}_i^{t+1} \leftarrow f(\mathcal{L}_i^t, \theta_i^{t-1})), i \in \mathcal{N}, i \neq n$
    for each possible direction θ **do**
      $\mathcal{C}_b \leftarrow \frac{1}{|\mathcal{N}|}\log_2(1+\frac{p_m g \eta(d(\theta))}{N_0})$
      for each user $u \in \mathcal{Q}_n$ **do**
        $C_u[u] \leftarrow Min(C_a(u,n), C_b)$
      end for
      $C[g] \leftarrow Sum(C_u)/|\mathcal{Q}_n|$
    end for
    $Performance[t] \leftarrow Max(C)$
    $\theta_n^t \leftarrow Index(Max(C))$
  end for
end procedure

B. SNR based

In this algorithm, only the knowledge of the associated users at each DBS is used for choosing heading directions. Each drone neither requires DBS-to-DBS communications, nor does it need to know the location of other users in the network.

Using only the knowledge of its own associated users, each DBS calculates the access link *SNR* (Signal to Noise Ratio) for every associated user and every candidate moving direction. Additionally, for each candidate moving direction, the SNR of backhaul link is also calculated. The minimum of the SNRs for backhaul and access link of candidate directions is then chosen to compute user capacity. After examining all the possible directions, the direction that provides the maximum average user capacity over all associated users is chosen as the DBS's heading direction in the next time slot. This algorithm is formulated in equation (10), where θ_n represents the selected direction for drone n. SNR_a, and SNR_b represent the SNR of access link and backhaul link, respectively.

$$
\begin{aligned}
\theta_n =& \arg\max \frac{\sum_{u\in\mathcal{Q}_n} min(SNR_a, SNR_b)}{|\mathcal{Q}_n|} \quad \forall\ n \in \mathcal{N} \\
s.t. \quad & \theta_n \in [-\theta_{max} \quad \theta_{max}]
\end{aligned} \tag{10}
$$

TABLE I: parameters and their adopted value

parameters	DBS/access link	MBS/backhaul
Height (h)	10 m	20 m
Frequency (f)	2 GHz	1.8 GHz
Transmission power (p)	24 dBm	30 dBm
A^L	$10^{-3.24-2log(f)}$	$10^{-2.8-2log(f)}$
N_0	-95 dBm	-95 dBm
A^{NL}	$10^{-2.24-2.13log(f)}$	–
α^L	2.1	2.2
α^{NL}	3.53	–
Area size (w)	500 m	
MBS distance(r)	[250 m 4000 m]	
Number of Users (U)	100	
Number of DBSs ($\lvert N\rvert$)	20	
User speed	[0.2 5] m/s	
Time slot duration	5 sec	

IV. Simulation

To understand the impact of two-hop communications on the network performance, and to compare the potential of proposed algorithms, we used Python to conduct computer simulations.

We choose a low altitude for our DBSs, as it was shown that due to high interference, reducing the height of drones will improve the performance [22]. Hence, the recommended height of 10 m [23] is selected for the DBSs in this section. Given the height of DBSs, the following probabilistic Line of Sight (LoS) model is employed to evaluate the probability of having a LoS connection between a user and a drone [6]:

$$
P^L = \begin{cases} 1, & d_{2D} \leq 18m \\ \frac{18}{d_{2D}} + \exp\Big(-\frac{d_{2D}}{36}\Big)\Big(1-\frac{18}{d_{2D}}\Big), & d_{2D} > 18m \end{cases} \tag{11}
$$

where d_{2D} denotes the distance between the ground user and the projection of drone location onto the ground. Clearly, the probability of having a NLoS (Non Line of Sight) connection is $P^{NL} = 1 - P^L$. Additionally, we assume that P^L is equal to 1 for the backhaul link.

In the considered area, with the width of 500 m, there exists 100 users moving according to RWP model with the speed range of [0.2-5] m/s. Initially, 20 DBSs are located at random locations to serve users during their movements. The parameters are as shown in Table I.

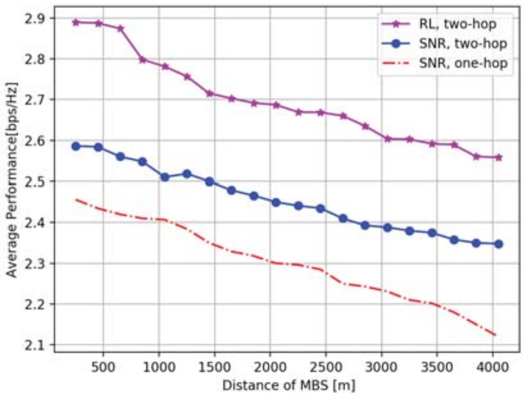

Fig. 3: performance of one-hop and two-hop algorithms as a function of MBS distance

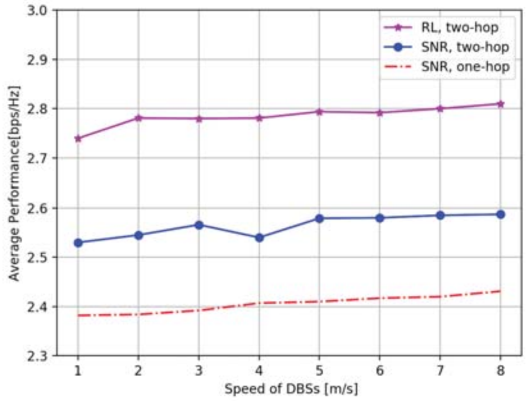

Fig. 4: performance of one-hop and two-hop algorithms as a function of DBSs' speed

To average out the randomness, we repeated the simulation over 20 times independently, each for a period of 500 seconds, and averaged the results to obtain an unbiased numerical result.

Moreover, we study optimizing the drones movement considering one-hop communication (access link), as the reference model. In this model, each drone only evaluates the performance of access link and changes its direction in a way to maximize the access link performance. In more detail, for the candidate directions, the average access link SNR of the associated users of each DBS are calculated. The heading direction that maximizes the average SNR of access link, ignoring the backhaul link, is selected for the next movement of each drone. It is worth noting that neither of two-hop SNR-based and one-hop SNR based models examines the interference to choose a heading direction.

We first evaluate the performance of the two-hop reinforcement learning and two-hop SNR models, and compare the results with the one-hop SNR model. The main goal of this comparison is to quantify the impact of considering joint optimization of the backhaul and access links on the system performance. In this study, all drones move at a fixed speed of 4m/s, while serving users. We discretize all possible heading directions into a finite set of 36 different heading directions, which are separated by 10 degrees. Each drone can turn at most 50 degrees. For example, if a drone is moving towards $0°$, it can choose a direction that belongs to $[-50°, -40°, \ldots, 0°, \ldots, +40°, +50°]$ for the next time slot. The distance of MBS, denoted by r, varies from 250 m of the center of considered network to 4000 m.

According to Figure 3, we observe the following results:

- Considering both the access link and the backhaul link, the drones can make more intelligent decisions, resulting in a higher performance. Taking into account the access link only will force the DBSs to fly closer to the users in order to improve the access link SNR, however, such movement will cause them moving away from the MBS. Increasing the distance between MSB and DBSs reduces the backhaul performance, and subsequently, decreases the end-to-end users performance. The performance gain of two-hop reinforcement learning model to one-hop reaches up to 17% according to Figure 3.
- We can observe that the reinforcement learning algorithm can achieve a better performance than the two-hop SNR maximization algorithm. An improvement of up to 10% in terms of the capacity performance can be obtained which is the result of using the history of drones' movements, and considering the interference power for controlling the drones' mobility.
- Another important note is that the system performance reduces as the distance of MBS increases. This is the direct result of the deteriorating signal power across large distances. However, it remains true that the two-hop communication algorithms outperform the one-hop one with a large MBS distance.

Additionally, to evaluate the impact of drone's speed, we display the results for different speeds, namely 1m/s to 8 m/s. The MBS is located at the distance of 1000 m from the center of considered network. The results are plotted in Figure 4. Not surprisingly, the system shows a higher performance with faster drones. It is because that drones are able to follow the environment changes faster and move to the selected location quicker. Additionally, similar to Figure 3, we can observe that the two-hop models outperform the one-hop model.

As we have mentioned in section III-A, the difference between the reward and the actual performance is measured as regret. Figure 5 exhibits the regret metric during the simulation.

The regret value, as shown in Figure 5, indicates that there is potential room for improving the reinforcement learning algorithm to increase the network performance. As mentioned in Section III-A, in this paper, each drone considers a probability of 1 for other DBSs to retain its previous direction, and assumes a probability of zero for any other direction. Following this assumption, each DBS selects a direction that

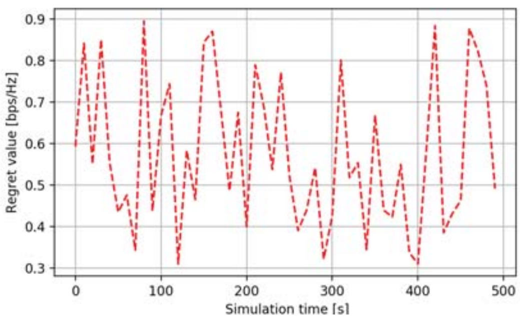

Fig. 5: regret during the simulation time.

results in a higher performance. A higher regret represents a higher error in predicting the movement of other drones. By revisiting this assumption, we may be able to reduce the regret value and enhance the system performance.

V. Conclusion

Based on reinforcement learning, we have solved a joint optimization problem where both the access link and the backhaul link are considered to design the trajectory of drone base stations. Simulation experiments have confirmed that the proposed joint optimization can significantly improve the UAV network performance compared to optimizations that consider only the access links. In our current work, we have assumed only a LoS model for backhaul links and leave the more advanced models that incorporate both LoS and NLoS as future work. The investigation of more advanced states and probabilities in the reinforcement learning algorithm is also left for future research.

References

[1] U. R. Mogili and B. B. V. L. Deepak, "Review on application of drone systems in precision agriculture," *International Conference on Robotics and Smart Manufacturing (RoSMa2018)*, vol. 133, pp. 502 – 509, 2018.

[2] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, May 2016.

[3] A. Fotouhi, M. Ding, and M. Hassan, "Dynamic base station repositioning to improve performance of Drone small cells," in *2016 IEEE Globecom Workshops*, Dec 2016, pp. 1–6.

[4] N. H. Motlagh, M. Bagaa, and T. Taleb, "UAV-Based IoT Platform: A Crowd Surveillance Use Case," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128–134, February 2017.

[5] E. P. Karan, C. Christmann, M. Gheisari, J. Irizarry, and E. N. Johnson, "A comprehensive matrix of unmanned aerial systems requirements for potential applications within a department of transportation," in *Construction Research Congress 2014: Construction in a Global Network*, 2014, pp. 964–973.

[6] 3GPP Technical Report 36.777, "Technical Specification Group Radio Access Network, Enhanced LTE support for aerial vehicles, Release 15 ," Dec. 2017.

[7] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV Cellular Communications: Practical Aspects, Standardization Advancements, Regulation, and Security Challenges," *IEEE Communications Surveys Tutorials (Early Access)*, pp. 1–1, 2019.

[8] H. Huang and A. V. Savkin, "An algorithm of efficient proactive placement of autonomous drones for maximum coverage in cellular networks," *IEEE Wireless Communications Letters*, pp. 1–1, 2018.

[9] A. Fotouhi, M. Ding, and M. Hassan, "Flying drone base stations for macro hotspots," *IEEE Access*, vol. 6, pp. 19 530–19 539, 2018.

[10] S. Sekander, H. Tabassum, and E. Hossain, "Multi-Tier Drone Architecture for 5G/B5G Cellular Networks: Challenges, Trends, and Prospects," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 96–103, MARCH 2018.

[11] Z. Xiao, P. Xia, and X. Xia, "Enabling uav cellular with millimeter-wave communication: potentials and approaches," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 66–73, May 2016.

[12] B. Galkin, J. Kibilda, and L. A. DaSilva, "Backhaul for low-altitude uavs in urban environments," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.

[13] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M. Alouini, "Fso-based vertical backhaul/fronthaul framework for 5g+ wireless networks," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 218–224, Jan 2018.

[14] B. Galkin, J. Kibilda, and L. A. DaSilva, "A stochastic geometry model of backhaul and user coverage in urban UAV networks," *CoRR*, vol. abs/1710.03701, 2017. [Online]. Available: http://arxiv.org/abs/1710.03701

[15] G. Zhang, H. Yan, Y. Zeng, M. Cui, and Y. Liu, "Trajectory optimization and power allocation for multi-hop uav relaying communications," *IEEE Access*, vol. 6, pp. 48 566–48 576, 2018.

[16] X. Ge, J. Ye, Y. Yang, and Q. Li, "User mobility evaluation for 5g small cell networks based on individual mobility model," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 528–541, March 2016.

[17] H. Li, M. Ding, D. Lpez-Prez, A. Fotouhi, Z. Lin, and M. Hassan, "Performance Analysis of the Access Link of Drone Base Station Networks with LoS/NLoS," *INISCOM*, August 2018.

[18] 3GPP, "3GPP TR 36.828, further enhancements to LTE time division duplex for downlink-uplink interference management and traffic adaptation," Tech. Rep., 2012.

[19] A. Fotouhi, M. Ding, and M. Hassan, "DroneCells: Improving 5G Spectral Efficiency using Drone-mounted Flying Base Stations," *ArXiv e-prints*, Jul. 2017.

[20] A. Fotouhi, M. Ding, and M. Hassan, "Understanding Autonomous Drone Maneuverability for Internet of Things Applications," in *2017 IEEE WoWMoM Workshops*, 2017.

[21] R. S. Sutton, A. G. Barto *et al.*, *Reinforcement learning: An introduction*. MIT press, 1998.

[22] M. M. Azari, F. Rosas, A. Chiumento, and S. Pollin, "Coexistence of terrestrial and aerial users in cellular networks," pp. 1–6, Dec 2017.

[23] M. Ding and D. L. Perez, "Please Lower Small Cell Antenna Heights in 5G," in *2016 IEEE GLOBECOM*, Dec 2016, pp. 1–6.