

Article

Channel Quality-Based Optimal Status Update for Information Freshness in Internet of Things

Fuzhou Peng , Xiang Chen * and Xijun Wang 

School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China; pengfzh@mail2.sysu.edu.cn (F.P.); wangxijun@mail.sysu.edu.cn (X.W.)

* Correspondence: chenxiang@mail.sysu.edu.cn

Abstract: This paper investigates the status updating policy for information freshness in Internet of things (IoT) systems, where the channel quality is fed back to the sensor at the beginning of each time slot. Based on the channel quality, we aim to strike a balance between the information freshness and the update cost by minimizing the weighted sum of the age of information (AoI) and the energy consumption. The optimal status updating problem is formulated as a Markov decision process (MDP), and the structure of the optimal updating policy is investigated. We prove that, given the channel quality, the optimal policy is of a threshold type with respect to the AoI. In particular, the sensor remains idle when the AoI is smaller than the threshold, while the sensor transmits the update packet when the AoI is greater than the threshold. Moreover, the threshold is proven to be a non-increasing function of channel state. A numerical-based algorithm for efficiently computing the optimal thresholds is proposed for a special case where the channel is quantized into two states. Simulation results show that our proposed policy performs better than two baseline policies.

Keywords: age of information; status update; channel quality



Citation: Peng, F.; Chen, X.; Wang, X. Channel Quality-Based Optimal Status Update for Information Freshness in Internet of Things. *Entropy* **2021**, *23*, 912. <https://doi.org/10.3390/e23070912>

Academic Editors: Anthony Ephremides and Yin Sun

Received: 18 June 2021
Accepted: 16 July 2021
Published: 18 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recently, the Internet of things (IoT) has been widely used in the field of industrial manufacturing, environment monitoring, and home automation. In these applications, the sensors generate and transmit new status updates to the destination, where the freshness of the status updates is crucial for the destination to track the state of the environment and to make decisions. Thus, a new information freshness metric, namely age of information (AoI), was proposed in [1] to measure the freshness of updates from the receiver's perspective. There are two widely used metrics, i.e., the average peak AoI [2] and the average AoI [3]. In general, the smaller the AoI is, the fresher the received updates are.

AoI was originally investigated in [1] for updating the status in vehicular networks. Considering the impact of the queueing system, the authors in [4] investigated the system performance under the M/M/1 and M/M/1/2 queueing systems with a first-come-first-served (FCFS) policy. Furthermore, the work of [5] studied how to keep the updates fresh by analyzing some general update policies, such as the zero-wait policy. The authors of [6] considered the optimal schedule problem for a more general cost that is the weighted sum of the transmission cost and the tracking inaccuracy for the information source. However, these works assumed that the communication channel is not error-prone. In practice, status updates are delivered through an erroneous wireless channel, which suffers from fading, interference, and noises. Therefore, the received updates may not be decoded correctly, which induces information aging and energy consumption.

There are several works that considered the erroneous channel [7,8]. The authors in [9] considered multiple communication channels and investigated the optimal coding and decoding schemes. The channel with an independent and identical packet error rate over time was considered in [10,11]. The work of [12] considered the impact of fading channels in packet transmission. A Markov channel was investigated in [13], where

threshold policy was proven to be optimal, and a simulation-based approach was proposed to compute the corresponding threshold. However, how the information of channel quality should be exploited to improve system performance in information freshness remains to be investigated.

Channel quality indicator (CQI) feedback is commonly used in wireless communication systems [14]. In block fading channels, the channel quality, generally reported by the terminal, is highly relevant to the packet error rate (PER) [15] or, namely, the block error rate (BLER). It is probable that a received packet fails to be decoded when the channel suffers from a poor condition. However, a transmitter with the channel quality information is able to keep idle when there is deep fading, thereby saving energy. The channel quantization was also considered in [12,13], where the channel was quantized into multiple states. However, the decision making was not dependent on the channel state in [12], while [13] did not consider the freshness of information. These motivate us to introduce the information of channel quality into the design of the updating policy.

In this paper, a status update system with channel quality feedback is considered. In particular, the channel condition is quantized into multiple states, and the destination feeds the channel quality back to the sensor before the sensor updates the status. Our problem is to investigate the channel quality-based optimal status update policy, which minimizes the weighted sum of the AoI and the energy consumption. Our key contributions are summarized as follows:

- An average cost Markov decision process (MDP) is formulated to model this problem. Due to the infinite countable states and unbounded cost of the MDP, which makes analysis difficult, the discounted version of the original problem is first investigated, and the existence of the stationary and deterministic policy to the original problem is then proven. Furthermore, it is proven that the optimal policy is a threshold structure policy with respect to the AoI for each channel state by showing the monotonic property of the value function. We also prove that the threshold is a non-increasing function of channel state.
- By utilizing the threshold structure, a structure-aware policy iteration algorithm is proposed to efficiently obtain the optimal updating policy. Nevertheless, a numerical-based algorithm which directly computes the thresholds by non-linear fractional programming is also derived. Simulation results reveal the effects of system parameters and show that our proposed policy performs better than the zero-wait policy and periodic policy.

The rest of this paper is organized as follows. In Section 2, the system model is presented and the optimal updating problem is formulated. In Section 3, the optimal updating policy is proven to be of a threshold structure, and a threshold-based policy iteration algorithm is proposed to find the optimal policy. Section 4 presents the simulation results. Finally, we summarize our conclusions in Section 5.

2. System Model and Problem Formulation

2.1. System Description

In this paper, we consider a status update system that consists of a sensor and a destination, as shown in Figure 1. Time is divided into slots. Without loss of generality, we assume that each time slot has an equal length, which is normalized to unity. At the beginning of each slot, the destination feeds the CQI back to the sensor. It is worth noting that the PER is different for different CQIs. Based on the CQI, the sensor decides in each time slot whether it should generate and transmit a new update to the destination via a wireless channel or keep idle for saving energy. These updates are crucial for the destination to estimate the states of the surrounding environment of the sensor and to make in-time decisions. Let a_t , which takes value from the action set $\mathcal{A} = \{0, 1\}$, denote the action that the sensor performs in slot t , where $a_t = 1$ means that the sensor generates and transmits a new update to the destination, and $a_t = 0$ represents that the sensor is idle. If the sensor transmits an update packet in slot t , an acknowledgment will be fed

back at the end of this time slot. In particular, an ACK is fed back when the destination successfully receives the update packet, and a NACK otherwise.

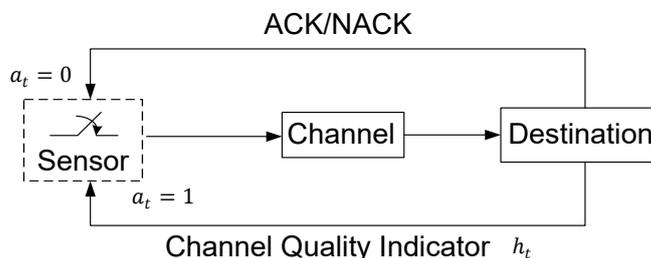


Figure 1. System model.

2.2. Channel Model

Suppose that the wireless channel is a block fading channel where the channel gain remains constant in each slot and varies independently over different slots. Let z_t denote the channel gain in slot t which takes value from $[0, +\infty)$. We quantize the channel gain into $N + 1$ levels which are denoted as $(z_0, z_1, \dots, z_i, \dots, z_N)$. The quantization levels are arranged in an increasing order where $z_0 = 0$ and $z_N = \infty$. Hence, the channel is said to be in state i if the channel gain z_t belongs to the interval $[z_i, z_{i+1})$. We denote by h_t the state of the channel in slot t , where $h_t \in \mathbb{H} \triangleq \{0, 1, 2, \dots, N - 1\}$. With the aid of CQI fed back from the destination, the sensor has knowledge of the channel state at the beginning of each time slot.

Let $p_z(z)$ denote the distribution of the channel gain. Then, the probability of the channel being in state i is

$$p_i = \int_{z_i}^{z_{i+1}} p_z(z) dz. \tag{1}$$

We assume that the signal-to-noise ratio (SNR) per information bit during the transmission remains constant. Then, the PER depends only on the channel gain. In particular, the PER for channel state i is given by

$$g_i = \int_{z_i}^{z_{i+1}} P_{\text{PER}}(z) p_z(z|i) dz, \tag{2}$$

where $P_{\text{PER}}(z)$ is the PER of a packet with respect to the channel gain. The success probability q_i of a packet transmitted over channel state i is $q_i = 1 - g_i$. According to [15], the success probability is a non-decreasing function of the channel state.

2.3. Age of Information

This paper uses the AoI as the freshness metric, which is defined as the time elapsed since the generation time of the latest update packet that is successfully received by the destination [1]. Let G_i be the generation time of the i th successfully received update packet. Then, the AoI in time slot t , Δ_t , is defined as

$$\Delta_t = t - \max\{G_i : G_i \leq t\}. \tag{3}$$

In particular, if an update packet is successfully received, the AoI decreases to one. Otherwise, the AoI increases by one. Altogether, the evolution of the AoI is expressed by

$$\Delta_{t+1} = \begin{cases} 1, & \text{if the transmission is successful,} \\ \Delta_t + 1, & \text{otherwise.} \end{cases} \tag{4}$$

An example of the AoI evolution is shown in Figure 2, where the gray rectangle represents a successful reception of an update packet, and the mesh rectangle represents a transmission failure.

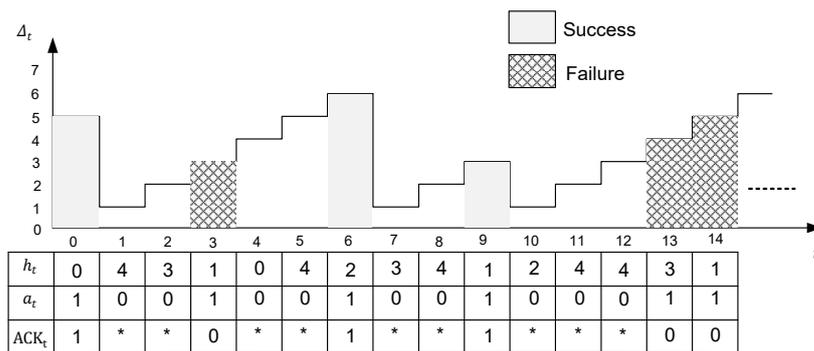


Figure 2. An example of the AoI evolution with the channel state h_t , the action a_t , and the acknowledgment ACK_t . The asterisk stands for no acknowledgment from destination when the sensor keeps idle.

2.4. Problem Formulation

The objective of this paper is to find an optimal updating policy that minimizes the long-term average of the weighted sum of AoI and energy consumption. A policy π can be represented by the sequence of actions, i.e., $\pi = (a_0, a_1, \dots, a_t, \dots)$. Let Π be a set of stationary and deterministic policies. Then, the optimal updating problem is given by

$$\min_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \mathbb{E}[\Delta_t + \omega a_t C_e], \tag{5}$$

where C_e is the energy consumption, and ω is the weighting factor.

3. Optimal Updating Policy

This section aims to investigate the optimal updating policy for the problem formulated in above section. In this section, our investigating problem is first formulated into an infinite horizon average cost MDP, and the existence of a stationary and deterministic policy that minimizes the average cost is proven. Then, the non-decreasing property of the value function is derived. Based on this property, we prove that the optimal update policy is of a threshold structure with respect to AoI, and the optimal threshold is a non-increasing function of the channel state. Aiming to reduce the computational complexity, a structure-aware policy iteration algorithm is proposed to find the optimal policy. Moreover, non-linear fractional programming is employed to directly compute the optimal thresholds in a special case where the channel is quantized into two states.

3.1. MDP Formulation

The Markov decision process (MDP) is typically applied to address the optimal decision problem when the investigation problem can be characterized by the evolution of the system state and the cost is per-stage. The optimization problem in (5) can be formulated as an infinite horizon average cost MDP, which is elaborated in the following.

- States: The state of the MDP in slot t is defined as $\mathbf{x}_t = (\Delta_t, h_t)$, which takes values in $\mathbb{Z}^+ \times \mathbb{H}$. Hence, the state space \mathcal{S} is countable and infinite.
- Actions: The set of actions a_t chosen in slot t is $\mathcal{A} = \{0, 1\}$.
- Transition Probability: Let $\Pr(\mathbf{x}_{t+1} | \mathbf{x}_t, a_t)$ be the transition probability that the state \mathbf{x}_t in slot t transits to \mathbf{x}_{t+1} in slot $t + 1$ after taking action a_t . According to the evolution of AoI in (4), the transition probability is given by

$$\begin{cases} \Pr(\mathbf{x}_{t+1} = (\Delta + 1, j) | \mathbf{x}_t = (\Delta, i), a_t = 0) = p_j, \\ \Pr(\mathbf{x}_{t+1} = (1, j) | \mathbf{x}_t = (\Delta, i), a_t = 1) = q_i p_j, \\ \Pr(\mathbf{x}_{t+1} = (\Delta + 1, j) | \mathbf{x}_t = (\Delta, i), a_t = 1) = g_i p_j. \end{cases} \tag{6}$$

- Cost: The instantaneous cost $C(\mathbf{x}_t, a_t)$ at state \mathbf{x}_t given action a_t in slot t is

$$C(\mathbf{x}_t, a_t) = \Delta_t + \omega a_t C_e. \tag{7}$$

For an MDP with infinite states and unbounded cost, it is not guaranteed to have a stationary and deterministic policy that attains the minimum average cost in general. Fortunately, we can prove the existence of stationary and deterministic policy in next sub-section.

3.2. The Existence of Stationary and Deterministic Policy

For rigorous mathematical analysis, this section is purposed to prove the existence of a stationary and deterministic optimal policy. According to [16], we first analyze the associated discounted cost problem of the original MDP. The expectation of discount cost with respect to discounted factor γ and initial state $\hat{\mathbf{x}}$ under a policy π is given by

$$V_{\pi, \gamma}(\hat{\mathbf{x}}) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t C(\mathbf{x}_t, a_t) | \mathbf{x}_0 = \hat{\mathbf{x}} \right], \tag{8}$$

where a_t is the decision made in state $\hat{\mathbf{x}}$ under policy π , and $\gamma \in (0, 1)$ is the discounted factor. We first verify that $V_{\pi, \gamma}(\hat{\mathbf{x}})$ is finite for any policy and all $\hat{\mathbf{x}} \in \mathcal{S}$.

Lemma 1. Given $\gamma \in (0, 1)$, for any policy π and all $\hat{\mathbf{x}} = (\hat{\mathbf{x}}, \hat{h}) \in \mathcal{S}$, we have

$$V_{\pi, \gamma}(\hat{\mathbf{x}}) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t C(\mathbf{x}_t, a_t) | \mathbf{x}_0 = \hat{\mathbf{x}} \right] < \infty. \tag{9}$$

Proof. By definition, the instantaneous cost in state $\mathbf{x}_t = (\Delta_t, h_t)$ given action a_t is

$$C(\mathbf{x}_t, a_t) = \begin{cases} \Delta_t, & \text{if } a_t = 0, \\ \Delta_t + \omega C_e, & \text{if } a_t = 1. \end{cases} \tag{10}$$

Therefore, $C(\mathbf{x}_t, a_t) \leq \Delta_t + \omega C_e$ holds. Combined with the fact that the AoI increases, at most, linearly at each slot for any policy, we have

$$\begin{aligned} & \sum_{t=0}^{\infty} \gamma^t C(\mathbf{x}_t, a_t | \mathbf{x}_0 = (\hat{\Delta}, \hat{h})) \\ & \leq \sum_{t=0}^{\infty} \gamma^t (\hat{\Delta} + t + \omega C_e) \\ & = \frac{1}{1 - \gamma} \left(\hat{\Delta} + \frac{\gamma}{1 - \gamma} + \omega \right) < \infty, \end{aligned} \tag{11}$$

which completes the proof. \square

Let $V_{\gamma}(\hat{\mathbf{x}}) = \min_{\pi} V_{\pi, \gamma}(\hat{\mathbf{x}})$ denote the minimum expected discounted cost. By Lemma 1, $V_{\gamma}(\hat{\mathbf{x}}) = \min_{\pi} V_{\pi, \gamma}(\hat{\mathbf{x}}) < \infty$ holds for every $\hat{\mathbf{x}}$ and $\gamma \in (0, 1)$.

According to [16] (Proposition 1), we have

$$V_{\gamma}(\hat{\mathbf{x}}) = \min_{a \in \mathcal{A}} \left\{ C(\hat{\mathbf{x}}, a) + \gamma \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \hat{\mathbf{x}}, a) V_{\gamma}(\mathbf{x}') \right\}, \tag{12}$$

which implies that $V_{\gamma}(\hat{\mathbf{x}})$ satisfies the Bellman equation. $V_{\gamma}(\hat{\mathbf{x}})$ can be solved via a value iteration algorithm. In particular, we define $V_{\gamma, 0}(\hat{\mathbf{x}}) = 0$, and for all $n \geq 1$, we have

$$V_{\gamma, n}(\hat{\mathbf{x}}) = \min_{a \in \mathcal{A}} Q_{\gamma, n}(\hat{\mathbf{x}}, a), \tag{13}$$

where

$$Q_{\gamma,n}(\hat{\mathbf{x}}, a) = \min_{a \in \mathcal{A}} \left\{ C(\hat{\mathbf{x}}, a) + \gamma \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \hat{\mathbf{x}}, a) V_{\gamma,n-1}(\mathbf{x}') \right\} \tag{14}$$

is related to the right-hand-side (RHS) of the discounted cost optimality equation. Then, $\lim_{n \rightarrow \infty} V_{\gamma,n}(\hat{\mathbf{x}}) = V_{\gamma}(\hat{\mathbf{x}})$ for every $\hat{\mathbf{x}}$ and γ .

Now, we can use the value iteration algorithm to establish the monotonic properties of $V_{\gamma}(\hat{\mathbf{x}})$

Lemma 2. For all Δ and i , we have

$$V_{\gamma}(\Delta, N - 1) \leq V_{\gamma}(\Delta, i), \tag{15}$$

and for all $\Delta_1 \leq \Delta_2$ and i , we have

$$V_{\gamma}(\Delta_1, i) \leq V_{\gamma}(\Delta_2, i). \tag{16}$$

Proof. See Appendix A. \square

Based on Lemmas 1 and 2, we are ready to show that the MDP has a stationary and deterministic optimal policy in the following theorem.

Theorem 1. For the MDP in (5), there exists a stationary and deterministic optimal policy π^* that minimizes the long-term average cost. Moreover, there exists a finite constant $\lambda = \lim_{\gamma \rightarrow 1} (1 - \gamma)V_{\gamma}(\mathbf{x})$ for all states \mathbf{x} , where λ is independent of the initial state, and a value function $V(\mathbf{x})$, such that

$$\lambda + V(\mathbf{x}) = \min_{a \in \mathcal{A}} \left\{ C(\mathbf{x}, a) + \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \mathbf{x}, a) V(\mathbf{x}') \right\} \tag{17}$$

holds for all \mathbf{x} .

Proof. See Appendix B. \square

3.3. Structural Analysis

According to Theorem 1, the optimal policy for the average cost problem satisfies the following equation

$$\pi^*(\mathbf{x}) = \arg \min_{a \in \mathcal{A}} Q(\mathbf{x}, a), \tag{18}$$

where

$$Q(\mathbf{x}, a) = C(\mathbf{x}, a) + \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \mathbf{x}, a) V(\mathbf{x}'). \tag{19}$$

Similar to Lemma 2, the monotonic property of the value function $V(\mathbf{x})$ is given in the following lemma.

Lemma 3. Given the channel state i , for any $\Delta_2 \geq \Delta_1$, we have

$$V(\Delta_2, i) \geq V(\Delta_1, i). \tag{20}$$

Proof. This proof follows the same procedure of Lemma 2, with one exception being that the value iteration algorithm is based on Equation (17). \square

Moreover, based on Lemma 3, the property of the increment of the value function is established in following lemma.

Lemma 4. *Given the channel state i , for any $\Delta_2 \geq \Delta_1$, we have*

$$V(\Delta_2, i) - V(\Delta_1, i) \geq \Delta_2 - \Delta_1. \tag{21}$$

Proof. We first examine the relation between the state-action value functions, i.e., $Q(\Delta_2, i, a)$ and $Q(\Delta_1, i, a)$. Specifically, based on Lemma 3, we have

$$\begin{aligned} Q(\Delta_2, i, 0) - (\Delta_2 - \Delta_1) &= \Delta_1 + \sum_{j=0}^{N-1} p_j V(\Delta_2 + 1, j) \\ &\geq \Delta_1 + \sum_{j=0}^{N-1} p_j V(\Delta_1 + 1, j) = Q(\Delta_1, i, 0), \end{aligned} \tag{22}$$

and

$$\begin{aligned} &Q(\Delta_2, i, 1) - (\Delta_2 - \Delta_1) \\ &= \Delta_1 + \omega C_e + q_i \sum_{j=0}^{N-1} p_j V(1, j) + g_i \sum_{j=0}^{N-1} p_j V(\Delta_2 + 1, j) \\ &\geq \Delta_1 + \omega C_e + q_i \sum_{j=0}^{N-1} p_j V(1, j) + g_i \sum_{j=0}^{N-1} p_j V(\Delta_1 + 1, j) \\ &= Q(\Delta_1, i, 1). \end{aligned} \tag{23}$$

Since $V(\mathbf{x}) = \min_{a \in \mathcal{A}} Q(\mathbf{x}, a)$, we complete the proof. \square

Our main result is presented in the following theorem.

Theorem 2. *For any given channel state i , there exists a threshold β_i , such that when $\Delta \geq \beta_i$, the optimal action is to generate and transmit a new update, i.e., $\pi^*(\Delta, i) = 1$, and when $\Delta < \beta_i$, the optimal action is to remain idle, i.e., $\pi^*(\Delta, i) = 0$. Moreover, the optimal threshold β_i is a non-increasing function of channel state i , i.e., $\beta_i \geq \beta_j$ holds for all $i, j \in \mathbb{H}$ and $i \leq j$.*

Proof. See Appendix C. \square

According to Theorem 2, the sensor will not update the status until the AoI exceeds the threshold. Moreover, if the channel condition is not good, i.e., channel state i is small, the sensor will wait for a longer time before it samples and transmits the status update packet so as to reduce the energy consumption because of a higher probability of transmission failure.

Based on the threshold structure, we can reduce the computational complexity of the policy iteration algorithm to find the optimal policy. The details of the algorithm are presented in Algorithm 1.

3.4. Computing the Thresholds for a Special Case

In the above section, we have proven that the optimal policy has a threshold structure. Given the thresholds $(\beta_0, \beta_1, \dots, \beta_{N-1})$, a Markov chain can be induced by the threshold policy. A special Markov chain is depicted in Figure 3, where the channel has two states. By leveraging the Markov chain, we first derive the average cost of the special case, which is summarized in the following theorem.

Algorithm 1 Policy iteration algorithm (PIA) based on the threshold structure.

- 1: Initialization: Set $k = 0$, iteration threshold ϵ , and initialize the value function $V_0(\mathbf{x}) = 0$ and policy $\pi(\mathbf{x}) = 0$ for all state $\mathbf{x} \in \mathcal{S}$
- 2: **repeat**
- 3: $k \leftarrow k + 1$.
- 4: Based on last iterative value function $V_{k-1}(\mathbf{x})$, compute the current value function $V_k(\mathbf{x})$ by calculating the following equations.
- 5:
$$V_k(\mathbf{x}) = \min_{a \in \mathcal{A}} \left\{ C(\mathbf{x}, a) + \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \mathbf{x}, a) V_{k-1}(\mathbf{x}') \right\}$$
- 6: **until** $|V_k(\mathbf{x}) - V_{k-1}(\mathbf{x})| \leq \epsilon$ for all $\mathbf{x} \in \mathcal{S}$
- 7: **for** $\mathbf{x} = (\Delta, i) \in \mathcal{S}$ **do**
- 8: **if** $\mathbf{x}' = (\Delta - 1, i) \in \mathcal{S}$ and $\pi(\mathbf{x}') = 1$ **then**
- 9: $\pi(\mathbf{x}) \leftarrow 1$.
- 10: **else**
- 11:
$$\pi(\mathbf{x}) \leftarrow \arg \min_{a \in \mathcal{A}} \left\{ C(\mathbf{x}, a) + \sum_{\mathbf{x}' \in \mathcal{S}} \Pr(\mathbf{x}' | \mathbf{x}, a) V(\mathbf{x}') \right\}$$
- 12: **end if**
- 13: **end for**
- 14: $\pi^* \leftarrow \pi$
- 15: **return** the optimal policy π^*

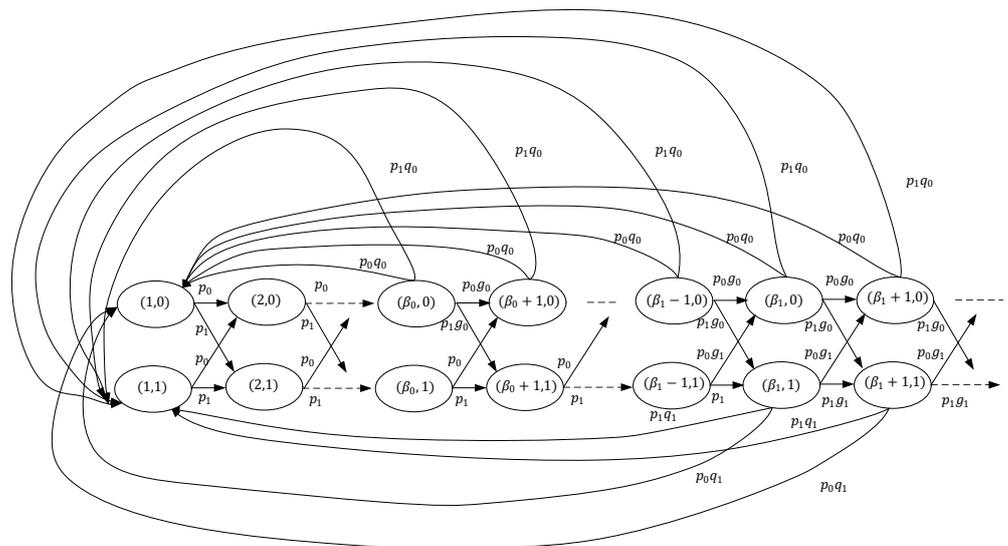


Figure 3. An illustration of established Markov chain with two channel states.

Theorem 3. Let $\varphi(\mathbf{x})$ be the steady state probability of state \mathbf{x} of the corresponding Markov chain with two states and β_0, β_1 be the threshold with respect to the channel state, respectively. The steady state probability is given by

$$\varphi(i, j) = \begin{cases} p_j \varphi_1, & \text{if } 1 \leq i \leq \beta_1, \\ p_j s_0^{i-\beta_1} \varphi_1, & \text{if } \beta_1 < i \leq \beta_0, \\ p_j s_0^{\beta_0-\beta_1} s_1^{i-\beta_1} \varphi_1, & \text{if } i > \beta_0, \end{cases} \quad (24)$$

where $\varphi_1 = \varphi(1, 0) + \varphi(1, 1)$, $s_0 = 1 - p_1 q_1$, $s_1 = 1 - p_0 q_0 - p_1 q_1$, and φ_1 satisfies following equation:

$$\varphi_1 = \frac{1}{\beta_1 + \frac{s_0 - s_0^{\beta_0 - \beta_1}}{1 - s_0} + s_0^{\beta_0 - \beta_1} \frac{s_1}{1 - s_1}}. \tag{25}$$

The average cost then is given by

$$\begin{aligned} & C_{mc}(\beta_0, \dots, \beta_{N-1}) \\ &= \varphi_1 \left(\frac{\beta_1(\beta_1 + 1)}{2} + A + B + \omega C_e E \right), \end{aligned} \tag{26}$$

where

$$A = \frac{s_0((\beta_1 + 1) - \beta_0 s_0^{\beta_0 - \beta_1})}{1 - s_0} + \frac{s_0^3 - s_0^{\beta_0 - \beta_1 + 1}}{(1 - s_0)^2}, \tag{27}$$

$$B = \frac{(\beta_0 + 1)s_1}{1 - s_1} + \frac{s_1^3}{1 - s_1}, \tag{28}$$

and

$$E = \left(s_0^{\beta_0 - \beta_1} \frac{1}{1 - s_1} + p_1 \frac{1 - s_0^{\beta_0 - \beta_1}}{1 - s_0} \right). \tag{29}$$

Proof. See Appendix D. □

Therefore, the closed form of the average cost is a function of thresholds. By linear search or gradient descent algorithm, the numerical solution of optimal thresholds can be obtained. However, computing its gradient directly requires a large amount of computation till convergence. Here, a nonlinear fractional programming (NLP) [17] based algorithm which can efficiently obtain the numerical solution is proposed.

Let $\mathbf{x} = (\beta_0, \beta_1)$. We can rewrite the cost function as a fractional form, where the numerator is denoted as $N(\mathbf{x}) = -C_{mc}(\mathbf{x}) / \varphi_1$, and the denominator term is $D(\mathbf{x}) = 1 / \varphi_1$. The solution to an NLP problem with the form in the following

$$\max \left\{ \frac{N(\mathbf{x})}{D(\mathbf{x})} \mid \mathbf{x} \in A \right\} \tag{30}$$

is related to the optimization problem (31)

$$\max \{ N(\mathbf{x}) - qD(\mathbf{x}) \mid \mathbf{x} \in A \}, \text{ for } q \in \mathbb{R}, \tag{31}$$

where the following assumption should also be satisfied:

$$D(\mathbf{x}) > 0, \text{ for all } \mathbf{x} \in A. \tag{32}$$

Define the function $F(q)$ with variable q as

$$F(q) = \max \{ N(\mathbf{x}) - qD(\mathbf{x}) \mid \mathbf{x} \in A \}, \text{ for } q \in \mathbb{R}. \tag{33}$$

According to [17], $F(q)$ is a strictly monotonic decreasing function and is convex over \mathbb{R} . Furthermore, we have $q_0 = N(\mathbf{x}_0) / D(\mathbf{x}_0) = \max \{ N(\mathbf{x}) - qD(\mathbf{x}) \mid \mathbf{x} \in A \}$ if, and only if,

$$F(q_0) = \max \{ N(\mathbf{x}) - q_0 D(\mathbf{x}) \mid \mathbf{x} \in A \} = 0. \tag{34}$$

Then, the algorithm can be described by two steps. The first step is to solve a convex optimization problem with a one dimensional parameter by a bisection method. The second step is to solve a high dimensional optimization problem by a gradient descent method.

According to [17], a bisection method can be used to solve the optimal q_0 , under the assumption that the value of function $F(q)$ can be obtained exactly for given q . We will actually use the gradient descent algorithm to obtain the numerical solution of $F(q)$ since the global search method may not perform in polynomial time. As a trick, we alternate the optimization variables of thresholds (β_0, β_1) by the variables of the decrement of thresholds, i.e., $\mathbf{x} = (\beta_0 - \beta_1, \beta_1)$. To summarize, the numerical-based method for computing the optimal thresholds is given by Algorithm 2.

Algorithm 2 Numerical computation of the optimal thresholds.

Input: Iteration time k , error threshold δ

Output: Numerical result \mathbf{x}^*

```

1: Let  $N(\mathbf{x}) = -C_{mc}(\mathbf{x})/\varphi_1$ , and  $D(\mathbf{x}) = 1/\varphi_1$ . Define  $F(q) = \max\{N(\mathbf{x}) - qD(\mathbf{x}) | \mathbf{x} \geq \mathbf{0}\}$ 
2: Let the iteration starts with  $i = 1$ , search range  $[a, b]$  of  $q$ .
3: while  $i \leq k$  do
4:    $m = \frac{a+b}{2}$ ;
5:   if  $F(m) * F(a) < 0$  then
6:      $b = m$ ;
7:   else
8:      $a = m$ ;
9:   end if
10:  if  $\frac{b-a}{2} < \delta$  then
11:     $\mathbf{x}^* = \arg \min_{\mathbf{x}} F(m)$ 
12:    break;
13:  end if
14:   $i = i + 1$ ;
15: end while

```

4. Simulation Results and Discussions

In this section, the simulation results are presented to investigate the impacts of the system parameters. We also compare the optimal policy with the zero-wait policy and periodic policy, where the zero-wait policy immediately generates an update at each time slot and the periodic policy keeps a constant interval between two updates.

Figure 4 depicts the optimal policy for different AoI and channel states, where the number of channel states is 5. It can be seen that, for each channel state, the optimal policy has a threshold structure with respect to the AoI. In particular, when the AoI is small, it is not beneficial for the sensor to generate and transmit a new update because the energy consumption dominates the total cost. We can also see that the threshold is non-increasing with the channel state. In other words, if the channel condition is better, the threshold is smaller. This is because the success probability of packet transmission increases with the channel state.

Figure 5 illustrates the thresholds for the MDP with two channel states with respect to the weighting factor ω , in which the two dashed lines are obtained by PIA and the other two solid lines are obtained by the proposed numerical algorithm. Both of the thresholds grow with the increasing of ω . Since the energy consumption has more weight, it is not efficient to update when the AoI is small. On the contrary, when ω decreases, the AoI dominates and the thresholds decline. In particular, both of the thresholds equal 1 when $\omega = 0$. In this case, the optimal policy reduces to the zero-wait policy. We can also see that the value of the threshold for channel state 1 of the numerical algorithm is close to the optimal solution. In contrast, the value of the threshold for channel state 0 gradually deviates from the optimal value.

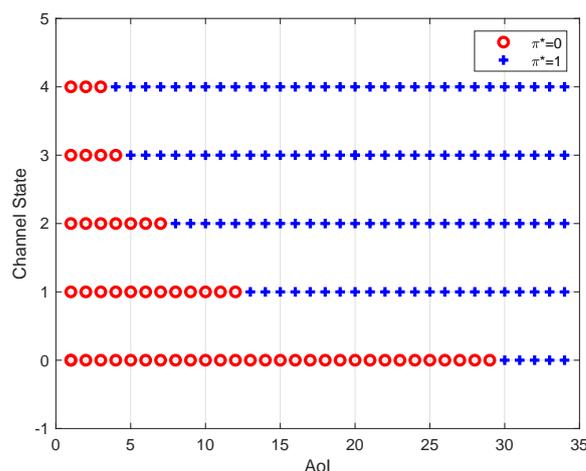


Figure 4. Optimal policy for different AoI and channel states ($q_0 = 0.1, q_1 = 0.2, q_2 = 0.3, q_3 = 0.4, q_4 = 0.5, p_0 = 0.1, p_1 = 0.1, p_2 = 0.3, p_3 = 0.3, p_4 = 0.2, \omega = 10, C_e = 1$).

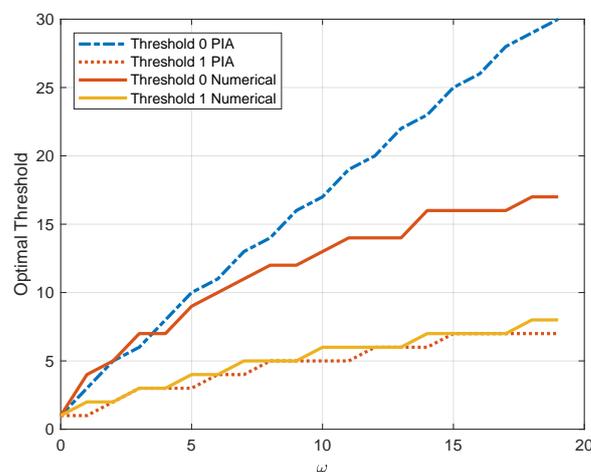


Figure 5. Optimal thresholds for two different channel states versus ω ($p_0 = 0.2, p_1 = 0.8, q_0 = 0.2, q_1 = 0.5, C_e = 1$).

Figure 6 illustrates the performance comparison of four policies, i.e., the zero-wait policy, the periodic policy, the numerical-based policy, and the optimal policy, with respect to the weighting factor ω . It is easy to see that the optimal policy has the lowest average cost. As we see in Figure 6, the zero-wait policy has the same performance with the optimal policy when $\omega = 0$. As ω increases, the average cost of all three policies increases. However, the increment of the zero-wait policy is larger than the periodic policy and the optimal policy due to the frequent transmission in the zero-wait policy. Although the thresholds obtained by the PIA and the numerical algorithm are not exactly the same as shown in Figure 5, the performance of the numerical-based algorithm also coincides with the optimal policy. This is because the threshold for channel state 1 exists in the quadratic term of the cost function, while the threshold for channel state 0 exists in the negative exponential term of the cost function. As a result, the threshold for channel state 1 has a much more significant effect on the system performance.

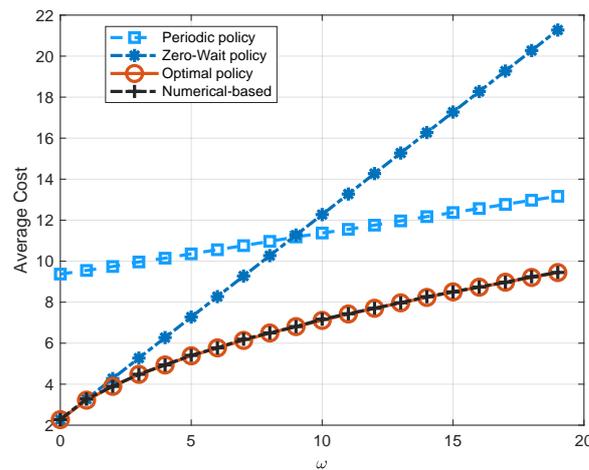


Figure 6. Comparison of the zero-wait policy, the periodic policy with period being 5, the numerical-based policy, and the optimal policy with respect to the weighting factor ω ($p_0 = 0.2, p_1 = 0.8, q_0 = 0.2, q_1 = 0.5, C_e = 1$).

Figure 7 compares the three policies with respect to the probability p_1 of the channel being in state 1. Since there is a higher probability that the channel has a good quality as p_1 increases, the average cost of all three policies decreases. We can see that, in the regime of p_1 , the optimal policy has the lowest average cost, because it can achieve a good balance between the AoI and the energy consumption. We can also see that the cost of the periodic policy is greater than the zero-wait policy first, and smaller later. To further demonstrate these curves, we separate the energy consumption term and AoI term into different figures, i.e., Figures 8 and 9. We see that the update cost of the zero-wait policy is smaller than that of the periodic policy, but the AoI of the zero-wait policy has a smaller decrease with respect to p_1 than the periodic policy.

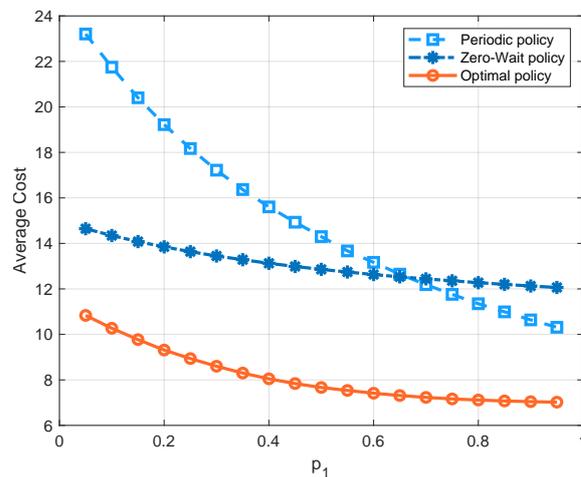


Figure 7. Comparison of the zero-wait policy, the periodic policy with period being 5, and the optimal policy with respect to p_1 ($q_0 = 0.2, q_1 = 0.5, \omega = 10, C_e = 1$).

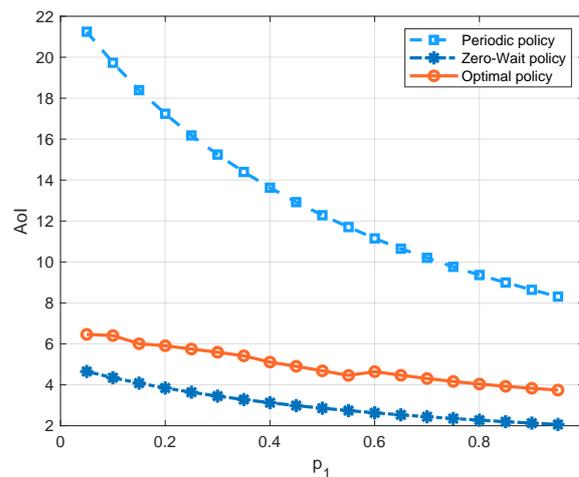


Figure 8. AoI comparison of the zero-wait policy, the periodic policy with period being 5, and the optimal policy with respect to p_1 ($q_0 = 0.2, q_1 = 0.5, \omega = 10, C_e = 1$).

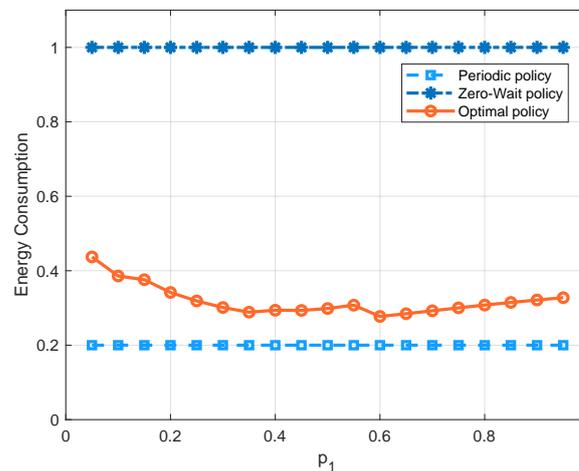


Figure 9. Energy consumption comparison of the zero-wait policy, the periodic policy with period being 5, and the optimal policy with respect to p_1 ($q_0 = 0.2, q_1 = 0.5, \omega = 10, C_e = 1$).

5. Conclusions

In this paper, we have studied the optimal updating policy in an IoT system, where the channel gain is quantized into multiple states and the channel state is fed back to the sensor before the decision making. The status update problem has been formulated as an MDP to minimize the long-term average of the weighted sum of the AoI and the energy consumption. By investigating the properties of the value function, it is proven that the optimal policy has a threshold structure with respect to AoI for any given channel state. We have also proven that the threshold is a non-increasing function of the channel state. Simulation results show the impacts of system parameters on the optimal thresholds and the average cost. Through comparisons, we have also shown that our proposed policy outperforms the zero-wait policy and the periodic policy. In our future research, the time-varying channel model will be further involved for guiding the future design of realistic IoT systems.

Author Contributions: Conceptualization, F.P., X.C., and X.W.; methodology, F.P. and X.W.; software, F.P.; validation, F.P., X.C., and X.W.; formal analysis, F.P. and X.W.; investigation, X.W.; writing—original draft preparation, F.P.; writing—review and editing, X.W. and X.C.; visualization, F.P. and X.W.; supervision, X.C. and X.W.; project administration, X.C.; funding acquisition, X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the State's Key Project of Research and Development Plan under Grants (2019YFE0196400), in part by Guangdong R&D Project in Key Areas under Grant (2019B010158001), in part by Guangdong Basic and Applied Basic Research Foundation under Grant (2021A1515012631), in part by Key Laboratory of Modern Measurement & Control Technology, Ministry of Education, Beijing Information Science & Technology University (KF20201123202).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Lemma 2

Based on the value iteration algorithm, the induction method can be employed in following proof. Firstly, we initial that $V_{\gamma,0}(\mathbf{x}) = 0$, where both Equations (15) and (16) hold for all $\mathbf{x} \in \mathcal{S}$.

Appendix A.1. Proof of Equation (16)

When $n = 1$,

$$Q_{\gamma,1}(\Delta_1, i, 0) - Q_{\gamma,1}(\Delta_2, i, 0) = \Delta_1 - \Delta_2 \leq 0, \quad (\text{A1})$$

and

$$Q_{\gamma,1}(\Delta_1, i, 1) - Q_{\gamma,1}(\Delta_2, i, 1) = \Delta_1 + \omega C_e - (\Delta_2 + \omega C_e) \leq 0, \quad (\text{A2})$$

hold due to $\Delta_1 \leq \Delta_2$, and we have $V_{\gamma,1}(\Delta_1, i) \leq V_{\gamma,1}(\Delta_2, i)$.

Suppose that $V_{\gamma,k}(\Delta_1, i) \leq V_{\gamma,k}(\Delta_2, i)$ holds for $k \leq K$. Considering the case of $k = K + 1$,

$$Q_{\gamma,K+1}(\Delta_1, i, 0) - Q_{\gamma,K+1}(\Delta_2, i, 0) = \Delta_1 - \Delta_2 \leq 0, \quad (\text{A3})$$

and

$$\begin{aligned} & Q_{\gamma,K+1}(\Delta_1, i, 1) - Q_{\gamma,K+1}(\Delta_2, i, 1) \\ &= (\Delta_1 - \Delta_2) + \gamma \sum_{j \in \mathbb{H}} p_j g_i (V_{\gamma,K}(\Delta_1 + 1, j) - V_{\gamma,K}(\Delta_2 + 1, j)) \\ &\leq 0, \end{aligned} \quad (\text{A4})$$

hold for all i according to $\Delta_1 \leq \Delta_2$. Therefore, we have $V_{\gamma,K+1}(\Delta_1, i) \leq V_{\gamma,K+1}(\Delta_2, i)$. Since $\lim_{n \rightarrow \infty} V_{\gamma,n} = V_{\gamma}$, we have $V_{\gamma}(\Delta_1, i) \leq V_{\gamma}(\Delta_2, i)$.

Appendix A.2. Proof of Equation (15)

By the definition of function $Q_{\gamma}(\mathbf{x}, a)$, we have

$$Q_{\gamma}(\Delta, i, 0) = \Delta + \gamma \sum_{j \in \mathbb{H}} p_j V_{\gamma}(\Delta + 1, j), \quad (\text{A5})$$

and

$$Q_{\gamma}(\Delta, i, 1) = \Delta + \omega C_e + \gamma \left(\sum_{j \in \mathbb{H}} p_j (g_i V_{\gamma}(\Delta + 1, j) + q_i V_{\gamma}(1, j)) \right). \quad (\text{A6})$$

Therefore,

$$Q_\gamma(\Delta, N - 1, 0) - Q_\gamma(\Delta, i, 0) \leq 0, \tag{A7}$$

and

$$Q_\gamma(\Delta, N - 1, 1) - Q_\gamma(\Delta, i, 1) \stackrel{(a)}{=} \gamma \sum_{j \in \mathbb{H}} p_j (q_{N-1} - q_i) (V_\gamma(1, j) - V_\gamma(\Delta + 1, j)) \leq 0, \tag{A8}$$

hold for all i , where step (a) is due to Equation (16). Hence, we have $V_\gamma(\Delta, N - 1) \leq V_\gamma(\Delta, i)$. This completes the whole proof.

Appendix B. Proof of Theorem 1

Theorem 1 can be proven by verifying the conditions given in [16]. The conditions are listed as follows:

- (1): For every state \mathbf{x} and discount factor γ , the discount value function $V_\gamma(\mathbf{x})$ is finite.
- (2): There exists a non-negative value L such that $-L \leq h_\gamma(\mathbf{x})$ for all \mathbf{x} and γ , where $h_\gamma(\mathbf{x}) = V_\gamma(\mathbf{x}) - V_\gamma(\hat{\mathbf{x}})$, and $\hat{\mathbf{x}}$ is a reference state.
- (3): There exists a non-negative value $M_{\mathbf{x}}$, such that $h_\gamma(\mathbf{x}) \leq M_{\mathbf{x}}$ for every \mathbf{x} and γ . For every \mathbf{x} , there exists an action $a_{\mathbf{x}}$ such that $\sum_{\mathbf{x}'} \Pr(\mathbf{x}'|\mathbf{x}, a_{\mathbf{x}}) M_{\mathbf{x}'} < \infty$.
- (4): The inequality $\sum_{\mathbf{x}'} \Pr(\mathbf{x}'|\mathbf{x}, a) M_{\mathbf{x}'} < \infty$ holds for all \mathbf{x} and a .

By Lemma 1, $V_\gamma(\hat{\mathbf{x}}) = \min_{\pi} V_{\pi, \gamma}(\hat{\mathbf{x}}) < \infty$ holds for every $\hat{\mathbf{x}}$ and γ . Hence, condition (1) holds. According to Lemma 2, by letting $\hat{\mathbf{x}} = (1, N - 1)$ and $L = 0$, we have $h_\gamma(\mathbf{x}) \geq 0$, which verifies condition (2).

Before verifying condition (3), a lemma is given as follows:

Lemma A1. *Let us denote $\hat{\mathbf{x}} = (1, N - 1)$ as the reference state and define the first time that an initial state \mathbf{x} transits to $\hat{\mathbf{x}}$ as $K = \min\{k : k \leq 1, \mathbf{x}_k = \hat{\mathbf{x}}\}$. Then, the expectation cost under the always-transmitting policy π_a , i.e., the sensor generates and transmits a new update in each slot, is*

$$C_{\mathbf{x}, \hat{\mathbf{x}}}(\pi_a) = \mathbb{E}_{\pi_a} \left[\sum_{t=0}^{K-1} \gamma^t C(\mathbf{x}_t, a_t) | \mathbf{x} \right], \tag{A9}$$

where $C_{\mathbf{x}, \hat{\mathbf{x}}}(\pi_a) < \infty$ holds for all \mathbf{x} .

Proof. Since $a_t = 1$ for all t , the probability that the state returns to $\hat{\mathbf{x}}$ from \mathbf{x} after exactly K slot is given by

$$\Pr(K = k | \mathbf{x} = (\Delta, j)) = \begin{cases} p_{N-1} q_j, & \text{if } k = 1, \\ p_{N-1} g_j (1 - \sum_{i \in \mathbb{H}} p_i q_i)^{k-2} (\sum_{i \in \mathbb{H}} p_i q_i), & \text{otherwise.} \end{cases} \tag{A10}$$

Then, the expectation return cost from \mathbf{x} to $\hat{\mathbf{x}}$ is expressed as

$$\begin{aligned}
 & C_{\mathbf{x},\hat{\mathbf{x}}}(\pi_a) \\
 &= \mathbb{E} \left[\sum_{t=0}^{K-1} \gamma^t C(\mathbf{x}_t, a_t) | \mathbf{x} \right] \\
 &\stackrel{(a)}{\leq} \sum_{k=1}^{\infty} \Pr(K = k | \mathbf{x} = (\Delta, j)) \left[\sum_{m=0}^{k-1} (\Delta + m + \omega C_e) \right] \\
 &< \infty,
 \end{aligned} \tag{A11}$$

where step (a) is due to the fact that $C(\mathbf{x}_t, a_t) \leq \Delta_t + \omega C_e$. \square

Considering a mixture policy π , in which it performs the always-transmitting policy π_a from initial state \mathbf{x} until it enters the reference state $\hat{\mathbf{x}}$, it later performs the optimal policy π_γ that minimizes the discounted cost. Therefore, we have

$$\begin{aligned}
 & V_\gamma(\mathbf{x}) \\
 &\leq \mathbb{E}_{\pi_a} \left[\sum_{t=0}^{K-1} \gamma^t C(\mathbf{x}_t, a_t) | \mathbf{x} \right] + \mathbb{E}_{\pi_b} \left[\sum_{t=K}^{\infty} \gamma^t C(\mathbf{x}_t, a_t) | \hat{\mathbf{x}} \right] \\
 &\leq C_{\mathbf{x},\hat{\mathbf{x}}}(\pi_a) + \mathbb{E}_\pi \left[\gamma^K V_\gamma(\hat{\mathbf{x}}) \right] \\
 &\leq C_{\mathbf{x},\hat{\mathbf{x}}}(\pi_a) + V_\gamma(\hat{\mathbf{x}}),
 \end{aligned} \tag{A12}$$

which implies that $h_\gamma(\mathbf{x}) \leq C_{\mathbf{x},\hat{\mathbf{x}}}(\pi_a)$. Hence, let $\hat{\mathbf{x}} = (1, N - 1)$ and $M_{\mathbf{x}} = C_{\mathbf{x},\hat{\mathbf{x}}}(\pi_a)$; condition (3) is verified.

On the other hand, $M_{\mathbf{x}} < \infty$ holds for all \mathbf{x} . The states that transit from \mathbf{x} are finite. Thus, the weighted sum of finite $M_{\mathbf{x}}$ is also finite, i.e., $\sum_{\mathbf{x}'} P(\mathbf{x}' | \mathbf{x}, a) M_{\mathbf{x}'} < \infty$ holds for all \mathbf{x} and a , which verifies condition (4). This completes the whole verification.

Appendix C. Proof of Theorem 2

Based on the definition of $Q(\Delta, i, a)$, we can obtain the difference between the state-action value function as follows:

$$\begin{aligned}
 & Q(\Delta, i, 0) - Q(\Delta, i, 1) \\
 &= \sum_{j=0}^{N-1} p_j V(\Delta + 1, j) - q_i \sum_{j=0}^{N-1} p_j V(1, j) - g_i \sum_{j=0}^{N-1} p_j V(\Delta + 1, j) - \omega C_e \\
 &= q_i \sum_{j=0}^{N-1} p_j (V(\Delta + 1, j) - V(1, j)) - \omega C_e \\
 &\stackrel{(a)}{\geq} q_i \Delta - \omega C_e.
 \end{aligned} \tag{A13}$$

where (a) is due to the property of the value function given in Lemma 4. We then discuss the difference between the state-action value function in two cases.

Case 1: $\omega = 0$.

In this case, $Q(\Delta, i, 0) - Q(\Delta, i, 1) \geq 0$ holds for any Δ and i . Therefore, the optimal policy is to update at each slot in spite of the channel state. In other words, the optimal thresholds are all equal to 1.

Case 2: $\omega > 0$.

We note that, given i , $q_i \Delta - \omega C_e$ increases linearly with Δ . Hence, there exists a positive integer $\hat{\beta}_i$, such that $\hat{\beta}_i$ is the minimum value that satisfies $q_i \hat{\beta}_i - \omega C_e \geq 0$. Therefore, if $\Delta \geq \hat{\beta}_i$, $Q(\Delta, i, 0) - Q(\Delta, i, 1) \geq q_i \hat{\beta}_i - \omega C_e \geq 0$ holds. This implies that there must exist a threshold β_i satisfying $1 \leq \beta_i \leq \hat{\beta}_i$. If $\Delta \geq \beta_i$, we have $Q(\Delta, i, 0) - Q(\Delta, i, 1) \geq 0$.

Altogether, the optimal policy has a threshold structure for $\omega \geq 0$. Then, we examine the non-increasing property of the thresholds. Firstly, we show that the difference between the state-action value function is monotonic with respect to the channel state by fixing the AoI. Assuming that $i, j \in \mathbb{H}$ and $i \leq j$, it is easy to obtain that

$$\begin{aligned} & Q(\Delta, j, 0) - Q(\Delta, j, 1) - (Q(\Delta, i, 0) - Q(\Delta, i, 1)) \\ &= (q_j - q_i) \sum_{l=0}^{N-1} p_l (V(\Delta + 1, l) - V(1, l)) \geq 0. \end{aligned} \quad (\text{A14})$$

Since $Q(\Delta, i, 0) - Q(\Delta, i, 1) \geq 0$ when $\Delta \geq \beta_i$, we have $Q(\Delta, j, 0) - Q(\Delta, j, 1) \geq 0$ according to (A14). This implies that the optimal threshold β_j corresponding to channel state j is no greater than β_i , i.e., $\beta_j \leq \beta_i$. This completes the whole proof.

Appendix D. Proof of Theorem 3

Assume that $\varphi(\mathbf{x})$ is the steady probability of state \mathbf{x} in a Markov chain. The steady state probability $\varphi(\mathbf{x})$ satisfies the following global balance equation [18], i.e.,

$$\varphi(\mathbf{x}) = \sum_{\mathbf{x}' \in \mathcal{S}} \varphi(\mathbf{x}') \Pr(\mathbf{x}|\mathbf{x}'). \quad (\text{A15})$$

Let $\varphi_1 = \varphi(1, 0) + \varphi(1, 1)$. We prove Equation (A23) by discussing three cases via mathematical induction.

Case 1: $1 < i \leq \beta_1$

Based on Equation (A15), we have

$$\begin{aligned} \varphi(2, j) &= \varphi(1, 0)p_j + \varphi(1, 1)p_j \\ &= p_j \varphi_1. \end{aligned} \quad (\text{A16})$$

Assuming that $\varphi(i, j) = p_j \varphi_1$ holds for all $i \leq k < \beta_1$, we examine $\varphi(k+1, j)$. We have

$$\begin{aligned} \varphi(k+1, j) &= \varphi(k, 0)p_j + \varphi(k, 1)p_j \\ &= p_j p_0 \varphi_1 + p_j p_1 \varphi_1 \\ &= p_j \varphi_1, \end{aligned} \quad (\text{A17})$$

which completes this segment of the proof.

Case 2: $\beta_1 < i \leq \beta_0$

Similarly, we have

$$\begin{aligned} \varphi(\beta_1 + 1, j) &= p_j(1 - q_1)\varphi(\beta_1, 1) + p_j\varphi(\beta_1, 0) \\ &= p_j\varphi_1((1 - q_1)p_1 + p_0) \\ &= p_j\varphi_1 s_0, \end{aligned} \quad (\text{A18})$$

where $s_0 = 1 - p_1 q_1$. Assuming that $\varphi(i, j) = p_j \varphi_1 s_0^{i-\beta_1}$ holds for all $\beta_1 < i \leq k < \beta_0$, we examine $\varphi(k+1, j)$. We have

$$\begin{aligned} \varphi(k+1, j) &= p_j(1 - q_1)\varphi(k, 1) + p_j\varphi(k, 0) \\ &= p_j\varphi_1((1 - q_1)p_1 + p_0)s_0^{k-\beta_1} \\ &= p_j\varphi_1 s_0^{k+1-\beta_1}, \end{aligned} \quad (\text{A19})$$

which completes this segment of the proof.

Case 3: $i > \beta_0$

Follow above discussion, we have

$$\begin{aligned} \varphi(\beta_0 + 1, j) &= p_j(1 - q_1)\varphi(\beta_0, 1) + p_j(1 - q_0)\varphi(\beta_0, 0) \\ &= p_j\varphi_1 s_0^{\beta_0 - \beta_1} ((1 - q_1)p_1 + (1 - q_0)p_0) \\ &= p_j\varphi_1 s_0^{\beta_0 - \beta_1} s_1, \end{aligned} \tag{A20}$$

where $s_1 = 1 - p_0q_0 - p_1q_1$. Assuming that $\varphi(i, j) = p_j\varphi_1 s_0^{\beta_0 - \beta_1} s_1^{i - \beta_0}$ holds for all $\beta_0 < i \leq k$, we examine $\varphi(k + 1, j)$. We have

$$\begin{aligned} \varphi(k + 1, j) &= p_j(1 - q_1)\varphi(k, 1) + p_j(1 - q_0)\varphi(k, 0) \\ &= p_j\varphi_1 s_0^{\beta_0 - \beta_1} ((1 - q_1)p_1 + (1 - q_0)p_0) s_1^{k - \beta_0} \\ &= p_j\varphi_1 s_0^{\beta_0 - \beta_1} s_1^{k + 1 - \beta_0}. \end{aligned} \tag{A21}$$

Altogether, we obtain the steady state probability with respect to an unknown parameter φ_1 . According to the fact that $\sum_{i=1}^{\infty} \sum_{j=0}^{N-1} \varphi(i, j) = 1$, we formulate an equation:

$$\begin{aligned} &\sum_{i=1}^{\infty} \sum_{j=0}^{N-1} \varphi(i, j) \\ &= \varphi_1 \left\{ \beta_1 + \sum_{i=\beta_1+1}^{\beta_0} s_0^{i-\beta_1} + \sum_{i=\beta_0+1}^{\infty} s_0^{\beta_0-\beta_1} s_1^{i-\beta_0} \right\} \\ &= \varphi_1 \left\{ \beta_1 + \frac{s_0 - s_0^{\beta_1-\beta_0}}{1 - s_0} + s_0^{\beta_1-\beta_0} \frac{s_1}{1 - s_1} \right\} \\ &= 1, \end{aligned} \tag{A22}$$

where the expression of φ_1 is obtained.

The average cost of a Markov chain is given by

$$C_{mc} = \sum_{\mathbf{x} \in \mathcal{S}} \varphi(\mathbf{x}) C(\mathbf{x}, \pi^*(\mathbf{x})). \tag{A23}$$

Substituting (24) into (A23), we have

$$\begin{aligned} &C_{mc}(\beta_0, \beta_1) \\ &= \sum_{i=1}^{\infty} i \sum_{j=0}^1 \varphi(i, j) + \sum_{i=\beta_0}^{\infty} \omega C_e \varphi(i, 0) + \sum_{i=\beta_1}^{\infty} \omega C_e \varphi(i, 1). \end{aligned}$$

Furthermore, the first term is given by

$$\begin{aligned} &\sum_{i=1}^{\infty} i \sum_{j=0}^1 \varphi(i, j) \\ &= \sum_{i=1}^{\beta_1} i \varphi_i + \sum_{i=\beta_1+1}^{\beta_0} i s_0^{i-\beta_1} + s_0^{\beta_0-\beta_1} \sum_{i=\beta_0+1}^{\infty} i s_1^{i-\beta_0} \\ &= \frac{\varphi_1 \beta_1 (\beta_1 + 1)}{2} + \varphi_1 A + \varphi_1 B, \end{aligned} \tag{A24}$$

where

$$A = \frac{s_0((\beta_1 + 1) - \beta_0 s_0^{\beta_0 - \beta_1})}{1 - s_0} + \frac{s_0^3 - s_0^{\beta_0 - \beta_1 + 1}}{(1 - s_0)^2}, \tag{A25}$$

and

$$B = \frac{(\beta_0 + 1)s_1}{1 - s_1} + \frac{s_1^3}{1 - s_1}. \quad (\text{A26})$$

Furthermore, the sum of last two terms is given by

$$\begin{aligned} & \sum_{i=\beta_0}^{\infty} \omega C_e \varphi(i, 0) + \sum_{i=\beta_1}^{\infty} \omega C_e \varphi(i, 1) \\ &= \varphi_1 \omega C_e \left(s_0^{\beta_0 - \beta_1} \sum_{i=\beta_0}^{\infty} s_1^{i - \beta_0} + p_1 \sum_{i=\beta_1}^{\beta_0 - 1} s_0^{i - \beta_1} \right) \\ &= \varphi_1 \omega C_e \left(s_0^{\beta_0 - \beta_1} \frac{1}{1 - s_1} + p_1 \frac{1 - s_0^{\beta_0 - \beta_1}}{1 - s_0} \right). \end{aligned} \quad (\text{A27})$$

This completes the proof.

References

1. Kaul, S.; Yates, R.; Gruteser, M. Real-time status: How often should one update? In Proceedings of the IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 2731–2735.
2. Abd-Elmagid, M.A.; Dhillon, H.S. Average peak age-of-information minimization in UAV-assisted IoT networks. *IEEE Trans. Veh. Technol.* **2018**, *68*, 2003–2008. [\[CrossRef\]](#)
3. Farazi, S.; Klein, A.G.; Brown, D.R. Average age of information for status update systems with an energy harvesting server. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, USA, 15–19 April 2018; pp. 112–117.
4. Costa, M.; Codreanu, M.; Ephremides, A. On the age of information in status update systems with packet management. *IEEE Trans. Inf. Theory* **2016**, *62*, 1897–1910. [\[CrossRef\]](#)
5. Sun, Y.; Uysal-Biyikoglu, E.; Yates, R.D.; Koksall, C.E.; Shroff, N.B. Update or wait: How to keep your data fresh. *IEEE Trans. Inf. Theory* **2017**, *63*, 7492–7508. [\[CrossRef\]](#)
6. Yun, J.; Joo, C.; Eryilmaz, A. Optimal real-time monitoring of an information source under communication costs. In Proceedings of the 2018 IEEE Conference on Decision and Control (CDC), Miami, FL, USA, 17–19 December 2018; pp. 4767–4772.
7. Imer, O.C.; Basar, T. Optimal estimation with limited measurements. In Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain, 15 December 2005; pp. 1029–1034.
8. Rabi, M.; Moustakides, G.V.; Baras, J.S. Adaptive sampling for linear state estimation. *Siam J. Control Optim.* **2009**, *50*, 672–702. [\[CrossRef\]](#)
9. Gao, X.; Akyol, E.; Başar, T. On remote estimation with multiple communication channels. In Proceedings of the 2016 American Control Conference (ACC), Boston, MA, USA, 6–8 July 2016; pp. 5425–5430.
10. Chakravorty, J.; Mahajan, A. Remote-state estimation with packet drop. *IFAC-PapersOnLine* **2016**, *49*, 7–12. [\[CrossRef\]](#)
11. Lipsa, G.M.; Martins, N.C. Optimal state estimation in the presence of communication costs and packet drops. In Proceedings of the 2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, USA, 30 September–2 October 2009; pp. 160–169.
12. Huang, H.; Qiao, D.; Gursoy, M.C. Age-energy tradeoff in fading channels with packet-based transmissions. In Proceedings of the IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 6–9 July 2020; pp. 323–328.
13. Chakravorty, J.; Mahajan, A. Remote estimation over a packet-drop channel with Markovian state. *IEEE Trans. Autom. Control* **2019**, *65*, 2016–2031. [\[CrossRef\]](#)
14. Chen, X.; Yi, H.; Luo, H.; Yu, H.; Wang, H. A novel CQI calculation scheme in LTE/LTE-A systems. In Proceedings of the 2011 International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, 9–11 November 2011; pp. 1–5.
15. Toyserkani, A.T.; Strom, E.G.; Svensson, A. An analytical approximation to the block error rate in Nakagami-m non-selective block fading channels. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 1543–1546. [\[CrossRef\]](#)
16. Sennott, L.I. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Oper. Res.* **1989**, *37*, 626–633. [\[CrossRef\]](#)
17. Dinkelbach, W. On nonlinear fractional programming. *Manag. Sci.* **1967**, *13*, 492–498. [\[CrossRef\]](#)
18. Chandry, K.M. *The Analysis and Solutions for General Queueing Networks*; Citeseer: Princeton, NJ, USA, 1974.