Beacon in the Air: Optimizing Data Delivery for Wireless Energy Powered UAVs

Huajian Jin[1], Jiangming Jin[2], and Yang Zhang[1(✉)]

[1] Wuhan University of Technology, Wuhan, China
{jinhuajian,yangzhang}@whut.edu.cn
[2] TuSimple, Beijing, China
jiangming.jin@tusimple.com

Abstract. UAV-aided Internet of Things (IoT) systems enable IoT devices to relay up-to-date information to base stations with UAVs, which extends the IoT network coverage and improves data transmission efficiency. To achieve a perpetual UAV data delivery system, simultaneous wireless data and power transfer (SWIPT) is employed for energy-constrained UAVs to harvest energy from wireless chargers to support data sensing and transmission from IoT devices (e.g., sensors) deployed at different locations. In this paper, the design objective is to pursue the optimal energy charging policy for each UAV considering the system states of location, the queue length and energy storage. We formulate and solve a Markov decision process for the UAV data delivery to optimally take the actions of energy charging, and data delivery to base stations. The performance evaluation shows that the proposed MDP scheme outperforms baseline schemes in terms of lower expected overall cost and high energy efficiency.

Keywords: Unmanned Aerial Vehicle · Wireless energy harvesting · Markov decision process

1 Introduction

Internet of Things (IoT) systems, e.g., wireless sensor networks (WSNs) [1], provide a spatially distributed cyber-physical approach to interconnect various components and enable efficient data transmission. To improve data transmission efficiency, using Unmanned Aerial Vehicle (UAV) as a relay in wireless sensor network introduced in this study. The UAV are deployed to assist the WSN and used to transfer data between sensors and base station. Unmanned aerial vehicle (UAV), also often knowns as drones, has been used in many areas ranging from agriculture, to military, and disaster scenarios [2], providing remote data collection and service providing. UAVs can be sent to different geographically

Supported by National Natural Science Foundation of China (Grant No. 61601336).

X. B. Zhai et al. (Eds.): MLICOM 2019, LNICST 294, pp. 173–185, 2019.
https://doi.org/10.1007/978-3-030-32388-2_15

locations to deliver on-site data in a real time manner. With the agile mobility and communication capability of UAVs, IoT communications can be efficiently extended.

Meanwhile, UAV-aided wireless communications can also introduce many new challenges. In particular, the energy consumption by the communication equipment of the UAV is substantial and reduce useful flying time by more than one-fifth [3]. Besides, once a UAV has not enough energy to transfer data stored in its queue to a base station, it will introduces high packet delivery delay. Fortunately, simultaneous wireless information and power transfer (SWIPT) as a upsurge of recent research topic can be a cost-effective way to replenish the energy of a UAV by radio frequency (RF) transmission [4]. Moreover, an efficient energy charging policy is highly desirable. RF is employed as a source of backscatter transmission in the work [5].

In this paper, we propose an optimal wireless energy charging policy focusing on the UAV transmission energy consumption with SWIPT. To achieve this goal, we model a UAV-aided wireless sensor networks from the perspective of the UAV as a relay. The UAV equipped with wireless charging facility can move among the locations to collect the data produced from sensors and send a request for energy transferring when it is at the location with an energy source (e.g.,a wireless charger). The UAV transfers data to a base station will consume units of energy, and the UAV's battery need to replenish energy for transferring by charging energy from energy sources at a certain cost. We formulate a Markov decision process to minimize the cost of the UAV consisting of the delay of storing data, the payments to the energy sources and the penalty cost of energy insufficiency. We conduct extensive simulations, which shows that the proposed MDP scheme greatly outperforms other baseline schemes.

2 Related Work

Several previous studies employing UAV in wireless networks have been proposed. For example, [2] illustrated typical use cases of UAV-aided wireless communications and surveyed several future challenges, including the energy constraint issues. Experimental analysis in [3] confirmed that energy constraint and power consumption can be one of the key research concerns in UAV applications.

How to improve energy efficiency of the UAV has became one of major research issue. An energy-efficient relaying scheme was introduced in [6] by decoupling the processes of energy balancing and data rate adjustment. Designing the trajectory of drones has been used to reduce energy consumption in [7–9]. The literature [7] studied the scenario in which a UAV-mounted energy transmitter broadcasts energy to distributed IoT devices as energy receivers on the ground. The Pareto boundary of the energy region has been characterized by optimizing the UAV's trajectory. In [8], the authors determined the optimal ground terminal transmit power and UAV trajectory by analytically deriving the energy consumption expressions of the UAV and ground terminal in a UAV-enabled data collection system. Energy-efficient UAV communication with a

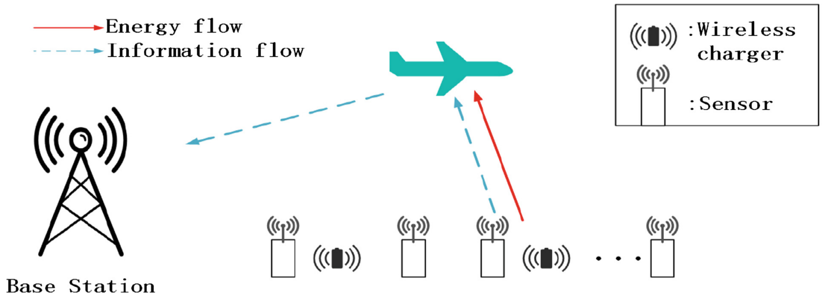


Fig. 1. System model.

ground terminal via trajectory optimization by considering the energy consumption of UAV has been studied in [9]. The authors in [10] maximized the spectrum efficiency and energy efficiency by jointly optimizing the UAV's relaying communication time allocation and the UAV trajectory in a UAV-enabled mobile relaying system. In [11] the authors proposed a novel design for energy-efficient data collection in UAV-enabled WSNs which jointly optimizes the wakeup schedules of the sensors and UAV's trajectory to minimize the maximum energy consumption of all sensors.

Instead of energy consumption management, wireless power transfer technique can be employed for perpetually replenishing energy in UAV-aided IoT systems. Using laser power as energy supply resource for UAV, the work in [12] maximized the UAV's communication throughput by jointly optimizing the UAV's trajectory and its transmit power allocation. The authors of [13] proposed an orthogonal frequency division multiplexing relaying based SWIPT protocol for energy-constrained UAV communication network. The authors of [14] solved the end-to-end cooperative throughput maximization problem by using UAV serves as an aerial mobile relay and its transmission capability powered with SWIPT.

To the best of our knowledge, employing wireless energy harvesting technology to replenish UAVs to provide perpetual wireless data collection and transmission has not been studied in recent literature. Moreover, wireless energy charging policy for UAV need to be studied for efficient data relaying.

3 System Model

We consider a UAV-aided IoT system which consists of four major components: End IoT devices, wireless chargers, UAVs, and an UAV base station, as shown in Fig. 1.

An IoT device sensor continues to sense and generate data to be potentially transmitted by the UAV to the UAV base station for further usage. The IoT devices are geographically distributed at different locations. A UAV is equipped with energy storage (e.g., a battery) and wireless communication components can charge energy from wireless energy charging sources (e.g., dedicated and ambient radio frequencies), and relay data between sensors and the UAV base station. However, delivering data for end IoT devices will consume stored energy of the UAV. There is a trade-off between the action of wireless energy charging and data relaying for the UAV to decide.

We assume that the UAV consumes units of energy for flighting per single-hop and the communication equipment has its own battery, separate from the UAV battery. We here consider the battery in charge of communication consumption for the model to be reasonable. In each location, data are kept generated for the UAV to sense. Without loss of generality, we assume the data update process is Poisson with rate λ, the UAV can receive several data from one sensor at a slot time. The data transition consumes certain units of energy stored in the battery of the UAV. The UAV can charge its battery from energy sources (i.e., wireless chargers) when it is in a location with wireless charging facility and pay a cost for charging.

The goal is that designing an optimal wireless charging policy for the drone to performs the mission of transferring data.

4 Optimization Problem Formation

To optimize the UAV-aided data delivery, we formulate the process that the UAV receives energy from wireless charger and relays data message to the UAV base station as a Markov decision process (MDP) [15]. We define the state and action spaces and derive the transition matrix.

4.1 State and Action Space

The state space of the UAV is defined as follows:

$$\mathbb{S} = \{\mathcal{L}, \mathcal{Q}, \mathcal{E}\} \tag{1}$$

where $\mathcal{S} \in \mathbb{S}$ is a composite state including all the system state variables $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{E}$. The state $\mathcal{L} \in \mathbb{L} = \{0, \cdots, L\}$ indicates the set of all the locations which the UAV can visit, the total number of locations is $L + 1$. $\mathcal{Q} \in \mathbb{Q} = \{0, \cdots, Q\}$ denotes the number of messages stored in the queue, respectively. The maximum capacity of the queue is Q, i.e., the UAV can store up to Q data messages. $\mathcal{E} \in \mathbb{E} = \{0, \cdots, E\}$ is the energy state (i.e., the current energy level of the battery) of the UAV, where E is the maximum capacity of the stored energy in the battery.

We divide the time into time slots. In each time slot when the system is in operation, the UAV takes and action defined as $\mathcal{A} \in \mathbb{A} = \{a_0, a_1, a_2\}$, where $\mathbb{A}$ is the action space. Action $\mathcal{A} = a_0$ denotes that the UAV is idle. $\mathcal{A} = a_1$ denotes that the UAV charges energy from energy charging devices. $\mathcal{A} = a_2$ is the action that the UAV deliveries the messages back to the base station.

4.2 Transition Matrix for Location States

With the dynamics of UAV and IoT systems, the system state in each time slot changes. For the ease of notations, we divide the locations into two regions (i.e., sets of locations) with respect to the existence of wireless energy charging facilities. At locations $\mathbb{L}_{es} \in \{1, \cdots, G\}$, there are wireless chargers for the UAV charging and at locations $\mathbb{L}_{node} \in \{G+1, \cdots, L\}$, there are no chargers. Therefore, $L = |\mathbb{L}_{es}| + |\mathbb{L}_{node}|$ where there are G locations in $\mathbb{L}_{es}$ region and $L - G$ locations in $\mathbb{L}_{node}$. The transition of the location state $\mathcal{L}$ of the UAV is as follows:

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{M}_{es,es} & \boldsymbol{M}_{es,node} \\ \boldsymbol{M}_{node,es} & \boldsymbol{M}_{node,node} \end{bmatrix} \tag{2}$$

In Eq. (2), $\boldsymbol{M}_{l,l'}$ is a submatrix denoting the transition when the current location is in the region $\mathbb{L}_l$, and the next location is in the region $\mathbb{L}_{l'}$, where the footnotes l and l' denote notations es or $node$. Each element $m_{i,j}$ in $\boldsymbol{M}_{l,l'}$ represents the probability that the location changes from location $i \in \mathbb{L}_l$ to location $j \in \mathbb{L}_{l'}$.

4.3 Transition Matrix for Queue States

Once the UAV arrives at a location $\mathcal{L}$, on-site data generated by the local IoT devices will be immediately transferred to the UAV for further process. The UAV can choose whether to carry the received data until returning to the UAV base station for relaying the data in a near field manner, or to directly relay the data back to the UAV base station using far field wireless transfer. Once the action $\mathcal{A} = a_2$ is taken, the UAV will remotely transfer the stored data. In this case, the queue state transitions can be be discussed in two conditions.

Increasing Queue State Case. As all the locations are equipped with IoT devices generating data messages for the UAV to help deliver, we assume that the update process is Poisson with rate λ. Intuitively, the queue state may increase when the UAV arrives at any location before the data messages are not delivered by the UAV, i.e., $\mathcal{A} = a_2$ is not taken. The transition matrix for such increasing queue state case is denoted as follows:

$$\boldsymbol{Q}^{+}(\mathcal{Q}, \mathcal{Q}') = \begin{bmatrix} P_{0,0} & P_{0,1} & \cdots & P_{0,Q-1} & \sum_{k=Q}^{\infty} P_{0,k} \\ & P_{1,0} & \cdots & P_{1,Q-2} & \sum_{k=Q-1}^{\infty} P_{1,k} \\ & & \ddots & & \vdots \\ & & & & 1 \end{bmatrix} \tag{3}$$

Each row of the matrix $\boldsymbol{Q}^{+}(\mathcal{Q}, \mathcal{Q}')$ in Eq. (3) represents the current queue state $\mathcal{Q}$ ranging from 0 to Q, and each column denotes the queue state of the next decision period $\mathcal{Q}'$. $P_{i,k}$, $k = 0, 1, \cdots, \infty$, denotes the transition probability that k data messages are generated from the local IoT devices in a decision period when the current queue state $\mathcal{Q} = i$. We can calculate the probability that the number of data messages received by the UAV during decision period according to Poisson distribution, i.e.,$P_{i,k} = \frac{\lambda^k e^{-\lambda}}{k!}$. The matrix $\boldsymbol{Q}^+(\mathcal{Q}, \mathcal{Q}')$ is a $(Q+1) \times (Q+1)$ upper triangular matrix.

Decreasing Queue State Case. The queue state can decrease by one message if the UAV takes the action $\mathcal{A} = a_2$ when there is at least one message in the queue. At the same time, there can still be new data messages generated in the current time slot. Without loss of generality, we assume that the delivering action is taken before the data messages are generated by the local IoT devices. The transition matrix for the decreasing queue state case is denoted as follows:

$$\boldsymbol{Q}^-(\mathcal{Q}, \mathcal{Q}') = \begin{bmatrix} P_{0,0} & P_{0,1} & \cdots & P_{0,Q-1} & \sum_{k=Q}^{\infty} P_{0,k} \\ P_{1,0} & P_{1,1} & \cdots & P_{1,Q-1} & \sum_{k=Q-1}^{\infty} P_{1,k} \\ & P_{2,0} & \cdots & P_{2,Q-2} & \sum_{k=Q-2}^{\infty} P_{2,k} \\ & & \ddots & \vdots & \vdots \\ & & & P_{Q,0} & \sum_{k=1}^{\infty} P_{Q,k} \end{bmatrix} \quad (4)$$

The first row of the matrix in Eq. (4) indicates that there is currently no message stored in the UAV data queue. The rest rows denote that the queue has at least one data message to be relayed. After the data message in the queue is transferred, there can be a message arrival. Note that here we assume that the message leaves the queue of the UAV if it takes the action of transferring the data back.

Overall Queue State Transition Matrix. As aforementioned, the queue state transition relies on the action taken at current state, as shown by $\boldsymbol{Q}^+$ and $\boldsymbol{Q}^-$. The overall transition matrix of the queue state $\mathcal{Q}$ is derived as follows:

$$\boldsymbol{W}(\mathcal{Q}|\mathcal{A}) = \begin{cases} \boldsymbol{Q}^+, & \mathcal{A} \in \{a_0, a_1\}, \\ \boldsymbol{Q}^-. & \mathcal{A} = a_2. \end{cases} \quad (5)$$

where the first condition in Eq. (5) is for the case that the UAV takes idle action $\mathcal{A} = a_0$ or charging action $\mathcal{A} = a_1$. The second condition is for the case that the UAV takes delivering back action $\mathcal{A} = a_2$, where the length of queue decreases.

4.4 Transition Matrix for Energy States

We derive the transition matrix of the energy state under different cases. We assume that the UAV increases and decreases energy after the action $\mathcal{A} = a_1$ and $\mathcal{A} = a_2$) taken. The energy state transition matrix can be divided into following three cases.

Increasing Energy State Case. The battery is able to store at most E units of energy. The transition matrix is expressed as follows:

$$\boldsymbol{E}^{+}(\mathcal{E},\mathcal{E}'|\mathcal{L}) = \begin{bmatrix} 1-\alpha & \alpha & & & \\ & \ddots & \ddots & & \\ & & & 1-\alpha & \alpha \\ & & & & 1 \end{bmatrix} \tag{6}$$

where each row of the matrix denotes the current energy sate $\mathcal{E}$, and each column represents the next energy state $\mathcal{E}'$. Let α denotes the successful probability of charging energy. The shape of $\boldsymbol{E}^{+}(\mathcal{E},\mathcal{E}'|\mathcal{L})$ is a $(E+1)\times(E+1)$.

Decreasing Energy State Case. When the UAV takes the delivering back action $\mathcal{A}=a_2$, one unit of energy stored in battery of the UAV will be consumed. The transition matrix is expressed as follows:

$$\boldsymbol{E}^{-}(\mathcal{E},\mathcal{E}'|\mathcal{L}) = \begin{bmatrix} 1 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \tag{7}$$

Unchanged Energy State Case. The energy state may not change if the UAV neither charges energy nor transfers messages. Under such situation, the energy transition matrix is denoted as the case $\boldsymbol{E}^{0} = \boldsymbol{I}_{E+1}$, where $\boldsymbol{I}$ is an $(E+1)\times(E+1)$ identity matrix.

Overall Energy State Transition Matrix. The energy state transition relies on the current location state $\mathcal{L}$ and the action $\mathcal{A}$ taken, as shown in $\boldsymbol{E}^{+}$, $\boldsymbol{E}^{+}$ and $\boldsymbol{E}^{0}$. Therefore, when the current state takes action $\mathcal{A}$, we define the composite transition matrix of the location state $\mathcal{L}$ and the energy state $\boldsymbol{E}$, i.e., $(\mathcal{L},\mathcal{E})$, as $\boldsymbol{W}(\mathcal{L},\mathcal{E}|\mathcal{A})$. When action $\mathcal{A}=a_0$ is taken, $\boldsymbol{W}(\mathcal{L},\mathcal{E}|\mathcal{A}=a_0)$ is defined as follows:

$$\boldsymbol{W}(\mathcal{L},\mathcal{E}|\mathcal{A}=a_0) = \boldsymbol{L}\bigotimes\boldsymbol{E}^{0}. \tag{8}$$

since when the UAV takes idle, it does not charge any energy from energy provider. The energy state is not changed.

When action $\mathcal{A}=a_1$ is taken, the UAV charges energy. $\boldsymbol{W}(\mathcal{L},\mathcal{E}|\mathcal{A}=a_1)$ is defined as follows:

$$\boldsymbol{W}(\mathcal{L},\mathcal{E}|\mathcal{A}=a_1) = \begin{bmatrix} \boldsymbol{M}_{es,es}\bigotimes\boldsymbol{E}^{+} & \boldsymbol{M}_{es,node}\bigotimes\boldsymbol{E}^{+} \\ \boldsymbol{M}_{node,es}\bigotimes\boldsymbol{E}^{0} & \boldsymbol{M}_{node,node}\bigotimes\boldsymbol{E}^{0} \end{bmatrix} \tag{9}$$

In Eq. (9), when the UAV arrives at a location $\mathcal{L}$ belongs to region $\mathbb{L}_{node}$ without energy charging devices, it can't supplement energy for battery. Consequently,

$\boldsymbol{E}^0$ is assigned to the second row respecting the current location in region $\mathbb{L}_{node}$. Besides, the UAV can charge energy, $\boldsymbol{E}^+$ is assigned to the first row corresponding to the location in region $\mathbb{L}_{es}$.

When action $\mathcal{A} = a_2$ is taken, $\boldsymbol{W}(\mathcal{L}, \mathcal{E}|\mathcal{A} = a_2)$ is defined as follows:

$$\boldsymbol{W}(\mathcal{L}, \mathcal{E}|\mathcal{A} = a_2) = \boldsymbol{L} \bigotimes \boldsymbol{E}^-. \tag{10}$$

since when the UAV takes a_2 action, it consumes energy for transferring messages, the energy capacity level of battery will decrease.

4.5 Overall Transition Matrix

In summary, given that the action $\mathcal{A}$, we denote the transition matrix of the overall state state space as $\boldsymbol{W}(\mathcal{S}, \mathcal{S}'|\mathcal{A})$ and combine the location state, queue state and energy state transition as follows:

$$\boldsymbol{W}(\mathcal{S}, \mathcal{S}'|\mathcal{A}) = \boldsymbol{W}(\mathcal{L}, \mathcal{E}|\mathcal{A}) \bigotimes \boldsymbol{W}(\mathcal{Q}|\mathcal{A}) \tag{11}$$

where $\bigotimes$ is the Kronecker product.

5 Optimization Formulation

5.1 Immediate Cost Function

Immediate cost of the MDP model is defined as the myopic reward of the current state when the UAV takes any particular action. We define immediate cost as a function $\boldsymbol{I}(\mathcal{S}|\mathcal{A})$ of the current state $\mathbb{S} = \{\mathcal{L}, \mathcal{Q}, \mathcal{E}\}$ and the action $\mathcal{A}$ taken by the UAV, as follows:

$$\boldsymbol{I}(\mathcal{S}|\mathcal{A}) = \begin{cases} \boldsymbol{F}(\mathcal{Q}) - \rho, & (\mathcal{A} = a_2) \text{ and } (\mathcal{E} \neq 0), \\ \boldsymbol{F}(\mathcal{Q}) + \tau, & (\mathcal{A} = a_2) \text{ and } (\mathcal{E} = 0), \\ \boldsymbol{F}(\mathcal{Q}) + \rho, & (\mathcal{A} = a_1) \text{ and } (\mathcal{L} \in \mathbb{L}_{es}), \\ \boldsymbol{F}(\mathcal{Q}), & \textit{otherwise}. \end{cases} \tag{12}$$

where $\boldsymbol{F}(\mathcal{Q})$ is the cost of delay caused by the messages stored in the queue, i.e., $\mathcal{Q}$. When the UAV has enough energy to transfer messages (i.e., $\mathcal{E} > 0$), if the UAV takes the delivering back action, it incurs the reward denoted by ρ. When the UAV'battery is empty, it will incur not only the cost of delay but the insufficient energy cost τ if the UAV takes the delivering back action. Moreover, if the UAV takes the charge action from charger, it need pay ρ as cost to charger.

5.2 Solving the Optimization

Then we solve the optimization problem to find an optimal policy $\pi^*(\mathcal{S})$ of the UAV that minimizing the expected discounted long term cost of the UAV, i.e., $V_\pi^*(\mathcal{S})$. The Bellman equation [15] is employed as follows:

$$V_\pi^*(\mathcal{S}) = \min_{a\in\mathcal{A}} \left(\boldsymbol{I}(\mathcal{S}|\mathcal{A}) + \gamma \sum_{\mathcal{S}'\in\mathbb{S}} (\boldsymbol{W}(\mathcal{S},\mathcal{S}'|\mathcal{A}))V_\pi^*(\mathcal{S}') \right) \tag{13}$$

$$\pi^*(\mathcal{S}) = \arg\min_{a\in\mathcal{A}} \left(\boldsymbol{I}(\mathcal{S}|\mathcal{A}) + \gamma \sum_{\mathcal{S}'\in\mathbb{S}} (\boldsymbol{W}(\mathcal{S},\mathcal{S}'|\mathcal{A}))V_\pi^*(\mathcal{S}') \right) \tag{14}$$

Value iteration algorithm [15] is applied to solve the Bellman equation, where $\boldsymbol{I}(\mathcal{S}|\mathcal{A})$ is the immediate cost function and $\sum_{\mathcal{S}'\in\mathbb{S}}(\boldsymbol{W}(\mathcal{S},\mathcal{S}'|\mathcal{A}))V(\mathcal{S}')$ is the expected future cost as defined in Sect. 5.1. $\gamma \in [0, 1)$ is a discount factor presenting value of expected future cost.

6 Numerical Results

6.1 Parameter Setting

We consider a UAV moving between 2 locations, where location $\mathcal{L} = 1$ belongs to the region $\mathbb{L}_{es}$, which can provides energy for the UAV charging energy. The transition matrix $\boldsymbol{L}$ with both rows of $[0.4, 0.6]$. The maximum capacity of the queue and the battery of the UAV are both set at 15 units. In the immediate cost function, the delay cost is set to be proportional to the number of messages stored in the queue, i.e., $\boldsymbol{F}(\mathcal{Q}) = \omega\mathcal{Q}$, where $\omega = 0.8$. The charging energy cost $\rho = 0.5$ and the immediate insufficient energy cost τ when the UAV takes $\mathcal{A} = a_1$ is 1. The probability of successfully charging energy is $\alpha = 0.95$. The discount factor γ is 0.9. The data update rate λ is 1.

For evaluating the performance of the proposed MDP policy, we consider three baseline policies consisting of a greedy policy where the UAV only minimizes the current immediate cost function, an location-aware policy where the UAV always charges energy from chargers and delivers messages back to base station in the region $\mathbb{L}_{es}$ and $\mathbb{L}_{node}$ respectively, a random policy where the UAV randomly selects an action from action sets $\mathcal{A}$.

6.2 Performance Analysis

We define following performance metrics which are evaluated and compared between the propose MDP policy and the baseline policies:

- Expected cost: We know state cost $\mathcal{C}_s$ incurred to the UAV is measured from any arbitrary initial state $\mathcal{S}$. We derive the expected cost $\mathcal{C}_s^-$as the average cost of all state cost, so $\mathcal{C}_s^- = \mathbb{E}(\sum_{s\in\mathbb{S}}\mathcal{C}_s^-)$, $\mathbb{E}$ denotes Expectation.
- Delay: The delay of messages at the UAV is equivalent to the queue length of the UAV.
- Energy inefficiency probability: The probability that the UAV is not able to transfer message owing to insufficient energy (i.e., $\mathcal{E} = 0$).

Impacts of Maximum Queue Capacity. We vary the maximum queue capacity Q from 0 to 20 and set the energy storage capacity of battery at 15 units. We then compare the results obtained from the MDP policy with the results from the baseline policies.

The expected cost of the UAV is shown in Fig. 2. As the maximum queue capacity Q increases, the costs of all policies tend to increase. The reason is that with a higher capacity Q, the UAV may accumulate more messages in its queue, causing large delay cost $\boldsymbol{F}(\mathcal{Q})$. The results of MDP policy outperform other baseline policies in terms of lowest cost, since the UAV can optimally take charging and delivering action to minimize the expected cost.

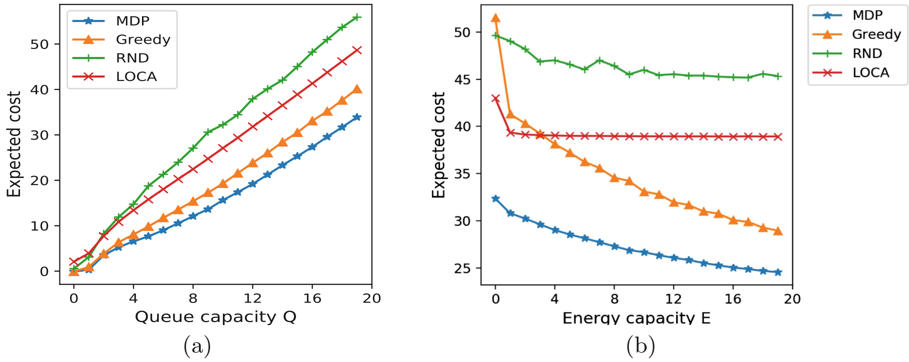


Fig. 2. Impacts to the expected cost by (a) the maximum queue capacity Q and (b) the maximum battery capacity Q.

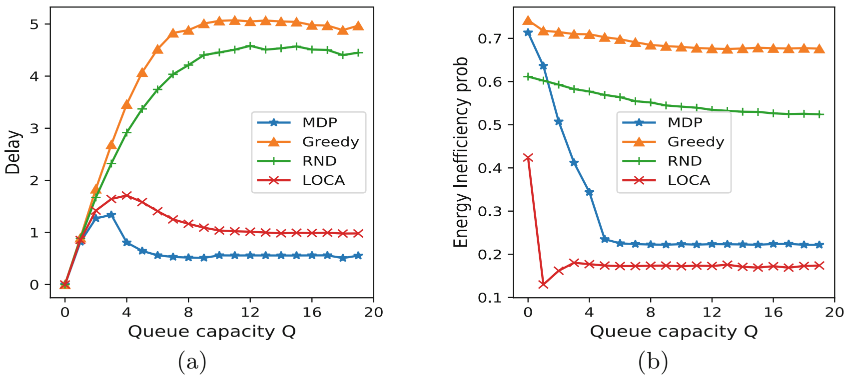
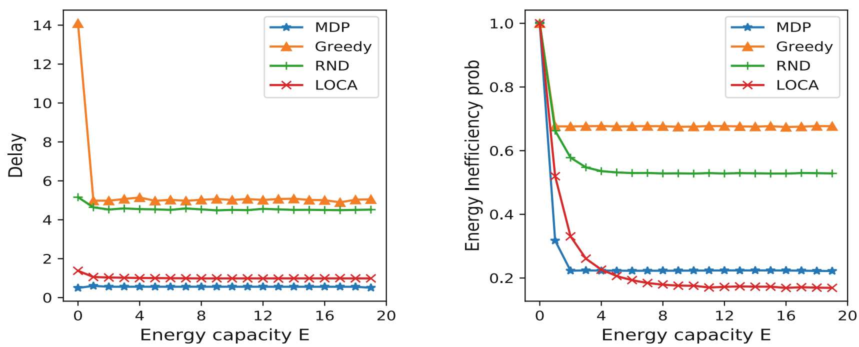


Fig. 3. Impacts of the maximum queue capacity Q to (a) delay and (b) energy insufficiency probability.

Figure 3(a) and (b) show the delay metric and the energy inefficiency probability of the steady states. Figure 3(b) shows that as Q increases, the delay of MDP policies tends to increases first. Because with a low capacity Q, the UAV tends to take idle action to reduce cost. However, after Q becomes large, e.g., $Q > 4$ for the MDP policy, the increasing delay cost make the UAV tend to take delivering back action. In contrast, the energy inefficiency probability of MDP policy tends to decreases first as Q increases, as show in Fig. 3(a). The reason is that the UAV optimally takes charging action to avoid frequent insufficient penalty. It is worth noting that the location-aware policy gets a better performance than MDP policy in terms of energy insufficient probability. Since adopting the location-aware policy, the UAV has enough energy supply by always charging energy from chargers. A trade-off exists between minimizing the delay and energy inefficiency probability.

From Figs. 3(a) and (b), we may obtain more meaningful parameter for the system. Figure 3(a) shows that when the value of Q increases to 10, the delay does not change. In addition, as shown in Fig. 3(b), increasing the queue capacity after $Q > 6$ does not increase the energy inefficiency probability. Therefore, we find that the best value of the queue capacity is 10.

Impacts of Energy Storage Capacity. We vary the maximum number of the energy storage capacity E of the battery in the UAV. When the maximum energy storage capacity from 0 to 20, as shows in Fig. 4(a), the expected cost of MDP policy decrease. The reason is that the increase battery capacity allows more energy units to be stored to support further messages transferring by the UAV. As the queue capacity is a constant, the expected cost of random policy and location-aware policy doesn't decrease after battery capacity increases to a large value.



(a) (b)

Fig. 4. Impacts of the maximum battery capacity E to (a) delay and (b) energy insufficiency probability.

As in Fig. 2(b) shows, the delay of the proposed MDP outperforms other policies, since the UAV optimally takes charging/delivering action to reduce the delay cost $\boldsymbol{F}(\mathcal{Q})$, resulting in low delay. Figure 4(b) shows the energy insufficiency probability of all policies decrease first as battery capacity increases, this is because the UAV has more energy to transfer messages, the frequency of insufficient energy is getting lower. Similarly, the location-aware policy outperforms the MDP policies due to the UAV always charging when it meets chargers. From Figs. 4(a) and (b), it is the best that the energy storage capacity is set as $E = 3$.

7 Conclusion

In this work, a SWIPT assisted optimal wireless charging scheme for UAV data transmission and energy management has been studied by employing an MDP approach. In the optimization, the overall expected cost of UAV has been minimized, including the delay of data storage, the payment for energy charging, as well as the occasional penalty cost due to energy insufficiency. Extensive numerical studies have been conducted to show the fact that the proposed MDP outperforms the baseline schemes.

References

1. Abu Alsheikh, M., Hoang, D., Niyato, D., Tan, H.P., Lin, S.: Markov decision processes with applications in wireless sensor networks: a survey. IEEE Commun. Surv. Tutor. **17**(3), 1239–1267 (2015)
2. Zeng, Y., Zhang, R., Lim, T.J.: Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun. Mag. **54**(5), 36–42 (2016)
3. Gupta, L., Jain, R., Vaszkun, G.: Survey of important issues in UAV communication networks. IEEE Commun. Surv. Tutor. **18**(2), 1132–1152 (2015)
4. Lu, X., Wang, P., Niyato, D., Kim, D.I., Han, Z.: Wireless networks with RF energy harvesting: a contemporary survey. IEEE Commun. Surv. Tutor. **17**(2), 757–789 (2015)
5. Gao, X., Wang, P., Niyato, D., Yang, K., An, J.: Auction-based time scheduling for backscatter-aided RF-powered cognitive radio networks. IEEE Trans. Wirel. Commun. **18**(3), 1684–1697 (2019)
6. Li, K., Ni, W., Wang, X., Liu, R., Kanhere, S., Jha, S.: EPLA: energy-balancing packets scheduling for airborne relaying networks. In: IEEE International Conference on Communications (2015)
7. Xu, J., Zeng, Y., Zhang, R.: UAV-enabled wireless power transfer: Trajectory design and energy region characterization. In: 2017 IEEE Globecom Workshops (GC Wkshps), Singapore, pp. 1–7 (2017)
8. Yang, D., Wu, Q., Zeng, Y., Zhang, R.: Energy trade-off in ground-to-UAV communication via trajectory design. IEEE Trans. Veh. Technol. **67**(7), 6721–6726 (2018)
9. Zeng, Y., Zhang, R.: Energy-efficient UAV communication with trajectory optimization. IEEE Trans. Wirel. Commun. **16**(6), 3747–3760 (2017)
10. Zhang, J., Zeng, Y., Zhang, R.: Spectrum and energy efficiency maximization in UAV-enabled mobile relaying. In: IEEE International Conference on Communications (2017)

11. Zhan, C., Zeng, Y., Zhang, R.: Energy-efficient information collection in UAV enabled wireless sensor network. IEEE Wirel. Commun. Lett. **7**(93), 328–331 (2017)
12. Ouyang, J., Che, Y., Xu, J., Wu, K.: Throughput maximization for laser-powered UAV wireless communication systems. In: IEEE International Conference on Communications Workshops (ICC Workshops), Kansas City, MO, pp. 1–6 (2018)
13. Lu, W., Fang, S., Gong, Y., Qian, L., Liu, X., Hua, J.: Resource allocation for OFDM relaying wireless power transfer based energy-constrained UAV communication network. In: IEEE International Conference on Communications Workshops (ICC Workshops), Kansas City, MO, 2018, pp. 1–6 (2018)
14. Yin, S., Zhao, Y., Li, L.: UAV-assisted cooperative communications with time-sharing SWIPT. In: IEEE International Conference on Communications (ICC), Kansas City, MO, pp. 1–6 (2018)
15. Bellman, R.: A Markovian decision process (1957)