Near-optimal Sleep Scheduling of a Rechargeable Sensor under Temporal Correlations

Neeraj Jaggi, Koushik Kar

Abstract

In this brief survey, we review and summarize our recent work on efficient sleep scheduling of a rechargeable sensor deployed to detect events that exhibit temporal correlations. We pose the dynamic sensor sleep scheduling question in a stochastic optimization framework. Under complete state observability, we show that the optimal sleep scheduling policy has a very simple structure. In a more realistic scenario where a sleeping sensor cannot observe the system state, we show that a simple policy using an appropriately chosen fixed sleep duration attains near-optimal performance. Finally, we argue that an aggressive wakeup policy is optimal if and only if there are no temporal correlations in the event process.

I. INTRODUCTION AND SYSTEM MODEL

We consider a sensor that has been deployed to detect interesting phenomena (“events”) that exhibit correlation in time. The sensor is assumed to be rechargeable, and the battery of the sensor is recharged according to a random recharging process. The rate of energy recharge is typically significantly lower than the rate of energy discharge, necessitating the sensor to deactivate itself (“go to sleep”) periodically. We consider the problem of optimal sleep scheduling (sensor activation) in this scenario with the goal of maximizing the time-average event detection probability.

We model a rechargeable sensor as an energy bucket that stores energy in units of a quantum. We assume a discrete time model, where in each time slot, a recharge event occurs with a probability $q$ and charges the sensor with a constant charge of $c$ quanta. The sensor in active state expends a charge of $\delta_1$ quanta (operational cost) during each time slot it is active, irrespective of occurrence of an event of interest during the time slot. An event of interest occurs randomly in a time slot and discharges the sensor (if active) by an additional charge of $\delta_2$ quanta (detection cost). We assume $\delta_1 \geq c$, and $\delta_2 \geq \delta_1$. Let $\rho = \frac{\delta_2}{\delta_1}$. A sensor is considered available for activation as long as it has sufficient energy ($\geq \delta_1 + \delta_2$) to provide coverage for at least one time slot.

The extent of temporal correlation, in events that the sensor is expected to detect, is specified using correlation probabilities $p_{c}^{\text{on}}$ and $p_{c}^{\text{off}}$, such that $\frac{1}{2} < p_{c}^{\text{on}}, p_{c}^{\text{off}} < 1$. If an interesting event occurred during time slot $t$, then in the next time slot $(t+1)$ a similar event occurs with probability $p_{c}^{\text{on}}$, while no event occurs with probability $1 - p_{c}^{\text{on}}$. Similarly, if no event occurred during the current time slot, no event occurs in the next time slot with probability $p_{c}^{\text{off}}$. The event occurrence process comprises of an alternating sequence of periods where events occur (On period) and do not occur (Off period). In practice, the Off periods are expected to be significantly longer than the On periods, which implies $p_{c}^{\text{off}} \geq p_{c}^{\text{on}}$, an assumption that we make in our analysis.

The sensor operating under an activation policy $\Pi$ takes activation decisions in each time slot depending upon (a) the current energy level of the sensor, and (b) the knowledge about the state of event occurrence in the system. Let $E_o(T)$ denote the total number of events that occur in the sensor’s sensing region during time interval $[0 \ldots T]$. Let $E_d(T)$ denote the total number of events detected during this interval by sensor operating under sensor activation policy $\Pi$. The quality of coverage (or the event detection probability) in the region, denoted $U(\Pi)$, is measured in terms of the time-average fraction of events detected, i.e.,

$$U(\Pi) = \lim_{T \to \infty} \frac{E_d(T)}{E_o(T)}.$$
In this paper, we address the question of how this rechargeable sensor node should be activated so that the event detection probability \( U \) is maximized. In other words, our goal is to choose the activation policy \( \Pi \) such that \( U(\Pi) \) is maximized.

The expected length of an \( \text{Off} \) period is given by \( \frac{1}{1-p_{\text{c}}^{\text{off}}} \) [3]. Similarly, the expected length of an \( \text{On} \) period is given by \( \frac{1}{1-p_{\text{c}}^{\text{on}}} \). Using Markov chain analysis, as \( T \to \infty \), the steady-state probability of event occurrence equals \( \pi_{\text{off}} = \frac{1-p_{\text{c}}^{\text{off}}}{2-p_{\text{c}}^{\text{on}}-p_{\text{c}}^{\text{off}}} \left( \pi_{\text{off}} = 1 - \pi_{\text{on}} \right) \). Note that, \( p_{\text{c}}^{\text{off}} \geq p_{\text{c}}^{\text{on}} \) implies \( \pi_{\text{off}} \geq \pi_{\text{on}} \).

II. METHODOLOGY AND CONTRIBUTIONS

We consider two different system observability scenarios, namely Completely Observable and Partially Observable. In the completely observable system, the sensor has perfect information about the event occurrence state in the system at all times. More specifically, the sensor is able to observe the system state even when it is inactive. In practice, once the sensor deactivates itself, it may not be able to observe the state of event occurrence in the system during the time slots in which it is inactive, leading to partial observability. Therefore, the sensor would be required to take activation decisions under imperfect information about the system state. The structure of the optimal policy with complete information is similar with that corresponding to incomplete information, and provides the important insights that help us develop simple, near-optimal activation policies for partially observable systems.

We formulate the sensor activation under complete observability as a Markov Decision Process [6] and solve for the optimal policy. We derive a tight bound on achievable performance measured in terms of \( U(\Pi) \) (Lemma 3.1) and provide a simple activation policy that achieves this bound and is therefore optimal (Theorem 3.2). Next, we formulate the sensor activation problem under partial observability as a Partially Observable Markov Decision Process (POMDP) [1]. The challenge here is to transform this problem into an equivalent, completely observable MDP with the same optimal reward and actions as the POMDP [2]. We provide a near-optimal solution to optimality equations (Lemma 4.1), and characterize some properties of the corresponding activation policy (Lemma 4.2). We empirically find the structure of the optimal policy for the POMDP, and focus on developing a simple and efficient near-optimal policy with performance guarantees (Theorem 4.4). In the process, we provide performance bounds for AW (Aggressive Wakeup) and CW (Correlation-dependent Wakeup) policies (Lemmas 4.3 and 4.5), and show that AW policy performs optimally in the absence of temporal correlations (Lemmas 5.1 and 5.2). In this paper, we provide an overview of the important results. More detailed analysis and proofs of these results can be found in [3], [4], [5].

III. ACTIVATION UNDER PERFECT STATE INFORMATION

Let the system state at time \( t \) be \( X_t = (L_t, E_t) \), where \( L_t \in \{0, 1, 2, \ldots \} \) represents the energy level of the sensor at (the end of) time \( t \), and \( E_t \in \{0, 1\} \) equals one if an event occurred at time \( t \); zero otherwise. Each time slot is a decision epoch (i.e., an activation decision is taken at the end of time \( t \), \( \forall t \)). The action taken at (the end of) time \( t \) is denoted by \( u_t \in \{0, 1\} \), where \( u_t = 0 \) (\( u_t = 1 \)) corresponds to sensor deactivation (activation) during time interval \( [t, t+1] \). For the sake of analysis, we assume that the recharge during a time slot takes effect at the end of a time slot, just before the activation decision for the next time slot is taken. Since the next state of the system depends only upon the current state and the action taken, the system constitutes a Markov Decision Process (MDP). The sensor gains a reward of one if \( u_t = 1 \), and \( E_{t+1} = 1 \).

Lemma 3.1: For any stationary activation policy \( \Pi^{\text{opt}} \), \( U(\Pi^{\text{opt}}) \) is upper bounded as follows:

(i) If \( qc \leq \pi_{\text{on}}(\delta_1 + p_{\text{c}}^{\text{on}}\delta_2) \), then \( U(\Pi^{\text{opt}}) \leq \left( \frac{1}{\pi_{\text{on}}} \right) \left( \frac{p_{\text{c}}^{\text{on}}qc}{\delta_1 + p_{\text{c}}^{\text{on}}\delta_2} \right) \);

(ii) If \( qc > \pi_{\text{on}}(\delta_1 + p_{\text{c}}^{\text{on}}\delta_2) \), then \( U(\Pi^{\text{opt}}) \leq \left( \frac{1}{\pi_{\text{on}}} \right) \left( \frac{qc(1 - p_{\text{c}}^{\text{off}}) + \pi_{\text{on}}\delta_1(p_{\text{c}}^{\text{on}} + p_{\text{c}}^{\text{off}} - 1)}{\delta_1 + \delta_2(1 - p_{\text{c}}^{\text{off}})} \right) \).
Theorem 3.2: The sensor activation policy $\Pi'$ with actions $u^*$ defined by

$$ u^* = \begin{cases} 
1 & \text{if } L \geq \delta_1 + \delta_2, \text{ and } E = 1, \\
1 \text{ with prob. } P^* & \text{if } L \geq \delta_1 + \delta_2, \text{ and } E = 0, \\
0 & \text{otherwise},
\end{cases} $$

is optimal where $P^* = \frac{\max[0, qc - (\delta_1 + p_{on}^c \delta_2)]}{\pi on(\delta_1 + p_{on}^c + \delta_2)}$. 

Note that if $qc \leq \pi on[\delta_1 + p_{on}^c \delta_2]$ (the “low recharge rate” case), then the policy described above simply reduces to $u^* = 1$ if $L \geq \delta_1 + \delta_2$, and $E = 1$; 0 otherwise. Thus the sensor activates itself if and only if the event process is in the On state at the current time. If $qc > \pi on[\delta_1 + p_{on}^c \delta_2]$ (the “high recharge rate” case), however, in addition to activating itself (with probability 1) when the event process in On, the sensor also activates itself with probability $P^*$ when the event process is Off. It is worth noting that $P^*$ in the optimal activation policy is chosen such that the rate of energy discharge equals the rate of energy recharge in the system. The near-optimal policy developed later in this paper for the case of imperfect state information is also intuitively motivated by this notion of energy balance.

IV. Activation under Imperfect State Information

Since the system state is not completely observable, the optimal action depends on the current and past observations, and on past actions. It has been shown that the POMDP can be formulated as a completely observable MDP [2], with the same finite action set. The solution to the equivalent MDP with complete information provides us with the optimal actions to take (in the POMDP) and with the optimal reward. The state space of the equivalent MDP, denoted $\Delta$, comprises of the space of probability distributions on the original state space, which may lead to a possibly uncountable or infinite state space. However, the structure of the original POMDP, in most cases, allows for the existence of solutions to the average cost (reward) optimality equation [2]. In the case of sensor activation under partial observations too, the structure of the POMDP leads to a countable state space for the equivalent MDP, guaranteeing existence of optimal solution.

The state of the equivalent MDP at time $t$, denoted $Z_t$, can be represented in the form $(L, E, i)$ [3]. The state $Z = (L, E, i)$ represents the following: (a) The sensor has been inactive for the last $i$ time slots, (b) The state of the event process observed when the sensor was last active is $E$, and (c) The current charge level of the sensor equals $L$. The POMDP is transformed to an equivalent completely observable MDP with state-space $\Delta$. The optimality equations for this MDP are given by [2]:

$$ \Gamma^* + h^*(Z) = \max_{u \in U} \left[ \bar{r}(Z, u) + \sum_{y \in Y} V(y, Z, u)h^*(W(y, Z, u)) \right], \forall Z \in \Delta, \quad (1) $$

where $V, W$ are defined in [3].

Lemma 4.1: The values of $\Gamma^*$, and $h^*$ given by $h^*((L, 0, i)) = h^*((L, 1, i)) = \alpha L \forall i \geq 0$, $\Gamma^* = \alphaqc$ where $\alpha = \frac{1}{\delta_2 + \delta_1 (1 + \frac{1}{\pi on})}$, satisfy the optimality equations for the POMDP within an error of $\epsilon \sim O \left( \frac{1}{\rho} \right)$.

In the above, $\epsilon \sim O \left( \frac{1}{\rho} \right)$ implies that the error $\epsilon$ approaches zero as $\rho = \frac{\delta_2}{\delta_1}$ becomes large. Thus the average reward attained per stage, given by $\Gamma^* = \frac{qc}{\delta_2 + \delta_1 (1 + \frac{1}{\pi on})}$, is within $O \left( \frac{1}{\rho} \right)$ of that attained by the optimal policy. As $T \to \infty$, $\frac{\Gamma^*}{\pi on T} \rightarrow \pi on$ and the fraction of events detected by the sensor in this case (which must also be within $O \left( \frac{1}{\rho} \right)$ of that attained by the optimal policy) is characterized as,

$$ \lim_{T \to \infty} \frac{\Gamma^*}{\pi on T} = \frac{\Gamma^*}{\pi on} = \frac{qc}{\pi on (\delta_2 + \delta_1) + \delta_1 (1 - p_{on}^c)}. \quad (2) $$
Lemma 4.2: The $\epsilon$-optimal policy whose solution is given by Lemma 4.1 satisfies the following properties, $\forall L \geq \delta_1 + \delta_2$ (where $K$ is the size of the sensor energy bucket):

(i) $\mu^*(((L, 1, 0))) = 1$ and $\mu^*((L, 0, 0)) = 0$ (for $L \leq K - c$),

(ii) $\mu^*((L, 1, i)) = 1 \Rightarrow \mu^*((L, 1, i - 1)) = 1, \forall i \geq 1,$ and

(iii) $\mu^*((L, 0, i)) = 1 \Rightarrow \mu^*((L, 0, i + 1)) = 1, \forall i \geq 0.$

Next, we develop a deterministic and memoryless activation algorithm and characterize its performance. A Correlation-dependent Wakeup (CW) policy satisfies the following criteria: (i) An active sensor with sufficient energy remains active if an event occurred during the previous time slot; (ii) An active sensor goes to sleep for an arbitrary duration of time if no event occurred during the previous time slot. At the end of the sleep duration, the sensor activates itself to poll the system state. Let $P_e$ denote the probability that the event occurrence process is On when the sensor wakes up from sleep and polls the system. An energy-balancing CW policy, $\Pi_{EB-CW}$, employs an energy-balancing sleep duration $SI^*$, which enables the sensor to spend as much energy during the course of its operation, as is gained by the sensor through recharge. It can be shown that [3],

$$SI^* = \frac{\delta_1}{qc} + \frac{P_e^*(\delta_1 + \delta_2 - qc)}{qc (1 - p_e^on)} - 1.$$  \hspace{1cm} (3)

where $P_e^*$ is the event detection probability ($P_e$) corresponding to $\Pi_{EB-CW}$.

Lemma 4.3: The maximum achievable performance for any CW policy $\Pi_{CW}$ is upper bounded as,

$$U(\Pi_{CW}) \leq \frac{q_c}{\pi^on(\delta_2 + \delta_1) + \delta_1 (1 - p_e^on)} = U_{CW}^*.$$  

Theorem 4.4: Energy-balancing CW policy, $\Pi_{EB-CW}$, achieves the following performance bounds:

(i) $U(\Pi_{EB-CW}) \geq \frac{\rho + 1}{\rho + \frac{1}{\pi^on}} \left(\frac{1}{\pi^on}\right) \left(\frac{p_e^onqc}{\delta_1 + p_e^on\delta_2}\right) \geq \frac{\rho + 1}{\rho + \frac{1}{\pi^on}} U_{CW}^*;$

(ii) $U(\Pi_{EB-CW}) \geq \frac{P_e^*}{\pi^on} U_{CW}^*.$

Thus $\Pi_{EB-CW}$ achieves near-optimal performance. Note that Theorem 4.4(i) implies that $U(\Pi_{EB-CW})$ approaches $U_{CW}^*$ as $\rho$ becomes large. When $\rho$ is relatively small, however, Theorem 4.4(ii) may provide a tighter performance bound.

Operating under the Aggressive-wakeup (AW) policy, the sensor switches itself on whenever possible. In other words, the sensor activates itself in a time slot if its current charge level is $\geq \delta_1 + \delta_2$.

Lemma 4.5: The maximum achievable performance for AW policy, $U_{AW}^* = \frac{qc}{\delta_1 + p\delta_2}$.

V. ABSENCE OF TEMPORAL CORRELATIONS

We briefly discuss the case where the occurrence of events of interest is not correlated in time. An event occurs in a time slot with probability $p$, independent of occurrence of events in previous time slots. Note that $p_e^on = p_e^off = 0.5$ corresponds to the case when $p = 0.5$. Note that the following results apply to both completely observable as well as partially observable systems. The average recharge rate of the sensor is $qc$, while the average discharge rate of an active sensor is $\delta_1 + p\delta_2$.

Lemma 5.1: In absence of temporal correlations across events, the performance for any activation policy is upper bounded by $\frac{qc}{(\delta_1 + p\delta_2)}$.

Lemma 5.2: AW policy achieves optimal performance in the absence of temporal correlations across events.
REFERENCES


