Network Association Strategies for an Energy Harvesting Aided Super-WiFi Network Relying on Measured Solar Activity

Jingjing Wang, *Student Member, IEEE*, Chunxiao Jiang, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*, Yong Ren, *Senior Member, IEEE* and Lajos Hanzo *Fellow, IEEE*

***Abstract*—The 'super-WiFi' network concept has been proposed for nationwide Internet access in the USA. However, the traditional mains power supply is not necessarily ubiquitous in this large-scale wireless network. Furthermore, the non-uniform geographic distribution of both the based stations and of the tele-traffic requires carefully considered user-association. Relying on the rapidly developing energy harvesting techniques, we focus our attention on the sophisticated access point selection strategies conceived for the energy harvesting aided super-WiFi network. Explicitly, we propose a solar radiation model relying on the historical solar activity observation data provided by the University of Queensland, followed a beneficial radiation parameter estimation method. Furthermore, we formulate both a Markov decision process (MDP) as well as a partially observable Markov decision process (POMDP) for supporting the users' decisions on beneficially selecting access points. Moveover, we conceived iterative algorithms for implementing our MDP and POMDP based AP-selection, respectively. Finally, our performance results are benchmarked against a range of traditional decision-making algorithms.**

***Index Terms*—Super WiFi network, access point selection algorithm, energy harvesting model, MDP, POMDP.**

I. Introduction

The concept of super-WiFi network was proposed by the United States Federal Communications Commission (FCC) for the creation of ubiquitous wireless Internet access [1], which aims for exploiting the lower-frequency white spaces in the television frequency band in order to create a nationwide wireless network [2], [3]. The super-WiFi network is expected to support sophisticated new services, such as intelligent transportation systems (ITS) [4] and telemedicine systems (TELEMED) [5], whilst relying on intelligent devices and inspiring new services.

However, the ambitious task of building a nationwide network faces numerous challenges. Specifically, it is unrealistic to expect the availability of mains power for all access points (AP), especially in rural areas. The provision of power supply for the super-WiFi network is one of its gravest challenges. Furthermore, how a user wishing to establish a session makes a well-informed AP selection is another open challenge, given the non-uniform AP and user distribution [6], [7], [8]. Each AP has its well-defined, unique state, including its energy condition, user access condition, quality of service (QoS) requirement, etc. The above-mentioned heterogeneous characteristics of the APs as well as of the users bursty tele-traffic demands require new wireless access protocols specifically designed for the super-WiFi system.

Despite the above-mentioned challenges, some solutions have been found, which led to successful experimental super-WiFi developments, demonstrating that it is feasible to tackle the aforementioned problems. The energy harvesting concept provides an ideal replacement for traditional rechargeable batteries or for the mains supply by harvesting energy from the surrounding ambient sources [9], [10], such as vibration, solar energy, thermal sources, etc. These alternatives eliminate the need for a mains power connection, making the AP truly tetherless and allowing flexibility in terms their positioning. As regards to AP selection, a range of sophisticated strategies have been conceived for large-scale wireless access networks relying for example on game theory [11], [12], on price theory [13], [14], on Markov decision processes (MDP) [15], etc. Most of them aim for balancing the traffic load of the APs, as well as for providing an improved QoS [16], [17], [18].

However, the aforementioned studies on solar energy radiation models focused their attention on the theoretical models, rather than on real solar activity observation data. Moreover, the users' access strategy has been oblivious of the near-instantaneouls energy harvesting conditions. Inspired by the above-mentioned open problems, in this paper, we conceive sophisticated AP selection strategies for an energy harvesting aided super-WiFi network. Our original contributions are summarized as follows:

- The historical solar radiation observation data provided by the University of Queensland [19] is invoked for conceiving a solar radiation model as well as an algorithm for estimating the available solar radiation power.
- Moreover, we construct a super-WiFi system model relying on a queue-based user access model as well as a dynamic battery state model. The system's state transition probability matrix is formulated as well.
- Relying on our solar radiation model and system model, both a MDP as well as a partially observable Markov decision process (POMDP) based AP selection algorithm is proposed, followed by our performance analyses. Fur-

J. Wang, C. Jiang and Y. Ren are with the Department of Electronic Engineering, Tsinghua University, Beijing, 100084, People Republic of China. E-mail: chinaeephd@gmail.com, {jchx, reny}@tsinghua.edu.cn.

Z. Han is with Electrical and Computer Engineering Department as well as Computer Science Department, University of Houston, Houston, TX, USA. Email: zhan2@uh.edu.

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ U.K. Email: lh@ecs.soton.ac.uk.

thermore, we conceive iterative MDP- and POMDP-based AP-selection algorithms.

The remainder of this article is structured as follows. The state-of-the-art in energy radiation models as well as in optimal access strategies is detailed in Section II. The model of the super-WiFi network and a MDP-based AP selection algorithm are discussed in Section III. In Section IV, a measurement-based solar radiation model is proposed, including its parameter estimation method. Moreover, an iterative POMDP-based AP selection algorithm as well as its implementation are also discussed. In Section V, simulation results are provided for characterizing both our MDP and POMDP AP selection algorithms, followed by our conclusions in Section VI.

II. State-of-the-art

A variety of energy radiation models have been proposed in the literature for solar-powered wireless access networks [20]-[24]. Generally, the solar energy radiation models can be categorized as deterministic and stochastic radiation models. In [20], Tacca *et al*. proposed a cooperative data link relying on automatic repeat request (ARQ) protocols for energy harvesting sensor networks based on deterministic solar radiation models, where they assumed that the transmitters are capable of accurately predicting the solar energy arrival instants and the resultant energy. By contrast, Niyato *et al*. [21] modeled the solar radiation state transitions as a continuous-time Markov chain and considered the impact of clouds on the intensity of solar radiation in terms of the cloud sizes and wind speed. Based on the solar radiation model, they presented an analytical queuing model to investigate the QoS performance of a solar-powered wireless sensor network (WSN). Moreover, a solar energy prediction model based on an exponentially weighted moving-average filter was proposed by Kansal *et al*. [22], where the harvested energy availability varied as a function of time in a non-deterministic manner. Lei *et al*. [23] constructed a birth-and-death chain in order to characterize the solar radiation state transition process and optimized their single-hop transmission policy of replenishing the sensors. Furthermore, an independent and identically distributed (i.i.d.) Bernoulli model was invoked for characterizing the energy harvesting process by Aprem *et al*. [24], who conceived minimum-outage power control policies for ARQ-aided wireless energy harvesting sensor nodes. However, all the aforementioned studies of solar energy radiation stipulated idealized simplifying hypotheses. However, in practice modeling mismatch may occur, when the solar energy radiation conditions change during the prediction interval. Therefore, it is imperative to construct a more precise solar radiation model based on real solar activity observation data, since a more accurate energy harvesting/consumption model facilitates an improved power management and access control performance. Accordingly, the new data-driven energy harvesting model proposed in this paper relies on accurate statistical analysis of the real solar radiation patterns, which is beneficial for the design of the large-scale super-WiFi system conceived and characterized in this treatise.

Given the specific energy supply pattern of solar-powered wireless access networks, it is necessary to develop new wireless protocols. Inspired by this requirement, some preliminary studies on how to conceive optimal access strategies are summarized as follows. In [25], Todd *et al*. highlighted the current shortcomings of IEEE 802.11 when used in solar-powered wireless mesh networks, which necessitates a cost-effective sustainable energy harvesting/consumption mechanism. Moreover, in [26], Xu *et al*. summarized the existing access strategies relying on the MDP family, namely on discrete time MDP (DTMDP) [27], continuous time MDP (CTMDP) [28] [29], POMDP [30] [15], etc. and analyzed their open challenges as well as potential solutions. Furthermore, since the access to multiple base stations can be readily formulated as a game, Chen *et al*. [30] introduced the fundamental concepts of game theory, with an emphasis on wireless access networks. To elaborate, the potential benefits of games in wireless access networks were originally revealed by Neel *et al*. [31], with the objective of optimizing the attainable network performance. Yang *et al*. [32] formulated the wireless network selection problem as a stochastic game with negative network externality[1] and proposed a modified value iteration algorithm for arriving at the optimal decision. In [33], a cooperative pricing model was presented, where the cooperation may involve either all or a limited subset of the service providers. Meanwhile, Niyato *et al*. [33] proposed a scheme for sharing the revenue among the service providers in a fair manner based on an N-person coalition game. However, none of the aforementioned studies on optimal access strategy have considered the realistic energy harvesting situation, where the APs of the super-WiFi system, considered are supplied by solar power. Therefore, in this paper, we focus our attention on the influence of both the supported users' states and the access points' battery states. Both a MDP as well as a POMDP AP selection strategy is proposed for our super-WiFi system, which are based on a realistic energy harvesting model derived from measured solar activity.

III. AP Selection Strategy Based on MDP

A. System Model

In this subsection, we focus our attention on mathematically modeling the super-WiFi system. As shown in Fig. 1, the super-WiFi system contains K densely and uniformly deployed APs supplied by solar energy. The APs are connected to the servers and support the users through the wireless backhaul. Each AP has its own coverage D and broadcasts a constant beacon signal to announce its presence at regular intervals. Meanwhile, one of the users receives the short beacon message from the surrounding APs and chooses the most appropriate one to establish the data-transmission connection. In our paper, we use a MDP model to formulate the users' AP selection strategy, where we make the assumption that each AP supports a maximum of U_M users. Hence, a new access request reaching the i^{th} AP will be denied, when the

[1]Negative network externality has the opposite effect on stability compared to positive network externality. For example, it may create a negative feedback and exponential decay, which drives the network towards equilibrium, and it is responsible for maintaining stability, hence preventing the network states from tending to infinity.

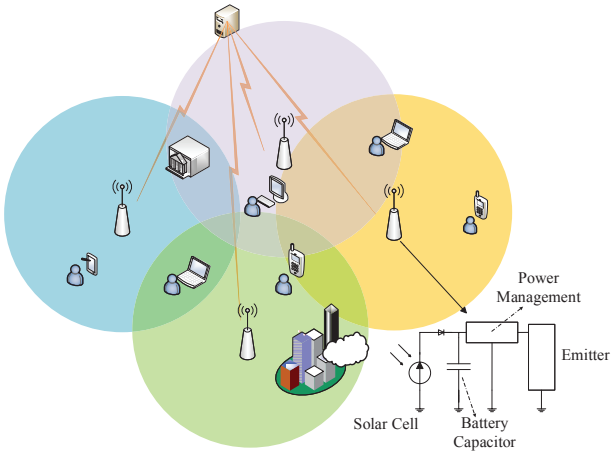


Fig. 1. The System Model of Super-WiFi Systems and the Internal Structure of Widely-used Solar Cells.

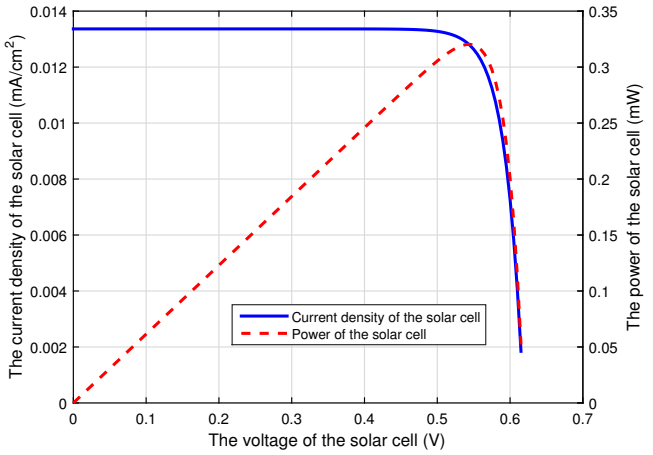


Fig. 2. The current density versus the voltage V of the solar cell, as well as the output power of the cell versus its voltage V (The short-circuit current density $J_{SC} = 13.364\text{mA/cm}^2$, the open-circuit voltage $V_{OC} = 0.7631\text{V}$, as well as the illuminated area $S_I = 0.05\text{cm}^2$ [35]).

number of its supported users is $U_i = U_M$. Moreover, the new arriving users are capable of accessing one of available APs within their coverage and stay connected until the relevant services have been satisfactorily completed. Fig. 1 has shown the internal structure of widely-used solar cells as well, where the APs' battery capacity is quantized into B_M discrete levels.

The QoS of the i^{th} AP is directly linked to the number of supported users, namely U_i, as well as with its battery capacity, B_i. According to the preliminary work of Ho and Zhang [34], the QoS of the i^{th} AP may be deemed to be proportional to the reward of its supported users, which is defined as follows:

$$R_i = \sum_{n=1}^{U_i} W_i \log\left(1 + \frac{P_T(n, B_i)/N_0}{\sum_{j\neq n} P_I(d_j)/N_0 + 1}\right), U_i \neq 0, \tag{1}$$

where we have $U_i \in \{u|u = 0, 1, ..., U_M\}$ and $B_i \in \{b|b = 0, 1, ..., B_M - 1\}$, and W_i represents the channel's bandwidth. When $U_i = 0$, then $R_i = 0$. $P_I(d_j)$ denotes the other users' interference power, which is assumed to be a function of the distance between user j and the AP d_j. The APs' dissipative transmission power $P_T(U_i, B_i) = \sum_{n=1}^{U_i} P_T(n, B_i)$ is a monotonically increasing function of the number of users U_i. Specifically, $P_T(U_i, B_i)$ only represents the dissipative power during the process of serving U_i users. For each end-to-end link, the data transmission power $P_T(n, B_i)$ is a fixed value, which may not impose any interference on the other users. Furthermore, $P_T(n, B_i)/N_0$ and $P_I(d_j)/N_0$ denote the signal-to-noise power ratio (SNR) of the accessing user n and interference-to-noise power ratio (INR) of the associated user j, respectively, where N_0 is the noise power. We assume that both the battery charge as well as the users' reward remains constant during each time slot. In the following, we will specifically analyze the battery state, the user access process, as well as the expected reward of the users.

B. Battery State Model

The battery state constitutes the critical consideration, when making an AP selection. Furthermore, we should strike a trade-off between the battery-level quantization accuracy and the AP selection algorithm complexity. For example, if four levels of battery states (BaS) are considered, we have $B_M = 4$, i.e. BaS0 (Empty), BaS1 (Not sufficient), BaS2 (Adequate), as well as BaS3 (Full). Given the short duration T of each time slot, again the battery state can only change between the adjacent states at the end of each time slot. Naturally, the battery charge is improved by the solar energy harvesting mechanism and it is reduced by transmitting packets.

The power harvested is influenced both by the solar illumination intensity and by the energy harvesting efficiency, and it determines the next time slot's battery state. Fig. 2 shows the relationship between the current density $J(V)$ and the solar cell's output voltage V in the classic photoelectric configuration depicted in the bottom right corner of Fig. 1. An idealized relationship can be inferred from [35]:

$$J(V) = J_{SC} - J_O\left(e^{\frac{qV}{KT_K}} - 1\right), \tag{2}$$

where Boltzmann's constant is $K = 1.38 \times 10^{-23}$ J/K, the temperature is $T_K = 300$ K, and the electron charge is $q = 1.6 \times 10^{-19}$C. Furthermore, J_{SC} represents the short-circuit current density and J_O is the saturation current density. Then the resultant output power, P_W, of the solar cell shown in Fig. 2, obeys:

$$P_W = S_I \cdot J(V) \cdot V, \tag{3}$$

where S_I denotes the solar cell's illuminated area at each AP, which is assumed to be a constant. Let us introduce the parameter η to indicate the energy harvesting efficiency, which is assumed to be constant. Then the harvested power P_H becomes:

$$P_H = \eta \cdot P_W = \eta \cdot S_I \cdot J(V) \cdot V. \tag{4}$$

The relationship between the battery energy E and the voltage in a capacitor having the capacity of C is $E = \frac{1}{2}CV^2$. Then the B_M battery states can be deduced by partitioning the maximum battery energy $E_{\max}$ into four intervals. Therefore,

we have an estimated battery voltage for the k^{th} AP given by:

$$\overline{V_k} = \sqrt{\frac{2B_k \lfloor E_{\max}/B_M \rfloor}{C}}, \tag{5}$$

where the symbol $\lfloor \cdot \rfloor$ represents the rounding function and B_k is the battery state of the k^{th} AP. Based on the aforementioned analysis, the harvested power P_H of Equation (4) is rewritten as:

$$P_H(B_k) = \eta \cdot \left[J_{SC} - J_O \left(\exp\left\{ \frac{q\sqrt{\frac{2B_k \lfloor E_{\max}/B_M \rfloor}{C}}}{KT_K} \right\} - 1 \right) \right] \cdot \sqrt{\frac{2B_i \lfloor E_{\max}/B_M \rfloor}{C}}. \tag{6}$$

Let us now focus our attention on the battery state transitions. In our battery model, $P_T(U_k, B_k)$ is the transmission power and a constant P_D represents the sum of all other power consumption , such as the wireless backhaul dissipation, the beacon signal power, the AP's signal-processing based dissipation, etc. During each time slot, an extra amount of $P_D T$ energy consumption is considered. Given that the harvested power $P_H(B_k)$ has been measured for the k^{th} AP, the battery state transition probability from BaS_i to BaS_j during each time slot obeys:

$$\Psi_k(j|i) = \begin{cases} \Pr\left(P_H(B_k) - P_T(U_k, B_k) - P_D > \frac{E_{\max}}{B_M T} \right), & \text{if } j = i+1, \\ \Pr\left(P_T(U_k, B_k) - P_H(B_k) - P_D > \frac{E_{\max}}{B_M T} \right), & \text{if } j = i-1, \\ 1 - \sum_{m=0, m\neq i}^{B_M - 1} \Psi_k(m|i), & \text{if } j = i, \\ 0, & \text{if } |j-i| > 1. \end{cases} \tag{7}$$

where we have $0 \leq j \leq B_M - 1$ as well as $0 \leq i \leq B_M - 1$. Considering a realistic case, two schemes have been conceived in order to reduce negligible error of battery discretization. Specifically, the APs are capable of automatically adjusting the discretization resolution B_M versus the time slot T, which guarantees satisfying an appropriate energy threshold and hence it is beneficial. Moreover, a self-inspection based battery state update mechanism is used by the APs with the objective of reducing the accumulated error.

C. User Access Model

In this subsection, we consider two different types of users. Specifically, the active mobile phone users are capable of deciding upon which one of the available APs within their reach they want to access. By contrast, the passive users, which can only access the specific AP assigned by the system control center. If the selected AP reaches its maximum user-access restriction, the access request of the next active user n will be denied and the user will be assigned to another available AP by a greedy access protocol, which aims for achieving the maximum possible immediate reward. Thus, the choice of the stand-by AP is determined by evaluating:

$$AP_i = \underset{i, B_i < B_M, U_i < U_M}{\arg\max} W_i \log\left(1 + \frac{P_T(n, B_i)/N_0}{\sum_{j\neq n} P_I(d_j)/N_0 + 1} \right). \tag{8}$$

The user-arrival process and user-service process of each AP can be represented by the classic M/M/U_M queueing model. When considering the k^{th} AP, its active as well as passive users' arrival process obeys a Poisson distribution having the same arrival rate of λ_k, $k = 1, 2, ..., K$, whilst the departure rate of all users is negative exponentially distributed with a departure-rate parameter of μ. During a sufficiently short time slot T, we may readily assume that only a maximum of one user can request access to or depart from a certain AP having adequate energy. Therefore, the state transition probability of the users associated with the k^{th} AP between adjacent time slots can be defined as:

$$\Phi_k(j|i, i \leq U_M - 1) = \begin{cases} \lambda_k, & \text{if } j = i+1, \\ i\mu, & \text{if } j = i-1, \\ 1 - \lambda_k - i\mu, & \text{if } j = i, \\ 0, & \text{if } |j-i| > 1, \end{cases} \tag{9}$$

or

$$\Phi_k(j|i, i = U_M) = \begin{cases} i\mu, & \text{if } j = i-1, \\ 1 - i\mu, & \text{if } j = i, \\ 0, & \text{if } |j-i| > 1. \end{cases} \tag{10}$$

Let us denote the super-WiFi system's user association state as $U = (U_1, U_2, ..., U_K)$, and its system battery state as $B = (B_1, B_2, ..., B_K)$. Moreover, we represent the super-WiFi system's state by $S =< U, B >$, which contains both the user association state and the battery state. Hence, the overall system state is represented by the $2K$-element vector $S = (U_1, B_1, U_2, B_2, ..., U_K, B_K)$. Given the independence of both the APs as well as of the pair of system sub-states, the system's state transition probability is given by:

$$\begin{aligned} P(S'|S) &= \prod_{k=1}^{K} \Phi_k(U_k^{'}|U_k, B_k, A) \Psi_k(B_k^{'}|U_k, B_k, A) \\ &= \prod_{k=1}^{K} \Phi_k(U_k^{'}|U_k, A) \Psi_k(B_k^{'}|B_k, A). \end{aligned} \tag{11}$$

D. MDP Decision-making Model

The active/passive users' AP selection is based on a decision-making process, which is governed by the specific utility value expected in the system's steady state after an infinite number of time slots. Furthermore, the decision-making during each time slot is simply based on the current system state, rather than on a set of past system states. Let $R^t(S^t)$ represent the system's total reward at the t^{th} time slot within the current system state S^t, which implies that we have $R^t(S^t) = \sum_{k=1}^{K} R_k^t$. Thus, the system's reward $V(S)$ expected in its steady state after the elapse of an infinite number of time slots can be expressed as:

$$V(S) = E[\sum_{t=0}^{\infty} \gamma^t R^t(S^t)], \tag{12}$$

where γ is a discount factor obeying $0 < \gamma < 1$, which reflects the life cycle[2] of the decision-making process.

Naturally, we have constructed a basic MDP model, which can be described as the four-tuple $< \mathbb{S}, \mathbb{A}, \mathbf{P}, \mathbb{R} >$ detailed below:

- System state set $\mathbb{S}$: a set of all the possible system states, i.e. $S \in \mathbb{S}$, where the $2K$-element state vector is $S = (U_1, B_1, U_2, B_2, ..., U_K, B_K)$;
- User action set $\mathbb{A}$: $A \in \mathbb{A}$, where A represents the users' actions in terms of which available APs they request association with;
- State transition matrix $\mathbf{P}$: $\mathbb{S} \times \mathbb{A} \to \mathbb{S}^{'}$, where the operand $'\times'$ represents the Cartesian product, while $S^{'}(U^{'}, B^{'}) \in \mathbb{S}^{'} = \mathbb{S}$ is the system's state during the next time slot. The transition probability is given by $P(S^{'}|S, A)$;
- The system's immediate reward set $\mathbb{R}$: $\mathbb{S} \times \mathbb{A} \to \mathbb{R}$, where $R(S, A) \in \mathbb{R}$ represents the immediate reward under the access action A, when the current state is S.

Given the above definitions, the AP selection strategy is based on a decision-making process. We denote the strategy by Π, which represents a mapping from the system's state set to the user action set, which is formulated as $\Pi : \mathbb{S} \to \mathbb{A}$. Given the selection strategy Π, the system's expected reward in its steady-state after an infinite number of time slots becomes:

$$V^{\Pi}(S|S^0) = E[\sum_{t=0}^{\infty} \gamma^t R^t(S^t, \Pi(S^t))|S^0], \tag{13}$$

where S^0 is the initial system state. Equation (13) can be rewritten as the following function to be evaluated:

$$V^{\Pi}(S \mid S^0) = R(S^0, \Pi(S^0)) + \gamma \sum_{S' \in \mathbb{S}} P(S^{'}|S^0, \Pi) V^{\Pi}(S^{'}|S^0). \tag{14}$$

Let us now define what we might refer to as an action value function (AVF) in the form of:

$$Q^{\Pi/A}(S, A) = R(S, A) + \gamma \sum_{S' \in \mathbb{S}} P^A(S^{'}|S, A) V^{\Pi}(S^{'}), \tag{15}$$

which indicates that an already communicating user takes action A during the system state S, while accepting the strategy Π, when it finds itself in other system states. Then, the optimal strategy can be deduced from:

$$\Pi^*(S) = \arg\max_{A \in \mathbb{A}} Q(S, A), \tag{16}$$

yielding:

$$\Pi^*(S) = \arg\max_{A \in \mathbb{A}} \{R(S, A) + \gamma \sum_{S' \in \mathbb{S}} P^A(S^{'}|S, A) V^{\Pi}(S^{'})\}. \tag{17}$$

Meanwhile, we have:

$$V^{\Pi^*}(S) = \max_{A \in \mathbb{A}} \{R(S, A) + \gamma \sum_{S' \in \mathbb{S}} P^A(S^{'}|S, A) V^{\Pi^*}(S^{'})\}. \tag{18}$$

Let $\Pi^*(S)$ represent the optimal strategy, while $V^{\Pi^*}(S)$ be the corresponding optimal expected reward. Because no other strategy exists, which would offer a higher expected reward, the optimal strategy can reach the Nash-equilibrium. For any other strategy $\Pi^{'}(S)$, which we refer to as an ε-optimal strategy, we have:

$$V^{\Pi^{'}}(S) - V^{\Pi^*}(S) \leq \varepsilon, \forall S \in \mathbb{S}. \tag{19}$$

In order to solve the MDP formulated in Equation (13), we propose the iterative **Algorithm** 1.

Algorithm 1 Iterative Algorithm

- **Initialization**

$V_k^0(S) \leftarrow 0$, for all $S \in \mathbb{S}$ and $k = 1, 2, ..., K$;
$\Pi^0(S) \leftarrow 1$, for all $S \in \mathbb{S}$.

- **Iteration**

while $\max_S \mid V_k^t(S) - V_k^{t+1}(S) \mid > \epsilon$ **do**
for all $S \in \mathbb{S}$ and $k = 1, 2, ..., K$,
$V_k^{t+1}(S) \leftarrow R_k(S) + \gamma \sum_{S'} P_k(S'|S, \Pi^t) V_k^t(S')$;
$\Pi^{t+1}(S) \leftarrow \arg\max_{A \in \mathbb{A}} V^{\Pi^{t+1}}(S)$,
where $S^{'}$ is the next time slot system state.
end while

- **Output**

$\Pi^*(S) \leftarrow \Pi^{t+1}(S)$;
$V^*(S) \leftarrow \max_{A \in \mathbb{A}} V^{\Pi^*}(S)$.

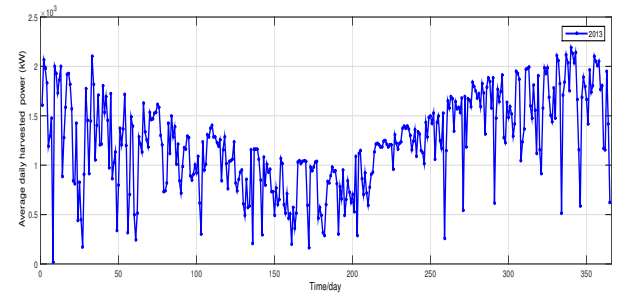
IV. POMDP AP Selection Strategy Based on a Solar-activity Measurement Aided Energy Harvesting Model

The POMDP models an agent-based decision-making process, in which the system dynamics are determined by a MDP, where the agent is unable to directly observe some of the underlying states. Specifically, it has to maintain a probability distribution over the set of legitimate states based on a set of observations and observation probabilities. In Section III, we assumed that both the battery states as well as the user access states are completely observable. However, Equation (4) represents an inaccurate estimate of the harvested solar power, which was formulated under the hypothesis of a fixed energy harvesting efficiency of η as well as under the idealized relationship of Equation (2). However, the solar radiation model has a grave influence on the battery states, thus critically affecting the final decision-making. Therefore, we first construct a solar radiation model relying on the historical solar activity observation data, followed by the conception of a POMDP AP selection strategy for the proposed super-WiFi system.

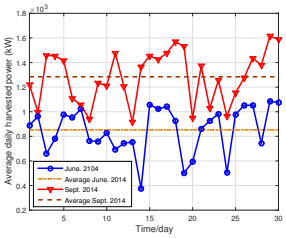
A. Data-driven Solar Radiation Model

Solar radiation intensity changes over time in a year, as influenced by the weather conditions. The radiation sensors have a limited sampling rate, which makes it hard to simultaneously record the solar radiation intensity and to accurately estimate the system's battery state. In this subsection, we focus our attention on modeling the solar radiation as well as on

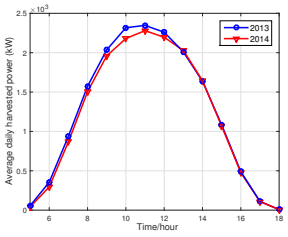
[2]In the process of decision making, the influence of environment should be considered. The impact on the target is reduced as a function of the time-elapse. The life cycle reflects the above phenomenon.



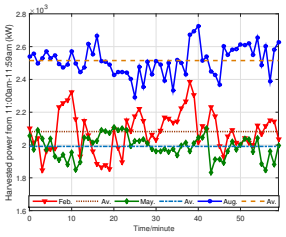
(a) Energy-Harvested power in a year.



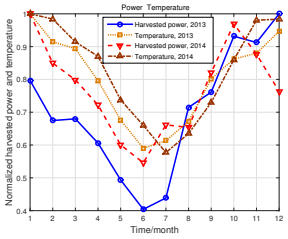
(b) Harvested power in a month.



(c) Harvested power in a day.



(d) Harvested power within an hour.



(e) Influence of temperature.

Fig. 3. Statistical analysis of historical solar radiation observation data provided by the University of Queensland, Australia [19].

estimating the real-time radiation intensity relying on historical solar radiation observation data.

The historical solar radiation observation data was provided by the University of Queensland, Australia [19], where the power harvested from solar radiation was recorded from 5:00am to 19:00pm, during the interval from 2013 to 2014. Moreover, the sampling frequency was 1/min. Fig. 3 presents some of the historical observation data. In subgraph 3(a), the time series of the average daily harvested power is sketched over the whole year of 2013. Subgraph 3(b) depicts the variation tendency of the average daily harvested power during both June as well as September. Subgraph 3(c) concentrates on the power variation tendency within a day from 5:00am to 19:00pm. Furthermore, the power harvested within the same one-hour period of different months is shown in subgraph 3(d), respectively, along with their hourly averages. Finally, subgraph 3(e) indicates the relationship between the normalized harvested power and the normalized temperature. We may conclude that as expected, the weather conditions have a significant impact on the harvested solar power. Moreover, the average daily harvested power fluctuates slowly within a short time interval, but varies substantially in a long term.

Therefore, in a short period of time, say, within an hour, the real-time harvested solar power can be modeled as: $P_H^t = \theta + \kappa$, where θ is constant for an hour, while κ is a small perturbation. Moreover, multiple factors, such as the effective irradiation area, the clouds' distribution, the sensors' operating status, etc. may independently affect the harvested power. Relying on the central-limit theorem, the perturbation κ can be regarded as being Gaussian distributed, i.e., we have $\kappa \sim \mathcal{N}(0, \sigma^2)$. Furthermore, the distribution of P_H can be written as $P_H \sim \mathcal{N}(\theta, \sigma^2)$. Given a set of observation data $X = [x_0, x_1, ..., x_{M-1}]$, it is critical to accurately estimate θ and σ^2. Upon introducing the notation of $\Lambda = [\theta, \sigma^2]^{\mathrm{T}}$, we have:

$$\begin{aligned} p(X; [\theta, \sigma^2]^{\mathrm{T}}) &= \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - \theta)^2 \right\} \\ &= \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left\{ -\frac{1}{2\sigma^2} \left(\sum_{i=0}^{M-1} {x_i}^2 - 2\theta \sum_{i=0}^{M-1} x_i + M\theta^2 \right) \right\} \\ &= g(T(X), [\theta, \sigma^2]^{\mathrm{T}}) \times h(X). \end{aligned} \tag{20}$$

Equation (20) represents a Neyman-Fisher factorization [36] of $p(X; [\theta, \sigma^2]^{\mathrm{T}})$, where we have both $h(X) = 1$, as well as the sufficient statistics of

$$T(X) = \begin{bmatrix} \sum_{i=0}^{M-1} x_i \\ \sum_{i=0}^{M-1} {x_i}^2 \end{bmatrix}. \tag{21}$$

For all the $\Lambda = [\theta, \sigma^2]^{\mathrm{T}}$ values we have:

$$\int \nu[T(X)] p(T(X); \Lambda) dT(X) = 0, \tag{22}$$

if and only if $\nu[T(X)] = 0$. In other words, $T(X)$ satisfies completeness. Upon introducing

$$T_1(X) = \sum_{i=0}^{M-1} x_i, \tag{23}$$

as well as

$$T_2(X) = \sum_{i=0}^{M-1} {x_i}^2, \tag{24}$$

we can construct the unbiased vector of:

$$\begin{aligned} y[T(X)] &= \begin{bmatrix} \frac{1}{M} T_1(X) \\ \frac{M}{M-1} (\frac{1}{M} T_2(X) - (\frac{1}{M} T_1(X))^2) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{M} \sum_{i=0}^{M-1} x_i \\ \frac{M}{M-1} (\frac{1}{M} \sum_{i=0}^{M-1} {x_i}^2 - (\frac{1}{M} \sum_{i=0}^{M-1} x_i)^2) \end{bmatrix}, \end{aligned} \tag{25}$$

where we have $E[y(T(X))] = [\theta, \sigma^2]^{\mathrm{T}}$. Therefore, the corresponding estimates may be formulated as:

$$\widehat{\theta} = \frac{1}{M} \sum_{i=0}^{M-1} x_i, \tag{26}$$

as well as

$$\widehat{\sigma}^2 = \frac{M}{M-1} \left(\frac{1}{M} \sum_{i=0}^{M-1} {x_i}^2 - \left(\frac{1}{M} \sum_{i=0}^{M-1} x_i \right)^2 \right). \tag{27}$$

The corresponding estimation errors are given by $\mathrm{var}(\widehat{\theta}) = \frac{\sigma^2}{M}$ and $\mathrm{var}(\widehat{\sigma}^2) = \frac{2\sigma^4}{M}$. Hence, we can obtain the distribution of the harvested solar power in the form of $P_H \sim \mathcal{N}(\widehat{\theta}, \widehat{\sigma}^2)$. Specifically, having more observation data as well as a lower perturbation contributes a more accurate harvested power distribution estimation.

In next subsection, we construct a POMDP decision-making model relying on the above measured solar radiation model.

B. Battery State Model

Given a short time slot T, E_C represents the energy consumption during a time slot. Hence, we have $E_C = P_T(U_k, B_k)T + P_D T = (U_k P_0 + P_D)T$, where P_0 is the unit transmission power and P_D represents all the other power consumption components, such as the wireless backhaul's dissipation, the beacon signal power, etc. Moreover, E_H indicates the energy harvested under the assumption that the fluctuation of the harvested power level is quasi-static over the information transmission interval, i.e. we have $E_H = \eta P_H(X)T$, where $X = [x_0, x_1, ..., x_{M-1}]$ represents the observation data of solar radiation. At the end of each time slot, an energy replenishing process takes place at each APs, hence we have:

$$E_R^{t+1} = \min\{E_R^t + E_H^t - E_C^t, E_{max}\}, \tag{28}$$

where E_R^{t+1} indicates the residual energy reserved for the next time slot. In light of the discrete states of the POMDP model, it is necessary to discretize the continuous battery charge value. Again, a rounding function $\lfloor \cdot \rfloor$ is invoked for linearly segmenting the battery energy value into B_M levels, i.e. we have $B_M = \lfloor E_{max}/(P_0 T) \rfloor + 1$. The battery state of the k^{th} AP at the end of time slot t can be expressed as $B_k^t = \lfloor E_R^{t+1}/(P_0 T) \rfloor$. In order to reduce negligible error of battery discretization, the APs are capable of automatically adjusting the discretization resolution B_M versus the time slot T, which guarantees satisfying an appropriate energy threshold. Moreover, a self-inspection based battery state update mechanism is used by the APs with the objective of reducing the accumulated error. Based on this, the users requesting access to it can judge, whether the battery charge is sufficient for reliable information transmission.

Therefore, we arrive at:

$$j \leq \frac{E_H - E_C}{P_0 T} \leq j + 1, \tag{29}$$

where we have $j = \lfloor \frac{E_H - E_C}{P_0 T} \rfloor$, where the integer j as well as $(j+1)$ represent the battery energy variation lower and upper bound, respectively. Let $\Delta B_k = B_k^{t+1} - B_k^t$ indicate the difference between the two battery states of the k^{th} AP. Thus, the probability that the battery state will change by an amount of ΔB_k in terms of the harvested energy as well as the users' action A can be formulated as:

$$\begin{aligned} &\Pr(\Delta B_k = i | E_H, A, j) \\ &= \begin{cases} \frac{E_H - E_C}{P_0 T} - j, & \text{if } i = j+1, \\ (j+1) - \frac{E_H - E_C}{P_0 T}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{30}$$

After some further manipulations, we arrive at:

$$\begin{aligned} &\Pr(\Delta B = i | A) = \\ &\int_{(i-1)E_0 + E_C}^{iE_0 + E_C} \Pr(\Delta B = i | E_H, A, i = j+1) \Gamma(E_H) dE_H + \\ &\int_{iE_0 + E_C}^{(i+1)E_0 + E_C} \Pr(\Delta B = i | E_H, A, i = j) \Gamma(E_H) dE_H, \end{aligned} \tag{31}$$

where the unit transmission energy of $E_0 = P_0 T$, as well as $\Gamma(E_H) \sim \mathcal{N}(T\widehat{\theta}, T^2\widehat{\sigma}^2)$ are derived from the solar radiation model of Equation (26) and (27). Then, the battery transition probability of the k^{th} AP from BaS_i to BaS_j can be calculated as:

$$\Psi_k(j|i) = \begin{cases} \Pr(\Delta B = j - i | A), & \text{if } j \leq B_M - 1, \\ \sum_{\Delta B = B_M - 1 - i}^{\Delta B^{\max}} \Pr(\Delta B | A), & \text{if } j = B_M - 1, \end{cases} \tag{32}$$

where $\Pr(\Delta B^{\max} | A)$ is a sufficiently small value. When we have $j = B_M - 1$, the battery is fully charged and all the extra harvested battery energy has to be discarded.

C. User Access Model

As formulated in Section III-C, the users' arrival/departure at/from a certain AP between adjacent time slots obeys a revised birth and death process, respectively. In order to simplify our discourse, new users requesting access to any of the super-WiFi APs obey the same arrival rate λ, while μ represents their identical departure rate. Under the hypothesis of adequate battery energy charge and a sufficiently short time slot T, for any candidate AP k, we can rewrite Equation (9) and (10) as:

$$\Phi_k(j|i, i \leq U_M - 1) = \begin{cases} \lambda, & \text{if } j = i+1, \\ i\mu, & \text{if } j = i-1, \\ 1 - \lambda - i\mu, & \text{if } j = i, \\ 0, & \text{if } |j - i| > 1, \end{cases} \tag{33}$$

and

$$\Phi_k(j|i, i = U_M) = \begin{cases} i\mu, & \text{if } j = i-1, \\ 1 - i\mu, & \text{if } j = i, \\ 0, & \text{if } |j - i| > 1, \end{cases} \tag{34}$$

where $\Phi_k(j|i)$ is the probability traversing from state i to adjacent state j. However, when energy depletion is encountered, i.e., we have $E_R^{t-1} + E_H^t - E_T^t < 0$, then we arrive at $j = 0$.

D. Transition Probabilities of the System

The K APs of our super-WiFi system have states constituted by both the user-association states as well as by the battery states. As defined in the MDP, let $U = (U_1, U_2, ..., U_K)$ denote the user-association states, while $B = (B_1, B_2, ..., B_K)$ represents the AP battery states, where $U_k \leq U_M$ and $B_k \leq B_M - 1$. Furthermore, the super-WiFi system state can be written as a $2K$-element vector $S = (U_1, B_1, U_2, B_2, ..., U_K, B_K)$, which includes both the K APs' user-association states and the K APs' battery states. Relying on Equation (32), (33) and (34), as well as assuming the independence of each AP's two sub-states, in contrast to Equation (11), the system's state transition

probability is given by

$$P(S^{'}|S) = \prod_{k=1}^{K} \Phi_k(U_k^{'}|U_k, A)\Psi_k(B_k^{'}|B_k, A). \tag{35}$$

E. POMDP Decision-making Model

In this subsection, we assume that the requesting access users only have partial knowledge of the entire super-WiFi system's state, which is refered to as the observed state O. Relying on the above definitions and hypotheses, we construct the POMDP decision-making model as a seven-tuple $< \mathbb{S}, \mathbb{A}, \mathbb{O}, \mathbf{P}, \Omega, \mathbf{F}, \mathbb{R} >$, which is detailed below:

- System state set $\mathbb{S}$: a set of all the possible system states, i.e., $S \in \mathbb{S}$, where the $2K$-element state vector is formulated as $S = (U_1, B_1, U_2, B_2, ..., U_K, B_K)$. $card(\mathbb{S}) = (U_M B_M)^K$;
- User action set $\mathbb{A}$: a set of all the possible users' actions, with $A \in \mathbb{A}$, where A represents the users' specific actions, in terms of which of the available APs they request access to;
- Observed state set $\mathbb{O}$: a set of all the possible observed system states, i.e., $O \in \mathbb{O}$;
- State transition matrix $\mathbf{P}$: $\mathbb{S} \times \mathbb{A} \to \mathbb{S}^{'}$, where the operand $'\times'$ represents the Cartesian product, while $S^{'}(U^{'}, B^{'}) \in \mathbb{S}^{'} = \mathbb{S}$ is the system's state during the next time slot. The transition probability is $P(S^{'}|S, A)$;
- Belief state vector Ω: $\Omega = \{\omega(S)|S \in \mathbb{S}\}$, where $\omega(S)$ is the belief probability vector reflecting the grade of similarity between the current observed state $O \in \mathbb{O}$ and the legitimate system state S;
- Observation function matrix $\mathbf{F}$: $\Omega \times \mathbb{A} \times \mathbb{S}^{'} \to \mathbb{O}$, where $F(O, \Omega, A, S^{'}) = \Pr(O|\Omega, A, S^{'})$ represents the probability that the observed state becomes O at the next time slot in terms of the user's action A;
- Immediate reward set of the system $\mathbb{R}$: $\mathbb{S} \times \mathbb{A} \to \mathbb{R}$, and $R(S, A) \in \mathbb{R}$, where $R(S|A) = \sum_{k=1}^{K} R_k(S|A)$ indicates the immediate reward as a consequence of access action A under current state S.

The belief probability $\omega(S^{'})$ for a certain observed state O can be updated based on the following Bayesian formula:

$$\begin{aligned}
&\omega(S^{'}) = \Pr(S^{'}|O, A, \Omega) \\
&= \frac{Pr(O|S^{'}, A, \Omega)Pr(S^{'}|A, \Omega)}{Pr(O|A, \Omega)} \\
&= \frac{Pr(O|S^{'}, A, \Omega)\sum\limits_{S\in\mathbb{S}} Pr(S^{'}|A, \Omega, S)Pr(S|A, \Omega)}{Pr(O|A, \Omega)} \\
&= \frac{Pr(O|S^{'}, A, \Omega)\sum\limits_{S\in\mathbb{S}} Pr(S^{'}|A, S)Pr(S|\Omega)}{Pr(O|A, \Omega)} \\
&= \frac{F(O, \Omega, A, S^{'})\sum\limits_{S\in\mathbb{S}} P(S^{'}|A, S)\omega(S)}{Pr(O|A, \Omega)},
\end{aligned} \tag{36}$$

where we have:

$$\begin{aligned}
\Pr(O|A, \Omega) &= \sum_{S^{'}\in\mathbb{S}} \Pr(O, S^{'}|A, \Omega) \\
&= \sum_{S^{'}\in\mathbb{S}} \Pr(S^{'}|A, \Omega)\Pr(O|S^{'}, A, \Omega) \\
&= \sum_{S^{'}\in\mathbb{S}}\sum_{S\in\mathbb{S}} \Pr(S^{'}|A, S)\omega(S)\Pr(O|S^{'}, A, \Omega) \\
&= \sum_{S^{'}\in\mathbb{S}} F(O, \Omega, A, S^{'})\sum_{S\in\mathbb{S}} P(S^{'}|A, S)\omega(S).
\end{aligned} \tag{37}$$

Thus, the POMDP can be reduced to a belief MDP in terms of the belief state vector Ω [37]. Therefore, the expected reward of the system relying on strategy Π after an infinite number of time slots can be written as:

$$V^{\Pi}(S|S^0) = E[\sum_{t=0}^{\infty} \gamma^t R^t(S^t, \Pi(S^t))\omega(S^t)|S^0], \tag{38}$$

where S^0 is the initial system state. After a number of simplifications, Equation (38) is reduced to:

$$\begin{aligned}
V^{\Pi}(S \mid S^0) = &\sum_{S^0\in\mathbb{S}} R(S^0, \Pi(S^0))\omega(S^0) \\
&+ \gamma \sum_{S^{'}\in\mathbb{S}} P(S^{'}|S^0, \Pi)V^{\Pi}(S^{'}|S^0).
\end{aligned} \tag{39}$$

Hence, the optimal strategy is given by:

$$\begin{aligned}
\Pi^*(S) = \arg\max_{A\in\mathbb{A}} \{&\sum_{S\in\mathbb{S}} R(S, A)\omega(S) + \\
&\gamma \sum_{S^{'}\in\mathbb{S}} P^A(S^{'}|S, A)V^{\Pi^*}(S^{'})\}.
\end{aligned} \tag{40}$$

Moreover, we have:

$$\begin{aligned}
V^{\Pi^*}(S) = \max_{A\in\mathbb{A}} \{&\sum_{S\in\mathbb{S}} R(S, A)\omega(S) + \\
&\gamma \sum_{S^{'}\in\mathbb{S}} P^A(S^{'}|S, A)V^{\Pi^*}(S^{'})\}.
\end{aligned} \tag{41}$$

The above formula can appear rather complex at the first sight. However, as shown in [38], $V^{\Pi^*}(S)$ is a piecewise linear and convex function, which is composed solely of line segments or hyperplanes. Hence, it can be simplified to:

$$V^{\Pi^*}(S) = \max_{k} \{\sum_{S\in\mathbb{S}} \omega(S)\alpha^k(S)\}, \tag{42}$$

for a set of vectors $\alpha = \{\alpha^1, \alpha^2, ...\}$, where we have $\alpha^k = [\alpha_1^k, \alpha_2^k, ..., \alpha_{card(\mathbb{S})}^k]$. Each vector represents the coefficients of one of the linear segments of a piecewise linear function. Relying on this α-vector based representation, we have found an appealing technique of representing the reward function for the belief state space. This simpler representation is the key element of the POMDP solution technique. Since each set of α^k vectors may be regarded as a set of parameters specifying a hyper-linear function, there is a dominated hyperplane structure in the model. Moreover, the partitions of the belief space during the time slot t can be calculated in terms of all the dominating vectors α_{t-1} of the previous time-slot.

Algorithm 2 Energy Function-based Algorithm

- **Initialization**

$\omega(S^0) \leftarrow \omega(S_0)$, where $\omega(S_0)$ is the initial belief vector; $\Pi(S^0) \leftarrow 1$, for all $S \in \mathbb{S}$.

- **Procedure**

repeat

1. Estimate the belief vector: $\omega(S^{t+1}) \leftarrow \Pr(S^{t+1}|O, A, \omega(S^t))$;
2. Calculate the energy function in the time slot t: $H[\omega(S^{t+1}), \Pi(S^t)]$;
3. Choose the suboptimal decision via: $\Pi^*(S^t) \leftarrow \arg\max_{\Pi} E[H(\omega(S^{t+1}), \Pi(S^t))]$, where $\omega(S^{t+1})$ is the corresponding belief vector if we choose decision Π;
4. Update the belief vector: $\omega(S^{t+1}) \leftarrow \Pr(S^{t+1}|O, \Pi^*(S^t), \omega(S^t))$.

until Terminated

Accordingly, the reward action that is optimal can be derived from (36), (37) and (42).

The optimal POMDP solution can be readily calculated in an off-line procedure, if the number of system states is small. However, the POMDP method has its limitations when K, U_M, and B_M are large, as well as when the parameters, such as the users' arrival rate λ, departure rate μ, etc. change rapidly. Let us now elaborate on how to reduce the complexity and how to expedite the calculations. The problem in Equation (42) can be reformulated as:

$$V^*(S) = \max_{\Pi} \{\sum_{t=0}^{\infty} E[H(\omega(S^t), \Pi(S^t))]\}, \tag{43}$$

where the expectation $E[\cdot]$ takes into account the probability of receiving different observations, while $H[\omega(S^t), \Pi(S^t)]$ represents the system's energy function. Thus, given the belief vector $\omega(S^t)$, our super-WiFi system's energy function is given by:

$$\begin{aligned} H[\omega(S^t),\Pi(S^t)] &= \sum_{S\in\mathbb{S}} \omega^t(S) \\ &\sum_{i=1}^{K} \min_{\Pi} \{E_{iR}^t + E_{iH}^t - E_{iC}^t, E_{max}\}. \end{aligned} \tag{44}$$

Then, a suboptimal solution can be conceived by maximizing the system's next-time-slot energy function, yielding:

$$\Pi^*(S) = \arg\max_{\Pi} E[H(\omega(S^{t+1}), \Pi(S^t))]. \tag{45}$$

This algorithm (43)∼(45), detailed in **Algorithm** 2, aims for optimizing the energy transfer function rather than the system's expected long-term reward. This philosophy relies on the assumption that the more energy the super-WiFi system harvests, the more rewards the rational users can glean. This algorithm is an altruistic one, which may sacrifice its own individual reward in order to arrive at the system's optimum state. Furthermore, this algorithm is capable of adapting to sudden environmental changes.

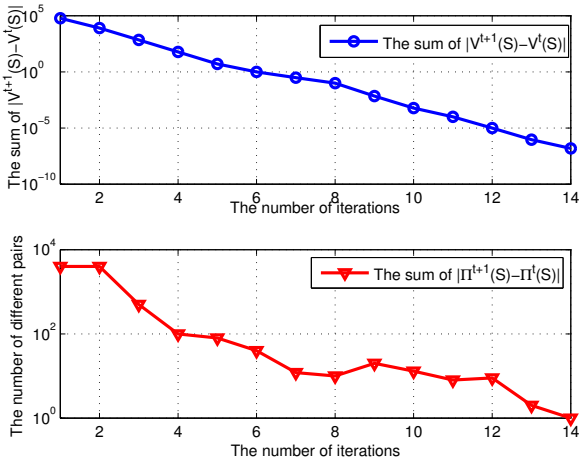


Fig. 4. Convergence analysis of the proposed MDP AP selection algorithm.

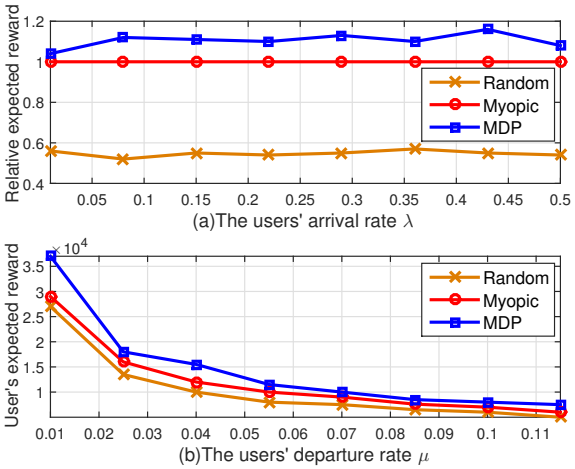


Fig. 5. Comparison of the expected reward versus both the arrival rate as well as departure rate. (Relative expected rewards are derived from the three user's expected rewards normalized by the reward with the myopic regime.)

V. Simulation Results

A. Simulation Results of the MDP AP Selection Strategy

In this subsection, the performance of the proposed AP selection strategy based on the MDP is illustrated in comparison to both the myopic strategy of [39] and to a random strategy. Specifically, the myopic strategy requires the users to access the specific AP providing the maximum immediate reward, which is formulated as $\Pi^M(S) = \arg\max_{i,1\le i\le K} R_i(S)$. By contrast, the random strategy grants an identical access probability for each requesting user, which is encapsulated in $Pr(\Pi^R(S) : \text{access } i) = 1/K$.

Based on the super-WiFi system model of Section III, it is necessary to carefully select the salient simulation parameters. First, we divided the battery charge into four discrete isometric battery levels, yielding $B_M = 4$. Furthermore, we set: P_I, P_T, P_H and N_0 as: $P_T/N_0 = 10$, $P_I/N_0 = 10$, $E_{max}/(B_M \times T) = 6$, as well as $\max P_H/N_0 = 6, 7, ..., 10$ in conjunction with a different energy harvesting efficiency.

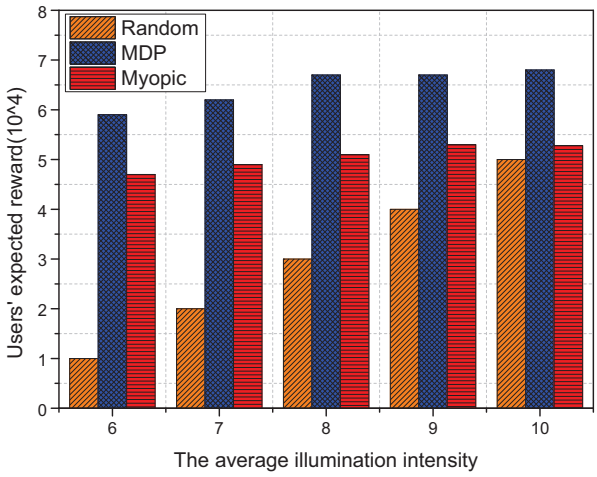


Fig. 6. The reward gleaned by three AP selection strategies under different illumination intensities (kw/m^2).

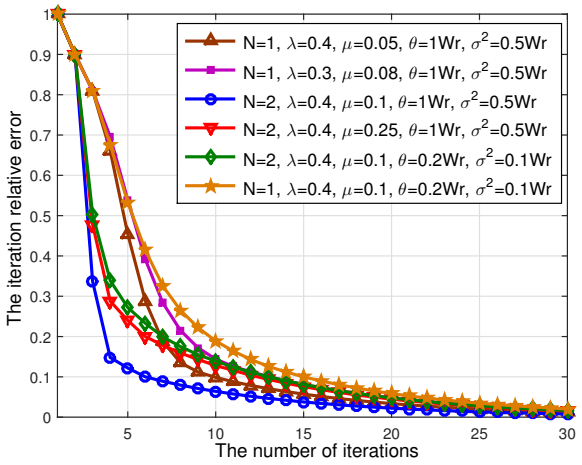


Fig. 7. The convergence analysis of proposed POMDP AP selection algorithm.

We denote the different energy harvesting efficiencies as $\eta = 6\eta_{unit}, 7\eta_{unit}, ..., 10\eta_{unit}$, respectively. The simulations were repeated 10000 times to avoid random errors.

The convergence analysis of the proposed MDP value iterative algorithm was illustrated in Fig. 4. Fig. 4(a) illustrates the sum of $\mid V_k^{t+1}(S) - V_k^t(S) \mid^2$, for all k, as well as $S \in \mathbb{S}$, while Fig. 4(b) shows the number of strategy errors during the subsequent iterations $\mid \Pi^{t+1}(S) - \Pi^t(S) \mid^2$. Beyond the 15^{th} iteration, the selection strategy may be deemed to be fully converged. Moveover, the figure also shows that our MDP value iterative algorithm converges according to a near-exponential trend, which appears as a linear trend on a logarithmic axis.

In Fig. 5, the expected reward versus the arrival rate λ_k of the requesting users, and versus the departure rate μ was evaluated. Let us consider the situation associated with $K = 3$, as well as with the arrival rates of $\lambda_1 = \lambda_2 = \lambda_3$. Furthermore, assuming $U_i = 0, 1, ..., 4$, i.e., $U_M = 4$ and $B_M = 4$, the energy harvesting efficiency becomes $\eta = 10\eta_{Unit}$. Fig. 5(a), where we have $\mu = 0.1$ and subgraph 5(b), where $\lambda_k = 0.1, k = 1, 2, 3$, showed that the proposed AP selection strategy based on MDP exhibited a clear advantage both over the myopic one, as well as over the random strategy in terms of the total expected reward. The random strategy performed the worst. Specifically, the myopic regime's total expected reward was normalized to 1 as a reference. Because of the limitation of $U_M = 4$, increasing the requesting users' arrival rate had nearly no impact on the reward. However, as depicted in Fig. 5(b), the total expected reward was reduced upon increasing the departure rate of the already communicating users, which indicated that users would stay shorter in the super-WiFi system. Moreover, when μ became high, the proposed strategy only had a modest advantage over the other two, which highlighted the important role of having an immediate reward.

Fig. 6 shows the performance of three AP selection algorithms under different average illumination intensities. Let us set $\lambda_1 = \lambda_2 = \lambda_3 = 0.1$, $\mu = 0.1$, $K = 3$, $B_M = 4$, as well as $U_i = 0, 1, ..., 4$. Observe in Fig. 6 that all the three strategies had a higher total expected reward in the presence of a more intense solar radiation. Moreover, the proposed AP selection algorithm based on the MDP exhibited a substantial advantage over the other two algorithms.

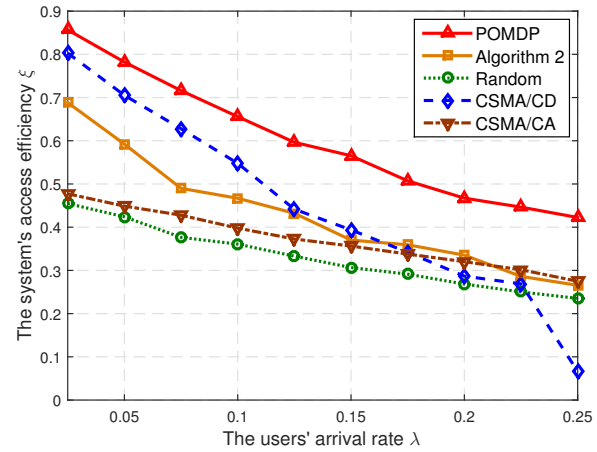
B. Simulation Results of the POMDP AP Selection Strategy

In this subsection, the convergence of the POMDP iterative algorithm is illustrated in Fig. 7, where the y-axis represents the α-vector error, which reflects the difference of the super-WiFi system's expected reward between two consecutive iterations. Moreover, the discount factor of Equation (38) is set to $\gamma = 0.9$. The Bellman stopping criterion of [40] is used to determine when to stop the iteration process. As we can see from Fig. 7, the POMDP model converges exponentially under certain conditions, depending on the system parameters. The reference benchmark solar intensity in the legend is given by $W_r = 1\text{kW}/\text{m}^2$.

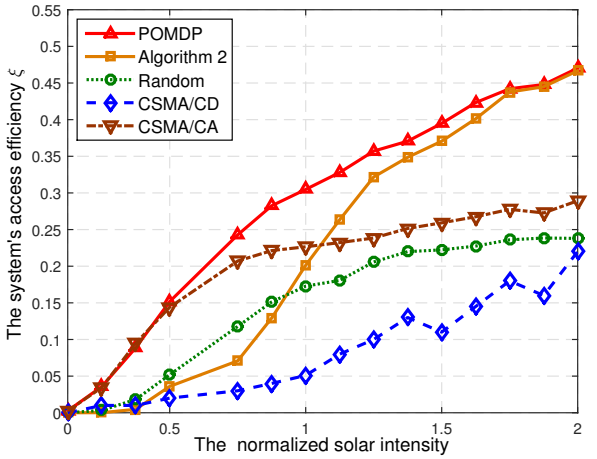
The efficiency of the AP selection algorithms is compared in terms of the system's access efficiency ξ, i.e., $\xi = N_S/T_S$, where N_S is the total number of successful access attempts during the entire simulation time T_S. Again, in this paper, we rely on the observed solar radiation data provided by the University of Queensland, Australia [19] and make the assumption that the simulation time is $T_S = 10,000\text{s}$, with each time slot having a length of $T = 200\text{ms}$. Moreover, the energy harvesting efficiency is $\eta = 75\%$ [41], while the transmission power requested for serving a single user is 40mW.

In order to fully characterize our AP selection efficiency, several algorithms are compared, including the classic carrier sensing multiple access/collision detection (CSMA/CD), the carrier sensing multiple access/collision avoidance (CSMA/CA)[3], as well as the random access algorithm. We use the same acronym CSMA/CD and CSMA/CA

[3]Strictly speaking, the CSMA/CD and CSMA/CA in this paper are different from the Ethernet's data link layer protocols. Here, both of them represent the access control mechanisms.

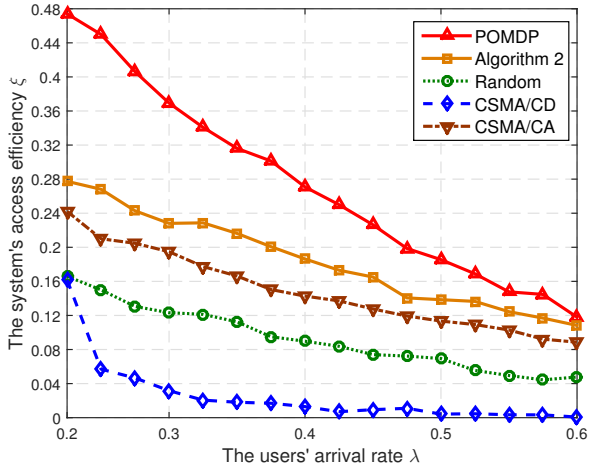


(a) Influence of the arrival rate

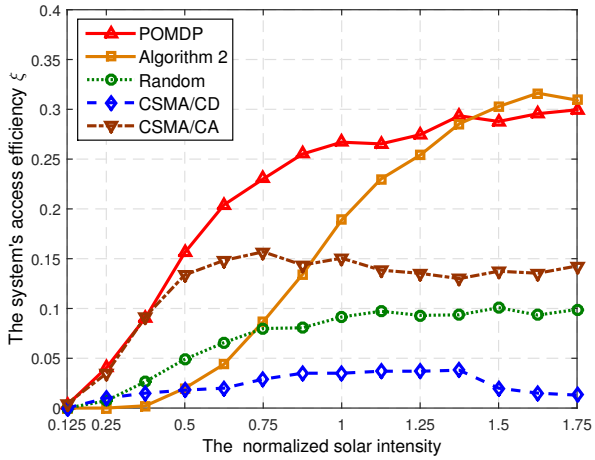


(b) Influence of the normalized solar radiation intensity

Fig. 8. The relationship between the system access efficiency and users' arrival rate as well as solar radiation intensity with one AP ($K = 1$, $U_M = 3$, $B_M = 8$).



(a) Influence of the arrival rate



(b) Influence of the normalized solar radiation intensity

Fig. 9. The relationship between the system access efficiency and users' arrival rate as well as solar radiation intensity with multiple APs ($K = 2$, $U_M = 1$, $B_M = 3$).

for convenience as in the Ethernet's data link layer, but it is important to note that "carrier sensing" here means that one of the users senses the surrounding APs by receiving the short beacon message from the APs, instead of checking the strength of the physical carriers. Our CSMA/CD method blocks any new requests when detecting a collision, and then the classic exponential back-off algorithm is used. Explicitly, after n failures, a gradually increasing number of sleeping time slots will be chosen from the range of 0 and $2^n - 1$. Furthermore, the CSMA/CA method would access the AP after the user receives the acknowledgement (ACK) from a certain AP in the most recent time slot and becomes aware that a successful service is available. In terms of the random access algorithm, the users simply select one of the surrounding APs with an equal probability.

In Fig. 8, we consider the situation having only a single AP, i.e., $K = 1$, as well as a maximum number of communicating users of $U_M = 3$, and a maximum number of battery states of $B_M = 8$. Fig. 8(a) shows the relationship between the arrival rate and the system's access efficiency for a departure rate of $\mu = 0.05$, as well as for the solar radiation parameters of $\theta = 1$Wr, $\sigma^2 = 0.5$Wr. Again, W_r represents the reference benchmark solar intensity. The performance of the proposed POMDP is superior to that of the others, while the overall performance **Algorithm** 2 is mediocre. Furthermore, we can see that the CSMA/CD has a good performance, when the system is not too busy, but the performance degrades sharply upon increasing the users' arrival rate. In Fig. 8(b), we focus our attention on the influence of the solar radiation intensity in conjunction with the users' arrival rate of $\lambda = 0.4$. A prominent point is that even when the solar intensity is high, the system's access efficiency tends to be saturated for the three traditional algorithms, instead of further improving their efficiency. This is because these algorithms do not make full use of the energy harvesting information available in the super-WiFi system. As demonstrated in Section IV-E, when the solar intensity is high, the performance of the proposed suboptimal **Algorithm** 2, which substantially reduces the complexity,

remains nearly identical to that of the POMDP.

In Fig. 9, multiple APs ($K = 2$) are considered with the maximum number of admitted users being $U_M = 1$, as well as having a maximum number of battery states given by $B_M = 3$. The departure rate is $\mu = 0.05$. We can conclude from Fig. 9(a) that a highly loaded system makes the CSMA/CD method almost useless, when the users' arrival rate reaches a certain value. As shown in Fig. 9(b), where $\lambda = 0.4$, the system's access efficiency recorded for all the AP selection algorithms only increases with the solar radiation intensity in a relatively small range. However, the performance of the CSMA/CD, CSMA/CA, as well as of the random selection algorithm remains unchanged, regardless of the increase of the solar radiation intensity. Moreover, the suboptimal **Algorithm** 2 is capable of outperforming the POMDP method at a strong solar radiation intensity, which may be deemed to be the result of the approximations and hypotheses inherent in the POMDP model.

VI. Conclusions

The super-WiFi system has a vast array of compelling applications in both civilian and military missions. In this paper, we focused our attention on the AP selection strategies of energy harvesting aided super-WiFi systems, as well as on a measurement-based solar radiation model. Furthermore, in order to construct a capable large-scale super-WiFi system, both a MDP as well as a POMDP AP selection algorithm was proposed. Additionally, we conceived iterative methods for the solution of both the MDP and POMDP. Finally, performance evaluations were provided. Through our extensive simulations, our proposed algorithms were shown to achieve significant gains with respect to the conventional approaches, which confirmed the efficiency of the MDP and POMDP AP selection algorithms.

References

[1] N. Cheng and X. S. Shen, "Opportunistic communication spectra utilization," in *Opportunistic Spectrum Utilization in Vehicular Communication Networks*. Springer, 2015, pp. 9–27.

[2] V. Gonçalves and S. Pollin, "The value of sensing for TV white spaces," in *2011 IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Aachen, Germany, May 2011, pp. 231–241.

[3] X. Feng, J. Zhang, and Q. Zhang, "Database-assisted multi-ap network on TV white spaces: Architecture, spectrum allocation and AP discovery," in *2011 IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Aachen, Germany, May 2011, pp. 265–276.

[4] G. Dimitrakopoulos and P. Demestichas, "Intelligent transportation systems," *IEEE Vehicular Technology Magazine*, vol. 5, no. 1, pp. 77–84, Mar. 2010.

[5] P. Taylor, "Evaluating telemedicine systems and services," *Journal of Telemedicine and Telecare*, vol. 11, no. 4, pp. 167–177, Apr. 2005.

[6] I. Soto, C. J. Bernardos, M. Calderon, A. Banchs, and A. Azcorra, "NEMO-enabled localized mobility support for internet access in automotive scenarios," *Communications Magazine, IEEE*, vol. 47, no. 5, pp. 152–159, May 2009.

[7] E. Bjornson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *2013 20th International Conference on Telecommunications*, Casablanca, Morocco, Oct. 2013.

[8] C. C. Coskun and E. Ayanoglu, "A greedy algorithm for energy-efficient base station deployment in heterogeneous networks," in *2015 IEEE International Conference on Communications ICC*, London, UK, May 2015, pp. 7–12.

[9] R. Torah, P. Glynne-Jones, M. Tudor, T. O'Donnell, S. Roy, and S. Beeby, "Self-powered autonomous wireless sensor node using vibration energy harvesting," *Measurement Science and Technology*, vol. 19, no. 12, pp. 202–233, Oct. 2008.

[10] A. Kansal and M. B. Srivastava, "An environmental energy harvesting framework for sensor networks," in *The 2003 International Symposium on Low Power Electronics and Design*, Seoul, Korea, Aug. 2003, pp. 481–486.

[11] C. J. Chang, T. L. Tsai, and Y. H. Chen, "Utility and game-theory based network selection scheme in heterogeneous wireless networks," in *IEEE Wireless Communications and Networking Conference*, Budapest, Hungary, Apr. 2009.

[12] R. Trestian, O. Ormond, and G.-M. Muntean, "Game theory-based network selection: solutions and challenges," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1212–1231, Feb. 2012.

[13] C. Yuen and P. Marbach, "Price-based rate control in random access networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 5, pp. 1027–1040, Oct. 2005.

[14] Q. Song and A. Jamalipour, "A network selection mechanism for next generation networks," in *The 2005 IEEE International Conference on Communications*, vol. 2, Seoul, Korea, May 2005, pp. 1418–1422.

[15] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[16] J. C. Chen, T. C. Chen, T. Zhang, and E. Van den Berg, "Wlc19-4: Effective AP selection and load balancing in IEEE 802.11 wireless lans," in *The 2006 IEEE Conference on Global Telecommunications Conference*, San Francisco, CA, Nov. 2006.

[17] C. C. Coskun and E. Ayanoglu, "Energy-Efficient base station deployment in heterogeneous networks," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 593–596, Oct. 2014.

[18] K. Davaslioglu and E. Ayanoglu, "Quantifying potential energy efficiency gain in green cellular wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2065–2091, Sep. 2014.

[19] "Database of solar radiation," http://solar.uq.edu.au/user/reportPower.php.

[20] M. Tacca, P. Monti, and A. Fumagalli, "Cooperative and reliable ARQ protocols for energy harvesting wireless sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2519–2529, Jul. 2007.

[21] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization," *IEEE Transactions on Mobile Computing*, vol. 6, no. 2, pp. 221–236, Feb. 2007.

[22] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 4, p. 32, Sep. 2007.

[23] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 547–551, Feb. 2009.

[24] A. Aprem, C. R. Murthy, and N. B. Mehta, "Transmit power control policies for energy harvesting sensors with retransmissions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 895–906, Oct. 2013.

[25] T. D. Todd, A. Sayegh, M. N. Smadi, D. Zhao *et al.*, "The need for access point power saving in solar powered WLAN mesh networks," *IEEE Network*, vol. 22, no. 3, pp. 4–10, May 2008.

[26] Y. Xu, A. Anpalagan, Q. Wu, L. Shen, Z. Gao, and J. Wang, "Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1689–1713, Apr. 2013.

[27] S. Yin, D. Chen, Q. Zhang, and S. Li, "Prediction-based throughput optimization for dynamic spectrum access," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1284–1289, Mar. 2011.

[28] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 785–796, Feb. 2008.

[29] X. Li, Q. Zhao, X. Guan, and L. Tong, "Optimal cognitive access of Markovian channels under tight collision constraints," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 746–756, Apr. 2011.

[30] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2053–2071, May 2008.

[31] J. Neel, R. M. Buehrer, J. H. Reed, and R. P. Gilles, "Game theoretic analysis of a network of cognitive radios," in *The 45th Midwest Symposium on Circuits and Systems*, Tulsa, Oklahoma, Aug. 2002, pp. 409–412.

[32] Y.-H. Yang, Y. Chen, C. Jiang, C.-Y. Wang, and K. R. Liu, "Wireless access network selection game with negative network externality," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 5048–5060, Sep. 2013.

[33] D. Niyato and E. Hossain, "A game theoretic analysis of service competition and pricing in heterogeneous wireless access networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5150–5155, Dec. 2008.

[34] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[35] H.-Y. Chen, J. Hou, S. Zhang, Y. Liang, G. Yang, Y. Yang, L. Yu, Y. Wu, and G. Li, "Polymer solar cells with enhanced open-circuit voltage and efficiency," *Nature Photonics*, vol. 3, no. 11, pp. 649–653, Oct. 2009.

[36] S. M. Kay, "Fundamentals of statistical signal processing, Volume.I: Estimation theory," *Signal Processing. Upper Saddle River, NJ: Prentice Hall*, 1998.

[37] A. L. Strehl and M. L. Littman, "An empirical evaluation of interval estimation for markov decision processes," in *The 16th IEEE International Conference on Tools with Artificial Intelligence*, Arlington, Virginia, Nov. 2004, pp. 128–135.

[38] A. R. Cassandra, "Optimal policies for partially observable Markov decision processes," Report CS-94-14, Brown Univ, Tech. Rep., 1994.

[39] M. Kosfeld, E. Droste, and M. Voorneveld, "A myopic adjustment process leading to best-reply matching," *Games and Economic Behavior*, vol. 40, no. 2, pp. 270–298, Aug. 2002.

[40] H. R. Maier, T. Sayed, and B. J. Lence, "Forecasting cyanobacterium anabaena spp. in the river murray, south australia, using b-spline neurofuzzy models," *Ecological Modelling*, vol. 146, no. 1, pp. 85–96, Dec. 2001.

[41] M.-H. Chang, J.-Y. Wu, W.-C. Hsieh, S.-Y. Lin, Y.-W. Liang, and H. Wei, "High efficiency power management system for solar energy harvesting applications," in *2010 IEEE Asia Pacific Conference on Circuits and Systems*, Kuala Lumpur, Malaysia, Dec. 2010, pp. 879–882.