

UAV-Assisted Wireless Charging for Energy-Constrained IoT Devices Using Dynamic Matching

Chunxia Su, *Student Member, IEEE*, Fang Ye[✉], *Member, IEEE*, Li-Chun Wang[✉], *Fellow, IEEE*, Li Wang[✉], *Senior Member, IEEE*, Yuan Tian, *Student Member, IEEE*, and Zhu Han, *Fellow, IEEE*

Abstract—In the emerging Internet-of-Things (IoT) paradigm, the lifetime of energy-constrained devices (ECDs) cannot be ensured due to the limited battery capacity. In this article, unmanned aerial vehicles (UAVs) are served as carriers of wireless power chargers (WPCs) to charge the ECDs. Aiming at maximizing the total amount of charging energy under the constraints of the UAVs and WPCs, a multiple-period charging process problem is formulated. To address this problem, bipartite matching with one-sided preferences is introduced to model the charging relationship between the ECDs and UAVs. Nevertheless, the traditional one-shot static matching is not suitable for this dynamic scenario, and thus the problem is further solved by the novel multiple-stage dynamic matching. Besides, the wireless charging process is history dependent since the current matching result will influence the future initial charging status, and consequently, the Markov decision process (MDP) and Bellman equation are leveraged. Then, by combining the MDP and random serial dictatorship (RSD) matching algorithm together, a four-step algorithm is proposed. In our proposed algorithm, the local MDPs for the ECDs are set up first. Next, using the RSD algorithm, all possible actions can be presented according to the current state. Then, the joint MDP is built based on the

local MDPs and all the possible matching results. Finally, the Bellman equation is utilized to select the optimal branch. Finally, simulation results demonstrate the effectiveness of our proposed algorithm.

Index Terms—Dynamic matching, energy-constrained Internet-of-Things (IoT) device, Markov decision process (MDP), unmanned aerial vehicle (UAV), wireless charging.

I. INTRODUCTION

THE PROMISING Internet of Things (IoT) has introduced enormous devices into the Internet environment, such as laptops, refrigerators, vehicles, and wireless sensors [1]. However, some of the devices are constrained by limited battery capacity [2], which restricts the development of IoT. As presented in [3], in order to collect critical information, the wireless sensors are distributed in certain harsh environments (e.g., desert and rain forest), but they will be powered off someday because of limited battery lives. Unfortunately, certain harsh places are not easily reached by the human beings for battery replacement. Another application restricted by limited energy is illustrated in [4]. As described in [4], some sensors are embedded into a bridge to monitor its health, and in that case, the bridge collapses disaster will not happen. Nonetheless, it is unrealistic to charge these embedded sensors by wired way or change battery manually. From the above two cases, the problem of charging the energy-constrained devices (ECDs) in IoT is nontrivial [5]. Furthermore, charging the ECDs by a periodic way makes it possible that the IoT network becomes a sustainable one [6].

As mentioned above, charging the ECDs in IoT by wires is unrealistic. Fortunately, one emerging technique, wireless power transfer (WPT), provides a promising way to charge these devices [7], as shown in [8], WPT can be utilized to power electric vehicle. Until now, WPT technologies have been researched extensively, e.g., inductive coupling, electromagnetic radiation, and magnetic resonant coupling [9]. For the inductive coupling technique, it is conducted over a short distance, which is usually less than a coil diameter. As for the electromagnetic radiation technique, the distance between the transmitter and the receiver can be very large [10]. Nevertheless, the charging efficiency is quite small [11]. Regarding the magnetic resonant coupling technique proposed

Manuscript received August 29, 2019; revised November 11, 2019; accepted January 12, 2020. Date of publication January 23, 2020; date of current version June 12, 2020. The work of Chunxia Su, Fang Ye, and Yuan Tian was supported in part by the National Natural Science Foundation of China under Grant 61701134; and in part by the Natural Science Foundation of Heilongjiang Province, China, under Grant F2017004. The work of Li-Chun Wang was supported by the Ministry of Science and Technology through Pervasive Artificial Intelligence Research Labs, Taiwan, under Grant MOST 108-2634-F-009-006 and Grant MOST 109-2634-F-009-018. The work of Li Wang was supported in part by the National Natural Science Foundation of China under Grant 61871416, and in part by the Fundamental Research Funds for the Central Universities under Grant 2018XKJC03. The work of Zhu Han was supported by the U.S. Multidisciplinary University Research Initiative under Grant 18RT0073, Grant NSF EARS-1839818, Grant CNS1717454, Grant CNS-1731424, Grant CNS-1702850, and Grant CNS-1646607. (Corresponding author: Fang Ye.)

Chunxia Su, Fang Ye, and Yuan Tian are with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China (e-mail: suchunxia33@gmail.com; yefang0923@126.com; tianyuan@hrbeu.edu.cn).

Li-Chun Wang is with the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: lichun@g2.nctu.edu.tw).

Li Wang is with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the Key Laboratory of the Universal Wireless Communications, Ministry of Education, Beijing 100876, China (e-mail: liwang@bupt.edu.cn).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: hanzhu22@gmail.com).

Digital Object Identifier 10.1109/JIOT.2020.2968346

in [12], it shows its superiority for our scenario, since it can transmit power within several meters with comparatively high efficiency.

On the other hand, assigning human beings to implement the wireless charging process for energy-constrained IoT devices is also impractical, especially in harsh environment, so unmanned aerial vehicles (UAVs) are leveraged in this article. Nowadays, UAVs are attracting much attention for their extensive potential applications, flexibility, and cost effectiveness [13]. In [14], the UAVs are utilized to deliver packages, which is a promising development. Meanwhile, the UAVs can also be regarded as data collector because of its mobility, as depicted in [15]. In addition, the UAVs are deployed as the mobile stations or mobile relays to support wireless communications, as shown in [16] and [17], respectively. Especially, the UAVs are served as relays for wireless sensor networks in [18], replacing the static relay with large power consumption. Except the typical use cases depicted in [7], allocating the UAVs to charge these ECDs by leveraging the WPT technique is becoming another promising research [19]. Griffin and Detweiler [20] showed the feasibility of utilizing the UAVs to power the sensors out of grid by experimental results. Nevertheless, the energy of a UAV is limited because of its small size and low cost as presented in [20], and so the UAV cannot be used as the source of energy transmitter. As a result, in this article, the UAVs are regarded as the carriers of energy transmitters.

Some previous works have dealt with the wireless charging problems, and we will analyze the literature and distinguish this article from the previous related works. In [4], the potential of utilizing the UAVs to charge wireless sensor networks by WPT is analyzed. In addition, Leng [4] evaluated the influence of different parameters on the lifetime of wireless sensor networks. In [21], the UAV is utilized to maximize the lifetime of the wireless sensor network by charging the sensors. Meanwhile, Basha *et al.* [22] answered the question of how much benefit can be obtained by leveraging the UAVs to power the sensors and what is the condition. However, only one-shot charging process is considered in both of the two literatures. In [13], the UAVs act as an energy source provider and charge the D2D pairs utilizing energy harvesting. In [13], the time slot is divided into the energy harvesting phase and information transmission phase, and the energy consumed in the information transmission phase is restricted by the energy received in the energy harvesting phase. Nonetheless, the resource allocation problem in [13] does not consider the dynamic characteristic. Another UAV-enabled charging system is described in [19], in which the sum energy maximization problem and the minimum received energy maximization problem are addressed separately. Nevertheless, Xu *et al.* [19] focused on the trajectory design and only one UAV is considered, as a consequence, the UAV allocation problem is not involved. Besides, the radio-frequency-based wireless charging problem is addressed for wireless sensor networks in [23]. Taking account of energy harvesting and information transmission together, both frequency-division multiplexing and time-division multiplexing are adopted in [23]. Nonetheless, the research in [23] focuses on power allocation problem while

this article aims to solve the association problem between the UAVs and energy-constrained IoT devices. Furthermore, the UAV-assisted WPT scenario in mobile-edge computing is applied in [24]. While taking the energy harvesting causal constraint and the UAV's speed constraint into account, the resource allocation problem is tackled. In [24], the UAVs can not only provide energy for the devices but also perform computation-intensive tasks for the devices by the equipped computing processor. Unfortunately, the energy constraint of the UAV is not considered in this reference.

As described above, few works focus on the charging association problem between the UAVs and ECDs. Besides, the UAVs can be utilized by one-time manner or periodic manner [25], nonetheless, previous works mainly focus on the one-shot allocation problem. To fill this gap and ensure sustainable energy for the ECDs, in this article, the UAV-assisted multiple period wireless charging problem is tackled. Accordingly, the house allocation (HA) model, bipartite matching with one-sided preferences, is leveraged. At the same time, it is a multiple stage dynamic matching, i.e., multiple matching results will be obtained during the whole charging process. It is worth noting that the multiple stage matching is not a simple repeating of one-shot matching. Intuitively, the charging process is history dependent, since the initial energy status of the current period is not only related to the initial energy status of the last period but also is affected by the matching result of the last period, which can be formulated as a Markov decision process (MDP). To the best of our knowledge, this is the first paper that deals with the multiperiod matching problem when considering the history-dependent wireless charging process.

The main contribution of this article is elaborated as follows.

- 1) A novel UAV-enabled multiperiod wireless charging process is presented to provide sustainable energy for the ECDs. In the proposed scenario, the Markov property of the charging process is taken into account, since the current state and action can affect the next period state. After considering the realistic scenario, the energy constraints of the UAVs and wireless power chargers (WPCs) are also considered.
- 2) To tackle the multiperiod wireless charging problem, the multistage dynamic bipartite matching with one-sided preference is proposed to model the interaction between the ECDs and UAVs. Quite different from the existing matching algorithm, this article focuses on multiple stage dynamic matching instead of one-shot matching. In order to conduct the dynamic matching, the Bellman equation-based evaluation function taking both the current reward and the future reward into account is leveraged to maximize the individual ECD's charging amount.
- 3) To achieve higher charging amount, the proactive four-step charging mechanism is designed. In the first step, the local MDPs are set up for all ECDs, which is crucial for predicting the future reward. In the second step, the random serial dictatorship (RSD) matching algorithm is proposed to take further action according to the current state. In the third step, the joint MDP is obtained based on the results in the first two steps. Finally, the Bellman

equation-based estimated reward is constructed to select the optimal branch.

- 4) To evaluate the performance of the proposed mechanism, numerical simulations are presented. The simulation results demonstrate the good performance of our proposed algorithm. Besides, the implementation process and parameter influence of the proposed algorithm are presented, which provide more insightful analyses.

The remainder of this article is organized as follows. In Section II, the UAV-assisted multiperiod wireless charging problem is presented, while taking into account of the MDP. Then some preliminaries regarding MDP and matching are introduced in Section III. In Section IV, the sequential decision problem is modeled by the multistage dynamic HA problem, and then is solved by a four-step solution. Simulation results shown in Section V evaluate the superior performance of the proposed charging mechanism. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, the multiperiod wireless charging process between the UAVs and ECDs is described in Section II-A. Then, the MDP-based states and actions are illustrated in Section II-B. Finally, the multiperiod charging process is formulated in Section II-C.

A. Charging Model

The multiperiod wireless charging model is depicted in Fig. 1. Assume at $t = 0$, several ECDs in IoT network denoted by $\mathcal{M} = \{1, \dots, i, \dots, M\}$ are distributed in different geographical areas, and the coordinates of ECD i is $\{x_{D_i}^0, y_{D_i}^0, 0\}$. These IoT devices have different missions. However, the battery capacity of each device is limited, and then a set of UAVs are employed to charge these ECDs. The set of UAVs is represented as $\mathcal{N} = \{1, \dots, j, \dots, N\}$. To be more practical, the multiple period wireless charging mechanism is considered, i.e., these UAVs will be allocated to charge the ECDs in a periodic duration, and T period is assumed in this article, i.e., $t = \{0, 1, \dots, T - 1\}$. Assume the position of these ECDs is known as prior information and they are constant, meanwhile the UAVs are located at different stations and all the UAVs are ready for the charging mission, i.e., they have been fully charged. In particular, the energy capacity of each UAV is limited due to the constrained size and cost, and each UAV has to return to the UAV station before its energy is exhausted. The initial 3-D coordinates of these UAVs are represented by $\{x_{U_j}^0, y_{U_j}^0, h\}$, where h is the height parameter.

To fully exploit the advantage of the UAVs, each UAV is equipped with a WPC, and the charging power of WPC equipped in UAV j is P_T^j . As a result, the energy of the UAV is mainly utilized for flying and hovering. In this article, the maximum energy of each UAV is represented by $E_{\max}^j$. To ensure each UAV can return to the initial position before they are powered off, the maximum charging time t_{ch}^{ij} of each UAV is calculated by

$$E_{\max}^j \geq \frac{2d_{ij}}{v} P_f + t_{ch}^{ij} P_h \quad (1)$$

where d_{ij} is the distance between ECD i and UAV j , which is denoted by $d_{ij} = \sqrt{(x_{D_i}^0 - x_{U_j}^0)^2 + (y_{D_i}^0 - y_{U_j}^0)^2}$. Besides, to simplify, the speed of the UAVs is assumed to be the same, and is denoted by v . Furthermore, P_f and P_h , respectively, indicate the UAVs' energy consumption when flying and hovering. Thus, the maximum charging time is obtained.

Moreover, the charging efficiency is another factor that is taken into account. Assuming the efficiency of the magnetic resonant coupling is η , and then the harvested power in IoT device i is given by

$$P_R^{ij} = \eta P_T^j. \quad (2)$$

For ECD i , the maximum energy storage is $S_{\max}^i$, and the remaining energy before the charging process at time t is $S_{re,t}^i$. If UAV j is allocated to charge ECD i , the charging result can be divided into three scenarios, which are

$$\begin{cases} S_{\max}^i - S_{re,t}^i > P_R^{ij} t_{ch}^{ij} \\ S_{\max}^i - S_{re,t}^i = P_R^{ij} t_{ch}^{ij} \\ S_{\max}^i - S_{re,t}^i < P_R^{ij} t_{ch}^{ij} \end{cases} \quad (3)$$

In the first scenario, even if the UAV leverages all the energy to help the charging process, the ECD cannot be fully charged. On the other hand, the ECD's energy reaches the maximum storage capacity before the UAV is powered off, and in this scenario, the UAV's energy is not fully utilized. Furthermore, the second inequality shows the critical state. Thus, the actual charging amount of ECD i when utilizing UAV j is expressed as

$$R_{ij}^t = \begin{cases} P_R^{ij} t_{ch}^{ij}, & \text{if } S_{\max}^i - S_{re,t}^i > P_R^{ij} t_{ch}^{ij} \\ S_{\max}^i - S_{re,t}^i, & \text{if } S_{\max}^i - S_{re,t}^i \leq P_R^{ij} t_{ch}^{ij} \end{cases} \quad (4)$$

B. States and Actions

To provide sustainable energy for the IoT devices, the UAVs will be allocated to charge them in a periodic manner. For the IoT device with a constant energy consuming rate, its remaining energy at time $t+1$ is related to the remaining energy and charging decision before time $t+1$, i.e., the charging process is history dependent. To simplify, the memory-less Markov-based model is adopted. In the Markov-based model, state S_{t+1} of time $t+1$ is only related to state S_t and action A_t at time t . Here, the state of an ECD indicates its demand for energy, and the ECD with lower energy will be more eager to be served by the UAV with high charging capacity. In Section IV-A, the state will be defined as the ECD's preference profile over the UAVs, and the UAV that can provide higher energy will be more preferred.

Now that the charging system needs to allocate the UAV to a certain ECD, the action can be expressed by $A_t = \{\alpha_i^t, \alpha_{-i}^t\}$, where α_i^t is the association vector of ECD i at time t , which can be further presented as $\alpha_i^t = \{\alpha_{i1}^t, \alpha_{i2}^t, \dots, \alpha_{iN}^t\}$. The variable α_{ij}^t indicates the charging allocation result, and $\alpha_{ij}^t = 1$ denotes that UAV j is allocated to charge ECD i , and *vice versa*. Besides, α_{-i}^t represents the association vector of all ECDs except i at time t . Assuming the underlying transition model is $P(S_{t+1}|S_t, A_t)$, each rational IoT device tries to maximize its individual charging amount. Since the charging

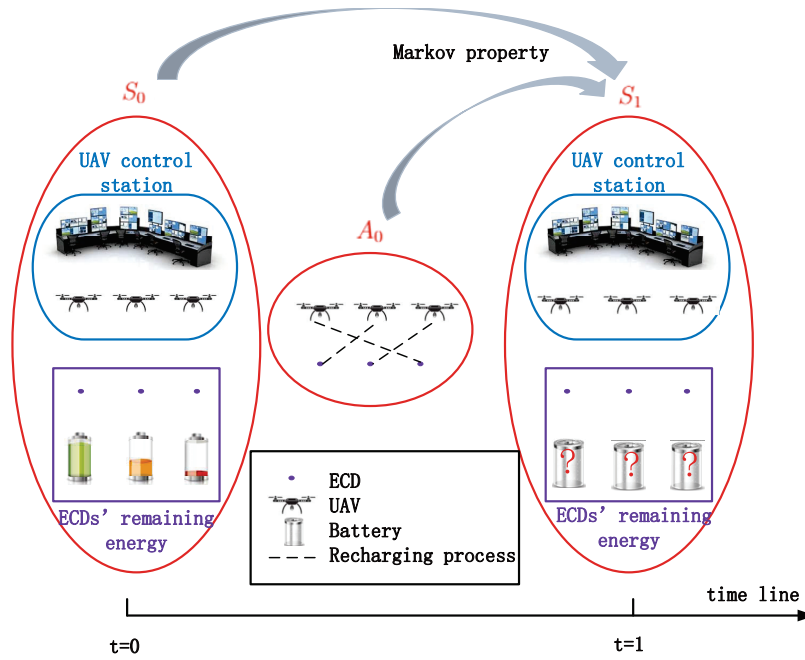


Fig. 1. System model of multiperiod wireless charging.

system is based on the Markov model, we not only consider the current time charging amount but also take account of the future possible charging amount.

C. Problem Formulation

By taking both the current charging amount and the future charging amount into account, the estimated possible charging amount of ECD i at time t is obtained by

$$U_i^\pi(S_t) = C_t^\pi(a_i^t, a_{-i}^t | S_t) + \gamma F_t^\pi(a_i^t, a_{-i}^t | S_t). \quad (5)$$

The first item means the current charging amount if action (a_i^t, a_{-i}^t) is adopted, and the second item indicates the possible future charging amount if action (a_i^t, a_{-i}^t) is taken. γ represents the discounting factor, indicating the different important levels of the current result and the future result, and $0 \leq \gamma \leq 1$ [26]. In addition, π is the charging mechanism, which will be designed in this article. More concretely, the estimated possible charging amount can be expressed by the Bellman equation, which is shown as

$$U_i^\pi(S_t) = f_i[\pi(A_t | S_t)] + \gamma \sum_{A_t} \sum_{S^{t+1}} \pi(A_t | S_t) P(S_{t+1} | S_t, A_t) U_i^\pi(S_{t+1}). \quad (6)$$

Here, $U_i^\pi(S_{t+1})$ is the estimated charging amount of time $t+1$, which can be obtained recursively. Besides, $f_i[\pi(A_t | S_t)]$ indicates the reward function at time t if the adopting mechanism is $\pi(A_t | S_t)$, more specifically, the charging amount function is further presented as

$$f_i[\pi(A_t | S_t)] = \sum_{j=1}^N \alpha_{ij}^t R_{ij}^t. \quad (7)$$

Since every rational and selfish ECD tries to maximize its own estimated possible charging amount, and then the problem

formulation is expressed by

$$\begin{aligned} \max_{\alpha_i^t, \alpha_{-i}^t} \quad & C_t^\pi(a_i^t, a_{-i}^t | S_t) + \gamma F_t^\pi(a_i^t, a_{-i}^t | S_t) \\ \text{s.t.} \quad & \text{C1: } \alpha_i^t = \{\alpha_{i1}^t, \alpha_{i2}^t, \dots, \alpha_{iN}^t\} \\ & \text{C2: } \alpha_{-i}^t = \{\alpha_1^t, \dots, \alpha_{i-1}^t, \alpha_{i+1}^t, \dots, \alpha_M^t\} \\ & \text{C3: } \alpha_{ij}^t \in \{0, 1\} \quad \forall i \quad \forall j \quad \forall t \\ & \text{C4: } \sum_j \alpha_{ij}^t \leq 1 \quad \forall i \quad \forall t \\ & \text{C5: } \sum_i \alpha_{ij}^t \leq 1 \quad \forall j \quad \forall t. \end{aligned} \quad (8)$$

Here, C1 and C2 describe the action set. C3 indicates the assignment element, while $\alpha_{ij}^t = 1$ represents that UAV j is assigned to ECD i at time t , and *vice versa*. Besides, C4 and C5 illustrate that the maximum association number of the ECD and UAV is one. The problem formulation above involves a series of optimization problem. At time $t = 0$, each ECD tries to maximize its estimated charging amount according to the initial state S_0 . Subsequently, at $t = 1$, all ECDs aim to maximize their estimated charging amount according to the state S_0 and action A_0 at $t = 0$. Thus, the charging mechanism design needs to be a sequential decision process.

III. PRELIMINARIES

To model the relationship between the ECDs and UAVs, the matching theory is introduced in this article [27]. Distinguished from the bipartite graph-based matching leveraged in [28], the matching models in [27] involve the notion of preference, and the matching results can be achieved in distributed manners. In this article, the bipartite matching with one-sided preferences is leveraged. Here, the UAVs are employed to accomplish the charging mission and they will not have preferences over the ECDs just as the houses will

not have preferences over the agents; while for the ECDs, they will have preferences over the UAVs because the charging capacities of different UAVs vary greatly, and all ECDs want to be served by the UAVs with higher charging capacities. In this section, the preliminaries of the MDP are presented in Section III-A. Next, the bipartite matching with one-sided preference is illustrated in Section III-B. Finally, the sequential matching with dynamic preferences is described in detail, as shown in Section III-C.

A. Markov Decision Process Basics

Distinguished from history-dependent process, the Markov property indicates that the next state depends only on the current state and current action, and is independent of the history state and action, which is a memory-less characteristic [29]. Mathematically, an MDP is a tuple instituted by $\langle S, A, P, R, \gamma \rangle$, where S denotes the set of all possible states while A represents the set of all possible actions. Besides, for state $s \in S$ and action $a \in A$, $P(s'|s, a)$ depicts the probability of transferring to state s' when action a is taken in state s . Additionally, $R(s, a)$ is the reward obtained when action a is taken in state s . Furthermore, γ is the discounting factor, and $0 \leq \gamma \leq 1$, indicating the importance of current reward and future reward. $\gamma = 0$ indicates the future reward that is neglected when making decision, which corresponds to the myopic greedy algorithm. In the meantime, $\gamma = 1$ considers that the future reward is as important as the current reward when making decision [30].

B. Bipartite Matching With One-Sided Preference

The matching theory has attracted much attention both from industry and academia for its distributed property. The most popular one in matching theory is the bipartite matching with two-sided preference, such as stable marriage and hospitals/residents problem. For this kind of matching, both sides have preferences over the other side, which indicates that both sides have the right to choose their partners. The Gale–Shapley (GS) algorithm has been regarded as an efficient algorithm for this problem. As another important representative of the matching theory, bipartite matching with one-sided preference describes that there are two sides in the matching problem and only one side has preference over the other side. More concretely, one intuitive application is the HA problem [27], and the two sides are, respectively, the agents and the houses (or alternatives). Obviously, in the HA problem, only the agents have preferences over the houses while the houses will not have preferences over the agents, i.e., only one side has the right to choose the partners, and this is the main difference between the two kinds of matching. In this case, the traditional GS algorithm is not suitable for bipartite matching with one-sided preferences anymore [31].

C. Sequential Matching With Dynamic Preference

In the practical application scenario, the static matching model is challenged by the ever-changing environments, and thus, the sequential matching problem with dynamic preference needs to be addressed. However, the sequential matching

with dynamic preference is not just a simple combination of a series of static matching problems, since it may not be reasonable to assume that the matching preference in the future time are independent of the matching preference and the matching result in the former time. Hence, the traditional matching algorithm cannot solve this kind of problem, to handle it, the multistage dynamic matching is introduced in this article. In addition, the sequential matching is a history-dependent decision process, i.e., the next-period matching preference is related to the preferences and the matching results of the former period. To simplify, the Markov property is considered, i.e., the next-period preference is only related to the current-period preference and the current matching result. Thus, MDP is leveraged in the multistage dynamic matching to model the sequential matching problem with the dynamic preference.

IV. DYNAMIC MATCHING ALGORITHM CONSIDERING FUTURE PAYOFF

The optimization problem presented in Section II-C is modeled as a sequential matching with dynamic preferences, and the matching is bipartite matching with one-sided preferences. Since MDP is taken into account, the state can be interpreted as the preference profiles of all ECDs, and meanwhile, the action can be illustrated as the matching result. The proposed algorithm is mainly constituted by four steps. First, to take into account of the future payoff, the local MDPs of different ECDs are constructed. Second, the RSD algorithm is implemented for different states. Third, the joint MDP is set up based on the results in the first two steps. Finally, the Bellman equation-based evaluation function considering the future payoff is utilized to select the best branch. It is worth noting that the central server is needed during the decision-making process, because the UAVs and ECDs cannot make decision by themselves. In practical implementation, the central server can collect information, such as local MDPs and charging capacity, from both the ECDs and UAVs and make charging decision for this system. The detailed algorithm procedure is elaborated in Algorithm 1.

A. Step 1: Set Up Local MDPs for ECDs

As illustrated before, we assume that the sequential matching process is based on MDP. As a result, one critical point is to build the local MDP for each ECD. To obtain the local MDP, the state of each period is achieved first. Here, the state of an ECD is defined as its preferences over the UAVs. Hence, the preference profile needs to be set up. In the following, based on the previous state and the possible matching result, the next possible state can be constructed. As shown in (4), the actual charging amounts of the UAVs differ according to the different scenarios. To achieve energy efficiency, the preference of ECD i over UAV j at time t is defined as

$$PL_{ij}^t = R_{ij}^t - \beta S_{w,t}^{ij} \quad (9)$$

where $S_{w,t}^{ij}$ denotes the energy waste if allocating UAV j to ECD i at time t , and it is derived as

$$S_{w,t}^{ij} = \max \left\{ P_{R_{ch}}^{ij} - (S_{\max}^i - S_{re,t}^i), 0 \right\}. \quad (10)$$

Algorithm 1 Proposed Algorithm

Require: $\mathcal{M}$, $\mathcal{N}$, $\{x_{D_i}^0, y_{D_i}^0, 0\}$, $\{x_{U_j}^0, y_{U_j}^0, h\}$, $S_{\max}^i$, $S_{re,i}^i$, η , $E_{\max}^j$, P_T^j , P_f , P_h and T ;

Ensure: Sequential action A_T ;

- 1: **Step 1: set up local MDPs**
- 2: Calculate the actual charging amount R_{ij}^0 according to (4) and then the preference can be obtained by (9) and (10);
- 3: Construct the preference profile for each ECD, and then initial state S_0 is achieved;
- 4: Assume different assignment results based on the initial state;
- 5: Calculate the actual charging amount R_{ij}^t from (4) and the preference can be obtained by (9) and (10);
- 6: Iterate the process to obtain all possible local MDPs for the ECDs;
- 7: **Step 2: implement RSD algorithm**
- 8: Try $M!$ kinds of priority orderings for all the ECDs and then carry out the RSD matching algorithm according to (12);
- 9: Obtain the probability distribution of different kinds of matching results $\tilde{u}_t$ according to the RSD algorithm;
- 10: **Step 3: build joint MDP based on state and RSD algorithm**
- 11: Calculate the next state according to (9) and (10), and then implement step 2 to get all the possible matching results at this time.
- 12: Repeat the procedure of calculating states and actions, and then the joint MDP for all the ECDs can be constructed;
- 13: **Step 4: select the optimal branch based on Bellman equation**
- 14: **for** $t = 0, \dots, T - 1$ **do**
- 15: Calculate the Bellman equation-based evaluation function as shown in (14);
- 16: Compare the evaluation values from different branches, and select the one that can achieve the maximum evaluation value according to (16);
- 17: **end for**
- 18: Terminate with sequential decision result A_T .

Here, β is a parameter which can be used to ensure energy efficiency. Furthermore, because the ECD is selfish and rational, it always prefers the UAV that can fully charge it than the one that cannot fully charge it, and thus β is quite small.

Now that the preference is given by (9), each ECD can sort the UAVs in a decreasing order, and then the preference profiles of all ECDs are achieved. In this article, the strict linear preference is considered. Assume two UAVs j and j' at time t , we can say that ECD i prefers UAV j to UAV j' , i.e., $j \succ_i^t j'$, if and only if $PL_{ij}^t > PL_{ij'}^t$. According to the preference profile PL^t at t , the state at time t is obtained. Subsequently, we try every possible matching result in an ergodic way, and then the state of the next period can be derived given the ECDs' power consuming rate. Leveraging the method, the local MDP of each ECD will be obtained. In the following, one example is given to explain the process. In this example, there are three ECDs, denoted by ECD 1, ECD 2, and ECD 3. At the

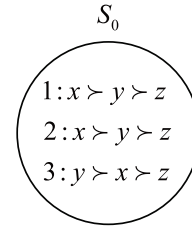


Fig. 2. Initial state.

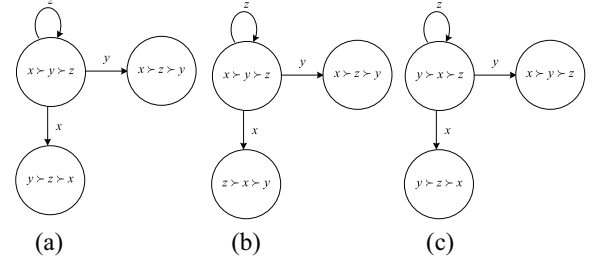


Fig. 3. Local MDPs. (a) ECD 1. (b) ECD 2. (c) ECD 3.

same time, three UAVs exist, which are UAV x , UAV y , and UAV z . Assume the initial remaining energy of the ECDs can be achieved, and then the initial state S_0 of the three UAVs can be obtained by (9) and (10), which is shown in Fig. 2. In the following, the UAVs try all possible actions, and then the next state can be predicted according to the ECDs' energy consuming rates, (9) and (10). Finally, the local MDPs of the ECDs can be achieved, as presented in Fig. 3.

B. Step 2: Take Actions According to the Current State and RSD Algorithm

Now that the state of the initial time is known, what kind of matching algorithm needs to be used to obtain matching result. Since in our matching model, only one side has preferences over the other side, the celebrated GS algorithm [31] is not applicable any more. Here, the assignment problem at time t is actually a triple with $(\mathcal{M}, \mathcal{N}, PL^t)$, i.e., the set of ECDs and UAVs, and the preference profile. Since a matching is a bijection from $\mathcal{N}$ to $\mathcal{M}$, which is denoted by u , and then we can get $u(N_j) \in \mathcal{M}$, indicating that UAV j is allocated to a certain ECD. As illustrated in [32], the assignment of the UAVs is a matching from the UAVs to ECDs, and the matching mechanism is a systematic procedure to select the matching for specific problems, and matching mechanism has been defined as π .

One of the widely researched matching mechanism is the serial dictatorship (SD) algorithm. In the SD algorithm, the priority ordering pri is defined first [33], and to be specific, $\text{pri} : \{1, 2, \dots, M\} \rightarrow \mathcal{M}$ is a one-to-one mapping that decides the priority ordering of different ECDs, i.e., ECD $\text{pri}(1)$ has the highest priority, ECD $\text{pri}(2)$ is ordered in the second, and so on. Given the SD mechanism and the priority ordering, the matching result for the ECD with priority $\text{pri}(i)$ is expressed by

$$u_{\mathcal{N}'}(\text{pri}(i)) = N_j \quad (11)$$

where $N_j \in \mathcal{N}'$, and $N_j \succ_{\text{pri}(i)} N_{j'}$ for all $N_{j'} \in \{\mathcal{N}' \setminus N_j\}$.

Based on (11), the matching result with priority order pri in the SD mechanism can be derived subsequently, as shown in

$$\begin{aligned}
 u^{\text{pri}}(\text{pri}(1)) &= u_{\mathcal{N}}(\text{pri}(1)) \\
 u^{\text{pri}}(\text{pri}(2)) &= u_{\mathcal{N} \setminus \{u^{\text{pri}}(\text{pri}(1))\}}(\text{pri}(2)) \\
 &\vdots \\
 u^{\text{pri}}(\text{pri}(i)) &= u_{\mathcal{N} \setminus \{\cup_{k=1}^{i-1} u^{\text{pri}}(\text{pri}(k))\}}(\text{pri}(i)) \\
 &\vdots \\
 u^{\text{pri}}(\text{pri}(M)) &= u_{\mathcal{N} \setminus \{\cup_{k=1}^{M-1} u^{\text{pri}}(\text{pri}(k))\}}(\text{pri}(M)). \quad (12)
 \end{aligned}$$

However, the SD mechanism is not fair for the ECDs, since the ECD with the first priority always can be allocated to its favorite UAV while the ECD with the last priority has to select whatever is remaining after all the other UAVs have made decisions. As a result, we propose to utilize the RSD algorithm to complete the matching procedure. In RSD algorithm, $M!$ kinds of priority orderings are tried with uniform distribution, and different priority orderings may lead to the same matching result [34]. Given all possible matching results $\mathcal{U} = \{u_1, u_2, \dots, u_q, \dots, u_Q\}$, the probability distribution over possible matching results $\bar{u}_t$ can be denoted by $\bar{u}_t = \{p_1, p_2, \dots, p_q, \dots, p_Q\}$, where $\sum_{q=1}^Q p_q = 1$ and $\bar{u}_t \in \Delta(\mathcal{U})$. It should be noted that the implementation of RSD algorithm is very easy. However, since all possible $M!$ kinds of priority orders need to be tried, this algorithm is suitable for the scenario with moderate ECD number. When the number of ECDs is huge, we can divide the ECDs into several clusters first according to the geographic position and then implement the RSD algorithm.

For the small example, all possible priority orders need to be listed, and the number of possible orderings is $3! = 6$. Regarding priority ordering $\text{pri}(1) = 1, \text{pri}(2) = 2$, and $\text{pri}(3) = 3$, ECD 1 has the highest priority, so ECD 1 can make decision first, and it chooses its favorite UAV x . Next, UAV x will be removed from the preferences of ECD 2 and ECD 3, and then ECD 2 chooses UAV y from the remaining UAVs. Finally, ECD 3 has to select UAV z since other options are not available. Here, the matching result xyz indicates that ECD 1 is associated with UAV x , ECD 2 is associated with UAV y , and ECD 3 is associated with UAV z . Implementing the above process for all priority orders, and all the possible matching results and the probability can be obtained as shown in Fig. 4.

C. Step 3: Set Up Joint MDP Based on Local MDPs and RSD Mechanism

According to the initial state and the RSD algorithm shown in Section IV-B, the state and all possible actions at time $t = 0$ can be obtained. Accordingly, the next state is achieved based on the last state and action. Then, all the possible matching results for the next state are obtained. Repeat the procedure of calculating states and actions, then the joint MDP will be obtained. The joint MDP for the small example is represented in Fig. 5. Given initial state S_0 , the possible matching results

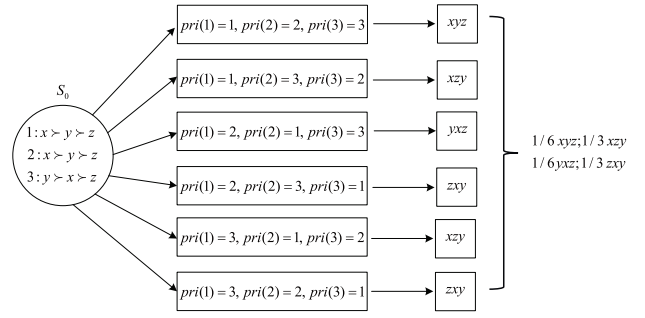


Fig. 4. Take action according to the RSD algorithm.

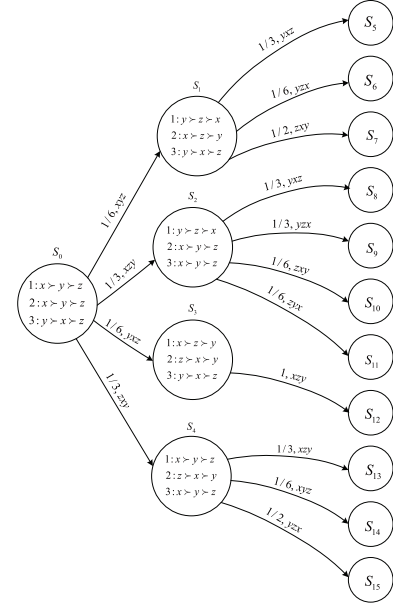


Fig. 5. Joint MDP.

and their probabilities, i.e., $(1/6)xyz$, $(1/3)xzy$, $(1/6)yxz$, and $(1/3)zxy$ can be obtained according to Section IV-B. Next, for state S_0 and matching result xyz , the preference of ECD 1 will become $y > z > x$ in state S_1 according to Fig. 3, since ECD 1 associates with UAV x . By this analogy, states S_1, S_2, S_3 , and S_4 can be achieved.

D. Step 4: Utilizing Bellman Equation-Based Evaluation Function to Select Branch

Now that the joint MDP is obtained, the last step is to select the branch that can maximize the system performance. In each period, every ECD tries to maximize its evaluated reward. Although the RSD algorithm gives several matching results and their probability distribution, the sequential decision cannot be achieved directly. In this section, the Bellman equation-based branch selection method is proposed.

Neglecting probability distribution, we describe the possible matching results at time t as u_t . Assume X optional branches at time t , the Bellman equation-based evaluation function for branch x is calculated as

$$U_B^x(S_t) = C_t(S_t) + \gamma F_t(S_t) \quad (13)$$

which can be further expressed by

$$U_B^x(S_t) = u_t^x \otimes S_t + \gamma \sum_{y=1}^Y p(\bar{u}_t^y) U_B^y(S_{t+1}). \quad (14)$$

Here, the estimated reward of current time is calculated based on the evaluated reward of next time, so the estimated reward is obtained by backward induction. Furthermore, Y is the total number of branches at time $t + 1$ if branch x is selected at time t . Besides, u_t^x indicates the specific matching result if selecting branch x is at time t . $p(\bar{u}_t^y)$ describes the possibility of branch y . Moreover, a special calculation is described as $u_x^t \otimes S_t$, which can be further calculated as

$$u_x^t \otimes S_t = \sum_{i=1}^M \sum_{j=1}^N \alpha_{ij}^{tx} R_{ij}^t \quad (15)$$

where α_{ij}^{tx} is the assignment matrix of branch x at time t , which is directly related to u_t^x , in the meanwhile, R_{ij}^t can be known from S_t . Accordingly, the optimal branch for time t can be selected, shown as

$$x^* = \max_x U_B^x(S_t). \quad (16)$$

As a consequence, the sequential decision can be made by (16). For the small example, obtain the instant utility for each charging decision at first and then calculate the Bellman equation-based evaluation function as shown in (14). Next, choose the branch with the largest evaluation function value sequentially. The elaborated algorithm procedure for the multiperiod charging process is shown in Algorithm 1.

E. Genie-Aided Policy as Performance Upper Bound

To validate the effectiveness of the proposed algorithm, the Genie-aided policy is presented as the upper bound. Aiming to maximize the sum of charging energy of all the UAVs in the T periodic duration, the system model based on the Genie-aided policy is formulated by

$$\max_{\alpha_{ij}^t} \sum_{t=0}^{T-1} \sum_{i=1}^M \sum_{j=1}^N \alpha_{ij}^t R_{ij}^t \quad (17)$$

subject to

$$\begin{cases} \alpha_{ij}^t = \{0, 1\} & \forall i \quad \forall j \quad \forall t \end{cases} \quad (18a)$$

$$\begin{cases} \sum_j \alpha_{ij}^t \leq 1 & \forall i \quad \forall t \end{cases} \quad (18b)$$

$$\begin{cases} \sum_i \alpha_{ij}^t \leq 1 & \forall j \quad \forall t. \end{cases} \quad (18c)$$

Here, the Genie-aided policy can obtain the global optimum from the view point of mathematics. However, the Genie-aided policy is difficult to be achieved from the perspective of reality. Before the implementation of the Genie-aided policy, the global information, including the current information and all possible future information, should be obtained. Nonetheless, it is unreasonable to assume that we can obtain all future information in advance. Thus, the Genie-aided policy is just deemed as the upper bound in this article.

TABLE I
SIMULATION PARAMETERS

Parameters	Values
UAVs' flying power consumption	121.91W
UAVs' hovering power consumption	92.28W
UAVs' velocity	within 10m/s ~ 20m/s
UAVs' maximum energy	within 10Wh ~ 50Wh
UAVs' maximum charging power	within 10W ~ 100W
ECDs' maximum energy storage	within 1.17Wh ~ 23.4Wh
ECDs' initial energy percentage	within 10% ~ 40%
ECDs' energy consumption rate	within 0.625mW ~ 3mW
WPT efficiency	0.3

V. SIMULATION RESULTS

To evaluate the effectiveness of our proposed algorithm, numerical simulation results are shown in this section. Besides, two other algorithms are adopted as comparison algorithms, i.e., the Genie-aided algorithm and the myopic greedy algorithm. The Genie-aided policy has been described in Section IV-E and it can be regarded as the upper bound. In the meantime, the myopic greedy algorithm indicates that the current reward is maximized at each decision-making process regardless of the future reward. In the simulation, three ECDs are randomly distributed within a 0.5 km $\times$ 0.5 km square area while three UAVs are served as mobile charging devices. In addition, two charging periods are considered while the ECDs are charged in every 60 days. The detailed simulation parameters are presented in Table I.

Given the simulation setup, the proposed algorithm is evaluated from four aspects. In Section V-A, the performance curves regarding the charging amount and energy usage efficiency are represented first. Next, the charging results in period 1 and period 2 are illustrated in Section V-B, both for the proposed algorithm and for the myopic greedy algorithm. The results distinguish the proposed algorithm from the myopic greedy algorithm, meanwhile, the results also validate the effectiveness of the proposed dynamic matching mechanism. Then, Section V-C depicts the influence of discounting factor on charging amount. Finally, the performance comparison taking account of different time spans is shown in Section V-D.

A. Performance Validation

1) *Actual Charging Amount*: In this part, the actual charging amount of the three algorithms are compared by varying different parameters. As can be seen from Fig. 6, the actual charging amounts of the three algorithms decrease with increasing distances, since the UAVs have to take more time to fly while have less time to hover above the ECDs, and thus the charging amounts are reduced. Furthermore, the Genie-aided algorithm always achieves the best charging amount, since it acquires all information in advance and then the global optimal solution can be ensured. It is obvious that the proposed algorithm achieves a higher charging amount than the myopic greedy algorithm, and at the same time, its charging amount is near to the global optimum. Compared with the myopic greedy algorithm, the proposed algorithm can obtain about 3% gain when the distance scale is equal to 3. For the myopic greedy

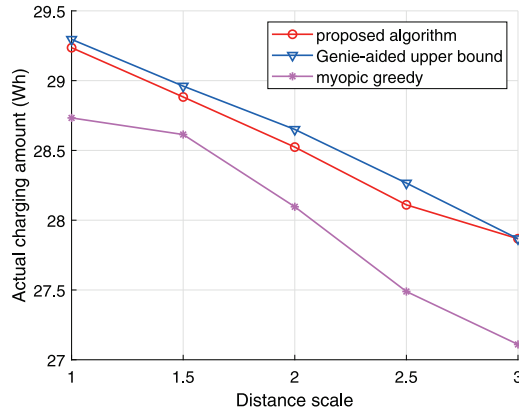


Fig. 6. Actual charging amount versus distance scale.

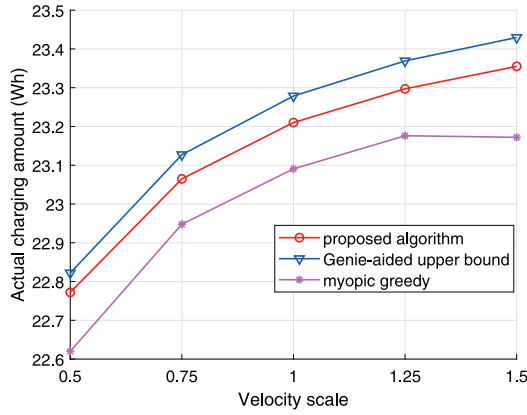


Fig. 7. Actual charging amount versus UAV's velocity scale.

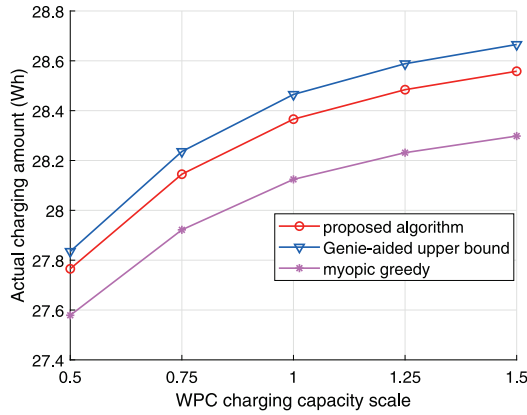


Fig. 8. Actual charging amount versus WPC charging capacity.

algorithm, it gets the worst performance since it only considers the instant reward when making decision while no future reward is taken into account. Meanwhile, the performance comparisons when changing UAV velocities and WPC charging capacities are given in Figs. 7 and 8, respectively. As depicted in Fig. 7, the larger the UAV's velocity is, the higher charging amount can be obtained, because more time can be utilized in charging process. Similarly, the larger WPC charging capacity will lead to a higher charging amount. Also, the proposed algorithm can achieve the near-optimal result while the myopic greedy algorithm obtain the worst performance.

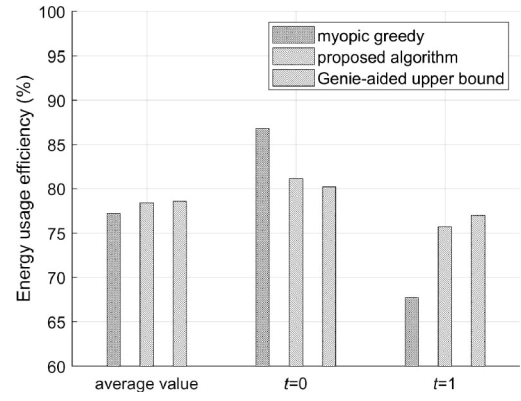


Fig. 9. Comparison of energy usage efficiency.

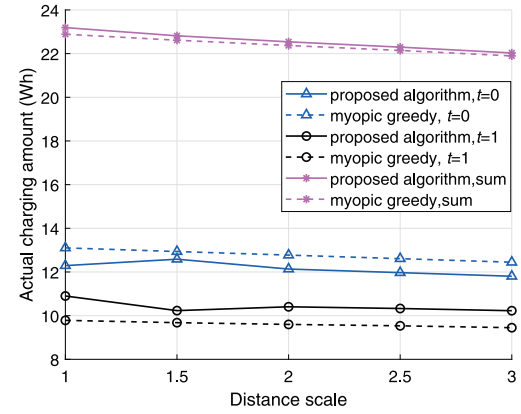


Fig. 10. Charging algorithm with increasing distances.

2) *Energy Usage Efficiency*: Then, the energy usage efficiency of the three algorithms is evaluated. Here, the total energy is estimated by assuming that the velocity of the UAV is infinite, i.e., all UAVs can make full use of their energy to hover above the ECDs. The comparison result is shown in Fig. 9. From Fig. 9, the myopic greedy algorithm always achieves the best energy usage efficiency at time $t = 0$ since it chooses the result greedily. Meanwhile, the myopic greedy algorithm will obtain the worst energy usage efficiency at time $t = 1$ because its average energy efficiency is the lowest. This indicates that utilizing myopic greedy policy is not rational in the history-dependent scenario. Besides, this figure presents that the proposed algorithm achieves nearly the same energy usage efficiency with the Genie-aided upper bound.

B. Mechanism Difference Between the Proposed Algorithm and Myopic Greedy Algorithm

In this part, the charging decision difference between the proposed algorithm and myopic algorithm is analyzed in detail, and the simulation results will verify that considering future reward when making decision is advisable.

First, the charging process is depicted in Fig. 10 when distance is changing. In the figure, $t = 0$ related curves indicate the charging amount at time $t = 0$, while $t = 1$ related curves are the charging amount at time $t = 1$. From Fig. 10, we can see that the myopic greedy algorithm always acquire the

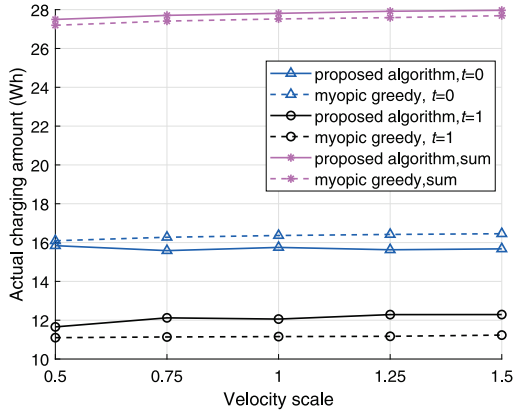


Fig. 11. Charging algorithm with increasing UAV's velocities.

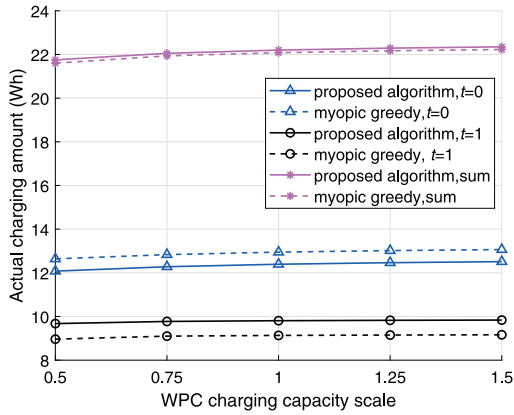


Fig. 12. Charging algorithm with increasing WPC charging capacities.

higher charging amount than the proposed algorithm when $t = 0$, since it can find the optimal value if only one time period is considered. However, the myopic greedy is not a rational decision in the history-dependent problem. Clearly, the myopic greedy policy leads to lower charging amount at time $t = 1$. In contrast, though the proposed algorithm gets lower charging amount at $t = 0$, the long-term consideration makes it achieve a higher total charging amount. Furthermore, the sum charging amount in the proposed algorithm is higher than the myopic greedy algorithm. Hence, the dynamic matching related algorithm is a wise choice when making wireless charging decision or other history-dependent decision.

Besides, the other two figures shown in Figs. 11 and 12 also validate the effectiveness of the proposed algorithm. At $t = 0$, the proposed algorithm has no advantage over the myopic greedy algorithm; however, it always can catch up with the myopic greedy algorithm at time $t = 1$ and achieves a higher charging amount.

C. Influence of Discounting Factor

Next, the influence of the discounting factor is analyzed. As we illustrated in Section II-C, the value of discounting factor γ is within $[0, 1]$. In Fig. 13, the charging amount curves are obtained with different γ . As can be seen from the figure, the higher γ will lead to a higher system performance, since the future payoff is becoming more and more important with

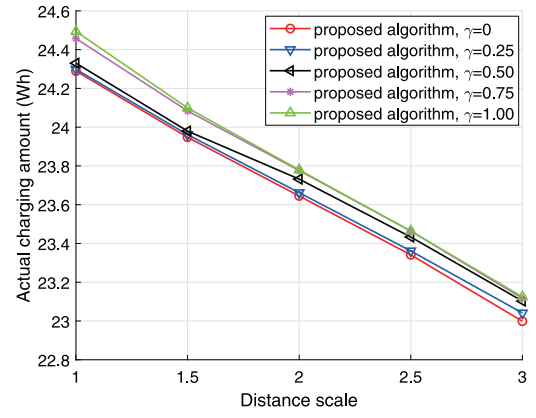


Fig. 13. Actual charging amount with different discounting factors.

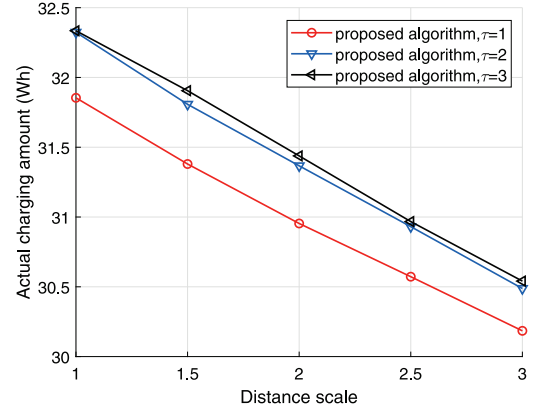


Fig. 14. Actual charging amount considering different time spans.

the increase of discounting factor γ . Obviously, actual charging amount is the highest when γ is equal to 1, which indicates treating future payoff as important as the current payoff is beneficial for performance improvement.

D. Influence of Time Span

Finally, the influence of time span τ is considered. Here, $\tau = 1$ indicates only the current reward is considered when the proposed matching algorithm is implemented. For $\tau = 2$, it means that we not only consider the current reward but also take account of the possible rewards in the next period when making the decision. Furthermore, $\tau = 3$ represents that two future periods will be taken into account except for the current period. The charging amount is presented in Fig. 14, from which we can see that the larger time span is considered, the better system performance will be achieved. However, compared to the result at $\tau = 2$, the performance improvement at $\tau = 3$ is slight while the complexity at $\tau = 3$ is much higher than the complexity at $\tau = 2$. Hence, this exactly explains why we choose parameter τ as 2 in the simulation.

VI. CONCLUSION

In this article, the UAV-assisted multiple period wireless charging problem involving history-dependent process is addressed. To model the dynamic wireless charging process, dynamic bipartite matching with one-sided preferences

is introduced as the novel solution. Then, a four-step algorithm combining MDP and Bellman equation is proposed. Next, the simulation results in Section V first validate the effectiveness of the proposed algorithm. Meanwhile, the simulation results also illustrate the charging mechanism difference between the proposed algorithm and myopic greedy algorithm. Furthermore, this article investigates the influence of different parameters (discounting factor and time span) on the charging amount from the perspective of simulations.

REFERENCES

- [1] S. F. Abedin, M. G. R. Alam, S. A. Kazmi, N. H. Tran, D. Niyato, and C. S. Hong, "Resource allocation for ultra-reliable and enhanced mobile broadband IoT applications in fog network," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 489–502, Jan. 2019.
- [2] L. Cui *et al.*, "Joint optimization of energy consumption and latency in mobile edge computing for Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4791–4803, Jun. 2019.
- [3] Y. Pang, Y. Zhang, Y. Gu, M. Pan, Z. Han, and P. Li, "Efficient data collection for wireless rechargeable sensor clusters in harsh terrains using UAVs," in *Proc. IEEE Glob. Commun. Conf. (GlobeCom)*, Austin, TX, USA, Dec. 2014, pp. 234–239.
- [4] J. Leng, "Using a UAV to effectively prolong wireless sensor network lifetime with wireless power transfer," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Nebraska-Lincoln, Lincoln, NE, USA, May 2014.
- [5] 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, Feb. 2019.
- [6] P. Ramezani, Y. Zeng, and A. Jamalipour, "Optimal resource allocation for multiuser Internet of Things network with single wireless-powered relay," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3132–3142, Apr. 2019.
- [7] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.
- [8] S. Li and C. C. Mi, "Wireless power transfer for electric vehicle applications," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 3, no. 1, pp. 4–17, Mar. 2015.
- [9] L. Xie, Y. Shi, Y. T. Hou, and A. Lou, "Wireless power transfer and applications to sensor networks," *IEEE Wireless Commun.*, vol. 20, no. 4, pp. 140–145, Aug. 2013.
- [10] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.
- [11] C. Mikeka and H. Arai, "Design issues in radio frequency energy harvesting system," in *Sustainable Energy Harvesting Technologies-Past, Present and Future*. London, U.K.: IntechOpen, Dec. 2011.
- [12] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, Jul. 2007.
- [13] H. Wang, J. Wang, G. Ding, L. Wang, T. A. Tsiftsis, and P. K. Sharma, "Resource allocation for energy harvesting-powered D2D communication underlying UAV-assisted networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 14–24, Mar. 2018.
- [14] G. Hoareau, J. J. Liebenberg, J. G. Musial, and T. R. Whitman, "Package transport by unmanned aerial vehicles," U.S. Patent 9 731 821, Aug. 2017.
- [15] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 328–331, Jun. 2018.
- [16] Q. Zhang, W. Saad, M. Bennis, X. Lu, M. Debbah, and W. Zuo, "Predictive deployment of UAV base stations in wireless networks: Machine learning meets contract theory," 2018. [Online]. Available: arXiv:1811.01149.
- [17] S. Yin, J. Tan, and L. Li, "UAV-assisted cooperative communications with wireless information and power transfer," 2017. [Online]. Available: arXiv:1710.00174.
- [18] S. Fu, L. Zhao, Z. Su, and X. Jian, "UAV based relay for wireless sensor networks in 5G systems," *Sensors*, vol. 18, no. 8, p. 2413, Jul. 2018.
- [19] 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.
- [20] B. Griffin and C. Detweiler, "Resonant wireless power transfer to ground sensors from a UAV," in *Proc. IEEE Int. Conf. Robot. Autom.*, Saint Paul, MN, May 2012, pp. 2660–2665.
- [21] J. Johnson, E. Basha, and C. Detweiler, "Charge selection algorithms for maximizing sensor network life with UAV-based limited wireless recharging," in *Proc. IEEE 8th Int. Conf. Intell. Sensors Sensor Netw. Inf. Process.*, Melbourne, VIC, Australia, Apr. 2013, pp. 159–164.
- [22] E. Basha, M. Eiskamp, J. Johnson, and C. Detweiler, "UAV recharging opportunities and policies for sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 8, pp. 1–10, Aug. 2015.
- [23] Q. Li, J. Gao, H. Liang, L. Zhao, and X. Tang, "Optimal power allocation for wireless sensor powered by dedicated RF energy source," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2791–2801, Mar. 2019.
- [24] 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.
- [25] J. Zhang, Y. Zeng, and R. Zhang, "UAV-enabled radio access network: Multi-mode communication and trajectory design," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5269–5284, Oct. 2018.
- [26] N. C. Luong *et al.*, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3133–3174, 4th Quart., 2019.
- [27] D. F. Manlove, *Algorithmics of Matching Under Preferences*, vol. 2. Singapore: World Sci., Apr. 2013.
- [28] A. Sultana, I. Woungang, L. Zhao, and A. Anpalagan, "Two-tier architecture for spectrum auction in SDN-enabled cloud radio access network," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 9191–9204, Sep. 2019.
- [29] Q. Li, L. Zhao, J. Gao, H. Liang, L. Zhao, and X. Tang, "SMDP-based coordinated virtual machine allocations in cloud-fog computing systems," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1977–1988, Jun. 2018.
- [30] D. C. Parkes and A. D. Procaccia, "Dynamic social choice with evolving preferences," in *Proc. 27th Conf. Artif. Intell. (AAAI)*, Bellevue, WA, USA, Jul. 2013, pp. 767–773.
- [31] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, vol. 69, no. 1, pp. 9–15, Jan. 1962.
- [32] A. Abdulkadiroğlu and T. Sönmez, "Random serial dictatorship and the core from random endowments in house allocation problems," *Econometrica*, vol. 66, no. 3, pp. 689–701, May 1998.
- [33] H. Hosseini, K. Larson, and R. Cohen, "Matching with dynamic ordinal preferences," in *Proc. 27th Conf. Artif. Intell. (AAAI)*, Austin, TX, USA, Jan. 2015, pp. 936–943.
- [34] S. Bade, "Random serial dictatorship: The one and only," *Math. Oper. Res.*, vol. 45, no. 1, pp. 353–368, Feb. 2020.



Chunxia Su (Student Member, IEEE) received the B.S. degree in communication engineering from Harbin Engineering University, Harbin, China, in 2014, where she is currently pursuing the Ph.D. degree.

She was a joint Ph.D. student with the Electrical and Computer Engineering Department, University of Houston, Houston, TX, USA, from 2017 to 2019. Her research interests include matching theory, contract theory, and resource allocation.



Fang Ye (Member, IEEE) received the B.S. and Ph.D. degrees in electrical information engineering from Harbin Engineering University, Harbin, China, in 2002 and 2006, respectively.

She has been a teacher with the Harbin Engineering University of China since 2002 and became an Associate Professor in 2007. From 2007 to 2008, she was a Visiting Scholar with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K. Her research interests include edge computing, resource allocation, and game theory.

Dr. Ye is a member of the China Institute of Communications and China Computer Federation.



Li-Chun Wang (Fellow, IEEE) received the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 1996.

From 1996 to 2000, he was with AT&T Laboratories, Florham Park, NJ, USA, where he was a Senior Technical Staff Member with the Wireless Communications Research Department. Since August 2000, he has been with the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan, where he is currently a Chair Professor and jointly appointed

by the Department of Computer Science and Information Engineering. He holds 21 U.S. patents, and have published over 300 journals and conference papers, and co-edited the book *Key Technologies for 5G Wireless Systems* (Cambridge University Press, 2017). His recent research interests are in the areas of cross-layer optimization for wireless systems, data-driven radio resource management, software-defined heterogeneous mobile networks, big data analysis for industrial Internet of Things, and AI-enabled unmanned aerial vehicular networks.

Dr. Wang won twice the Distinguished Research Award of Ministry of Science and Technology, Taiwan (2012–2017). He was a co-recipient of the IEEE Communications Society Asia-Pacific Board Best Award in 2015, the Y. Z. Hsu Scientific Paper Award in 2013, and the IEEE Jack Neubauer Best Paper Award in 1997. He was elected to the IEEE Fellow in 2011 for his contributions to cellular architectures and radio resource management in wireless networks.



Yuan Tian (Student Member, IEEE) received the B.S. and M.S. degrees from Harbin Engineering University, Harbin, China, in 2000 and 2006, respectively.

She is currently a Lecturer with Harbin Engineering University. Her research interests include cognitive radio network and radio resource management.



Li Wang (Senior Member, IEEE) received the Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2009.

She is currently a Full Professor with the School of Electronic Engineering, BUPT, where she heads the High Performance Computing and Networking Laboratory. She is also a Member of the Key Laboratory of the Universal Wireless Communications, Ministry of Education, Beijing, and an Associate Dean of the School of Software

Engineering, BUPT. She also held visiting positions with the School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA, USA, from December 2013 to January 2015, and the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, from August to November 2015 and July to August 2018. She has authored/coauthored almost 50 journal papers and two books. Her current research interests include wireless communications, distributed networking and storage, vehicular communications, social networks, and edge AI.

Prof. Wang was a recipient of the 2013 Beijing Young Elite Faculty for Higher Education Award, the Best Paper Award at ICCTA 2011, the Best Paper Runner Up from WASA 2015, the Best Paper Award from IEEE ICC 2017, the Demo Award from IEEE ICC 2018, the Best Paper Award from IEEE GLOBECOM 2018, the Best Paper Award from IEEE WCSP 2019, and the Beijing Technology Rising Star Award in 2018. She was selected as a Distinguished Young Investigator by the China Academy of Engineering in 2018. She currently serves on the Editorial Boards for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, IEEE ACCESS, *Computer Networks*, and *China Communications*. She was the Symposium Chair of IEEE ICC 2019 on Cognitive Radio and Networks Symposium and a Tutorial Chair of IEEE VTC 2019-Fall. She also chairs the special interest group on Social Behavior Driven Cognitive Radio Networks for IEEE Technical Committee on Cognitive Networks. In 2019, she was funded by the National High Level Young Investigator Support Plan. She has served on a TPC Member for multiple IEEE conferences, including IEEE Infocom, Globecom, International Conference on Communications, IEEE Wireless Communications and Networking Conference, and IEEE Vehicular Technology Conference in recent years.



Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was Research and Development Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with

Boise State University, Boise, ID, USA. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, Houston, TX, USA. He is also a Chair Professor with National Chiao Tung University, Hsinchu, Taiwan. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid.

Dr. Han received NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, the IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He has been 1% Highly Cited Researcher according to Web of Science since 2017. He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018. He has been an AAAS Fellow since 2019 and the ACM Distinguished Member since 2019.