Ahead-Me Coverage (AMC): On Maintaining Enhanced Mobile Network Coverage for UAVs

Hamed Hellaoui[1], Bin Yang[2], Tarik Taleb[3], and Jukka Manner[1]

[1]Aalto University, Communications and Networking Department, Finland. Email: {firstname.lastname}@aalto.fi
[2]Chuzhou University, School of Computer and Information Engineering, China. Email: yangbinchi@gmail.com
[3]University of Oulu, Centre for Wireless Communications, Finland. Email: tarik.taleb@oulu.fi

***Abstract*—This paper proposes the concept of Ahead-Me Coverage (AMC) aiming to get the coverage of a cellular network ahead of the mobile users for maintaining enhanced Quality-of-Service (QoS) in cellular-connected unmanned aerial vehicle (UAV) networks. In such networks, each base station (BS) with an intelligent logic can automatically tilt the direction of its radio antennas based on the trajectory of UAVs. For this purpose, we first formulate AMC as an integer optimization problem for maximizing the minimum transmission rate of UAVs by jointly optimizing the angles of the different radio antenna, the resource allocation and the selection of the appropriate serving BS for the UAVs throughout their path. For this complex optimization problem, we then propose a solution based on Deep Reinforcement Learning (DRL) to solve it. Under this solution, we adopt a multi-heterogeneous agent-based approach (MHA-DRL) including two types of agents, namely the UAV agents and the BS agents. Each agent implements an Advantage Actor Critic (A2C) to learn optimal policies. Specifically, the BS agents aim to tilt their antennas to get ahead of the UAVs throughout their mobility, and the UAV agents target selecting the appropriate serving BSs along with resource allocation. Performance evaluations are presented to validate the effectiveness of the proposed approach.**

***Index Terms*—Ahead-Me Coverage (AMC), Cellular Networks, Unmanned Aerial Vehicles (UAVs), Deep Reinforcement Learning (DRL), Multi-Heterogeneous Agent-based DRL (MHA-DRL).**

I. Introduction

The next generation of mobile networks are envisioned to ensure limitless connectivity with higher throughput, lower relay and stronger security. It is notable that 5G/6G mobile networks are expected to be the key infrastructure enabler to support a wide range of applications, such as virtual/augmented reality, the Internet of Things (IoT), autonomous vehicles, and unmanned aerial vehicles (UAVs). Such networks are deployed to mainly serve ground users. This poses a new challenge on the mobile networks to maintain the enhanced Quality-of-Service (QoS) for UAV users flying in the sky.

This paper introduces the concept of Ahead-Me Coverage (AMC), where the goal is to get network coverage ahead of mobile users. Specifically, we consider the scenario of UAVs using a cellular network to ensure network connectivity. Field evaluations have shown that the radio antennas of the base stations (BSs) are usually tilted to serve ground users, which might not be advantageous for flying UAVs. one way to overcome this is by dynamically tilting the direction of the radio antennas in a way to anticipate the mobility of the users and ensure better coverage. As shown in Fig. 1, a radio antenna of BS equipped with an electrical motor can take different position by tilting its angle, which leads to the change of the radio coverage area.

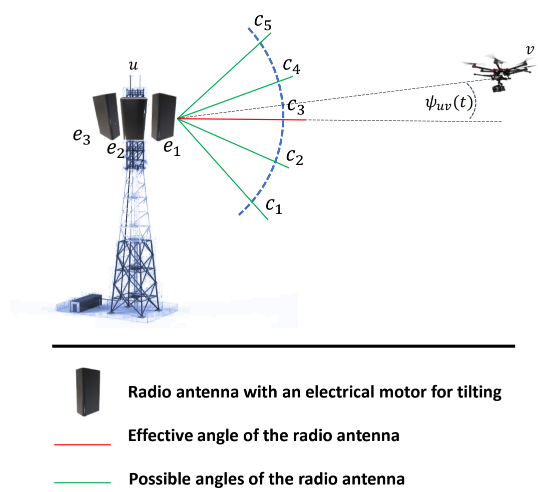


Fig. 1: Tilting the radio antenna: a radio antenna with an electrical motor can take several positions by tilting its angle.

In the literature, different works have been proposed to enhance the QoS of mobile users. However, it has not been investigated for the problem of automatically adjusting the antennas to maintain the radio coverage ahead of the flying UAVs. To the best of the authors' knowledge, this is the first work to propose the AMC concept. To this end, we first formulate AMC as an integer optimization problem for maximizing the minimum transmission rate of UAVs by jointly optimizing the angles of the different radio antenna, the resource allocation and the selection of the appropriate serving BS for the UAVs throughout their path. This is a a non-linear and non-convex optimization problem, and then we propose a solution based on Deep Reinforcement Learning (DRL) to solve it. Under this solution, we adopt a multi-heterogeneous agent-based approach (MHA-DRL) including two types of agents, namely the UAV agents and the BS agents. Each agent implements an Advantage Actor Critic (A2C) to learn optimal policies. Specifically, the BS agents aim to tilt their antennas to get ahead of the UAVs throughout their mobility, and the

UAV agents target selecting the appropriate serving BSs along with resource allocation.

The rest of this paper is organized as follows. Section II presents related works. We present the system model and problem formulation in Section III. The proposed MHA-DRL approach for the AMC is introduced in Section IV. Performance evaluations are provided in Section V. Finally, Section VI concludes this paper.

II. Related Works

Different works have been proposed in the literature to enhance the QoS for mobile users connected to cellular networks. In [1], the authors proposed Follow-Me Cloud (FMC), i.e., a design tailored to user mobility. The framework enables services to follow the mobility of the users based on the implementation of service migration between cloud servers. In [2], the authors provided analysis of the handover procedure in the FMC scheme. Service migration to follow the mobility of users was also the focus of standardization bodies. The European Telecommunications Standards Institute (ETSI) provided a specification for end-to-end Multi-Access Edge Computing application mobility support in a multi-access edge system. The underlying specifications are documented in the group specification ETSI GS MEC 021 [3].

Besides supporting the mobility of users in terms of migrating the associated application between cloud servers, the interest has also been manifested in terms of enhancing the coverage of the cellular network. In UAV applications, this has mainly been materialized in optimizing the deployment of UAVs acting as flying base stations (UAV-BSs). In [4], the authors investigated the problem of energy-efficient 3D placement of a UAV-BS that is capable of tilting its directional antenna. In another work [5], the authors investigated the problem of energy-efficient UAV-BS coverage deployment, which includes coverage maximization and power control. The authors in [6] interested in using a UAV to detect coverage hole, and then to act as UAV-BS to server users affected by the detected coverage hole. The article [7] also investigated the issue of 3D deployment of a heterogeneous set of UAV-BSs that provide maximum wireless coverage for ground users in a given geographical area. In another work [8], the authors considered the use of UAV-BSs to provide coverage for vehicles entering a highway that is not covered by other networks. The proposed solution uses DRL to find the trajectories that would ensure effective coverage.

Some works have also considered the effect of the elevation angle on users [9], [10]. Nevertheless, the issue of implementing an intelligent logic allowing to automatically tilt the radio antenna depending on the mobility of the users has not been investigated in the literature. This underpins the focus of this article, whereas we advance the concept of AMC that aims to get the coverage of a cellular network ahead of mobile users to maintain enhanced QoS. The next section introduces the the system model and the problem formulation.

III. System Model and Problem Formulation

This section introduces a formulation of the AMC problem. We consider a mobile network providing network connectivity to flying UAVs. We focus in this paper on enhancing the QoS for UAVs. Without losing in generality, this can also be extended to enhancing the QoS for ground UEs. We use the notations $\mathcal{U}$ and $\mathcal{V}$ to denote the set of BSs and the set of UAVs, respectively. Each BS $u \in \mathcal{U}$ has a set $\mathcal{B}$ of RBs and is equipped with a number of radio antennas. We denote by $\mathcal{E}$ the set radio antennas for a BS u. As shown in Fig. 1, the radio antenna of a BS can take several positions by changing its angle. We denote by $\mathcal{C} = [c_1, \ldots, c_C]$ the C ordered positions for a radio antenna $e \in \mathcal{E}$.

We consider the downlink scenario in which data is transmitted from the BSs of the mobile network to the flying UAVs. Let us denote by $p_u = [p_u^b]_{b \in \mathbf{B}(u,v,t)}$ and $h_{uv}(t) = [h_{uv}^b(t)]_{b \in \mathbf{B}(u,v,t)}$ the vectors of the transmission power and the channel gain between the BS $u \in \mathcal{U}$ and the UAV $v \in \mathcal{V}$ over the set of selected RBs $\mathbf{B}(u,v,t)$ at timestamp t. The channel gain depends on the angle of the radio antenna and the position of the UAV. It can be approximately modeled as [4]

$$h_{uv}^b(t) = \begin{cases} g_{uv}^b(t) & \text{if } \psi_{uv}(t) \leq \phi, \\ g \approx 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $g_{uv}^b(t)$ is the gain from the main lobe of the antenna. Each deployed BS in the network operates an Orthogonal Frequency Division Multiple Access (OFDMA) technique, and thus intra-cell interference is neglected. The transmission rate between the BS u and the UAV v can be expressed as

$$\begin{aligned} r_v(t) &= \sum_{b \in \mathbf{B}(u,t)} r_{uv}^b(t) \\ &= \sum_{b \in \mathbf{B}(u,t)} W \log_2 \left(1 + \frac{p_u^b h_{uv}^b(t)}{I_{uv}^b(t) + W N_0} \right), \end{aligned} \tag{2}$$

where W refers to the bandwidth of a RB, $I_{uv}^b(t) = \sum_{u' \in \mathcal{U} \setminus \{u\}} p_{u'}^n h_{u'v}^n(t)$ is the interference impact from non-serving BSs, and N_0 represents the noise power.

The trajectory of the UAVs is predefined and we aim to optimize the network in such a way that the coverage is ensured throughout their paths. The availability of a predefined paths is due to the fact that operating UAVs requires sending an operational flight plan to a traffic regulation entity (e.g., Unmanned aerial system Traffic Management - UTM) for validation, which can also be communicated to the mobile network to ensure the required QoS. As shown in Fig. 2, based on the the trajectory of the UAV, the BS adjusts the angle of their radio antennas. In Fig. 2 (a), u_1 has directed a radio antenna to serve the UAV, while u_2 is preparing. In Fig. 2 (b), u_1 and u_2 directed their antennas to serve the UAV, while u_3 is preparing. The later has directed its antenna to serve the UAV when it reached it range, as shown in Fig. 2 (c). Although

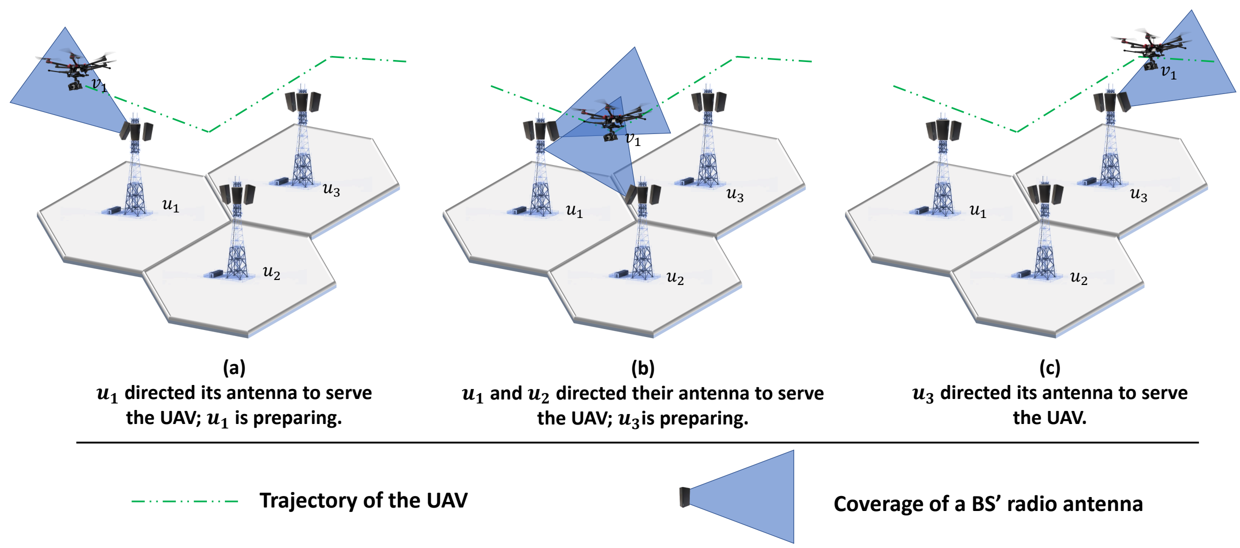


Fig. 2: Illustration of AMC principle: BSs dynamically adjust the direction of their radio antennas so to cover the trajectory of the UAV.

Fig. 2 illustrates the case of one UAV, we consider in this paper several UAVs, which makes the optimization more complex.

Let us denote by $\mathcal{T} = [1, \ldots, T]$ the T timestamps in which the optimization should be performed. In order to characterize the angle taken by a radio antenna $e \in \mathcal{E}$ of the BS $u \in \mathcal{U}$ at timestamp $t \in \mathcal{T}$, we define the boolean variable $\mathcal{X}^t_{u,e,c}$ as

$$\mathcal{X}^t_{u,e,c} = \begin{cases} 1 & \text{if the BS } u \in \mathcal{U} \text{ chooses the angle } c \in \mathcal{C} \text{ for its radio antenna } e \text{ at timestamp } t \in \mathcal{T}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

As the UAV is moving, the associated serving BS will change throughout the path in such a way to maintain an optimized QoS. In order to characterize the the selected BS for the UAV v, we define the boolean variable $\mathcal{Y}^t_{v,u}$ as

$$\mathcal{Y}^t_{v,u} = \begin{cases} 1 & \text{if the UAV } v \in \mathcal{V} \text{ connects to the BS } u \in \mathcal{U} \text{ at timestamp } t \in \mathcal{T}, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Furthermore, changing the serving BS along with the mobility of the UAVs implies an adequate selection of the RBs. In order to characterize the RB $b \in \mathcal{B}$ assigned to the UAV $v \in \mathcal{V}$, we define the boolean variable $\mathcal{Z}^t_{v,b}$ as

$$\mathcal{Z}^t_{v,b} = \begin{cases} 1 & \text{if the UAV } v \in \mathcal{V} \text{ uses the RB } b \in \mathcal{B} \text{ at timestamp } t \in \mathcal{T}, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Based on the previous variables and the expression of the transmission rate, we can formulate the AMC as follows:

$$\underset{\{\mathcal{X}^t_{u,e,c}\},\{\mathcal{Y}^t_{v,u}\}}{\text{maximize}} \min_{v \in \mathcal{V}} \sum_{t=1}^{T} \sum_{u \in \mathcal{U}} \mathcal{Y}^t_{v,u} \sum_{e \in \mathcal{E}} \sum_{c=c_1}^{c_C} \mathcal{X}^t_{u,e,c} \sum_{b \in \mathcal{B}} \mathcal{Z}^t_{v,b} r_v(t), \tag{6}$$

s.t.

$$\forall u \in \mathcal{U}, \forall v \in \mathcal{V}, \forall t \in \mathcal{T}; \quad \mathcal{Y}^t_{v,u} \in \{0,1\}, \tag{7}$$

$$\forall u \in \mathcal{U}, \forall e \in \mathcal{E}, \forall c \in \mathcal{C}, \forall t \in \mathcal{T}; \quad \mathcal{X}^t_{u,e,c} \in \{0,1\}, \tag{8}$$

$$\forall v \in \mathcal{V}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \mathcal{Z}^t_{v,b} \in \{0,1\}, \tag{9}$$

$$\forall v \in \mathcal{V}, \forall t \in \mathcal{T}; \quad \sum_{u \in \mathcal{U}} \mathcal{Y}^t_{v,u} = 1, \tag{10}$$

$$\forall u \in \mathcal{U}, \forall e \in \mathcal{E}, \forall t \in \mathcal{T}; \quad \sum_{c \in \mathcal{C}} \mathcal{X}^t_{u,e,c} = 1, \tag{11}$$

$$\forall u \in \mathcal{U}, \forall e \in \mathcal{E}, \forall c_i \in \mathcal{C}, \forall t \in \mathcal{T}; \quad \mathcal{X}^t_{u,e,c_i} = 1 \implies \mathcal{X}^{t+1}_{u,e,c_i} + \mathcal{X}^{t+1}_{u,e,c_{i-1}} + \mathcal{X}^{t+1}_{u,e,c_{i+1}} = 1, \tag{12}$$

$$\forall v \in \mathcal{V}, \forall t \in \mathcal{T}; \quad \sum_{b \in \mathcal{B}} \mathcal{Z}^t_{v,b} \geq 1, \tag{13}$$

$$\forall u \in \mathcal{U}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \sum_{v \in \mathcal{V}} \mathcal{Y}^t_{v,u} \mathcal{Z}^t_{v,b} \leq 1. \tag{14}$$

The objective of the above optimization problem is to maximize the minimum transmission rate for the UAVs $v \in \mathcal{V}$ throughout their trajectories (as expressed in (6)). Conditions (7), (8) and (9) limit the values of the boolean variables $\mathcal{Y}^t_{v,u}$, $\mathcal{X}^t_{u,e,c}$ and $\mathcal{Z}^t_{v,b}$ to the set $\{0,1\}$. Condition (10) ensures that a UAV $v \in \mathcal{V}$ will select one and only one serving BS at each timestamp $t \in \mathcal{T}$. Condition (11) forces each radio antenna $e \in \mathcal{E}$ of a BS $u \in \mathcal{U}$ to take one position at each timestamp $t \in \mathcal{T}$. As for Condition (12), it states that if a radio antenna e of a BS u chooses the position c_i at timestamp t, it can only select one of the neighboring positions (c_{i-1}, c_i or c_{i+1}) at next timestamp $t+1$. This imposes a realistic constraint in changing the angle of a radio antenna between two consecutive timestamp. This condition can be substituted by the following constraint:

$$\forall u \in \mathcal{U}, \forall e \in \mathcal{E}, \forall c_i \in \mathcal{C}, \forall t \in \mathcal{T}; \\ \mathcal{X}^t_{u,e,c_i} \leq \mathcal{X}^{t+1}_{u,e,c_i} + \mathcal{X}^{t+1}_{u,e,c_{i-1}} + \mathcal{X}^{t+1}_{u,e,c_{i+1}}, \quad (15)$$

which imposes that $\mathcal{X}^{t+1}_{u,e,c_i} + \mathcal{X}^{t+1}_{u,e,c_{i-1}} + \mathcal{X}^{t+1}_{u,e,c_{i+1}}$ is equal to 1 if $\mathcal{X}^t_{u,e,c_i}$ is equal to 1 (note that $\mathcal{X}^{t+1}_{u,e,c_i} + \mathcal{X}^{t+1}_{u,e,c_{i-1}} + \mathcal{X}^{t+1}_{u,e,c_{i+1}}$ can not be greater than 1 as per Condition (11)). On the other hand, Condition (13) imposes that at least one RB should be assigned to each UAV $v \in \mathcal{V}$ at each timestamp $t \in \mathcal{T}$. As for the last one, Condition (14) ensures that a RB $b \in \mathcal{B}$ of a BS $u \in \mathcal{U}$ is only assigned to one of its served UAVs, at most. This constraint is expressed as the produce of boolean variables and can be linearized by defining a boolean variable $\omega^t_{u,v,b} = \mathcal{Y}^t_{v,u}\mathcal{Z}^t_{v,b}$ and submitting (14) by the following constraints:

$$\forall u \in \mathcal{U}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \sum_{v \in \mathcal{V}} \omega^t_{u,v,b} \leq 1, \quad (16)$$

$$\forall u \in \mathcal{U}, \forall v \in \mathcal{V}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \omega^t_{u,v,b} \leq \mathcal{Y}^t_{v,u} \quad (17)$$

$$\forall u \in \mathcal{U}, \forall v \in \mathcal{V}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \omega^t_{u,v,b} \leq \mathcal{Z}^t_{v,b} \quad (18)$$

$$\forall u \in \mathcal{U}, \forall v \in \mathcal{V}, \forall b \in \mathcal{B}, \forall t \in \mathcal{T}; \quad \omega^t_{u,v,b} \geq \mathcal{Y}^t_{v,u} + \mathcal{Z}^t_{v,b} - 1. \quad (19)$$

Note that in the above conditions, (17) and (18) force $\omega^t_{u,v,b}$ to 0 if $\mathcal{Y}^t_{v,u}$ or $\mathcal{Z}^t_{v,b}$ is equal to 0, while (19) forces $\omega^t_{u,v,b}$ to 1 if both $\mathcal{Y}^t_{v,u}$ and $\mathcal{Z}^t_{v,b}$ are equal to 1.

However, the above optimization problem is not linear, which is due to the objective function. In order to enable network optimization so that the radio coverage can get ahead of the UAVs to enhance their QoS, we propose a solution based on DRL. The proposed solution adopts an heterogeneous agent-based approach including two types of agents, namely the UAVs and the BSs. Here, the UAVs aim to select the serving BSs along with the radio resources, and the BSs are to adjust the positions of their radio antennas. The next section introduces the proposed solution.

IV. An Heterogeneous Agent-based Deep Reinforcement Learning Approach for the AMC Problem

This section introduces the proposed solution for the AMC problem. This solution is based on the framework of DRL, where a multi-heterogeneous agent-based approach is adopted. The general DRL framework is depicted in Fig. 3. As we can see, two types of agents are considered, namely the UAV agents and the BS agent. At timestamp t, each agent $\mathbf{g}$ gets the system state s^t (step 1 of Fig. 3) and decides the action $s^t_{\mathbf{g}}$ to execute (step 2 of Fig. 3). The taken action depends on the agent type; a UAV agent aims to select the serving BS along with the allocation of resources, whereas a BS agent targets tilting its radio antennas. After the execution of the action, the agents get the reward value along with the next state s^{t+1} (step 3 in Fig. 3). Each agent $\mathbf{g}$ also gets the actions, $a^t_{-\mathbf{g}}$, performed by the other agents, which will be used to learn the model. Indeed, implementing DRL in a multi-agent environment requires sharing/aggregating the experiences/models between the agents. A replay memory is considered in our design to store the experiences that will be used to train the neural network (step 4 in Fig. 3). We further detail in the rest of this section the proposed design, mainly in terms of the system state, the action space, the system reward and learning process.

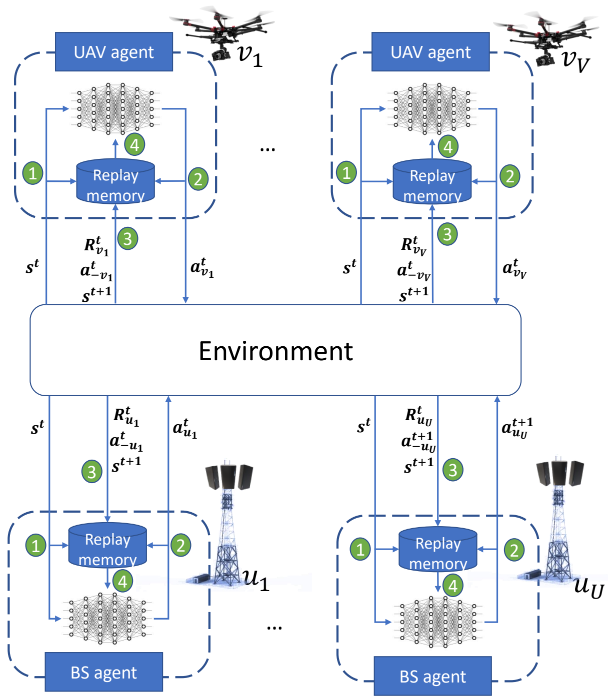


Fig. 3: MHA-DRL framework: two types of heterogeneous agents are considered (UAV and BS agents).

A. System state

We define the system state in a way to capture the characteristic of the network and that of the mission performed by the UAVs (materialized by the trajectory). To this end, let us denote by $\mathcal{L}_v(t,n)$ the locations of the UAV $v \in \mathcal{V}$ at timestamps $[t, \ldots, t+n]$. $\mathcal{L}_v(t,n)$ allows to capture a portion of the trajectory that can be used to prepare the network in advance and be ahead of the mobile users. The system state s^t at timestamp t is described as

$$s^t = [p_u, h_{uv}(t), \mathcal{L}_v(t,n), c_{ue}]_{u \in \mathcal{U}, e \in \mathcal{E}, v \in \mathcal{V}}, \quad (20)$$

where c_{ue} is the angle position of the radio antenna e of the BS u.

B. Action space

The action taken by an agent depends on its type. For a BS agent, the aim is to tilt the direction of the radio antennas and the underlying action taken by an agent u at timestamp t is described as

$$a^t_u \in \{incr, decr, keep\}^{|\mathcal{E}|}, \quad (21)$$

which refers to the action of increasing, decreasing of keeping the angle of each radio antenna $e \in \mathcal{E}$ of the BS u. As for a UAV agent v, the taken action aims to select the serving BS along with the allocation of resources. This can be described as

$$a_v^t \in \mathcal{U} \times \{0,1\}^{|\mathcal{B}|}. \tag{22}$$

C. System reward

After executing an action, the agent gets a reward from the environment corresponding to the value of the selected action given the current system state. The reward function is therefore defined to promote the actions leading to maximize the objective function. The reward function for the two type of agents is based on transmission rate of the UAVs. For a UAV agent v, the reward function corresponding to the application of the actions a_v^t and a_{-v}^t (respectively from the agent v and by other agent than v) on the system state s^t is defined as

$$\mathcal{R}_v(s^t, a_v^t, a_{-v}^t) = r_v(t). \tag{23}$$

Therefore, increasing the reward for a UAV agent is translated into enhancing the associated transmission rate. As for a BS agent u, the corresponding reward function from applying the actions a_u^t and a_{-u}^t (respectively by the agent u and by other agent than u) on the system state s^t is defined as

$$\mathcal{R}_u(s^t, a_u^t, a_{-u}^t) = \frac{1}{|\mathcal{C}_u^t|} \sum_{v \in \mathcal{C}_u^t} r_v(t), \tag{24}$$

where $\mathcal{C}_u^t$ is the set of UAVs using u as the serving BS at timestamp t.

D. Learning process

The two agent types aim to learn optimal policies allowing to maximize the objective function. To this end, each agent $\mathbf{g} \in \mathcal{U} \cup \mathcal{V}$ implements an A2C algorithm that directly parameterizes the policy $\pi_{\mathbf{g}}$ to learn. Indeed, A2C is a policy-based algorithm that employs two deep neural networks: the actor network (to approximate the policy) and the critic network (to approximate the value function). For the agent $\mathbf{g}$, we denote by $\dot{\theta}_{\mathbf{g}}^t$, respectively $\ddot{\theta}_{\mathbf{g}}^t$, the parameters of the actor network, respectively the critic network, at timestamp t. The parameters are updated in the direction $\hat{\gamma}_{\mathbf{g}}^t$ (which is an unbiased estimation of $\gamma_{\mathbf{g}}^t$) defined as

$$\hat{\gamma}_{\mathbf{g}}^t = \nabla log(\pi_{\mathbf{g}}(a_{\mathbf{g}}^t|s^t)) A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t), \tag{25}$$

$$\gamma_{\mathbf{g}}^t = \nabla \mathbb{E}[\sum_{k=0}^{\infty} \tau_{\mathbf{g}}^k \mathcal{R}_{\mathbf{g}}(s^{t+k}, a_{\mathbf{g}}^{t+k}, a_{-\mathbf{g}}^{t+k})], \tag{26}$$

where $\pi_{\mathbf{g}}(a_{\mathbf{g}}^t|s^t)$ is the probability of taking the action $a_{\mathbf{g}}^t$ by the agent $\mathbf{g}$ when the state is s^t, $\tau_{\mathbf{g}} \in [0,1]$ is a discount factor and $A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t)$ is the advantage value which is defined as

$$A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t) = Q_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t) - V_{\pi_{\mathbf{g}}}(s^t). \tag{27}$$

The function $Q_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t)$ in the above equation refers to the Q-function, which defines the value of the taken action, while $V_{\pi_{\mathbf{g}}}(s^t)$ refers to the V-function. By considering the Bellman equation, (27) can be formulated as

$$A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t) = \mathcal{R}_{\mathbf{g}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t) + \tau_{\mathbf{g}} V_{\pi_{\mathbf{g}}}(s^{t+1}) - V_{\pi_{\mathbf{g}}}(s^t). \tag{28}$$

The parameters of the actor network are learned by minimizing the negative log likelihood scaled by the advantage as

$$\mathcal{L}^a(\dot{\theta}_{\mathbf{g}}^t) = \mathbb{E}[A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t) log(\pi_{\mathbf{g}}(a_{\mathbf{g}}^t|s^t))]. \tag{29}$$

As for the critic network, the parameters are learned by minimizing the error of the value function as

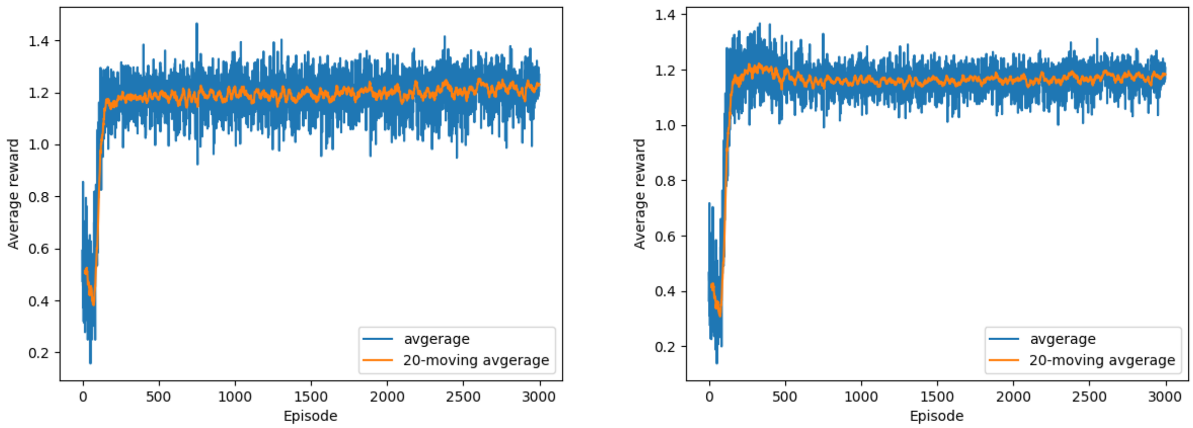
$$\mathcal{L}^c(\ddot{\theta}_{\mathbf{g}}^t) = \mathbb{E}[A_{\pi_{\mathbf{g}}}(s^t, a_{\mathbf{g}}^t, a_{-\mathbf{g}}^t)^2]. \tag{30}$$

V. Performance Evaluations

This section introduces the performance evaluations of the proposed solution for the AMC concept. The simulation has been preformed in a $1000m \times 1000m$ area consisting of 4 BSs. The simulation also considered 5 UAVs, where each UAV has a trajectory of 7 waypoints. Each BS has 2 radio antennas that can be tilted from $0°$ (horizontal level) to $50°$. The threshold angle ϕ is set to $15°$. We considered a noise power N_0 of $-130dBm$ and a RB bandwidth of $180kHz$. The DRL model has been implemented using Pytorch [11]. In addition, the discount factor $\tau_{\mathbf{g}}$ is set to 0.9 and the learning rate is set to 0.001.

We have first evaluated the proposed MAH-DRL in terms of learning optimal solutions for preparing the network to get ahead of the mobility of the UAVs. The obtained results are depicted in Fig. 4. As we can see, the two types of agents are able learn policies allowing to maximize the reward values. For a UAV agent (Fig. 4 (a)), this is translated into enhancing the transmission rate for the associated UAV. Indeed, the reward function for a UAV agent is based on the value of the transmission rate (see equation (23)). As for a BS agent (Fig. 4 (b)), increasing the reward value implies enhancing the transmission rate for the UAVs connected to this BS (see equation (24)).

Furthermore, we have also compared the obtained results of the proposed approach against a baseline solution. The latter is materialized in the case where the BSs do not change the angle of their radio antennas. This reflects the situation of today's networks. The comparison is made considering different angles of the radio antennas of the BSs for the baseline solution. The underlying scenario consists of 5 UAVs, where each UAV has a path of 7 waypoints (corresponding to 7 timestamps). The same scenario has been considered for the proposed AMC approach and the baseline solution, and the obtained results are depicted in Fig. 5. As we can see from, the proposed AMC approach ensures better transmission rate for the UAVs throughout their path compared to the baseline solution. Indeed, adjusting the angle of the radio antennas in a way to get ahead of the mobility of the UAVs allows to ensure better coverage. Furthermore, it also allows to control the interference from the non-serving BSs. The results of this evaluation prove the effectiveness of the AMC approach.



(a) Average reward of UAV agents (b) Average reward of BS agents

Fig. 4: Evaluation of the proposed MHA-DRL approach.

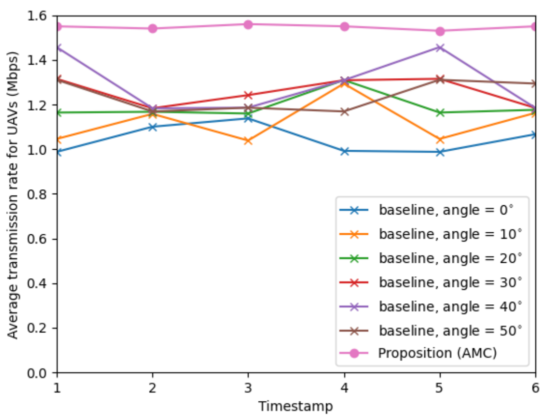


Fig. 5: Comparison of the proposed approach against baseline solutions (fixed angles of the BSs' radio antennas).

VI. Conclusion

This paper investigated the problem of maintaining an enhanced mobile network coverage for UAVs. It introduced the concept of Ahead-Me Coverage (AMC), that aims to dynamically tilt the direction of the radio antennas of the BSs in a way to get the coverage ahead of the UAVs. The formulation of the problem resulted in a non-linear and non-convex optimization. To address this issue, we proposed a solution based on DRL. More precisely, a multi-heterogeneous approach is adopted, where two types of agents are considered (namely, UAV agents and BS agents). The performance evaluation showed that the proposed approach provides better results compared to the baseline solutions. It also proves that preparing the network to get ahead the flying UAVs ensures better coverage and provides a new mean for controlling the interference.

Acknowledgment

This work was supported by the European Union's Horizon 2020 Research and Innovation Program through the 5G!Drones Project under Grant No. 857031.

References

[1] T. Taleb, A. Ksentini, and P. A. Frangoudis, "Follow-Me Cloud: When Cloud Services Follow Mobile Users," *IEEE Transactions on Cloud Computing*, vol. 7, no. 2, pp. 369–382, 2019.

[2] R. Bifulco and R. Canonico, "Analysis of the handover procedure in Follow-Me Cloud," in *2012 IEEE 1st International Conference on Cloud Networking (CLOUDNET)*, 2012, pp. 185–187.

[3] ETSI GS MEC 021 V2.1.1. (2020) Multi-access Edge Computing (MEC); Application Mobility Service API. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/021/02.01.01_60/gs_MEC021v020101p.pdf

[4] J. You, S. Jung, J. Seo, and J. Kang, "Energy-Efficient 3-D Placement of an Unmanned Aerial Vehicle Base Station With Antenna Tilting," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1323–1327, 2020.

[5] L. Ruan, J. Wang, J. Chen, Y. Xu, Y. Yang, H. Jiang, Y. Zhang, and Y. Xu, "Energy-efficient multi-UAV coverage deployment in UAV networks: A game-theoretic framework," *China Communications*, vol. 15, no. 10, pp. 194–209, 2018.

[6] S. A. Al-Ahmed, M. Z. Shakir, and S. A. R. Zaidi, "Optimal 3D UAV base station placement by considering autonomous coverage hole detection, wireless backhaul and user demand," *Journal of Communications and Networks*, vol. 22, no. 6, pp. 467–475, 2020.

[7] N. Namvar, A. Homaifar, A. Karimoddini, and B. Maham, "Heterogeneous UAV Cells: An Effective Resource Allocation Scheme for Maximum Coverage Performance," *IEEE Access*, vol. 7, pp. 164 708–164 719, 2019.

[8] M. Samir, D. Ebrahimi, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Leveraging UAVs for Coverage in Cell-Free Vehicular Networks: A Deep Reinforcement Learning Approach," *IEEE Transactions on Mobile Computing*, vol. 20, no. 9, pp. 2835–2847, 2021.

[9] M. Simunek, F. P. Fontán, and P. Pechac, "The UAV Low Elevation Propagation Channel in Urban Areas: Statistical Analysis and Time-Series Generator," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 7, pp. 3850–3858, 2013.

[10] A. AL-Hourani, S. Chandrasekharan, G. Kaandorp, W. Glenn, A. Jamalipour, and S. Kandeepan, "Coverage and rate analysis of aerial base stations [Letter]," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 3077–3081, 2016.

[11] "PyTorch pytorch website," https://pytorch.org/, accessed: 2021-03-30.