UAV-Enabled SWIPT for IoT Networks

Journal:	*IEEE Wireless Communications Magazine*
Manuscript ID	WCM-19-00556
Topic or Series:	Open Call Article
Date Submitted by the Author:	27-Dec-2019

Complete List of Authors:	Feng, Wanmei; South China University of Technology, School of Electronic and Information Engineering Tang, Jie; South China University of Technology, Yu, Yu; South China University of Technology, School of Electronic and Information Engineering Song, Jingru; South China University of Technology, School of Electronic and Information Engineering Zhao, Nan; Dalian University of Technology, School of Information and Communication Engineering Chen, Gaojie; University of Leicester, Department of Engineering Wong, Kai-Kit; University College London, Department of Electronic and Electrical Engineering Chambers, Jonathon; University of Leicester, Department of Engineering

Key Words:	Deep reinforcement learning, Internet of Things (IoT), simultaneous wireless information and power transfer (SWIPT), trajectory optimization, unmanned aerial vehicle (UAV)

UAV-Enabled SWIPT for IoT Networks

Wanmei Feng, Jie Tang, *Senior Member, IEEE*, Yu Yu, Jingru Song,

Nan Zhao, *Senior Member, IEEE*, Gaojie Chen, *Senior Member, IEEE*,

Kai-Kit Wong, *Fellow, IEEE* and Jonathon Chambers, *Fellow, IEEE*

Abstract

Energy-limited devices and connectivity in complicated environments are two challenges for Internet of Things (IoT)-enabled mobile networks. Unmanned aerial vehicle (UAV)-enabled simultaneous wireless information and power transfer (SWIPT) is emerging as a promising technique to tackle the above problems, and thus provide cost-effective connectivity for IoT devices. In this paper, a unified framework of UAV-enabled SWIPT for IoT networks is established. First, three-dimensional (3D) trajectory and charging strategy of UAVs are investigated for maximizing the sum received power of energy receivers (ERs), where an antenna array provided by a UAV is used to generate multi-beams to charge the ERs simultaneously. Then, a trajectory planning scheme for multiple UAVs based on deep-Q-learning (DQL) is studied to maximize the minimum throughput of users by jointly optimizing the trajectory and resource scheduling of UAVs. In addition, a deep deterministic policy gradient (DDPG)-based algorithm is proposed to achieve the correct prediction of the service requirement of devices (i.e., data transmission and battery charging) and, then, a dynamic path planning scheme is designed to maximize the energy efficiency (EE) of the system. Simulation results demonstrate the effectiveness of the above schemes. Finally, potential research directions and challenges are also discussed.

Index Terms

Deep reinforcement learning, Internet of Things (IoT), simultaneous wireless information and power transfer (SWIPT), trajectory optimization, unmanned aerial vehicle (UAV)

I. Introduction

INTERNET of Things (IoT)-enabled mobile networks are generally composed of heterogeneous smart devices which collect, transmit, exchange and process massive information, and face particular challenges including limited battery energy and connectivity in complicated environments. To overcome these two challenges, on one hand, simultaneous wireless information and power transfer (SWIPT) is considered as a prominent technology for prolonging the lifetime of power constrained IoT devices; this technology enables energy and information to be simultaneously transferred. On the other hand, an unmanned aerial vehicle (UAV) can act as a flying base station (BS), and can be utilized for rapid deployment in rural and geographically constrained areas due to advantages of autonomy, flexibility and mobility [1].

Recently, the notion of UAV-enabled SWIPT has been proposed, and it has attracted much attention in both academia and industry [2]–[7]. In [2], a successive hover-and-fly trajectory strategy was proposed to maximize the sum received power of energy receivers (ERs) by jointly optimizing transmit power and trajectory in UAV-enabled wireless power transfer (WPT) networks. Motivated by [2], the authors in [3] extended the flying trajectory planning into UAV-enabled wireless powered communication networks (WPCNs), where a transmission scheme for throughput maximization was studied which considered the constraints of flying time and speed. In [4], a point-to-point throughput maximization problem was studied in a UAV-assisted cooperative communication system with SWIPT, in which the decision profile and power profile in addition to flight trajectory of the UAV are optimized, and solved by the alternating optimization algorithm. An on-board deep Q-network (DQN) method was proposed in [5] to provide the online decisions of wireless services (i.e., power transfer and data collection) for UAVs, in order to minimize data packet loss as well as preventing battery drainage and data queue overflow of ground devices. In [6], partial and binary computation offloading modes were proposed for investigating the computation rate maximization problem in a UAV-enabled mobile-edge computing (MEC) wireless-powered system. In [7], the authors studied a secure transmission scheme based on millimeter wave (mmWave) SWIPT for UAV-based relay networks, in which a closed-form expression for the lower bound of the average secrecy rate was derived.

Although excellent research has been conducted on UAV-enabled SWIPT, very few works

TABLE I: Summary of Important Acronyms.

Acronym	Definition	Acronym	Definition
BS	base station	CSI	channel state information
DQL	deep-Q-learning	DDPG	deep deterministic policy gradient
DQN	deep-Q-network	DNN	deep neural network
ER	energy receiver	ET	energy transmitter
EE	energy efficiency	IoT	Internet of Things
MEC	mobile-edge computing	mm Wave	millimeter wave
MOEA/D	multiobjective evolutionary algorithm based on decomposition	NOMA	non-orthogonal multiple access
QoS	quality of service	QL	Q-learning
SWIPT	simultaneous wireless information and power transfer	SLL	side-lobe level
TDMA	time division multiple access	UAV	unmanned aerial vehicle
UPA	uniform planar array	WPT	wireless power transfer
WET	wireless energy transfer	WIT	wireless information transfer
WPCN	wireless powered communication network	3D	three-dimension

have focused on UAV-enabled SWIPT for IoT networks [8], [9]. In [8], a novel time division multiple access (TDMA) scheme based on a workflow mode was proposed to minimize the energy consumption in UAV-enabled mobile computing systems, where the tasks of the UAV included WPT, communication and computation. In [9], a battery charging policy and interference mitigation scheme were designed for use in UAV-aided wireless power IoT networks, in which a machine learning framework based on echo state networks was exploited to predict the energy consumption of nodes. However, a systematic study of UAV-enabled SWIPT for IoT networks is missing in the aforementioned research works.

In this paper, a unified framework for UAV-enabled SWIPT for IoT networks is established. First, the three-dimension (3D) trajectory design and charging strategy are studied to maximize the sum received power of the ERs in UAV-enabled WPT networks, where a UAV mounted antenna array generates multi-beams to serve several ground users simultaneously. Then, a multiple trajectory planning scheme based on deep-Q-learning (DQL) for UAVs is designed to maximize the minimum throughput of users for a UAV-enabled WPCN by jointly taking into account the trajectory design and resource scheduling of UAVs. Furthermore, a deep deterministic policy gradient (DDPG)-based method is proposed to achieve the correct prediction of the service requirements of the devices (i.e., data transmission and battery charging) and accordingly, a dynamic path planning scheme is proposed for the UAV to maximize the energy efficiency (EE) of the data transmission and energy transfer.

The remainder of this paper is organized as follows. In the next section, the framework of

UAV-enabled SWIPT for IoT networks is first presented. The 3D trajectory design and charging strategy are studied. Then, the trajectory planning scheme for multiple UAVs based on DQL is designed. In addition, a DDPG-based method is proposed to predict the service requirements of devices and, then, a dynamic path planning scheme is established. Finally, potential research directions and challenges are discussed, and then the conclusions are presented in the final section. The definition of the acronyms in this paper are summarized in TABLE I.

II. Framework of UAV-enabled SWIPT for IoT Networks

The framework of UAV-enabled SWIPT for IoT networks can be described as follows:

- Scenario 1: In the scenario with several serving areas, a UAV exploits a multi-beam antenna array to transfer energy to IoT devices simultaneously in one serving area and, then, the UAV flies to another hovering location according to the flight planning. In this case, the trajectory and beam pattern of the UAV are optimized to maximize the sum received power of the receivers.
- Scenario 2: In the scenario with multiple UAVs, each UAV first charges a set of IoT devices via the WPT downlink and, then, users send their independent information to the UAV via the WIT uplink. In this case, trajectory planning of the UAVs and an interference management scheme are jointly designed to maximize the minimum throughput of the IoT devices.
- Scenario 3: Energy supply and data transmission of IoT devices in both scenario 1 and scenario 2 are predicted correctly in a timely manner. Then, based on the prediction of service requirements, dynamic path planning of the UAV is designed, in order to provide effective service for the IoT devices.

In the follow sections, these key scenarios for UAV-enabled SWIPT for IoT networks are analyzed in detail.

III. Joint 3D trajectory and beam pattern design

When IoT devices are distributed in disaster areas or remote mountain areas, it is not efficient for BSs to deliver energy to ERs due to the severe path loss of receivers over long distance. Fortunately, a UAV can be deployed as an energy transmitter (ET) to provide wireless energy. In [2], the UAV trajectory was studied to maximize the energy transfer while guaranteeing the

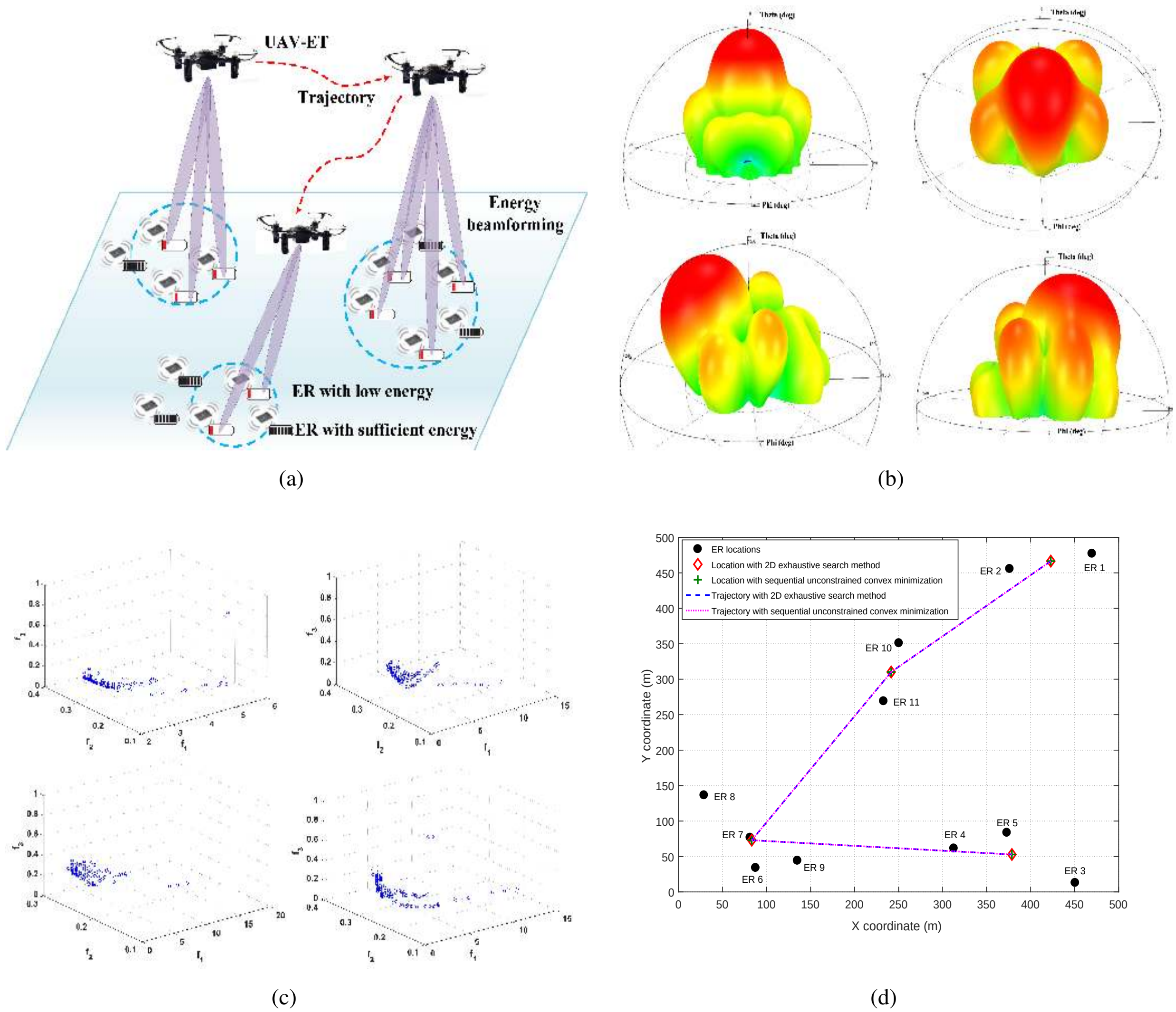


Fig. 1: Joint 3D trajectory and beam pattern design for a UAV-enabled WPT system with IoT devices: (a) Illustration of a UAV-enabled WPT system with multi-beams generated by an antenna array; (b) 3D realized gain patterns of the antenna array; (c) An example convergence performance of the MOEA/D for 3-objective optimization problem; (d) The 3D trajectory design for a UAV-enabled WPT system with K=11 ERs.

maximum flying speed; however, the flight altitude is not considered which is related to the ground coverage area. Therefore, in this section, the altitude of UAV is taken into account. To further improve the sum received power of ERs, multi-beams formed by a phased antenna array mounted on the UAV should be designed and optimized, which is discussed in this section.

A. Problem Formulation

The UAV-enabled WPT architecture is illustrated in Fig. 1(a), where a rotary-wing UAV is dispatched as a wireless charger that flies above Γ serving areas according to the trajectory planning, generates multi-beams and, then, delivers energy to K IoT devices. A uniform planar array (UPA) with $M \times N$ array elements is available at the UAV, while a single antenna is assumed at each IoT device. The UPA at the UAV is split into multiple sub-arrays, wherein the phases of each sub-array are controlled by the phase shifters and, independent steering beams are thereby generated. In this case, it is assumed that the mutual coupling between two sub-arrays is negligible, however, the mutual coupling of adjacent array elements within a sub-array is considered.

Define the horizontal location of the UAV as $z_u = (x_u, y_u)$, and its flight altitude as h. The locations of the ERs $k \in \{1, 2, \cdots K\}$ on the ground are given by $z_k = (x_k, y_k)$. In addition, the desired geographical area covered by the UAV is a circle with radius $h \tan \Theta$, where 2Θ indicates effective illumination angle. Thus, the multi-beams can be generated by an antenna array flexibly serving ERs within the coverage area of the UAV; that is, the beam pattern $\mathbf{E}(\theta, \phi)$ with the elevation θ and azimuth ϕ angles should be properly designed. The charging time of the γ-th ($\gamma \in \{1, 2, \cdots \Gamma\}$) serving area is τ_γ, and $\sum_{\gamma=1}^{\Gamma} \tau_\gamma$ is equal to the whole period T.

To prolong the battery-life of the IoT devices and improve the EE of the UAVs, an energy harvested optimization problem is studied by sequentially optimizing the 3D position of the UAV (x_u, y_u, h), beam pattern $\mathbf{E}(\theta, \phi)$, charging time τ_γ, with the constraints of flight altitude h, charging time τ_γ, and the coverage radius of the UAV. However, this optimization problem is non-convex due to the coupling variables, and it is extremely difficult to solve directly.

To tackle this problem, we first fix the UAV altitude h, beam pattern $\mathbf{E}(\theta, \phi)$ and charging time τ_γ, but the objective function of the problem, with respect to the 2D position of the UAV, is still non-convex. To solve this problem, a sequential unconstrained convex minimization based algorithm [10] is adopted to obtain the optimal solutions. Then, the optimal altitude can be calculated through monotonic optimization theory with the fixed beam pattern $\mathbf{E}(\theta, \phi)$ and charging time τ_γ. Subsequently, when the charging time τ_γ is fixed, the problem thereby becomes a function of beam pattern $\mathbf{E}(\theta, \phi)$. To generate a steerable beam pattern, the phases of the antenna array are controlled to adjust the antenna gain, side-lobe level (SLL) and beamwidth, which can be constructed as a multiobjective optimization problem with decision variables $\boldsymbol{\beta}$ (i.e., the phases of the array elements). To handle this problem, the multiobjective evolutionary algorithm based on decomposition (MOEA/D) based algorithm is proposed, which uses Tchebycheff's approach [11] to decompose the optimization problem into a number of scalar sub-problems and, then, optimizes all sub-problems to approximate the Pareto front via an iteration process. Finally, the problem becomes a function of charging time τ_γ with the solved variables, however, a fairness issue remains, i.e., as the total received power of all ERs increases in one serving area so too must the allocated charging time. To address this problem, the original problem is converted into a max-min problem that can be solved by standard optimization techniques. Furthermore, to reduce the flight distance, we use the branch and bound method to design the 3D flight trajectory of the UAV.

Through the joint 3D trajectory and beam pattern design scheme for the UAV, the sum received power of ERs can be maximized. This is achieved by jointly considering the beam angles and flight altitude, which relates to the UAV coverage radius, and thus affects the performance of energy harvesting.

B. Discussion

In this scenario, a rotary-wing UAV with an antenna array acts as the energy charger to serve multiple ERs simultaneously. Nevertheless, the case when the UAV transmits data to ground users is available in UAV-enabled communications networks. In this situation, the UAV can adopt the non-orthogonal multiple access (NOMA) technique to serve multiple users, and thus the optimization problem can be modelled as a sum-rate maximization problem. To tackle this problem, a joint 3D placement, beam pattern and power allocation scheme for the UAV can be properly designed.

C. Simulation Results

The performance of the joint 3D trajectory and beam design scheme is analyzed in Fig. 1. First, the multi-beam gain performance using the MOEA/D solution is investigated. An 8×8 antenna array is separated into four sub-arrays, where each sub-array is of size 4×4. We set the amplitude and spacing of the antenna array to 1A and 5.5mm, and the maximum effective illustration angle 2Θ to $80°$. In Fig. 1(b), we can see that the four beams have different directions, i.e., ($-10°$, $0°$), ($30°$, $0°$), ($30°$, $270°$) and ($20°$, $90°$), and the gains of the main beams are much greater than the initial sidelobes. Fig. 1(c) illustrates that the best solutions found by the MOEA/D algorithm through 300 iterations lie on the true Pareto-optimal front. Then, the performance of the 3D trajectory design in the UAV-enabled WPT system is investigated. We set the minimum altitude $h_{min} = 21$m, maximum altitude $h_{max} = 120$m, $T = 20$s, and the carrier frequency to 25 GHz. In Fig. 1(d), the 3D placements of the UAV lie at the center of the IoT devices that are distributed over the same serving area. This is due to the fact that the location of the UAV in the middle of the receivers can maximize the harvested energy of the ERs. Thus, the sum received power of the IoT devices can be maximized with the flight altitude and coverage radius of the UAV guaranteed through the proper management of the joint 3D trajectory and beam pattern design scheme.

IV. Minimum throughput maximization for multi-UAV enabled WPCN

The energy-limited IoT devices may have communication requirement for the above considered scenario, such as uploading the collected information. Hence, in this section, we consider a multi-UAV enabled WPCN, in which the UAVs charge the IoT devices in the downlink and, then, IoT devices send their collected information in the uplink by utilizing the harvested energy. Since the co-channel interference between UAVs is inevitable, a joint trajectory and resource allocation scheme is properly designed, which is discussed in this section.

A. Problem Formulation

A multi-UAV assisted WPCN is shown in Fig. 2. The IoT devices are divided into L clusters by the K-means algorithm, and one UAV belongs to a single cluster. The UAVs and IoT devices are each equipped with a single antenna. For any cluster $l \in \mathcal{L} = \{1, 2, \cdots, L\}$, we consider a TDMA scheme. In particular, the flight period of the UAVs T can be discretized into $N + 1$ time slots, where the 0-th time slot is assigned to the downlink wireless power transfer (WET) and the n-th time slot, $n \in \mathcal{N} = \{1, 2, \cdots, N\}$, is allocated to the uplink wireless information transmission (WIT). The location of the l-th UAV in the n-th time slot is $q_l[n] = (x_l[n], y_l[n], h_l[n])$, and its

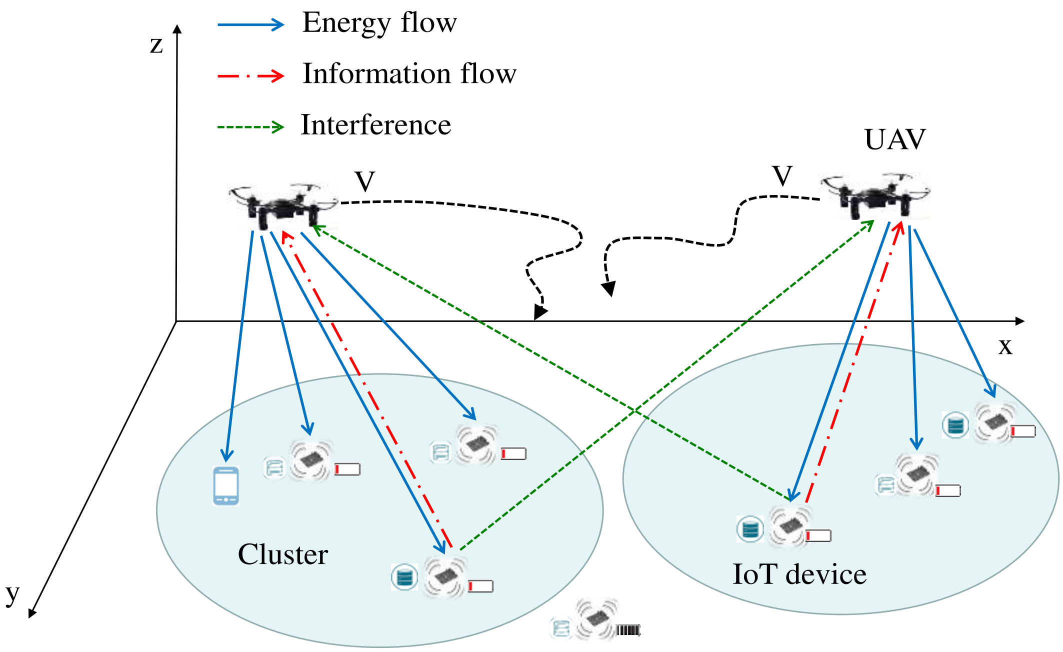


Fig. 2: Illustration of multi-UAV enabled WPCN with trajectory planning and resource scheduling.

maximum speed is V_{max}. Thus, the location of the UAVs should satisfy $q_l[n]-q_l[n-1] \leq V_{max}\cdot\delta_N$, where δ_N denotes the length of each time slot for the uplink. The IoT device $k_l \in \{1, 2, \cdots, K_l\}$ served by the l-th UAV is located at $(x_{k_l}, y_{k_l}, 0)$. To illustrate the scheduling of the WET and WIT, we exploit the binary variables $a_l[0]$ and $a_{l,k_l}[n]$. $a_l[0]$ equaling 1 or 0 means that the energy is transferred or not transferred by the l-th UAV; $a_{l,k_l}[n]$ equaling 1 or 0 means that ground user k_l is served or not served by the l-th UAV. In addition, the time resource allocation should satisfy $\sum_{k_l \in K_l} a_{l,k_l}[n] = 1$. To guarantee the transmission reliability, the uplink power consumption is less than the harvested energy.

In order to improve the quality of service (QoS) in the UAV-enabled WPCN, we maximize the minimum average throughput of devices by jointly optimizing the trajectories of the UAVs $\{q_l[n]\}$, time resource allocations $\{a_l[0], a_{l,k_l}[n]\}$ and uplink power $\{P^U_{k_l}[n]\}$, with the constraint of maximum flight speed and flight range. This problem is a mixed-integer non-convex problem and is difficult to solve due to the problem with coupled variables and a binary variable.

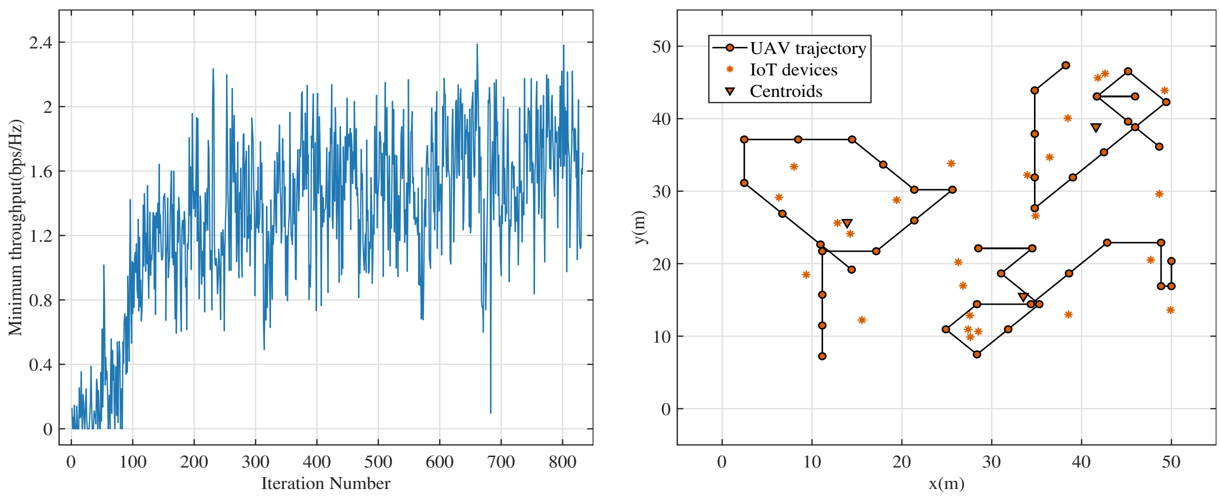
To solve this problem, we propose an effective reinforcement learning-based trajectory planning and resource scheduling scheme. Instead of exploiting Q-learning (QL), here we use deep-Q-learning [12]. Since the UAVs fly flexibly in a 3D environment which changes over time, the DQL method, utilizing the deep neural network (DNN), can better handle the dynamic time variant environment compared with the QL method, and thus provide an ideal strategy for the dynamic flight control of the UAVs. States, actions and rewards are the basic elements in the DQL strategy, which are defined as follows:

- State $s[n] = [q_l[n], a_{l,k_l}[n], R_{k_l}[n]]$: The l-th UAV acts as an agent and observes the states of the system, including the 3D location $q_l[n]$, the number of times $a_{l,k_l}[n]$ that device k_l communicates with the l-th UAV and the data rate of device k_l in the l-th cluster $R_{k_l}[n]$.
- Action $a[n]$: Based on the observed states, the UAVs take actions. Specifically, the action space for the UAVs is denoted by (x,y,z), varying from (-1,-1,-1) to (1,1,1). $x = -1$ means turning left; $x = 1$ indicates flying towards the right; $y = -1$ implies flying backward; $y = 1$ means flying forward; $z = -1$ represents descending; $z = 1$ means rising; $(x, y, z) = (0, 0, 0)$ indicates staying static.
- Reward $r[n]$: The UAVs get rewards or penalties according to their own actions. On one hand, the UAVs receive a penalty of -1 for flying beyond the borders, intersecting flight trajectory of two UAVs and decreasing the minimum average throughput of all devices. On the other hand, the UAVs are rewarded with 2 for increasing the minimum average throughput of all devices.

Based on the above definitions, the DQL-based algorithm dynamically updates the DQN based training samples [13] through interacting with the environment iteratively and, then, designs the flight trajectory of the UAVs and allocates the time resources. Through the path planning scheme for the UAVs based on DQL algorithm, the decision-making of the UAVs can improve with experience, aiming to maximize the minimum throughput of the IoT devices.

B. Discussion

In the above demonstration, a multi-UAV enabled WPCN is considered, in which UAVs and IoT devices are equipped with single antennas. However, the robustness of communication can not be guaranteed when the number of IoT devices increases. To better manage interference and improve the performance gain of the system, a beamforming technique can be considered in



(a)

(b)

Fig. 3: Performance of joint UAV trajectory design and time resource allocation scheme: (a) An example convergence behavior of the DQL strategy in terms of minimum throughput; (b) Trajectory design for a UAV-enabled WPCN system with 3 UAVs.

the UAV-enabled WPCN, however, the optimization problem becomes more complex. Thus, an effective resource allocation scheme should be designed to maximize the minimum throughput of users by jointly optimizing beam selection and trajectory design, which are further analyzed.

C. Simulation Results

For our simulation, the flight altitude of the UAVs is set in the range $[10, 20]$m, the flight period $T = 1$s, and the maximum speed $V_{max} = 6$m/s. There are 25 IoT devices uniformly distributed within a $50m \times 50m$ area. The downlink transmission power and the uplink transmission power are $P^{DL} = 40$ dBm and $P^{UL}_{max} = -20$ dBm, respectively. The noise power of the UAVs σ^2 is set to -110 dBm. The energy conversion efficiency of devices is set to $\eta = 0.1$. First, an example convergence performance of the proposed joint trajectory design and time resource allocation algorithm is shown in Fig. 3(a). It is observed that the optimal solutions of the proposed algorithm remain stable in a range $[1, 2.4]$ after 400 iterations. Then, we can see in Fig. 3(b) that the locations of the UAVs are close to the centroids of the clusters, in order to cover all devices in their cluster. Moreover, the UAVs hover sufficiently close to the IoT devices but keep away from each other, in order to improve the quality of communications and reduce interference. Thus, the optimized trajectory of UAVs tends to make a balance between communication quality and interference, and thus improve the system performance effectively.

V. Dynamic Path Planning of the UAV Based on User Service Requirements

In the above two sections, the trajectory planning and resource allocation strategies for the UAVs are proposed based on the IoT devices with constant service requirements. However, in practice, massive IoT devices are widely deployed to monitor environments, so the data buffer length and residual battery level will change over time. Due to the limited capacity of the data buffers and batteries in IoT devices, the UAVs may not be able to provide the data collection and charging service for devices in time, so data may overflow and energy supply may be exhausted. Thus, it is of great importance for UAVs to provide efficient and timely service for IoT devices, thereby avoiding data overflow and battery expiry. To this end, in this section, a system consisting of sensor devices with dynamic service requirements is first described. Then, the dynamic path planning scheme for UAVs based on the DDPG method is proposed.

A. System Description

In Fig. 4, we consider a SWIPT-based IoT communication system, where a UAV and a certain number of IoT devices are randomly distributed in a 2D area. The UAV flies horizontally to provide data collection and charging service and is aware of its own location. The IoT devices have different data transmission and charging requirements, which vary over time. $\lambda_d(t)$ and $\lambda_e(t)$ represent data accumulation rate and energy consumption rate of the IoT devices, where $t = 1, 2, ..., T$ represents the number of slots. It is noted that the former one obeys a Poisson distribution, and the latter obeys a Gaussian distribution. $Q_d(t)$ represents the data transmission requirements, which is defined as the ratio of the current data buffer length $b_d(t)$ to the maximum storage C_d, and $b_d(t)$ is updated according to $b_d(t+1) = b_d(t) + \lambda_d(t)$. Similarly, $Q_e(t)$ denotes the charging requirements and is defined as the ratio of the consumed energy $b_e(t)$ to the battery capacity C_e, and $b_e(t)$ is updated according to $b_e(t+1) = b_e(t) + \lambda_e(t)$. $N_D(t)$ and $N_B(t)$ denote the number of users with data overflow and battery drainage, respectively. The horizontal position of the UAV is denoted as $[x_u(t), y_u(t)]$. In addition, $d(t) \in [0, v_{\max}]$ indicates the flying distance, where $v_{\max}$ is the maximum flying speed, and $\theta(t) \in [0, 2\pi]$ denotes the flying angle. We assume

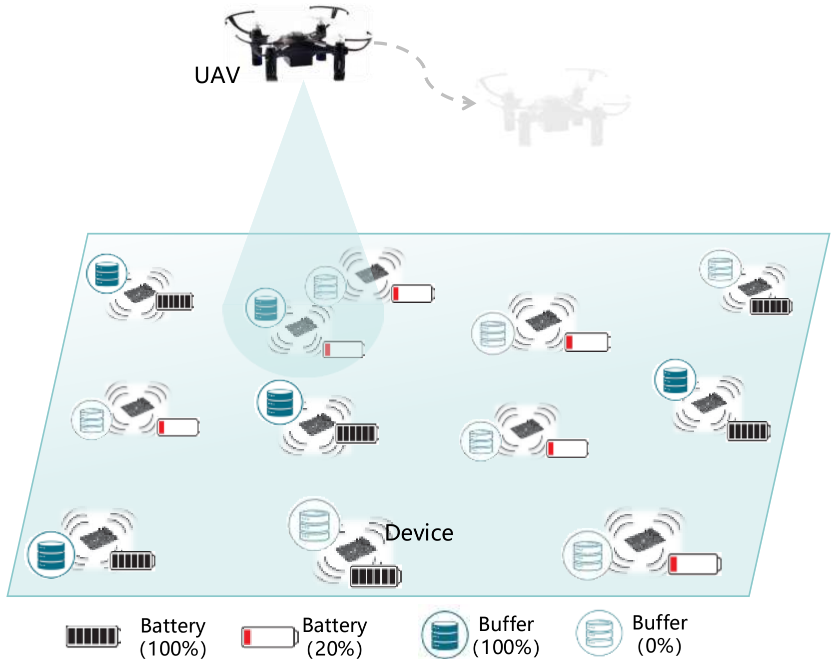


Fig. 4: Illustration of SWIPT-based IoT communication system with the prediction of dynamic service requirements.

that the flying energy consumption $E_c(t)$ increases linearly with the flying distance. The UAV is restricted to a designated area. Once it tries to fly out of range, it is forced to hover at the original location. $N_F(t)$ is given to record the number of times that the UAV tries to go out of bounds.

B. DDPG-based Dynamic Path Planning of the UAV

Based on the above system model, we propose an efficient algorithm to control the flight planning of a UAV based on the real-time state of the system. Instead of the DQL method proposed in Section IV, here we use the DDPG method for UAV control. This is due to the fact that the former one can only handle tasks with a low-dimensional discrete action space. In our model, we assume that the UAV can fly at any continuous speed within $v_{\max}$, and thus UAV control is a continuous control task. DDPG has been proved as an effective reinforcement learning method for the continuous control task. The state, action and reward of the DDPG agent are defined as follows:

- State $\mathbf{s}(t) = [\Delta d_x(t), \Delta d_y(t), x_u(t), y_u(t), N_F, N_D, N_B]$: We assume that the UAV is aware of the location of the IoT device with the highest service requirements and thus is set as the target location. $[\Delta d_x(t), \Delta d_y(t)]$ is the distance between the target device and the UAV. Once the target device is covered, another device will be selected as the new target according to the state of the system at the time.

- Action $\mathbf{a}(t) = [d(t), \theta(t)]$: According to the observed state, the UAV flies to the next waypoint using the calculated flying distance and angle.
- Reward $R(t) = r_1 Q_d(t) + r_2 Q_e(t) - p_1 E_c(t) - p_2 N_F(t) - p_3 N_D(t) - p_4 N_B(t)$: On one hand, we reward the UAV based on service requirements of devices; r_1, r_2 are the reward weight parameters. On the other hand, we punish the UAV for its flight energy consumption, flying out of the designated area, data overflow and battery expiry of devices, where p_1, p_2, p_3, p_4 are the corresponding weight parameters, respectively.

Based on the above setup, the UAV path planing is designed based on the DDPG-based algorithm by iteratively interacting with the environment. Through dynamically updating training samples, the UAV achieves intelligent flight control, and thereby provides timely and efficient services for IoT devices.

C. *Discussion*

Multi-UAV Collaboration: In this scenario, we only consider one single UAV flying in the designated area in providing data collection and charging services to the IoT devices based on their requirements. However, in practice, IoT devices are widely distributed over a large area; therefore multi-UAV collaboration should be investigated to achieve wide coverage, and thus communication and collision avoidance between multiple UAVs are jointly considered.

D. *Simulation Results*

The performance of our proposed algorithm is shown in Fig. 5. We assume that 100 IoT devices are distributed in a 400×400m area, where the data accumulation rate $\lambda_d(t)$ is randomly assigned from $\{4, 6, 10, 18\}$ and energy consumption rate $\lambda_e(t)$ is randomly generated from a Gaussian distribution with expectation selected from $\{0.2, 0.4, 0.6, 0.8\}$ and variance 0.05. The maximum storage C_d is set as 2000 and the battery capacity C_e is set as 100. The radius of the

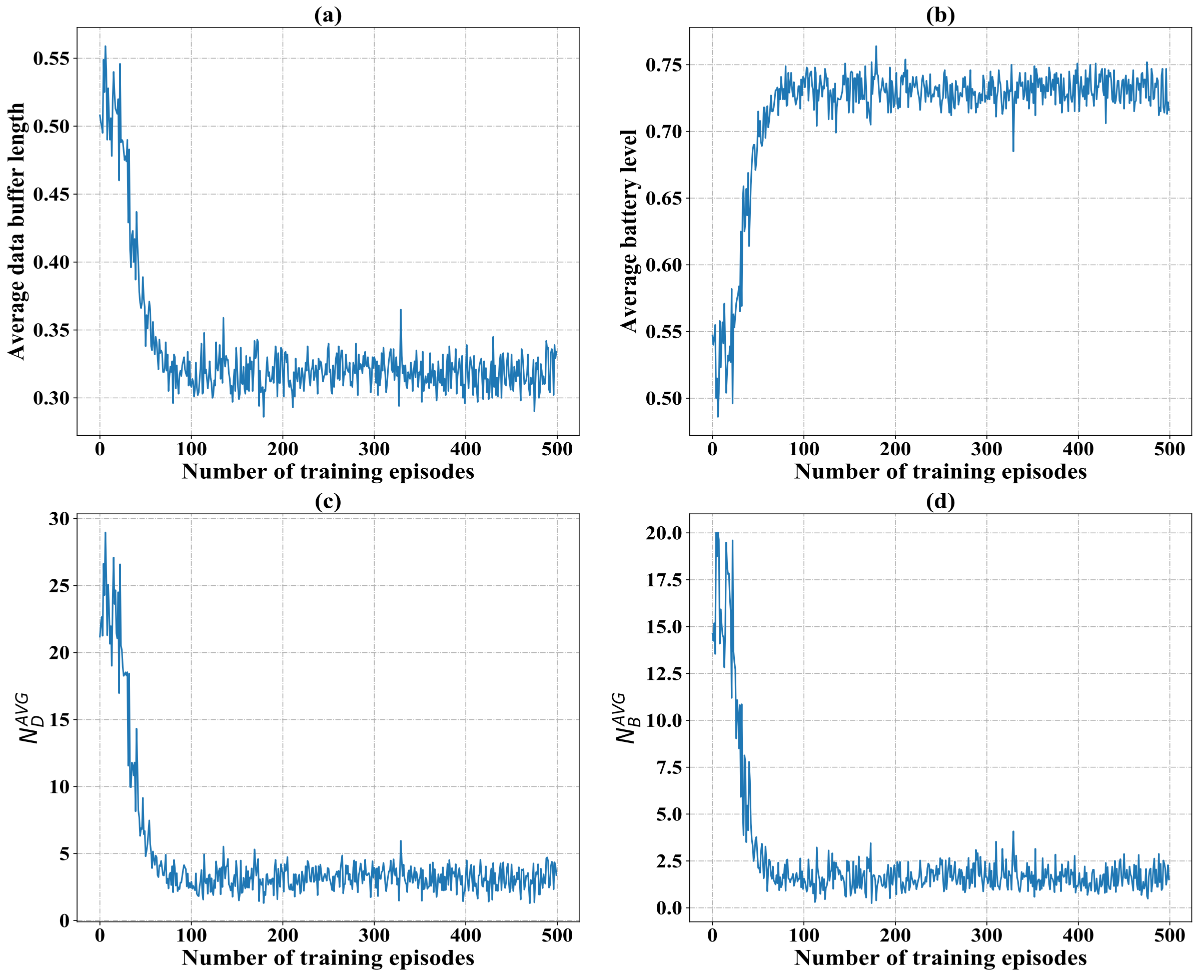


Fig. 5: Network status over time during training: (a) Average data buffer length, (b) Average residual battery level, (c) Average number of devices with data overflow, (d) Average number of devices with expired batteries.

UAV coverage is 20m, and the maximum flying speed $v_{\max} = 40$m/s. The energy consumption for each unit distance is set to 1 unit. The reward parameters: $r_1 = 100, r_2 = 100, p_1 = 1, p_2 = p_3 = p_4 = 10$. Our simulation runs are performed with Tensorflow 1.13.1 and Python 3.7. It can be seen that in Fig. 5 the training process has converged at about 100 episodes, and it remains stable afterwards. It has been shown that as training episodes increase, the average data storage of all IoT devices in each step remains at a low level and the average residual power level remains high. Besides, the number of devices with data overflowed and expired batteries converge to about 3 and 2, respectively. By continuously learning from the interaction with the environment, the UAV achieves timely and efficient service for the IoT devices and avoids data overflow and battery expiry of devices.

VI. Open Research Issues and Challenges

Secure transmission: Secure communication is one key challenge for UAV-enabled wireless networks because of the openness of the wireless transmission medium. Traditionally, cryptographic encryption algorithms are the effective solutions. Moreover, power allocation and channel state information (CSI) of the UAV can also affect communication security. Therefore, the joint trajectory planning and power allocation of UAV-enabled wireless secure communications can be further analyzed.

Mobile edge computing: Internet of Things can support IoT devices to achieve intelligent applications, including automatic navigation, face recognition, and unmanned driving, however, the computing capability of IoT devices hinders them from performing these services. This challenge can be solved by the MEC technique. Thus, a UAV-enabled MEC system should be further designed to provide intelligent services to IoT devices.

Cache-enabled UAVs: The content requirements of edge IoT devices may not be properly met by ground BSs. Fortunately, UAVs can be deployed as the flying BSs to cache the popular contents and, then, deliver them to the users. Thus, UAV-enabled communications with caching should be further studied to provide better wireless services as well as caching for IoT devices.

Joint Resource Allocation: In our scenario, we focus on the prediction of service requirements of IoT devices and path planning of a UAV. After the UAV updates its position, the management of wireless resources should be further analyzed with the consideration of varying channel condition.

VII. Conclusion

UAV-enabled SWIPT can be utilized to overcome the low power of IoT devices and connectivity in complicated environments for IoT networks. In this paper, a unified framework of UAV-enabled SWIPT for IoT networks has been established. First, the 3D trajectory and charging strategy of a UAV have been investigated to maximize the harvested energy of ERs, as a result, a multi-beam generation scheme has been proposed that enables a UAV to serve multiple ERs simultaneously. Then, a trajectory planning scheme for multiple UAVs has been established to maximize the minimum throughput of ground users via jointly optimizing the trajectory and resource scheduling of UAVs. In addition, a DDPG-based method and dynamic path planning were proposed to estimate the devices' service requirements and maximize the EE of the system, respectively. Finally, some potential research directions and challenges have also been discussed.

REFERENCES

[1] N. Zhao, W. Lu, M. Sheng, Y. Chen, J. Tang, F.R. Yu, and K.-K. Wong, "UAV-assisted emergency networks in disasters," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 45–51, Feb. 2019.

[2] J. Xu, Y. Zeng, and R. Zhang, "UAV-enabled wireless power transfer: Trajectory design and energy optimization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5092–5106, Aug. 2018.

[3] L. Xie, J. Xu, and R. Zhang, "Throughput maximization for UAV-enabled wireless powered communication networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1690–1703, Apr. 2019.

[4] S. Yin, Y. Zhao, L. Li, and R. Yu, "UAV-assisted cooperative communications with time-sharing information and power transfer," *IEEE Trans. Veh. Technol.*, to be published.

[5] K. Li, W. Ni, E. Tovar, and A. Jamalipour, "On-board deep Q-network for UAV-assisted online power transfer and data collection," *IEEE Trans. Veh. Technol.*, to be published.

[6] F. Zhou, Y. Wu, R.Q. Hu, and Y. Qian, "Computation rate maximization in UAV-enabled wireless-powered mobile-edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1927–1941, Sep. 2018.

[7] X. Sun, W. Yang, Y. Cai, Z. Xiang, and X. Tang, "Secure transmissions in millimeter wave SWIPT UAV-based relay networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 785–788, Jun. 2019.

[8] Y. Du, K. Yang, K. Wang, G. Zhang, Y. Zhao, and D. Chen, "Joint resources and workflow scheduling in UAV-enabled wirelessly-powered MEC for IoT systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 187–10 200, Oct. 2019.

[9] L. Li, Y. Xu, Z. Zhang, J. Yin, W. Chen, and Z. Han, "A prediction-based charging policy and interference mitigation approach in the wireless powered Internet of Things," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 439–451, 2018.

[10] H. Tuy and F.A. Al-Khayyal, "Global optimization of a nonconvex single facility location problem by sequential unconstrained convex minimization," *J. Global Optim.*, vol. 2, no. 1, pp. 61–71, Mar. 1992.

[11] Q. Zhang and H. Li, "MOEA/D: A multi-objective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[12] C.H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, Sep. 2018.

[13] C. Yu, M. Zhang, F. Ren, and G. Tan, "Multiagent learning of coordination in loosely coupled multiagent systems," *IEEE Trans. Cybern.*, vol. 45, no. 12, pp. 2853–2867, Dec. 2015.