Value of Information Aware Opportunistic Duty Cycling in Solar Harvesting Sensor Networks

Jianhui Zhang, *IEEE Member*, Zhi Li, Shaojie Tang

***Abstract*—Energy-harvested Wireless Sensor Networks (WSNs) may operate perpetually with extra energy supply from natural energy, such as solar energy. Nevertheless, harvested energy is often too limited to support perpetual network operation with full duty cycle. To achieve perpetual network operation and process the data with high importance, measured by Value of Information (VoI), sensor nodes have to operate under partial duty cycle and to improve the efficiency of harvested energy. A challenging problem is how to deal with the stochastic feature of natural energy and variable data VoI. We consider the energy consumption during the energy storage and the diversity of the data process including sampling, transmitting and receiving, which consume different power levels. The problem is then mapped as a budget-dynamic Multi-Arm Bandit (MAB) problem by treating harvested energy as budget and the data process as arm pulling. This paper proposes an Opportunistic Duty Cycling (ODC) scheme to improve the energy efficiency while satisfying perpetual network operation. ODC chooses some proper opportunities to store harvested energy or to spend it on the data process based on historical information of energy harvesting and VoI of the processed data. With this scheme, each sensor node only needs to estimate ambient natural energy in short term so as to reduce computation cost and storage capacity for the historical information. It then can adjust its own duty cycle distributively with its local historical information. This paper conducts extensive theoretical analysis for the performance of our scheme ODC on the regret, which is the difference between the optimal scheme and ours. Our experimental results also manifest the promising performance of ODC.**

***Index Terms*—Opportunistic Duty-cycling; Energy Harvesting; Wireless Sensor Networks; Multi-armed Budget**

I. Introduction

As a promising technique, the great success of Wireless Sensor Networks (WSNs) has been witnessed over a variety of critical applications in recent years [1][2][3]. One common constraint, impeding the wider application of this kind of networks, is the limited energy supply. To extend network life or even to support perpetual network operation, two major techniques have been severally applied to WSNs: energy harvesting [2][4] and duty cycling [5]. Energy harvesting can supply sensor nodes with extra power from natural energy while the duty cycling technique can save energy so as to extend network lifetime. But most tiny energy-harvesting modules in solar sensor networks cannot harvest enough energy to support full duty cycle operation [2]. Some existing works combine the two techniques to achieve permanent network operation, *i.e.*, to meet the *energy neutral operation* [6][7]. These existing works estimate the amount of active time in a period previously, such as at the initialization phase of the period [4][7], or the average amount of active time for some periods over a long duration, such as a season [3]. However, there are several facts ignored by the existing works.

1) Imperfect charge efficiency. In practice, the charge efficiency of rechargeable battery in the solar powered sensor node is often less than 75% [8], which means that it indirectly wastes 25% energy to store harvested energy. Another choice, capacitor, suffers high leakage [9].

2) Variable data importance. In WSNs, the data redundancy is a common phenomenon. Meanwhile, if introducing the concept of Value of Information (VoI), the more important data has higher VoI. It can obtain higher energy efficiency to process the more important data.

3) Random natural energy. Some natural energy, such as solar or wind power, is shown to be random [2][10], so as hard to predict their profiles for long term accurately because of unpredictable weather and disturbance.

Section II illustrates some detailed technical evidences and examples to illustrate above observations. We find that it is still an open problem to improve the efficiency to exploit natural energy.

Notice that the energy consumption caused by imperfect charge efficiency can be decreased if harvested energy is directly used rather than stored in the battery. Considering data VoI, a sensor node can arrange right moments to process important data and to sleep so as to improve the energy efficiency, which is defined as the average VoI obtained per unit energy consumption in this paper. To do this, we propose an Opportunistic Duty Cycling (ODC) scheme to catch some features: the dynamic profile of harvested energy, the variable data VoI , and the easiness to estimate the harvested energy in short term. In this paper, the data process includes three actions: data sampling, transmitting and receiving, which consume different power so they have much impact on the energy efficiency. We then map the opportunistic duty cycling as a gambling game: Multi-Arm Bandit (MAB) [11]. In this game, a sensor node is treated as a gambler. The gambler decides its next action (sampling, receiving, transmitting or

Manuscript received Nov. 29, 2014; Accepted for publication Dec. 5, 2015. This work is under the support of the Major Program of the National Natural Science Foundation of China (NSFC) under Grants No.61190113, the General Program of NSFC 61473109 and 61003298, and the Zhejiang Province Natural Science Foundation under Grant No.LY14F020042.

J. Zhang (corresponding author, Email: jhzhang@ieee.org) and Z. Li are with College of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018 China.

S. Tang is with Jindal School of Management, University of Texas at Dallas, TX 75080 USA.

storing energy) step by step based on its estimation for natural energy and data VoI in the subsequent time.

In the real applications of energy-harvested WSNs, the data process and energy harvesting are highly dynamic. Under the MAB game, each sensor node can determine its next action according to its historical information in short term so as to deal with the dynamic feature. The goal of the gambling game is to maximize the energy efficiency for each sensor node. Clearly, in order to achieve this goal, each sensor node should carefully decide its next action while adhering to the energy constraint. Notice that to meet the energy neutral operation and to improve the energy efficiency usually contradict to each other when adjusting duty cycle. The former goal requires each sensor node to short its duty cycle while the later requires longer one to obtain the overall VoI as much as possible. To achieve the bi-criteria object, this paper adjusts a VoI threshold according to historical information.

Contributions. The contributions of this paper include:

1) This paper adjusts duty cycle by considering imperfect charge efficiency and data VoI while meeting the energy neutral operation. We map the new duty cycling problem as the budget-dynamic MAB problem. To our best knowledge, this is the first work to formulate and study the problem.

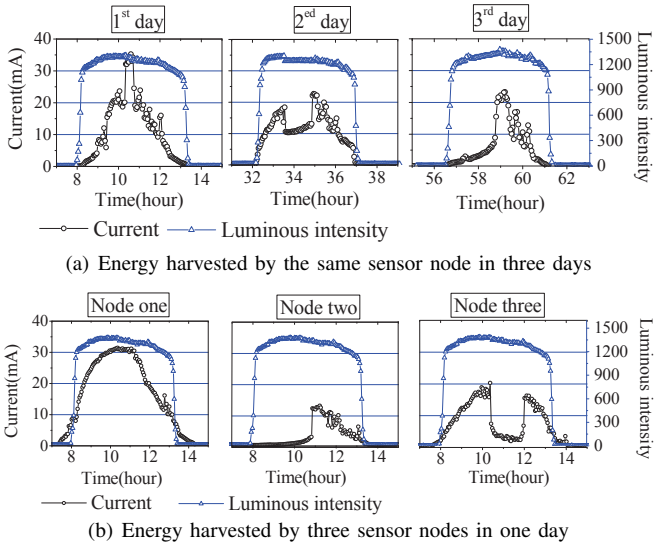
2) This paper designs ODC scheme to achieve the bi-criteria object. With ODC, each sensor node can distributively determine the action for the next time slot by playing the MAB game under the constraint of limited and dynamic harvested energy. An algorithm, called ODC, is designed to implement the ODC scheme. We theoretically analyze the performance of ODC by measuring a regret, the difference between the optimal scheme and ODC.

3) Extensive experiments are also conducted to evaluate the performance of our scheme. In the experiments, because of the hardness to find the optimal scheme, we propose two baseline approaches: a Centralized and Off-line duty cycling Algorithm (COA), and a Simple Duty Cycling (SDC). COA has the complete knowledge of the natural energy and the data VoI in advance. SDC predicts natural energy and calculates the duty cycle in advance as the algorithm given in the reference [4]. The experimental results show that the average energy efficiency achieved by our scheme is only 16.02% lower than that of COA, and 69.09% higher than that of SDC.

Road map. The following context of the paper is organized as follows. Section II describes the motivation based on our preliminary experiments, and formulates the opportunistic duty cycling problem in Section III. The problem is mapped as the budget-dynamic MAB problem, and ODC is presented in Section IV with its performance analysis in Section V, while the experimental results are discussed in Section VI. In Section VII, we review the related works on the energy harvesting module and the duty cycling schemes for WSNs and conclude this paper in Section VIII.

II. Preliminary Experiments and Motivation

This work is motivated by the following observations. Firstly, the inherent hardware property of the energy harvesting module leads to time varying charge efficiency. In practice, the average charge efficiency of the battery for the solar powered sensor node is often less than 75% [8]. Secondly, the random environmental factors, such as the shadow of clouds, can also decrease the charge efficiency. Thirdly, the data VoI varies over time and is different among the nodes. These observations leave the existing duty cycling schemes unsuitable, and motivate us to design a new duty cycling scheme.



(a) Energy harvested by the same sensor node in three days

(b) Energy harvested by three sensor nodes in one day

Fig. 1. Current indicates the amount of the harvested energy. (a) Energy profile changes with time. (b) Different sensor nodes have different profiles in one day.

A. Dynamic Energy Harvesting and Storage

The unpredictable environmental factors lead to diverse energy profiles among sensor nodes as examples in Figure 1. The experiment results in Figure 1(a) indicate that the same sensor node usually has different energy profiles in several days even under the similar weather conditions. More so, the energy profiles for several different sensor nodes vary a lot during the same day because of their different locations as shown in Figure 1(b). Similar phenomenon was also observed in previous works [9]. Some works model the solar energy harvesting as a first-Markov random process [10].

The time to consume or store harvested energy has great impact on the energy efficiency. Due to imperfect charge efficiency, denoted by λ, the relation between the harvested energy e^h and the actual stored energy e^s is $e^s = \lambda e^h$ for the charge efficiency $\lambda < 0.75$. The solar panels on the most existing solar modules, such as SolarMote [2] and Prometheus [4], have the rated current of about 20 mA. Meanwhile, the working current of the existing sensor nodes, such as TelosB, is 20 mA for receiving and 19 mA or more for transmission. If a sensor node powers its antenna with the harvested energy (20 mA) directly, then its antenna can work normally. Otherwise, if it stores the harvested energy with the power 20 mA, the actual stored energy is 20×0.75=15 mA given $\lambda = 0.75$, which means that 5 mA harvested energy is wasted. The power of the stored energy is thus too low to support its normal operation. Notice that this paper uses

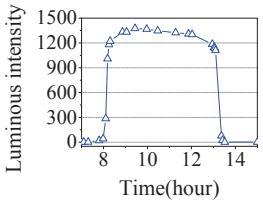


Fig. 2. Data process is reduced greatly while a little VoI is lost.

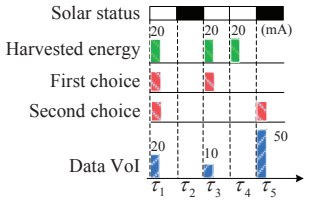


Fig. 3. Example for different data process choice.

the electric current to measure the power since the fault rated voltage for the existing sensor nodes is constant, such as 3 V for the TelosB and MICA nodes.

B. VoI of Data

The limited harvested energy compels each sensor node to preferentially process the data with high VoI. According to Information Theory, the data importance can be indicated by VoI, denoted by I [12]. The Kullback-Leibler (KL) divergence measure can calculate VoI by qualifying the difference between two probability distributions: $p_1(t)$ and $p_2(t)$ as follows.

$$I_{KL}(p_1(t), p_2(t)) = \int p_1(t) \log \frac{p_1(t)}{p_2(t)} \tag{1}$$

With the concept of VoI, a sensor node chooses the important data (*i.e.* with high VoI) to process. The times to process data then can be decreased so as to save much energy while preserving the overall VoI. For example, when reducing the times to sample the luminous intensity from Figure 1(b) to Figure 2, about 92% energy is saved while the overall VoI lose is preserved under 5%.

C. Call for Online Energy Allocation

Since both of the data process and energy harvesting are random processes, each sensor node can make online decision to allocate the harvested energy. The example in Figure 3 illustrates the necessity of the online energy allocation to maximize the overall VoI by carefully scheduling the energy consumption. In this example, the sensor node v_i can harvest 20 mA energy at the time slots marked with "white" color solar status, and cannot harvest energy at the "black" time slots. Suppose that v_i requires at least 20 mA energy to support its normal operation at each time slot, and that the charge efficiency $\lambda = 0.75$. When time t goes to τ_1, v_i can use the harvested 20 mA energy directly to process the first data with 20 unit VoI. After t goes to τ_3, v_i has two choices. The first choice is that v_i uses the harvested energy at τ_3 to process the second data, and then obtains 10 unit VoI. At τ_4, v_i stores the harvested 20 mA energy, and obtains 15 mA energy because $\lambda = 0.75$. At τ_5, v_i cannot process data since the stored energy is not sufficient. The VoI per unit energy that v_i obtained by the first choice is $\frac{20+10}{20+20} = 0.75$. The second choice is that v_i stores the 40 mA energy harvested at τ_3 and τ_4 and obtains 30 mA energy. It then processes the second data at τ_5, and obtains 50 unit VoI. The VoI per unit energy that v_i obtained by the second choice is $\frac{20+50}{20+30} = 1.4$. Obviously, the second choice can result in higher energy efficiency, *i.e.*, the VoI per unit energy, than the first one.

D. Opportunistic Duty Cycling

From the above facts, we find that the processes of the data process and energy harvesting are highly dynamic. It can greatly improve the energy efficiency to wake up the sensor node to process data and to hibernate them for storing energy at proper moments. These facts motivate us to propose the novel opportunistic duty cycling scheme, under which the sensor nodes can catch the right opportunities to process data or to store the harvested energy. Existing works on duty cycling adjust only the amount of active time in a period as shown in Figure 4. Under the opportunistic duty cycling, the active slots are also considered as the example in Figure 5, where the period composes of 8 slots. The set of active slots may be different as the cases a and b in Figure 5 although the duty cycles under both cases are same, *i.e.*, $\frac{3}{8}$. The reason to adjust the duty cycle in this way is that they may result in different energy efficiency. The goal of the opportunistic duty cycle is to adjust the duty cycle and to arrange the active time slots so that the energy efficiency can be improved under the constraint of the energy neutral operation.

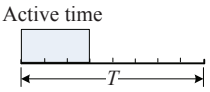
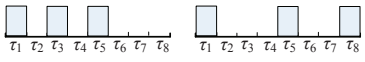


Fig. 4. Previous duty cycling.



(a) Case a (b) Case b

Fig. 5. Opportunistic duty cycling.

Most symbols used in this paper are summarized in Table I.

TABLE I
SYMBOL AND MEANING

Sym.	Description	Sym.	Description
T	Period	a	Arm of bandit machine
v	Node	Ψ	# of pulling all arms
s	Set of slots	φ	# of pulling one arm
K	# of arms	I	Reward/VoI
χ	Pull variable	$\bar{I}$	Estimation of I
Ł	Route	I_d	Threshold of VoI
Δ	Reward difference	$\hat{I}$	Upper-bound of VoI
R	Regret	p	Probability
X	Scheme	c	Cost/Energy consumption
θ	Solar state	e	Energy processed in slot τ
λ	Charge efficiency	E	Energy processed till slot τ
$\mathbb{E}$	Expectation	M	VoI of remaining data
α, β, γ	Coefficients	A, B	Coefficient vectors

III. System Model and Problem Formulation

A. Network and Energy Model

Given a network with a sink and some sensor nodes v_i, $i = 1, 2, \cdots$, each node is assumed to have at least one stable route connecting with the sink. A period T composes of $|T|$ time slots τ_i, $i = 1, \cdots, |T|$. Each node is equipped with a micro-scale energy-harvesting module, and cannot receive and transmit data at same time because of its half-duplex antenna. It is equipped with one battery to store energy with the initial energy level e_0. Because of the limited hardware, the battery cannot support the operation of the sensor node when it is being charged by harvested energy [2][9]. Meanwhile,

the power of the micro-solar panel is too low to support the normal operation of the sensor node and its battery charging simultaneously in most time as the experimental result in Figure 1. We thus assume that the limited harvested power cannot support the normal operation of the sensor and antenna simultaneously.

For each sensor node v_i, the different power levels are required to support data sampling, receiving, transmitting and storing the harvested energy, respectively denoted by c^s, c^r, c_i^t and c_i^g. c^s and c^r are constant and same over all sensor nodes. The VoI, denoted by $I_i(\tau)$, is measured by Equation (1). Denote the amount of energy harvested by a single sensor node at time slot τ by $e^h(\tau)$. The harvested energy $e^h(\tau)$, $\tau \in T$, over a period can be modelled as the first-order stationary Markov process [10]. The data arrived over a sequence time slots is assumed to be a Markov process. Each solar panel can support its node's normal operation or can charge its node's battery if and only if its harvested energy is over a threshold e_t. Let $\theta = 1$ if the power of harvested energy is over the threshold, and 0 otherwise.

B. Opportunistic Duty Cycling Problem

The opportunistic duty cycling can be formalized as the optimization problem. The goal of ODC is to maximize the overall VoI collected at the sink as given in Equation (2), while satisfying the energy neutral operation under the constraint of the energy harvesting randomness in Equation (3).

$$\max \sum_{\tau \in T} I_{sink}(\tau) \tag{2}$$

where $I_{sink}(\tau)$ denotes the VoI received by the sink at τ. At the time slots in the sets s^s, s^r and s^t, the sensor node v_i samples, receives and transmits data respectively. At the time slots in the set s^g, v_i stores the harvested energy into its battery and thus $\theta = 1$ at every slot in s^g. To meet the energy neutral operation, the consumed energy should be less than the harvested.

$$|s^s|c^s + |s^t|c_i^t + |s^r|c^r + |s^g|c_i^g \leq \sum_{\tau \in T} e_i^h(\tau) \tag{3}$$

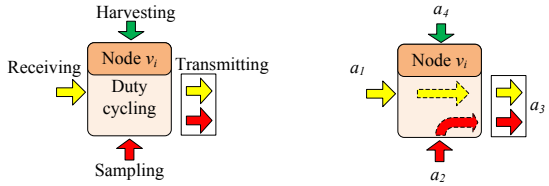
According to the assumption in the subsection III-A, the antenna is half-duplex so the sets s^r, s^t has no common element. Meanwhile, the four sets: s^g, s^s, s^r, and s^t have no common element because of the limited hardware and harvested energy as the assumption in Section III-A. The four sets thus satisfy the following condition.

$$\begin{cases} s^g \cup s^s \cup s^r \cup s^t = T \\ s^r \cap s^t = \varnothing; \text{ and } s^r \cap s^s = \varnothing; \text{ and } s^t \cap s^s = \varnothing \\ s^g \cap s^s \cup s^r \cup s^t = \varnothing \end{cases} \tag{4}$$

The core of ODC scheme is to find these four subsets: s^s, s^r, s^t and s^g, so as to solve the optimal problem in Equation (2) under the constraint in Equation (3) and (4).

IV. Opportunistic Duty Cycling

This section formulates the opportunistic duty cycling as the budget-dynamic MAB problem [13], and then presents our duty cycling scheme: ODC.



(a) Node (b) Gambler

Fig. 6. (a) Node v_i has four actions: sampling, receiving, transmitting and storing. (b) Mapping v_i to a gambler with four arms, a_1, a_2, a_3 and a_4. Four actions become four arms respectively after mapping a node to a gambler.

A. Budget-dynamic MAB Problem

Let us look into the detailed process of the opportunistic duty cycling in energy-harvested WSNs. For the harvested energy, each node has two ways to deal with it: consuming or storing it. To store it results some energy consumption because of imperfect charge efficiency, *i.e.*, $\lambda < 1$. Otherwise, it spends the harvested energy on data process. When no energy to harvest, it must consume its battery power on data process, or sleep so as to lose the chance to process data. Obviously, each node has to choose one of the four actions: sampling, receiving, transmitting data and storing energy (*i.e.* sleeping), as shown in Figure 6(a), by consuming the harvested or stored energy at each time slot. To maximize the energy efficiency, each node need choose the best action by learning its historical information of the energy harvesting and data process. Since the energy harvesting and data process are the Markov process, the conditional probability (given the historical information) that the harvested energy and VoI of the data are at certain levels at the beginning of slot τ is a sufficient statistic for the design of the optimal actions in the slot τ [14]. Each node thus need not record the long historical information, and can estimate VoI for next time slot by counting the probability that the power and data VoI are at certain levels during the previous time slots in short term.

If treating a sensor node as a gambler, the harvested energy is the budget of the gambler and its four actions represent the four arms of the bandit machine as shown in Figure 6, the opportunistic duty cycling can be formulated as the budget-dynamic MAB problem. Pulling the arms a_1, a_2, a_3 and a_4 represents the four actions: data receiving, sampling, transmitting and energy storing. In the MAB problem, the gambler pulls one of the bandit machine's arms by costing some budget. The bandit machine then returns it with some reward each time. For simplicity, we take the VoI of the processed data as the reward. For example, a sensor node receives a data, whose VoI is I, and then the reward is I. The goal of the gambler is to maximize the overall reward under its budget constraint by a series of arm pullings. In this paper, the harvested energy, *i.e.* the budget, is dynamic, so the problem studied in this paper is a new variation of the classical stochastic MAB problem: the budget-dynamic MAB problem. By mapping the opportunistic duty cycling problem to the MAB problem, the goal to maximize the energy efficiency is equivalent to maximizing the reward given dynamic and limited budget.

Since a sensor node is treated as a gambler in the MAB

problem, it means that the solution to the problem is implemented distributively. A challenge to solve the problem is to prove the distributive scheme can guarantee the global maximization of the overall VoI. Notice that all data received by the sink is sent by each single node. Thus, a straightforward idea is to maximize the VoI of each single node. To support the idea, the following context considers a more general case than that given in Equation (4). Notice that the case in Equation (4) is covered by the following statements. Denote the VoI obtained by the three actions: sampling, receiving and transmission, by I^s, I^r and I^t respectively. Let $M(\tau)$ denote the overall VoI of the data remaining in v_i's memory till the end of time slot τ. Recall that each node cannot receive and transmit data simultaneously as the constraint in Equation (4). When the node takes the action to transmit data in τ, there is a balance that is $M(\tau)=M(\tau-1)+I^s(\tau)-I^t(\tau)$ at time slot τ, where $I^s(\tau)$ is the VoI of the sampled data at the slot τ. We have the following equation:

$$I^t(\tau) = M(\tau-1) - M(\tau) + I^s(\tau) \tag{5}$$

Similarly, we have the following updating equation when the node takes the receiving action.

$$M(\tau) = I^r(\tau-1) + M(\tau-1) + I^s(\tau-1) \tag{6}$$

where $I^r(\tau-1)$ and $I^s(\tau-1)$ are the VoI of received and sampled data at time slot $\tau-1$ respectively. They may be zero since the action: data transmitting or energy storing, may be taken. Considering the special case that only one of the four items: $I^t(\tau)$, $I^s(\tau)$, $I^r(\tau)$ and $I^s(\tau)$ can be the value over zero, Equation (5) and (6) satisfy the constraints in Equation (4). It meas that the situation given in Equation (4) is a special case of the following statement.

Recall that each node has at least one routing connecting with the sink as the statement in Section III-A. Let Ł_k denote the set of nodes that are k hops away from the sink, $k = 1, 2, \cdots$. The overall reward of the whole network can be calculated as $\sum_{\tau\in T} I_{sink}(\tau) = \sum_{\tau\in T}\sum_{v_i\in \text{Ł}_1} I_i^t(\tau)$ in the period T, where $I_i^t(\tau)$ is the VoI of the data transmitted by the node v_i at the time slot τ. The following theorem proves that $\sum_{\tau\in T} I_{sink}(\tau)$ can be maximized by maximizing the overall reward of each single node. This paper decomposes the overall reward of the sink to that of each node by the following theorem.

Theorem 1: Assume each node has at least one route connecting with the sink. The total reward of all nodes accumulated in a period equals to the total reward received by the sink in the period.

Proof: The intuitive idea of the proof is that all of the data received by the sink must be sent or relayed by the intermediate nodes in the network. Let v_0 denote the sink, and suppose that the network starts at the time slot $\tau = 0$. When $\tau = 0$, *i.e.*, the network does not begin to run, each node v_i does not receive or sample any data so $M_i(\tau_i = 0) = 0$. In an arbitrary time slot $\tau > 0$, the VoI of the data received by the sink is that the relay node $v_i \in \text{Ł}_1$ transmits at the same slot. That is

$$I_{sink}(\tau) = \sum_{v_i\in \text{Ł}_1} I_i^t(\tau)$$

Thus, to maximize $I_{sink}(\tau)$ is equivalent to maximizing the data traffic of each node away one-hop from the sink at time slot τ. According to Equation (5), the right side of the above equation can be rewritten as follows:

$$I_i^t(\tau) = M_i(\tau-1) - M_i(\tau) + I_i^s(\tau), v_i \in \text{Ł}_1 \tag{7}$$

Notice that any data sampled or received at time slot τ can be transmitted after τ. The transmitted data $I_i^t(\tau)$ must come from the remaining data $M_i(\tau-1)$. The last two items $M_i(\tau)$ and $I^s(\tau)$ have no contribution to $I_i^t(\tau)$. Before the time slot τ, v_i ($v_i \in \text{Ł}_1$) must receive or sample data and store it in $M_1(\tau^t - 1)$. Otherwise, it has no data to transmit in τ. The data that the sensor node chooses to transmit at time slot τ must be received or sampled at some earlier time slot τ', *i.e.*, $\tau' < \tau$. When the sensor node transmits the data in τ, the time $\tau-1$ or τ' ($\tau' < \tau-1$) at which the data is received or sampled has no affection on the transmission of the data. To understand the proof easily, we can assume that the data that the sensor node chooses to transmit at time slot τ is received or sampled at $\tau-1$. Meanwhile, the data received by the sensor nodes in the layer Ł_k must be transmitted by those in the layer Ł_{k+1} so we have the following equation:

$$\sum_{v_i\in \text{Ł}_k} I_i^r(\tau) = \sum_{v_j\in \text{Ł}_{k+1}} I_j^t(\tau) \tag{8}$$

According to Equation (6) and (7), the VoI of the data received by the sink till time slot τ is:

$$\begin{aligned} I_{sink}(\tau) &= \sum_{v_i\in \text{Ł}_1} I_i^t(\tau) = \sum_{v_i\in \text{Ł}_1} [M_i(\tau-1) - M_i(\tau) + I_i^s(\tau)] \\ &= \sum_{v_i\in \text{Ł}_1} M_i(\tau-1) - \sum_{v_i\in \text{Ł}_1} [M_i(\tau) + I_i^s(\tau)] \\ &= \sum_{t=0}^{\tau-1}\sum_{v_j\in \text{Ł}_2} I_j^t(t) + \sum_{t=0}^{\tau}\sum_{v_i\in \text{Ł}_1} I_i^s(t) - \sum_{v_i\in \text{Ł}_1} M_i(\tau) \end{aligned} \tag{9}$$

In the last equality of the above equation, the first item is the sum of the traffic of the sensor nodes in the layer Ł_2, which contributes to the VoI of the data received by the sink, *i.e.*, $I_{sink}(\tau)$ at time slot $\tau-1$. In other words, the VoI of each sensor node $v_j \in \text{Ł}_2$ must be maximized at $\tau-1$ before the overall VoI $I_{sink}(\tau)$ can be maximized at time slot τ since the last two items have no contribution to $I_{sink}(\tau)$ in $\tau-1$ according to the statement below Equation (7).

Similarly, we can deduce $I_{sink}(\tau)$ in Equation (9) back to the sum of the VoI of the data transmitted by the sensor nodes in the layer Ł_k during time slot $\tau-k+1$. Therefore, the overall VoI of the sink in the period T, *i.e.* $\sum_\tau^T I_{sink}(\tau)$, can be maximized by maximizing the VoI of the data transmitted by each sensor node in each layer over a series of time slot τ, $\tau \in T$. ■

B. ODC

This block presents the detailed design of our scheme: ODC. In order to achieve the energy neutral operation, a parameter I_d, called VoI threshold, is introduced to control the amount of energy that each sensor node can consume on average. Because of the randomness of natural energy, I_d should be updated continuously. An Adaptive VoI Adjustment (AVA) algorithm is designed to update the threshold I_d in this paper.

1) ODC algorithm: Recall that the goal of ODC is to maximize the VoI of each sensor node, *i.e.* to solve the budget-dynamic MAB problem, so that the overall VoI can be maximized according to Theorem 1. Imagine that taking an action corresponds to placing an item into a knapsack. The expected reward by taking the action equals to the item's value and the energy consumption for the action is the item's weight. The total harvested energy till τ is then the weight capacity of the knapsack at τ. Therefore, the budget-dynamic MAB can be reduced to the unbounded knapsack problem at each time slot τ. We borrow the idea of the density-ordered greedy algorithm [15] to solve the problem.

During solving the budget-dynamic MAB problem by the density-ordered greedy algorithm, the key step is to estimate the VoI that each action will obtain at the next time slot τ, so that the sensor node v_i can take those actions with the highest energy efficiency. Auer introduced the Upper Confidence Bound (UCB) to calculate the estimated VoI of each action [16]. The most popular UCB, called UCB-1, relies on the upper-bound VoI $\bar{I}'_j(\tau) + \delta'_j(\tau)$ obtained by taking the action a_j, where $\delta'_j(\tau)$ is a padding function. A standard expression of the function is $\delta'_j(\tau) = \hat{I}\sqrt{\frac{\varepsilon' \ln \Psi(\tau)}{\varphi_j(\tau)}}$, where $\hat{I}$ is the upper-bound on the reward/VoI, $\varepsilon' > 0$ is an appropriate constant. $\varphi_j(\tau)$ is the number of taking action a_j till τ. $\Psi(\tau)$ is the overall number of actions that the sensor node v_i has taken till τ, and $\bar{I}'_j(\tau)$ is the estimation of the action a_j's expected reward for the slot τ at the end of the slot $\tau - 1$. In order to improve the energy efficiency, the upper-bound VoI per unit cost can be calculated as $\bar{I}_j(\tau) + \delta_j(\tau) = (\bar{I}'_j(\tau) + \delta'_j(\tau))/c_j$ by taking the cost c_j into consideration. We have $\bar{I}_j(\tau) = \bar{I}'_j(\tau)/c_j$ and $\delta_j(\tau) = \hat{I}\sqrt{\frac{\varepsilon_j \ln \Psi(\tau)}{\varphi_j(\tau)}}$, where $\varepsilon_j = \varepsilon'/c_j^2$. Notice that the remaining energy $E(\tau)$ till time slot τ composes of the energy remained in its battery $E(\tau)$ and possibly harvested energy at τ, *i.e.*, $E(\tau) = E(\tau - 1) + \theta(\tau)e^h(\tau)$. Thus, the unbounded knapsack problem can be formulated as the following problem with the time-dependent energy bound $E(\tau)$.

$$\max \sum_{j=1}^{K} \chi_j(\tau)(\bar{I}_j(\tau) + \varepsilon_j) \tag{10}$$

$$s.t. \sum_{j=1}^{K} \chi_j(\tau)c_j \leq E(\tau), \forall j, \tau : \chi_j(\tau) \in \{0, 1\} \tag{11}$$

where $\chi_j(\tau)$ is a bool indicator. $\chi_j(\tau) = 1$ if the action a_j is taken at τ, and 0 otherwise. c_j is the energy consumption to pull the arm a_j once. The constraint in Equation (11) means that the energy consumption at time slot τ is constrained by $E(\tau)$. $\bar{I}_j(\tau)$ can be calculated as the average reward received by pulling arm a_j till $\tau - 1$.

$$\bar{I}_j(\tau) = \sum_{t=1}^{\tau-1} \frac{\chi_j(t)I_j(t)}{c_j\varphi_j(\tau - 1)} \tag{12}$$

The problem defined in Equation (10) is NP-hard so this paper uses the density-ordered greedy method [15] to find a near-optimal selection of the sets s^s, s^t and s^r, *i.e.* to find the integer $\chi_j(\tau)$ so that Equation (10) is maximized (see step 12 in Algorithm 1).

The memory capacity of a sensor node is limited. Each sensor node thus should keep balance between its output: the transmitted data and its input: the received and sampled data in the long term. In other words, the times to pull the arm a_3 is expected to equal to the sum of the times to pull the arms a_1 and a_2. To do this, we assign each action with some probability. Let $\chi^*_j(\tau)$ be the solution to the problem in Equation (10) by the density-ordered greedy method at the time slot τ. ODC takes the next action $a(\tau)$ with some probability, which is determined by the following equation (see step 13 in Algorithm 1).

$$p(a(\tau) = a_j) = \begin{cases} \chi^*_j(\tau)/\sum_{j=1}^{K} \chi^*_j(\tau), & j = 1, 2 \\ 2\chi^*_j(\tau)/\sum_{j=1}^{K} \chi^*_j(\tau), & j = 3 \end{cases} \tag{13}$$

where K is the number of the arms of the bandit machine. Notice that the arm with the higher upper bound VoI will have higher probability in Equation (13) since the times that it is pulled is higher than others. ODC is presented in Algorithm 1, and its performance will be theoretically analyzed on its regret bound in the next section. In this algorithm, $c(\tau)$ is the energy consumed at time slot τ. For example, if the arm a_j is pulled and the consumed energy is c_j in τ, then $c(\tau) = c_j$.

2) AVA: The intuitive idea behind AVA is that each sensor node dynamically estimates the VoI threshold for the next time slot according to the harvested energy and the consumed energy in the previous time slots. The energy neutral operation condition requires each sensor node to consume energy less than the remaining one, *i.e.* $E^h(\tau) \geq E^c(\tau)$, while the sensor node v_i has to consume energy as much as possible to maximize the total reward in the period. The best choice is to keep the balance between the harvested and consumed energy in a period, *i.e.* $E^h(T) = E^c(T)$, where $E^h(T) = \sum_{\tau \in T} e^h(\tau)$ and $E^c(T) = \sum_{\tau \in T} c(\tau)$. We define the following function as a metric to find the balance point.

$$\lim_{T \to \infty} \frac{1}{|T|} \sum_{\tau=1}^{T} [E(\tau) - c(\tau)]^2 \tag{14}$$

Denote the VoI threshold updated at τ by $I_d(\tau)$. A proper $I_d(\tau)$ ensures that each sensor node can minimize the average squared deviation of the harvested energy from the consumed energy by Equation (14). To find the proper I_d, we adopt the adaptive control theory in Algorithm 2, transforming the

Algorithm 1 The ODC Algorithm

Input: $c(1) = 0$ and $I_d(1) = \bar{I}(1) = 0$;
Output: A sequence of actions;

```
Initialize: τ = 0 and E(τ) = e_0;
while τ+ = 1, and τ ≤ |T| do
    Update the remaining energy E(τ) till τ;
    Input e^h(τ) and c(τ) into Algorithm 2 to update I_d(τ+1);
    if Ī(τ) < I_d(τ) then
        Pull arm a_4 to store energy;
        E(τ) = E(τ−1) + λθ(τ)e^h(τ), and go to the step 2;
    end if
    if τ ≤ K then
        Initial phase: pull the arms a_i, i = 1, 2, 3 one by one;
    else
        Calculate χ*_j(τ) by solving the knapsack problem in Equation (10);
        Take the action a_j(τ) with the highest probability p(a(τ) = a_j) given in Equation (13);
        E(τ) = E(τ−1) + (θ(τ) − 1)c_j;
    end if
    Update the upper bound VoI Î_j of the action a_j(τ);
    Update Ī(τ+1) = max_{a_j : j=1,···,K} Ī_j(τ) by Equation (12);
end while
```

threshold determining problem as the linear-quadratic tracking problem. More formally, this paper argues that a first order, discrete-time, linear dynamical system with colored noise for the problem. This system can be described by the following equation:

$$c(\tau+1) = \alpha c(\tau) + \beta I_d(\tau) + \gamma\omega_\tau + \omega_{\tau+1} \tag{15}$$

In this system, $c(\tau+1)$ is refer to the output of the system, I_d is the control. ω is mean zero input noise. α, β, γ are real-valued coefficients. The optimal output of the system is to keep the metric in Equation (14) as small as possible in the period T. The optimal control law to minimize the tracking error is [17]:

$$I_d(\tau) = [e^h(\tau) - (\alpha+\beta)c(\tau) + \gamma e^h(\tau)]/\beta \tag{16}$$

The coefficients α, β and γ are not known in advance, and can be estimated online in our problem by using the standard gradient descent techniques [17]. Firstly, we define a parameter vector $A_\tau \triangleq (\alpha+\gamma, \beta, \gamma)^T$, and a feature vector $B_\tau \triangleq (c(\tau), I_d(\tau), -e^h(\tau))^T$. By the two vectors, the optimal control law in Equation (16) can be expressed as $B_\tau^T A = e^h(\tau)$. The estimated parameter vector $\hat{A}$ for A then can be defined by the gradient descent update rule as given by

$$\hat{A}_{\tau+1} = \hat{A}_\tau + \mu B_\tau (c_{\tau+1} - B_\tau^T \hat{A}_\tau)/(B_\tau^T B_\tau) \tag{17}$$

where μ is a positive constant step-size parameter.

Because each sensor node need store its harvested energy in its battery, the initial energy level e_0 would better be about half of the battery capacity. The choice of the $\hat{A}_\tau$'s initial value $\hat{A}_0$ greatly affects the converge speed of the parameter estimation in Equation (17). $\hat{A}_0$ can be set preciously according to preliminary experimental results. Examining the system in Equation (15), the increment of the control I_d results in less data being received or sampled, so less energy consumption. b should be negative. Set $B_0 = (c_0, I_d(0), -e^h(\tau))$.

Algorithm 2 AVA

Input: The harvested energy $e^h(\tau)$ and the consumed energy $c(\tau)$ of v_i till τ. Let $\tau = 0$.
Output: The updated threshold $I_d(\tau+1)$.

```
if τ = 0 then
    Â_τ = Â_0 and set B_0;
end if
Update the parameter vector Â_{τ+1} by Equation (17);
Update the feature and parameter vectors B_τ, A_τ;
Output I_d(τ+1) using Equation (16);
```

Considering a special case in which each sensor node can harvest enough solar energy. Thus, the harvested energy can support its operation at every time slot. When it cannot harvest sufficient energy, a high threshold $I_d(\tau)$ prevents it from working at every time slot, *i.e.* by reserving some energy at some time slots. So the harvested energy is stored and will not be consumed completely at every time slot, *i.e.*, $E(\tau) \geq 0$.

C. Common Activity

Recall that ODC is implemented distributively so a concerned issue is how about the common active time among neighboring nodes under it. By Algorithm 1, each sensor node chooses the transmitting and receiving arms with some probability and thus it may have common active time, *i.e.* simultaneous waking up, with others at each time slot. This section shows the probability that one sensor node has common active time with its neighbor theoretically and experimentally. If the node can communicate with at least one neighbor, we say that its common active time is nonzero. Figure 7 illustrates the theoretical probability that some neighboring nodes have common active time. When each node has some probability to wake up, *i.e.* active probability, the common active probability can be easily computed as the y-coordinate. More neighbors the node has or higher probability it wakes up, it has higher probability to communicate with its neighbor in Figure 7. Figure 8 illustrates the experimental results when one node has two neighbors. The experimental setting is given in Section VI. In the experiment, the common active probability tends to 0.22, and the average data VoI obtained by each action tends to about 0.57. In each time slot, the node can guarantee a certain probability to communicate with its neighbors. The probability is not quite high but the obtained VoI is not low since the node transfer the important data. Next section analyzes that VoI difference of the data processed by the optimal solution and our scheme ODC.

V. Performance Analysis

This section analyzes the theoretical performance of ODC by the metric: regret. Let $I_X(E)$ be the total VoI returned

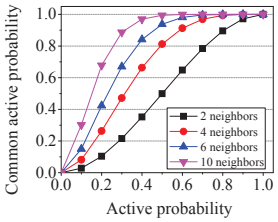


Fig. 7. Theoretical common active time.

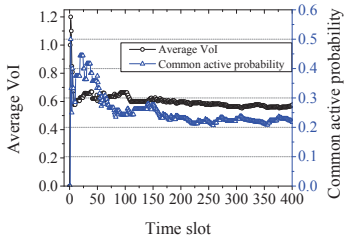


Fig. 8. Experimental common active time.

by a given algorithm X under the constraint of the variable harvested energy E in a period T. The expectation of $I_X(E)$ is denoted by $\mathbb{E}[I_X(E)]$. This paper always sticks the superscript "*" to any instance that is the optimum. Suppose that X^* is the optimal algorithm for our problem, *i.e.*

$$X^* = \arg\max_X \mathbb{E}[I_X(E)] \tag{18}$$

Thus, the regret $R_X(E)$ of the algorithm X can be formally defined as [16]:

$$R_X(E) = \mathbb{E}[I_{X^*}(E)] - \mathbb{E}[I_X(E)] \tag{19}$$

which is represented by the expectation of the arm with the maximal reward, *i.e.* $\mathbb{E}[I_{X^*}(E)] = \max_{a_j=1,2,3} \mathbb{E}[I_j(E)]$, because of the hardness to find the optimal scheme.

Before analyzing the regret of our scheme ODC, we introduce the Hoeffding inequality as follows:

The Hoeffding inequality—Let $x_1, \cdots, x_n$ be random variables with common range [0,1] such that $\mathbb{E}[x_t|x_1, \cdots, x_{t-1}] = \mu$. Let $S_n = \frac{1}{n}(x_1 + \cdots + x_n)$. Then, we have the probability $p(S_n \geq \mu + a) \leq e^{-2na^2}$ and $p(S_n \leq \mu - a) \leq e^{-2na^2}$ for the constant $a > 0$.

Recall that the harvested energy power must be higher than the threshold e_t. It then can support the normal operation of the sensor node. Denote by T' the time slot set in which the power is higher than the threshold e_t. T' is determined by the energy harvesting process, and its expectation can be determined easily if its state transition probability is previously known. By Algorithm 2, the VoI threshold is continuously adjusted so each sensor node may choose to sleep (*i.e.* to store the harvested energy) in some slots when the harvested energy is higher than the threshold e_t. Because of the charge efficiency $\lambda < 1$, the amount of the time slots, denoted by $|T_a|$, in which the harvested energy can support the normal operation of the sensor node under Algorithm 2 must not be higher than $|T'|$. Thus, we have $|T_a| \leq |T'| \leq |T|$.

Firstly, we analyze the expected times that the arm a_j, $j = 1, 2$ or 3 is pulled. The arm a_4 (storing energy) is not included since it does not return any reward. This is given in Lemma 2. We prove the following lemma based on the idea of the reference [16] while considering the cost of each arm c_j, $j = 1, 2, 3$.

Lemma 2: For an arbitrary arm a_j, $j = 1, 2$ or 3, the expected times to be pulled in a period T satisfy:

$$\mathbb{E}[\varphi_j(T)] < (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_j^2} + 2 \tag{20}$$

where Δ_j is the difference of the expected reward between the optimal algorithm X^* and the arm a_j. $c_{max} = \max_{j=1,2,3} c_j$ an $c_{min} = \min_{j=1,2,3} c_j$.

Proof: Recall that the step 9 and 10 of Algorithm 1 indicates that each arm a_j, $j = 1, 2, 3$, is pulled once in the first K slots. Thus, the times to pull a_j is $\varphi_j(T_a) = 1 + \sum_{\tau=K+1}^{T_a} \chi_j(\tau)$, where $\tau \in T_a$. Since Algorithm 1 is a greedy algorithm, the selected arm has the higher upper-band VoI per unit cost over other arms including the optimal one in each slot $\tau \in T_a$. So we have the following condition: $\bar{I}_j(\tau) + \delta_j(\tau) \geq \bar{I}^*(\tau) + \delta^*(\tau)$, *i.e.* $(\bar{I}'_j(\tau) + \delta'_j(\tau))/c_j \geq (\bar{I}^{*'}(\tau) + \delta^{*'}(\tau))/c^*$. In order to satisfy the condition with high probability, at least one of the following inequalities must be satisfied.

$$\bar{I}^*(\tau) + \delta^*(\tau) \leq u^* \tag{21}$$

$$\bar{I}_j(\tau) + \delta_j(\tau) \geq u_j \tag{22}$$

$$u^{*'}/c^* < (u'_j + \delta'_j(\tau))/c_j \tag{23}$$

where $u^{*'}$ and u'_j are the reward expectation of the optimal algorithm and the arm a_j by our algorithm, which is unknown to the sensor node. $u^* = u^{*'}/c^*$ and $u_j = u'_j/c_j$. By using the Hoeffding inequality, the probability that the inequalities in Equation (21) and (22) are satisfied is given as follows:

$$\begin{aligned} p(\bar{I}^*(\tau) \leq u^* - \delta^*(\tau)) &\leq e^{-4\ln \Psi(\tau)} = \Psi(\tau)^{-4} \\ p(\bar{I}_j(\tau) \geq u_j - \delta_j(\tau)) &\leq e^{-4\ln \Psi(\tau)} = \Psi(\tau)^{-4} \end{aligned} \tag{24}$$

Recall that $c_j > 0$ and $u^{*'}, u'_j \geq 0$, and then the inequality in Equation (23) implies:

$$\begin{aligned} c_j u^{*'} < c^*(u'_j + \delta'_j(\tau)) &\Rightarrow \varphi_j(\tau) < \frac{c^{*2}\varepsilon' \ln \Psi(\tau)}{(c_j u^{*'} - c^* u'_j)^2} \\ &\Rightarrow \varphi_j(\tau) < \begin{cases} \dfrac{\varepsilon' \ln \Psi(\tau)}{\Delta_j^2} & c_j \geq c^* \\ \dfrac{c^{*2}\varepsilon' \ln \Psi(\tau)}{(c_j \Delta_j)^2} & c_j < c^* \end{cases} \end{aligned} \tag{25}$$

where $\Delta_j = u^{*'} - u'_j$.

By Equation (24) and (25), the expectation of the times to pull the arm a_j thus can be given as follows:

$$\begin{aligned} &\mathbb{E}[\varphi_j(T)] = 1 + \sum_{\tau=K+1}^{T_a} \chi_j(\tau) \\ &< 1 + \max\{\frac{\varepsilon' \ln \Psi(T_a)}{\Delta_j^2}, \frac{c^{*2}\varepsilon' \ln \Psi(T_a)}{(c_j\Delta_j)^2}\} \\ &+ \sum_{\tau=K+1}^{T_a} \{p(\bar{I}^*(\tau) \leq u^* - \delta^*(\tau)) + p(\bar{I}_j(\tau) \geq u_j + \delta_j(\tau))\} \\ &\leq 1 + \frac{c^{*2}\varepsilon' \ln \Psi(T_a)}{(c_j\Delta_j)^2} + \sum_{\tau=K+1}^{T_a} 2\Psi(\tau)^{-4} \\ &\leq (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln \Psi(|T_a|)}{\Delta_j^2} + 2 \qquad (\text{as } |T_a \to |\infty) \\ &\leq (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_j^2} + 2 \qquad (|T_a| \leq |T'|) \end{aligned} \tag{26}$$

where $c_{max} = \max_{a_j:j=1,2,3} c_j$ and $c_{min} = \min_{a_j:j=1,2,3} c_j$ and $\varepsilon' \geq 1$. Notice that $\Psi(T')$ is the total number of times to pull all arms and only one arm can be pulled in each time slot so $\Psi(T') = |T'|$. ∎

Similarly, we can obtain that the expected times to pull the arm a_4.

Lemma 3: The expected times to pull the arm a_4 in a period T satisfy:

$$\mathbb{E}[\varphi_4(T)] < (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_4^2} + 1 \tag{27}$$

where $\Delta_4 = \min_{j=1,2,3} \Delta_j$.

Proof: According to the step 9 of Algorithm 1, the arm a_4 will be pulled when $\bar{I}(\tau) < I_d(\tau)$, which means that at least one of the following inequalities must be satisfied with high probability.

$$\bar{I}_j(\tau) + \delta_j(\tau) < u_j, \qquad \forall j = 1,2,3 \tag{28}$$

$$u^{*'}/c^* > (u_j' - \delta_j'(\tau))/c_j, \qquad \forall j = 1,2,3 \tag{29}$$

By using the Hoeffding inequality, the probability that the inequality in Equation (28) is satisfied is given as follows:

$$p(\bar{I}_j(\tau) < u_j - \delta_j(\tau)) < e^{-4\ln \Psi(\tau)} = \Psi(\tau)^{-4} \tag{30}$$

The inequality in Equation (29) implies:

$$\begin{aligned} c_j u^{*'} > c^*(u_j' - \delta_j'(\tau)) &\Rightarrow \varphi_4(\tau) < \frac{c^{*2}\varepsilon' \ln \Psi(\tau)}{(c_j u^{*'} - c^* u_j')^2} \\ &\Rightarrow \varphi_4(\tau) < \begin{cases} \dfrac{\varepsilon' \ln \Psi(\tau)}{\Delta_j^2} & c_j \geq c^* \\ \dfrac{c^{*2}\varepsilon' \ln \Psi(\tau)}{(c_4 \Delta_j)^2} & c_j < c^* \end{cases} \end{aligned} \tag{31}$$

According to the step 17 in Algorithm 1, the conditions given in Equation (28) and (29) should be satisfied for all arms a_j, $j = 1,2,3$ simultaneously. Therefore, by Equation (30) and (31), the expectation of the times to pull the arm a_4 thus can be given as follows:

$$\begin{aligned} \mathbb{E}[\varphi_4(T)] &= \sum_{\tau=K+1}^{T_a} \chi_4(\tau) \\ &< \max_{j=1,2,3} \max\{\frac{\varepsilon' \ln \Psi(T_a)}{\Delta_j^2}, \frac{c^{*2}\varepsilon' \ln \Psi(T_a)}{(c_j\Delta_j)^2}\} \\ &+ \sum_{\tau=K+1}^{T_a} \prod_{j=1}^{3} p(\bar{I}_j(\tau) < u_j + \delta_j(\tau)) \\ &\leq \max_{j=1,2,3} \frac{c^{*2}\varepsilon' \ln \Psi(T_a)}{(c_j\Delta_j)^2} + \sum_{\tau=K+1}^{T_a} \Psi(\tau)^{-16} \\ &\leq (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln \Psi(|T_a|)}{\Delta_4^2} + 1 \qquad (\text{as } |T_a \to |\infty) \\ &\leq (\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_4^2} + 1 \qquad (|T_a| \leq |T'|) \end{aligned} \tag{32}$$

where $a_4 = \min_{j=1,2,3} \Delta_j$, and $\varepsilon' \geq 1$. ∎

Recall that the harvested energy can support the normal operation of the sensor node in at most $|T'|$ slots, $T' \subseteq T$. By the lemma 2 and 3, we can analyze the reward regret of Algorithm 1.

Theorem 4: For the dynamic energy budget $E(T) > 0$, the expectation of ODC's regret is at most:

$$\sum_{j=1}^{K-1} [(\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_j} + 2\Delta_j] + u^{*'} [(\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln \Psi(|T'|)}{\Delta_4^2} + 1] \tag{33}$$

where $\varepsilon' \geq 1$ is a constant, and $u^{*'}$ is the reward expectation of the optimal algorithm.

Proof: Algorithm 1 can operate at the time slots in the set T_a, where $|T_a| \leq |T'|$. Suppose that $T_a = T_1 \cup T_2$. In the period T_1, the arm a_j, $j = 1,2,3$, are pulled, and in the period T_2, the arm a_4 is pulled. Suppose that the optimal algorithm operates at the time slots in the set T^*. $|T^*| \leq |T'|$ and $|T_a| \leq |T'|$ because the charge efficiency $\lambda < 1$. The reward regret of ODC is:

$$\begin{aligned} R_{ODC}(E) &= \mathbb{E}[I_{X^*}(E)] - \mathbb{E}[I_{ODC}(E)] \\ &= \mathbb{E}[\sum_{\tau=1}^{T} I^*(\tau)] - \mathbb{E}[\sum_{\tau=1}^{T} I(\tau)] = \mathbb{E}[\sum_{\tau=1}^{T^*} I^*(\tau)] - \mathbb{E}[\sum_{\tau=1}^{T_a} I(\tau)] \\ &= \mathbb{E}[\sum_{\tau=1}^{T_1} \sum_{j=1}^{K-1} (I^*(\tau) - I_j(\tau))] + \mathbb{E}[\sum_{\tau=1}^{T_2} (I^*(\tau) - I_4(\tau))] \\ &\quad + \mathbb{E}[\sum_{\tau=1}^{T^*-T_1-T_2} I^*(\tau)] \\ &\leq \mathbb{E}[\sum_{\tau=1}^{T_1} \sum_{j=1}^{K-1} (I^*(\tau) - I_j(\tau))] + \mathbb{E}[\sum_{\tau=1}^{T^*-T_1} (I^*(\tau) - I_4(\tau))] \\ &= \mathbb{E}[\sum_{\tau=1}^{T_1} \sum_{j=1}^{K-1} \Delta_j p(a(\tau) = a_j)] + \mathbb{E}[\sum_{\tau=1}^{T^*-T_1} I^*(\tau)\varphi_4(\tau)] \\ &= \mathbb{E}[\sum_{\tau=1}^{T_1} \sum_{j=1}^{K-1} \Delta_j \varphi_j(\tau)] + \mathbb{E}[\sum_{\tau=1}^{T^*-T_1} I^*(\tau)\varphi_4(\tau)] \\ &< \sum_{j=1}^{K-1} [(\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_j} + 2\Delta_j] \\ &\quad + u^{*'} [(\frac{c_{max}}{c_{min}})^2 \frac{\varepsilon' \ln |T'|}{\Delta_4^2} + 1] \end{aligned}$$

where $a(\tau)$ denotes the arm pulled at τ, and the reward of the arm a_4 is I_4, which is zero since to store energy cannot process data. This finishes the proof. ∎

VI. Experiment Results

This section depicts our experiments established on the real data obtained from the real solar harvesting module: SolarMote [2]. A series of experiments are designed and implemented to validate the performance of our scheme ODC by comparing with two baseline algorithms: COA and SDC, which are designed because of the hardness to find the optimal algorithm for the opportunistic duty cycling. The strong assumption behind COA is that data VoI and harvested energy can be previously known while no extra energy is consumed on the energy storage. COA is a centralized and off-line

algorithm. Thus, the performance of COA should be closer to the optimal algorithm than ODC and SDC. SDC predicts the amount of the energy to harvest and then calculates duty cycle in advance as the typical algorithm given in the reference [3]. In the following context, two scenarios: single sensor node and a network, are established to evaluate the performance of these algorithms. For the algorithm ODC and SDC, the charge efficiency $\lambda = 80\%$. The time slot τ is set to be 60 seconds. The energy threshold is set to be 20 mA. All experiments in this section are simulated on the network simulation platform OMNeT++ 4.1 (http://www.omnetpp.org/).

A. Single Node Scenario

This subsection simulates the scenario consisting of only one sensor node v_1 and the sink. v_1 samples data from its surrounding, and transmits its data to the sink. The scenario contains four experiments and is set up to evaluate the impact of the chance to harvest energy by excluding the impact of other factors occurring in large scale networks, such as the packet loss. Each experiment evaluates the reward performance of the three algorithms: COA, ODC and SDC.

In the first experiment, a certain amount 1 mAh of energy is previously assigned in the phase from the time slots 0 to 10. In the second experiment, 1 mAh is divided into two equivalent parts. One part is assigned to the phase from the time slots 0 to 5 while the other is assigned to the phase from the time slots 90 to 95. In the third experiment, 1 mAh is divided into 180 units, which are uniformly and randomly distributed into the period from the slot 0 to 200. In the three experiments, there is one data available in each time slot, and its VoI is assumed to follow the Gaussian probability distribution with the expectation 1 and the variance 0.5. In the fourth experiment, the data to process and the energy to harvest are the real data collected by the energy harvesting module SolarMote [2]: the luminous intensity and harvested energy in the first sub-figure of Figure 1(b). Assume that there is 20 mAh initial energy in the sensor node's battery. The simulations for each of the experiments are repeatedly run for 100 times so each data point in Figure 9∼12 is the average of 100 trials.

The results of the first and second experiments are respectively illustrated in Figure 9 and 10. These experiment results indicate the impact of the energy harvesting access on the total VoI. When the energy is assigned to certain phases, the sensor nodes tend to spend the energy timely at these phases by ODC since some extra energy must be consumed to store the energy. In Figure 9, the growth rate of v_1's total VoI by ODC is higher than those of COA and SDC at the initial phase. Although the total VoI under ODC slows down its growth in the first experiment in Figure 9 after the initial phase, the finally total VoI of ODC is 28.25% higher than SDC, and 28.86% lower than that of COA. In the second experiment, the fixed energy is assigned to two phases. During the two phases, the growth rate of total VoI by ODC suddenly increases since the two phases are considered to be good chance to use the energy by ODC. The finally total VoI of ODC 34.62% higher than SDC, and 24.02% lower that of COA as shown in Figure 10.

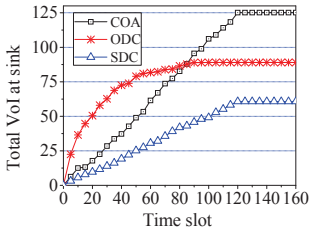


Fig. 9. A certain amount of energy is assigned to one phase.

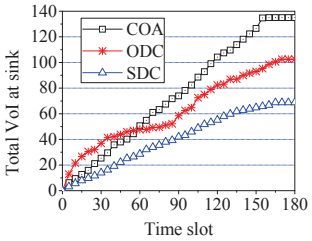


Fig. 10. A certain amount of energy is assigned equally to two phases.

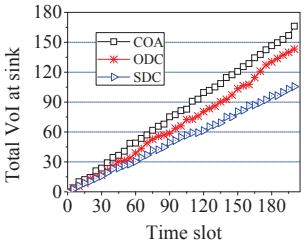


Fig. 11. Uniformly and randomly distributed data and energy.

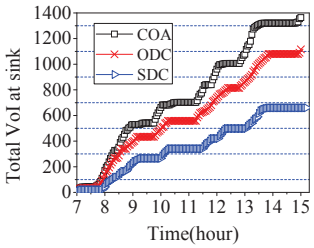


Fig. 12. The real data of luminous intensity and harvested energy.

The results of the third experiment is illustrated in Figure 11. The total VoI of all algorithms grows almost linearly with time. Till the time slot 200, the total VoI obtained by COA, ODC and SDC are 166.14, 143.195 and 105.509 respectively. ODC is 16.02% lower than COA, and 35.72% higher than SDC. Because the energy 1 mAh is distributed in the whole period from the slot 0 to 200 uniformly, there are much more chances that each sensor node need not store the harvested energy but to consume it directly. The three algorithms thus can obtain much more VoI than those in the first and second experiments.

In the fourth experiment, we adopt the real data sampled by SolarMote [2] including the luminous intensity and the harvested energy as shown in the first subgraph of Figure 1(b). Thus, the data that the sensor node v_1 will process is the luminous intensity. In each time slot, v_1 can harvest different amount of energy and luminous intensity, and cannot know the exact information of the luminous intensity in the future time slots. As shown in Figure 12, the experiment results illustrate that the finally total VoI of ODC is 18.18% lower than that of COA, and 69.09% higher than that of SDC. Different from the previous experiments, the accesses to harvest energy and the data are inhomogeneous over time. Compared to the results in the third experiment, the performance of ODC is better than SDC. Because of the inhomogeneous accesses, ODC loses some better chances to process the data and its performance decreases compared to the result in the third experiment.

In the four experiments, the occasions that the sensor node v_1 can harvest energy increase from one short phase as the first experiment to multiple moments as the third and fourth ones. In the third experiment, v_1 has uniformly possibility to harvest energy at each time slot while having the heterogeneous possibility in the fourth experiment. Through the four experiments, we find that the distribution of the accesses to harvest energy and to process data have much impact on the performance of the duty cycling scheme. The

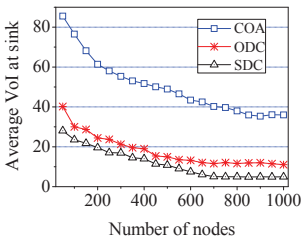


Fig. 13. A certain amount of energy is assigned to the initial phase of each sensor node.

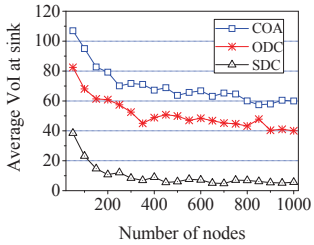


Fig. 14. A certain amount of energy is equally divided and randomly allocated to each time slot.

overall performance of ODC is closer to that of the centralized algorithm COA than SDC when the access to harvest energy and to process data distribute more uniformly. In the four experiments, the reward regret of ODC are given by comparing to the performance of COA because of the hardness to find the optimal scheme. The numerical values of the regret are presented in the form of the percentage in the above analysis.

B. Network Scenario

The node density also has obvious impact on the performance of these algorithms. This block simulates the network scenario composed of several numbers of the sensor nodes with the fixed deploy area of 100×100 m^2. Each sensor node runs the IEEE 802.15.4 protocol to assign wireless channel and to deal with the interference among the sensor nodes. All sensor nodes have the same receiving and minimal transmitting range as far as 50 meters. Each data point represents the average VoI when the network scale is a specific number of sensor nodes. For example, the most left data point in Figure 13 represents the average VoI of 50 sensor nodes. The simulation time is set to be 200 time slots and every slot is one minute. Two experiments are designed and implemented. In the first one, the energy 1 mAh is assigned to each sensor node at the initial phase from the time slot 0 to 10. In the second one, the energy 1 mAh is divided equally into 180 units and distributed to the 200 time slots uniformly and randomly. The simulation results of the two experiments are respectively shown in Figure 13 and 14. In these figures, each data point is the average of 10 trials.

From the results of both experiments in Figure 13 and 14, it is easy to notice that the increase of the node density results in the much decrease of VoI per sensor node. In the first experiment, the average VoI under the algorithms: COA, ODC and SDC drop 82.35%, 72.54% and 57.92% respectively when the number of the sensor nodes increases from 50 to 1000. Similarly, the average VoI drops 82.35%, 48.95% and 43.897% respectively in the second experiment. In spite of the VoI degression with the increasing of the node density, the performance of ODC is over that of the SDC. The average VoI of ODC can be 47. 01% of COA while SDC is at most 32.08% of COA in the first experiment. The average VoI of ODC is at least 67.89% of that by COA while SDC is at most 35.96% of COA in the second experiment.

The section evaluates the performance of our algorithm ODC by comparing with the centralized algorithm COA and the algorithm SDC representing the typical duty cycling scheme. The results show that the ODC has stable performance over the existing scheme. Although its performance is not as good as the centralized algorithm COA, ODC has acceptable performance as a distributed online algorithm. Through the experiments in the section, we found that the access to process data and to harvest energy have much impact on the performance of the duty cycling scheme, and thus the energy efficiency. ODC shows the impact under the requirement of the energy neutral operation, and has promising performance.

VII. Related Work

This section reviews the existing energy harvesting modules and the duty cycling schemes in energy-harvested WSNs. Energy harvesting technique and its applications in WSNs have been widely studied, and some modules were designed to harvest energy [18][19][20].

Energy harvesting module. Some typical modules were designed for sensor node to harvest solar, vibration and wind energy [21][22][23]. Most of the existing modules can harvest solar energy by the micro-scale photovoltaic power system. Actually, duty cycling is still necessary in energy-harvested WSNs because sensor nodes have no energy sufficient enough to sustain their continuous full duty cycle operation although they could harvest natural energy [4]. For example, the solar panel in the module Prometheus requires at least 4 hour hard light each day if the node's duty cycle is 10% [4]. Some modules were designed to store the harvested energy into their batteries while others took capacitor as the primary buffer and battery as the second buffer [4]. Battery suffers from low charge efficiency and long charging duration, and capacitor has high leakage [9]. For example, the 2000F ultra-capacitor has high leakage rate up to 43.8% during the first month [9].

Although the sensor network can obtain natural energy continuously, the harvested energy is not enough to support full duty cycle. This paper notices the phenomenon and argues to spend harvested energy on proper moments so that the energy efficiency can be improved while each node has enough energy to support the energy neutral operation.

Duty cycling. Duty cycling has been constantly researched as a promising technique to improve energy efficiency and prolong network life because of the energy limitation [5][24]. We can group the previous techniques into two classes: classical and adaptive duty cycling, according to the way under which the networks are powered. In WSNs, the classical duty cycling based on the preliminary assumption that each sensor has limited energy, *i.e.*, no extra energy is supplied [24][25]. So the goal of the classical duty cycling is to save energy as much as possible. In energy-harvested WSNs, each sensor node can be supplied with extra energy continuously by some energy harvesting modules [6][26]. Only a few works were engaged in adjusting duty cycle according to the weather conditions in order to achieve high energy efficiency [6][26][3]. However, they did not consider that natural energy arrives randomly [10], and the energy profile of each sensor node is time-varying and different from others [9]. Furthermore, these previous works adjusted duty cycle only by predicting the amount of

harvested energy in a duration, and did not consider the impact of network demand and the occasion to implement the demand on duty cycling.

Different from the previous works, this paper addresses the duty cycling under dynamic harvested energy to avoid the complex prediction or rough estimation of the duty cycle. We consider the spatiotemporal dynamic of harvested energy and our scheme ensures that each sensor node adjusts its duty cycle according to its local information on energy harvesting and data process.

VIII. Conclusion

This paper investigates the process of energy harvesting in solar sensor networks, and finds the phenomenon that it can greatly improve the energy efficiency of harvested energy to catch right chance to use it. We formulate it as the budget-dynamic MAB problem, and propose the new scheme: ODC, to exploit the harvested energy fully while satisfying the energy neutral operation. The theoretical performance for the scheme are analyzed, and the experimental analysis is also designed and implemented. We are the first to study the phenomenon, and will go on studying the problem in the following aspects. Because of the simplicity of sensor node hardware, it needs more simple and feasible schemes. Furthermore, the scheme of this paper involves some frequent updating for the probability to choose arms. Our future work will reduce the computation frequency and design even simple scheme requiring a few frequency to update the probability. Thirdly, the synchronization and communication coordination among nodes are quite helpful to improve the channel utilization. In the coming work, we will consider the coordination among sensor nodes.

References

[1] Q. Yang, S. He, J. Li, J. Chen, and Y. Sun. Energy-efficient probabilistic area coverage in wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 64(1):367–377, Jan 2015.

[2] X. Shen, C. Bo, J. Zhang, S. Tang, X. Mao, and G. Dai. EFCon: Energy flow control for sustainable wireless sensor networks. *Elsevier Ad Hoc Networks*, 11(4):1421–1431, 2013.

[3] B. Buchli, F. Sutton, J. Beutel, and L. Thiele. Dynamic power management for long-term energy neutral operation of solar energy harvesting systems. In *Proceeding of the 12th Conference on Embedded Networked Sensor Systems (SenSys)*, pages 31–45, Memphis, TN, USA, Nov. 3-6 2014. ACM.

[4] X. Jiang, J. Polastre, and D. Culler. Perpetual environmentally powered sensor networks. In *Proceeding of the 4th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 463–468, (UCLA) Los Angeles, CA, USA, April 25-27 2005. ACM/IEEE.

[5] G. Ghidini and S. K. Das. Energy-efficient markov chain-based duty cycling schemes for greener wireless sensor networks. *ACM Journal on Emerging Technologies in Computing Systems*, 8(4):29:1–29:32, November 2012.

[6] A. Kansal, J. Hsu, S. Zahedi, and M.B. Srivastava. Power management in energy harvesting sensor networks. *ACM Transactions on Embedded Computing Systems*, 6(4):32, 2007.

[7] C. Moser, L. Thiele, D. Brunelli, and L. Benini. Adaptive power management for environmentally powered systems. *IEEE Transactions on Computers*, 59(4):478–491, 2010.

[8] Y. Ding, R. Michelson, and C. Stancil. Battery state of charge detector with rapid charging capability and method, July 25 2000. US Patent 6,094,033.

[9] T. Zhu, Z. Zhong, Y. Gu, T. He, and Z.L. Zhang. Leakage-aware energy synchronization for wireless sensor networks. In *Proceeding of the 7th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*, pages 319–332, Krakw, Poland, June 22-25 2009. ACM.

[10] W.H.R. Chan, P. Zhang, I. Nevat, S.G. Nagarajan, A.C. Valera, H. Tan, and N. Gautam. Adaptive duty cycling in sensor networks with energy harvesting using continuous-time markov chain and fluid models. *IEEE Journal on Selected Areas in Communications*, 33(12):2687–2700, Dec 2015.

[11] K. Liu and Q. Zhao. Distributed learning in multi-armed bandit with multiple players. *IEEE Transactions on Signal Processing*, 58(11):5667 – 5681, 2010.

[12] P. Padhy, R.K. Dash, K. Martinez, and N.R. Jennings. A utility-based adaptive sensing and multihop communication protocol for wireless sensor networks. *ACM Transactions on Sensor Networks*, 6(3):27, 2010.

[13] J. Gittins, K. Glazebrook, and R. Weber. *Multi-armed bandit allocation indices*. Wiley, 2011.

[14] R. D. Smallwood and E. J. Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.

[15] R. Kohli, R. Krishnamurti, and P. Mirchandani. Average performance of greedy heuristics for the integer knapsack problem. *European Journal of Operational Research*, 154(1):36–45, 2004.

[16] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2):235–256, 2002.

[17] P. R. Kumar and P. Varaiya. *Stochastic systems: estimation, identification and adaptive control*. Prentice-Hall, Inc., 1986.

[18] S. He, J. Chen, D. K. Yau, H. Shao, and Y. Sun. Energy-efficient capture of stochastic events under periodic network coverage and coordinated sleep. *IEEE Transactions on Parallel and Distributed Systems*, 23(6):1090–1102, 2012.

[19] S. He, J. Chen, F. Jiang, D. K. Yau, G. Xing, and Y. Sun. Energy provisioning in wireless rechargeable sensor networks. *IEEE Transactions on Mobile Computing*, 12(10):1931–1942, 2013.

[20] Y. Zhang, S. He, and J. Chen. Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks. *IEEE/ACM Transactions on Networking*, PP(99):1–15, 2015. DOI: 10.1109/TNET.2015.2425146.

[21] R.S. Liu, K.W. Fan, Z. Zheng, and P. Sinha. Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Transactions on Networking*, 19(4):947–960, 2011.

[22] O. Ozel and S. Ulukus. Achieving AWGN capacity under stochastic energy harvesting. *IEEE Transactions on Information Theory*, 58(10):6471–6483, 2012.

[23] A. Castagnetti, A. Pegatoquet, Trong Nhan Le, and M. Auguin. A joint duty-cycle and transmission power management for energy harvesting wsn. *IEEE Transactions on Industrial Informatics*, 10(2):928–936, May 2014.

[24] Z. Li, M. li, and Y. Liu. Towards energy-fairness in asynchronous duty-cycling sensor networks. In *Proceeding of the 31st Annual IEEE International Conference on Computer Communications (INFOCOM)*, pages 801–809, Orlando, Florida USA, March 25-30 2012. IEEE.

[25] S. Guo, L. He, Y. Gu, B. Jiang, and T. He. Opportunistic flooding in low-duty-cycle wireless sensor networks with unreliable links. *IEEE Transactions on Computers*, 63(11):2787–2802, 2014.

[26] C.M. Vigorito, D. Ganesan, and A.G. Barto. Adaptive control of duty cycling in energy-harvesting wireless sensor networks. In *Proceeding of the 4th Annual Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 21–30, San Diego, California, USA, June 18-21 2007. IEEE.