Applied Energy 259 (2020) 114204

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Green wireless power transfer system for a drone fleet managed by reinforcement learning in smart industry

Giuseppe Faraci, Angelo Raciti, Santi Agatino Rizzo*, Giovanni Schembra

DIEEI – University of Catania, Italy

HIGHLIGHTS

- A FANET based on a fleet of drones provides the connectivity to IoT-enabled devices.
- A battery charge station supplied by a wind generator enables persistent mission without any human intervention.
- Reinforcement Learning optimally manage the use of the fleet to meet bandwidth request.
- Reinforcement Learning reduces the wasted energy and enables the use of the wind generator.
- Wireless power transfer based charge station has been used to face hostile environmental conditions.

ARTICLE INFO

Keywords:
Artificial intelligence
Drone
Industry 4.0
Internet of Things
Wind generator
Wireless power transfer

ABSTRACT

The optimal management of a fleet of drones is proposed in this paper for providing connectivity to sensors and actuators in Industrial Internet of Things (IIoT) scenarios. The persistent mission without any human intervention on the battery charge is obtained by means of an on-field wind generator supplying a charge station that adopts resonant wireless power transfer. The objective of the fleet management is to provide the best connectivity over the time considering the variability of both the bandwidth request and the wind energy availability. The optimal management is performed by a system controller adopting reinforcement learning (RL) for deciding the number of drones to take off and, consequently, the instantaneous provided bandwidth. A constant charge time of drone battery represents a key element of the system because this enables to strongly reduce the complexity of the system controller task. To this purpose, an adaptive current control for the charge station is introduced to compensate charge time variabilities due to the coupling factor changes caused by misalignments that can occur between a pad and a drone. The results have highlighted that the RL provides good performance improvement in case of green generation. An important aspect arose from this study is the ability of RL to increase the saved energy even if it is not considered as a target of the controller.

1. Introduction

With the widespread deployment of wireless sensor networks and embedded systems, as well as the introduction of the Internet of Things (IoT) in industrial application scenarios, industrial automation is undergoing a tremendous change, shifting the manufacturing companies' environment toward smart factories. This trend is fostering the rise of Industrial IoT (IIoT) [1–3] as the first step towards the implementation of Industry 4.0 [4,5]. For example, smart agriculture is a challenging IIoT scenario that needs smart control and intelligent decision making based on accurate real-time field data and smart warehouse management [6,7]. Nevertheless, there are many cases in which providing connectivity for IoT devices is challenging for the wide deployment of these technologies. A typical case is when the devices are dispersed in very large areas without infrastructure coverage due to, for example, severe shadowing by mountainous terrain [8], and often not covered by grid power supply. Some solutions analyzed in the past are based on the use of high-altitude platforms (HAPs), such as balloons, which usually operate tens of kilometers above the Earth's surface. Unfortunately, sources of information in IoT scenarios are typically small battery-limited devices, so typically unable to transmit over a long distance due to their energy constraints [8]. To this purpose, a very promising solution that has received significant attention recently in such scenarios is the use of unmanned aerial vehicles (UAVs), also commonly referred

* Corresponding author at: viale A. Doria 6, 95125, Catania, Italy.
E-mail addresses: giuseppe_1989@hotmail.it (G. Faraci), angelo.raciti@dieei.unict.it (A. Raciti), santi.rizzo@dieei.unict.it (S.A. Rizzo), giovanni.schembra@dieei.unict.it (G. Schembra).

https://doi.org/10.1016/j.apenergy.2019.114204
Received 5 August 2019; Received in revised form 14 November 2019; Accepted 18 November 2019
Available online 29 November 2019

to as drones [9–11]. According to both the specific peculiarities and performance requirements of the considered scenarios, UAVs can be organized in fleet. In this way, they can be used as wireless relays to create flying wireless communication platforms aiming to achieve connectivity between ground wireless devices [12], or act as mobile aerial base stations to provide reliable downlink and uplink communications for ground users and boost the capacity of wireless communications networks (WCN) [13–16].

The previous scenarios ask for persistent mission without any human intervention. Consequently, on-field autonomous power systems for charging drones' batteries are necessary [17–19]. Different strategies could be adopted in the charge station: charging batteries left from drones that have substituted them with previously charged ones [20], or directly charging on-board the drones' battery. The first strategy presents the advantage of reducing the drone parking time, provided that a full battery is available when it lands on the station. On the other hand, this would require standardizing the battery types, the communications protocol and so on, with the additional difficulty of ensuring hot-battery swapping. Moreover, the change may be difficult due to hostile environmental conditions. On the contrary, on-board battery charge enables to overcome some of these problems at cost of a longer parking period. The battery can be charged either through contactors [21,22], or by a wireless power transfer (WPT) converter [23–28]. The former strategy presents the disadvantage of contactor corrosions, while the inductive WPT [26–28] represents the best contactless charging strategy because it avoids the safety issues arising in case of capacitive WPT [25].

Another important peculiarity in the considered scenarios is the lack of the main electrical network, which implies the use of on field generation. To this purpose, using conventional generators (e.g. diesel) would be the best solution from the "island" adequacy point of view, but this would lead to environmental issues [29]. In this paper, an IIoT framework constituted by a huge number of sensors and actuators for application scenarios concerning Industry 4.0 has been considered. They are located in a large area where telecommunication infrastructure is not available, and a Flying Ad-hoc NETwork (FANET) made of a fleet of drones realizes the connectivity platform [30]. A charge station adopting WPT for recharging drones' batteries and supplied by a wind generator (WG) has been considered. A System Controller (SC) decides how many fully-charged drones are enabled to fly in order to satisfy the time-variant bandwidth demand of the IoT devices while accounting for the time-variant power availability from the WG. A constant charge time of drone battery has been considered because this enables to strongly reduce the complexity of the SC task, so making feasible the proposed management approach. This has been obtained by implementing a current tuning mechanism on each pad able to compensate for the coupling factor variability due to different misalignment occurrence. The SC also manages the pads to properly implement the current tuning mechanism. Limited economical investments are required by using only one inverter to supply all the pads. Therefore, the adaptive power control has to be designed properly for enabling the battery charging in a fixed time chosen a priori for any drone and whatever the pad of the charge station. A current tuning mechanism is proposed in this work to achieve this objective. Finally, other UAV applications may be supported by a similar system [31,32].

The main contributions of this paper can be synthesized as follows. First, the novelty of the whole drone-based IIoT system framework, a platform for IoT-enabled devices exploiting the connectivity provided by a fleet of drones. Second, the application of Reinforcement Learning (RL) to optimally manage the use of the fleet considering the variability of both the connectivity demand and the wind energy availability. Third, the design of the aforesaid single-inverter multi-pad charge station using resonant inductive WPT in adaptive converters. The paper is structured as follows. Section 2 describes the system under study focusing on targets and constraints. Section 3 presents the system modeling and controlling strategy. The single-inverter multi-pad station

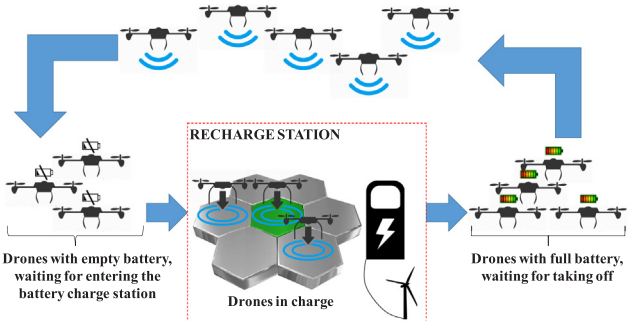

Fig. 1. Drone lifecycle phases in the system under study.

is described in Section 4. The results of the system optimal management as well as the simulation of the single-inverter-multi-pad charge station are reported in Section 5.

2. System targets, features and operation challenges

A FANET realized by a fleet of drones provides connectivity to a huge number of IoT devices distributed on a territory that is subdivided in adjacent zones. Each zone is served by a WPT station supplied by an on-field WG. Drones of the fleet periodically land in the station for charging. In the following, the model of one zone of the system (sketched in Fig. 1) is described. It is worth to note that the model is valid for each zone but the optimal management policy of each zone may differ from the others even when there is a small difference on the variability of the connectivity request, on the availability of power from the on-field WG, and on the number of drones and pads.

The drones serving a given zone are assumed to be equal to each other. Moreover, it is assumed that each drone provides the same connectivity contribution to the IoT devices deployed on the ground. Since the IoT devices operating in that zone need a time-variant bandwidth, then the number of drones necessary to guarantee the requested connectivity changes over the time. A SC decides how many drones have to take off among the ones fully charged, accounting for the required bandwidth trend and the variability of the available power from the WG.

The management challenge arises from the trade-off between the connectivity request occurring at the instant the choice has to be performed and the future requests. The problem difficulty is increased from the stochastic availability of drones with a charged battery due to the variability of wind power production [33]. RL is a well-suited approach for this kind of problems [34]. This is the reason why we adopted it as the core of the SC in order to select the optimal management policy from a long-term point of view. An appropriate model of the system elements is the key for a successful use of the RL because a very complex model could prevent RL to converge to a solution in a limited time. As described in the next section, the RL problem is formalized by using a Markov decision process (MDP) framework. At the planning stage of the SC, the best policy is found by means of RL in off-line mode by modeling the behavior of the system components and their interactions (historical or forecasted data) by using a MDP [35]. The information about the states of the discrete model with the RL choice for each state will be stored in the SC. At operation stage, the SC works online: it observes system conditions (power available from the WG, connectivity request and so on), then identifies the most similar state of the discrete model and, finally, commands some drones to take-off. The number of drones that receive the command is found by RL at planning stage as the optimal policy for the specific state identified by the SC. Notwithstanding, SC could also apply on-line RL starting from the results obtained in off-line mode.

Although the MDP is subject to the "curse of dimensionality" problem [36], assuming a fixed charge time of drone battery allows the use

of a discrete-time MDP, thus strongly reducing the complexity of the model without affecting its exactness. Therefore, a fixed charge time is chosen according to the electric characteristics of the station and drones. To this purpose, the pads of the charge station adaptively change the absorbed power from the WG to meet this time constraint regardless the coupling factor value that depends on the different misalignments between the pad and the drone landed on it.

A Lithium Ion battery (LIB) is the best candidate for the considered application since it provides high energy density and light weight [37]. Typically, LIB charging is initially performed at constant current and increasing voltage. When the voltage reaches the nominal value, it is kept constant and the current rapidly decreases. Consequently, the characteristic of the absorbed power changes during the charge interval. On the other hand, the average value of this characteristic, P_{AVG}, must be fixed and assigned a priori to ensure the aforesaid target, i.e. a constant charge interval. To reach this target the average charge power must be independent from the different coupling between drone and pad because it must be equal to P_{AVG}. The use of adaptive absorption at the pad stage instead of at the inverter stage enables to meet the constant-charge-interval constraint with the use of only one inverter. The single-inverter multi-pad station will be described in Section 4.

3. System modeling and controlling strategy

As said in the previous section, the management role of the considered framework is played by the SC. In Section 3.1 we will define a Markov model of the whole system. This model will be used in Section 3.2 to introduce the optimal management policy that is based on a RL approach. Table 1 summarizes the notation adopted in this section,

3.1. Model of the Markov Decision Process (MDP)

Let the battery *charge time*, Δ_{EF}, i.e. the period needed to change the battery status from Empty to Full, be constant thanks to the technique presented in Section 4.2. Let Δ_{FE} be the mean flight time that includes take-off, climb, cruise, descendent and landing. This interval will be also referred to as the *discharging time*. As discussed so far, in order to reduce the complexity of the model to be applicable by the SC, we use a discrete-time MDP. The charging time, Δ_{EF}, is assumed as the system time slot, Δ, while Δ_{FE} is assumed, on average, equal to $H{\cdot}\Delta$, with H a positive integer.

Table 1
Notation of the Markov Decision Process.

Symbol	Meaning
Δ_{EF}	Period needed to change the battery
Δ_{FE}	Mean flight time
$B(n)$	time-variant bandwidth requested by the IoT devices
$G(n)$	number of drones that can be concurrently supplied for battery charging
$\underline{S}^{(D)}(n)$	Battery state of the drones in the three states E, F and A
$S^{(B)}(n)$	Underlying Markov chain of $B(n)$
$S^{(G)}(n)$	Underlying Markov chain of $G(n)$
$S^{(\Sigma)}(n)$	State of the whole system
$\mathfrak{I}^{(B)}$, $\mathfrak{I}^{(G)}$, $\mathfrak{I}^{(D)}$	State space of the processes $S^{(B)}(n)$, $S^{(G)}(n)$ and$S^{(D)}(n)$
$\mathfrak{I}^{(\Sigma)}$	State space of the whole system
$\mathfrak{I}^{(A)}(n)$	Action space
N	Number of drones in the fleet
$f_R(d)$	Reward function
$P^{(B)}$, $P^{(G)}$, $P^{(D)}(s'_G,\ a)P^{(\Sigma)}(a)$	Transition probability matrix of the processes $S^{(B)}(n)$, $S^{(G)}(n)$ and$S^{(D)}(n)$
$r^{(\Sigma)}_{[s'_\Sigma,s'_\Sigma]}(a)$	Expected immediate reward
$R(n)$	Reward achieved at the slot n
$\widetilde{v}(s'_\Sigma)$	State-value function
$\widetilde{v}_{\rho^*}(s_\Sigma)$	Optimal state-value function
γ	Discount factor

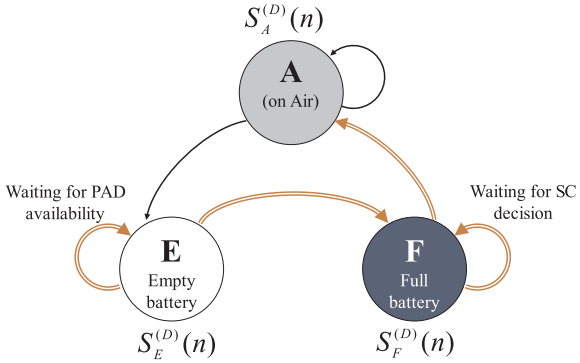

Fig. 2. Drone battery states.

The time-variant bandwidth requested by the IoT devices, indicated as $B(n)$, is assumed to be an independent stochastic process. Instead, the number of drones, $G(n)$, that can be concurrently supplied for battery charging depends on the WG. A *first in, first out* (FIFO) policy is adopted for choosing the drones to be charged.

The drone behavior can be modeled by a 3-state Markov model, as depicted in Fig. 2, where some state transitions, highlighted with '⇒', depend on external conditions, as explained below. When a flying drone on air (state A) reaches a low state of charge (SOC), it lands on a pad of the charge station and waits for charge. This is represented by the permanence on the state "Empty Battery" (E), where it remains until the pad is not supplied. When the pad is supplied, in the successive slot the drone enters the state "Full Battery" (F). The drone will stay in this state until the SC does not authorize its take off, thus the cycle restarts. Therefore, the transitions $E \Rightarrow E$ and $E \Rightarrow F$ depend on the available power from the WG, while the transitions $F \Rightarrow F$ and $F \Rightarrow A$ depend on the SC decision.

The processes $G(n)$ and $B(n)$ are modeled with Switched Batch Bernoulli Processes (SBBP) [38], which are the most general event-occurrence processes whose behavior is modulated by Markov chains in the discrete-time domain. The system state at the generic slot n is represented as:

$$S^{(\Sigma)}(n) = (S^{(G)}(n),\ S^{(B)}(n),\ \underline{S}^{(D)}(n)) \tag{1}$$

where $S^{(G)}(n)$ and $S^{(B)}(n)$ are the underlying Markov chains of the SBBP processes $G(n)$ and $B(n)$, respectively, whereas $\underline{S}^{(D)}(n)$ represents the battery state of all the drones. It is defined as an array of three Markov chains representing the number of drones in each of the three states described in Fig. 2, that is:

$$\underline{S}^{(D)}(n) = [S_E^{(D)}(n),\ S_F^{(D)}(n),\ S_A^{(D)}(n)] \tag{2}$$

As described in [38], the processes $G(n)$ and $B(n)$ are characterized by the transition probability matrix of their modulating Markov chains, and by the matrix containing, on its rows, the event-occurrence probability distribution array for each state of the Markov chain. Let $\mathfrak{I}^{(G)}$ and $\mathfrak{I}^{(B)}$ be the state space of the processes $S^{(G)}(n)$ and $S^{(B)}(n)$, respectively. Moreover, if we indicate the number of drones in the fleet as N, the state space of the processes $S_E^{(D)}(n)$, $S_F^{(D)}(n)$ and $S_A^{(D)}(n)$ is $[0, N]$, and

$$S_E^{(D)}(n) + S_F^{(D)}(n) + S_A^{(D)}(n) = N \tag{3}$$

Therefore, the state space is:

$$\mathfrak{I}^{(\Sigma)} = \mathfrak{I}^{(G)} \times \mathfrak{I}^{(B)} \times \mathfrak{I}^{(D)} \tag{4}$$

where

$$\Im^{(D)} = \left\{ (j_1, j_2, j_3): \begin{cases} j_1, j_2, j_3 \in [0, N] \\ j_1 + j_2 + j_3 = N \end{cases} \right\} \tag{5}$$

The action space is state-dependent since it is defined as $\Im^{(A)}(n) = [0, S_F^{(D)}(n)]$. Its upper limit is given by the state variable $S_F^{(D)}(n)$, representing the maximum number of drones that the SC can select for taking off.

Now we can define the transition probability matrix for the Markov chain $S^{(\Sigma)}(n)$. To this purpose, let us indicate the generic state at the slot n as $s'_\Sigma = S^{(\Sigma)}(n) = (s'_G, s'_B, \underline{s}'_D)$, and the state at the slot $n+1$ as $s_\Sigma^{''} = S^{(\Sigma)}(n+1) = (s^{''}_G, s^{''}_B, \underline{s}^{''}_D)$. The elements $\underline{s}'_D$ and $\underline{s}^{''}_D$ are as defined so far, that is:

$$\underline{s}'_D = S^{(D)}(n) = [s'_E, s'_F, s'_A]$$

$$\underline{s}^{''}_D = S^{(D)}(n+1) = [s^{''}_E, s^{''}_F, s^{''}_A]$$

The system state evolution depends on the bandwidth requested by the IoT devices and the power availability from the WG, as well as on the decisions taken by the SC, which applies RL to choose how many drones have to take off among the ones on ground with full battery. This decision is taken adopting a *policy* that maximizes the mean reward based on a reward function $f_R(d)$. This function, which is an input of the problem, associates a reward to each 2-uple constituted by the current system state, $S^{(\Sigma)}(n)$, and the action a chosen for that state by the considered policy. More specifically, we consider $f_R(d)$ as a function of the variable d defined as the difference between the number of drones on air, $S_A^{(D)}(n) + a$, and the mean value of the required bandwidth by the IoT devices:

$$d = S_A^{(D)}(n) + a - b \tag{6}$$

where $a = A(n)$ is the number of drones that take off due to the decision taken by the SC at the slot n, and b is defined as:

$$b = E\{B(n) |\, S^{(B)}(n)\} \tag{7}$$

$\mathrm{E}\{\cdot\backslash\}$ being the expected-value operator.

The reward function has to be chosen according to the target of the system to be optimized, then $f_R(d)$ is defined in the case study, more specifically in (21).

The generic element of the transition probability matrix for a given action a can be defined as follows:

$$P^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a) = P^{(G)}_{[s'_G, s^{''}_G]} \cdot P^{(B)}_{[s'_B, s^{''}_B]} \cdot P^{(D)}_{[\underline{s}'_D, \underline{s}^{''}_D]}(s'_G, a) \tag{8}$$

$P^{(G)}_{[s'_G, s^{''}_G]}$, $P^{(B)}_{[s'_B, s^{''}_B]}$ and $P^{(D)}_{[\underline{s}'_D, \underline{s}^{''}_D]}(s'_G, a)$ being the transition probabilities of the processes $S^{(G)}(n)$, $S^{(B)}(n)$ and $\underline{S}^{(D)}(n)$, respectively. The first two matrices are known as input, while the last one depends on s'_G, i.e. the state of the WG-based charge station, and a, i.e. the number of drones decided by the SC to take off. It is derived in the Appendix A.

3.2. Reinforcement learning based optimal management

The optimal policy is obtained by means of the maximization of the cumulative reward based on the reward function $f_R(d)$ introduced in the previous section. To this end, in this paper we use a RL approach in offline mode, by solving a system of equations, called Bellman optimality equations [39], derived by the Markov model defined so far.

To this purpose, we define the *expected immediate reward* as follows:

$$r^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a) = \mathrm{E}\left\{ R(n+1) \,\middle|\, \begin{matrix} S^{(\Sigma)}(n+1) = s^{''}_\Sigma, S^{(\Sigma)}(n) = s'_\Sigma, \\ A(n+1) = a \end{matrix} \right\} \tag{9}$$

where $R(n+1)$ is the reward achieved through the function $f_R(d)$, and corresponding to the transition from the state s'_Σ at the slot n to the state $s_\Sigma^{''}$ at the slot $n+1$, when the action a is performed by the SC.

The optimal policy for the system is achieved by solving the linear equation system whose generic equation is the *Bellman optimality equation* that allows evaluating the *optimal state-value function* $\widetilde{v}(s'_\Sigma)$ in the generic state s'_Σ as follows [39]:

$$\widetilde{v}(s'_\Sigma) = \max_{a \in \Im^{(A)}_{s'_\Sigma}} \left(\sum_{s^{''}_\Sigma \in \Im^{(\Sigma)}} P^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a) \cdot [r^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a) + \gamma\, \widetilde{v}(s'_\Sigma)] \right) \tag{10}$$

where $P^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a)$ is the generic element of the transition probability matrix defined in (8) from the state s'_Σ to the state $s_\Sigma^{''}$, while $r^{(\Sigma)}_{[s'_\Sigma, s^{''}_\Sigma]}(a)$ is the *expected immediate reward* defined in (9). The term $\gamma \in [0, 1[$ is the so-called discount-rate parameter, used in the RL to define how far ahead in time the algorithm looks.

Actually, (10) is a system of equations, one for each state. The solution of this system gives the optimal state-value function $\widetilde{v}_{\rho^*}(s_\Sigma)$, for each state $s_\Sigma \in \Im^{(\Sigma)}$. Once the optimal state-value function $\widetilde{v}_{\rho^*}(s_\Sigma)$ is known for all the states, we have derived the optimal policy by using asynchronous dynamic programming [39]. For each state s_Σ there will be one or more actions at which the maximum is obtained in the Bellman optimality equation. Any policy that assigns non-zero probability only to these actions is an optimal policy. However, let us note that, since $\widetilde{v}_{\rho^*}(s_\Sigma)$ already takes into account the reward consequences of all possible future behavior, by means of $\widetilde{v}_{\rho^*}(s_\Sigma)$, the optimal expected long-term return is turned into a quantity that is locally and immediately available for each state. Hence, a one-step-ahead search yields the (discounted) long-term optimal actions.

Let us note that it is not possible to express the best policy in closed form. Indeed, it is constituted by the set of the best actions, i.e. the number of charged drones that can take off among the ones already charged, for all the system states in the state space, and this is evaluated numerically. Instead, as anticipated in the previous section, we can define the action space $\Im^{(A)}(n)$. It is the set of all possible actions that the SC can perform, and depends on the current system state because it is given by the interval $[0, S_F^{(D)}(n)]$.

4. WPT system enabling drone-based IIoT

The use of RL for optimally facing the management issues arising from the considered emerging application has the additional advantage of avoiding any human intervention. Therefore, an autonomous charge station is also necessary in view of a persistent mission without any human intervention [17–19]. In this perspective, as discussed in the Introduction section, two strategies can be adopted: substitute the empty battery with one previously charged [20], or directly charge the drone battery. The latter solution is the most popular and can be performed either through contactors [21,22], or using different WPT solutions: beam-based WPT [23,24], capacitive [25] and inductive WPT [26–28].

An overview of the charge station based on resonant inductive WPT is presented in Section 4.1. Then, Section 4.2 describes the design of the single-inverter multi-pad configuration, with the theoretical bases of the adaptive technique enabling constant charging time.

4.1. Charge station based on resonant inductive WPT

The considered charge station is supplied by a variable-speed wind turbine configuration with full-scale power converters [40]. The WG supplies the pads that, in turns, charge the drones according to the current sharing policy described in the next section. The block diagram of the overall system is drawn in Fig. 3, where the area delimited by a dotted line represents one pad, while the area enclosed in a dashed rectangle refers to one drone.

Fig. 4a details the electrical circuit of a generic pad structure where the coupled coils are partially represented by both their primary circuit self-inductance, L_{T1}, and winding resistance, R_{T1}. The MOSFET connected to the input terminals of the pad, named M_1, is switched on by the SC when it desires to charge the drone located on the pad. On the

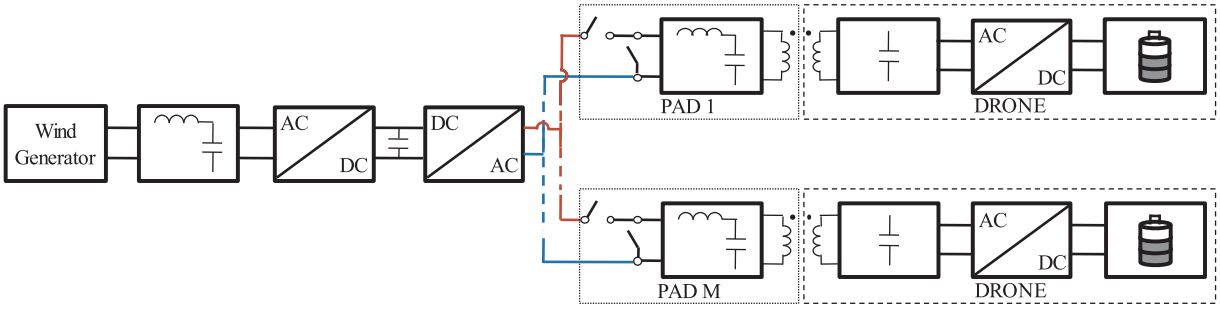

Fig. 3. Charge station main blocks with the representation of M pad-drone subsystems.

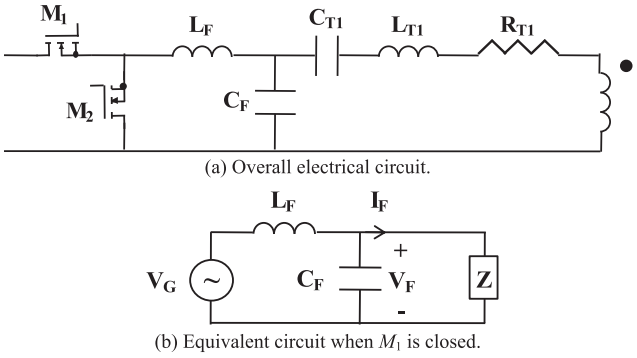

Fig. 4. Circuit model of the generic pad.

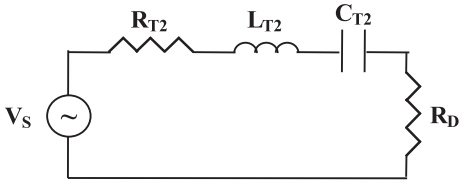

Fig. 5. Simplified circuit model of a drone.

other hand, it is switched off when the drone battery is full or when the SC can not to charge it (e.g. there is not sufficient power). Moreover, M_1 is forced to stay open when there is no any drone on the pad. The control signal on the MOSFET M_2 is complementary to the one applied to M_1 in order to short the terminals when M_1 is switched off. Finally, L_F-C_F is a filter operating at resonance, and C_{T1} is the compensation capacitor of the primary circuit self-inductance.

Without loss of generality, in the following it is assumed that each drone is charged at a constant power equal to P_{AVG}. This assumption simplifies the description of the current sharing mechanism adopted to ensure that a drone is charged in a time interval Δ regardless of the coupling factor. Notwithstanding, the idea behind this mechanism can be mixed with charging systems accounting for the actual charging profile [41]. In other words, a given power characteristic with a variable power during the charge process can be achieved by performing the tuning during the whole charge period, considering that this characteristic is fixed a priori according to the actual charging profile and the charge time interval Δ.

4.2. Single-inverter-multi-pad Current Sharing Control

Fig. 4b shows the pad when it is supplied by closing M_1 (and concurrently opening M_2), where V_G is the input voltage and Z represents the impedance "seen" downstream from the filter.

The current flowing towards Z can be expressed as:

$$I_F = \frac{V_G - V_F}{j\varpi L_F} - j\varpi C_F V_F \tag{11}$$

The current value is independent from Z at resonance [42]:

$$I_{FR} = -j\frac{V_G}{\varpi_R L_F} \text{with} \varpi_R = \frac{1}{\sqrt{L_F C_F}} \tag{12}$$

Fig. 5 shows the drone model where V_S is an equivalent voltage source whose value depends only on I_{FR} and the mutual inductance. In the figure, L_{T2} and R_{T2} represent the secondary circuit self-inductance and winding resistance, respectively. Finally, C_{T2} is the compensation capacitor and R_D is the drone equivalent resistance seen at the rectifier input port.

By assuming the self-inductance of both primary and secondary circuits equal to L_T, and considering the coupling factor k, V_S can be expressed as:

$$V_s = j\varpi_R k L_T I_{FR} = k\frac{L_T}{L_F} V_G \tag{13}$$

Then, the power absorbed by the drone can be expressed as follows:

$$P_D = \frac{R_D}{(R_D + R_{T2})^2} V_S^2 = \frac{R_D}{(R_D + R_{T2})^2} k^2 \left(\frac{L_T}{L_F}\right)^2 V_G^2 \tag{14}$$

Moreover, by considering that all the pads are equivalent to each other as well as the drones, the absorbed power depends on the coupling factor only. The coupling factor changes as the landing position changes, thus the drones well aligned with the pad can be charged in a time interval smaller than the misaligned ones. The drone could be moved as proposed in [19] to avoid misalignment so also obtaining an equal charging time. However, such a solution requires the use of actuators and additional hardware in each pad and, in the considered scenario, has the additional disadvantage of consuming part of the power available from the WG.

Therefore, the use of a variable value for L_F is the solution adopted to ensure the same charging time for all drones. More specifically, given a reference value of inductance, L_{FR}, the SC increases the inductance in the pads where the drones are well aligned, while the SC reduces the inductance in case of low coupling factor. The variation of the inductance leads to a change in the current along the various pads, then the current sharing control is adopted to transmit the same power regardless of the coupling factor.

For a given probability density function (pdf) related to the values of k, the average value of the coupling factor, k_{AVG}, can be obtained. Then, the reference value of L_F is calculated from (14) as:

$$L_{FR} = \sqrt{\frac{R_D}{P_{AVG}}} \frac{1}{R_D + R_T} k_{AVG} L_T V_G \tag{15}$$

and, for a given k, the inductance L_F has to be changed according to the following expression to ensure a charging power equal to P_{AVG}:

$$L_F = \frac{k}{k_{AVG}} L_{FR} \tag{16}$$

The fundamentals and basic operating principles of some variable inductors have been reviewed in [43]. Some methods specifically designed for WPT systems have been also proposed in the last years

[44,45]. Finally, a tunable capacitor [46] is also necessary, thus the capacitance can be modified in order to maintain the resonance condition.

Eq. (16) has the merit to highlight that the filter inductance has to be modified proportionally as the coupling factor changes. On the other hand, a relation providing the criterion to change the inductance according to the measure of the absorbed power, P_D, is more effective. Therefore, from an operating point of view, the following expression, derived from (14), is considered:

$$L_F = \sqrt{\frac{P_D}{P_{AVG}}} L_{FR} \tag{17}$$

Therefore, the SC increases the inductance at the pad terminals to reduce the current and, consequently, the transmitted power when the one measured towards the drone is greater than the target value (i.e. $P_D > P_{AVG}$ since $k > k_{AVG}$). Otherwise, when the absorbed power is too low (i.e. $P_D < P_{AVG}$ since $k < k_{AVG}$), the inductance is reduced to increase the current in the primary circuit in order to transmit a greater power to the secondary one. It is worth to note that this mechanism can be easily readapted when the actual charging profile is considered instead of using the average charging power P_{AVG}. Finally, it is to be noticed that the implementation of the proposed current sharing mechanism requires a wireless communication system between the pad and the drone [47] or the adoption of techniques not requiring any communication neither information about the coupling [48], but this is out of the scope of this paper.

5. System simulation

In this section, firstly, the ability of RL in optimally managing the use of the fleet is investigated in different cases by means of MATLAB® simulations. Proved the benefits deriving from the RL application, the ability of the power system to charge in a fixed time whatever the drone pad misalignment is shown through simulations performed in the Simscape Power Systems™ module of Simulink®.

5.1. Drone fleet optimal management: some case studies

The reference case considers a zone served by a charge station with M=30 pads, and covered by a fleet of $N = 50$ technologically advanced drones, with a battery charge time of 20 min and an average flight autonomy of 1 h. Accordingly, we considered a slot duration $\Delta = 20$ min. Moreover, it has been assumed that, on average, 300 W from the WG are necessary to charge a drone, and that a drone provides a bandwidth of 10 Mbit/s. The requested connectivity has been modeled with a 2-state SBBP model, where the states B_L and B_H represent the periods with low- and high- bandwidth requests, respectively. The duration of these states are assumed geometrically distributed random variables with mean values of 5 h (15 slots) and 1 h 40 m (5 slots), respectively. The bandwidth request during B_L is assumed geometrically distributed in the interval [190,250] Mbit/s, while the bandwidth requested during the B_H state ranges in the interval [310,370] Mbit/s. As far as the WG power, we considered a real trace derived from historical log files [29], with an average generation power of 3537 W and a peak of 9000 W. Remembering that charging a drone requires a power of 300 W (on average) from the WG, we can express the power generation as a function of the number of drones that can be simultaneously charged. So, we have an average generation power of 11.79 drones (i.e. 3537/300) and a peak of 30 drones (i.e. 9000/300), equal to 39.3% and 100% of the pads in the reference case, respectively. We modeled the generation process with a 4-state SBBP obtained by applying the inverse eigenvalue problem [49] to the real trace. The four states represent four different power availability levels that, expressed in number of drones that can be charged simultaneously, are: [0,10], [8,18], [16,26] and [22,30].

Let us note that, according to (1) and (2), the state space cardinality, i.e. the total number of states of the Markov model of the system is:

$$U_\Sigma = U_G \cdot U_B \cdot U_D \tag{18}$$

where U_G and U_B are the number of states of the underlying Markov chains of SBBP processes $G(n)$ and $B(n)$, respectively, while U_D is the number of battery states of all the drones. From (2), and considering that each of the three components of $\underline{S}^{(D)}(n)$ ranges between 0 and N, and that their sum is equal to N, it is easy to demonstrate that:

$$U_D = \sum_{k=0}^{k_{\max}} (-1)^k \binom{3}{k} \binom{N+2-(N+1)\cdot k}{N-(N+1)\cdot k} \tag{19}$$

with $k_{\max} = \lfloor N/(N+1) \rfloor$, where $\lfloor x \rfloor$ is the floor operator, defined as the largest integer less than or equal to x. In our case, the cardinality of the state space is:

$$U_\Sigma = 4 \cdot 2 \cdot 1326 = 10.608 \tag{20}$$

Actually, let us note that, after the optimization process obtained by RL, some states result not reachable due to the chosen action set. More specifically, in our case the state space is reduced to 8.008 states.

In the following, we will carry out two kinds of analysis: the first, presented in Section 5.1.1, is obtained from the reference case described above, by varying the range of B_H. This will be done by indicating this range as $[\mu_{B,H} \pm 30]$ Mbit/s, and varying $\mu_{B,H}$ between 250 Mbit/s and 400 Mbit/s. Therefore $\mu_{B,H}$=340 Mbit/s gives the reference case. The second analysis, presented in Section 5.1.2, is done by varying the number of pads in the charge station between 5 and 50.

In order to evaluate the gain obtained by applying the optimization management achieved by using the RL-based policy, in the following referred to as *best policy* (BP), we compare it with the case when *all-charged drones take off*, in the following indicated as ACTOP. For the BP case, we considered $\gamma = 0.95$ as the discount factor, and we assumed the following reward function:

$$f_R(d) = \begin{cases} -d^2 & \text{if } d \leqslant 0 \\ \sqrt{d} & \text{if } d > 0 \end{cases} \tag{21}$$

where d has been defined in (6).

For the reference case, Fig. 6 reports in red the number of drones available for taking off at each system state, $S_F^{(D)}(n)$, that it is equal to the ones taking off when the SC operates according to the ACTOP. For each state is also reported in blue the number of drones that the BP chooses to take off. The state are firstly sorted in ascending order according to $S_F^{(D)}(n)$, and then sorted in ascending order according to the number of drones related to the BP. From the figure, it is evident that the BP acts similar to ACTOP in the state where the number of available drones is small (bottom-left in the figure). On the other hand, it is interesting to note that, in some of these states, all drones in the station are left there, even when they are ready for takeoff being their battery completely charged.

The BP acts differently from ACTOP in the state where the number of available drones is wide (up-right in the figure). In fact, the number of drones that the BP keep at the station increases as the available number of drones increases.

5.1.1. System performance versus the high-bandwidth request

Fig. 7 shows the mean reward obtained against the mean bandwidth request, $\mu_{B,H}$, during high-bandwidth request periods. The analysis has been carried out for the system described before as the reference case, where the charge station is supplied by a WG. Moreover, the case in which the charge station is supplied by a conventional generator (CG) with a constant power output, e.g. diesel, has been also considered. More specifically, we have analyzed a subcase where the CG power output is equal to the mean power (CGmean) of the WG, and another one where it is equal to the WG maximum power (CGpeak). The results show that the BP always improves the SC performance with respect to

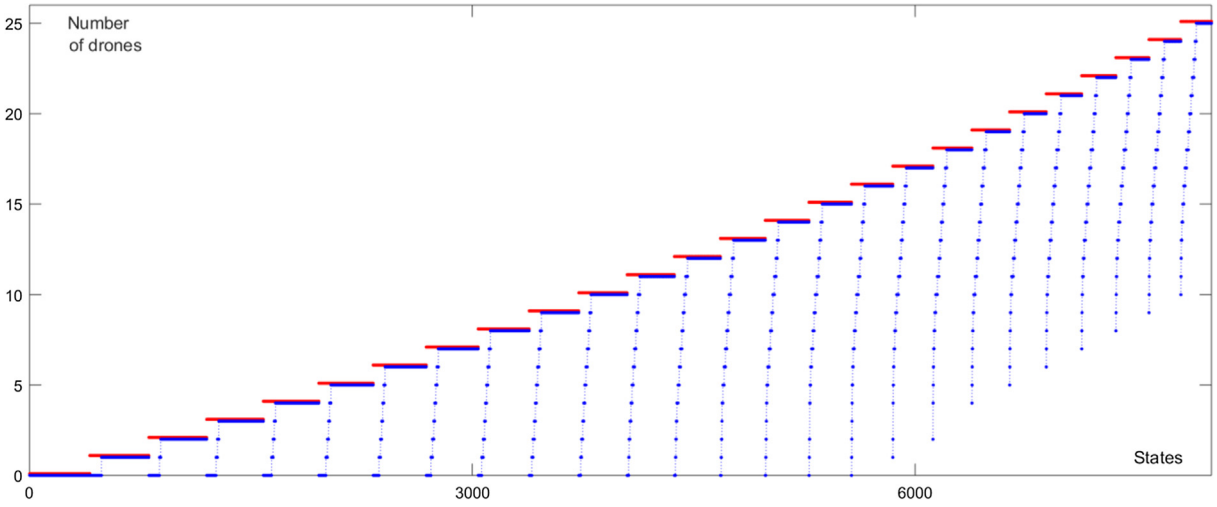

Fig. 6. Number of drone available to take off (red) and selected by BP to take off (blue).

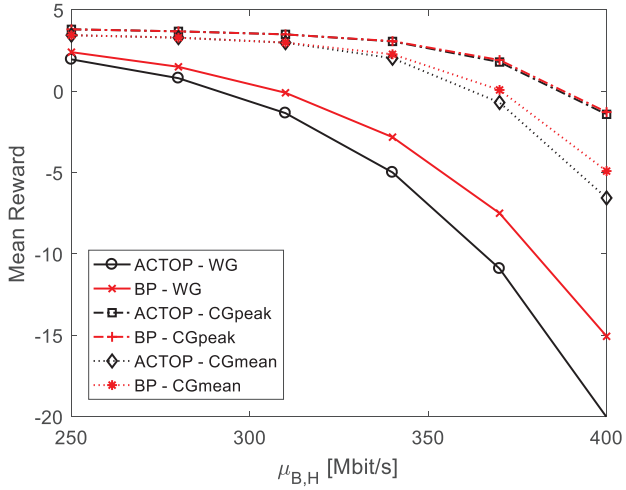

Fig. 7. Mean reward versus the mean high-bandwidth request.

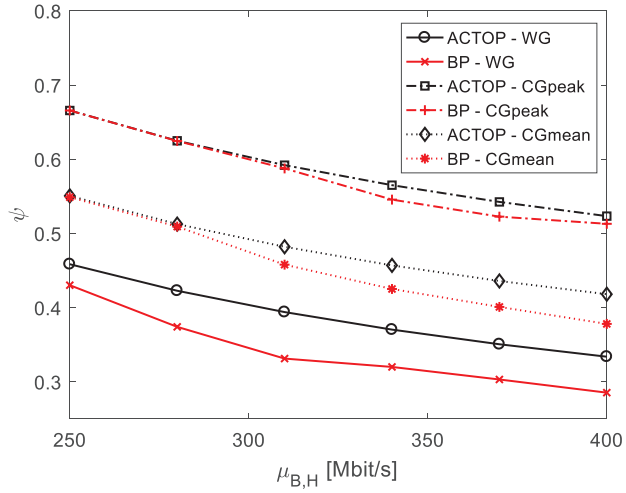

Fig. 8. Normalized bandwidth surplus.

the ACTOP, and the impact of using the RL is more relevant when the WG is considered. In other words, the RL is very suitable in mitigating the variability of WG power output accounting also for the bandwidth request variability. Finally, as expected, the use of conventional generators increases the reward, but we have to take into account that this is obtained at cost of environmental pollution [50]. Fig. 8 presents the average normalized bandwidth surplus, ψ. It is defined as the expected value of the difference between the offered bandwidth and the required bandwidth, normalized with the latter. It can be derived as follows:

$$\psi = E\left\{\frac{B_{OFF}(n) - B_{REQ}(n)}{B_{REQ}(n)}\right\} = \\ = \sum_{\forall s'_{\Sigma}} \pi^{(\Sigma)}_{[s'_{\Sigma}]} \cdot \sum_{\forall b} \frac{s'_A + a - b}{b} \cdot B^{(B)}_{[s'_B, b]} \tag{22}$$

In all the cases under analysis, the ACTOP provides a greater surplus. This result is strictly related to the average bandwidth provided by the SC: when using ACTOP, a higher value is obtained since all available drones take off regardless they are necessary or not. On the other hand, the SC using the BP leaves some drones at the station when they are not necessary on air. Therefore, this “store of drones” can meet future higher bandwidth requests and, consequently, the BP achieves better average reward at cost of providing lower average bandwidth. In other words, ACTOP maximizes the average bandwidth, while BP maximizes the ability of SC to meet the bandwidth request.

5.1.2. System performance versus the number of pads in the charge station

In this section, starting from the reference case, the analysis is performed against the number of pads $M \in [5–50]$. Three power supplies have been considered:

- WG50: a wind generator with a peak power that is able to charge 50 drones simultaneously; in this case, the mean power output is equal to the power necessary to charge 19.65 drones (the WG power trace has been modified proportionally to the reference case);
- WGpad: a wind generator with a peak power that is able to supply all the pads simultaneously, that is, the peak power varies with M (once again, the WG power trace of the reference case is modified accordingly);
- CG50: a conventional generator with a power able to charge always 50 drones simultaneously.

As in the previous section, we compare the BP against the ACTOP. Fig. 9 presents the obtained statistics regarding the rewards. More specifically, Fig. 9.a reports the mean reward obtained by the BP and ACTOP in the aforesaid three cases, while Fig. 9b shows the improvement, in terms of mean reward increment, achieved thanks to the BP. Fig. 9a reveals that there is not any improvement in the SC performance when a charge station with more than 12–13 pads is adopted in case of oversized WG (i.e. WG50) and conventional generator (i.e. CG50). In other words, from the point of view of the charge station design, 12–13

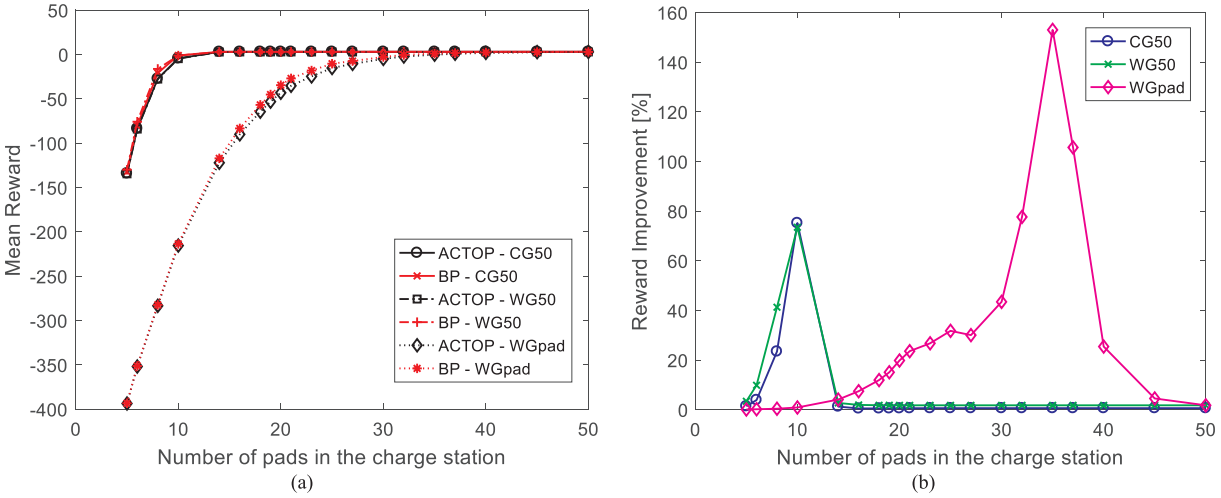

Fig. 9. Reward statistics. (a) Mean reward for different generation units; (b) Improvement obtained thanks to the best policy indicated by RL.

pads are enough when an oversized WG or a conventional generator are adopted. This is due to the number of drones as well as to the mean time on air. In fact, the ACTOP, at each time slot, tends to equalize the number of drones landing and the ones taking off. Indicating the number of drones on air as N_A, we have that N_A/H is the expected number of landing drones, and these drones will be replaced by the same number of charged drones, N_C. Therefore, we have:

$$N = N_A + N_C = N_A + \frac{N_A}{H} \tag{23}$$

and, consequently:

$$N_C = \frac{N}{H+1} \tag{24}$$

Thus, since we have $N = 50$ and $H = 3$, during a time slot, about 3/4 of drones are on air (37–38), while the remaining 1/4 are in charge (12–13) when using ACTOP in the case a CG is adopted. Considering that 37 drones are sufficient for satisfying the maximum bandwidth request of the reference case, no more than 13 pads are necessary in case of CG. Such a consideration is confirmed by Fig. 9a. Therefore, the BP improves the performance of ACTOP only when the number of pads is less than 13 since it enables to account for the pdf of the bandwidth request (Fig. 9b).

Referring to Fig. 9a, similar considerations are valid in case of the WG50, with the difference that the BP accounts for both bandwidth request variability and (oversized) WG power output pdf. Therefore, the margin of performance improvement is maximum when the number of pads is about 10 for both CG and WG50 cases (Fig. 9b). On the other hand, when the number of pads is smaller, the BP provides slightly greater improvement in case of (oversized) WG with respect to CG, as highlighted in Fig. 9b, since the BP faces also the wind variability. When the number of pads is too small (5–6), only few drones are charged and available at each time slot, thus also the BP requires to take off all of them, that is, the BP coincides with ACTOP. When the number of pads is greater than 13, the BP does not provide any advantage in case of CG since, as highlighted before, it coincides with ACTOP. Similarly, the improvement is negligible when considering an oversized WG.

In case of WG with power output changed proportionally with the number of pads (case labeled as WGpad), similar considerations are valid for the curves in Fig. 9a. More specifically, when there are about 31–33 pads in the charge station, the SC performance converges (Fig. 9a). The similarity with the previous cases is due to the fact that 31–33 pads imply a WGpad mean power equal to the power necessary to supply about 12–13 pads.

When the number of pads is less than 15, the WGpad can simultaneously supply less than 6 pads on average. In this case, the BP coincides with ACTOP. It is worth to note that such result is similar to the one obtained when considering CG or WG50 with less than 6 pads (Fig. 9b). The improvement achieved by the BP increases with the number of pads. In fact, there is more power available from the WGpad when the number of pads increases, and the BP optimizes the management of both bandwidth request and power output variability (Fig. 9b). On the other hand, when there are many pads, WGpad tends towards WG50. Consequently, the improvement due to the BP decreases for the same reasons reported for WG50 (Fig. 9b).

When the drones to be charged are less than the ones that can be charged, the available power surplus (saved power) may be used to supply other local loads. The difference between the available power to charge drones and the actual number of drones on the ground to be charged is referred to as mean saved power (MSP). It is computed by weighing this difference on all the states of the system Markov model, and then dividing by the mean value of the available power provided by the generator, that is:

$$MSP = \frac{\sum_{s'_\Sigma} \pi^{(\Sigma)}_{[s'_\Sigma]} \sum_{s'_\Sigma} P^{(\Sigma)}_{[s'_\Sigma, s'_\Sigma]} \sum_{\forall w \in \mathfrak{I}^{(G)}} B^{(W)}_{[s'_W, w]} \cdot \max\{w - s'_E, 0\}}{\sum_{s_\Sigma} \pi^{(\Sigma)}_{[s_\Sigma]} \sum_{\forall w \in \mathfrak{I}^{(G)}} B^{(W)}_{[s_W, w]} \cdot w} \tag{25}$$

Fig. 10a shows such a quantity as a function of the number of pads. The difference in the MSP among the three cases (CG50, WG50 and WGpad) is mainly due to the different size of generators. Therefore, it is more interesting the comparison between SC policies, for each generation unit. Such a comparison is performed in Fig. 10b where the saved power gain, η, is shown:

$$\eta = MSP_{BP} - MSP_{ACTOP} \tag{26}$$

The CG can supply 50 pads but no more than 12–13 drones are charged in the station according to the previous considerations (Fig. 9a). Therefore, the difference represents the power remained available for other loads, which is about 75% of CG power output, as also confirmed by Fig. 10a in the interval [12–50]. When the number of pads is less than 12, all of them are used for charging the drones landed in the station. As the number of pads decreases, the percentage of power surplus linearly increases. When the number of pads is greater than 12, the improvement obtained thank to the BP is constant (Fig. 10b). The improvement is due to the fact that the ACTOP imposes all available drones take off regardless they are necessary or not, while the BP leaves some drones at the station when they are not necessary on air and, consequently, it saves energy. As said before, when the number of pads is less than 12, the BP tends towards ACTOP as the number of pads decreases, then the advantages obtained by RL become negligible

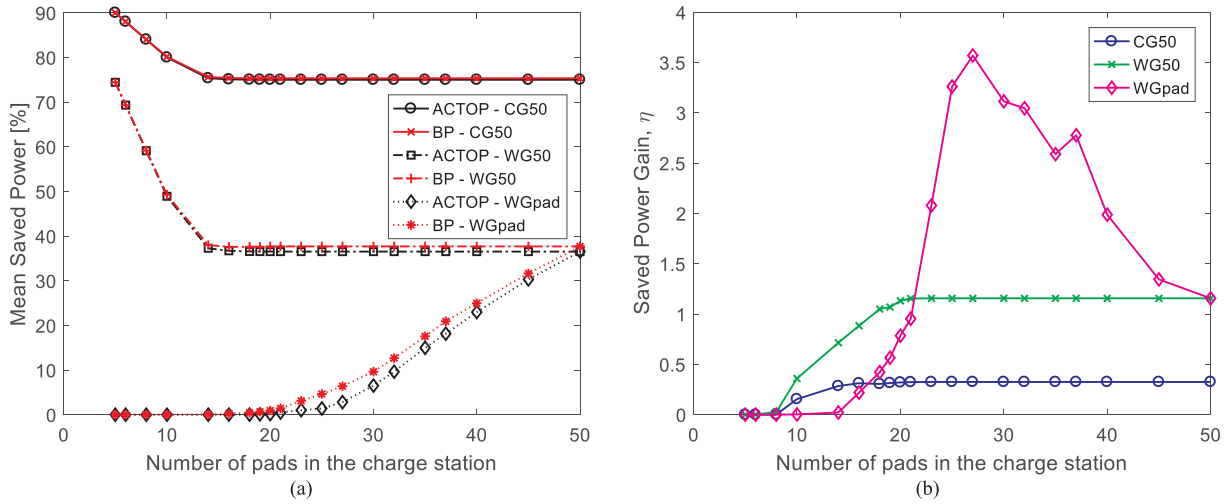

Fig. 10. Saved power: (a) Mean value for different generation units; (b) Improvement obtained thanks to the best policy indicated by RL.

(Fig. 10b). Also, when an oversized WG is considered (WG50), if the number of pads is greater than 12, MSP converges to 40% (Fig. 10a). With a reasoning similar to CG, such a value is due to the difference with the mean power output of WG50 (equivalent to 19.65 drones). Moreover, when the number of pads is less than 12, MSP increases as the number of pads decreases likewise CG50 (Fig. 10a). Recalling that the WGpad power output is proportional to the number of pads, a low power is available from the WGpad when there are few pads. Therefore, there is no saved power if there are few pads since it is used only for battery charging (Fig. 10a). For a higher number of pads, the BP enables to save more power since, as said before, the BP leaves some drones at the station when they are not necessary on air.

Therefore, such an approach has the double positive effect of having a "store of drones" for satisfying higher bandwidth request and saving energy. This fact is more evident in case of a WG designed according to the actual number of pads (Fig. 10b).

5.2. Single-inverter-multi-pad charge station

Fig. 11 shows the analyzed circuit, with evidence of the main blocks of the converter, while Table 2 reports the values of the components and the other considered quantities. According to the considered single-inverter multi-pad structure (Fig. 3), the inverter in the green box (I) in Fig. 11 is shared by all the pads, then the circuitry downstream from it represents a single pad in Fig. 11.

The voltage at its output terminals (corresponding to V_G in Fig. 4) is equal to the one of the other pads because they are connected in parallel. Consequently, a pad can be simulated independently from the others if the inverter is able to set its output voltage regardless of the load level (i.e. irrespective of the number of pads and their absorption). Then, the total current absorption can be obtained by summing the one obtained by simulating each pad. Therefore, the simulation of only one pad is sufficient to demonstrate the suitability of the adopted current tuning techniques, provided that the voltage input is independent of the load.

The coupling factor variation has been emulated by means of multiple coupling coils whose primary (and secondary) circuits are connected in series (gold box, IV, in Fig. 11). The upper coupling coil emulates the lowest value of coupling factor. Two ideal switches are

Table 2
Main quantities used for WPT system simulation.

Quantity	value	Quantity	Value
V_{DC}	48 V	a.m.r.	0.6
f_R	1 MHz	$R_{T1} - R_{T2}$	0.5 Ω
$L_{T1} - L_{T2}$	5μH	$C_{T1} - C_{T2}$	5.07 nF
k_{AVG}	0.1	k_1	0.05
k_2	0.05	k_3	0.1
L_{FR}	222 nH	C_{FR}	114 nF
$L_{F1} - L_{F2}$	444 nH	$C_{F1} - C_{F2}$	57 nF
L_{F3}	222 nH	C_{F3}	114 nF
R_B	10 Ω	C_B	1 μF

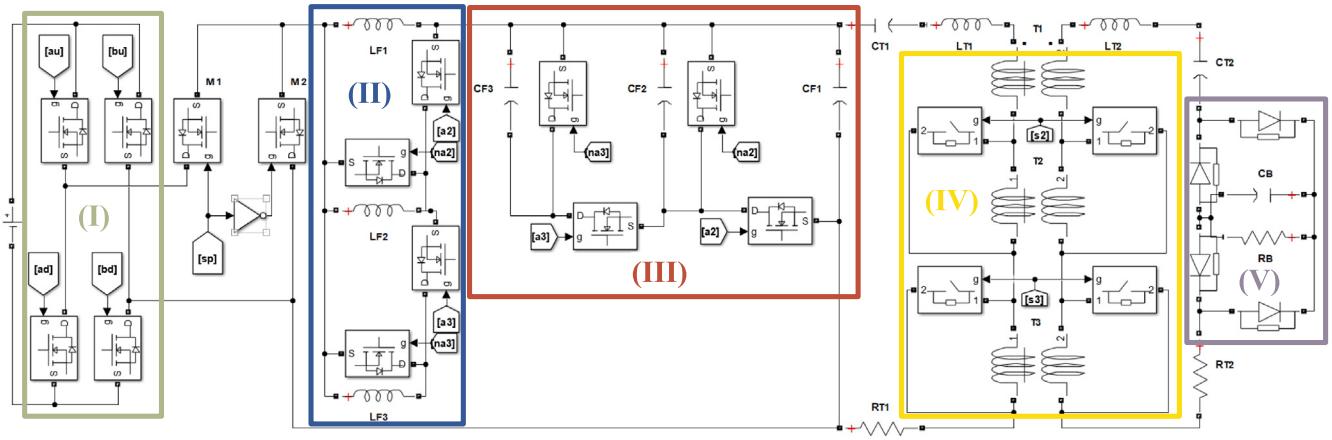

Fig. 11. Circuit simulated by Simscape Power Systems in MATLAB-Simulink. The inverter (I) is in the green box; the variable inductance (II) and capacitance (III) are in the blue and red boxes, respectively; the variable coupling (IV) is emulated in the gold box, the rectifier and the battery (V) are in the violet box.

placed in parallel with each additional coupling coils. When these switches are set to the closed status, the coupling coils are short-circuited and then they do not contribute to the overall coupling factor. Otherwise, the status of the switches is set to open when the contribution on the coupling factor of the related coupling coils has to be considered.

The variable inductance (blue box, II, in Fig. 11) and capacitance (red box, III, in Fig. 11) are obtained by parallel connection of inductors and capacitors, respectively. Two MOSFETs are associated to each additional inductor to enable its connection or disconnection from the main circuit. More specifically, one MOSFET is connected in series with the inductor and the other in parallel. The SC operates these switches by complementary control signals. When the first is turned on, the second one is turned off, then the inductor is connected to the main circuit enabling to increase the current towards the primary circuit of the coupling coils. On the other hand, when a current reduction is necessary, the MOSFET in series is turned off to disconnect the inductor, and the other MOSFET is turned on to enable inductive energy storage dissipation. A capacitor is associated to each inductor to ensure that the filter always operates at the resonance condition. Therefore, a MOSFET is connected in series with the capacitor and another one in parallel, and they are operated through the same signal controlling the MOSFETs of the related inductor. More specifically, C_{F2} and L_{F2} are both connected in parallel with, respectively, C_{F1} and L_{F1} when signal $a2 = 1$. Similarly, C_{F3} and L_{F3} are also connected when signal $a3 = 1$ (in this case it is necessary – but it is not sufficient – that $a2 = 1$).

When the pad is supplied by closing M_1, there is no knowledge of the alignment condition. Therefore, any configuration of L_F-C_F may be adopted, provided that resonance is guaranteed. The configuration L_{FR}-C_{FR} (i.e. the one to be set when a coupling factor equal to the average one occurs, hereafter called basic scenario) has been chosen as the initial one. This configuration is obtained by connecting L_{F2} and C_{F2} while leaving both L_{F3} and C_{F3} disconnected (signal $a2 = 1$ and $a3 = 0$).

Three scenarios have been emulated. The basic one has been obtained by keeping open the upper couple of ideal switches in the gold box (IV) of Fig. 11 (i.e. $s2 = 0$), while the lower ones are closed (i.e. $s3 = 1$ in Fig. 11). The scenario with a worse alignment has been emulated by keeping closed both couples of ideal switches (i.e. $s2 = s3 = 1$). When the SC detects this scenario by means of elaboration of the received information about the power delivered to the drone, it connects L_{F3} and C_{F3} to increase the current through the primary circuit, thus counterbalancing the poor alignment. The SC reaches this result by setting $a3 = 1$. The scenario with a better alignment than the basic one has been emulated by leaving open both couples of ideal switches (i.e. $s2 = s3 = 0$). Unlike the previous scenario, in this case, the SC disconnects L_{F2} and C_{F2} to reduce the current when it detects a power transfer exceeding the reference one. In this case it sets $a2 = 0$ (consequently, $a3 = 0$).

In the following, as mentioned before, it has been assumed that the battery charge is performed with a charging power equal to P_{AVG}. This is obtained by setting the voltage across the battery equal to the nominal one, V_{NOM}:

$$P_{AVG} = \frac{V_{NOM}^2}{R_B} \tag{27}$$

In this case study a nominal voltage equal to 48 V has been considered. Then, according to the quantities in Table 2 and Eq. (27), the power absorbed by the battery is about 230 W. The simulations have been initially performed considering ideal MOSFETs and ideal passive components. Therefore, the power that the WG has to provide to charge the battery of a drone is greater than 230 W considering the efficiency of the overall conversion system as well as the actual coupling. In view of an overall efficiency of about 77% for the system, the assumption of 300 W necessary on average for charging a drone is valid.

It is worth to note that the value of the DC voltage supplying the inverter is also 48 V. Therefore, given the quantities related to the coupling coils and the battery in Table 2, the values of inductance and capacitance in the filter have been chosen in order to obtain the aforesaid voltage equality.

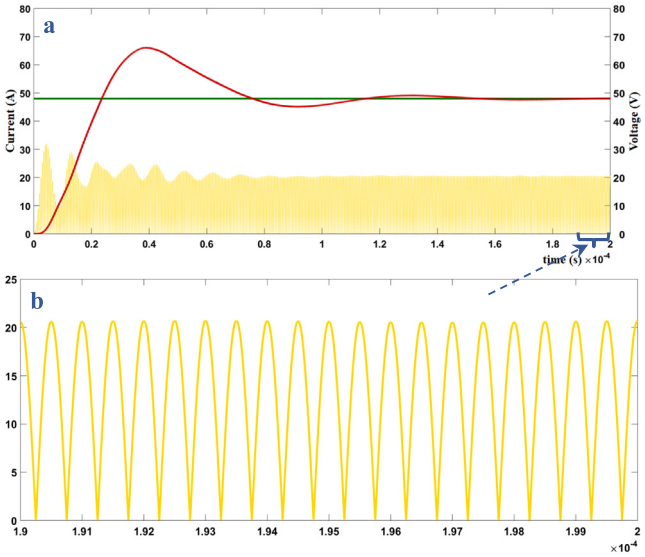

Fig. 12. Basic scenario: V_{DC} and battery nominal voltage (green), voltage across the battery (red), amplitude of the current in the primary circuit, I_F, (gold).

Figs. 12–14 show the amplitude of the current in the primary circuit of the coupling coils (gold waveform), the DC voltage at the inverter input terminals (green waveform) and the voltage across the battery (red waveform).

Fig. 12 shows the basic scenario, $k = 0.1$, where it is not necessary any change in the primary current because the coupling factor equals the average value, k_{avg}. More specifically, the voltage across the battery tends towards the desired value.

Fig. 13 shows the case with a poor alignment, $k = 0.05$. When the SC detects a voltage across the battery less than the nominal one, it reduces the inductance (and increases the capacitance accordingly) of the filter to increase the current in the primary circuit of the coupled coils to obtain the desired voltage level.

Fig. 13b and c report the amplitude of this current, respectively,

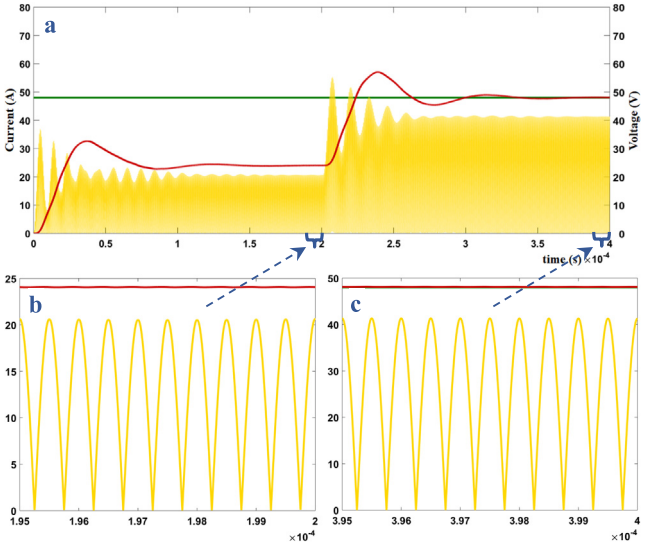

Fig. 13. Poor alignment: V_{DC} and battery nominal voltage (green), voltage across the battery (red), amplitude of the current in the primary circuit, I_F, (gold).

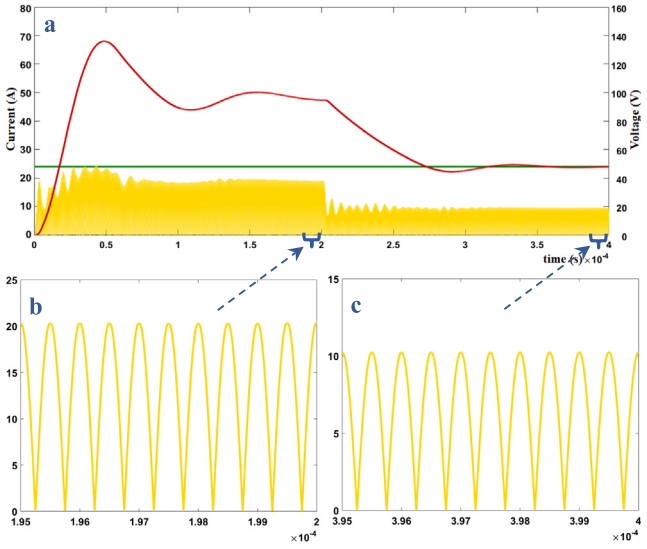

Fig. 14. Good alignment: V_{DC} and battery nominal voltage (green), voltage across the battery (red), amplitude of the current in the primary circuit, I_F, (gold).

before and after the change. By comparing Fig. 13b with Fig. 12b, it can be noted that the initial current in the primary circuit is equal to the basic scenario, thus confirming that this current does not depend on the circuit downstream from the filter. Moreover, in Fig. 13c it is shown that the current has to be increased in order to deliver the desired power. By comparing the current amplitude of Fig. 13b and c, it is also evident that the current amplitude has been doubled to keep constant the product with the coupling factor (which is half of k_{avg}).

The amplitudes obtained for the scenario with a better alignment, $k = 0.2$, are reported in Fig. 14. Once again, the results confirm the expected behavior of the charge system.

6. Conclusions

The adoption of a fleet of drones that provides connectivity to a large number of distributed IoT devices is an innovative scenario towards the Industry 4.0 framework. A charge station based on wireless power transfer is necessary to avoid any human intervention for battery charging. Moreover, the use of on field renewable generators to supply the station is recommended for green Industry 4.0 development.

In this paper, RL has been adopted in the system controller to optimally manage the fleet usage considering the variability of both the bandwidth demand and the green power availability. The results obtained in different test cases and scenarios have highlighted the effectiveness of the proposed tool. The results have highlighted that the RL provides positive effects, more evident in case of wind generators. In the framework of green power generation, the improvement obtained thanks to the RL is greater in case of generator power output proportional to the number of pads instead of oversized generation. An important aspect is the ability of RL to increase the saved energy even when it is not considered as a target of the RL.

A single-inverter-multi-pad power converter has been also adopted to guarantee a constant charge time with the goal of making the system controller task easier. The inductance and capacitance of the filter in a pad are tuned to obtain, through a wireless power transfer link, the same current towards its drone battery regardless of the misalignment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Let us derive the generic element of the matrix $P^{(D)}(s'_G,\ a)$, used in (8). From the definition of the state $\underline{S}^{(D)}(n)$ in (2), and applying the total probability theorem on both the number β of pads that can be supplied simultaneously among the M that are available in the charge station, and the number c of drones that can land in a slot because needing to be charged among the s'_A drones that are flying, we have:

$$P^{(D)}_{[\underline{s}'_D,\underline{s}''_D]}(s'_G,\ a)= = \sum_{\beta=0}^{M} B^{(G)}_{[s'_G,\beta]}\left(\sum_{c=0}^{s'_A} p^{(Land)}_{[c]}(s'_A)\cdot P^{(D_E)}_{[s'_E,s''_E]}(\beta, s'_A, c)\cdot \cdot P^{(D_F)}_{[s'_F,s''_F]}(a, \beta, s'_E)\cdot P^{(D_A)}_{[s'_A,s''_A]}(a, c)\right) \quad (28)$$

The term $B^{(G)}_{[s'_G,\beta]}$ represents the probability that β pads can be supplied when the underlying Markov chain of the WG is $S^{(G)}(n) = s'_G$. The terms $P^{(D_E)}_{[s'_E,s''_E]}(\beta, s'_A, c)$, $P^{(D_F)}_{[s'_F,s''_F]}(a, \beta, s'_E)$ and $P^{(D_A)}_{[s'_A,s''_A]}(a, c)$ are the transition probabilities of the component states of $\underline{S}^{(D)}(n)$.

Finally, the term $p^{(Land)}_{[c]}(s'_A)$ is the probability that c drones, among the s'_A ones that are flying, land in a slot because needing to be charged. In order to calculate it, since the drone battery SOCs are independent of each other, we can model the landing probability for a drone in a slot as a Bernoulli process with parameter ρ_{Down}. This parameter depends on the mean time on air. Assuming that in the slot when a drone takes off does not land, the per-slot landing probability starting from the second slot on air is $\rho_{Down} = 1/(H-1)$. Therefore, considering that s'_A represents the number of drones that are already on air at the beginning of the slot, they potentially may land with probability $p^{(Land)}_{[c]}(s'_A)$ following a Binomial distribution:

$$p^{(Land)}_{[c]}(s'_A) = \binom{s'_A}{c}\cdot(\rho_{Down})^c\cdot(1-\rho_{Down})^{s'_A-c} \quad (29)$$

References

[1] Wang S, Wan J, Zhang D, Li D, Zhang C. Towards the smart factory for industry 4.0: a self-organized multi-agent system assisted with big data based feedback and coordination Elsevier computer networks. Comput Netw 2016;101:158–68.
[2] Wang S, Wan J, Li D, Zhang C. Implementing smart factory of industry 4.0: an outlook. Int J Distrib Sens Netw 2016; 2015 art. no. 3159805.
[3] Chen F, Deng P, Wan J, Zhang D, Vasilakos A, Rong X. Data mining for the Internet of things: Literature review and challenges. Int J Distrib Sens Netw 2015; 2015. art. no. 431047.
[4] Li X, Li D, Wan J, Vasilakos A, Lai C, Wang S. A review of industrial wireless networks in the context of industry 4.0. Wireless Netw 2015:1–19.
[5] Liu Q, Wan J, Zhou K. Cloud manufacturing service system for industrial-cluster-

oriented application. J Internet Technol 2014;15(3):373–80.
[6] Gondchawar N, Kawitkar RS. IoT based Smart agriculture. Int J Adv Res Comput Commun Eng 2016;5(6).
[7] Rajalakshmi P, Mahalakshmi SD. IOT based crop-field monitoring and irrigation automation. In: 10th International conference on intelligent systems and control, 7–8 January 2016.
[8] Zeng Y, Zhang R, Lim TJ. Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun Mag 2016;54(5):36–42.
[9] Mozaffari M, Saad W, Bennis M, Debbah M. Unmanned aerial vehicle with underlaid device-to-device communications: performance and tradeoffs. IEEE Trans Wireless Commun 2016.;15(6):3949–63.
[10] Hourani A, Kandeepan S, Jamalipour A. Modeling air-to-ground path loss for low altitude platforms in urban environments. IEEE Global Communications Conference, Austin, TX, USA, December 2014.
[11] Yaliniz R, El-Keyi A, Yanikomeroglu H. Efficient 3-D placement of an aerial base station in next generation cellular networks. IEEE Int. Conference on Communications, Kuala Lumpur, Malaysia, May 2016.
[12] Dawy Z, Saad W, Ghosh A, Andrews JG, Yaacoub E. Toward massive machine type cellular communications. IEEE Wirel Commun 2017;24(1):120–8.
[13] Mozaffari M, Saad W, Bennis M, Debbah M. Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. IEEE Commun Lett 2016;20(8):1647–50.
[14] Mozaffari M, Saad W, Bennis M, Debbah M. Unmanned aerial vehicle with underlaid device-to-device communications: performance and tradeoffs. IEEE Trans Wireless Commun 2016;15(6):3949–63.
[15] Mozaffari M, Saad W, Bennis M, Debbah M. Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications. IEEE Trans Wireless Commun 2017;16(11):7574–89.
[16] Faraci G, Grasso C, Schembra G. Reinforcement-learning for management of a 5G network slice extension with UAVs. In: Proc. of IEEE SMILING 2019, Paris, France, 29 April, 2019.
[17] Di Puglia Pugliese L et al. Modelling the mobile target covering problem using flying drones. Optim Lett 2016.
[18] Junaid A, Konoiko A, Zweiri Y, Sahinkaya M, Seneviratne L. Autonomous wireless self-charging for multi-rotor unmanned aerial vehicles. Energies 2017;10(6):803.
[19] Choi CH et al. Automatic wireless drone charging station creating essential environment for continuous drone operation. Int Conf on Control Automation and Inform Sciences, Ansan; 2016. p. 132–136.
[20] Suzuki KAO, Kemper Filho P, Morrison JRJ. Automatic battery replacement system for UAVs: analysis and design. J Intell Rob Syst 2012;65:563–86.
[21] Augugliaro F, et al. The flight assembled architecture installation: cooperative construction with flying machines. IEEE Control Syst 2014;34(4):46–64.
[22] Compagnin A et al. Autoport project: a docking station for planetary exploration drones. In: 55th AIAA Aerospace Sciences Meeting, AIAA SciTech Forum, (AIAA 2017-1680), 2017.
[23] Gómez-Tornero JL, Poveda-García M, Guzmán-Quirós R, Sánchez-Arnause JC. Design of Ku-band wireless power transfer system to empower light drones. IEEE WPT Conference, Aveiro; 2016. p. 1–4.
[24] Song C, et al. Matching network elimination in broadband rectennas for high-efficiency wireless power transfer and energy harvesting. IEEE Tr Ind Electron 2017;64(5):3950–61.
[25] Mostafa TM, Muharam A, Hattori R. Wireless battery charging system for drones via capacitive power transfer. IEEE PELS Workshop on Emerging Technologies: WPT Chongqing 2017:1–6.
[26] Song C et al. Low EMF three phase resonant magnetic field charger for drone with high Q reactive loop shielding. In: IEEE Wireless Power Transfer Conference, Aveiro, 2016. p. 1–4.
[27] Jeong S, Bito J, Tentzeris MM. Design of a novel wireless power system using machine learning techniques for drone applications. IEEE Wireless Power Transfer Conference, Taipei; 2017. p. 1–4.
[28] Bi Z, Keoleian GA, Ersal T. Wireless charger deployment for an electric bus network: a multi-objective life cycle optimization. Appl Energy 2018;225:1090–101.
[29] Conti S, et al. Effect of islanding and telecontrolled switches on distribution system reliability considering load and green-energy fluctuations. Appl Sci 2016;6(5–138).
[30] Zafar W, Muhammad Khan B. Flying ad-hoc networks: technological and social implications. In: IEEE Technology and Society Magazine 35(2); 2016:67–74.
[31] Chiang W-C, Li Y, Shang J, Urban TL. Impact of drone delivery on sustainability and cost: realizing the UAV potential through vehicle routing optimization. Appl Energy 2019;242:1164–75.
[32] Raciti A, Rizzo SA, Susinni G. Drone charging stations over the buildings based on a wireless power transfer system. In: IEEE/IAS 54th Industrial and Commercial Power Systems Technical Conference (I&CPS), Niagara Falls, ON, May 7–10, 2018.
[33] Khalid M, Aguilera RP, Savkin AV, Agelidis VG. On maximizing profit of wind-battery supported power station based on wind power and energy price forecasting. Appl Energy 2018;211:764–73.
[34] Conti S, Faraci G, Nicolosi AR, Rizzo SA, Schembra G. Battery management in a green fog-computing node: a reinforcement-learning approach. IEEE Access 2017;5:21126–38.
[35] Nguyen DD, Rajagopalan A, Lim CC. Online versus offline reinforcement learning for false target control against known threat. In: Chen Z, Mendes A, Yan Y, Chen S, editors. Intelligent Robotics and Applications. ICIRA 2018. Lecture Notes in Computer Science, vol 10985. Springer.
[36] Arel I, Rose DC, Karnowski TP. Deep machine learning – a new frontier in artificial intelligence research [Research Frontier]. IEEE Comput Intell Mag 2010;5(4):13–8.
[37] Davis S. Military operations command more from mobile power sources. Power Electron Technol 2010;36(3):32–6.
[38] Lombardo A, Morabito G, Schembra G. Modeling intramedia and intermedia relationships in multimedia network analysis through multiple time-scale statistics. IEEE Trans Multimedia 2004;6(1).
[39] Sutton Richard S, Barto Andrew G. Reinforcement learning: an introduction. The MIT Press Cambridge, Massachusetts London, England; 2012.
[40] Blaabjerg F, Liserre M, Ma K. Power electronics converters for wind turbine systems. IEEE Trans Ind Appl 2012;48(2). pp. 708–719.
[41] Liu M, Zhao C, Song J, Ma C. Battery charging profile-based parameter design of a 6.78-MHz class E^2 wireless charging system. IEEE Tr Ind Electron 2017;64(8):6169–78.
[42] Wang Y, Yao Y, Liu X, Xu DG, Cai L. An LC/S compensation topology and coil design technique for wireless power transfer. In: IEEE Tr. on Ind. Electronics, doi: 10.1109/TPEL.2017.2698002.
[43] Perdigão MS, et al. A review on variable inductors and variable transformers: applications to lighting drivers. IEEE Trans Ind Appl 2016;52(1). pp. 531–547.
[44] Aldhaher S, Luk PCK, Whidborne JF. Electronic tuning of misaligned coils in wireless power transfer systems. IEEE Tr Power Electron 2014;29(11):5975–82.
[45] Lee ES, et al. Wide-range adaptive IPT using dipole-coils with a reflector by variable switched capacitance. IEEE Trans Power Electron 2017;32(10):8054–70.
[46] Lee E-S, et al. LED driver with TRIAC dimming control by variable switched capacitance for power regulation. J Power Electron 2015;15(2):555–66.
[47] Miller JM, Onar OC, Chinthavali M. Primary-side power flow control of wireless power transfer for electric vehicle charging. IEEE J Emerg Select Top Power Electron 2015;3(1):147–62.
[48] Zhong W, Hui SYR. Charging time control of wireless power transfer systems without using mutual coupling information and wireless communication system. IEEE Trans Ind Electron 2017;64(1):228–35.
[49] Lombardo A, Morabito G, Schembra G. An accurate and treatable markov model of MPEG-video traffic. In: Proc IEEE Infocom '98, San Francisco, CA, USA, March 29–April 2, 1998.
[50] Lin B, Chen Y. Impacts of policies on innovation in wind power technologies in China. Appl Energy 2019,;247. pp. 682–691, pp. 549–561.