Energy optimization and age of information enhancement in multi-UAV networks using deep reinforcement learning

Jeena Kim,[1] Seunghyun Park,[2]
and Hyunhee Park[1,✉]
[1] Department of Information and Communication Engineering, Myongji University - Natural Science Campus, Yongin, The Republic of Korea
[2] Division of Computer Engineering, Hansung University, Seoul, The Republic of Korea
✉ E-mail: hhpark@mju.ac.kr

This letter introduces an innovative approach for minimizing energy consumption in multi-unmanned aerial vehicles (multi-UAV) networks using deep reinforcement learning, with a focus on optimizing the age of information (AoI) in disaster environments. A hierarchical UAV deployment strategy that facilitates cooperative trajectory planning, ensuring timely data collection and transmission while minimizing energy consumption is proposed. By formulating the inter-UAV network path planning problem as a Markov decision process, a deep Q-network (DQN) strategy is applied to enable real-time decision making that accounts for dynamic environmental changes, obstacles, and UAV battery constraints. The extensive simulation results, conducted in both rural and urban scenarios, demonstrate the effectiveness of employing a memory access approach within the DQN framework, significantly reducing energy consumption up to 33.25% in rural settings and 74.20% in urban environments compared to non-memory approaches. By integrating AoI considerations with energy-efficient UAV control, this work offers a robust solution for maintaining fresh data in critical applications, such as disaster response, where ground-based communication infrastructures are compromised. The use of replay memory approach, particularly the online history approach, proves crucial in adapting to changing conditions and optimizing UAV operations for both data freshness and energy consumption.

Introduction: In recent years, the use of unmanned aerial vehicles (UAVs) has expanded extensively across civilian, commercial, and military domains [1–4]. Particularly in environments where ground-based stations are unreliable, UAVs can act as aerial base stations to provide communications during disasters [5–9]. Recent studies have focused on energy optimization and age of information (AoI) improvement in UAV networks, proposing various algorithms and approaches [10, 11]. In a dynamically changing disaster environment that requires real-time status updates, it is crucial to maintain the freshness of the collected data for immediate response and action. (AoI is a critical metric that measures the freshness of data, defined as the time elapsed since the most recent data packet was generated until it is received by the end user [12]. AoI considers the overall temporal aspect of data from its generation to its delivery. In contrast, delay refers to the total time a data packet takes to travel from the source to the destination, including processing, transmission, and propagation delays.

In reference [13], traditional dynamic programming (DP) algorithms and ant colonies (AC) are used to minimize AoI in UAV systems. In reference [14], an explicit formulation for the average AoI in Hamiltonian and non-Hamiltonian cycles using a graph-theoretical approach is presented, providing a mechanism to improve AoI on a given flight path by creating new cycles around specific IoT devices. However, as the number of constraints increases, optimizing the trajectory design of UAVs while minimizing AoI becomes more complex and leads to an NP-hard problem. To address these challenges, deep reinforcement learning(DRL) has been proposed as an effective approach.

In reference [15], a deep Q-network (DQN) is applied to optimize UAV scouting in an edge computing environment, considering energy efficiency and AoI. Specifically, reference [16] explores path design to minimize AoI through cooperative sensing and transmission in the cellular Internet of UAVs, introducing a scheduling method and proposing a composite action actor-critic (CA2C) algorithm based on DRL.

DRL algorithms leverage experience replay memory to store experiences gained over episodes, enabling the agent to take actions that maximize future rewards. Here, we explore whether the properties of experience replay memory can ensure AoI while reducing the energy consumption of UAVs in disaster environments. The main contributions of this thesis are summarized as follows:

- First, we propose a hierarchical UAV deployment structure based on their respective roles for cooperative trajectory planning in disaster environments.
- Second, we propose a scheduling method to ensure AoI while minimizing energy consumption. To this end, we define the inter-UAV network path planning problem as a Markov decision process (MDP) and apply DQN to support real-time decision making.
- Finally, we conduct extensive experimental analysis to evaluate the performance of the proposed approach. Using the average AoI performance metric values, we conduct a simulation analysis to find the appropriate parameters for the learning model. The results suggest that the UAV AoI and energy consumption can be optimized.

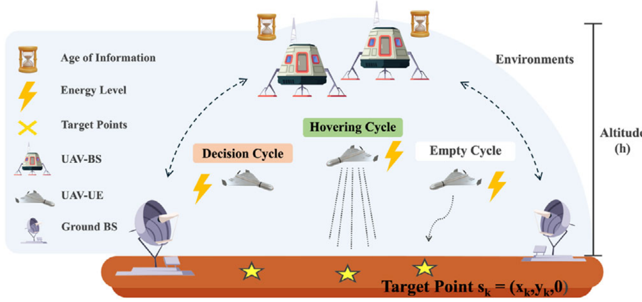

Fig. 1 *Overview of system model*

System model: Here, we address disaster scenarios occurring in rural and urban areas of size $N \times N$ m^2. As illustrated in Figure 1, our model incorporates a hierarchical UAV structure consisting of UAV-base station (BS) linked to operational ground base stations and UAV-user equipment (UE) deployed to collect data over dispersed target points. In this hierarchical structure, the UAV-base station (BS) acts as the central node responsible for the primary calculations and decision-making processes. It coordinates the data collection activities of the UAV-UEs and processes the incoming data to ensure optimal performance and energy efficiency. This central node manages the replay memory, storing the history of all UAVs and utilizing it to dynamically adjust UAV paths based on real-time environmental feedback, thereby overcoming the memory limitations of individual UAVs.The locations where data is generated, positioned at $s_k \in \mathbb{R}^2$ for $1 \leq k \leq K$, are expected to initiate data production at $t_1 \leq 0$. Upon data generation at the kth target point, the UAV-BS identifies the genesis location of data. Positioned at $q_b = (x_b, y_b, h_b) \in \mathbb{R}^3$ for $b = 1, 2, \ldots, B$, each UAV-BS operates at the maximum altitude $h_{\max}$ and acts as a relay node covering area R_b. Subsequent to data generation, the UAV-BS designates the nearest UAV-UE to the target point. During any given time slot t, only a single UAV-UE is allowed to navigate to the target point s_k location. To determine the UAV-UE positioned at the minimum distance to the target point, we calculate the Euclidean distance as follows:

$$d_{3D}(t) = \sqrt{(x_i(t) - x_k(t))^2 + (y_i(t) - y_k(t))^2 + (h_i(t) - h_k(t))^2} \quad (1)$$

We assume the channel model between the ith UAV-UE and the UAV-BS encompasses both large-scale and small-scale fading. Based on a 3D map simulated in reference [15], it precisely discerns the presence of a line-of-sight (LoS) or non-line-of-sight (NLoS) connection. To minimize path loss, the UAV-UE must fly below altitude $h_i < h_{\max}$. If the distance from the UAV-UE's location s_k does not exceed the safety distance δ, the UAV-UE can directly transmit the collected data to the UAV-BS. This safety distance δ, as defined in Equation (2), indicates that the data uploading location falls within the UAV-BS's coverage area. Otherwise, the UAV-UE must re-navigate within a safe distance to the area covered by the UAV-BS R_b.

$$\min_{1 \leq b \leq B} \{\|s_k - q_b\|\} \leq \delta \quad \text{for } 1 \leq k \leq K \quad (2)$$

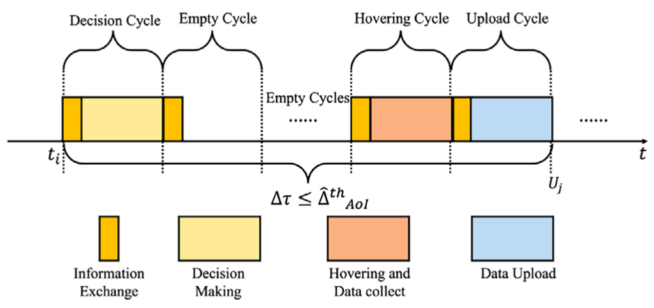

Fig. 2 *Scheduling cycle*

Problem formulation: We introduce AoI energy aware scheduling to efficiently coordinate UAVs under contextual constraints, namely trajectory design, AoI, and energy consumption constraints. As depicted in Figure 2, it consists of five stages. The duration of each cycle is represented by $\tau(t)$.

1) **Information exchange cycle**: Each cycle begins with the exchange of information between UAV-BS and UAV-UE. The information exchanged includes the current position of the ith UAV-UE at time slot t, denoted as $q_{(i,t)} = (x_i, y_i, h_i)$, and the energy consumption during the previous cycle, represented as $E^{\text{cmp}}_{i,\tau(t-1)}$. The energy consumption used for flying the UAV-UE is denoted by $p_{\text{move}}(t) = \sqrt{(\Delta x + \Delta y + \Delta h)}$. Here, Δx, Δy, and Δh represent the distance travelled by the UAV in the x-, y-, and h-axis (altitude) directions, respectively. The variable $o_i(t) \in \{0, 1\}$ indicates whether there is a collision with obstacles. The energy consumption for hovering and uploading is represented by $P_h(t)$ and $\hat{P}_{(b,i)}(t)$, respectively. The AoI, calculated as $\Delta_{\text{AoI}}\tau^i_\tau$, is included. The total duration of five cycles is denoted by $\Delta\tau(t)$, and $U_i(t-1)$ represents the last upload time. This information forms the basis for the next cycle decisions.

$$E_{\text{cmp}} = \frac{1}{I}\sum_{i=1}^{I}\big(\hat{P}_{\text{move}}(t)\cdot o_i(t)\cdot\tau_e(t) + \hat{P}_h(t)\cdot\tau_h + \hat{P}_{(b,i)}(t)\cdot\tau_{\text{tx}}(t)\big) \tag{3}$$

$$\Delta\text{AoI}_i = t - U_j(t-1) \tag{4}$$

In other words, Equation (3) represents the energy consumption E_{cmp} of the ith UAV at a specific time t. It includes the total energy consumption of the ith UAV, which consists of the energy consumed for travelling, hovering, and data transmission.

2) **Decision cycle**: The decision cycle begins when the UAV-BS identifies the location of the requested target point, $s_k(t)$, and, considering the current position and state of each UAV-UE, selects the UAV-UE that is closest to the requested target point. The selected UAV-UE must adhere to the energy constraint equation $\epsilon_i(t) = \frac{e_i^{\text{cmp}}(t)}{e_i^{\max}(t)}$, where the current energy $\Delta e_i^{\text{cmp}}(t)$ and the maximum energy capacity $e_i^{\max}(t)$ must meet the condition. If the selected UAV-UE does not satisfy the condition, the UAV-BS must reselect a new UAV-UE that is the closest within its area R_b.

$$\epsilon_i(t) = \begin{cases} 1, & \text{if } \Delta {e_i}^{\text{cmp}}(t) \leq {e_i}^{\max} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

3) **Empty cycle**: An empty cycle represents the state in which the UAV-UE is en route to the target point $s_k(t)$ but has not yet arrived. During this phase, the UAV-BS continuously monitors the UAV-UE and considers the estimated flight time $\tilde{\tau}_e$. If necessary, the path of the UAV-UE can be adjusted to minimize energy consumption and ensure the AoI.

4) **Hovering cycle**: The hovering cycle occurs when the UAV-UE reaches the designated target point $s_k(t)$ and stops for data collection. At this time, the remaining energy of the ith UAV-UE must satisfy the energy constraint condition $\epsilon_i(t)$

5) **Upload cycle**: After the UAV-UE completes the hovering cycle, it starts the upload cycle. During this cycle, the UAV-UE can upload the collected data to the UAV-BS. It must transmit the data to a UAV-BS that covers the area R_b, which is within the transmission range of the UAV-UE. After the transmission is complete, the UAV-UE can record the time step U_j indicating the completion of all cycles. Subsequently, the UAV-UE flies to an area R_b within its transmission range for re-upload.

$$\zeta_i(t) = \begin{cases} \tau(t+1) = \tau_d, & \text{if } q_i \in R_b \text{ and } U_j \\ \tau(t+1) = \tau_{tx}, & \text{otherwise} \end{cases}, \tag{6}$$

where τ_d is the duration of the decision cycle. The UAV-UE selects a UAV-BS that covers an area R_b within the transmission distance.

$$\mu_i(t) = \begin{cases} 1, & \text{if } \sum_{j=1}^{U_j}\Delta\tau_i(t) \leq \hat{\Delta}^{\text{th}}_{\text{AoI}} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

To ensure the AoI for the ith UAV-UE, the scheduling total duration $\tau(T)$ must not exceed the threshold $\hat{\Delta}^{\text{th}}_{\text{AoI}}$ as stipulated in Equation (7). If the total scheduling duration exceeds $\hat{\Delta}^{\text{th}}_{\text{AoI}}$, the data will be discarded. The variable j represents the time it takes for the UAV to perform the selected task, and U_j denotes the total time until the task is completed, which is used in the calculation of AoI.

To solve the problem, we apply the DQN [16], which is a combination of deep neural networks and reinforcement learning algorithms. A DQN can be defined as an MDP represented by a tuple $< S, A, R, S_{t+1} >$. An agent decides on an action a in a given state s. The agent receives a reward R and builds a policy π that takes into account a discount factor γ for the cumulative future reward. The proposed DQN approach consists of:

1. A deep neural network to reduce the dimensionality of the state space used to extract contextual features.
2. An experience replay memory to store the state transitions observed by the UAV-BS agent and the UAV-UE agent.
3. A reinforcement learning framework to find the optimal trajectory policy by solving constraints (9–11) to have a unique target area for each UAV-UE [17]

State: The state can be represented as $S_i(t) = [q_i(t), e_i^{\text{cmp}}(t), c_i(t)]$, which represents three key elements at time t. The position of the UAV-UE, $q_i(t) = (x_i(t), y_i(t), h_i(t))$, accurately tracks the spatial location of the UAV and is used to plan the next movement. $e_i^{\text{cmp}}(t)$ represents the current energy level of the UAV-UE, which can be expressed as the remaining operational energy $e_i^{\text{cmp}}(t) \in \mathbb{R}$. This directly impacts the sustainable operation and mission execution capability of the UAV [18]. Lastly, $c_i(t)$ indicates the current cycle in which the UAV-UE is located. The possible states include {"Decision", "Empty", "Hovering", "Transmission"}, and this information is used to determine the next action of the UAV.

Action: Action is defined by Equation (8), which describes the mobility of the UAV-UE in a given state. If it is hovering, it does not move.

$$A_i(t) = \begin{cases} q_i(t+1) = (x_i(t) + \Delta x, y_i(t) + \Delta y, h_i(t) + \Delta h), & \text{Moving} \\ q_i(t+1) = (x_i(t), y_i(t), h_i(t)), & \text{Hovering} \end{cases} \tag{8}$$

Reward: When the learning agent, namely the UAV-UE, executes action $a_i(t)$, it transitions to a new state $s_i(t+1)$ and receives an immediate reward $r_i(t)$ associated with the state transition $s_i(t), a_i(t), s_i(t+1)$. The reward can be defined as follows in Equation (9), where ϵ_i represents the energy constraint, and $r_{\text{cmp}}(t)$ signifies the reward for saving energy. The energy reward $r^i_{\text{energy}} = \Delta e_i(t)$ is defined by $\Delta e_i(t) = e_i(t) - \Delta e_i(t-1)$, which represents the energy consumed due to action a^i_t. $\mu_i(t)$ indicates the AoI constraint, and $r_{\text{AoI}}(t) = \Delta U_i(t)$ is expressed as $\Delta U_i(t) = U_i(t) - \Delta U_i(t-1)$. This provides a higher reward

Table 1. Simulation hyperparameter values

Hyperparameter	Value
The number of UAV-BS	2
The number of UAV-UE	1–10
Episode	1000
Learning rate (α)	0.0005
Discount factor (γ)	0.99
Mini-batch size	32
Size of memory ($\mathcal{M}$)	10,000

for the UAV-UE's continuous upload of fresh data. Lastly, $o_i(t)$ indicates whether there is a collision with obstacles.

$$R_i(t) = \epsilon_i(t) \times r_{\mathrm{cmp}}(t) + \mu_i(t) \times r_{\mathrm{AoI}}(t) + o_i(t) \tag{9}$$

The learning agent, UAV-UE, aims to maximize future rewards over T time slots as defined in Equation (10). $\gamma = [0, 1]$ reflects the balance between the importance of immediate and future rewards, allowing convergence to the optimal policy π^{opt}, which is a strategy that enables the UAV-UE to choose the optimal behaviour given a set of conditions to minimise energy consumption, maintain the freshness of information (AoI), and avoid obstacles.

$$\hat{R}(s, a, t) = \sum_{t_0=0}^{T} \gamma^{t_0} \times r_i(t - t_0) \tag{10}$$

Therefore, we can update the Q-function to derive the optimal policy π^{opt} as follows (11).

$$Q_{t'}(s, a) = Q_t(s, a) + \alpha \left[R + \gamma \max_{a'} q(s', a') - Q_t(s, a) \right] \tag{11}$$

Here, α is the learning rate that regulates the speed of the Q-function update. Additionally, $t' = t + 1$, and a' represents all actions considered during the maximization process.

Simulation results: We propose a replay memory-based approach to find the appropriate AoI $\Delta_{\mathrm{AoI}}^{\mathrm{th}}$ within the proposed method, ensuring AoI through the use of replay memory. Initially, replay memory represents the (s, a, r, s_{t+1}) obtained by the agent interacting with the environment during the learning process. We addressed the memory limitations of UAVs by placing the replay memory in a central node, which manages the history of all UAVs and uses it to plan the optimal route. It exists in the following types:

- **Replay history**: Stores all past experiences and randomly selects them for learning, contributing to the learning process.
- **Online history**: Stores real-time or the most recent experiences, contributing to immediate learning.
- **Prioritized history**: Selects experiences for learning based on their importance, contributing to the learning process by choosing specific experiences.

We conducted experiments in two distinct scenarios to test UAVs in various environments. The first was a rural environment with a large area (1600 * 1600 m^2) and four obstacles, with UAVs initially positioned at (800, 800). The second was an urban environment with a smaller area (800 * 800 m^2) and eight obstacles, with UAVs initially positioned at (400, 400) (see Table 1).

Figure 3 presents an analysis of the reward acquisition and the learning performance of the memory approach in reinforcement learning. The findings demonstrate that the online history memory approach exhibits consistent and stable learning performance, ultimately achieving the most effective reduction in energy consumption and AoI.

We conducted experiments in a rural scenario characterized by few obstacles and a relatively large area, and an urban scenario with many obstacles and a relatively small area, in order to test UAVs in various environments. Figure 4a,b represents the AoI results according to each memory access approach, facilitating the search for the appropriate $\Delta_{\mathrm{AoI}}^{\mathrm{th}}$. In the rural scenario of Figure 4a, the lowest average AoI was observed to be 94.63 s when applying the priority history memory access approach, which was 30.33 s shorter than the approach without memory usage. In the urban scenario of Figure 4b, the application of the online history memory access approach resulted in the lowest average AoI of 82.23 s, which was a reduction of 43.11 s compared to the non-memory approach.

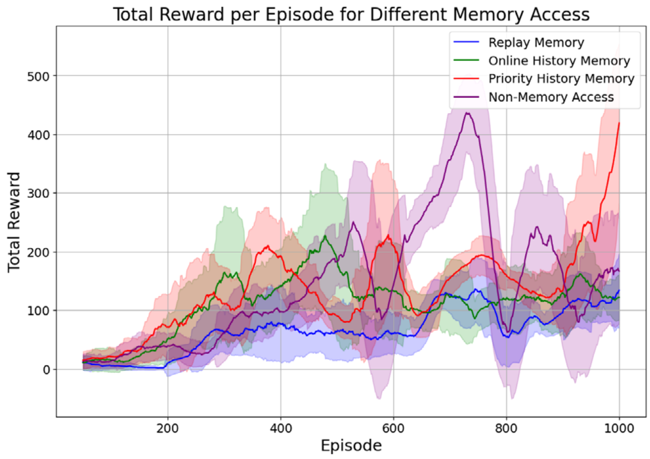

Fig. 3 *Reward per episode for different memory*

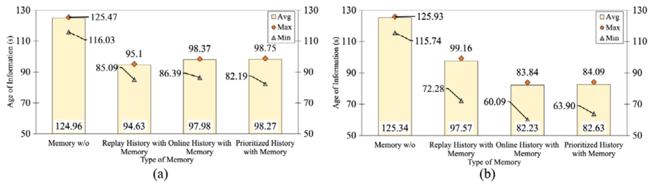

Fig. 4 *Age of information in (a) rural scenario and (b) urban scenario*

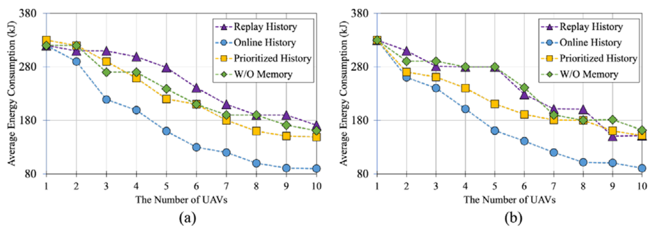

Fig. 5 *Energy consumption in (a) rural scenario and (b) urban scenario*

After setting the average AoI to $\Delta_{\mathrm{AoI}}^{\mathrm{th}}$ in each scenario, we proceeded with energy consumption experiments. Figure 5a indicates that, in the rural scenario with five UAVs deployed, the online history memory access approach shows the lowest energy consumption, which is up to 33.25% lower compared to the non-memory approach. Similarly, Figure 5b shows that in the urban scenario, also with five UAVs, the online history approach results in the lowest energy consumption, showing up to a 74.20% reduction compared to the non-memory approach. These results suggest that the online history approach can adapt in real time to relatively dynamic environments. On the other hand, both the replay history memory access method and the non-memory approach show comparatively higher energy consumption.

To examine the energy consumption of UAVs based on different memory access approaches, we visualized the trajectories of two UAVs as shown in Figures 6 and 7. Figure 6 illustrates that in the rural scenario with the online history memory access approach applied, the UAVs fly in divided areas, suggesting that they reach the target points and collect data. In contrast, without the memory access approach, the UAVs overlap in their flight paths and fail to reach the target points. Therefore, the results indicate that the absence of a memory access approach leads to increased energy consumption due to overlapping flight paths and a failure to occupy distinct flying zones.

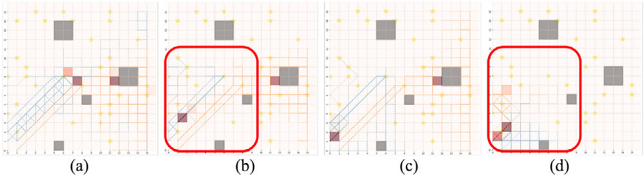

Fig. 6 *Comparison trajectory in rural scenario using different memory access approaches. (a) Replay history. (b) Online history. (c) Prioritised history. (d) Non-memory*

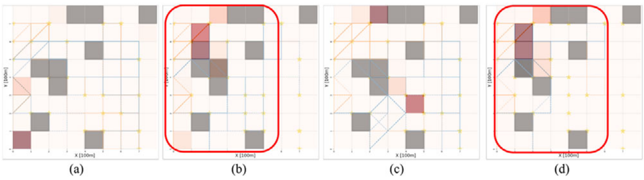

Fig. 7 *Comparison trajectory in urban scenario using different memory access approaches. (a) Replay history. (b) Online history. (c) Prioritised history. (d) Non-memory*

Conclusion: Here, we proposed a hierarchical deployment structure and an energy consumption minimization scheduling method centered around the AoI for efficient UAV operations. The results of applying a memory access approach-based DQN demonstrated significant reductions in energy consumption, up to 33.25% in rural scenarios and up to 74.20% in urban scenarios, while maintaining data freshness. However, the applicability of our approach in real-world scenarios and its potential challenges, such as computational overhead and integration with existing systems, require further investigation.

Author contributions: **Jeena Kim**: Investigation; visualization; writing—original draft; writing—review and editing. **Seunghyun Park**: Project administration; software; writing—review and editing. **Hyunhee Park**: Conceptualization; methodology; supervision; writing—review and editing.

Acknowledgements: This work was supported in part by the National Research Foundation of Korea (NRF) under Grant 2022R1A2C2005705, and this research was financially supported by Hansung University for Prof. Seunghyun Park.

Conflict of interest statement: The authors declare no conflicts of interest.

Data availability statement: To access the data supporting the findings of this study, please contact the corresponding author directly. Specific requests will be considered on a case-by-case basis.

Received: *7 March 2024* Accepted: *7 October 2024*
doi: 10.1049/ell2.70063

References

1 Sabuj, S.R., Ahmed, A., Cho, Y., Lee, K., Jo, H.: Cognitive UAV-aided URLLC and mMTC services: Analyzing energy efficiency and latency. *IEEE Access* **9**, 5011–5027 (2020)

2 Li, B., Fei, Z., Zhang, Y.: UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet Things J.* **6**(2), 2241–2263 (2018)

3 Zeng, Y., Wu, Q., Zhang, R.: Accessing from the sky: A tutorial on UAV communications for 5G and beyond. *Proc. IEEE* **107**(12), 2327–2375 (2019)

4 Gupta, L., Jain, R., Vaszkun, G.: Survey of important issues in UAV communication networks. *IEEE Commun. Surv. Tutorials* **18**(2), 1123–1152 (2015)

5 binti Burhanuddin, L.A., Liu, X., Deng, Y., Challita, U., Zahemszky, A.: QoE optimization for live video streaming in UAV-to-UAV communications via deep reinforcement learning. *IEEE Trans. Veh. Technol.* **71**(5), 5358–5370 (2022)

6 Erdelj, M., Natalizio, E.: UAV-assisted disaster management: Applications and open issues. In: 2016 International Conference on Computing, Networking and Communications (ICNC), pp. 1–5. IEEE, Kauai, HI (2016)

7 Zhao, N., Lu, W., Sheng, M., et al.: UAV-assisted emergency networks in disasters. *IEEE Wireless Commun.* **26**(1), 45–51 (2019)

8 Panda, K.G., Das, S., Sen, D., Arif, W.: Design and deployment of UAV-aided post-disaster emergency network. *IEEE Access* **7**, 102985–102999 (2019)

9 Gao, Y., Xiao, L., Wu, F., Yang, D., Sun, Z.: Cellular-connected UAV trajectory design with connectivity constraint: A deep reinforcement learning approach. *IEEE Trans. Green Commun. Networking* **5**(3), 1369–1380 (2021)

10 Abubakar, A.I., Zegura, E.W., Ammar, M.H.: A survey on energy optimization techniques in UAV-based cellular networks: From conventional to machine learning approaches. *Drones* **7**(3), 214 (2023)

11 Shi, L., Chen, H., Gao, Y.: Age of information optimization with heterogeneous UAVs based on deep reinforcement learning. In: 2022 14th International Conference on Advanced Computational Intelligence (ICACI), pp. 364–369. IEEE, Wuhan, China (2022)

12 Kahraman, İ., Köse, A., Koca, M., Anarím, E.: Age of information in internet of things: A survey. *IEEE Internet Things J.* **11**(6), 9896–9914 (2024)

13 Hu, H., Xiong, K., Qu, G., Ni, Q., Fan, P., Letaief, K.B.: AoI-minimal trajectory planning and data collection in UAV-assisted wireless powered IoT networks. *IEEE Internet Things J.* **8**(2), 1211–1223 (2020)

14 Ahani, G., Yuan, D., Zhao, Y.: Age-optimal UAV scheduling for data collection with battery recharging. *IEEE Commun. Lett.* **25**(4), 1254–1258 (2020)

15 Abedin, S.F., Munir, M.S., Tran, N.H., Han, Z., Hong, C.S.: Data freshness and energy-efficient UAV navigation optimization: A deep reinforcement learning approach. *IEEE Trans. Intell. Transp. Syst.* **22**(9), 5994–6006 (2020)

16 Hu, J., Zhang, H., Song, L., Schober, R., Poor, H.V.: Cooperative internet of UAVs: Distributed trajectory design by multi-agent deep reinforcement learning. *IEEE Trans. Commun.* **68**(11), 6807–6821 (2020)

17 Wang, Y., Gao, Z., Zhang, J., et al.: Trajectory design for UAV-based Internet of Things data collection: A deep reinforcement learning approach. *IEEE Internet Things J.* **9**(5), 3899–3912 (2021)

18 Bonecker, M.P., Zhu, Y.: Deep Q-network based decision making for autonomous driving. In: 2019 3rd International Conference on Robotics and Automation Sciences (ICRAS), pp. 154–160. IEEE, Wuhan, China (2019)