Online Age-minimal Sampling Policy for RF-powered IoT Networks

Mohamed A. Abd-Elmagid, Harpreet S. Dhillon, and Nikolaos Pappas

***Abstract*—In this paper, we study a real-time Internet of Things (IoT)-enabled monitoring system in which a source node (e.g., IoT device or an aggregator located near a group of IoT devices) is responsible for maintaining the freshness of information status at a destination node by sending update packets. Since it may not always be feasible to replace or recharge batteries in all IoT devices, we consider that the source node is powered by *wireless energy transfer* (WET) by the destination. For this system setup, we investigate the optimal online sampling policy that minimizes the long-term average Age-of-Information (AoI), referred to as the *age-optimal policy*. The age-optimal policy determines whether each slot should be allocated for WET or update packet transmission while considering the dynamics of battery level, AoI, and channel state information (CSI). To solve this optimization problem, we model this setup as an average cost Markov Decision Process (MDP). After analytically establishing the monotonicity property of the value function associated with the MDP, the age-optimal policy is proven to be a threshold-based policy with respect to each of the system state variables. We extend our analysis to characterize the structural properties of the policy that maximizes average throughput for our system setup, referred to as the *throughput-optimal policy*. Afterwards, we analytically demonstrate that the structures of the age-optimal and throughput-optimal policies are different. We also numerically demonstrate these structures as well as the impact of system design parameters on the optimal achievable average AoI.**

***Index Terms*—Age-of-Information, Internet of things, RF energy harvesting, Markov Decision Process.**

I. Introduction

The performance of many real-time IoT-enabled applications is driven by how fresh the collected data measurements of the IoT devices are when they reach the destination nodes [1]. The timely delivery of the measurements to the destination nodes is greatly restricted by the energy-constrained nature of the IoT devices. To enable a self-perpetuating operation of IoT networks, radio-frequency (RF) energy harvesting has emerged as an appealing solution for charging low-power IoT devices due to its ubiquity and cost efficient implementation. This necessitates the design of efficient transmission policies for freshness-aware RF-powered IoT networks, which is the main objective of this paper.

Related work. To quantify freshness of information at the destination node, we use AoI as a performance metric. The authors of [2] introduced the concept of AoI and characterized average AoI for a simple queueing-theoretic model. Building on this, a series of works [3]–[6] focused on characterizing the average AoI and its variations (e.g., Peak Age-of-Information [4] and Value of Information of Update [6]) for adaptations of the queueing model studied in [2]. Another direction of research [7]–[20] focused on applying various tools from optimization theory to characterize age-optimal transmission policies for different communication systems that deal with time critical information.

The offline/online age-optimal policy for an energy harvesting source node was investigated under various system settings in [13]–[20]. Note that the online age-optimal policy is obtained when a causal knowledge of energy arrivals is assumed. A common model of the energy harvesting process in [13]–[18] is an external point process (e.g., Poisson process) independent from all system design parameters. In contrast, when the source is powered by RF energy harvesting, as considered in this paper, the harvested energy is a function of the temporal variation of the CSI. This, in turn, means that the proposed age-optimal policies in [13]–[18] are not directly applicable to such system settings since one needs to explicitly incorporate the statistics of CSI in the process of decision-making. Hence, the analysis of characterizing the age-optimal policies becomes more challenging. Before going into more details about our contributions, it is instructive to note that the problem of age-optimal policy in wireless powered communication systems has been studied very recently in [19], [20]. However, neither of the proposed policies took into account the evolution of the battery level at the source and the variation of CSI over time in the process of decision-making. Different from these, this paper makes the first attempt to: *1) characterize the online age-optimal sampling policy while considering the dynamics of battery level, AoI and CSI, and 2) establish analytically key differences between the structures of the online age-optimal and throughput-optimal polices.*

Contributions. This paper studies a real-time monitoring system in which an RF-powered source node transmits status update packets to a destination node over time to keep its information status as fresh as possible. For this setup, we study the long-term average AoI minimization problem in which WET and scheduling of update packet transmissions are jointly optimized. We model the problem as an average cost MDP for which we prove the monotonicity property of its associated value function analytically. Using this, we show that the age-optimal policy is a threshold-based policy with respect to each of the system state variables, i.e., the battery level, AoI, and channel power gains. We further study the average throughput maximization problem for our system setup, and demonstrate analytically the difference between the structures of the age-optimal and throughput-optimal polices. Our numerical results verify the analytical findings and further demonstrate the impact of system design parameters on the optimal achievable average AoI.

M. A. Abd-Elmagid and H. S. Dhillon are with Wireless@VT, Department of ECE, Virginia Tech, Blacksburg, VA (Email: {maelaziz, hdhillon}@vt.edu). N. Pappas is with the Department of Science and Technology, Linköping University, SE-60174 Norrköping, Sweden (Email: nikolaos.pappas@liu.se). The support of the U.S. NSF (Grant CPS-1739642) is gratefully acknowledged.

II. System Model

A. Network Model

We consider a monitoring system in which a source node is deployed to observe some physical process, such as temperature or humidity. The source node is supposed to keep the information status of its observed process at a destination node fresh by sending status update packets. In the context of IoT networks, the source node could refer to an aggregator located near a group of IoT devices, which transmits update packets collected from them to the destination node (for instance, a cellular base station). We use the concept of AoI to quantify freshness of information at the destination node. Formally, AoI is defined as the time elapsed since the recently received update packet at the destination was generated at the source [2]. The destination node is assumed to have a stable energy source whereas the source node is equipped with an RF energy harvesting circuitry as its only source of energy. Specifically, the source harvests energy from the RF signals transmitted by the destination in the downlink. The energy harvested at the source is stored in a battery with finite capacity $B_{\max}$ joules. The source and destination nodes are assumed to have a single antenna each, and operate over the same frequency channel. Hence, the source cannot simultaneously harvest wireless energy in downlink and transmit data in uplink.

Without loss of generality, the time horizon is partitioned into slots of unit length such that slot $k = 0, 1, \ldots$ corresponds to the time duration $[k, k+1)$. We denote by $B(k)$ and $A(k)$ the amount of available energy in the battery at the source and the AoI at the destination, respectively, at the beginning of time slot k. We assume that $A(k)$ is upper bounded by a finite value $A_{\max}$ which can be chosen to be arbitrarily large, i.e., $A(k) \in \{1, 2, \cdots, A_{\max}\}$. When AoI reaches $A_{\max}$, it means that the information is too stale to be of any use at the destination node. Let $g(k)$ and $h(k)$ denote the downlink and uplink channel power gains between the source and destination nodes over slot k, respectively. The channels are assumed to be affected by quasi-static flat fading, i.e., they remain constant over a time slot but change independently from one slot to another. The destination node has perfect knowledge about the channel power gains in the current time slot, and only a statistical knowledge for future slots.

B. State and Action Spaces

At the beginning of an arbitrary time slot k, the state of the system $s(k)$ is characterized by the battery level at the source, the AoI value at the destination, and the uplink and downlink channel power gains, i.e., $s(k) \triangleq (B(k), A(k), g(k), h(k)) \in \mathcal{S}^a$; where $\mathcal{S}^a$ is the state space which contains all the combinations of $B(k), A(k), g(k)$, and $h(k)$, and the superscript a indicates that it is defined for the average AoI minimization problem. Based on $s(k)$, the action taken at slot k is given by $a(k) \in \mathcal{A} \triangleq \{H, T\}$. When $a(k) = H$, slot k is dedicated for wireless energy transfer where the destination broadcasts RF energy signal in the downlink to charge the battery at the source. The amount of energy harvested by the source can be expressed as $E^{\mathrm{H}}(k) = \eta P g(k)$, where η is the efficiency of the energy harvesting circuitry and P is the average transmit power by the destination. On the other hand, when $a(k) = T$, slot k is allocated for information transmission where the source sends an update packet about its observed process to the destination. We consider a *generate-at-will policy* [7], where whenever a time slot is scheduled for information transmission, the source generates an update packet at the beginning of that time slot. According to Shannon's formula, when the energy consumed by the source to transmit an update packet of size S in slot k is $E^{\mathrm{T}}(k)$, its maximum transmission rate is $\log_2(1+\frac{h(k)E^{\mathrm{T}}(k)}{\sigma^2})$ bits/Hz (recall that the slot length is unity), where σ^2 is the noise power at the destination. Hence, the action T can only be decided if the battery level at the source satisfies the following condition

$$B(k) \geq E^{\mathrm{T}}(k) = \frac{\sigma^2}{h(k)} \left(2^{S/W} - 1\right), \tag{1}$$

where W is the channel bandwidth.

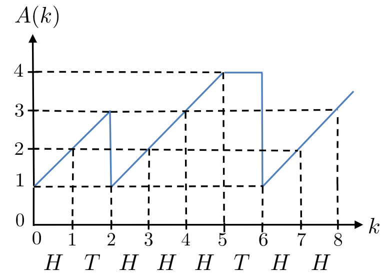

Fig. 1. AoI evolution vs. time when $A_{\max} = 4$.

In every time slot, the battery level at the source and the AoI at the destination are updated based on the action decided. Specifically, if $a(k) = T$, then the battery level decreases by $E^{\mathrm{T}}(k)$, and AoI becomes one (recall that a generate-at-will policy is employed); otherwise, the battery level increases by $E^{\mathrm{H}}(k)$ and AoI increases by one. Hence, the evolution of the battery level and AoI can be expressed, respectively, by

$$B(k+1) = \begin{cases} B(k) - E^{\mathrm{T}}(k), & \text{if } a(k) = T, \\ \min\left\{B_{\max}, B(k) + E^{\mathrm{H}}(k)\right\}, & \text{otherwise.} \end{cases} \tag{2}$$

$$A(k+1) = \begin{cases} 1, & \text{if } a(k) = T, \\ \min\left\{A_{\max}, A(k) + 1\right\}, & \text{otherwise.} \end{cases} \tag{3}$$

To help visualize (3), Fig. 1 shows the AoI evolution as a function of actions taken over time when $A_{\max} = 4$.

III. Problem Formulation and Structural Properties of the Age-optimal Policy

A. Problem Formulation

Our objective is to obtain the optimal policy, which specifies the actions taken at different states of the system over time, achieving the minimum long-term average AoI at the destination. Particularly, a policy $\pi = \{\pi_0, \pi_1, \cdots\}$ is a sequence of action probability measures over the state space. For instance, the probability measure π_k specifies the probability of taking action $a(k)$, conditioned on the sequence s^k which includes the past states and actions, and the current state, i.e., $s^k \triangleq \{s(0), a(0), \cdots, s(k-1), a(k-1), s(k)\}$. Formally, π_k specifies $\mathbb{P}(a(k) \mid s^k)$ such that $\sum_{a(k)\in\mathcal{A}(s(k))} \mathbb{P}(a(k) \mid s^k) = 1$, where $\mathcal{A}(s(k))$ is the set of possible actions at state $s(k) \in \mathcal{S}^a$. The policy π is said to be stationary when $\mathbb{P}(a(k) \mid s^k) = \mathbb{P}\left(a\left(k\right) \mid s\left(k\right)\right), \forall k$, and is called deterministic if $\mathbb{P}(a(k) \mid s^k) = 1$ for some $a(k) \in \mathcal{A}(s(k))$. Under a policy π, the long-term average AoI at the destination starting from an initial state $s(0)$ can be expressed as

$$\bar{A}^{\pi} \triangleq \limsup_{K\to\infty} \frac{1}{K+1} \sum_{k=0}^{K} \mathbb{E}\left[A(k) \mid s(0)\right], \tag{4}$$

where the expectation is taken with respect to the channel conditions and the policy. Our goal is to find the online age-optimal policy π^* such that

$$\pi^\star = \arg\min_{\pi} \bar{A}^{\pi}. \tag{5}$$

B. Proposed Solution

Due to the nature of evolution for the battery level and AoI, as described by (2) and (3), and the independence of channel power gains over time, the problem can be modeled as a Markov Decision Process (MDP). In order to obtain the age-optimal policy by applying standard optimization techniques such that the Value Iteration Algorithm (VIA) or the Policy Iteration Algorithm (PIA) [21], we discretize the battery level and the channel power gains. In particular, we denote by $b(k) \in \{0, 1, \cdots, b_{\max}\}$ the discrete battery level at slot k, where $b_{\max}$ represents the maximum amount of energy quanta that can be stored in the battery such that each energy quantum contains $\frac{B_{\max}}{b_{\max}}$ joules. In this case, the quantities $E^{\mathrm{T}}(k)$ and $E^{\mathrm{H}}(k)$ in (2) should be replaced by two integer variables expressed in terms of energy quanta. By letting, $e^{\mathrm{T}}(k) \triangleq \left\lceil \frac{b_{\max}}{B_{\max}} E^{\mathrm{T}}(k) \right\rceil$ and $e^{\mathrm{H}}(k) \triangleq \left\lfloor \frac{b_{\max}}{B_{\max}} E^{\mathrm{H}}(k) \right\rfloor$, the dynamics of the battery for the discrete model can be expressed as

$$b(k+1) = \begin{cases} b(k) - e^{\mathrm{T}}(k), & \text{if } a(k) = T, \\ \min\left\{b_{\max}, b(k) + e^{\mathrm{H}}(k)\right\}, & \text{otherwise}, \end{cases} \tag{6}$$

where we used the ceiling and floor in the definitions of $e^{\mathrm{T}}(k)$ and $e^{\mathrm{H}}(k)$ to obtain a lower bound to the performance of the continuous system. Clearly, an upper bound to the performance of the continuous system can be obtained by reversing the use of the floor and ceiling in the definitions of $e^{\mathrm{T}}(k)$ and $e^{\mathrm{H}}(k)$. Similarly, if the channel power gains are modeled by continuous random variables, we divide them into a finite number of intervals with the same probability according to the probability density function (PDF) of the fading gain. Each interval is then represented by a discrete level of channel power gain which has the same probability as that of this interval. In this sense, the problem is modeled as a *finite-state finite-action MDP* with state $s(k) \triangleq (b(k), A(k), g(k), h(k)) \in \mathcal{S}_{\mathrm{d}}^{a}$ (the state space of the discrete model), for which there exists an optimal stationary deterministic policy [21]. Therefore, in the remaining, we investigate this age-optimal stationary deterministic policy and omit the time index. Given a stationary deterministic policy π, the transition probability from state $s = (b, A, g, h)$ to state $s' = (b', A', g', h')$ is given by

$$\begin{aligned} \mathbb{P}\left(s' \mid s, \pi(s)\right) &\triangleq \mathbb{P}\left(b', A', g', h' \mid b, A, g, h, \pi(s)\right) \\ &\overset{(a)}{=} \mathbb{P}\left(b', A' \mid b, A, g, h, \pi(s)\right) \mathbb{P}(g')\mathbb{P}(h') \\ &\overset{(b)}{=} \mathbb{P}\left(b' \mid b, g, h, \pi(s)\right) \mathbb{P}\left(A' \mid A, \pi(s)\right) \mathbb{P}(g')\mathbb{P}(h'), \end{aligned} \tag{7}$$

where $\pi(s)$ denotes the action taken at state s according to π. Step (a) follows from the independence of the channel power gains over time and from other random variables; where $\mathbb{P}(g')$ and $\mathbb{P}(h')$ denote the probability mass functions for the downlink and uplink channel power gains (after discretization if they are modeled originally by continuous random variables), respectively. Step (b) follows since given s and $\pi(s)$, the next battery level b' and value of AoI A' can be obtained deterministically in a separable manner. Specifically, b' only depends on the current battery level and channel power gains, i.e., (b, g, h), and A' is only function of its current value A. Thus, from (3) and (6), b' and A' can be determined, respectively, as

$$\mathbb{P}(b' \mid b, g, h, \pi(s)) = \begin{cases} \mathbb{1}\left(b' = b - e^{\mathrm{T}}\right), \text{ if } \pi(s) = T, \\ \mathbb{1}\left(b' = \min\left\{b_{\max}, b + e^{\mathrm{H}}\right\}\right), \text{ otherwise}, \end{cases} \tag{8}$$

$$\mathbb{P}(A' \mid A, \pi(s)) = \begin{cases} \mathbb{1}\left(A' = 1\right), \text{ if } \pi(s) = T, \\ \mathbb{1}\left(A' = \min\left\{A_{\max}, A + 1\right\}\right), \text{ otherwise}, \end{cases} \tag{9}$$

where $\mathbb{1}(\cdot)$ is the indicator function. The optimal policy $\pi^\star$ satisfying (5) can be evaluated by solving the following Bellman's equation for average cost MDPs [21]

$$\bar{A}^\star + V(s) = \min_{a\in\mathcal{A}(s)} Q(s, a), s \in \mathcal{S}_{\mathrm{d}}^{a}, \tag{10}$$

where $\bar{A}^\star$ is the achievable optimal average AoI by $\pi^\star$ which is independent of the initial state $s(0)$, $V(s)$ is the value function and $Q(s, a)$ is the expected cost resulting from taking action a in state s, which is given by

$$Q(s, a) = A + \sum_{s'\in\mathcal{S}_{\mathrm{d}}^{a}} \mathbb{P}(s' \mid s, a)V(s'), \tag{11}$$

where $\mathbb{P}(s' \mid s, a)$ is evaluated using (7). In addition, the optimal action taken at state s is given by

$$\pi^\star(s) = \arg\min_{a\in\mathcal{A}(s)} Q(s, a). \tag{12}$$

The value function $V(s)$ can be evaluated iteratively using the VIA [21]. Particularly, according to the VIA, the value function at iteration m, $m = 1, 2, \cdots$, is computed as

$$\begin{aligned} V(s)^{(m)} &= \min_{a\in\mathcal{A}(s)} Q(s, a)^{(m-1)} \\ &= \min_{a\in\mathcal{A}(s)} \left\{ A + \sum_{s'\in\mathcal{S}_{\mathrm{d}}^{a}} \mathbb{P}(s' \mid s, a)V(s')^{(m-1)} \right\}, \end{aligned} \tag{13}$$

where $s \in \mathcal{S}_{\mathrm{d}}^{a}$. Hence, the optimal policy at iteration m is given by

$$\pi^{\star(m)}(s) = \arg\min_{a\in\mathcal{A}(s)} Q(s, a)^{(m-1)}. \tag{14}$$

As per the VIA, under any initialization of value function $V(s)^{(0)}$, the sequence $\left\{V(s)^{(m)}\right\}$ converges to $V(s)$ which satisfies the Bellman's equation in (10), i.e.,

$$\lim_{m\to\infty} V(s)^{(m)} = V(s). \tag{15}$$

In the next subsection, we explore the structural properties of the age-optimal policy $\pi^\star$ obtained using the VIA. Note that the obtained analytical results can be derived using the Relative VIA (RVIA) as well [21].

C. Structural Properties of the Age-optimal Policy

Lemma 1. *The value function $V(b, A, g, h)$, satisfying the Bellman's equation in (10) and corresponding to the age-optimal policy $\pi^\star$, is non-increasing with respect to the battery level b, the downlink channel power gain g and the uplink channel power gain h. In contrast, $V(b, A, g, h)$ is non-decreasing with respect to the AoI A.*

Proof: First, to prove that $V(b,A,g,h)$ is non-increasing with respect to b, let us define $s_1=(b_1,A,g,h)$ and $s_2=(b_2,A,g,h)$ where $b_1\le b_2$. Hence, the objective is to show that $V(s_1)\ge V(s_2)$. According to (15), it is then sufficient to show that $V(s_1)^{(m)}\ge V(s_2)^{(m)},\forall m$, which we prove using mathematical induction. Particularly, the relation holds by construction for $m=0$ since it corresponds to the initial values for the value function which can be chosen arbitrary. Now, we assume that $V(s_1)^{(m)}\ge V(s_2)^{(m)}$ holds for some m, and then show that it holds for $V(s_1)^{(m+1)}\ge V(s_2)^{(m+1)}$ as well. Particularly, according to (13) and (14), $V(s_2)^{(m+1)}$ and $V(s_1)^{(m+1)}$ can be expressed, respectively, as

$$\begin{aligned}V(s_2)^{(m+1)}&=A+\sum_{s_2'\in\mathcal{S}_{\rm d}^a}\mathbb{P}(s_2'\mid s_2,\pi^{\star(m)}(s_2))V(s_2')^{(m)}\\&\overset{(a)}{\le}A+\sum_{s_2'\in\mathcal{S}_{\rm d}^a}\mathbb{P}(s_2'\mid s_2,\pi^{\star(m)}(s_1))V(s_2')^{(m)}\\&\overset{(b)}{=}A+C\sum_{g_2'}\sum_{h_2'}V(\bar{b}_2,\bar{A},g_2',h_2')^{(m)},\end{aligned}\tag{16}$$

$$\begin{aligned}V(s_1)^{(m+1)}&=A+\sum_{s_1'\in\mathcal{S}_{\rm d}^a}\mathbb{P}(s_1'\mid s_1,\pi^{\star(m)}(s_1))V(s_1')^{(m)}\\&=A+C\sum_{g_1'}\sum_{h_1'}V(\bar{b}_1,\bar{A},g_1',h_1')^{(m)},\end{aligned}\tag{17}$$

where $C=\mathbb{P}(g_i')\mathbb{P}(h_i'),\forall i\in\{1,2\}$. Step (a) follows since it is not optimal to take action $\pi^{\star(m)}(s_1)$ in state s_2, and step (b) follows from (7)-(9) where, for a given $\pi^{\star(m)}(s_1)$, $\bar{A}$ is evaluated based on (9), and $\bar{b}_2$ and $\bar{b}_1$ are determined using (8). Since $b_1\le b_2$, we can observe from (8) that $\bar{b}_1\le\bar{b}_2$ for $\pi^{\star(m)}(s_1)\in\mathcal{A}$, and hence $V(\bar{b}_1,\bar{A},g_1',h_1')^{(m)}\ge V(\bar{b}_2,\bar{A},g_2',h_2')^{(m)}$. Therefore the expression in (16) is less than or equal to $V(s_1)^{(m+1)}$ which makes $V(s_1)^{(m+1)}\ge V(s_2)^{(m+1)}$ and indicates that the value function is non-increasing with respect to b. Using the same approach, we can show that $V(b,A,g,h)$ is non-decreasing with respect to A. Finally, note that increasing g (h) increases $e^{\rm H}$ (reduces $e^{\rm T}$) which leads to a larger amount of energy in the battery at the next time slot and hence a lower value function. This proves that $V(b,A,g,h)$ is non-increasing with respect to g and h. ■

Based on Lemma 1, the following Lemma characterizes some structural properties of the age-optimal policy $\pi^\star$.

Lemma 2. *For any $s_1=(b_1,A_1,g_1,h_1)$ and $s_2=(b_2,A_2,g_2,h_2)$, the age-optimal policy $\pi^\star$ has the following structural properties:*
(i) When $s_1\preceq s_2$ and $b_1\ge\max\{b_{\max}-e_1^{\rm H},e_1^{\rm T}\}$, if $\pi^\star(s_1)=T$, then $\pi^\star(s_2)=T$.
(ii) When $s_1\succeq s_2$ and $b_2\ge\max\{b_{\max}-e_2^{\rm H},e_2^{\rm T}\}$, if $\pi^\star(s_1)=H$, then $\pi^\star(s_2)=H$.
Note that the symbols $\preceq$ and $\succeq$ represent the element-wise inequalities.

Proof: First, we note that proving that $\pi^\star(s_1)=a$ leads to $\pi^\star(s_2)=a$ is equivalent to showing that

$$Q(s_2,a)-Q(s_2,a')\le Q(s_1,a)-Q(s_1,a'),\forall a'\ne a,\tag{18}$$

where this holds since if a is optimal in state s_1, then we have $Q(s_1,a)-Q(s_1,a')\le 0,\forall a'$, which leads to $Q(s_2,a)\le Q(s_2,a'),\forall a'$, i.e., taking action a is optimal in state s_2. Hence, (i) is proven ((ii) is proven) if (18) holds when $a=T$ and $a'=H$ ($a=H$ and $a'=T$). Therefore, in the remaining, we focus on the proof of (i) while (ii) can be proven similarly. Particularly, from (7)-(9) and (11), we have

$$Q(s_i,T)=A_i+C\sum_{g_i'}\sum_{h_i'}V(b_i-e_i^{\rm T},1,g_i',h_i'),\tag{19}$$

$$Q(s_i,H)=A_i+C\sum_{g_i'}\sum_{h_i'}V(b_{\max},\min\{A_{\max},A_i+1\},g_i',h_i'),\tag{20}$$

where $i\in\{1,2\}$ and the next battery level in (20) is equal to $b_{\max}$ since $b_1+e_1^{\rm H}\ge b_{\max}$ and $b_1\le b_2$. Since $s_1\preceq s_2$ and based on Lemma 1, we have $V(b_1-e_1^{\rm T},1,g_1',h_1')\ge V(b_2-e_2^{\rm T},1,g_2',h_2')$ ($e_1^{\rm T}\ge e_2^{\rm T}$) and $V(b_{\max},\min\{A_{\max},A_2+1\},g_2',h_2')\ge V(b_{\max},\min\{A_{\max},A_1+1\},g_1',h_1')$. Hence, (18) holds for $a=T$ and $a'=H$, which completes the proof of (i). ■

Remark 1. *Note that according to Lemma 2, the age-optimal policy $\pi^\star$ has a threshold-based structure over the set of states $\mathcal{S}_{\rm d}^{{\rm th},a}\triangleq\{s\in\mathcal{S}_{\rm d}^a:b\ge\max\{b_{\max}-e^{\rm H},e^{\rm T}\}\}$. Particularly, $\pi^\star$ is a threshold-based policy with respect to each of the system state variables, i.e., b,A,g, and h. For instance, for a fixed (b,g,h), if $A_{\rm th}$ is the minimum value of AoI for which it is optimal to take an action $a=T$, then for all states $s\in\mathcal{S}_{\rm d}^{\rm th}$ such that $A\ge A_{\rm th}$, the optimal decision is T as well. This is also intuitive, since when the value of AoI becomes large, it is optimal to update the status of information at the destination by sending a new status update. In addition, if there exists a state $s_{\rm th}=(b_{\rm th},A_{\rm th},g_{\rm th},h_{\rm th})$, where $b_{\rm th},g_{\rm th}$, and $h_{\rm th}$ are defined similar to $A_{\rm th}$, then $\pi^\star(s)=T,\forall s\in\mathcal{S}_{\rm d}^{\rm th}$, such that $s\succeq s_{\rm th}$.*

IV. Age-optimal Policy vs. Throughput-optimal Policy

In this section, we aim at comparing the age-optimal policy with the throughput-optimal one. Towards this objective, we first formulate the average throughput maximization problem for the system setup presented in Section II. Afterwards, we investigate the structural properties of the throughput-optimal policy using which we highlight key differences between the structures of the age-optimal and throughput-optimal polices.

A. Average Throughput Maximization Formulation and Proposed Solution

When the objective is to maximize the average throughput, the system state at slot k is defined as $s(k)=\{b(k),g(k),h(k)\}\in\mathcal{S}_{\rm d}^r$, where $\mathcal{S}_{\rm d}^r$ is the state space of the discrete model for the throughput maximization problem, i.e., when the battery and channel power gain are discretized. Note that the AoI is not included now in the state of the system. The action space $\mathcal{A}$ is defined similarly as in Section II, where the source node can either harvest energy or transmit a packet of size S at each time slot. The evolution of the battery is then given by (6). Hence, the average throughput maximization problem is modeled as a finite-state finite-action MDP for which there exists an optimal stationary deterministic policy

[21]. Particularly, under a policy μ, the long-term average throughput is defined as

$$\bar{R}^{\mu} \triangleq \liminf_{K\to\infty} \frac{1}{K+1} \sum_{k=0}^{K} \mathbb{E}\left[\mathbb{1}\left(a(k)=T\right) S \mid s(0)\right], \quad (21)$$

where the system receives some reward equal to S in an arbitrary time slot only if this slot is allocated for data transmission to the destination node. Our goal is then to characterize the throughput-optimal policy μ^{*} which maximizes the long-term average throughput, i.e., $\mu^{\star} = \arg\max_{\mu} \bar{R}^{\mu}$. Under a stationary deterministic policy μ, the probability of moving from state s to state s' can be expressed as

$$\mathbb{P}\left(s' \mid s, \mu(s)\right) = \mathbb{P}\left(b' \mid b, g, h, \mu(s)\right) \mathbb{P}(g')\mathbb{P}(h'), \quad (22)$$

where $\mathbb{P}\left(b' \mid b, g, h, \mu(s)\right)$ can be expressed as in (8). The optimal policy $\mu^{\star}$ can then be obtained by solving the following Bellman's equation using the VIA (similar to (13) and (14))

$$\bar{R}^{\star} + V(s) = \max_{a\in\mathcal{A}(s)} Q(s,a), s \in \mathcal{S}_{\rm d}^{r}, \quad (23)$$

where $\bar{R}^{\star}$ is the optimal average throughput achievable by $\mu^{\star}$ and $Q(s,a)$ can be expressed as

$$Q(s,a) = \mathbb{1}\left(a=T\right) S + \sum_{s'\in\mathcal{S}_{\rm d}^{r}} \mathbb{P}(s' \mid s,a)V(s'), \quad (24)$$

where $\mathbb{P}(s' \mid s,a)$ is computed by (22) and $\mu^{\star}(s)$ is given by

$$\mu^{\star}(s) = \arg\max_{a\in\mathcal{A}(s)} Q(s,a). \quad (25)$$

B. Structural Properties of the Throughput-optimal Policy

Lemma 3. *The value function $V(b,g,h)$, corresponding to the throughput-optimal policy $\mu^{\star}$, is non-decreasing with respect to the battery level b, the downlink channel power gain g, and the uplink channel power gain h.*

Proof: By using (22), the result can be obtained using the same approach used in the proof of Lemma 1, i.e., by applying mathematical induction to the iterations of the VIA. ■

Using Lemma 3, some structural properties of the throughput-optimal policy are presented by the following Lemma.

Lemma 4. *For any $s_1 = (b_1, g_1, h_1)$ and $s_2 = (b_2, g_2, h_2)$, the throughput-optimal policy $\mu^{\star}$ has the following structural properties:*
(i) When $s_1 \preceq s_2$ and $b_1 \geq \max\left\{b_{\max} - e_1^{\rm H}, e_1^{\rm T}\right\}$, if $\mu^{\star}(s_1) = T$, then $\mu^{\star}(s_2) = T$.
(ii) When $s_1 \succeq s_2$ and $b_2 \geq \max\left\{b_{\max} - e_2^{\rm H}, e_2^{\rm T}\right\}$, if $\mu^{\star}(s_1) = H$, then $\mu^{\star}(s_2) = H$.

Proof: This result can be proven using the same approach used in the proof of Lemma 2. Note that since this is a maximization problem, proving that $\mu^{\star}(s_1) = a$ leads to $\mu^{\star}(s_2) = a$ is now equivalent to showing that

$$Q(s_2,a) - Q(s_2,a') \geq Q(s_1,a) - Q(s_1,a'), \forall a' \neq a. \quad (26)$$

■

Remark 2. *Similar to Remark 1 and based on Lemma 4, we observe that the throughput-optimal policy has a threshold-based structure over the set of states $\mathcal{S}_{\rm d}^{{\rm th},r} = \left\{s \in \mathcal{S}_{\rm d}^{r} : b \geq \max\{b_{\max} - e^{\rm H}, e^{\rm T}\}\right\}$.*

Remark 3. *Our results in Lemmas 2 and 4 clearly demonstrate that the structures of the age-optimal and throughput-optimal policies are different, which will also be verified in the numerical results section. Specifically, let us consider a state $\bar{s} = (\bar{b}, \bar{g}, \bar{h}) \in \mathcal{S}_{\rm d}^{{\rm th},r}$ such that $\mu^{\star}(\bar{s}) = T$. Note that since $s \in \mathcal{S}_{\rm d}^{{\rm th},r}$, the set of states $\bar{\mathcal{S}}_{\rm d}^{{\rm th},a} = \{(b,A,g,h) : (b,g,h) = \bar{s}, 1 \leq A \leq A_{\max}\}$ belongs to $\mathcal{S}_{\rm d}^{{\rm th},a}$. Similar to the definition of $A_{\rm th}$ in Remark 1, let us define $\bar{A}^{\rm th} = \min\left(\left\{A : \pi^{\star}(\bar{b}, A, \bar{g}, \bar{h} = T)\right\}\right)$. Now, for a given state $s \in \bar{\mathcal{S}}_{\rm d}^{{\rm th},a}$ such that $A < \bar{A}^{\rm th}$, according to Lemma 4, we note that $\pi^{\star}(s) = H$. This indicates that $\mu^{\star}(\bar{s})$ and $\pi^{\star}(s)$ are different even though the states s and $\bar{s}$ have the same combination $(\bar{b}, \bar{g}, \bar{h})$ which demonstrates the difference between the structures of the age-optimal and the throughput-optimal polices.*

V. Numerical Results

The downlink and uplink channel power gains between the source and destination nodes are modeled as $g = h = \theta\psi^2 d^{-\beta}$; where θ is the signal power gain at a reference distance of 1 meter, $\psi^2 \sim \exp(1)$ denotes the small-scale fading gain, and $d^{-\beta}$ represents standard power law path-loss with exponent β. In addition, the channel power gains are discretized into 10 levels. In the following, we use $h = i$ to refer to the value of the channel power gain at its i-th level. Unless otherwise specified, we use the following values for different system parameters: $W = 1$ MHz, $P = 37$ dBm, $d = 25$ meters, $\eta = 0.5$, $\sigma^2 = -95$ dBm, $\theta = 4 \times 10^{-2}$, $\beta = 2$, $S = 12$ Mbits, $B_{\max} = 0.3$ mjoules, $A_{\max} = 10$ and $b_{\max} = 9$.

Verification of Analytical Results. In Figs. 2a and 2b, we demonstrate the structures of the age-optimal and throughput-optimal polices. Particularly, each point in both figures represents a potential state of the system where a red circle point (a blue square point) indicates that the optimal action at this state is T (H). The points located inside the solid polygon refer to the states for which it is possible to take T action, i.e., for each of those states $b \geq e^{\rm T}$. Furthermore, the points located inside the dotted polygon represent the sets $\mathcal{S}_{\rm d}^{{\rm th},a}$ and $\mathcal{S}_{\rm d}^{{\rm th},r}$, where the dotted polygon is the same as the solid one in Fig. 2a. We can check that the analytical structural properties of the optimal polices, derived in Lemmas 2 and 4, are satisfied. For instance, in Fig. 2a, since the optimal action at the point $(2,2)$ is T, we observe that the optimal action at all the points (x,y), where $x \geq 2$ and $y \geq 2$, is T as well (Lemma 2, (i)). In addition, in Fig. 2b, the optimal action at the point $(3,10)$ is H, and hence, we observe that it is optimal to take action H at all the states (x,y) located inside the dotted polygon such that $x \leq 3$ and $y \leq 10$ (Lemma 4, (ii)).

Comparison between the Structures of the Age-optimal and Throughput-optimal Policies. The structure of the age-optimal policy is plotted for $g = h = 10$ in Fig. 2a. Hence, the difference between structures of the age-optimal and throughput-optimal polices can be captured by comparing the optimal actions at the points $(x,10)$ in Fig. 2b with the actions at the points (x,A) in Fig. 2a, where $x \geq 1$ and $1 \leq A \leq 10$ is a fixed value of AoI. Specifically, according to the value of A, we have two different regimes: (i) when A is small (for

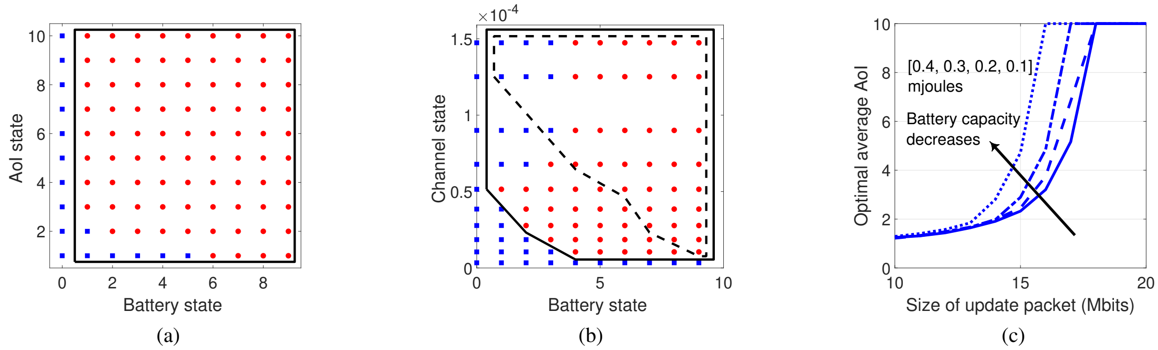

Fig. 2. (a) Structure of age-optimal policy, (b) Structure of throughput-optimal policy, and (c) Impact of battery capacity and size of update packet on $\bar{A}^{\star}$.

instance, $A = 1$), it is optimal to keep taking action H till a larger value of battery state in the age-optimal policy than the case for the throughput-optimal policy, and (ii) when A is large ($A \geq 3$), different from the throughput-optimal policy, it is always optimal to take action T regardless of the amount of available energy in the battery according to the age-optimal policy. This is intuitive since if AoI is small, it is wise to save the current energy in battery for future update packet transmissions when the AoI becomes large.

Impact of System Design Parameters on Optimal Average AoI. Fig. 2c shows the impact of the capacity of battery and size of update packets on the optimal achievable average AoI $\bar{A}^{\star}$, satisfying the Bellman's equation in (10). It is observed that the achievable average AoI monotonically decreases as the size of update packets decreases and/or the capacity of battery increases. This is due to the fact that decreasing the size of update packets reduces the amount of energy needed to transmit an update packet, and increasing the battery capacity allows to store more harvested energy inside the battery. This, in turn, increases the likelihood that the battery has enough energy for update packet transmissions when the AoI is large, and hence the achievable average AoI is reduced.

VI. Conclusion

This paper proposed an implementable age-optimal sampling strategy for designing freshness-aware RF-powered IoT networks. Particularly, the long-term AoI minimization problem was formulated for a real-time IoT-enabled monitoring system, in which a source node is powered by wireless energy transfer by a destination node. To obtain the age-optimal policy, the problem was then modeled as an average cost MDP for which the monotonicity property of its value function, with respect to the system state, was analytically established. Afterwards, to inspect the difference between the age-optimal and throughput-optimal policies for our system setup, we extended our analysis to the average throughput maximization problem. Multiple system design insights were drawn from our results. For instance, they demonstrated that the age-optimal and throughput-optimal policies are threshold-based policies with significantly different structures. They also revealed that the optimal average AoI is a monotonically increasing (decreasing) function with respect to size of update packets (battery capacity).

References

[1] M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "On the role of age-of-information in internet of things," *IEEE Commun. Magazine*, to appear. Available online: arxiv.org/abs/1812.08286.

[2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc., IEEE INFOCOM*, 2012.

[3] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc., IEEE ISIT*, 2012.

[4] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Trans. on Info. Theory*, vol. 62, no. 4, pp. 1897–1910, 2016.

[5] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc., IEEE ISIT*, 2015.

[6] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *Proc., IEEE ISIT*, 2017.

[7] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. on Info. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[8] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimizing data freshness, throughput, and delay in multi-server information-update systems," in *Proc., IEEE ISIT*, 2016.

[9] M. A. Abd-Elmagid and H. S. Dhillon, "Average peak age-of-information minimization in UAV-assisted IoT networks," *IEEE Trans. on Veh. Technology*, vol. 68, no. 2, pp. 2003–2008, Feb 2019.

[10] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc., IEEE INFOCOM*, 2018.

[11] R. Talak, S. Karaman, and E. Modiano, "Optimizing age of information in wireless networks with perfect channel state information," in *Proc., WiOpt*, 2018.

[12] M. A. Abd-Elmagid, A. Ferdowsi, H. S. Dhillon, and W. Saad, "Deep reinforcement learning for minimizing age-of-information in UAV-assisted networks," in *Proc., IEEE Globecom*, 2019.

[13] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc., IEEE ISIT*, 2015.

[14] A. Arafa and S. Ulukus, "Timely updates in energy harvesting two-hop networks: Offline and online policies," *IEEE Trans. on Wireless Commun.*, vol. 18, no. 8, pp. 4017–4030, Aug. 2019.

[15] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc., Information Theory and its Applications (ITA)*, 2015.

[16] S. Feng and J. Yang, "Age of information minimization for an energy harvesting source with updating erasures: With and without feedback," 2018, available online: arxiv.org/abs/1808.05141.

[17] A. Baknina, S. Ulukus, O. Oze, J. Yang, and A. Yener, "Sending information through status updates," in *Proc., IEEE ISIT*, 2018.

[18] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. on Green Commun. and Networking*, vol. 2, no. 1, pp. 193–204, 2018.

[19] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and K. B. Letaief, "Optimal online transmission policy in wireless powered networks with urgency-aware age of information," 2019, available online: arxiv.org/abs/1901.09232.

[20] I. Krikidis, "Average age of information in wireless powered sensor networks," *IEEE Wireless Commun. Letters*, 2019.

[21] D. P. Bertsekas, "Dynamic programming and optimal control 3rd edition, volume ii," *Belmont, MA: Athena Scientific*, 2011.