Minimizing Age of Information in UAV-Assisted Data Collection With Limited Charging Facilities

Yunkai Wei, Yuxin Lu, Pengcheng Zhao, Supeng Leng, *Senior Member, IEEE*, and Kun Yang, *Fellow, IEEE*

***Abstract*—Age of Information (AoI) has become a new metric for data freshness in Unmanned Aerial Vehicle (UAV)-assisted data collection. Whereas, the UAVs are usually energy-limited. When they leave to be charged, the AoI may increase additionally. Current studies of optimizing the AoI usually consider the situation where charging facilities are sufficient, which is not true in many scenarios. For the situation where only one UAV can be charged at one time, we propose a multi-agent cooperative Deep Q-Network algorithm with delayed reward (DR-MADQN) to minimize the AoI. Simulation results show DR-MADQN can effectively decrease the AoI and charge waiting time.**

***Index Terms*—Unmanned aerial vehicle (UAV), age of information (AoI), Internet of Things (IoT), wireless charging, deep Q-network (DQN).**

I. Introduction

UNMANNED Aerial Vehicles (UAVs) have been widely used to collect sensory data in various Internet of Things (IoT) applications, such as environmental monitoring [1], disaster management [2] and military reconnaissance. In such UAV-assisted data collection, Age of Information (AoI) is usually used to measure the freshness of the sensory data, which is an important metric in latency sensitive applications. Consequently, scheduling the data collection strategy of the UAVs to optimize AoI has become a crucial problem.

Studies have been conducted to optimize the AoI in UAV-assisted data collection. Zhou et al. [3] proposed an online AoI-based trajectory planning (A-TP) algorithm to decrease the AoI. Literature [4] optimized the AoI of the road-side IoT devices. Zhang et al. [5] proposed a twin delayed deep deterministic policy gradient (TD3) algorithm to minimize the AoI of massive IoT devices. However, these studies have not considered the energy limitation in the UAVs. To take the energy into consideration, Dai et al. [6] introduced a relational graph convolutional network (RGCN) mechanism for extracting UAV-user correlations and ultimately minimized the AoI with efficient use of constrained energy reserve. Literature [7] and [8] proposed schemes to jointly optimize the AoI of devices and the energy consumption of UAV to prevent UAV from running out of energy.

The above studies either neglected energy consumption or aimed to reduce it. In order to make the UAVs continuously work for a sufficiently long period, some studies explored recharging them. Literature [9] proposed to deploy multiple charging facilities, with each UAV accessing a fixed charging station when needed. Li et al. [10] proposed a concept of the nondisruptive wireless rechargeable UAV network (WRUN), in which UAVs can be charged while flying. These studies assumed sufficient charging facilities. However, this is not necessarily true in many scenarios, where the UAVs may have to wait in line to be charged.

To address the above issue, in this letter, we propose an AoI optimization model for energy-limited multiple UAV-assisted IoT data collection, where a high-altitude platform (HAP) can charge only one UAV simultaneously. HAP can be rapidly deployed to the mission area, including inaccessible regions for human, disaster and military fields, where ground charging facilities are hard to be deployed [11]. And it can be deployed at optimal position based on the objective environment [12]. By carrying solar panels, HAP can continuously provide charging services to the UAVs [13], [14]. Then, a multi-agent cooperative Deep Q-Network algorithm with delayed reward (DR-MADQN) is proposed to minimize the average AoI. The performance of the algorithm is verified by the simulation. The contributions can be summarized as follows:

- We model the issue of multiple data-collecting UAVs with one charging facility into an optimization problem, so as to jointly optimize the UAVs' trajectory and charging duration, and minimize the average AoI of IoT devices.
- We form the above optimization model into a Markov problem, and present a DR-MADQN algorithm to find the optimal solution. Unlike traditional MADQN, the DR-MADQN provides a delayed reward upon reaching the target instead of an instant reward in each time slot.
- Simulations are conducted to verify the performance of DR-MADQN. The results show that DR-MADQN obtains much lower average AoI and charge waiting time than the benchmark schemes.

The rest of this letter is organized as follows. Section II presents the system models and formulates the optimization model. In Section III, the DR-MADQN is proposed to find the optimal solution of the model. Simulation results and

Yunkai Wei, Yuxin Lu, Pengcheng Zhao, and Supeng Leng are with the Yangtze Delta Region Institute (Quzhou) and the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: ykwei@uestc.edu.cn; 202122011938@std.uestc.edu.cn; 202221011129@std.uestc.edu.cn; spleng@uestc.edu.cn).

Kun Yang is with the School of Information and Communication Engineering and the Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: kyang@ieee.org).

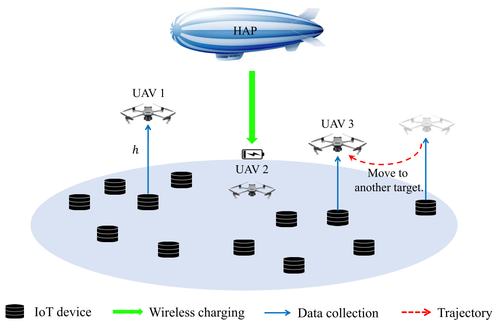


Fig. 1. UAV-assisted data collection for IoT devices.

performance evaluations are provided in Section IV. Finally, Section V concludes this letter.

II. Preliminary and Problem Formulation

As shown in Fig. 1, we consider a UAV-assisted data collection scenario with n IoT devices $\{d_1, d_2, \ldots, d_n\}$ randomly distributed in the target area, generating sensory data periodically. m energy-limited UAVs $\{v_1, v_2, \ldots, v_m\}$ $(m \ll n)$ flying above these devices to collect the data. A HAP hovering above this area can charge only one UAV simultaneously. The UAV will land on the ground vertically below the HAP to be charged or waiting to be charged. Assume the time is split into time slots $\{1, 2, \ldots, T\}$ with equal slot duration t_0.

A. Channel Model

The link between the UAV and the devices may be line-of-sight (LoS) or non-line-of-sight (NLoS) subject to some possibility distribution. Therefore, we use a probabilistic air-to-ground (A2G) path loss channel model which is the possibility statistics of the LoS and NLoS [15]. This model takes into account the uncertainties such as terrain features and building vegetation. The LoS probability can be shown as [16]:

$$P_{\mathrm{LoS}}(t) = \frac{1}{1 + \varphi \exp[-\psi(\theta - \varphi)]}, \tag{1}$$

where φ and ψ are two environment-dependent constants, θ is the elevation angle between the UAV and IoT device. Accordingly, the occurrence probability of NLoS is $P_{\mathrm{NLoS}}(t) = 1 - P_{\mathrm{LoS}}(t)$. The path loss of LoS and NLoS link is:

$$PL_{\mathrm{LoS}}(t) = 20\log_{10}\left(\frac{4\pi f_c h}{c}\right) + \xi_{\mathrm{LoS}}, \tag{2}$$

$$PL_{\mathrm{NLoS}}(t) = 20\log_{10}\left(\frac{4\pi f_c h}{c}\right) + \xi_{\mathrm{NLoS}}, \tag{3}$$

where f_c and c are the carrier frequency and the speed of light, ξ_{LoS} and ξ_{NLoS} are the additional path loss under the LoS and NLoS, respectively, related to the environment. When collecting data, the UAV hovers over the IoT device, so the distance between the UAV and device is the flying altitude h of UAV. Then the average path loss can be expressed as:

$$\overline{PL(t)} = P_{\mathrm{LoS}}(t) \times PL_{\mathrm{LoS}}(t) + P_{\mathrm{NLoS}}(t) \times PL_{\mathrm{NLoS}}(t). \tag{4}$$

Consequently, the data rate can be expressed as (5), where B is the channel bandwidth, p is the transmission power of the UAV, and N_0 is the noise power.

$$\mathcal{R}(t) = B\log_2\left(1 + \frac{p}{\overline{PL(t)} \times N_0}\right). \tag{5}$$

B. Energy Model

The laser charging efficiency of the HAP is:

$$\eta_i = \frac{p_i^r}{p_0} = \eta_i^{lt}\eta_i^{le} = e^{-\alpha(H-h)}\eta_i^{le} \tag{6}$$

where p_i^r is the charging power received by v_i, p_0 is the source laser power emitted by the HAP, η_i^{lt} and η_i^{le} represent the transmission efficiency of laser in air and the conversion efficiency of laser to electricity, respectively, H is the altitude of the HAP, and α is the laser attenuation coefficient.

$$\alpha = \frac{\sigma}{\kappa}\left(\frac{\lambda}{\chi}\right)^{-\rho}, \tag{7}$$

where σ and χ are two constants, κ is the visibility, λ is the wavelength, and ρ is the size distribution of the scattering particles [13]. Let $c_i(t)$ denote the charging status of v_i, where $c_i(t) = 1$ means v_i is being charged.

In general, the UAV's energy consumption consists of propulsion energy and communication-related energy. Communication-related functions, including circuits, data transmission and processing, consume energy represented by the transmission power $e_{com} = p$ of the UAV. Let e_{fly} denote the flying power of the UAV, and e_{hov} for that of the hovering power. Then, the remaining energy of the UAV v_i is:

$$E_i(t) = E_i^0 + p_i^r T_i^{hc} - (e_{com} + e_{hov})T_i^{hd} - e_{fly}T_i^{fl}, \tag{8}$$

where E_i^0 is the initial energy of v_i. T_i^{hc}, T_i^{hd} and T_i^{fl} denote the time for charging, collecting and flying of v_i, respectively.

C. AoI of IoT Devices

In UAV-assisted data collection, AoI is used to measure the freshness of the sensory data, which is defined as the time elapsed between the generation of a data packet at the source node and its update at the destination. Thus, the AoI of d_j at time-slot t is $A_j(t) = t - U_j(t)$, where $U_j(t)$ refers to the generation time.

Each device generates new data at the beginning of each time slot, which will overwrite the old data. So, $A_j(t)$ is incremented whenever the new data of d_j is not collected by a UAV. Otherwise, when a UAV collects the new data and updates it at the end of this time slot, $A_j(t)$ will reset to 1. The AoI of d_j can be expressed as:

$$A_j(t) = \begin{cases} 1, & \text{if } \exists\ b_i^j(t) = 1, \\ A_j(t-1) + 1, & \text{otherwise}, \end{cases} \tag{9}$$

where $b_i^j(t)$ denotes the collecting status of v_i and d_j. When the data of d_j is being collected by v_i, there is $b_i^j(t) = 1$. A UAV can only serve at most one IoT device at the same time, consequently, there is $\sum_{j=1}^{n} b_i^j(t) \leq 1$.

Figure 2 shows the AoI of d_j over 8 time-slots. In this figure, d_j generates new data at the beginning of each time slot, which is collected at the end of time slot 3, 4, and 8.

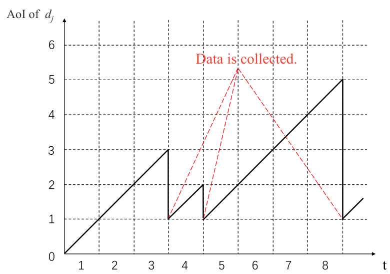


Fig. 2. Associate AoI of d_j in 8 time-slots.

The average AoI of all IoT devices in T time slots is given by the following equation:

$$\overline{A(t)} = \frac{\sum_{t=1}^{T} \sum_{j=0}^{n} \omega_j A_j(t)}{n \times T}, \tag{10}$$

where ω_j means the priority of devices, depending on the importance and type of the data generated by d_j.

D. Problem Formulation

To minimize the average AoI in scenario with a single charging facility and only one UAV can be charged at one time, the flying trajectory and charging schedule of the UAVs need to be jointly optimized. Therefore, we propose an optimization model to minimize the average AoI:

$$\min_{L_i, T_i^{hc}} \quad \overline{A(t)} \tag{11}$$

$$\text{s.t.} \quad x_{\min} < x_i < x_{\max}, \tag{11a}$$

$$y_{\min} < y_i < y_{\max}, \tag{11b}$$

$$\sum_{j=1}^{n} b_i^j(t) \leq 1, \tag{11c}$$

$$E_i(t) > 0, \tag{11d}$$

$$\sum_{i=1}^{m} c_i(t) \leq 1, \tag{11e}$$

where L_i and T_i^{hc} denote the position and charging duration of UAV, respectively. Eqs. (11a) and (11b) define the flying area of the UAVs. Eq. (11c) restricts v_i to serving at most one device at the same time. The remaining energy of the UAV must be greater than zero at any time in (11d). The HAP can only charge at most one UAV at one time in (11e).

III. Multi-Agent Cooperative Deep Q-Network Algorithm With Delayed Reward

To solve this optimization problem with low complexity, an appropriate solution is to model it as a Markov problem with a large state space. Based on the multi-agent deep reinforcement learning framework [17], a DR-MADQN model is constructed with UAV as the agent. The state space is S, action space is A, and the reward function is R. In each training episode, the UAV observes the current environment to get s, selects action a, and subsequently reaches the IoT device or HAP in a defined time step post-action. It then receives a reward r, and progresses to the next state s'. The specific state space, action space and reward function will be given in detail below.

A. State Space

The state space of the system is composed of the states of IoT devices and UAVs. IoT devices are randomly distributed on the ground and periodically generate sensory data. The AoI will increase linearly with time until the data is collected and the AoI is reset to one. Therefore, the state space of d_j is represented by its AoI $A_j(t)$ and the data collection state $b_i^j(t)$.

The UAV exhibits dynamic positional changes and experiences continuous variations in residual energy during flight, data collection hover, and charging phases. Therefore, the state space of v_i should contain position $L_i(t)$, remaining energy $E_i(t)$ and charging states $c_i(t)$. Then, the state space can be represented by $S = \{A_j(t), b_i^j(t), L_i(t), E_i(t), c_i(t)\}$.

B. Action Space

Each UAV can choose one of the three actions at one time: flying to a target IoT device or the HAP, hovering above an IoT device to collect data, or landing to be charged. Specifically, each action can be expressed as follows:

- A_f: The UAV is flying to a target IoT device or the HAP. Let $A_f = \{A_f^1, \ldots, A_f^j, \ldots, A_f^n, A_f^c\}$ denote the action space of this part, where $A_f^j (j \in \{1, 2, \ldots, n\})$ and A_f^c denote the UAV chooses to fly to d_j or HAP, respectively.
- A_c: The UAV is landing on the platform of the HAP, being charged or waiting to be charged.
- A_d: The UAV is hovering above an IoT device and collecting the sensory data of this device.

Then, the action space of a UAV is $A = \{A_f, A_c, A_d\}$. When a UAV is choosing its next action, the remaining energy after this action must be enough to fly to the HAP. Consequently, the UAV will select action A_c when $E_i(t) < E_{thr}$. Where $E_{thr} = \frac{\|L_i^{k+1} - L_i^k\| + \|L_i^{k+1}\|}{v} e_{fly} + \frac{D_i^{k+1}}{R(t)}(e_{com} + e_{hov})$ in (11d). L_i^k and L_i^{k+1} define the locations of the k-th and $(k+1)$-th IoT device served by v_i, D_i^{k+1} is the size of data needs to be collected of the $(k+1)$-th device, and v is the flying speed of the UAVs.

C. Reward Function

The UAV's reward comprises two components: a reward for decreasing the average AoI, and an incentive for aligning recharging times with the HAP's availability to minimize charging conflict probabilities. Consequently, there is

$$R_i(t) = R_{AoI}(t) + R_i^{charge}(t), \tag{12}$$

where the reward of the average AoI is denoted as $R_{AoI}(t)$:

$$R_{AoI}(t) = \overline{A(t-1)} - \overline{A(t)}, \tag{13}$$

where $\overline{A(t)}$ and $\overline{A(t-1)}$ denote the average AoI in time slot t and $t-1$, respectively.

$$R_i^{charge}(t) = \begin{cases} r_c, & \text{if } \sum_{i=1}^{m} c_i(t) = 0 \; and \; a_i(t) = A_f^c, \\ 0, & \text{otherwise}, \end{cases} \tag{14}$$

where r_c is a constant, $a_i(t)$ is the action selected by v_i.

Algorithm 1 DR-MADQN Algorithm

Initialize task area $\mathcal{A}$, the positions of HAP $\mathcal{L}_H$ and devices $\mathcal{L}_D$, the initial positions of the UAVs $\mathcal{L}_U$, experience replay buffer *ER*, Q-network $Q(\theta)$, target Q-network $Q'(\theta')$, and $\theta' = \theta$.
for each episode **do**
  Update ε in action policy.
  for t in $1, 2, \dots, T$ **do**
    for each UAV **do**
      Observe the environment and get the state s.
      if $E_i(t) < E_{thr}$ **then**
        Select the action A_c.
      else
        Select a with $\varepsilon - greedy$.
      end if
      if arrive at the IoT device or HAP **then**
        Calculate the reward r and get s'.
        Store $< s, a, r, s' >$ in ER.
        Sample random experiences from buffer.
        Calculate target $y = r + \gamma maxQ(s', a'; \theta')$.
        Train the Q-network parameter θ.
      end if
    end for
    Every N steps, update $\theta' = \theta$.
  end for
end for

D. Framework of the DR-MADQN

For the optimization model of UAV-assisted data collection with one HAP, we present a DR-MADQN algorithm to minimize the average AoI. The framework of the DR-MADQN consists of three parts: the Q-network $Q(\theta)$, the target Q-network $Q'(\theta')$, and the experience replay. The Q-network is responsible for interacting with the environment in real time and evaluating to get the $Q(\theta)$ corresponding to s. The target Q-network computes $Q'(\theta')$ using s', which has the same structure as the Q-network but with different parameters. To break the correlation between the training data, an experience replay is used to store the historical data.

The specific process of the DR-MADQN is summarized as Algorithm 1. Each UAV independently observes the environment to get the s. A $\varepsilon - greedy$ action policy with a decreasing parameter ε is adopted to guide the agent to choose actions:

$$a = \begin{cases} random\ action, & \varepsilon, \\ argmax\ Q(s, a; \theta), & 1 - \varepsilon. \end{cases} \tag{15}$$

Different from the traditional MADQN, the DR-MADQN will give a delayed reward r when the UAV reaches the target, and then store the quadruple $< s, a, r, s' >$ into the experience replay buffer. This design in DR-MADQN enables the UAV to fly straight towards the target to save energy. Additionally, through the design of delayed rewards, it avoids the chaos of the reward system, and can significantly improve the training efficiency. Random samples are taken from the buffer, to train the neural network by minimizing the loss function:

$$J(\theta) = \left(r + \gamma maxQ\left(s', a; \theta'\right) - Q(s, a; \theta)\right)^2, \tag{16}$$

TABLE I
SIMULATION PARAMETERS

Parameter	Description	Value
f_c	Carrier frequency	2.5GHz
(φ,ψ,ξ_{LoS},ξ_{NLoS})	Urban environment	(9.61,0.16,1,20)
h	Altitude of UAVs	10m
H	Altitude of HAP	1km
N_0	Noise power	-100dBm
B	Communication bandwidth	1MHz
p	Transmission power	1W
p_0	Laser beam power	1kW [18]
η	Charging efficiency	70% [13]
e_{hov}	UAV hovering power	170W
e_{fly}	UAV flying power	140W
e_{com}	UAV communication power	1W
v	Flying velocity	5m/s

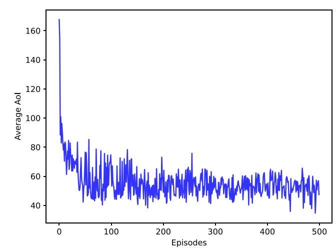


Fig. 3. Convergence of DR-MADQN algorithm.

where γ is a discount factor ($0 < \gamma < 1$). When the Q-network has updated for N steps, its updated parameter is copied to the target network to update θ'.

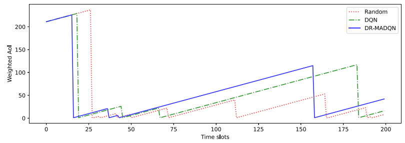
IV. SIMULATION RESULTS

In this section, we conduct numerical simulations with Python 3.9.11 to verify DR-MADQN's effectiveness in a 50m $\times$ 50m square area, with simulation parameters detailed in Table I. Both Q-network and target Q-network feature two fully connected hidden layers, each with 256 neurons utilizing the Rectified Linear Unit (ReLU) as the activation function. We assign priority ω_j to d_j based on the urgency of the information generated: high priority ($w_j = 3$), medium priority ($w_j = 2$), and low priority ($w_j = 1$). We then compare DR-MADQN's performance against traditional DQN and a random-target-selection scheme.

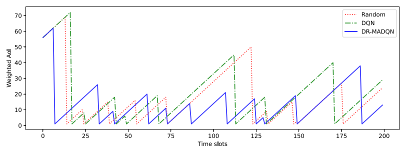
Figure 3 shows the convergence performance of DR-MADQN when 3 UAVs are used to collect data from the devices. Initially, the UAVs explore the environment with random actions, resulting in a high average AoI. Then, after about 300 episodes, DR-MADQN converges to stable solutions.

In Fig. 4, we use the term $weighted\_AoI = \omega_j \times A_j$ to illustrate the influence of the IoT devices with different priorities. Lower priority devices have lower collection frequency, resulting in higher average AoI at times. However, higher priority devices consistently achieve smaller average AoI compared to benchmark algorithms. Overall, DR-MADQN significantly reduces average AoI for all IoT devices.

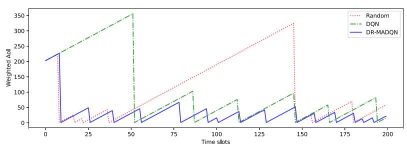
In Figure 5(a), DR-MADQN consistently achieves the lowest average AoI across 8-16 IoT devices, outperforming benchmark algorithms by reducing the average AoI up to 31.32% and 45.08% compared to traditional DQN and



(a) AoI of IoT devices with low priority.

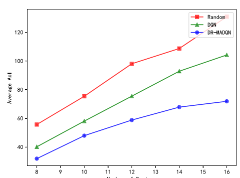


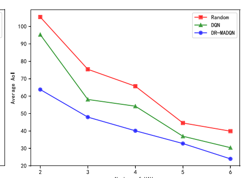
(b) AoI of IoT devices with medium priority.



(c) AoI of IoT devices with high priority.

Fig. 4. AoI of IoT devices with different priority.





(a) Average AoI versus the number of IoT devices.

(b) Average AoI versus the number of UAVs.

Fig. 5. Average AoI versus the IoT device and UAV number.

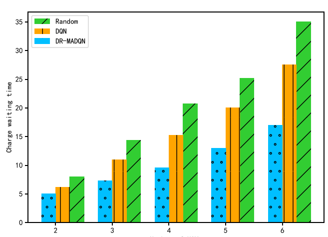


Fig. 6. Average charge waiting time versus the UAV number.

random-target-selection scheme, respectively. In Figure 5(b), increasing UAV density results in a decrease in average AoI, with DR-MADQN consistently maintaining the lowest average AoI.

Figure 6 shows the average charge waiting time of the UAVs when the number of the UAVs increases. Compared with the benchmark schemes, the DR-MADQN can decrease the average charge waiting time by up to 38.20% and 51.43%. Meanwhile, DR-MADQN has a relatively smaller increase in charge waiting time when the number of the UAVs increases.

V. Conclusion

In this letter, we address energy-limited UAVs in IoT data collection, where only one UAV can be charged at one time slot. We have formulated this problem into an optimization model to minimize the average AoI of IoT devices. Our proposed DR-MADQN significantly decreases the average AoI and the charge waiting time compared to benchmark schemes. For future work, we can consider a time-varying priority of devices, whereby the location of the HAP could be optimized based on the changing priority of IoT devices.

References

[1] H. Zhang, L. Dou, B. Xin, J. Chen, M. Gan, and Y. Ding, "Data collection task planning of a fixed-wing unmanned aerial vehicle in forest fire monitoring," *IEEE Access*, vol. 9, pp. 109847–109864, 2021.

[2] N. Lin, Y. Liu, L. Zhao, D. O. Wu, and Y. Wang, "An adaptive UAV deployment scheme for emergency networking," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2383–2398, Apr. 2022.

[3] C. Zhou et al., "Deep RL-based trajectory planning for AoI minimization in UAV-assisted IoT," in *Proc. 11th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2019, pp. 1–6.

[4] R. Han, J. Wang, L. Bai, J. Liu, and J. Choi, "Age of information and performance analysis for UAV-aided IoT systems," *IEEE Internet Things J.*, vol. 8, no. 19, pp. 14447–14457, Oct. 2021.

[5] J. Zhang, K. Kang, M. Yang, H. Zhu, and H. Qian, "AoI-minimization in UAV-assisted IoT network with massive devices," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2022, pp. 1290–1295.

[6] Z. Dai et al., "AoI-minimal UAV crowdsensing by model-based graph convolutional reinforcement learning," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2022, pp. 1029–1038.

[7] O. S. Oubbati, M. Atiquzzaman, H. Lim, A. Rachedi, and A. Lakas, "Synchronizing UAV teams for timely data collection and energy transfer by deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6682–6697, Jun. 2022.

[8] Z. Wei et al., "UAV-assisted data collection for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15460–15483, Sep. 2022.

[9] Y. Zhang, Z. Mou, F. Gao, L. Xing, J. Jiang, and Z. Han, "Hierarchical deep reinforcement learning for backscattering data collection with multiple UAVs," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3786–3800, Mar. 2021.

[10] M. Li, L. Liu, Y. Gu, Y. Ding, and L. Wang, "Minimizing energy consumption in wireless rechargeable UAV networks," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3522–3532, Mar. 2022.

[11] (Glob. Syst. Mobile Commun. Assoc., London, UK). *High Altitude Platform Systems: Towers in the Skies*, (2021). [Online]. Available: https://www.gsma.com/futurenetworks/resources/high-altitude-platform-systems-haps-whitepaper-2021/

[12] J. Holis and P. Pechac, "Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas," *IEEE Trans. Antennas Propag.*, vol. 56, no. 4, pp. 1078–1084, Apr. 2008.

[13] Q. Zhang, W. Fang, Q. Liu, J. Wu, P. Xia, and L. Yang, "Distributed laser charging: A wireless power transfer approach," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3853–3864, Oct. 2018.

[14] Y. Du, K. Wang, K. Yang, and G. Zhang, "Trajectory design of laser-powered multi-drone enabled data collection system for smart cities," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[15] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[16] J. Li et al., "Joint optimization on trajectory, altitude, velocity, and link scheduling for minimum mission time in UAV-aided data collection," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1464–1475, Feb. 2020.

[17] R. Zhong, X. Liu, Y. Liu, and Y. Chen, "Multi-agent reinforcement learning in NOMA-aided UAV networks for cellular offloading," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1498–1512, Mar. 2022.

[18] M.-A. Lahmeri, M. A. Kishk, and M.-S. Alouini, "Laser-powered UAVs for wireless communication coverage: A large-scale deployment strategy," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 518–533, Jan. 2023.