Age optimal sampling for unreliable channels under unknown channel statistics

Hongyi He[1,2], Haoyue Tang[3], Jiayu Pan[4], Jintao Wang[1,2,5], Jian Song[1,2,6], Leandros Tassiulas[3],
[1]Beijing National Research Center for Information Science and Technology (BNRist)
[2]Department of Electronic Engineering, Tsinghua University, Beijing, China
[3]Department of Electronic Engineering, Yale University, New Haven, USA
[4] Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, USA
[5]Research Institute of Tsinghua University in Shenzhen, Shenzhen, China
[6]Shenzhen International Graduate School, Tsinghua University, Shenzhen, China
Email: hehy22@mails.tsinghua.edu.cn, wangjintao@tsinghua.edu.cn

***Abstract*—In this work, we study a system with a sensor forwarding status update to the receiver through an error-prone channel, and the receiver sends the transmission results to the sensor via a reliable link. We assume both transmission links suffer from random delays. We use Age of Information (AoI) to measure the freshness of the status information at the receiver. Our goal is to design a sampling policy that minimizes the expected time average AoI when the channel statistics are unknown. The problem is reformulated into a renewal-reward process optimization, and an online algorithm based on the Robbins-Monro algorithm is proposed. We prove that when the forward and backward transmission delays are bounded, the AoI difference between the online algorithm and the optimal policy decays with rate $\mathcal{O}(\ln K/K)$, where K is the number of successful transmissions. Simulation results validate the performance of our proposed algorithm.**

***Index Terms*—Age-of-Information, Online learning, Renewal-Reward Process, Unreliable Transmissions**

I. Introduction

With the widespread application of real-time control systems, such as autonomous driving and intelligent production, timely status update is becoming increasingly important to support the demand for monitoring real-time status [1]. To measure the freshness of the status update, Age of Information (AoI) has been proposed [2]. AoI is the time difference between the current time and the generation time of the freshest packet [2]. Maintaining a smaller AoI ensures more timely status updates from the remote source and thus improves the performance of the monitoring system.

Sampling for AoI minimization has received a lot of attention from researchers [3]–[12]. When the transmission statistics, such as delay distributions, and packet-loss probabilities are known in advance, the minimization of average AoI can be formulated into a Markov decision process (MDP). To reduce the age, we may require the sensor to wait before taking a new sample, i.e., the zero-wait policy may not be the optimum [10]. Optimal sampling policies for the status update system with one-way delay, i.e., there is no feedback delay, and with two-way delay, i.e., both forward and feedback delays may be nonzero, are studied in [12], [13], respectively. These optimal policies exhibit a similar threshold structure where the threshold depends on the optimal objective value and the constraint of sampling rate. The previous works assume reliable transmissions. However, because of fading and interference in the real environment, the channel is unreliable and the packet will encounter transmission failures. In [8], unreliable transmission is considered and an optimal sampling policy is derived. For unreliable channels, the sensor only waits if the last transmission is successful, otherwise, it will take a new sample immediately.

When the channel statistics are unknown, designing age-optimal sampling policies can be formulated as a sequential decision problem. One way to solve this problem is online learning, which has low complexity and easy analysis. Online learning has been used to design sampling strategies in reliable channels. Tang et al. used Robbins-Monro to obtain the age-optimal sampling policy adaptively and performed a performance analysis for a one-way delay model [14], [15]. Furthermore, a similar online algorithm is derived to minimize the MSE when sampling for a wiener process in [16]. In [17], the authors proposed an online algorithm for sampling in a system with two-way delay but without a sampling rate constraint. However, unreliable transmission is not considered in the literature above. In [18], [19], the authors studied optimal sampling policy for unreliable channels with unknown channel statistics. Nonetheless, theoretical analysis such as the convergence rate for the optimality gap, i.e., the cumulative AoI difference between the proposed algorithm and the optimal policy is not provided in [17]–[19].

To address the previously mentioned challenges, in this paper, we aim to minimize AoI with unknown channel statistics, under one of the most general channel settings in the literature: unreliable transmissions with random two-way delay. Note that the above channel settings are similar to that of [8], but without access to channel statistics. Our objective is to minimize the expected time-averaged AoI under a frequency constraint. Then, based on Robbins-Monro algorithm, we propose an online algorithm to adaptively learn the optimal

The work of J. Song was supported by National Key Research and Development Program of China under Grant 2022YFE0101700.
Corresponding author: Jintao Wang.

ISBN 978-3-903176-55-3 © 2023 IFIP

sampling policy without channel statistics. Additionally, we provide a theoretical analysis of the convergence rate when there is no frequency constraint. We show that the gap between the expected cumulative AoI of the online algorithm and the AoI of the optimal policy decays with rate $O(\frac{\ln K}{K})$. Finally, simulations are conducted to validate the performance of our proposed online algorithm.

II. Problem Formulation

A. System Model

We consider a system as Fig. 1 shows. The system comprises a sensor, a receiver, a forward sensor-to-receiver channel, and a backward receiver-to-sensor channel. The sensor takes a sample of the latest system state and submits the fresh sample to the channel. Due to fading and interference that exists in the real environment, we assume the forward transmission is unreliable with a random delay. After the transmission is complete, the receiver immediately sends 1-bit feedback through the reliable backward channel to indicate whether the transmission was successful (ACK) or not (NACK). We assume that the backward transmission time is random.

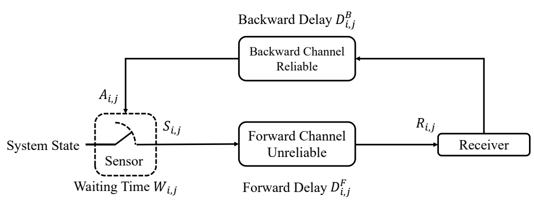

Fig. 1. System Model

Due to possible transmission failures, we may need several attempts to successfully transmit a sample. Therefore, to describe easily, we introduce the notion of epoch. Denote $i \in \{1, 2, \cdots\}$ as the number of successfully transmitted packets. Then, the i-th epoch represents the time interval between the sampling time of the i-th successful transmission and the sampling time of the $(i+1)$-th successful transmission. We assume that the forward transmission experiences i.i.d. transmission failures with a probability of α. Therefore, the total number of transmissions attempts in the i-th epoch, denoted by M_i, follows a geometric distribution with parameter $1-\alpha$. We use j to denote the number of attempted samples in the i-th epoch, where we have $1 \le j \le M_i$. Specifically, when $j = 1$, the previous sample is successfully delivered to the receiver. Throughout the paper, we will denote the tuple (i, j) as the index of j-th sampled packet in the i-th epoch.

At the i-th epoch, the sensor takes the j-th sample at time $S_{i,j}$, then the sample experiences a random delay of $D^F_{i,j}$ in the forward channel before arriving at the receiver. The reception time is denoted as $R_{i,j}$. At $R_{i,j}$, the receiver sends immediate feedback that undergoes a backward random delay $D^B_{i,j}$ and arrives at the sensor at time $A_{i,j}$. The sensor waits for a time period $W_{i,j}$ before taking the next sample. The waiting time $W_{i,j}$ is decided by our policy, and we assume that $W_{i,j}$ is bounded. We assume that the forward delay and the backward delay are mutually independent and follow their independent and identically distributed probabilities $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$, respectively. We provide a further technical assumption on $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$:

Assumption 1: The probability measure $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$ are both absolutely continuous on $[0, \infty)$. Their expectations are upper bounded by $\overline{D^F}_{\text{ub}}$ and $\overline{D^B}_{\text{ub}}$, respectively, and are lower bounded by $\overline{D^F}_{\text{lb}}$ and $\overline{D^B}_{\text{lb}}$, respectively. Their second moments $\overline{H^F} \triangleq \mathbb{E}_{\mathbb{P}_{\text{BD}}}[(D^F)^2]$ and $\overline{H^B} \triangleq \mathbb{E}_{\mathbb{P}_{\text{BD}}}[(D^B)^2]$ are upper bounded by $\overline{H^F}_{\text{ub}}$ and $\overline{H^B}_{\text{ub}}$, respectively, and lower bounded by $\overline{H^F}_{\text{lb}}$ and $\overline{H^B}_{\text{lb}}$, respectively.

B. Age of Information

We measure how fresh the data is at the receiver using the metric Age of Information (AoI). By definition [20], AoI equals the time difference between the current time and the generation time of the freshest sample. Let $Z_t = \max_i\{S_{i,1} : R_{i,1} \le t\}$. Note that only the first package in an epoch is successfully delivered. Then, the AoI $A(t)$ of the current time t is defined as

$$A(t) \triangleq t - Z_t. \tag{1}$$

A sample path of AoI evolution is depicted in Fig. 2. After a successful delivery, the AoI decreases to the transmission delay $D^F_{i+1,1}$ of the first sample at the $(i+1)$-th epoch. Otherwise, the AoI grows linearly.

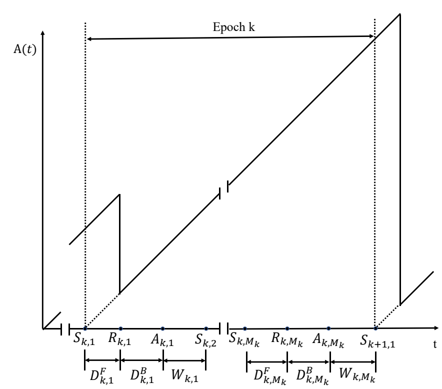

Fig. 2. AoI Evolution

C. Problem Formulation

Our objective is to design a sampling policy $\pi \triangleq \{W_{i,j}\}$, to minimize the average AoI under a sampling frequency constraint when the delay distributions $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$ and the packet loss probability α are unknown. We only consider the causal policies Π in which the waiting time is selected based on the history information.

Problem 1:

$$\begin{aligned} \text{AoI}_{\text{opt}} &= \inf_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T A(t)\mathrm{d}t\right], \\ &\text{s.t.} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}[C(T)] \le f_{\max}. \end{aligned} \tag{2}$$

Here, $C(T)$ is the total number of samples taken in $[0, T]$.

III. Solution to the Problem

In this section, we first reformulate the problem into a renewal-reward process optimization. Then, we will review the optimal sampling policy $\pi^\star$ when the delay distribution $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$ are known. Based on the optimal offline policy, when $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$ are unknown, we will propose an online algorithm that adaptively learns the waiting policy $\pi^\star$.

A. A Renewal-Reward Process Reformulation

A policy $\pi \in \Pi$ is *stationary* and *deterministic* if the waiting time $W_{i,j}$ is a deterministic mapping from the previous transmission delay information. Let $\Pi_{SD} \subseteq \Pi$ be the set of stationary deterministic policies.

Denote the cumulative AoI in the i-th epoch as $F_i \triangleq \int_{S_{i,1}}^{S_{i+1,1}} A(t)\mathrm{d}t$ and the epoch length as $L_i \triangleq S_{i+1,1} - S_{i,1}$. Because the policy π is stationary, and the delays are i.i.d, F_i and L_i are both i.i.d variables and follow a regenerative process. By renewal theory, similar to [8, Section V.A], we can reformulate Problem 1 as follows

Problem 2:

$$\text{AoI}_{\text{opt}} = \inf_{\pi\in\Pi} \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left[F_i\right]}{\sum_{i=1}^n \mathbb{E}\left[L_i\right]}, \tag{3a}$$

$$\text{s.t.} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[L_i\right] \geq \frac{\sum_{i=1}^n \mathbb{E}\left[M_i\right]}{f_{\max}}. \tag{3b}$$

Notice that F_i is non-decreasing in $W_{i,j}$. Since $W_{i,j}$ is bounded, i.e., $W_{i,j} \in [0, W_{\text{ub}}]$, we assume

$$\mathbb{E}[F_i | W_{i,j} = W_{\text{ub}}] < \infty, \tag{4}$$

which implies $\mathbb{E}[F_i] < \infty$ for all $W_{i,j} \in [0, W_{\text{ub}}]$.

Since $A(t) : [0,\infty) \mapsto [0,\infty)$ is non-decreasing and if (4) is satisfied, we have the following theorem.

Theorem 1: [12, Theorem 2, Restated] There exists a *stationary* and *deterministic* policy $\pi^\star \in \Pi_{SD}$ that is optimal to Problem 2.

Then, we search for $\pi^\star$ in the policy space Π_{SD} without losing optimality. The following corollary reveals the structure of the AoI minimum sampling policy, which further reduce our search space [8].

Corollary 1: Problem 2 can be represented by a function $w : \mathbb{R}^2 \to \mathbb{R}$, where the waiting time $W_{i,j}$ is selected by:

$$\begin{aligned} W_{i,1} &= w(D_{i,1}^F, D_{i,1}^B), \\ W_{i,j} &= 0 \quad j = 2, 3, \cdots. \end{aligned} \tag{5}$$

Then, the expected time-average AoI of each policy that satisfies corollary 1 with waiting time selection function w can be computed by:

Problem 3 (Renewal-Reward Process Optimization Reformulation):

$$\text{AoI}_{\text{opt}} = \inf_{\pi\in\Pi_{\text{SD}}} \left(\mathbb{E}[D^F] + \frac{\mathbb{E}\left[\left(D^B + D^F + w + D'\right)^2\right]}{2\mathbb{E}\left[D^B + D^F + w + D'\right]} \right), \tag{6a}$$

$$\text{s.t.}\ \mathbb{E}\left[D^B + D^F + w + D'\right] \geq \frac{\mathbb{E}\left[M\right]}{f_{\max}}, \tag{6b}$$

where $D^F \sim \mathbb{P}_{\text{FD}}$, $D^B \sim \mathbb{P}_{\text{BD}}$, and $D' = \sum_{j=2}^{M_i}(D_j^F + D_j^B)$, $D_j^F \overset{i.i.d}{\sim} \mathbb{P}_{\text{FD}}, D_j^B \overset{i.i.d}{\sim} \mathbb{P}_{\text{BD}}$ is the equivalent transmission delay after the first sample. And w is the abbreviation of $w(D^F, D^F)$. The detailed derivation is in Appendix B of [21].

Define $Q_k := \frac{1}{2}(D_{k,1}^F + D_{k,1}^B + W_k + D_k')^2$ and $L_k := D_{k,1}^F + D_{k,1}^B + W_k + D_k'$. Q_k can be seen as the reward in the k-th epoch and L_k is the length of the k-th epoch. Because the delays D^F and D^B are independent of each other, Q_k and L_k are also i.i.d variables. Then, the problem can be seen as a renewal-reward process optimization.

B. Offline optimal policy with known channel statistics

To satisfy the frequency constraint, we will focus on searching in the set Π_{cons}.

$$\Pi_{\text{cons}} = \left\{ \pi \in \Pi_{\text{SD}} \mid \mathbb{E}\left[D^B + D^F + w + D'\right] \geq \frac{\mathbb{E}\left[M\right]}{f_{\max}} \right\}. \tag{7}$$

Notice that each stationary deterministic policy π can be represented by a function w, then finding the optimal policy $\pi^\star$ is equivalent to finding the optimal function $w^\star$. Because $w^\star$ achieves AoI_{opt}, the AoI under any other policy w is larger than AoI_{opt}. Denote $\overline{A}_w$ as the average AoI under policy w, we have the following inequality

$$\overline{A}_w = \mathbb{E}[D^F] + \frac{\mathbb{E}\left[\left(D^B + D^F + w + D'\right)^2\right]}{2\mathbb{E}\left[D^B + D^F + w + D'\right]} \geq \text{AoI}_{\text{opt}}. \tag{8}$$

For simplicity, let $\beta^\star = \text{AoI}_{\text{opt}} - \mathbb{E}[D^F]$. Deducting $\mathbb{E}[D^F]$ on both sides of inequality (8), and multiplying $\mathbb{E}\left[D^B + D^F + w + D'\right]$ on both side, we have

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\left[\left(D^B + D^F + w + D'\right)^2\right] \\ &\quad - \beta^\star \mathbb{E}\left[D^B + D^F + w + D'\right] \geq 0, \forall w \in \Pi_{\text{cons}}. \end{aligned} \tag{9}$$

Notice that (9) takes equality if and only if policy w achieves AoI minimum. Therefore, when $\beta^\star$ is known, we can obtain the optimum policy $w^\star$ by solving the following functional optimization problem:

$$\begin{aligned} \theta_{\text{opt}} \triangleq \min_{w\in\Pi_{\text{cons}}} &\frac{1}{2}\mathbb{E}\left[\left(D^B + D^F + w + D'\right)^2\right] \\ &- \beta^\star \mathbb{E}\left[D^B + D^F + w + D'\right], \end{aligned} \tag{10a}$$

$$\text{s.t.}\ \mathbb{E}\left[D^B + D^F + w + D'\right] \geq \frac{\mathbb{E}[M]}{f_{\max}}. \tag{10b}$$

Inequality (9) shows by using the optimum policy, $\theta_{\text{opt}} = 0$. To obtain the policy that achieves $\theta_{\text{opt}} = 0$, we can place the sampling frequency constraint (10b) into the objective function (10a) using a dual optimizer $\nu \geq 0$. Define $\gamma \triangleq \beta - \mathbb{E}[D']$. The Lagrange function is as follows

$$\mathcal{L}(\gamma, \nu, w) := \frac{1}{2}\mathbb{E}\left[\left(D^B + D^F + w + D'\right)^2\right]$$

$$
\begin{aligned}
&-(\gamma+\nu)\mathbb{E}\left[D^B+D^F+w+D'\right] \\
&-\mathbb{E}[D']\mathbb{E}\left[D^B+D^F+w+D'\right]+\nu\frac{\mathbb{E}[M]}{f_{\max}}.
\end{aligned} \tag{11}
$$

Proposition 1: The optimum policy $w^\star_{\gamma,\nu}$ that minimizes the Lagrange function (11) as follows

$$
w^\star_{\gamma,\nu}(D_1^F, D_1^B) = \left(\gamma+\nu-(D_1^B+D_1^F)\right)^+. \tag{12}
$$

The proof for Proposition 1 is in Appendix C of [21].

Let $\nu^\star := \underset{\nu\geq 0}{\operatorname{argsup}} \inf_{w\in\Pi_{\text{cons}}} \mathcal{L}(\gamma^\star,\nu,w)$ be the dual optimizer that resolves the Lagrange function when $\gamma=\gamma^\star$. Define function g as follows

$$
\begin{aligned}
&g(\gamma,\nu,D^F,D^B,D_1',D_2') \\
=&\frac{1}{2}\left(D^F+D^B+(\gamma+\nu-D^F-D^B)^++D_1'\right)^2 \\
&-\gamma(D^F+D^B+(\gamma+\nu-D^F-D^B)^++D_1') \\
&-D_2'(D^F+D^B+(\gamma+\nu-D^F-D^B)^++D_1').
\end{aligned} \tag{13}
$$

D_1' and D_2' have the same distribution as D'. And for simplicity, we denote the expectation of function g as

$$
\overline{g}(\gamma,\nu) = \mathbb{E}_{D^F,D^B,D_1',D_2'}[g(\gamma,\nu,D^F,D^B,D_1',D_2')].
$$

Then we have the necessary condition on $\gamma^\star$:

$$
\overline{g}(\gamma^\star,\nu^\star) = \theta_{\text{opt}} = 0. \tag{14}
$$

The function $\overline{g}(\gamma) \triangleq \overline{g}(\gamma,0)$ is monotonic decreasing and concave. The proof is in Appendix D of [21]. Because of the monotonic decreasing and concave property, when the delay distributions $\mathbb{P}_{\text{FD}}$ and $\mathbb{P}_{\text{BD}}$ are known, we can search for the optimum $(\gamma^\star+\nu^\star)$ using the bisection method [12].

Furthermore, denote $\overline{D}_{\text{ub}} = \overline{D^F}_{\text{ub}}+\overline{D^B}_{\text{ub}}+\overline{D'}_{\text{ub}}$ and $\overline{H}_{\text{ub}} = \overline{(D^F+D^B+D')^2}_{\text{ub}}$, we can bound $\gamma^\star$ as follows. The proof is in Appendix E of [21].

Lemma 1: The optimum $\gamma^\star$ can be bounded by $\gamma_{\text{lb}} \leq \gamma \leq \gamma_{\text{ub}}$, where

$$
\begin{aligned}
\gamma_{\text{lb}} &:= \max\{\frac{1}{2}(\overline{D^F}_{\text{lb}}+\overline{D^B}_{\text{lb}}-\overline{D'}_{\text{ub}}),0\}, \\
\gamma_{\text{ub}} &:= \frac{\frac{1}{2}\overline{H}_{\text{ub}}+\overline{D}_{\text{ub}}\frac{1}{f_{\max}}+\frac{1}{f_{\max}^2}}{\overline{D}_{\text{ub}}+\frac{1}{f_{\max}}}-\overline{D'}_{\text{lb}}.
\end{aligned}
$$

C. Online policy with unknown channel statistics

When the channel statistics, such as delay distributions and the packet loss probability are unknown, we can approximate $\gamma^\star$ that resolves (14) using stochastic approximation. Because (14) takes expectation with respect to two D' with the same distribution, we use the sample of D' in the last epoch and the current epoch to update γ and ν. We will use k,j to denote the index of epoch and attempted samples in the k-th epoch, respectively, where $k\in\{1,2,\cdots\}, j\in\{1,2,\cdots,M_k\}$.

To guarantee the sampling frequency constraint, we use $\nu_k=\frac{1}{V}U_k$ as the dual optimizer in epoch k, where $V>0$ is a fixed constant. Then, by assuming that $\nu^\star=\nu_k$, we can approximate $\gamma^\star$ in epoch k using the Robbins-Monro algorithm.

We start by initializing $\gamma_0 \in \text{Uni}([\gamma_{\text{lb}},\gamma_{\text{ub}}])$ and $U_0=0$. In each epoch k, the sampling and updating rules are as follows.

1) Update γ_k: After the k-th ACK of is received, we update the γ_k and ν_k. To search for the root γ of the equation (14), we update γ_k using the Robbins-Monro algorithm. Denote

$$
B_{k-1} = g(\gamma_{k-1},\nu_{k-1},D^F_{k-1,1},D^B_{k-1,1},D'_{k-1},D'_{k-2}) \tag{15}
$$

as the i.i.d sample of (14). We first compute B_{k-1} as (13).

The Robbins-Monro algorithm operates by

$$
\gamma_k = [\gamma_{k-1}+\eta_{k-1}B_{k-1}]_{\gamma_{\text{lb}}}^{\gamma_{\text{ub}}}, \tag{16}
$$

where $[\gamma]_a^b = \min\{b,\max\{\gamma,a\}\}$ and $\{\eta_k\}$ is a set of sequence that guarantees the convergence. It is selected to be:

$$
\eta_k = \begin{cases} \frac{1}{2\overline{D}_{\text{lb}}}, & k=1; \\ \frac{1}{(k+2)\overline{D}_{\text{lb}}}, & k\geq 2, \end{cases} \tag{17}
$$

where $\overline{D}_{\text{lb}} = \overline{D^F}_{\text{lb}}+\overline{D^B}_{\text{lb}}+\overline{D'}_{\text{lb}}$.

2) Update U_k: To guarantee that the sampling frequency constraint is not violated, we update the violation U_k at the end of each epoch.

$$
\begin{aligned}
U_k = \Bigg(&U_{k-1}+\Bigg[\frac{M_{k-1}}{f_{\max}} \\
&-(D^B_{k-1,1}+D^F_{k-1,1}+W_{k-1,1}+D'_{k-1})\Bigg]\Bigg)^+,
\end{aligned} \tag{18}
$$

where M_{k-1} is the number of attempted samples in the $(k-1)$-th epoch.

3) Sampling: After the k-th ACK of is received and the update of γ_k and ν_k, we select the waiting time $W_{k,1}$ by

$$
W_{k,1} = (\gamma_k+\nu_k-(D^F_{k,1}+D^B_{k,1}))^+. \tag{19}
$$

Then, after each NACK is received, the waiting time is selected as zero, i.e.,

$$
W_{k,j} = 0 \quad j=2,3,\cdots. \tag{20}
$$

In each epoch, we record the delay information, the times of transmission, and the waiting time $D^F_{k,1}, D^B_{k,1}, D'_k, W_{k,1}$.

The algorithm is summarized in Algorithm. 1.

D. Theoretical Analysis

Because the AoI evolution of the proposed online algorithm is a function of t, it is hard to analyze in general. As an alternative, we use the ratio

$$
\tilde{A}_K := \frac{\mathbb{E}\left[\int_{t=0}^{S_{K+1}} A(t)\mathrm{d}t\right]}{\mathbb{E}[S_{K+1}]} = \frac{\mathbb{E}[\sum_{k=1}^K F_k]}{\mathbb{E}[\sum_{k=1}^K L_k]}, \tag{21}
$$

Algorithm 1: Proposed Online Algorithm

Input: Frequency Constraint $f_{\max}$, Time T, hyper-parameter V.

Output: Time-Averaged AoI

Initialize $\gamma \in \mathrm{Uni}[\gamma_{\mathrm{lb}}, \gamma_{\mathrm{ub}}]$, $k = 1$, $U = 0$

while $t \leq T$ **do**

  Sample and Choose waiting time as (19) and (20)

  Record the waiting time and delay information

  if *Receive the* k*-th ACK* **then**

    Update t: $t = t + D^F_{k-1,1} + D^B_{k-1,1} + W_{k-1,1} + D'_{k-1}$

    Compute the AoI accumulation in this epoch

    Compute B_{k-1} as (15)

    Update γ_k: $\gamma_k = (\gamma_{k-1} + \eta_{k-1} B_{k-1})^{\gamma_{\mathrm{ub}}}_{\gamma_{\mathrm{lb}}}$

    Update U_k as (18)

    Set $\nu_k = \frac{1}{V} U_k$

  $k = k + 1$

Compute the time-averaged AoI accumulation.

which is the average AoI up to the $(K+1)$-th epoch of the online algorithm.

Denote $w_K = (\gamma_K + \frac{1}{V} U_K - D^F_1 - D^B_1)^+$ as the policy used in the K-th epoch. For simplicity, we use $\overline{A}_{w_K}$ to represent the expected time-averaged AoI under policy w_K. We can measure the performance of the time-averaged accumulation of AoI: $\tilde{A}_K - \mathrm{AoI}_{\mathrm{opt}}$ and the time-averaged AoI under policies w_K and optimal policy: $\overline{A}_{w_K} - \mathrm{AoI}_{\mathrm{opt}}$. When there is no frequency constraint, the main result is as follows.

Theorem 2: If the upper bound of the transmission delay D^F, D^B and the maximum transmission times in an epoch are known, i.e., $D^F < D^F_{\mathrm{ub}} < \infty, D^B < D^B_{\mathrm{ub}} < \infty, M < M_{\mathrm{ub}} \leq \infty$, the approximation error of $\gamma^\star$ up to epoch K can be bounded by:

$$E[(\gamma_K - \gamma^\star)^2] \leq \frac{2}{K} \frac{L^4_{\mathrm{ub}}}{\overline{D}^2_{\mathrm{lb}}} = \mathcal{O}(\frac{1}{K}), \tag{22a}$$

where $\overline{D}_{\mathrm{lb}} = \overline{D^F}_{\mathrm{lb}} + \overline{D^B}_{\mathrm{lb}} + \overline{D'}_{\mathrm{lb}}, L_{\mathrm{ub}} = \gamma_{\mathrm{ub}} + M_{\mathrm{ub}}(D^F_{\mathrm{ub}} + D^B_{\mathrm{ub}})$.

Also, the difference between the expected time-averaged cumulative AoI up to epoch K, i.e., $\tilde{A}_K$ and the expected time-averaged AoI under the optimal policy $w^\star$, i.e., $\overline{A}_{w^\star}$ can be upper bounded by

$$\tilde{A}_K - \overline{A}_{w^\star} \leq 2 \frac{L^4_{\mathrm{ub}}}{D D^2_{\mathrm{lb}}} \times \frac{1 + \ln K}{K} = \mathcal{O}(\frac{\ln K}{K}). \tag{22b}$$

Furthermore, the difference of expected time-average AoI under policy w_K and the optimal policy $w^\star$ can be upper bounded by

$$\overline{A}_{w_K} - \overline{A}_{w^\star} \leq \frac{2}{K} \times \frac{L^4_{\mathrm{ub}}}{D D^2_{\mathrm{lb}}} = \mathcal{O}(\frac{1}{K}). \tag{22c}$$

The proof of Theorem 2 is in Appendix F of [21].

IV. Simulations

In this section, we provide simulation results to validate the performance of our algorithm. We consider that the forward and backward transmission delay follow the log-normal distribution parameterized by μ and σ, i.e., the density function is

$$p(x) := \frac{\mathbb{P}_D(\mathrm{d}x)}{\mathrm{d}x} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \tag{23}$$

The probability of transmission failure is set to be $\alpha = 0.1$.

Since the zero-wait policy may not satisfy the sampling frequency constraint, we compare the proposed online algorithm with the following two policies:

1) A constant wait policy w_{const} that selects the waiting time by $W_{i,1} = \max\{\frac{M}{f_{\max}} - \overline{D^F} - \overline{D^B} - \overline{D'}, 0\}$.
2) The optimum policy
$$w^\star_{\gamma,\nu}(D^B_1, D^F_1) = \left(\gamma^\star + \nu^\star - (D^B_1 + D^F_1)\right)^+.$$
The $\gamma^\star + \nu^\star$ is computed by [8].

Fig. 3 studies the asymptotic average AoI performance as a function of time using different policies when there is frequency constraint, i.e., $f_{\max} = \frac{1}{10(\overline{D^F} + \overline{D^B})}$. The parameters of the delays are set to be $\mu = 1$ and $\sigma = 1.4$. The confidence interval of the average AoI of the proposed online algorithm is also plotted. From Fig. 3, it can be seen that the constant waiting policy has a larger AoI than the proposed online algorithm, and this shows the superiority in obtaining data freshness using the proposed online algorithm. In addition, when time t goes to infinity, the average AoI of the online algorithm converges to the minimum AoI.

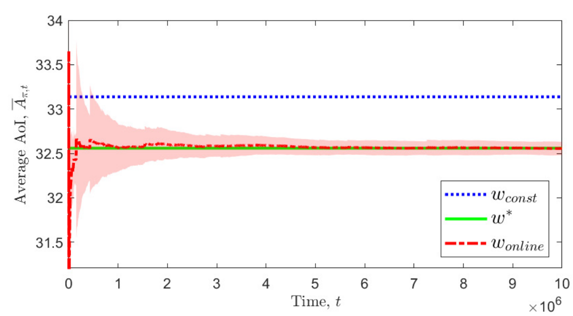

Fig. 3. The expected time-average AoI evolution under log-normal(1,1.4) with frequency constraint

Fig. 4 displays the evolution of γ_k as a function of epoch number k. The confidence interval of γ_k in the 20 runs is also plotted. From the figure, γ_k converges to the optimal $\gamma^\star$ as the epoch k goes to infinity.

Fig. 5 evaluates the evolution of the average sampling interval under different values of V. When the number of epochs increases to infinity, the averaged sampling interval keeps larger or equal to $1/f_{\max}$, which means the frequency constraint is not violated. In addition, by using a larger V, the average AoI of the online algorithm converges faster to the optimal average AoI, while by choosing a smaller V, the

ISBN 978-3-903176-55-3 © 2023 IFIP

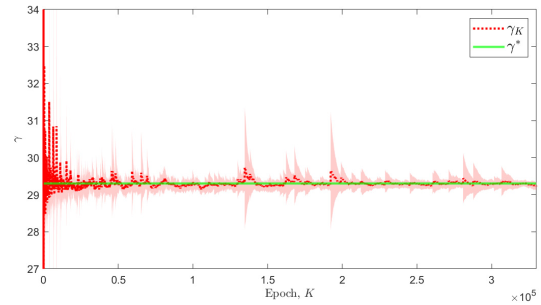

Fig. 4. The evolution of γ_k under log-normal(1,1.4)

sampling constraint can be satisfied in a shorter time, which is similar to the queueing length-utility trade-off in network utility maximization [22].

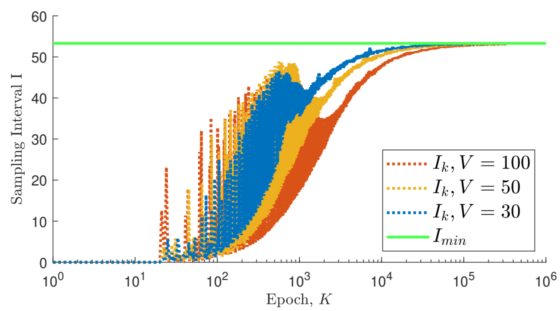

Fig. 5. The evolution of averaged sampling-interval under log-normal(1,1.4)

V. Conclusion

In this paper, we studied a status update system with a sensor sending status updates to a receiver through an unreliable channel with delayed feedback. We aimed to minimize the average AoI at the receiver while meeting the sampling frequency constraint of the sensor with unknown channel statistics. The problem was first reformulated into a renewal-reward process optimization and we proposed a stochastic approximation algorithm that can learn the AoI minimum sampling policy adaptively. Theoretical analysis and simulation results validate the convergence and performance of our algorithm.

References

[1] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Guest Editorial Age of Information," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1179–1182, 2021.

[2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*. Orlando, FL, USA: IEEE, 2012, pp. 2731–2735.

[3] B. Wang, S. Feng, and J. Yang, "When to preempt? Age of information minimization under link capacity constraint," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 220–232, 2019.

[4] B. Zhou and W. Saad, "Joint Status Sampling and Updating for Minimizing Age of Information in the Internet of Things," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7468–7482, 2019.

[5] H. Tang, J. Wang, L. Song, and J. Song, "Minimizing Age of Information With Power Constraints: Multi-User Opportunistic Scheduling in Multi-State Time-Varying Channels," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 854–868, 2020.

[6] M. A. Abd-Elmagid, H. S. Dhillon, and N. Pappas, "A Reinforcement Learning Framework for Optimizing Age of Information in RF-Powered Communication Systems," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4747–4760, 2020.

[7] E. T. Ceran, D. Gunduz, and A. Gyorgy, "A Reinforcement Learning Approach to Age of Information in Multi-User Networks With HARQ," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1412–1426, 2021.

[8] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimal sampling for data freshness: Unreliable transmissions with random two-way delay," *IEEE/ACM Transactions on Networking*, vol. 31, no. 1, pp. 408–420, 2023.

[9] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal Sampling and Scheduling for Timely Status Updates in Multi-Source Networks," *IEEE Transactions on Information Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.

[10] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *2015 IEEE International Symposium on Information Theory (ISIT)*. Hong Kong, Hong Kong: IEEE, 2015, pp. 3008–3012.

[11] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal scheduling over hybrid channels," *IEEE Transactions on Mobile Computing*, pp. 1–17, 2022.

[12] 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 Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[13] C.-H. Tsai and C.-C. Wang, "Age-of-Information Revisited: Two-way Delay and Distribution-oblivious Online Algorithm," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 1782–1787.

[14] H. Tang, Y. Chen, J. Wang, P. Yang, and L. Tassiulas, "Age Optimal Sampling Under Unknown Delay Statistics," *IEEE Transactions on Information Theory*, vol. 69, no. 2, pp. 1295–1314, 2023.

[15] H. Tang, Y. Chen, J. Sun, J. Wang, and J. Song, "Sending Timely Status Updates through Channel with Random Delay via Online Learning," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. London, United Kingdom: IEEE, 2022, pp. 1819–1827.

[16] H. Tang, Y. Sun, and L. Tassiulas, "Sampling of the wiener process for remote estimation over a channel with unknown delay statistics," in *Proceedings of the Twenty-Third International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. Seoul Republic of Korea: ACM, 2022, pp. 51–60.

[17] C.-H. Tsai and C.-C. Wang, "Distribution-oblivious Online Algorithms for Age-of-Information Penalty Minimization," *Department of Electrical and Computer Engineering Technical Reports*, 2022.

[18] E. T. Ceran, D. Gunduz, and A. Gyorgy, "Average age of information with hybrid ARQ under a resource constraint," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. Barcelona: IEEE, 2018, pp. 1–6.

[19] E. T. Ceran, D. Gündüz, and A. György, "Reinforcement learning to minimize age of information with an energy harvesting sensor with harq and sensing cost," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 656–661.

[20] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of Information: An Introduction and Survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[21] H. He, H. Tang, J. Pan, J. Wang, J. Song, and L. Tassiulas, "Age optimal sampling for unreliable channels under unknown channel statistics," https://www.overleaf.com/read/kygznmdyrzpn, 2023.

[22] M. J. Neely, E. Modiano, and C.-P. Li, "Fairness and Optimal Stochastic Control for Heterogeneous Networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 2, pp. 396–409, 2008.

ISBN 978-3-903176-55-3 © 2023 IFIP