IEEE TRANSACTIONS ON
EMERGING TOPICS
IN COMPUTING

Received 17 April 2013; revised 24 November 2013 and 12 February 2014; accepted 19 February 2014. Date of publication 13 April 2014; date of current version 30 July 2014.

Digital Object Identifier 10.1109/TETC.2014.2316518

A Reinforcement Learning-Based ToD Provisioning Dynamic Power Management for Sustainable Operation of Energy Harvesting Wireless Sensor Node

ROY CHAOMING HSU[1], (Member, IEEE), CHENG-TING LIU[2], AND HAO-LI WANG[2]

[1]Department of Electrical Engineering, National Chiayi University, Chiayi City 60004, Taiwan
[2]Department of Computer Science and Information Engineering, National Chiayi University, Chiayi City 60004, Taiwan
CORRESPONDING AUTHOR: H. WANG (haoli@mail.ncyu.edu.tw)

This work was supported by the National Science Council of Taiwan under Grant NSC-99-2221-E-415-013 and NSC-98-2221-E-415-011.

ABSTRACT In this paper, a reinforcement learning-based throughput on demand (ToD) provisioning dynamic power management method (RLTDPM) is proposed for sustaining perpetual operation and satisfying the ToD requirements for today's energy harvesting wireless sensor node (EHWSN). The RLTDPM monitors the environmental state of the EHWS and adjusts their operational duty cycle under criteria of energy neutrality to meet the demanded throughput. Outcomes of these observation-adjustment interactions are then evaluated by feedback/reward that represents how well the ToD requests are met; subsequently, the observation-adjustment-evaluation process, so-called reinforcement learning, continues. After the learning process, the RLTDPM is able to autonomously adjust the duty cycle for satisfying the ToD requirement, and in doing so, sustain the perpetual operation of the EHWSN. Simulations of the proposed RLTDPM on a wireless sensor node powered by a battery and solar cell for image sensing tasks were performed. Experimental results demonstrate that the achieved demanded throughput is improved 10.7% for the most stringent ToD requirement, while the residual battery energy of the RLTDPM is improved 7.4% compared with an existing DPM algorithm for EHWSN with image sensing purpose.

INDEX TERMS Reinforcement learning, wireless sensor node, energy harvesting, energy neutrality, dynamic power management, throughput on demand.

I. INTRODUCTION

POWER-aware computing for the sustained operation of handheld devices has become one of the primary concerns of embedded systems design due to the increasing popularity of battery-powered portable and wearable communication devices. Dynamic power management (DPM), an energy utilization technique [1]–[3] for adaptively controlling components' power states by trading off performance while reducing energy consumption, is considered critical in power-aware embedded system design. Today, DPM for embedded systems powered by a renewable energy source, such as energy harvesting wireless sensor node (EHWSN), presents yet another challenge not only for power-aware design, but also for sustainable operation for embedded systems. Recent studies of DPM for energy harvesting embedded systems [5]–[10] have concentrated on how to adaptively and effectively utilize harvested energy for perpetual operation while maximizing system performance under such a nondeterministic energy harvesting environment. In [5] and [6], Kansal et al. proposed a harvesting theory for energy harvesting embedded systems, especially for wireless sensor nodes, and suggested that under conditions of energy neutrality, i.e., the consumed energy is less than or equal to the harvested energy, perpetual operation of a sensor node could be achieved. A dynamic duty cycling adaption (DDCA) method [7]

was then proposed to decrease the duty cycle of the sensor node in times of low harvested energy, and increase the duty cycle when the harvested energy is high for energy neutral operation. An approximate multi-parametric programming algorithm was proposed in [8] for the adaptive power management of an energy harvesting embedded system. Similar to computer system design, reliability [9], [10] and quality of service (QoS) [11]–[14] are the two major concerns in designing power management for embedded systems, such as a WSN. Basically, meeting QoS requirements and reducing energy consumption are opposite in nature; most DPM methods have emphasized minimizing energy consumption or maximizing system lifetime in satisfying certain required QoS criteria, such as throughput. Because many different applications in WSN exist, their QoS requirements will vary [15]. In [16], the authors suggested using measurements of throughput, latency, reliability, security, adaptability, and affordability to define QoS for wireless industrial sensor networks. In the present study, the QoS requirement of concern is the query-driven throughput-on-demand (ToD) from the sink node for a sensor node. In other words, we investigate the data volume a sensor node collects and transmits to the sink node within a sensing period, which results from the exercised duty cycle of the sensor node within a predetermined sensing period.

Research into applying intelligent and machine learning methods for power management has only been considered recently [17], [18], with our earlier studies [19]–[21] being among those specifically targeting the area of DPM for energy harvesting embedded systems. In this study, a reinforcement learning [22] (RL)-based ToD provisioning dynamic power management method, named RLTDPM, is proposed for the perpetual operation of an EHWSN. The framework of our study is analogous to the MDP in [14]; however, the problem affecting ToD provision power management is solved using on-line reinforcement learning instead of dynamic programming. In the RL framework, a learner, referred to as an agent, interacts with the environment and autonomously determines required actions; and then, is rewarded by the reward function to respond to different environmental states. In the RLTDPM, after observing the environmental states of the harvesting system, the RLTDPM agent determines the operational duty cycle of a sensor node and receives a reward by the rewarding function. By receiving rewards, the RLTDPM is encouraged to select (state, action) pairs with positive rewards; hence, a series of actions, receiving with positive rewards, is generated iteratively such that a (state, action) pair selection strategy is gradually achieved after the learning phase. Via the RLTDPM, the EHWSN is capable of autonomously adjusting the appropriate duty cycle to simultaneously maintain energy neutrality and satisfy the given ToD requirement for perpetual operation of the EHWSN.

This paper extends our earlier studies [20], [21] with formally derived rewarding functions designed specifically for EHWSN, a refined occurrence pattern for the ToD requirement, and comprehensive simulations. The major contributions of this paper are:

- A reinforcement learning-based ToD provisioning dynamic power management method, RLTDPM, for sustaining perpetual operation of EHWSN, which is novel in the area of sustainable computing for EHWSN.
- Reward functions depending on the degree of energy neutrality and current battery storage levels with sigmoid and Mexican hat shapes, enabling autonomous learning of RLTDPM to maintain energy neutrality for sustaining perpetual operation of EHWSN.
- RLTDPM not only achieves energy neutrality, but also ToD provisioning by incorporating the correlation between the required and provided throughput in designing the reward strategy to meet the ToD requirements in data gathering.
- Comprehensive experimental results exhibit the advantage of the RLTDPM in self-learning and self-adapting to seasonal changes of the harvested energy for sustaining perpetual operation of EHWSN.

II. SYSTEM MODEL

In this paper, dynamic power management for EHWSN is approached by first defining an appropriate system model and introducing the energy neutrality theory; then, the ToD provisioning dynamic power management method is formulated for EHWSN.

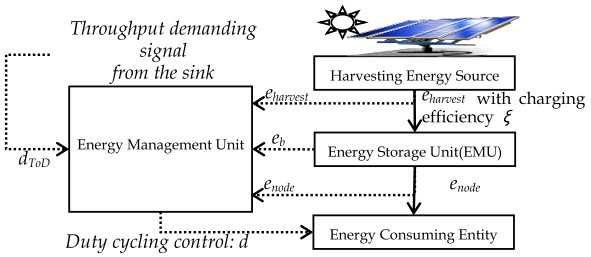

FIGURE 1. System model of ToD provisioning power management method for the energy harvesting sensor node.

A. SYSTEM MODEL OF EHWSN

Figure 1 shows the system model of the EHWSN. On the right side of the figure is the hardware layer, which consists of the harvesting energy source, energy storage unit, and energy consumption entity. In this model, an uncontrolled but predictable solar energy source is considered. Although the solar energy is uncontrolled, its behavior is subject to diurnal and seasonal cycles that can be modeled. The energy storage unit is charged by the harvesting energy source with a charging efficiency ξ, where $0 < \xi < 1$, with some energy lost through leakage. The energy consuming entity, the sensor node, consumes different levels of energy from the energy storage unit by exercising its demanded mode of sensing operations and transmitting the sensed data per the ToD request from the sink node. In a typical sensor node, the sensing and transmission of sensed data each consume far more

energy than other operation modes such as standby, receiving demand, and acknowledgement to the sink node. In this study, energy consumed from both sensing and transmitting the sensed data is considered as the dominating activity of the energy consuming entity, and is duty cycle adjustable with upper and lower bounds. The energy consumed from standby, receiving demands and acknowledgement to the sink node is accommodated in the lower limit of the duty cycle. The energy management unit (EMU), shown in the left part of Fig. 1, receives observable environment signals of the sensor node's energy consumption, energy provided by the energy harvesting source, residual energy of the energy storage unit, and the requested ToD signal, which are respectively denoted as e_{node}, e_{harvest}, e_{b}, and d_{ToD}, and shown in dashed lines. Moreover, the EMU decides the duty cycle of the energy consuming entity according to its management strategies.

B. THE ENERGY-NEUTRAL OPERATION OF EHWSN

In contrast with a battery-powered sensor node, when the battery is exhausted in an EHWSN, it will operate again at the next energy harvesting opportunity. Hence, the concept of energy-neutral operation is addressed for situations in which the energy consumed by the EHWSN system is less than the energy harvested from the environment [7]. Consider a harvesting system where the energy harvesting and energy consuming profiles, respectively, are characterized by $(\rho_1, \sigma_1, \sigma_2)$ and (ρ_2, σ_3) functions, while the energy storage is characterized by charging efficiency, ξ, and leakage ρ_{leak},. In such a case, the following conditions are sufficient for the system to achieve energy-neutrality [7]:

$$\rho_2 \leq \xi\rho_1 - \rho_{\text{leak}} \tag{1}$$

$$B_0 \geq \xi\sigma_2 + \sigma_3 \tag{2}$$

$$B \geq B_0 \tag{3}$$

where ρ_1 and ρ_2, are the average rates of the entity's energy source harvesting and energy consumption, respectively, over long durations; the burstiness of harvesting energy is bounded by σ_1 and σ_2, where σ_3 is the lower bound of the burstiness of the energy consuming entity; and, B and B_0, respectively denote the capacity of and the initial energy stored in the energy storage. According to [7], to ensure energy neutral operation, the first step is to obtain the parameters characterizing the energy harvesting and energy consumption of the entity such that the performance levels for perpetual operation can be determined using (1)–(3). The energy management scheme then attempts to adjust the performance level, i.e., ρ_2, to respond to temporal variations in harvested energy and minimize energy wastes, such as charge inefficiency and leakage. In this study, without knowing the exact parameters characterizing the harvesting energy source and the energy consuming entity, the proposed RLTDPM with designed reward function learns autonomously from the observable environment variables and adjusts the duty cycle of the energy consuming entity to respond to temporal variations in realizing sustainable operation of the EHWSN.

III. THE PROPOSED METHOD

Suppose the time axis is discretized into slots of duration ΔT, and the energy management strategy is carried out over a window of T time slots. The observable environment variables are defined with the index i ranging over $\{1, \ldots, T\}$ as follows:

- $d(i)$: the controlled and exercised duty cycle used in slot i for the sensor node, value of which is to be controlled and determined by the DPM strategy.
- $e_{\text{node}}(i)$: energy consumption of the sensor node in the i^{th} slot. Without loss of generality, $e_{\text{node}}(i)$ can be written as

$$e_{\text{node}}(i) = d(i)Es \tag{4}$$

  where Es is the energy consumed for full duty cycle.
- $e_{\text{harvest}}(i)$: the energy supplied by the energy harvesting source in slot i. Total energy within T sensing time slots represents the accumulation of energy at each time slot as

$$e_{\text{totalHarvest}} = \sum_{\forall i \in T} e_{\text{harvest}}(i) \tag{5}$$

- $\boldsymbol{\Delta} e_{\text{neutral}}(i)$: the deviation from energy neutrality in slot i. $\boldsymbol{\Delta} e_{\text{neutral}}(i)$ can be used in evaluating the degree to which energy neutrality is achieved, and can be obtained by

$$\Delta e_{\text{neutral}}(i) = e_{\text{harvest}}(i) - e_{\text{node}}(i) \tag{6}$$

- $e_{\text{b}}(i)$: the residual energy of the energy storage unit in slot i. $e_{\text{b}}(i)$ can be calculated by

$$\begin{aligned} e_{\text{b}}(i) &= e_{\text{b}}(i-1) + e_{\text{harvest}}(i) - e_{\text{node}}(i) \\ &= e_{\text{b}}(i-1) + \Delta e_{\text{neutral}}(i) \end{aligned} \tag{7}$$

  where $e_{\text{b}}(i-1)$ is the residual energy of the $(i-1)$ time slot.
- $d_{\text{ToD}}(i)$: the requested ToD signal, which, depending on the sensor's application, is defined by different operation levels as

$$d_{ToD}(i) \in Q = \{q_0, q_1, \ldots, q_{n-1}\} \tag{8}$$

  where Q is the space of the required ToD levels, for all $q_i, q_j \in Q$, $q_i < q_j$ for $i < j$, and, no ToD is required when $q_0 = 0$.

As mentioned, the EMU receives the observable environment variables of $e_{\text{node}}(i)$, $e_{\text{harvest}}(i)$, $e_{\text{b}}(i)$, and $d_{\text{ToD}}(i)$ at each sensing period and decides the operational duty cycle, $d(i)$, to be exercised according to the RLTDPM, as described in the following.

A. REINFORCEMENT LEARNING

Reinforcement learning (RL) is a heuristic learning method [22] that has been applied in many different areas, such as agent society systems, power management for embedded systems [20], robot control [23]–[25], and image processing [26], etc. In RL, there is a decision-making agent, termed the RL agent, which observes the environment states, $\boldsymbol{S}$, takes actions, $\boldsymbol{A}$, and receives rewards, $\boldsymbol{r}$, for its actions in trying to

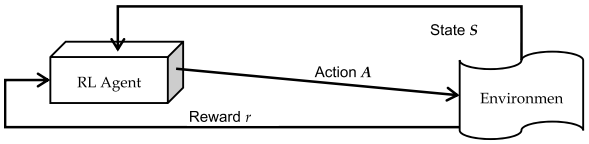

FIGURE 2. Diagram of RL agent's interaction with the environment.

solve a problem in an environment, as shown in Fig. 2. After a certain amount of trial-and-error steps, the RL agent learns the best policy, which is the sequence of actions that maximizes the total reward. For unknown environments, the RL agent continuously counts rewards during the early learning phase, with the optimized policy being obtained when learned. In the beginning of the learning process, the RL agent is unable to decide the most appropriate action; as such, an exploration strategy is used to select the action that provides the highest reward for each state. After the agent has learned, an exploitation strategy is then adopted for deciding the action offering the highest reward and probability on certain states. For action selection, the *soft-max* function [22] is often utilized as the exploration-exploitation strategy in most applications. In the case of a nondeterministic environment or state transition model, where the transition probability is unknown or uncertain, a particular action for the next state cannot be decided precisely by observing the current state. In such a case, the Q-learning algorithm is frequently utilized to calculate the accumulative reward and decide the best policy [22]. In Q-learning, the accumulative reward, $Q(s, a)$, is a function of state, s, and action, a, and the agent iteratively updates the Q-values of $Q(s, a)$ using the equation below,

$$Q(s_t, a_t) := (1-\eta)Q(s_t, a_t) + \eta\left[r_{t+1} + \gamma \max_{\forall a_{t+1} \in A} Q(s_{t+1}, a_{t+1})\right] \tag{9}$$

where $Q(s_t, a_t)$ is the accumulative reward at state s and an action a taken at time t; r_{t+1} is the reward obtained by taking action a_t and then making the transition from state s_t to state s_{t+1}; and, the parameters η and γ, respectively, are the learning factor and discount rate with values between 0 and 1. The learning factor, η, as normally used in the delta rule, controls the convergence speed of the learning, the value of which is gradually decreased in time for the best convergence. The discount rate, γ, is used to weight near-term rewards more heavily than distant future rewards. More specifically, the closer γ is to 1, the greater the weight of future rewards; however, if γ is 0, only the immediate reward counts. As normal, the Q-values are stored in a Q-table as a reference for the agent to determine the next action.

B. THE RLTDPM FOR THE EHWSN

In the RLTDPM, the agent receives the observable environmental signals of $e_{\text{node}}(i)$, $e_{\text{harvest}}(i)$, $e_{\text{b}}(i)$ and $d_{\text{ToD}}(i)$, formulated as a *state* vector, and adaptively decides and executes the desired operational duty cycle $d(i)$, characterized as *action*. After the selected action is executed, a reward signal, $r(i)$, is calculated and granted to the agent; and accordingly, the agent then evaluates the performance of the state-action interaction, i.e., *learning*. By receiving rewards, the agent is encouraged to select the action with the best reward. This leads to a series of actions with the best rewards being iteratively generated such that better power management performance is gradually achieved after the learning phase. The state, action, and reward of the RLTDPM for the energy harvesting sensor node are sequentially defined in the following:

***States*:** In RLTDPM, the state vector is denoted as

$$S = [S_{\text{D}}, S_{\text{H}}, S_{\text{B}}, S_{\text{ToD}}] \subseteq \boldsymbol{S} \tag{10}$$

where $\boldsymbol{S}$ is the space of all possible environmental state vectors with elements transformed from the environment variables, while S_{D}, S_{H}, S_{B}, and S_{ToD} respectively represent the state of distance to energy neutrality, the state of harvested energy, the state of energy storage level, and the state of the required ToD of the sensor node at any time slot. S_{D} is defined as the difference between $e_{\text{harvest}}(i)$ and $e_{\text{node}}(i)$, as below

$$S_{\text{D}} \in \{\Delta e_{\text{neutral}}(i) | \Delta e_{\text{neutral}}(i) = e_{\text{harvest}}(i) - e_{\text{node}}(i), 1 \le i \le T\} \tag{11}$$

S_{H} is obtained by normalizing the harvested energy with the maximum harvested energy, calculated by

$$S_{\text{H}} \in \{S_{\text{h}}(i) | S_{\text{h}}(i) = \frac{e_{\text{harvest}}(i)}{e_{\text{maxharvest}}} \times 100\%, \quad 1 \le i \le T\} \tag{12}$$

where $e_{\text{maxharvest}}$ is the maximum harvested energy according to [10]. S_{B}, the state of current energy storage, is defined by normalizing the value of the current energy storage with the maximum capacity of the energy storage, denoted as E_{B}, as

$$S_{\text{B}} \in \left\{S_{\text{b}}(i) | S_{\text{b}}(i) = \frac{e_{\text{b}}(i)}{E_{\text{B}}} \times 100\%, \quad 1 \le i \le T\right\} \tag{13}$$

The requested ToD of the sensor node is usually demanded by the sink node [16]. The state of the requested ToD, S_{ToD}, is defined according to different degrees of required throughput, as the following

$$S_{\text{ToD}} \in Q(i) = \left\{q_0, q_1, \ldots, q_j, \ldots, q_{n-1}\right\} \tag{14}$$

where $Q(i)$ is the requested ToD in any time slot i, and, with a larger j, q_j means that a higher throughput is required.

***Actions*:** The action of the RLTDPM agent is defined as the controllable variable of operational duty cycle, $d(i)$, of the EHWSN, and the state of the agent's action, A_{dc}, is denoted by

$$A_{\text{dc}} \in A = \{d_0, d_1, \ldots, d_{\max}\} \tag{15}$$

where $\boldsymbol{A}$ is a set of all the sensor node's controllable operational duty cycles with a maximum value of $d_{\max}$. Hence, an action with a higher value means a higher operational duty cycle is assigned to the sensor node to satisfy the ToD requirement; in such a case, more energy is consumed during data sensing.

***Rewards*:** In RLTDPM, rewards are incorporated to accomplish the goals of maintaining energy neutrality and autonomously satisfying the EHWSN's ToD requirement. For a sensor node without an energy storage unit, the deviation from energy neutrality, $\Delta e_{\text{neutral}}(i)$, should be greater than or equal to 0 to maintain energy neutral operation. For a sensor node with an energy storage unit, the immediate reward, denoted as r, can be intuitively defined as a function of $\Delta e_{\text{neutral}}(i)$ and the current energy storage level $e_{\text{b}}(i)/E_{\text{B}}$ in any time slot, as in the following

$$r_{i+1} = \left(1 - 2\frac{e_{\text{b}}(i)}{E_{\text{B}}}\right) \cdot \frac{\Delta e_{\text{neutral}}(i)}{e_{\text{maxharvest}}} \tag{16}$$

where E_{B} indicates the energy storage capacity. $\Delta e_{\text{neutral}}(i)$ should be large and positive when the residual energy in the storage unit is low, i.e., smaller $e_{\text{b}}(i)/E_{\text{B}}$, such that $\Delta e_{\text{neutral}}(i)$ can be transferred to the storage unit for later use. On the other hand, when $e_{\text{b}}(i)/E_{\text{B}}$ is very high, $\Delta e_{\text{neutral}}(i)$ should be large and negative to prevent overcharging the energy storage unit. However, (16) is a nonlinear function of $\Delta e_{\text{neutral}}(i)$ and $e_{\text{b}}(i)/E_{\text{B}}$ with singularity at $e_{\text{b}}(i) = 0.5\ E_{\text{B}}$. Hence, to overcome the reward function shortcomings of (16), the sigmoid and Mexican hat type of reward functions are defined with dependence on $\Delta e_{\text{neutral}}(i)$ for high, low, and moderate energy storage levels, respectively, as below

$$r_{i+1} = \begin{cases} k_1\left(\frac{2}{1+e^{-k_2\Delta e_{\text{neutral}}(i)}} - 1\right), & \frac{e_{\text{b}}(i)}{E_{\text{B}}} \in \text{low state} \\ k_1\left(\frac{2}{1+e^{k_2\Delta e_{\text{neutral}}(i)}} - 1\right), & \frac{e_{\text{b}}(i)}{E_{\text{B}}} \in \text{high state} \\ k_1\left(\frac{8}{(1+e^{-k_2\Delta e_{\text{neutral}}(i)})(1+e^{k_2\Delta e_{\text{neutral}}(i)})} - 1\right), & \text{otherwise} \end{cases} \tag{17}$$

where k_1 and k_2, respectively control the amplitude and slope of the sigmoid and Mexican hat functions. The designed reward functions can be seen in Fig. 3, where k_1 and k_2 are assigned with 2 and 2.20. In (17), a large reward is granted to larger positive/negative $\Delta e_{\text{neutral}}(i)$ when the energy storage level is correspondingly low/high to maintain energy neutrality; however, in the opposite case, where larger positive/negative $\Delta e_{\text{neutral}}(i)$ occurs when the energy storage level is high/low, a punishment is applied to avoid overcharging/deep discharge. On the other hand, when the energy storage level is moderate, rewarding should gradually reach a maximum when $\Delta e_{\text{neutral}}(i)$ is close to zero.

***Rewards with requested ToD*:** To satisfy the ToD requirements, a positive reward is granted to the RLTDPM agent if the agent's action $d(i)$ meets $Q(i)$ in any time slot; otherwise, no reward is given. A simple reward function could be intuitively defined as

$$r_{i+1} = \delta(d(i), Q(i)) \tag{18}$$

where $\delta(\bullet)$ is the Kronecker delta function. However, it is very rare for the agent's action to be exactly the same as the required throughput in any time slot; as such, to obtain better learning results, a negative reward might explicitly represent punishment compared with a zero reward for the agent's failure in meeting the ToD requirement. Hence, (18) can be relaxed as the required ToD is met if $d(i) \geqq Q(i)$ in any time slot and generalized by utilizing negative rewards when these ToD requirements are not met, as below

$$r_{i+1} = \begin{cases} k_3\delta(d_i, q_i), & d_i \subseteq q_i \\ -k_4, & \text{otherwise} \end{cases} \tag{19}$$

where k_3 and k_4, respectively denote weighting factors in maximizing the difference between reward and punishment for better learning. When no ToD requirement is presented for the sensor node in any time slot, that is $S_{\text{ToD}} = q_0 = 0$, the reward function defined in (17) for energy neutral operation is used instead. Alternatively, the reward function in satisfying a ToD request depends on the degree of correlation between the required ToD request and the agent's action, where a positive reward is granted when the agent's action meets the required ToD level, i.e., positive correlation, while negative rewards in monotonic decreasing fashion are granted otherwise. In the RLTDPM experiment, to decrease the computational complexity of on-line learning for an EHWSN, the reward functions of (19) are simplified with efficient lookup tables. The reward table for the experiment is elaborated in V.

In this study, solar energy is used as the energy harvesting source, the harvested energy strength of which fluctuates according to the sun's position on a diurnal basis. Hence, a run for the execution of the RLTDPM is defined as the duration of one day. The algorithm for the RLTDPM is shown in Fig. 4.

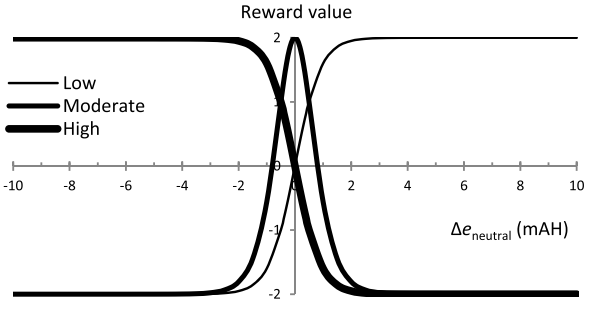

FIGURE 3. Designed reward function for the RLTDPM.

1 **initialize** states and actions according to the given mapping range
2 **initialize** Q table as an 2-D array $Q(.)$ with size $|S|\times|A|$
3 **initialize** iteration count i:=*0*, all Q value in $Q(s,a)$:= 0,
discount rate γ:=*0.8*, learning rate η:=*0.5*,
maximum time slot number *MTS*:=*23*
4 **Start** a run by **observing** the first state s_0
5 **while** i<=*MTS* **do**
6 **select** duty-cycle as action a according to Q table **using** Soft-max function
7 **adjust** duty-cycle to $A_{dc}(a)$ for this slot i
8 **set** $i := i + 1$, for next time slot
9 **assess** r' and **observe** state s'
10 **update** Q value by Eq. 9
11 **end** *while* loop

FIGURE 4. The RLTDPM algorithm for the EHWSN.

IV. CONFIGURATIONS OF EXPERIMENT

Validation of RLTDPM effectiveness in maintaining energy neutrality for sustaining perpetual operation and satisfying the ToD requirements of EHWSN was conducted via two computer simulated experiments. A basic experiment was performed to determine whether RLTDPM could achieve the energy neutrality criterion; then, an advanced experiment examined whether RLTDPM satisfies the different ToD requirements. Configurations of the experiments are described in the following.

A. ENVIRONMENT CONFIGURATION FOR EHWSN

In this study, the available harvested energy source is the sun, and the intensity of sunlight and the solar energy converted from the sunlight, are respectively calculated using the incidence and azimuth angles of sunlight [27], [28] and the ideal solar energy equation [29]. The intensity of sunlight from sunrise to sunset diminishes according to a normal distribution of 0.5 mean and 0.1 variance on a diurnal basis, i.e., with the peak value at noon, while the probability of sun occultation is assumed to be 0.25 to accommodate weather conditions such as occasional cloudy and seasonal rainy days, etc. Fig. 5 shows harvested power generated from 4 BR-160334C solar cells [30] in parallel. In particular, Fig. 5(a) shows the recorded solar cell output for 7 days, starting from the beginning of the spring equinox. Fig. 5(b) shows the recorded solar cell daily output for 90 days over 4 seasons, which started from the beginning of the spring equinox, the summer solstice, the autumn equinox and the winter solstice. The average harvested power, ρ_1, within the seven days can be calculated from Fig. 5(a), and found to be about 150mW. The total harvested energy of a day, σ_2, is obtained using ρ_1 as 12,960J. The sensor node adopted in this study was the MICAz mote [31], which operates at 3.4V with an image sensor [32] and consumes a maximum power of 396mW including the accommodated overhead, such as CPU power on/off, dissipation, etc. The initial battery capacity, B_0, can hence be calculated by (1), (2), and (3), and yields 1100mAH. Taking leakage, and battery aging and degradation from multiple charge-discharge cycles into consideration, a full battery capacity of 2000mAH was adopted; hence, three 2000mAH AA-sized NiMH rechargeable batteries in series were provided for the EHWSN with image sensing purpose.

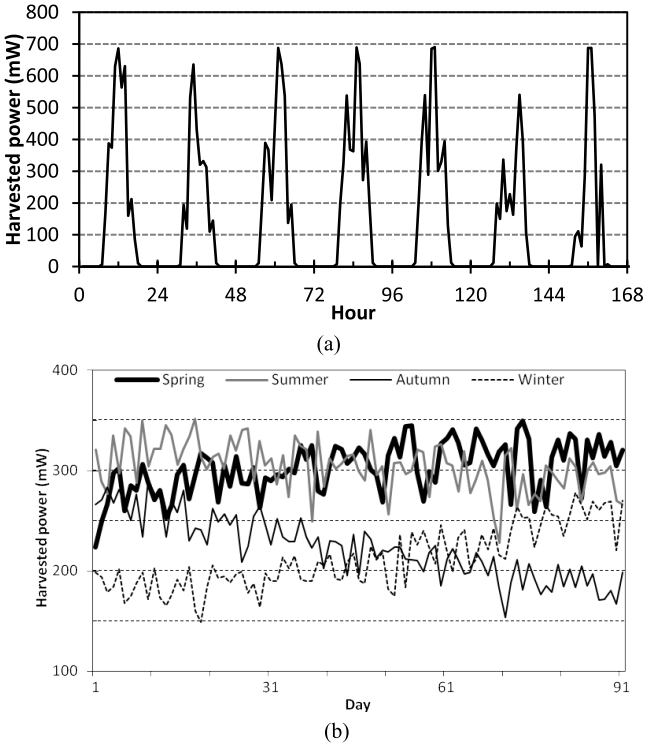

FIGURE 5. Graphs of harvested power from the solar cells: (a) for the first seven days; (b) daily average over 4 seasons.

B. CONFIGURATION FOR THE BASIC AND ADVANCED EXPERIMENT

For the basic experiment, the environment state of distance to energy neutrality, S_D, is defined as being either in a negative, neutral, or positive state, which respectively represents that the consumed energy is larger than, equal to, or smaller than the harvested energy. The state of harvested power, S_H, comprises three states, with each respectively representing that the harvested power is in the ranges of 0–70mW, 70–315mW, or 315–700mW. The state of current battery energy, S_B, consists of low, moderate, and high states, which respectively represent 0–40%, 40–60% and 60–100% of full storage capacity. Four actions of the agent's exercised duty cycle are defined: 0%, 35%, 60%, and 100% sensor node duty cycle with the corresponding power consumed by the sensing node being 0mW, 139mW, 238mW, and 396mW, respectively. The basic experiment's reward rule for the RLTDPM is defined in Table 1, which is a simplified form of the reward function in (17) by applying a hard-limiting operation to the waveform in Fig. 3. Taking the moderate residual battery energy level as an example, when the deviation from energy neutrality is around 0, a reward value of 2 is provided, while a negative reward value of -2 is given for other cases. Table 1 also shows the different rewards given to the different ranges of deviation from energy neutrality under different states of current battery energy. For example, a higher positive distance to energy neutrality has no benefit in maintaining neutrality if the current battery energy is high, as in the second row, which also

TABLE 1. Reward rule of RLTDPM without ToD.

Distance toEnergy Neutrality / Battery State	<-0.5mAH	-0.5~0.5mAH	>0.5mAH
High	2	0	-2
Moderate	-2	2	-2
Low	-2	0	2

may lead to overcharging and damage the battery. Hence, a positive deviation from energy neutrality that continues to widen obtains a negative reward when a high state of current battery energy exists. On the other hand, at a high storage state, a negative distance to energy neutrality should be encouraged with positive rewards to increase the duty cycle and consume more power.

The advanced experiment is conducted to test whether the RLTDPM can maintain energy neural operation and satisfy the ToD requirement simultaneously. In a query-driven wireless sensor network, the ToD requirement for sensor node data sensing is usually requested by the base station or the sink node [20]. For an EHWSN with an image sensing task, such as industrial streaming surveillance video, different throughputs might be required from the sink node under different circumstances. For instance, in any time slot, the sink node might request the EHWSN to collect data with a throughput of 50%; as a consequence, the node will operate the image sensor at a 50% duty-cycle to accomplish the image sensing task under the requested ToD. In this manner, the RLTDPM shall be able to realize the goal of energy neutrality for the perpetual operation of an EHWSN and concurrently satisfies the requested ToD. In this study, the state of required ToD level, S_{ToD}, is defined as

$$S_{\text{ToD}} \in Q = \{low, medium, high\} \tag{20}$$

where a low, medium, or high state of required ToD level in a given time slot represents the requested sensing duty cycle falling within the ranges of 5%–40%, 40%–70%, and 70%–100%, respectively. For experiment purposes, three occurrence patterns of the requested ToD – sparse, moderate, and intensive – are designed and generated according to Table 2 as below

TABLE 2. Occurrence patterns of the required throughput.

Occurrence Pattern	Sparse	Moderate	Intensive
Next required ToD occurring time / Prob. of the required duty cycle	N (8,1.5)	N (6, 1.5)	N (4,1.5)
P_{Low}	0.7	0.2	0.2
P_{Medium}	0.2	0.6	0.3
P_{High}	0.1	0.2	0.5

The occurrence pattern of the next required ToD is determined by the normally distributed expected next ToD-required time slot with a combination of the probabilities of three different required ToD durations, namely, low, medium, and high. Whenever the requested low, medium, or high required ToD duration is presented, the sensor node is required to operate at a duty cycle at or higher than the required ToD for the upcoming one, two, and three time slots, respectively. As an example, the sparse occurrence pattern in the second column of Table 2 means that the next ToD-required time slot will arrive on average after 8 time slots with a variance of 1.5 time slots, and that the probability of requesting low, medium, or high required ToD durations is 0.7, 0.2, and 0.1, respectively. The state of exercised duty cycle, i.e., the action of the RLTDPM agent, is defined as zero, low, moderate, or high, which represents that the exercised duty cycle falls within ranges of 0%–5%, 5%–40%, 40%–70%, and 70%–100%, respectively. The reward rule for the RLTDPM, in conducting the advanced experiment with required ToD, is defined in Table 3. This reward rule is an extension of the reward rule of (19) and takes into consideration the degree of correlation between the exercised duty cycle and the required ToD level. The reward rule gives a positive reward when the RLTDPM exercised duty cycle matches the required ToD level, i.e., with positive correlation. For instance, in the second row of Table 3, when the required ToD level is low in a time slot, if the exercised duty cycle is low, which matches the required ToD, a reward of +5 is given to the agent, while for the others, smaller positive (+2) and negative (−2, and −4) values are given to the agent according to the state of the exercised duty cycle. If the required ToD is absent in any given time slot, the reward rule for the basic experiment on energy neutrality defined in Table 1 is used instead.

TABLE 3. Reward rule of RLTDPM for advanced experiment with required ToD.

EDC / Occurrence Pattern	zero	low	moderate	high
Low	-2	+5	+2	-4
Medium	-4	-2	+5	+2
High	-8	-4	-2	+5

C. EVALUATING METRICS FOR THE EXPERIMENT PERFORMANCE

To compare the performance of the RLTDPM with other existing algorithms, measurements of residual battery energy (RBE) and exercised duty cycle (EDC) were recorded for the basic experiment. The RBE was obtained by measuring the battery's storage status at the end of each time slot: the larger the RBE, the better battery utilization the DPM performed. The length of each time slot was selected as one hour in the experiment. The EDC of the sensor node is also considered a meaningful metric and indicates the data sensing performance at a certain time slot. In conducting the advanced experiment with required ToD, two other metrics, termed offset to the required ToD (OTRT) and ToD achievability, were considered. The OTRT is obtained by subtracting the power management exercised duty cycle from the required sensor node duty cycle in any time slot. Once the OTRT becomes positive in every sensing time slot, the required ToD is considered met by the power management algorithm;

however, a larger OTRT also means more power is consumed. The ToD achievability, A_{ToD}, is defined as the following

$$\lambda_{OTRT} = \begin{cases} 1, & \text{if}(OTRT \geq 0) \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

$$A_{ToD} = \frac{\sum_{i=1}^{T} \lambda_{OTRT}}{T} \times 100\% \quad (22)$$

where the ToD achievability is obtained by dividing the number of time slots with positive OTRT by the total number of sensing time slots, T, and is represented in percentage form.

V. EXPERIMENTAL RESULTS

For long-term experiment purposes, the experiment time starts from the spring equinox of one year and continues until the spring equinox of the following year. For comparison purposes, the dynamic duty cycle adaptation (DDCA) method [7] was also employed in the experiments.

A. RESULTS OF BASIC EXPERIMENT

In the simulation, three initial battery energy levels, low, middle, and high with battery capacities of 20%, 50%, and 80%, respectively, were used as the initial battery conditions. The state of current battery energy were labeled as low, moderate, and high, which are defined as the residual battery falls into the range of 0%-40%, 40%-60%, and 60%-100% of full capacity. The reward rule for the basic simulation of the RLTDPM is defined in Table 1. Experimental results for the first 10 simulated days are illustrated in Fig. 6, where Figs. 6(a) and (b), respectively show the RBE and EDC. For the three levels of initial battery energy, the RBEs of Fig 6(a) first oscillate up and down according to the solar strength within a day, and finally converge to around 65% of full capacity at day 10. By examining the EDC results with respect to RBE for the first 4 days in Fig. 6(b), one can see that for the low initial battery condition (solid dark line), the EDC is suppressed by the RLTDPM agent with a lower EDC to maintain the energy neutral criterion such that the RBE can be increased. On the contrary, for a high initial battery condition (dark dashed line), the RLTDPM agent decides to exercise a larger duty cycle to reduce the RBE and maintain the energy neutral criterion. By inspecting both Figs. 6(a) and (b), the RLTDPM demonstrates its ability to adjust the appropriate duty cycle in data sensing and maintain RBE at a level higher than 50% after only 4 days of operation for different initial conditions of battery energy.

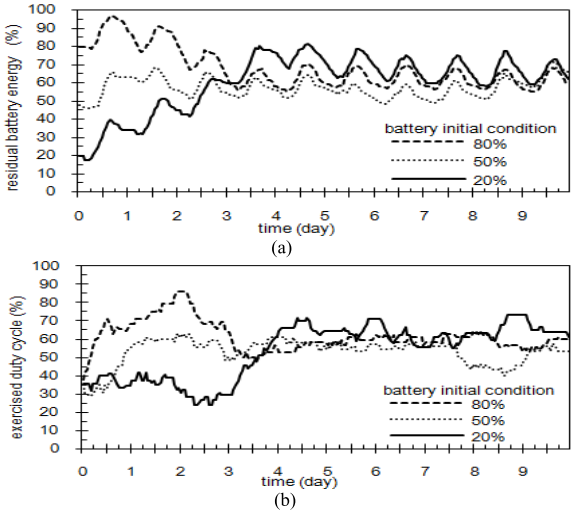

FIGURE 6. Experimental results of the RLTDPM for the basic experiment within the first 10 experiment days in terms of (a) RBE, and (b) EDC.

Qualitative results of the basic experiment using average RBE and EDC are shown in Table 4 for the first and second years. As can be seen from the average RBE and its standard deviation, given in the second and third rows, respectively, the average RBEs of the second year are all above 60% with a standard deviation much lower than those of the first year. Satisfactory EDCs higher than 50% and 57%, for the first and second years, respectively, are obtained after continuous operation of the EHWSN, as shown in the bottom row of Table 4. This again demonstrates that under long-term operation, the RLTDPM is able to effectively utilize harvested energy with satisfactory data sensing capabilities for perpetual operation of EHWSN.

TABLE 4. Qualitative results of basic experiment using average RBE and EDC.

	First year			Second year		
Init. bat. energy levels	20%	50%	80%	20%	50%	80%
Average RBE	59.8%	56.7%	68.7%	61.7%	63.0%	63.0%
Sta. dev. of ave. RBE	16.2%	15.1%	11.3%	3.4%	4.7%	4.2%
Average EDC	50.3%	53.0%	61.0%	57.7%	58.1%	54.4%

B. COMPARISON WITH DDCA ALGORITHM FOR BASIC EXPERIMENT

To understand how well the RLTDPM performs in comparison with other existing energy neutral-based DPM algorithms on a long-term scale, experimental results of the RLTDPM are compared with the DDCA method [7] for the basic experiment, as shown in Fig. 7. Figs. 7(a) and (b) respectively offer a performance comparison in terms of RBE and EDC for the RLTDPM and DDCA method both with an initial battery condition of 50% full capacity and experiment starting time from the spring equinox for one year. By examining the results of the DDCA method in Fig. 7(a), one can see that it slowly increases its RBEs and reaches a maximum of 76% around day 271, as shown by the dashed line. By contrast, the RLTDPM, shown in the dark solid line, quickly reaches its RBE maximum of 86% at day 15 and maintains the RBE within 75–85% of full capacity afterwards, which falls within the range of the defined high battery state. In Fig. 7(b), EDC traces of both the RLTDPM

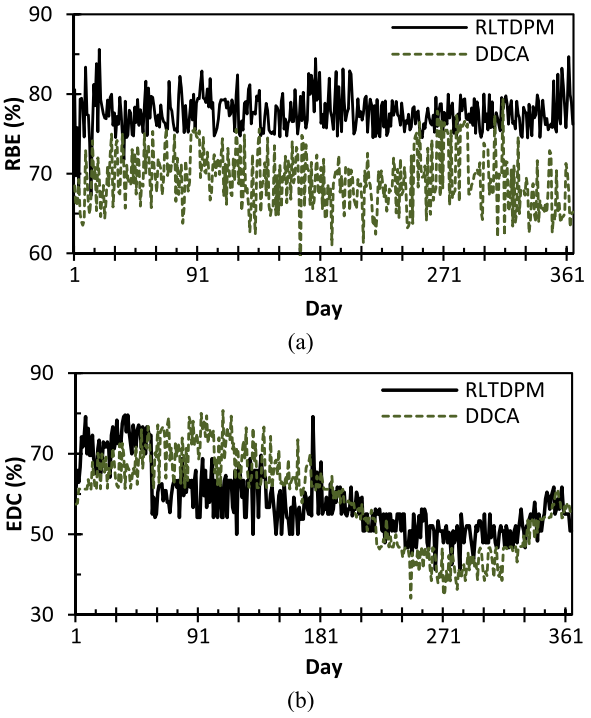

FIGURE 7. Experimental results of RLTDPM and DDCA methods in terms of (a) RBE, and (b) EDC for the basic experiment within one year's operation.

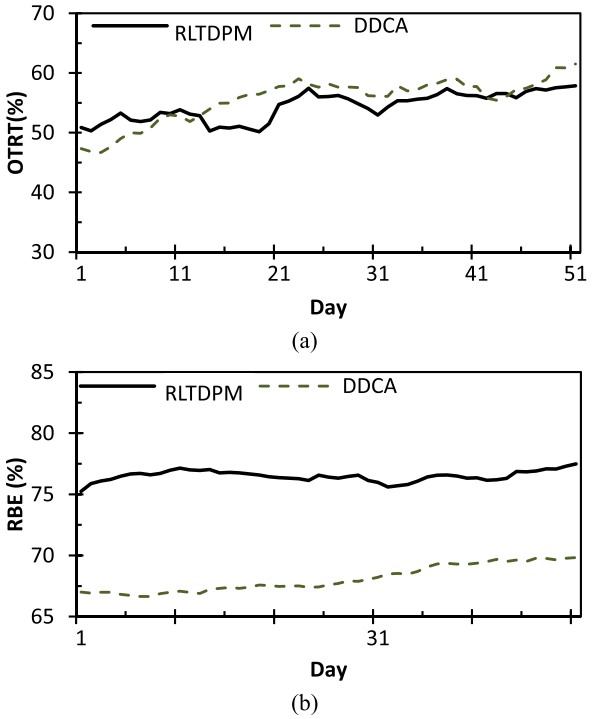

FIGURE 8. Experiment results of RLTDPM method in compared with DDCA method for the advanced experiment with required ToD. (a) OTRT, and (b) RBE for moderate ToD requirements for 50 days.

and DDCA methods exhibit trends of gradual decrease in response to the seasonal variation of harvested power, as was illustrated in Fig. 5(b). However, in comparing EDC traces with respect to the RBEs of Fig. 7(a), it can be seen that the RLTDPM exercises a more stable duty cycle for data sensing than the DDCA method, yet obtains higher RBEs than the DDCA method on most occasions. This indicates that the RLTDPM achieves better utilization of harvested energy by adaptively responding to seasonal variations better than does the DDCA method.

TABLE 5. Qualitative comparison of RLTDPM and DDCA method for basic experiment.

Comparing items	Methods		Percentage increase
	DDCA	RLTDPM	
Ave. RBE (%)	69.2	77.6	+7.4
Ave. EDC (%)	58.3	58.2	-0.1

Qualitative comparisons of the RLTDPM and DDCA methods for the basic experiment under the aforementioned variables and conditions are shown in Table 5. Table 5 demonstrates that both the RLTDPM and DDCA methods achieve almost the same average EDC, yet the RLTDPM obtains a 7.4% higher increase in average RBE than does the DDCA method. Statistics again exhibit that the RLTDPM method ensures more effective and adaptive utilization of harvested energy compared to the DDCA method. As such, one can conclude that for EHWSN deployed in remote areas for wild animal or insect data sensing, the RLTDPM provides a better energy management mechanism in both utilization of harvested energy and data sensing in responding to seasonal changes of solar energy, especially for the spring and summer time, when the activity of wild animals and insects is high.

C. RESULTS OF ADVANCED EXPERIMENT OF RLTDPM WITH REQUIRED TOD IN COMPARING WITH DDCA

The advanced experiment was conducted to contrast the RLTDPM against the DDCA method with a 50% initial battery condition and spring equinox experiment starting time for one year to understand how the ToD requirement is satisfied by the RLTDPM as opposed to the DDCA method in a long-term scale. Three different occurrence patterns of required ToD, as described in Table 2, were generated for the advanced experiment. Fig. 8 shows the results of the advanced experiment with ToD requirement, but for demonstration purposes, only the experimental results of moderate ToD requirements are presented. Fig. 8 (a) shows the OTRT of the moderate required ToD for 51 experiment days starting from the spring equinox and lasting for a year. By examining Fig. 8 (a), the resulting OTRTs of the RLTDPM are positive with values lower than those of the DDCA method for a majority of the time. This demonstrates that the ToD requirements are better satisfied by the RLTDPM with smaller EDC than with the DDCA method which consumes less energy. Fig. 8 (b) shows the RBE for the moderate ToD requirements. By examining

Fig. 8 (b) with respect to the OTRT of Fig. 8 (a), the RBEs of RLTDPM are found to be much larger than those of the DDCA method after one year of continuous operation. This indicates that the RLTDPM not only exercises more duty cycles in better satisfying the ToD requirements, but achieves better energy utilization as well.

TABLE 6. Qualitative comparisons of RLTDPM and DDCA algorithm for advanced experiment, all numbers shown are in percentages(%).

Occurrence pattern Of required ToD / Comparison item	RLTDPM			DDCA		
	Sparse	Moderate	Intensive	Sparse	Moderate	Intensive
Ave. ToD achievability	99.3	99.5	98.9	97.4	95.2	88.2
Ave. OTRT	54.9	53.6	50.9	62.4	58.8	53.2

Qualitative comparisons of the RLTDPM and DDCA methods for the advanced experiment with the required ToD for one year are shown in Table 6. By examining the average ToD achievability, as shown in the first data row of Table 6, we can see that the RLTDPM delivers a much higher achievable average ToD than does the DDCA method for the sparse, moderate, and high ToD requirements by 1.9%, 4.3%, and 10.7%, respectively. Further, the DDCA method obtains a higher average OTRT than does the RLTDPM, which means that the DDCA method consumes more power by exercising larger duty cycles than the required ToD. Inspection of the average ToD achievability and OTRT demonstrates that although the requested ToD requirements may vary, higher ToD achievability with less power consumption can still be realized by the RLTDPM method for better satisfying the ToD requirements than by the DDCA method.

VI. CONCLUSIONS

In this study, a novel ToD provisioning dynamic power management method for sustaining perpetual operation of energy harvesting wireless networks utilizing reinforcement learning, named RLTDPM, was proposed. Experimental results demonstrated that the proposed RLTDPM method can not only achieve energy neutrality between the harvested and consumed energy for sustaining perpetual operation of EHWSN, but can also yield ToD provisioning in satisfying ToD requirements in imaging data gathering. Moreover, experimental results confirmed the advantage of the proposed RLTDPM in self-learning and self-adapting to seasonal changes in energy harvesting for sustaining perpetual operation. With appropriate modeling of the energy harvesting environment and embedded system, the proposed RLTDPM could possibly be widely applied to the DPM of other types of embedded energy harvesting systems, such as solar-powered vehicles, space rovers, and electric vehicles with renewable/regenerative power.

REFERENCES

[1] L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 3, pp. 299–316, Jun. 2000.

[2] S. Irani, G. Singh, S. K. Shukla, and R. K. Gupta, "An overview of the competitive and adversarial approaches to designing dynamic power management strategies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 12, pp. 1349–1361, Dec. 2005.

[3] A. Rowe, K. Lakshmanan, H. Zhu, and R. Rajkumar, "Rate-harmonized scheduling and its applicability to energy management," *IEEE Trans. Ind. Informat.*, vol. 6, no. 3, pp. 265–275, Aug. 2010.

[4] C.-H. Hwang and A. C.-H. Wu, "A predictive system shutdown method for energy saving of event-driven computation," in *IEEE/ACM Int. Conf. Comput.-Aided Design, Dig. Tech. Papers*, Nov. 1997, pp. 28–32.

[5] A. Kansal, D. Potter, and M. B. Srivastava, "Performance aware tasking for environmentally powered sensor networks," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 32, no. 1, pp. 223–234, 2004.

[6] A. Kansal, J. Hsu, M. Srivastava, and V. Raghunathan, "Harvesting aware power management for sensor networks," in *Proc. ACM/IEEE Des. Autom. Conf.*, Jul. 2006, pp. 651–656.

[7] A. Kansal, J. Hsu, S. Zahedi, and M. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embedded Comput. Syst.*, vol. 6, no. 4, p. 32, Sep. 2007.

[8] C. Moser, L. Thiele, D. Brunelli, and L. Benini, "Adaptive power management for environmentally powered systems," *IEEE Trans. Comput.*, vol. 59, no. 4, pp. 478–491, Apr. 2010.

[9] D. Zhu, H. Aydin, and J.-J. Chen, "Optimistic reliability aware energy management for real-time tasks with probabilistic execution times," in *Proc. 29th IEEE RTSS*, Nov./Dec. 2008, pp. 313–322.

[10] D. Zhu and H. Aydin, "Reliability-aware energy management for periodic real-time tasks," *IEEE Trans. Comput.*, vol. 58, no. 10, pp. 1382–1397, Oct. 2009.

[11] R. Iyer and L. Kleinrock, "QoS control for sensor networks," in *Proc. IEEE Int. Conf. Commun.*, vol. 1. May 2003, pp. 517–521.

[12] A. Koubaa, R. Severino, M. Alves, and E. Tovar, "Improving quality-of-service in wireless sensor networks by mitigating 'hidden-node collisions,'" *IEEE Trans. Ind. Informat.*, vol. 5, no. 3, pp. 1444–1452, Mar. 2008.

[13] K. T. Phan, R. Fan, H. Jiang, S. A. Vorobyov, and C. Tellambura, "Network lifetime maximization with node admission in wireless multimedia sensor network," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3640–3646, Sep. 2009.

[14] A. Fallahi and E. Hossain, "A dynamic programming approach for QoS-aware power management in wireless video sensor networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 843–854, Feb. 2009.

[15] D. Chen and P. K. Varshney, "QoS support in wireless sensor networks: A survey," in *Proc. Int. Conf. Wireless Netw.*, Las Vegas, NV, USA, Jun. 2004, pp. 1–7.

[16] I. Howitt, W. W. Manges, P. T. Kuruganti, G. Allgood, J. A. Gutierrez, and J. M. Conrad, "Wireless industrial sensor networks: Framework for QoS assessment and QoS management," *ISA Trans.*, vol. 45, no. 3, pp. 347–359, Jul. 2006.

[17] D. Ernst, M. Glavic, and L. Wehenkel, "Power systems stability control: Reinforcement learning framework," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 427–435, Feb. 2004.

[18] K. S. Jeong, W. Y. Lee, and C. S. Kim, "Energy management strategies of a fuel cell/battery hybrid system using fuzzy logics," *J. Power Sources*, vol. 145, no. 2, pp. 319–326, 2005.

[19] R. C. Hsu and C.-T. Liu, "A reinforcement learning agent for dynamic power management in embedded systems," *J. Internet Technol.*, vol. 9, no. 4, pp. 347–353, 2008.

[20] R. C. Hsu, C.-T. Liu, and W.-M Lee, "Reinforcement learning-based dynamic power management for energy harvesting wireless sensor network," in *Next-Generation Applied Intelligence*. Berlin, Germany: Springer-Verlag, 2009, pp. 399–408.

[21] R. C. Hsu, C.-T. Liu, K.-C. Wang, and W.-M. Lee, "QoS-aware power management for energy harvesting wireless sensor network utilizing reinforcement learning," in *Proc. Int. Conf. Comput. Sci. Eng.*, Aug. 2009, pp. 537–542.

[22] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.

[23] L. Hu, C. Zhou, and Z. Sun, "Estimating biped gait using spline-based probability distribution function with Q-learning," *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 1444–1452, Mar. 2008.

[24] K.-S. Hwang, Y.-J. Chen, and C.-H. Lee, "Reinforcement learning in strategy selection for a coordinated multirobot system," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 6, pp. 1151–1157, Nov. 2007.
[25] C.-F. Juang and C.-H. Hsu, "Reinforcement ant optimized fuzzy controller for mobile-robot wall-following control," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 3931–3940, Oct. 2009.
[26] P.-Y. Yin, "Maximum entropy-based optimal threshold selection using deterministic reinforcement learning with controlled randomization," *IEEE Trans. Signal Process.*, vol. 82, no. 7, pp. 993–1006, Jul. 2002.
[27] C. B. Honsberg, R. C. Corkish, and S. P. Bremner, "A new generalized detailed balance formulation to calculate solar cell efficiency limits," in *Proc. 17th Eur. Photovolt. Solar Energy Conf.*, 2001, pp. 22–26.
[28] I. Reda and A. Andreas, "Solar position algorithm for solar radiation applications," *Solar Energy*, vol. 76, no. 5, pp. 577–589, 2004.
[29] M. Cucumo, D. Kaliakatsos, and V. Marinelli, "General calculation methods for solar trajectories," *Renew. Energy*, vol. 11, no. 2, pp. 223–234, 1997.
[30] (2014, Apr. 18). *Panasonic Solar Cells Technical Handbook* [Online]. Available: http://www.solarbotics.net/library/datasheets/sunceram.pdf
[31] N. A. Ali, M. Drieberg, and P. Sebastian, "Deployment of MICAz mote for wireless sensor network applications," in *Proc. IEEE ICCAIE*, Dec. 2011, pp. 303–308.
[32] OmniVision, Santa Clara, CA, USA. (2014, Apr. 18). *VGA Image Sensors* [Online]. Available: http://www.ovt.com/products/category.php?id=15

ROY CHAOMING HSU (S'93–M'01) received the M.S. and Ph.D. degrees in electrical engineering from Pennsylvania State University, PA, USA, and the M.S. degree in engineering management from National Tsing Hua University, Hsinchu, Taiwan, in 1991, 1995, and 2001, respectively. He was with National Space Organization, Taiwan, from 1995 to 2000, where he was involved in satellite flight software, and with Loop Telecom from 2000 to 2001, where he was involved in telecom-internet management software. He was an Associate Professor with National Formosa University from 2001 to 2005. He is currently a Professor and serves as the Chairperson of the Department of Electrical Engineering with National Chiayi University, Chiayi City, Taiwan. His expertise is in machine learning, image processing, pattern recognition, and embedded system.

CHENG-TING LIU received the B.S. degree in aeronautical engineering from National Formosa University, Yunlin, Taiwan, and the M.S. degree in computer science and information engineering from National Chiayi University, Chiayi City, Taiwan, in 2005 and 2008, respectively, where he is currently pursuing the Ph.D. degree with the Department of Computer Science and Information Engineering. His research interests include artificial intelligence, embedded systems, and image processing.

HAO-LI WANG received the B.E. and M.S. degrees from National Cheng Kung University, Taiwan, in 1997 and 1999, respectively, and the Ph.D. degree in computer and electrical engineering from Iowa State University, USA. He is currently a Faculty of the Department of Computer Science and Information Engineering with National Chiayi University, Taiwan. His research interests include wireless sensor networks and wireless LANs.