Performance Analysis and 3D Position Deployment for V2V-assisted UAV Communications in Vehicular Networks

Biling Zhang, *Senior Member, IEEE,* Zixuan He, Yaoyu Feng, and Zhu Han *Fellow, IEEE* .

***Abstract*—Deploying unmanned aerial vehicles (UAVs) as flying base stations (BSs) is a promising solution to alleviate the burden of communication infrastructure. However, few works have been done to theoretically analyze the system performance, especially for the vehicle-to-vehicle (V2V) assisted vehicular network, where vehicle users (VUs) can request contents from either the UAV or other VUs that the UAV has successfully served. We first consider the single hot-spot scenario and analyze the system performance in terms of the VU's successful service probability (SSP). Then we extend our analysis to the multi-hot-spot scenario where the hot-spots merge and split due to VU mobility. For this scenario, the SSP and the average number of successfully served VUs (SSVs) are theoretically derived, based on which a UAV's 3D position deployment problem is formulated. Due to the complicated formulation of SSP and the dynamic environment including hot-spot distribution and UAV's battery state, the closed-formed solutions to the UAV's deployment are intractable. To obtain the sub-optimal positions of the UAV, the proposed problem is reformulated into a deep reinforcement learning (DRL) framework, and a pre-trained deep Q-network (DQN) based scheme is proposed. Simulation results demonstrate that compared to existing schemes, our proposed scheme achieves a high number of successfully served users with moderate energy consumption.**

***Index Terms*—UAV 3D position deployment, successful service probability, deep reinforcement learning, V2V-assisted UAV communication, vehicular network.**

I. Introduction

VEHICULAR networks have advanced significantly in parallel with the advent of intelligent transportation systems. As a typical application of the Internet of Things (IoT) technology, they are expected to enable reliable and efficient large-scale connectivity for vehicle users (VUs) with diversified services [1]. To address the problem of surging traffic in the vehicular networks, unmanned aerial vehicles (UAVs), which have high mobility and a significant likelihood of establishing line-of-sight (LoS) communication links, are frequently utilized as flying base stations (BSs) for data delivery, traffic monitoring, and improving connectivity [2].

Existing works related to UAV-aided vehicular networks have mainly investigated issues such as communication and resource allocation [3]- [5], edge computing [6] [7], and UAV position deployment or trajectory design [8] [9]. To optimize the system performance, the UAV's position or trajectory in most of these works was deployed to reduce the flight time of each task so that the UAV could carry out the tasks as much as possible. However, without modeling and analyzing how the UAV position or trajectory is affected by energy consumption, the deployment schemes are not optimal when the UAV's onboard battery capacity is concerned.

In some works, the UAV's position deployment is based on the energy consumption modeling [10]- [12]. For instance, considering the UAV energy constraint, the authors in [10] studied how to adjust the UAV 3D position and maximize energy harvesting in wireless power transfer. Taking the transmit power and motion power into consideration, the authors in [11] maximized the users' quality of experience by optimizing the UAV's 3D positions in emergency information collection and transmission. However, when deploy the UAVs, these schemes assumed that the users' distribution is deterministic and static, which is not true in reality.

Recently, some UAV position deployment schemes have been proposed for the scenarios where users are moving. For example, taking into account their communication coverage, flying range, and energy limitations, the authors in [13] introduced an approach based on deep reinforcement learning (DRL) to optimize the UAVs' positions. While in [14], the authors deployed the optimum 3D Cartesian coordinate of the UAV over each time slot, in which the VU's mobility was considered. In these works, the UAVs were assumed to serve the VUs within only one hot-spot. In reality, however, as the VUs move, the hot-spot will split or merge dynamically [15] [16]. In such a case, one UAV has to serve the VUs in multiple hot-spots when no more UAVs join to provide services. Nevertheless, few work has pay attention to such a scenario where the users' distribution is dynamic and uneven and the UAV has to adjust its positions accordingly for better service. What's more, in these works, only the gains offered by the UAV in the UAV-to-vehicle (U2V) transmission mode are exploited, while the gains offered by the vehicle-to-vehicle (V2V) communication are neglected.

There are a few works that studied the V2V communications in the UAV-aided vehicular networks [17] [18]. In [17], the authors studied the V2V information diffusion in the disaster scenarios with assistance of UAVs. In [18], the UAV was adopted as a shared relay node to extend the cooperative V2V transmission from inter-cell to multiple cells. Although these works have achieved the gains of both the U2V and V2V communications, the performance of the systems is not theoretically analyzed, and several critical aspects in the vehicular networks, such as vehicle mobility and social proximity, have not been utilized to fully explore the network capacity.

Note that many works focusing on the UAV position de-

B. Zhang and Y. Feng are with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China (email: bilingzhang@bupt.edu.cn).

Z. He is with the Communication Systems Department, EURECOM, 06410 Sophia-Antipolis, France

Z. Han is with the Department of Electrical and Computer Engineering at the University of Houston, Houston, TX 77004, USA (email: hanzhu22@gmail.com)

ployment schemes based on traditional optimization methods which require complete and accurate information of the networks [19]. Due to the high network dynamic caused by the mobility of VUs, however, such a requirement is difficult to meet. Furthermore, the UAV position deployment problems for the vehicular networks are often modeled as optimization problems with nonlinear constraints that are difficult to solve by traditional optimization methods. Fortunately, nowadays most of the UAVs have intelligence which enable them the ability of learning and reasoning to extract valuable knowledge and make adaptive decisions [20]. Noticing DRL is an emerging artificial intelligence technique suitable for dynamic scenarios and is appropriate for long-term optimization problems, and deep Q-network (DQN) is one of the most basic DRL method, it is adopted here to learn the UAV's 3D position during the peak traffic hours.

Therefore, in this paper, we theoretically analyze the performance of V2V-assisted UAV communications in vehicular networks, based on which the UAV's near optimal positions are obtained by DQN. Part of the work, i.e., the UAV development and the performance analysis for the scenario of a single hot-spot, has been published in [21]. Here, the analysis is extended to the scenario of multi-hot-spot where the dynamics of the VUs and the energy consumption of the UAV are taken into account. The main contributions are summarized as follows:

- Although quite a few existing works have exploited the gains of U2V and V2V communications in vehicular networks, its theoretical performance analysis has not been given out in literature. Considering the key factors such as vehicle mobility and social proximity, in this paper, the performance of V2V-assisted UAV communications in terms of successful service probability (SSP) is theoretically derived for the single hot-spot scenario.
- Different from most previous works that assumed there is only one stationary hot-spot within the service area of the UAV, we consider the VUs' mobility and study the multi-hot-spot scenario where the hot-spots may merge and split as time passes by. The system performance of the multi-hot-spot scenario is also theoretically analyzed here, based on which the UAV's 3D position is optimized to maximize the average number of successfully served VUs (SSVs).
- While most previous works only consider the UAV's transmission energy consumption when deploying its position, we also take into account its energy consumption for moving, such that the UAV's limited energy can be fully utilized when it adjusts its position in each time slot to serve more VUs.
- Due to the complicated expression of SSP and the dynamic environment such as UAV battery state and hot-spot distribution, the closed-formed solutions to the proposed problem are intractable, and a pre-trained DQN based DRL scheme is proposed to find the sub-optimal positions of the UAV.
- Simulation results are shown to validate the effectiveness and efficiency of our proposed scheme. Compared to existing schemes in the literature, our proposed scheme achieves the highest number of SSVs with moderate energy consumption.

The remainder of this paper is organized as follows. Section II introduces the system model. In Section III, we analyze the system performance for the single static hot-spot scenario. In Section IV, we analyze the system performance for the dynamic multi-hot-spot scenario, based on which the UAV 3D position deployment problem is formulated. Section V proposes a pre-trained DQN based scheme to find the sub-optimal solutions, and Section VI provides the numerical results to demonstrate the performance of the proposed scheme. Finally, the paper is concluded in Section VII.

II. System Model

As depicted in Fig. 1, we focus on a vehicular network in an urban area during the peak traffic hours, where the hot-spots split or merge as the VUs move. To alleviate the load on the micro base station (MBS) and roadside units (RSUs), one UAV equipped with a wireless transmission module and cache units is deployed to serve the VUs in the considered area. Popular contents, such as safety information and infotainment, have been cached in the UAV in advance.

Suppose the duration of the peak traffic hours is T_e, and it is partitioned evenly into I time slots with duration τ. In each time slot $i \in \mathcal{I} = \{1, ..., I\}$, there are N_i hot-spots in the considered area. The distribution of the VUs in each hot-spot is assumed to follow the homogeneous Poisson point process (HPPP) with identical density λ.

In the considered scenario, once a VU has been successfully served in a U2V mode, it will announce its role of assisting VU (AVU). Then the neighbors, who have failed to receive the content from the UAV and become help needed VUs (HVUs), can communicate with this AVU in a V2V mode if they happen to be located within the V2V-proximity area of this AVU. In the following, we will first discuss the U2V transmission model and the V2V transmission model in Section II-A and Section II-B, respectively. Then we will discuss the UAV's energy consumption model in Section II-C. The detailed procedure for how the AVUs and the HVUs are pairing is not discussed here because it is beyond the scope of this paper, and the interested readers may refer to [22] and [23] for more information.

A. U2V Transmission Model

Assume that at time slot i, the UAV is deployed at a height of h_i meters from the ground with a horizontal position $\boldsymbol{w}_i = [ux_i, uy_i]$. The communication links between the VUs and the UAV can be either LoS or non-LoS (NLoS). For a VU whose distance from the UAV's project position on the ground is r, the gain of the LoS and the NLoS link can, respectively, be given as

$$g_{\text{LoS},i}(r) = \left(r^2 + h_i^2\right)^{-\alpha_u/2}, \tag{1}$$

and

$$g_{\text{NLoS},i}(r) = \eta \left(r^2 + h_i^2\right)^{-\alpha_u/2}, \tag{2}$$

where α_u is the path loss exponent over the U2V transmission, and η is the additional attenuation factor about the NLoS. Due

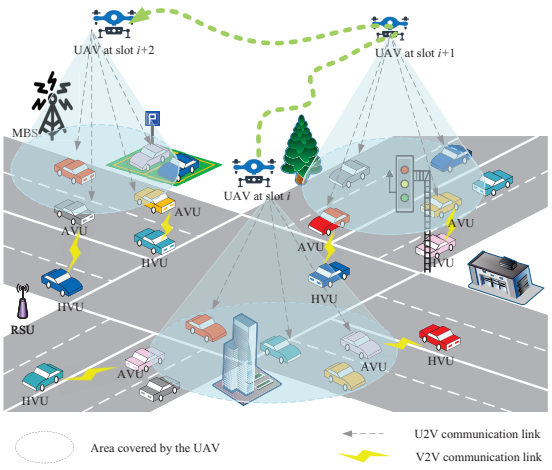

Fig. 1. An illustration of V2V-assisted UAV communications with multiple hot-spots.

to the obstruction of buildings, the different locations of VUs, and the altitude of UAV, the probability that a user connects to UAV through the LoS link is [24]

$$p_{\text{LoS},i}(r) = \frac{1}{1 + a\exp\left(-b\left[\theta_i - a\right]\right)}, \tag{3}$$

while the probability that a user connects with UAV through the NLoS link is $p_{\text{NLoS},i}(r) = 1 - p_{\text{LoS},i}(r)$. Here, a and b are constants which are determined by the environment, and $\theta_i = \tan^{-1}\left(\frac{h_i}{r}\right)$ is the elevation angle. Then the average U2V channel power gain can be derived as

$$g_{\text{U2V},i}(r) = p_{\text{LoS},i}(r)\, g_{\text{LoS},i}(r) + p_{\text{NLoS},i}(r)\, g_{\text{NLoS},i}(r). \tag{4}$$

Suppose the UAV is allocated a dedicated spectrum to broadcast popular content to the VUs, and the spectrum is equally and orthogonally allocated to each VUs. Hence, the signal-to-noise ratio (SNR) of the considered user is given by

$$\gamma_{\text{U2V},i}(r) = \frac{P_u g_{\text{U2V},i}(r)}{\sigma^2}, \tag{5}$$

where P_u is the UAV's transmit power and σ^2 is the noise power.

B. V2V Transmission Model

For V2V transmissions, we assume that all VUs utilize a separate dedicated spectrum. Hence, a VU receiving data through a V2V link will receive interference from other V2V pairs. We further assume that the V2V channels are characterized by Rayleigh fading, and the distribution of the V2V pairs also follows HPPP, then according to [21], the signal-to-interference-plus-noise ratio (SINR) of a VU served via V2V transmission will be

$$\gamma_{v,i}(D) = \frac{P_v g_v D^{-\alpha_v}}{I_{v,i} + \sigma^2}. \tag{6}$$

Here, P_v is the transmission power, g_v is the channel power gain, D is the distance between the VU and its V2V transmitter, α_v is the path loss exponent, and $I_{v,i}$ is the total received interference from other V2V transmitters at time slot i.

C. UAV Energy Consumption Model

Note that when the UAV adjusts its 3D position or stays still to transmit content, it will consume energy. Since the energy of the UAV is limited, it should be allocated carefully. Therefore, two parts of the rotary-wing UAV energy consumption [11] [25], i.e., serving energy consumption and moving energy consumption, are considered in this paper.

1) Energy consumption in Moving : When the UAV adjusts its position, it cannot serve the VUs and the energy is consumed for its propulsion. Here, we leverage the rotary-wing UAV propulsion energy consumption model which is given by

$$E_{mv,i} = \left(P_{w,i} + P_{h,i}\right)\tau. \tag{7}$$

In (7), $P_{w,i}$ denotes the horizontal motion power at time slot i as

$$P_{w,i} = c_1 + c_2\left(\frac{||\dot{\boldsymbol{w}}_i||}{\tau}\right)^2, \tag{8}$$

where $||\dot{\boldsymbol{w}}_i||$ is the UAV's velocity, and c_1 and c_2 are constants related to the air density, blade drag and angular coefficient, rotor radius, and UAV mass. While $P_{h,i}$ denotes the vertical motion power at time slot i as

$$P_{h,i} = c_3\frac{h_i - h_{i-1}}{\tau} + \frac{h_i - h_{i-1}}{\|h_i - h_{i-1}\|}\sqrt{\left(\frac{h_i - h_{i-1}}{\tau}\right)^2 + c_4}, \tag{9}$$

where c_3 and c_4 are also constants related to the UAV mass, the UAV blade radius, and the air density.

2) Energy Consumption in Service: When the UAV stays still and serves the VUs at slot i, the energy it consumes will be

$$E_{sv,i} = P_u\tau + c_1\tau. \tag{10}$$

In (10), the first part is the energy for transmission, and the second part is the energy for hovering which is obtained from the horizontal motion energy consumption model (8) when the velocity is set to zero.

III. Analysis for the Single Hotspot Static Scenario

In this section, we study the single static hot-spot scenario, for which the SSP of U2V transmission and V2V transmission are analyzed in Section III-A and Section III-B, respectively, and then the average SSP of users in V2V-assisted UAV communication is derived in Section III-C. Since the hot-spot is stationary, the subscript i indicating time slot i in the equations in Section II is removed when the equations are used in this case.

A. SSP of U2V Transmission

Assume that the UAV is hovering at the center of the hot-spot with a height of h meters from the ground, with the maximum transmission coverage radius being $R_{\max}$ which is determined by its height and antenna azimuth. For the VU whose distance from the projection of the UAV is r, the SNR of U2V transmission will be $\gamma_{\text{U2V}}(r)$. According to the definition, if a user's SNR is larger than the threshold, say

γ_0, it is said to be successfully served. Therefore, let $R_{\text{LoS}} = \sqrt{\left(\frac{P_u}{\gamma_0 \sigma^2}\right)^{2/\alpha_u} - h^2}$ and $R_{\text{NLoS}} = \sqrt{\left(\frac{P_u \eta}{\gamma_0 \sigma^2}\right)^{2/\alpha_u} - h^2}$, the SSP of the considered user in the U2V transmission can be calculated as

$$\begin{aligned} p_{succ,\text{U2V}}(r) &= \Pr\left[\gamma_{\text{U2V}}(r) \geq \gamma_0\right] \\ &= p_{\text{LoS}}(r) \cdot \Pr\left[r \leq R_{\text{LoS}}\right] + p_{\text{NLoS}}(r) \cdot \Pr\left[r \leq R_{\text{NLoS}}\right]. \end{aligned} \tag{11}$$

Supposing the VUs in the hot-spot are uniformly distributed, the probability density function (PDF) of their locations is $f(r) = \frac{2r}{R_{\max}^2}$, $(r < R_{\max})$. Then for any VU in the hot-spot, the average SSP of the U2V transmission will be

$$\begin{aligned} \overline{p}_{succ,\text{U2V}} =& \mathbb{E}_r\left[p_{succ,\text{U2V}}(r)\right] \\ =& \int_0^{\min[R_{\text{LoS}}, R_{\max}]} p_{\text{LoS}}(r) \frac{2r}{R_{\max}^2} \mathrm{d}r \\ &+ \int_0^{\min[R_{\text{NLoS}}, R_{\max}]} p_{\text{NLoS}}(r) \frac{2r}{R_{\max}^2} \mathrm{d}r. \end{aligned} \tag{12}$$

B. SSP of V2V Transmission

In this subsection, we calculate the SSP of a VU in V2V transmission, taking into account the vehicle's mobility and social proximity.

1) Vehicle Mobility and Connection Probability: The V2V links are often unstable and unreliable due to the mobility of vehicles. For instance, a transmission may be disrupted if the vehicles move beyond their maximum communication range. To assess the quality of links in V2V transmission, we adopt a similar approach proposed in [21] and [27] and calculate the connection probability as follows.

Suppose that the initial headway distance of a V2V pair is D. Noticing that this distance varies over time, we denote it as $D(t)$, with $D(t) \leq 0$ meaning the VU is ahead of its V2V transmitter and $D(t) > 0$ meaning it is behind. Additionally, the maximum communication range is denoted by $D_{\max}$. The duration of the V2V transmission, known as the connection time, is approximated by the mean first passage time T which starts at $t = 0$ and ends when the users move out of their maximum communication range, i.e.,

$$T = \{\min t | D(0) = D, -D_{\max} < D(x) < D_{\max}, 0 \leq x \leq t\}. \tag{13}$$

In (13), T is a random value and depends on D and the relative velocities of the V2V pair. Hence, $D(t)$ can be modeled as a Wiener process [27] with drift $\beta = v_R - v_T$ and variance $\sigma_v^2 = \sigma_R^2 - \sigma_T^2$. Here, v_T (v_R) and σ_T^2 (σ_R^2) are the mean and variance of V2V transmitter (receiver) velocity, respectively. Within the infinitesimal interval Δt, the increment of $D(t)$, i.e., $\Delta D(t)$, can be calculated as

$$\Delta D(t) = D(t + \Delta t) - D(t) = \beta \Delta t + G\sqrt{\sigma_v^2 \Delta t}, \tag{14}$$

where G is a random variable that follows the unit normal distribution.

Next, let $p(x|D,t)$ represent the PDF of the user's time-varying velocity. Given that (14) is a Wiener process, $p(x|D,t)$ satisfies the Kolmogorov equation [27] as

$$\begin{aligned} \frac{1}{2}\sigma_v^2 \frac{\partial^2}{\partial x^2} p(x|D,t) + \beta \frac{\partial}{\partial x} p(x|D,t) = \frac{\partial}{\partial t} p(x|D,t), \\ -D_{\max} < x < D_{\max}. \end{aligned} \tag{15}$$

Let $\delta(.)$ be a Dirac delta function, the initial and boundary conditions on the headway distance for (15) will be given as

$$\begin{aligned} &p(x|D,0) = \delta(D), \text{and} \\ &p(D_{\max}|D,t) = p(-D_{\max}|D,t) = 0, t > 0. \end{aligned} \tag{16}$$

Combining (15) and (16), we have

$$\begin{aligned} p(x|D,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \sum_{y=-\infty}^{\infty} [\exp\left\{\frac{4y\beta D_{\max}}{\sigma_v^2}\right. \\ \left. - \frac{[(x-D) - 4yD_{\max} - \beta t]^2}{2\sigma_v^2 t}\right\} - \exp\left\{\frac{2\beta D_{\max}(1-2y)}{\sigma_v^2}\right. \\ \left. - \frac{[(x-D) - 2D_{\max}(1-2y) - \beta t]^2}{2\sigma_v^2 t}\right\}]. \end{aligned} \tag{17}$$

Finally, the cumulative distribution function (CDF) of the connection time T can be derived as

$$F_{rt}(t, D) = \Pr\left[T \leq t\right] = 1 - \int_{-D_{\max}}^{D_{\max}} p(x|D,t) \mathrm{d}x. \tag{18}$$

2) Social Proximity: To assess the probability that a potential V2V transmitter possesses the requested content, we investigate it from the social perspective and evaluate it based on social proximity.

Assuming the UAV has pre-cached K popular contents of equal normalized size, and the content popularity is represented by $\mathcal{F} = \{f_1, ..., f_k, ..., f_K\}$, which follows the Zipf law as

$$f_k = \frac{(1/k)^z}{\sum_{j=1}^{K} (1/j)^z}. \tag{19}$$

Here, f_k represents the popularity of the kth content, and z is the skewness of the content popularity.

From (19), we can see that if a VU desires content k, the probability of obtaining it from other VUs is determined by its popularity f_k. In other words, the popularity of a content reflects the social proximity between two VUs who are interested in the same popular content. The higher the desired content's popularity, the greater the social proximity between these VUs.

Since we lack exact information about the content a particular VU is interested in, all content must be considered when estimating the probability. Given that the VUs' requests are independent, the average probability of a VU obtaining its desired content, i.e., the social proximity, will be

$$\overline{p}_{sp} = 1 - \prod_{k \in K} (1 - f_k). \tag{20}$$

3) SSP of V2V: Similar to the U2V transmission, a VU is considered to be successfully served via V2V transmission if the SINR surpasses a predetermined threshold γ_0. Hence, inspired by [24], the SSP of V2V can be derived as

$$\begin{aligned}\Pr\left[\gamma_v\left(D\right)\geq\gamma_0\right]&=\Pr\left[\frac{P_v\cdot g_v\cdot D^{-\alpha_v}}{I_v+\sigma^2}\geq\gamma_0\right]\\&=\exp\left(-\frac{2\pi^2\lambda_v\gamma_0^{2/\alpha_v}D^2}{\alpha_v\sin\frac{2\pi}{\alpha_v}}-\frac{\gamma_0D^{\alpha_v}\sigma^2}{P_v}\right).\end{aligned}\tag{21}$$

We assume that the potential V2V transmitters are distributed uniformly around a VU, and the PDF of the distance is given by $f(D)=\frac{2D}{D_{\max}^2}, D<D_{\max}$. By considering the user's mobility, social proximity, and physical transmission, the average SSP for V2V transmission can be derived as

$$\begin{aligned}\overline{p}_{succ,\mathrm{V2V}}&=\mathbb{E}_D\left[\Pr\left[\gamma_v\left(D\right)\geq\gamma_0\right]\cdot F_{rt}(t,D)\cdot\overline{p}_{sp}\right]\\&=\int_0^{D_{\max}}\exp\left(\frac{-2\pi^2\lambda_v\gamma_0^{2/\alpha_v}D^2}{\alpha_v\sin\frac{2\pi}{\alpha_v}}-\frac{\gamma_0D^{\alpha_v}\sigma^2}{P_v}\right)\\&\qquad\cdot\overline{p}_{sp}F_{rt}(t,D)\frac{2D}{D_{\max}^2}\mathrm{d}D.\end{aligned}\tag{22}$$

C. Average SSP

With (12) and (22), we calculate the average SSP for VUs who are served by V2V-assisted UAV communication as

$$\overline{p}_{succ}=\overline{p}_{succ,\mathrm{U2V}}+\left(1-\overline{p}_{succ,\mathrm{U2V}}\right)\cdot\overline{p}_{succ,\mathrm{V2V}}\cdot\overline{p}_{succ,\mathrm{U2V}}.\tag{23}$$

In (23), the first component represents the probability that a user can be served successfully by the UAV through U2V transmission, and the second component indicates the probability that a user can be served successfully through V2V transmission when the U2V transmission fails.

IV. Analysis for the Dynamic Multi-Hot-spot Scenario

Noticing that the hot-spots split or merge as the VUs move, the UAV adjusts its position to serve the VUs better. Therefore, in this section, we will extend our analysis to the dynamic multi-hot-spot scenario. In the following, we first derive the average SSP at a single time slot in Section IV-A to Section IV-C. Then we derive the average number of SSVs during the peak traffic hours in Section IV-D, based on which the UAV's positions is optimized.

A. SSP of U2V at Single Time Slot

Recall that at time slot i, the UAV is deployed at a height of h_i with horizontal position $\boldsymbol{w}_i=[ux_i,uy_i]$. Suppose there are N_i hot-spots observed at time slot i, the set of which is $\boldsymbol{H}_i=\{H_{i,1},...,H_{i,n},...,H_{i,N_i}\}$. For any hot-spot $H_{i,n}\in\boldsymbol{H}_i$, it can be described by the center $C_{i,n}=(x_{i,n},y_{i,n})$ and the radius $R_{i,n}$, where $x_{i,n}$ and $y_{i,n}$ are the horizontal and vertical coordinate of center $C_{i,n}$, respectively. In this subsection, since we discuss the SSP for U2V at a single time slot, we omit subscript i for simplicity.

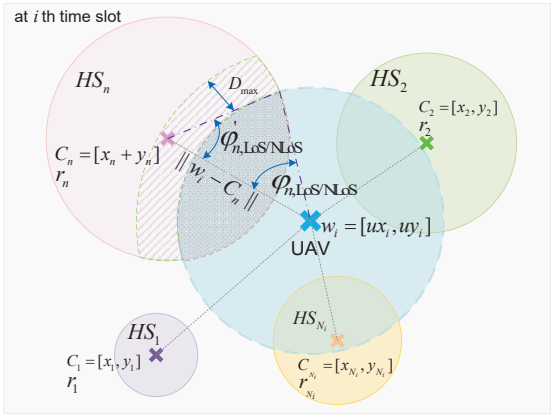

Fig. 2. An illustration of 2D hot-spots distribution and UAV transmission coverage.

We first calculate the coverage area of the UAV when multiple hot-spots appear. As illustrated in Fig. 2, angles $\phi_{n,\mathrm{LoS}}$ and $\phi'_{n,\mathrm{LoS}}$ can be calculated as

$$\phi_{n,\mathrm{LoS}}=\arccos\frac{R_{\mathrm{LoS}}^2+||\boldsymbol{w}-C_n||^2-R_n^2}{2\,||\boldsymbol{w}-C_n||\,R_{\mathrm{LoS}}},\tag{24}$$

and

$$\phi'_{n,\mathrm{LoS}}=\arccos\frac{R_n^2+||\boldsymbol{w}-C_n||^2-R_{\mathrm{LoS}}^2}{2\,||\boldsymbol{w}-C_n||\,R_n}.\tag{25}$$

Similarly, we can calculate angle $\phi_{n,\mathrm{NLoS}}$ and $\phi'_{n,\mathrm{NLoS}}$. Then, the area covered by the UAV through LoS link and NLoS link, i.e., $s_{n,\mathrm{LoS}}$ and $s_{n,\mathrm{NLoS}}$, can be respectively obtained as

$$s_{n,\mathrm{LoS}}=\begin{cases}0, & ||\boldsymbol{w}-C_n||>R_{\mathrm{LoS}}\\ \arccos\phi_{n,\mathrm{LoS}}\cdot R_{\mathrm{LoS}}^2+\arccos\phi'_{n,\mathrm{LoS}}\cdot R_n^2,\\ \quad-||\boldsymbol{w}-C_n||\,R_{\mathrm{LoS}}\cdot\sin\phi_{n,\mathrm{LoS}}, & \text{else},\end{cases}\tag{26}$$

and

$$s_{n,\mathrm{NLoS}}=\begin{cases}0, & ||\boldsymbol{w}-C_n||>R_{\mathrm{NLoS}}\\ \arccos\phi_{n,\mathrm{NLoS}}\cdot R_{\mathrm{NLoS}}^2+\arccos\phi'_{n,\mathrm{NLoS}}\cdot R_n^2,\\ \quad-||\boldsymbol{w}-C_n||\,R_{\mathrm{NLoS},i}\cdot\sin\phi_{n,\mathrm{NLoS}}, & \text{else}.\end{cases}\tag{27}$$

Combining (3), (26) and (27), the average SSP in U2V transmission at this time slot will be

$$\begin{aligned}\overline{p}_{succ,\mathrm{U2V}}&=\mathbb{E}_r\left[p_{succ,\mathrm{U2V}}\left(r\right)\right],\\&=\frac{\sum_{n=1}^{N}s_{n,\mathrm{LoS}}}{\pi R_{\max}^2}\int_0^{\min[R_{\mathrm{LoS}},R_{\max}]}p_{\mathrm{LoS}}(r)\frac{2r}{R_{\max}^2}\mathrm{d}r\\&\quad+\frac{\sum_{n=1}^{N}s_{n,\mathrm{NLoS}}}{\pi R_{\max}^2}\int_0^{\min[R_{\mathrm{NLoS}},R_{\max}]}p_{\mathrm{NLoS}}(r)\frac{2r}{R_{\max}^2}\mathrm{d}r.\end{aligned}\tag{28}$$

B. SSP of V2V at Single Time Slot

In this subsection, we derive the VU's SSP of V2V transmission at a single time slot for the multi-hot-spot scenario.

1) Vehicle Mobility and Social Proximity: Recall that the duration of a time slot is τ. Substituting τ into (18), we have the probability that a V2V pair can keep in connection at a single time slot as

$$F_{rt}(\tau, D) = \Pr\left[T \leq \tau\right] = 1 - \int_{-D_{\max}}^{D_{\max}} p(x|D,\tau)\mathrm{d}x. \quad (29)$$

As to the social proximity, i.e., the average probability derivation of a VU having its interested content, it can still be expressed as (20).

2) Physical Transmission Model: From Fig. 2 we can see that in hot-spot HS_n, the VUs who need V2V transmission can be divided into two groups: One group is located in the overlapping zone of HS_n and the area within UAV's maximum transmission coverage, the other group is located in the "annular sector" zone, i.e., the green shadow in Fig. 2, with the maximum V2V transmission range $D_{\max}$ as width. Both groups can be served by the VUs who have been successfully served through U2V transmission in the overlapping zone. In the following, we will analyze these two groups respectively.

First, similar to the single hot-spot scenario, we have the area covered and uncovered by the UAV through the LoS link as

$$s_{n,\text{LoS},in} = \begin{cases} 0, & ||\boldsymbol{w} - C_n|| > R_n + R_{\text{LoS}} \\ \arccos\phi_{n,\text{LoS},in} \cdot R_{\text{LoS}}^2 + \arccos\phi_{n,\text{LoS},in}^{'} \cdot R_n^2, & \\ \quad - ||\boldsymbol{w} - C_n|| \, R_{\text{LoS}} \cdot \sin\phi_{n,\text{LoS},in}, & \text{else}, \end{cases} \quad (30)$$

and

$$s_{n,\text{LoS},out} = \begin{cases} 0, & ||\boldsymbol{w} - C_n|| > R_n + R_{\text{LoS}} + D_{max} \\ \arccos\phi_{n,\text{LoS},out} \cdot R_{\text{LoS}}^2 + \arccos\phi_{n,\text{LoS},out}^{'} \cdot R_n^2, & \\ \quad - ||\boldsymbol{w} - C_n|| \, R_{\text{LoS}} \cdot \sin\phi_{n,\text{LoS},out}, & \text{else}, \end{cases} \quad (31)$$

where

$$\phi_{n,\text{LoS},in} = \arccos\frac{R_{\text{LoS}}^2 + ||\boldsymbol{w} - C_n||^2 - R_n^2}{2\,||\boldsymbol{w} - C_n||\, R_{\text{LoS}}}, \quad (32)$$

$$\phi_{n,\text{LoS},in}^{'} = \arccos\frac{R_n^2 + ||\boldsymbol{w} - C_n||^2 - R_{\text{LoS}}^2}{2\,||\boldsymbol{w} - C_n||\, R_n}, \quad (33)$$

$$\phi_{n,\text{LoS},out} = \arccos\frac{R_{\text{LoS}}^2 + ||\boldsymbol{w} - C_n||^2 - R_n^2}{2\left(||\boldsymbol{w} - C_n|| + D_{\max}\right) R_{\text{LoS}}}, \quad (34)$$

$$\phi_{n,\text{LoS},out}^{'} = \arccos\frac{R_n^2 + ||\boldsymbol{w} - C_n||^2 - R_{\text{LoS}}^2}{2\left(||\boldsymbol{w} - C_n|| + D_{\max}\right) R_n}. \quad (35)$$

Note that the area covered and uncovered by the UAV through the NLoS link, i.e., $s_{n,\text{NLoS}}$ and $s_{n,\text{NLoS}}^{'}$, can be obtained similarly but is omitted here due to page limitation.

Then we have the average area of "annular sector" as

$$\begin{cases} ss_{n,\text{LoS}} = ss_{n,\text{LoS},out} - ss_{n,\text{LoS},in}, \\ ss_{n,\text{NLoS}} = ss_{n,\text{NLoS},out} - ss_{n,\text{NLoS},in}. \end{cases} \quad (36)$$

For the VUs in the overlapping zone and the "annular sector", their average SSP of V2V transmission can be respectively calculated as

$$\begin{aligned} &P_n(\text{V2V succ}|\text{overlapping U2V fail}) \\ &= [1 - \frac{s_{n,\text{LoS}}}{\pi R_{\max}^2} \int_0^{\min[R_{\text{LoS}}, R_{\max}]} p_{\text{LoS}}(r)\frac{2r}{R_{\max}^2}\mathrm{d}r \\ &\quad - \frac{s_{n,\text{NLoS}}}{\pi R_{\max}^2} \int_0^{\min[R_{\text{NLoS}}, R_{\max}]} p_{\text{NLoS}}(r)\frac{2r}{R_{\max}^2}\mathrm{d}r] \\ &\quad \cdot \int_0^{D_{\max}} F_{rt}(\tau, D)\overline{p}_{sp}\Pr\left[\gamma_v\left(D\right) \geq \gamma_0\right]\frac{2D}{D_{\max}^2}\mathrm{d}D, \end{aligned} \quad (37)$$

and

$$\begin{aligned} &P_n\left(\text{V2V succ}|\text{annular sector}\right) \\ &= \frac{ss_n}{\pi R_n^2} \int_0^{D_{\max}} F_{rt}(\tau, D)\overline{p}_{sp}\Pr\left[\gamma_v\left(D\right) \geq \gamma_0\right]\frac{2D}{D_{\max}^2}\mathrm{d}D. \end{aligned} \quad (38)$$

Finally, for the VUs in hot-spot HS_n, we can derive their average SSP of V2V transmission as

$$\begin{aligned} \overline{p}_{n,succ,\text{V2V}} &= P_n\left(\text{V2V succ}|\text{annular sector}\right) \\ &\quad + P_n\left(\text{V2V succ}|\text{overlapping U2V fail}\right)\overline{p}_{succ,\text{U2V}}. \end{aligned} \quad (39)$$

Taking all the hot-spots in this area into consideration, we have the average SSP at the ith time slot as

$$\overline{p}_{succ,\text{V2V}} = \frac{\sum_{n=1}^{N} \lambda\pi R_n^2 \overline{p}_{n,succ,\text{V2V}}}{\sum_{n=1}^{N} \lambda\pi R_n^2}. \quad (40)$$

C. Average SSP at Single Time Slot

With (28) and (40), we derive the average SSP of VUs at a specific time slot in the multi-hot-spot scenario as

$$\overline{p}_{succ} = \overline{p}_{succ,\text{U2V}} + \left(1 - \overline{p}_{succ,\text{U2V}}\right)\overline{p}_{succ,\text{V2V}}. \quad (41)$$

D. Average Number of SSVs During Peak Traffic Hours

According to (41), we have the VUs' average SSP at time slot i, i.e., $\overline{p}_{i,succ}$, when the UAV's 3D position $(\boldsymbol{w}_i, h_i)$ is given. Then let $\mathcal{W} = \{\boldsymbol{w}_1, ..., \boldsymbol{w}_I\}$ and $\mathcal{H} = \{h_1, ..., h_I\}$, the average number of SSVs during the peak traffic hour T_e will be

$$AV(\mathcal{W}, \mathcal{H}) = \sum_{i=1}^{I} \overline{p}_{i,succ}\left(\lambda \sum_{n=1}^{N_i} \pi R_n^2\right). \quad (42)$$

Since the hot-spot distribution and the UAV's battery state change as time passes, to maximize the total number of successful service users, the UAV is required to optimize its 3D position overall time slots under the battery constraint. Thus, the UAV's 3D position deployment problem can be formulated as

$$\mathcal{P}: \max_{\mathcal{W},\mathcal{H}} AV(\mathcal{W}, \mathcal{H}), \quad (43)$$

$$\text{s.t.} \sum_{i=1}^{I} \psi_i E_{sv,i} + \sum_{i=1}^{I} \left(1 - \psi_i\right) E_{mv,i} \leq E_{\max}, \forall i \in \mathcal{I}, \quad (44)$$

$$||\boldsymbol{w}_i - \boldsymbol{w}_{i-1}|| \leq ||\dot{\boldsymbol{w}}_i||\tau, \forall i \in \mathcal{I}, \quad (45)$$

$$h_i - h_{i-1} \leq \dot{h}_i\tau, \forall i \in \mathcal{I}. \quad (46)$$

In (44), the UAV's energy is constrained with the maximum battery capacity $E_{\max}$, and an auxiliary binary variable ψ_i is introduced to indicate whether the UAV stays still to serve the VUs ($\psi_i = 0$) or moves to adjust its 3D position ($\psi_i = 1$). Moreover, since we assume that the UAV's position adjustment should be completed within a time slot τ, the UAV's 3D position between two consecutive time slots, i.e., $(\boldsymbol{w}_i, h_i)$ and $(\boldsymbol{w}_{i-1}, h_{i-1})$, should be constrained by its horizontal velocity $\dot{\boldsymbol{w}}_i$ and vertical velocity $\dot{h}_i$ as in (45) and (46), respectively.

However, the UAV's optimal strategy is not easy to obtain due to the following two reasons: Firstly, the SSP for each transmission is not a continuous function with regard to the UAV's 3D position, which makes it difficult to obtain the UAV's optimal position deployment even at a single time slot; Secondly, the hot-spot distribution and the UAV battery state are dynamic and coupled between time slots, which makes the closed-form solution intractable. Noticing that DRL provides a framework in which an agent can interact with the environment and learn the best behavioral decisions for each step through trial and error, in the following, we adopt DRL to address the aforementioned issues and find the sub-optimal UAV deployment solution to the proposed problem.

V. DRL Based Solutions to UAV 3D Position Deployment

To obtain the solutions to the UAV position deployment problem, in the following, we first reformulate the proposed problem to a DRL framework in Section V-A, within which the UAV deployment scheme is then elaborated in Section V-B.

A. DRL Framework

In the framework of DRL, the UAV in the considered scenario acts as an agent, while the hot-spot distribution and the UAV's energy state are regarded as the environment. As depicted in Fig. 3, at each time slot, say time slot i, the UAV observes an environment state s_i from the state space $\mathcal{S}$, and then takes an action a_i from its action space $\mathcal{A}$ based on the policy π accordingly. The policy π is determined by a state-action value function, i.e., the Q-function of the reinforcement learning, to obtain a Q-value $Q(s_i, a_i)$. After the action, the UAV receives a reward $\mathcal{R}_i$ from the environment, and the system transitions to a new state s_{i+1} at the next time slot. In the following, we will explain in detail how to reformulate the proposed optimization problem to such a DRL-based framework.

1) State: The environment state s_i the UAV observes from the system at time slot i is comprised of the hot-spot distribution and its remaining energy, i.e.,

$$s_i = \{\boldsymbol{HS}_i, E_i\} \in \mathcal{S}. \tag{47}$$

Here, E_i is the UAV's remaining energy which can be obtained by rewriting (44) as

$$E_i = E_{\max} - \sum_{j=1}^{i} \psi_j E_{sv,j} + \sum_{j=1}^{i} (1 - \psi_j) E_{mv,j}. \tag{48}$$

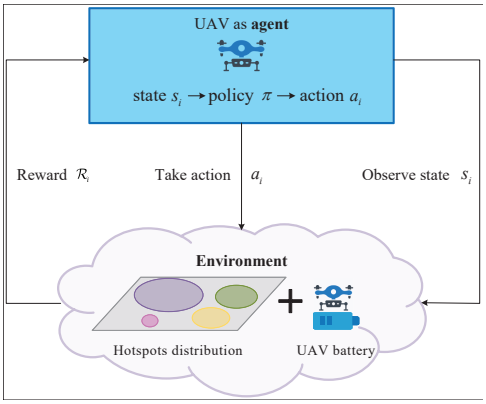

Fig. 3. An Illustration of Reinforcement Learning Framework.

2) Action: Given the observation s_i, the UAV takes an action a_i of adjusting its 3D position based on policy π, which is denoted as

$$a_i = \{W_i, H_i\} \in \mathcal{A}. \tag{49}$$

Here, W_i is the 2D position adjustment option out of a total of 9 options: *East, West, South, North, Southeast, Northeast, Southwest, Northwest,* and *Stay*, with the increment in each direction of $||\Delta W||$. H_i is the altitude adjustment option out of a total of 3 options: *Up, down* and *stay*, with the increment in each direction of $||\Delta H||$. Hence, there are in total 27 combinations of UAV's 3D position adjustment in the action space $\mathcal{A}$.

3) Reward Function: Recalling that the goal of the formulated problem $\mathcal{P}$ is to maximize the total number of successful service users in the peak time T_e under the UAV's energy constraint, the reward function here can be designed as

$$\mathcal{R}_i = \omega_p \left(\overline{p}_{i,succ} \lambda \sum_{n=1}^{N_i} \pi R_n^2 \right) - \omega_e \left(E_{\max} - E_i \right). \tag{50}$$

In (50), the first part denotes the number of successful service users, the second part is the UAV's remaining energy which is represented as a penalty, and ω_p and ω_e are the weights of these two parts, respectively. Since the goal of reinforcement learning is to maximize a long-term reward, the expected cumulative discounted reward can be measured as

$$R_i = \mathbb{E}\left[\sum_{n=0}^{\infty} \gamma^n \mathcal{R}_{i+n} \right], \tag{51}$$

where $\gamma \in [0, 1]$ is the discount factor, and the larger the value of γ is, the more the reinforcement learning model focuses on the future rewards.

B. Pre-trained DQN based Deployment Scheme

In the DRL framework established in the previous subsection where the action space is modeled as a finite discrete data set, the hot-spot distribution and the power state of the UAV are high-dimensional and complex. Using a traditional Q-table to store such state-action Q-value will cause dimension disaster. What's more, even if such a large-scale Q-table can be created, searching for a specific state-action pair's Q-value will also be time-consuming.

Under such circumstances, DQN which combines Q-learning with deep neural network (DNN) to evaluate the Q-function, is leveraged. The DQN adopts two neural networks with the same architectures: one is the evaluate Q network which is used to approximate the Q-value of the current state-action pair, and the other is the target Q network which is used to update the target Q-value. In the DQN, the state and the actions with parameter θ_{NN} are input into and analyzed through the DNN, creating a mapping from the environment to the action.

Then, for a state-action pair (s_i, a_i), the DQN model can evaluate the Q-value $Q(s_i, a_i)$ to reflect the quality of a decision. Once the Q-value of the actions are approximated, the UAV can take the appropriate action based on

$$a_i = \arg\max_{a \in \mathcal{A}} Q(s_i, a). \tag{52}$$

Without knowing the environment information in advance, the optimal Q-value i.e., Q^*, under the optimal policy can be obtained through iteration as

$$Q(s_i, a_i) = Q(s_i, a_i) + \alpha[\mathcal{R}_{i+1} + \gamma \max_{a \in \mathcal{A}} Q(s_{i+1}, a) - Q(s_i, a_i)], \tag{53}$$

where α is the learning rate of the UAV.

With the determined architectures, the neural network can be trained to minimize the loss

$$L_{DQN}(\theta_{NN}) = E[(Y - Q(s, a; \theta_{NN}))^2], \tag{54}$$

where $Y = \mathcal{R}_i + a \max_{a \in \mathcal{A}} Q(s_i, a; \theta_{NN})$ is the target Q network, and parameter θ_{NN} is updated with the gradient descent as

$$\begin{aligned} \theta_{NN} = \theta_{NN} + \nabla Q(s, a; \theta_{NN})[\mathcal{R}_i \\ + \gamma \max_{a \in \mathcal{A}} Q(s_{i+1}, a; \theta_{NN}) - Q(s_i, a_i; \theta_{NN})]. \end{aligned} \tag{55}$$

Noticing that there are many parameters in the DNN, and randomly initializing the parameters takes a relatively long time for learning to converge, in this paper, we add a pre-trained process to the aforementioned DQN framework, helping it converge quickly and improve the performance. To this target, a cache memory $\mathcal{D}^{demo}$ is set up to store the training data in form of (s_i, a_i, r_i, s_{i+1}) which can be obtained by algorithms such as ant colony optimization algorithm. Then a random mini-batch is sampled from $\mathcal{D}^{demo}$ to train the DNN.

Different from traditional DQN, the loss function L_{pre} of pre-trained based DQN is a combination of L_{DQN}, supervised loss L_s, and regular loss L_r as

$$L_{pre} = L_{DQN} + \lambda_s L_s + \lambda_r L_r. \tag{56}$$

Here, λ_s and λ_r are the weight, L_{DQN} has the same expression as in (54), L_r is used to prevent overfitting during pre-trained, while L_s gives the DNN a direction to update its parameters by fine-tuning the original DQN loss as

$$L_s = \max_{a \in \mathcal{A}}[Q(s, a) + l(s, a_E, a)] - Q(s, a_E). \tag{57}$$

In (57), a_E is the action taken under state s during pre-trained, and $l(\cdot)$ is a margin function to make sure some state-action pairs that are not available in $\mathcal{D}^{demo}$ still have a reasonable value and converge in the same direction as the pre-trained data. Specifically, when $a = a_E$, there is $l(\cdot) = 0$ and then $L_s = 0$. While in other cases, there is $l(\cdot) > 0$ and then $L_s > 0$, forcing the value of the action in the pre-trained data to be higher than the value of the other action in the current state.

Algorithm 1 Pre-trained DQN based Deployment Scheme

input: $\mathcal{A}$, $\mathcal{R}_1$, α,γ, I and episode number K, M, pre-trained data sample portion p_{pre}
output: Q-value of each action

1: Initialize pre-trained data cache $\mathcal{D}^{demo}$, replay memory unite $\mathcal{D}$, and the evaluate and the target Q Networks with random weights.
Pre-training:
2: **for** episode $= 1 : K$ **do**
3: Sample a mini-batch (s_j, a_j, r_j, s_{j+1}) from $\mathcal{D}^{demo}$
4: Perform a gradient descent step on L_{pre} based on (56)-(57) for the evaluate Q networks
5: Every N steps update the target Q network with the evaluate Q networks
6: **end for**
Training:
7: **for** episode $= 1 : M$ **do**
8: Initialize state sequence s_1
9: **for** $i = 1 : I$ **do**
10: Select a random action a_i with probability ϵ
11: Otherwise select a_i based on (52)
12: Execute the selected action a_i in the emulator
13: Observe reward r_i and update to new state s_{i+1}
14: Store data (s_i, a_i, r_i, s_{i+1}) into $\mathcal{D}$
15: Sample a mini-batch (s_j, a_j, r_j, s_{j+1}) from $\mathcal{D}$
16: Set $Y_j = \begin{cases} \mathcal{R}_j, & \text{terminal } s_{j+1} \\ \mathcal{R}_j + \max\limits_{a' \in \mathcal{A}} Q\left(s_{j+1}, a'; \theta_{NN}\right), & \text{otherwise} \end{cases}$
17: Perform a gradient descent step on $(Y_j - Q(s_j, a_j; \theta_{NN}))^2$ based on (54)-(55)
18: Every N steps update the target Q network with the evaluate Q networks
19: **end for**
20: **end for**
21: **return** the target Deep Q Network

The whole training process of the pre-trained based DQN scheme for UAV 3D position deployment, including pre-training and training, is summarized in Algorithm 1. In Algorithm 1, the neural network is pre-trained (line 2-6), and a greedy strategy is used to select the actions (line 10 and 11). Then the existing state, action, reward, and next state are obtained and deposited into the replay memory unite (line 12-14). Finally, a mini batch of data is sampled from the replay memory unite to update the Q-network (line 15-18). With Algorithm 1, the UAV is able to interact with the environment without any prior knowledge continuously and learn near-optimal current deployment strategies efficiently and intelligently to improve the average number of SSVs.

Remark: The time complexity of Algorithm 1 is determined

TABLE I
SIMULATION PARAMETERS

Parameter	Value
UAV transmit power P_u [26]	0.5 W
V2V transmit power P_v [26]	0.5 W
Additional attenuation factors about NLoS η	20 dB
Parameters for dense urban environment a, b [27]	11.95, 0.136
Path loss exponent for U2V transmission α_u [21]	3
Path loss exponent for V2V transmission α_v [21]	3.5
Noise power σ^2 [21]	-120 dB/Hz
V2V maximum communication range $D_{\max}$ [27]	100 m
Content popularity (Zipf law) skewness z	2
UAV horizontal motion power parameters c_1, c_2 [27]	272.6, 0.02175
UAV vertical motion power parameters c_3, c_4 [27]	19.6, 127.32
UAV maximum battery capacity $E_{\max}$ [27]	100,000J
DQN learning rate α	0.01
DQN discount factor γ	0.9
DQN weight of reward function ω_p, ω_e	0.9, 0.1

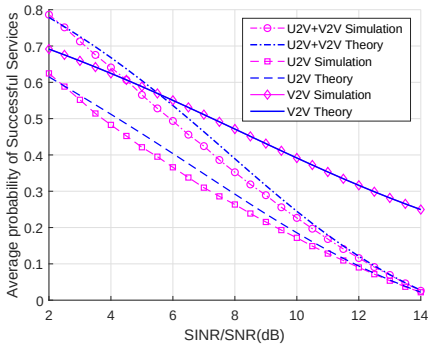

Fig. 4. Successful service probability.

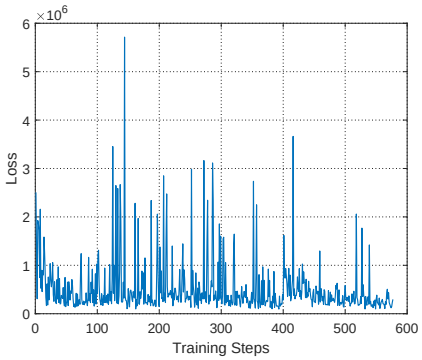

Fig. 5. Loss of training.

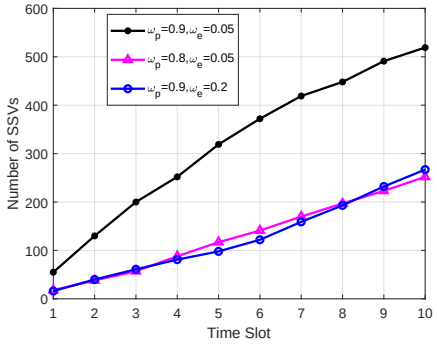

Fig. 6. Average number of successfully served users over time (VU density=16/1000m^2, SINR threshold=6 dB, distance of hot-spot centers=600m).

by the process of pre-training and training. Since the time complexity of pre-training is $\mathcal{O}(K)$ where K is the number of episodes, while the time complexity of training is $\mathcal{O}(MI)$ where M is the number of episodes and I is the number of steps in each episode, the time complexity of Algorithm 1 should be $\mathcal{O}(K) + \mathcal{O}(MI)$.

VI. NUMERICAL SIMULATION

In this section, simulations are performed to validate our theoretical analysis and assess the performance of the proposed framework. In the simulations, the vehicular network is located within an area of $1,200m \times 1,200m$ in a dense city, where the mobility of the users follows the homogeneous Poisson point process with parameter $\lambda = 48$. The peak traffic period T_e is assumed to sustain 30 minutes, and the number of time slots I is set to 10. Within the area, a UAV is dispatched to provide auxiliary communication. The parameters of UAV communication and energy consumption are shown in Table I. What's more, the parameters of the proposed DQN scheme as well as other default parameter settings are shown in Table I.

A. Features of Proposed Scheme

We first verify our theoretical analysis for the single hot-spot scenario. From the results shown in Fig. 4, we can see that for all the SSP of U2V transmission, V2V transmission, and V2V-assisted UAV communications, which decrease as the thresholds of SNR or SINR increase, the lines of theoretical results and the lines of simulation results are very close or even overlapping. This means that our theoretical analysis matches the simulation perfectly, and thus is soundness. From Fig. 4, we can also see that the thresholds of SINR and SNR are small, $\overline{p}_{succ,\text{U2V}}$ and $\overline{p}_{succ,\text{V2V}}$ are relatively high, making $\overline{p}_{succ}$ high. As the thresholds of SINR and SNR increase, $\overline{p}_{succ,\text{U2V}}$ and $\overline{p}_{succ,\text{V2V}}$ decrease, making $\overline{p}_{succ}$ drop rapidly. When $\overline{p}_{succ}$ drops and is smaller than $\overline{p}_{succ,\text{V2V}}$, the intersection of lines appears. What's more, when the thresholds of SINR and SNR are very large, the number of AVUs will become very small. In such a case, the V2V mode is almost unavailable and $\overline{p}_{succ}$ is approaching to $\overline{p}_{succ,\text{U2V}}$.

Next, we verify the convergence of the proposed scheme for the multi-hot-spot scenario. The DQN model is trained with 300 episodes, and in each episode, the UAV undergoes position deployment adjustments during 10 time slots. From the results shown in Fig. 5, we can see that as the training progresses, the training error gradually decreases and eventually reaches a converging trend by the end of the training, which verifies the convergence of the proposed scheme.

Finally, we investigate how the SSP of V2V-assisted UAV communications is affected by the setting parameters such as ω_p and ω_e. It can be observed from Fig. 6 that when ω_p is given, increasing ω_e leads to a gradual decrease in the number of SSVs. This is because the scheme weighs more on energy consumption and gains energy savings at the cost of service degradation. Conversely, when ω_e is given, increasing ω_p results in a gradual increase in the number of SSVs, since

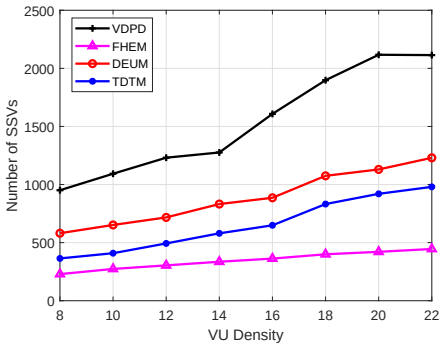

Fig. 7. Total number of SSVs vs. VU density (SINR threshold=6 dB, distance of hot-spot centers=600m).

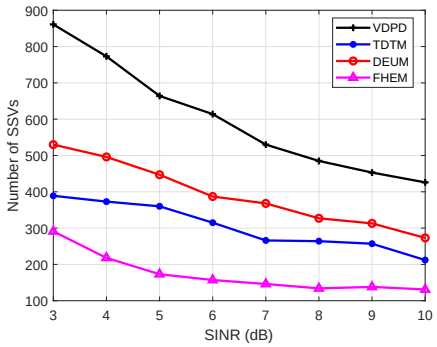

Fig. 8. Total number of SSVs vs. SINR threshold (VU density=16/1000m^2, distance of hot-spot centers=600m).

the service quality weighs more in this case.

B. Performance Evaluation

To verify the performance of our proposed V2V-assisted 3D Position Deployment (VDPD) scheme in the multi-spot scenario, we compare it with three existent algorithms, i.e., the Trajectory Design for Throughput Maximization (TDTM) scheme [28], the Fly-Hover Energy Minimization (FHEM) scheme [30], and the UAV Deployment with Energy and User Mobility awareness (DEUM) scheme [29]. In the TDTM scheme, the hovering position of the UAV is optimized to maximize the coverage of the UAV and the total throughput of each mission. In the FHEM scheme, the hovering locations and durations as well as the flying trajectory is optimized to minimize the total UAV energy consumption. While in the DEUM scheme, the 3D position and flying trajectory of the UAV is optimized to maximize the coverage of the UAV while minimizing the UAV energy consumption.

Trajectory Design for Throughput Maximization in UAV-Assisted Communication System

We first verify the performance of the algorithms in terms of the total number of SSVs. From Fig. 7, we can see that as the density of VUs in the hot-spots increases, the total number of SSVs of all the algorithms grows. Obviously, the performance of the proposed VDPD scheme has superior advantages to that of DEUM, TDTM, and FHEM, with an increase by 76%, 139%, and 350%, respectively. This is because both the TDTM and FHEM schemes are planning the UAV's trajectory in a two-dimensional plane without considering the height of the UAV. Although the DEUM scheme obtains the 3D position of the UAV, it only considers the single hot-spot scenario. When encountering a multi-spot scenario, it purely serves the VUs in the spot with the highest density, leaving VUs in other spots unserved. What's more, these algorithms do not exploit the V2V communication which can further improve the performance.

Fig. 8 shows the number of SSVs vs. the SINR threshold. We can see that as the SINR threshold increases, the number of SSVs will decrease. Moreover, due to the same reasons as having been stated, there are more VUs in the coverage of the proposed VDPD scheme as the SINR threshold changes. Thus the number of SSVs in VDPD gains an increase by 105%, 221%, and 360% when compared to DEUM, TDTM, and FHEM, respectively.

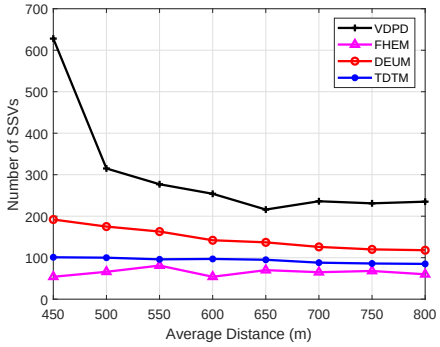

Fig. 9. Total number of SSVs vs. distance of hot-spot centers (VU density=16/1000m^2, SINR threshold=6 dB).

Fig. 9 shows the number of SSVs vs. the average distance of centers of multiple hot-spots. We can see that as the average distance of centers of multiple hot-spots increases, the number of SSVs of all the algorithms decreases. However, the VDPD scheme has the largest number of SSVs in each case when compared to other algorithms. This is because the position of the UAV in the VDPD scheme is regulated in time according to the movement of VUs and the hot-spot distribution, while the TDTM and FHEM schemes do not consider the mobility of the VUs and the DEUM scheme purely considers the single hot-spot scenario, thus the UAV in these schemes cannot be located in the optimal positions.

Next, we verify the performance of the schemes in terms of energy consumption, where the energy consumption is observed to evolve according to the density of VUs in the hot-spot, the SINR threshold, and the average distance of centers of multiple hot-spots. It can be seen from Fig. 10 to Fig. 12 that the energy consumption of the DEUM scheme, which is followed by the TDTM scheme, is the highest in three different cases due to the uncertain position of the UAV and the hover points between time slots having too large a distance, although it takes the energy consumption into consideration. While the VDPD scheme and FHEM scheme which take the energy consumption into account when adjusting the position of UAV and have similar performance, consume only about 32% and 18% of the energy of DEUM and TDTM, respectively.

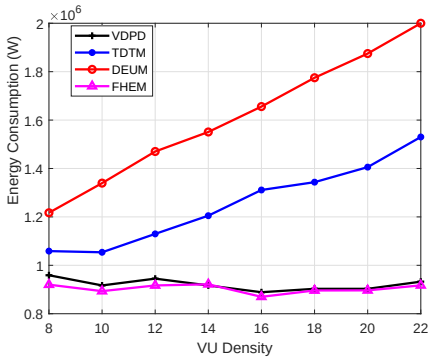

Fig. 10. Energy consumption vs. VU density (SINR threshold=6 dB, distance of hot-spot centers=600m).

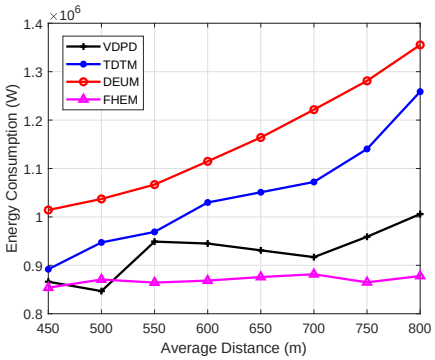

Fig. 12. Energy consumption vs. distance of hot-spot centers (VU density=16/1000m^2, SINR threshold=6 dB).

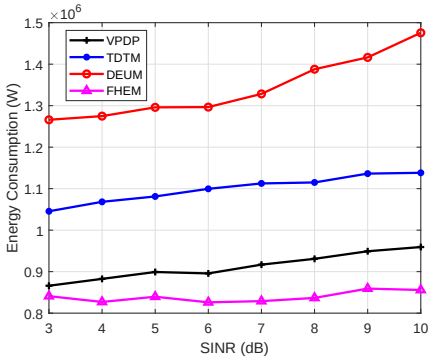

Fig. 11. Energy consumption vs. SINR threshold (VU density=16/1000m^2, distance of hot-spot centers=600m).

In general, the TDTM, FHEM, and DEUM schemes are proposed for the UAV in a single hot-spot scenario where the difference in VUs density is insignificant, and they do not consider the communication between V2V. In our VDPD scheme, the UAV moves according to the VU's distribution, taking energy consumption into consideration and exploiting V2V communication. As a result, the VDPD scheme performs well in terms of the number of SSVs and energy consumption.

VII. Conclusions

In this paper, we investigate a V2V-assisted UAV communications framework where the VUs can request contents from either the UAV or the VUs that have been served successfully by the UAV. For the single hot-spot scenario in which one UAV is deployed in the center of hot-spot, we analyze the SSP of U2V and V2V transmission when the vehicle mobility and social proximity are involved. Furthermore, considering the dynamic of the users and the hot-spots, we extend our analysis to the multi-hot-spot scenario, where the UAV needs to adjust its 3D position to serve more users with better services under the constraint of energy. A pre-trained based DQN scheme is proposed to find the sub-optimal solution. The simulation results demonstrate that our theoretical findings align closely with the simulation results. Moreover, our proposed approach uses only 69% energy to achieve an increase in the number of SSVs by 168% on average when compared to the existing algorithms in the literature. Based on the achievements in this paper, it is more interesting to study the scenario where more than one UAVs are evolved. Considering the costs and rewards in providing service in a dynamic environment, the 3D position deployment of multiple UAVs can be formulated as a decision making process of multiple agents. Such an interesting direction will be pursued in our future work.

References

[1] R. Liu, A. Liu, Z. Qu and N. Xiong, "An UAV-Enabled Intelligent Connected Transportation System With 6G Communications for Internet of Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2045-2059, Feb. 2023.

[2] J. Hu, C. Chen, L. Cai, M. R. Khosravi, Q. Pei and S. Wan, "UAV-Assisted Vehicular Edge Computing for the 6G Internet of Vehicles: Architecture, Intelligence, and Challenges," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 12-18, Jun. 2021.

[3] J. Li, Y. Niu, H. Wu, B. Ai, R. He, N. Wang, and S. Chen, "Joint Optimization of Relay Selection and Transmission Scheduling for UAV-Aided mmWave Vehicular Networks," *IEEE Transactions on Vehicular Technology*, Vol. 72, no. 5, Apr. 2023.

[4] P. Chen, X. Zhou, J. Zhao, F. Shen, and S. Sun, "Energy-Efficient Resource Allocation for Secure D2D Communications Underlaying UAV-Enabled Networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7519-7531, Jul. 2022.

[5] Z. Su, "Energy-Efficiency Optimization for D2D Communications Underlaying UAV-Assisted Industrial IoT Networks with SWIPT," *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 1990-2002, Feb. 2023.

[6] B. Li, W. Xie, Y. Ye, L. Liu, and Z. Fei, "FlexEdge: Digital Twin-Enabled Task Offloading for UAV-Aided Vehicular Edge Computing," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 11086-11091, Aug. 2023.

[7] Z. Wu, Z. Yang, C. Yang, J. Lin, Y. Liu, and X. Chen, "Joint Deployment and Trajectory Optimization in UAV-assisted Vehicular Edge Computing Networks," *Journal of Communications and Networks*, vol. 24, no. 1, pp. 47-58, Feb. 2022.

[8] P. Wu, F. Xiao, H. Huang, and R. Wang, "Load Balance and Trajectory Design in Multi-UAV Aided Large-Scale Wireless Rechargeable Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13756-13767, Nov. 2020.

[9] M. M. Islam, M. M. Saad, M. T. Raza Khan, and S. H. A. Shah, "Proactive UAVs Placement in VANETs," in *IEEE International Conference on Communications*, Seoul, South Korea, May 2022.

[10] W. Feng, "Joint 3D Trajectory Design and Time Allocation for UAV-Enabled Wireless Power Transfer Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9265-9278, Sep. 2020.

[11] Z. Huang, C. Chen, and M. Pan, "Multiobjective UAV Path Planning for Emergency Information Collection and Transmission," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6993-7009, Aug. 2020.

[12] B. Zhu, E. Bedeer, H. H. Nguyen, R. Barton, and J. Henry, "UAV Trajectory Planning in Wireless Sensor Networks for Energy Consumption Minimization by Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 9540-9554, Sep. 2021.

[13] T. Yuan, C. E. Rothenberg, K. Obraczka, C. Barakat, and T. Turletti, "Harnessing UAVs for Fair 5G Bandwidth Allocation in Vehicular Communication via Deep Reinforcement Learning," *IEEE Transactions on Network and Service Management*, vol. 18, no. 4, pp. 4063-4074, Dec. 2021.

[14] M. Hosseini and R. Ghazizadeh, "Stackelberg Game-Based Deployment Design and Radio Resource Allocation in Coordinated UAVs-Assisted Vehicular Communication Networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 1, pp. 1196-1210, Jan. 2023.

[15] J. Wang, X. Zhou, H. Zhang and D. Yuan, "Joint Trajectory Design and Power Allocation for UAV Assisted Network with User Mobility," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 10, pp. 13173-13189, Oct. 2023.

[16] R. Zhou, H. Chen, H. Chen, E. Liu and S. Jiang, "Research on Traffic Situation Analysis for Urban Road Network Through Spatiotemporal Data Mining: A Case Study of Xi'an, China," *IEEE Access*, vol. 9, pp. 75553-75567, May 2021.

[17] Y. Kawamoto, T. Mitsuhashi, and N. Kato, "UAV-Aided Information Diffusion for Vehicle-to-Vehicle (V2V) in Disaster Scenarios," *IEEE Transactions on Emerging Topics in Computing*, vol. 10, no. 4, pp. 1909-1917, Oct.-Dec. 2022.

[18] Q. Wang, Y. Chen, S. Yin, L. Tian, and Y. Guo, "Pricing based Power Control for Inter-cell UAV-assisted Vehicle-to-vehicle Underlay Communication," *China Communications*, vol. 16, no. 1, pp. 57-68, Jan. 2019.

[19] L. Wang, K. Wang, C. Pan, W. Xu, N. Aslam and L. Hanzo, "Multi-Agent Deep Reinforcement Learning-Based Trajectory Planning for Multi-UAV Assisted Mobile Edge Computing," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 73-84, Mar. 2021.

[20] H. Peng and X. Shen, "Multi-Agent Reinforcement Learning Based Resource Management in MEC- and UAV-Assisted Vehicular Networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 131-141, Jan. 2021.

[21] Z. He, B. Zhang, J. Wang, L. Wang, Y. Ren, and Z. Han, "Performance Analysis and Optimization for V2V-assisted UAV Communications in Vehicular Networks," in *Proceedings of IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun. 2020.

[22] L. Zhao and H. Song, "Low-Latency D2D Pairing Strategy Based on Joint Attributes," in *2023 IEEE International Conference on Smart Internet of Things*, Xining, China, pp. 184-189, Aug. 2023.

[23] W. Song, Y. Zhao and W. Zhuang, "Stable Device Pairing for Collaborative Data Dissemination With Device-to-Device Communications," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1251-1264, Apr. 2018.

[24] C. Yan, L. Fu, J. Zhang, and J. Wang, "A comprehensive survey on UAV communication channel modeling," *IEEE Access*, vol. 7, pp. 107769-107792, Aug. 2019.

[25] S. Chai and V. K. N. Lau, "Online Trajectory, Radio Resource Optimization of Cache-Enabled UAV Wireless Networks With Content and Energy Recharging," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1286-1299, Apr. 2020.

[26] O. M. Bushnaq, A. Celik, H. Elsawy, M. Alouini, and T. Al-Naffouri, "Aeronautical Data Aggregation and Field Estimation in IoT Networks: Hovering and Traveling Time Dilemma of UAVs," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4620-4635, Oct. 2019.

[27] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Social Big-data-based Ccontent Dissemination in Internet of vehicles," *IEEE Trans. on Industrial Informatics*, vol. 14, no. 2, pp. 768-777, Feb. 2018.

[28] N. Gupta, S. Agarwal and D. Mishra, "Trajectory Design for Throughput Maximization in UAV-Assisted Communication System," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1319-1332, Sept. 2021.

[29] D. N. Anwar, M. Peer, K. Lata, A. Srivastava and V. A. Bohara, "3-D Deployment of VLC Enabled UAV Networks With Energy and User Mobility Awareness," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 4, pp. 1972-1989, Dec. 2022.

[30] Y. Zeng, J. Xu and R. Zhang, "Energy Minimization for Wireless Communication With Rotary-Wing UAV," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329-2345, Apr. 2019.